# Modeling the growth effects of regional knowledge production: The GMR-Europe model and its applications for EU Framework Program policy impact simulations


**Attila Varga**
Faculty of Business and Economics, University of Pécs
Rákóczi u. 80
Email: vargaa@ktk.pte.hu

**Péter Járosi**
Faculty of Business and Economics, University of Pécs
Rákóczi u. 80
Email: jarosip@ktk.pte.hu

**Tamás Sebestyén**
Faculty of Business and Economics, University of Pécs
Rákóczi u. 80
Email: sebestyent@ktk.pte.hu



**Abstract**
This paper introduces the Geographic Macro and Regional (GMR) model for NUTS-2 regions of the Euro zone. This model consists of three blocks: the TFP, the SCGE and the MACRO blocks. The model is built for impact analysis of policies targeting intangible assets in the forms of R&D, human capital and social capital. The analysis can be done both at the regional and the EU macroeconomic levels. Policy simulations on the growth impacts of the 6$^{th}$ European Framework Program illustrate the capabilities of the complex model system.



**Keywords:** TFP, SCGE models, DSGE models, impact analysis R&D, human capital, social capital

**JEL:** O31, H41, O40

**Acknowledgments**: The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 216813. We especially thank Raffaele Paci, Barbara Dettori and Emanuela Marrocu for sharing with us the data on human capital, social capital and TFP developed at CRENOS.

# Modeling the growth effects of regional knowledge production: The GMR-Europe model and its applications for EU Framework Program policy impact simulations

## 1 Introduction

The geographic macro and regional modeling (GMR) framework has been established and continuously improved to better support development policy decisions by ex-ante and ex-post scenario analyses. Policy instruments targeting the development of knowledge economies (such as R&D subsidies, human capital development, entrepreneurship policies or instruments promoting more intensive public-private collaborations in innovation) are in the focus of the GMR-approach[1].

Models frequently applied in development policy analysis are neither geographic nor regional. They either follow the tradition of macroeconometric modeling (like the HERMIN model - ESRI 2002 or the QUEST II model - Veld 2007), the tradition of macro CGE modeling (like the ECOMOD model – Bayar 2007) or the most recently developed DSGE approach (QUEST III - Ratto, Roeger and Veld 2009). They also bear the common attribute of national level spatial aggregation. The novel feature of the GMR-approach is that it incorporates geographic effects (e.g., agglomeration, interregional trade, migration) while both macro and regional impacts of policies are simulated. Why does geography get such an important focus in the system? Why is the system called "regional" and "macro" at the same time?

Geography plays a critical role in development policy effectiveness for at least four major reasons. First, interventions happen at a certain point in space and the impacts might spill over to proximate locations to a considerable extent. Second, the initial impacts could significantly be amplified or reduced by short run (static) agglomeration effects. Third, cumulative long run processes resulting from labor and capital migration may further amplify or reduce the initial impacts in the region resulting in a change of the spatial structure of the economy (dynamic agglomeration effects). Forth, as a consequence of the above effects different spatial patterns of interventions might result in significantly different growth and convergence/divergence patterns.

"Regions" are spatial reference points in the GMR-approach. They are sub-national spatial units ideally at the level of geographic aggregation which is appropriate to capture proximate relations in innovation. Besides intraregional interactions the model captures interregional connections such as knowledge flows exceeding the regional border (scientific networking or spatially mediated spillovers), interregional trade connections and migration of production factors.

The "macro" level is also important when the impact of development policies is modeled: fiscal and monetary policy, national regulations or various international effects are all potentially relevant factors in this respect. As a result the model system simulates the effects of policy interventions both at the regional and the macroeconomic levels. With such an approach different scenarios can be compared on the basis of their impacts on (macro and regional) growth and interregional convergence.

The GMR-framework is rooted in different traditions of economics (Varga 2006). While modeling the spatial patterns of knowledge flows and the role of agglomeration in knowledge

[1] The framework and its roots in economics are explained in Varga (2006, 2008). The first realization of the system is the EcoRet model (Schalk and Varga 2004, Varga and Schalk 2004) which was further developed in the GMR-Hungary model (Varga 2007, Varga, Schalk, Koike, Járosi and Tavasszy 2008).

transfers it incorporates insights and methodologies developed in the geography of innovation field (e.g., Anselin, Varga and Acs 1997, Varga 2000). Interregional trade and migration linkages and dynamic agglomeration effects are modeled with an empirical general equilibrium model in the tradition of the new economic geography (e.g., Krugman 1991, Fujita, Krugman and Venables 1999). Specific macroeconomic theories are followed while modeling macro level impacts.

In order to simulate policy impacts at the regional and macro levels while incorporating geographical effects three model blocks are integrated in the GMR-system: a regional productivity (TFP) block, a spatial computable general equilibrium (SCGE) block and a macroeconomic (MACRO) block (Varga 2008). This paper introduces the GMR-Europe model system as well as provides examples of policy simulations on the macro and regional effects of the EU 6$^{th}$ Framework Program. The paper has the following structure. Section two describes the applied GMR-Europe model in five sub-sections. Section 4 presents policy impact analyses. Appendices provide further details for readers who are interested in additional technical details.

## 2 Model structure

### *2.1 Model overview*

The GMR-system integrates three sub-models which are organized in three model-blocks. The initial regional impacts of policies on total factor productivity (TFP) are modeled in the TFP block. The resulting regional level changes in quantities and prices of inputs and outputs as well as further modifications in TFP (implied by factor migration) are simulated in the spatial computable general equilibrium (SCGE) block. The SCGE model is thus responsible for estimating the effects of geography (including agglomeration forces and factor migration). However the applied SCGE model is static and as such does not account for temporal changes in labor, capital and technology in an endogenous manner. What it does is that for any given aggregate level of labor, capital and technology it calculates their equilibrium spatial distributions. As highlighted above dynamism in technology is modeled in the TFP block. Dynamic effects of interventions on labor and capital are simulated in the MACRO block. With this block QUEST III the DSGE model for the Euro zone is incorporated into the system. The three model blocks are interconnected and run subsequently until the aggregate regional impacts from the regional sub-models approach very closely the EU-level impacts estimated in the macroeconomic model.

The model system uses data from various sources. Some of them are publicly available from the EUROSTAT web-page (such as the New Cronos database for regional patents, R&D, technology employment and data for most of the macro level variables) and some of them are developed for the European Commission (such as the regional FP5 and FP6 databases and the regional publication database). The model system includes 144 NUTS-2 regions of the Euro zone. Estimation of the equations in the TFP block is carried out in SpaceStat. The GMR-system is programmed and run in Matlab.

The following sub-sections describe the three model blocks and their integration. Sub-section 2.2 explains the TFP block, 2.3 focuses on the SCGE block, 2.4 highlights those features of the MACRO block which are relevant for the impact analyses and 2.5 discloses the manner the three sub-models are integrated.

### *2.2 The regional TFP block*

The function of the TFP sub-model is to generate initial TFP changes as a result of policy interventions. Thus this model block (such as the whole GMR-system) is not designed for forecasting purposes but for policy impact analysis. In the followings the knowledge production equations and the TFP equation are introduced subsequently.

#### **2.2.1 The knowledge production equations**

Economically useful new technological knowledge is measured by patent counts spatially allocated according to the addresses of inventors (and distributed proportionally in case of multiple inventors). Shortcomings of patent data in measuring new technologies is well known in the innovation literature (e.g., Griliches 1990) however it has also been shown that this measure proxies innovation closely in the regional knowledge production function environment (Acs, Anselin and Varga 2002). The level of analysis (as throughout the two regional sub-models) is NUTS-2 European regions. The knowledge production equations are empirically estimated and explained in details in Varga, Pontikakis and Chorafakis (2010). Further information about the empirical analysis can be found in the regression tables shown in the Appendix.

Following Romer (1990) and Jones (2002) technological change is explained by the size of research and the level of already existing technological knowledge. The corresponding empirical relationship is estimated by the following regional knowledge production function.

$$(1)\ Log(PATENTS) = 1.325381*(-2.3006 + BETAPAT*Log(GRD(-2)) + 0.1804*\ Log(PSTCK_N(-2)) + 0.4614*\ PAHTCORE) + U_1$$

where GRD is gross research and development expenditures (including both private and public expenditures) $PSTCK_N$ is national level stock of patents (measuring already accumulated knowledge at the country level), PATHCORE is a dummy representing regions with high patenting activity (i.e., regions where the number of patents is two standard deviations higher than the average in the sample). Each estimated parameter is multiplied by 1.325381 which is the spatial multiplier[2].

BETAPAT measures regional productivity of research. It is an elasticity representing the impact of research expenditures on patents. It is assumed that regional research productivity is not constant over space but varies according to the agglomeration of knowledge necessary in innovation in the region (Varga 2000). Thus regions where considerable amount of complementary knowledge is accumulated at innovative manufacturing and service firms or public organizations are assumed to use R&D expenditures more efficiently in knowledge production than those regions where knowledge is less agglomerated.

The estimated equation of BETAPAT is:

$$(2)\ BETAPAT = [(0.7088 + 0.1439*Log(\delta(-2))]$$

where agglomeration of knowledge is measured by the following index:

[2] The spatial multiplier represents the indirect knowledge inputs from spatially proximate regions. That is for example not only R&D carried out in the region effects regional knowledge production but also interactions of the region with spatially proximate regions (via formal collaborations, learning or pure knowledge spillovers) affect it indirectly. The value of the spatial multiplier is calculated based on the spatial lag coefficient in Table 1 following Anselin 1988.

(3) $\delta_i = [(EMPKI_i / EMPKI_{EU}) / (EMP_i / EMP_{EU})] / [(1 - \sum_j (EMPKI_{i,j} / EMPKI_{j,EU})][1 - (EMP_i / EMP_{EU})]$

Equation 3 is an index of relative regional specialization of knowledge intensive employment with a correction for the size of the regional economy[3].

R&D is not constant over time. It is assumed that regions with high R&D productivity attract further research activities. The following equation shows the empirical relationship between changes in regional R&D and research productivity:

(4) $(GRD-GRD(-3)) = -299.107 + 351.824*BETAPAT(-3) + 190.322*BETAPUB(-3) + 360.98*RDHCORE + U_3$

RDHCORE is a dummy variable[4]. BETAPAT has already been explained. BETAPUB is productivity of pre-competitive research in the region to produce scientific publications. The BETAPUB equation is a result of a related empirical model exhibited in Table 2 in the Appendix and has the following form:

(5) BETAPUB = [0.4317 + 0.0003* WFP5_Log(RD(-2))]

where WFP5_Log(RD) is the sum of (the log of) R&D expenditures of partner regions in the $5^{th}$ Framework program. While BETAPAT represents "agglomeration effects" in research productivity in patenting BETAPUB reflects the significant impact of formal interregional research collaborations on the productivity of research in producing publications. This second effect is termed "network effect" in regional research productivity.

It is also assumed that agglomeration of knowledge intensive employment partly follows the spatial distribution of R&D. The following empirical equation exhibits this relationship formally:

(6) $(EMPKI-EMPKI(-3)) = 11168.3 + [(0.0262 + 5.624E\text{-}06*GRD(-3))]*EMPKI(-3) + 21321.1*RDCORE + U_4$

where EMPKI is regional knowledge intensive employment as before and RDCORE is a dummy variable[5]. Equation (6) shows that changes in the agglomeration of knowledge are to a large extent a path dependent phenomenon. However, R&D also plays a role: attraction of knowledge intensive employment to regions with considerable R&D activities is more intensive than otherwise.

Equations 1 to 6 reflect the dynamic nature of R&D support policy impacts. In a relatively short run this support affects patenting directly while in the longer run it also strengthens concentration of research and knowledge intensive employment in the region which further impacts knowledge production indirectly (via additional R&D and increased values of the parameter BETAPAT). This dynamic feature is represented by Figure 1 where the first 7 time periods are shown (without continuing the impacts throughout additional periods).

[3] EMPKI is employment in knowledge intensive economic sectors (high and medium high technology manufacturing, high technology services, knowledge intensive market services, financial services, amenity services – health, education, recreation) and EMP is total employment.
[4] RDHCORE is 1 for regions with more than two standard deviations higher than average R&D expenditures and 0 otherwise.
[5] RDCORE is 1 for regions with R&D expenditures above one standard deviation of the sample mean and 0 otherwise.

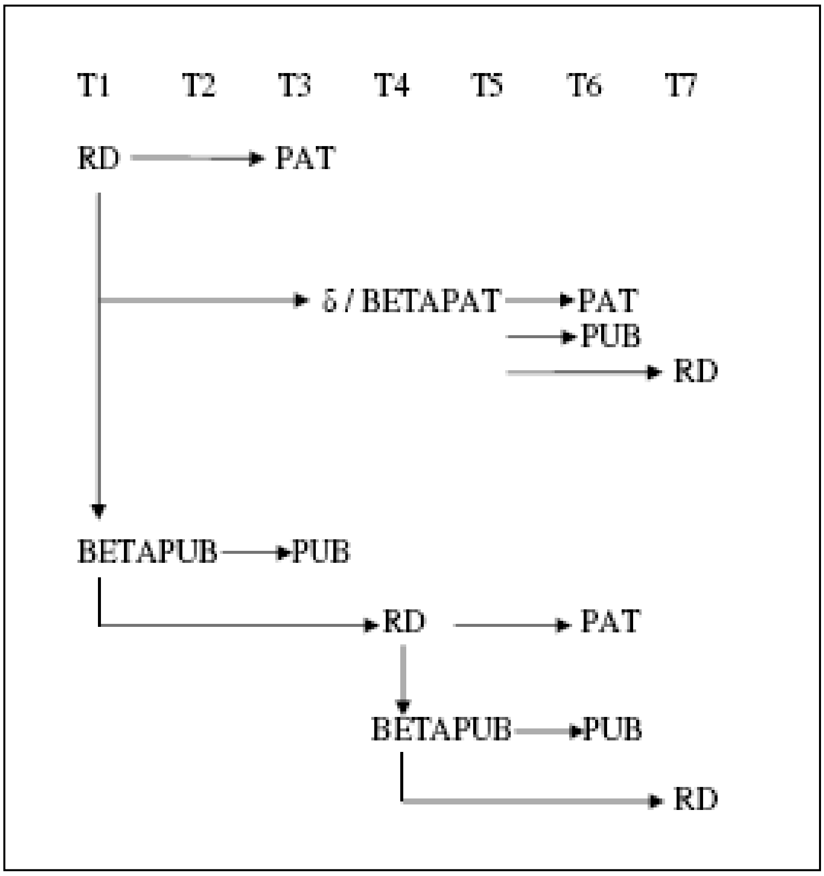


Figure 1: The dynamic impacts of R&D promotion (followed only for the first seven periods)

<u>2.2.2 The TFP equation</u>

R&D is a very important intangible asset of a region however it is not the only one that might be critical for regional (and aggregate) growth. Dettori, Marocu and Paci (2009) draw attention to the role of human capital and social capital in this context. In the followings we introduce the estimated regional TFP equation that plays a central role in channeling policy effects into the rest of the GMR model system. Data on regional human capital, social capital and TFP is kindly provided by CRENOS. Human capital is measured by the number of people that has attained at least a university degree. The proxy for social capital is the share of population over total population that has taken part at least once in the last 12 months in social activities (such as voluntary service, unions and cultural associations meetings). TFP is estimated within a regression context in Dettori, Marocu and Paci (2009). TFP is calculated for 2004 whereas the rest of the variables are collected for 2002 in order to account for a reasonable time lag between inputs and the resulting TFP level.

Table 1 provides details on the regression analysis. The idea behind the estimated model is that human capital and accumulated technological knowledge are the main inputs to regional productivity. However, the human capital effect on TFP is largely influenced by the level of social capital in the region. That is regions where substantial levels of trust, willingness to collaborate and knowledge sharing are present utilize their human capital in a more effective

**Table 1. The TFP equation. Regression Results for Log (A) for 135 Eurozone regions, 2004**

| Model | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Estimation | OLS | OLS | OLS | OLS | IV (2SLS) Spatial Lag (INV1) |
| Constant | 3.6425*** (0.2105) | 4.0850*** (0.0460) | 3.9331*** (0.0425) | 3.9832*** (0.0385) | 3.9309*** (0.0414) |
| Log(HUMCAP(-2)) | 0.0722*** (0.0175) | | | | |
| Log(HUMCAP(-2))*SOCKAP(-2) | | 0.0008*** (7.9577E-5) | 0.0003*** (8.7574E-05) | 0.0004*** (7.5823E-5) | 0.0004*** (7.4023E-5) |
| Log(PATSTCK(-2)) | | | 0.0623*** (0.0078) | | |
| Log(PATSTCK(-2))*Log(DENS(-2)) | | | | 0.0073*** (0.0008) | 0.0054*** (0.0010) |
| W_Log(A) | | | | | 0.0015*** (0.0005) |
| $R^2$-adj | 0.11 | 0.41 | 0.60 | 0.63 | |
| Sq. Corr. | | | | | 0.65 |
| Multicollinearity condition number | 22 | 6 | 9 | 7 | |
| White test for heteroskedasticity | 8.8335** | 11.1798*** | 10.5357* | 7.7393 | |
| LM-Err | | | | | |
| INV1 | 154.48*** | 57.35*** | 1.57 | 3.27* | 1.38 |
| INV2 | 19.56*** | 9.00*** | 0.61 | 0.02 | |
| LM-Lag | | | | | |
| INV1 | 52.47*** | 38.11*** | 14.98*** | 7.67*** | |
| INV2 | 29.31*** | 22.03 | 11.33*** | 3.09* | |

Notes: HUMCAP is human capital, year 2002; SOCKAP is social capital, year 2002; PATSTCK is cumulated number of patents (1991-2002), year 2002; DENS is employment (in thousands) per area of the region, year 2002. Data sources: TFP, HUMCAP, SOCKAP (CRENOS), PATSTCK (EPO), DENS (EUROSTAT); Estimated standard errors are in parentheses; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; For Model 4 the Durbin-Wu-Hausman test for Log(HUMCAP)*SOCKAP and Log(PATSTCK)*Log(DENS) does not reject exogeneity; The 3-Group method was followed in instrument selection for the D-W-H test; W_Log(TFP) is the spatially lagged dependent variable where W stands for the weights matrix INV1; Instruments in Model 5 are the spatially lagged exogenous variables calculated with weights matrix INV1; The average value of regional spatial multipliers is 1.03035; *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates significance at $p < 0.1$.

way than regions with lower levels of social capital. Similarly it is assumed that regionally accumulated technological knowledge (proxied by patent stock) impacts TFP more productively in regions where industry shows a considerable concentration (measured by the density of employment). The reason behind this is that industrial concentration enhances opportunities for the application of locally accumulated knowledge as well as it provides better possibilities for formal and informal interactions.

Regression results support the hypothesized relationships described in the previous paragraph. Though both human capital (HUMCAP) and patent stock (PATSTCK) enter the equation with highly significant parameters cross products of human capital and social capital (SOCKAP) and patent stock and density (DENS) remain highly significant but result in estimated equations with better regression fits (i.e., model 1 vs. model 2 and model 3 vs. model 4).

Multicollinearity is not an issue (the Multicollinearity condition number is well below the threshold value of 30) however spatial dependence remained a problem in model 4 in the form of spatial lag dependence. The weights matrix is INV1 which is an inverse distance matrix. Though the left hand side variables lag two years behind the dependent variable and as such no endogenous relationship is expected in the equation, data errors might be the source of correlation between the explanatory variables and the error term (Dettori, Marocu and Paci 2009). However the D-W-H test does not reject exogeneity for the left hand side variables. Given that error terms are not normal the appropriate regression is the spatial lag model estimated with the instrumental variables methodology (2SLS). In the final model (model 5) no remaining spatial error dependence is found.

Equation 7 is the estimated form of the TFP equation:

$$(7)\ A = 57.42*(HUMCAP(-2))^{0.0004*SOCKAP(-2)}\,(PATSTCK(-2))^{0.0056*\ln(DENS(-2))}$$

## *2.3 The regional SCGE block*

To model dynamic agglomeration effects of policy interventions in the GMR-system a spatial computable general equilibrium (SCGE) model is integrated. CGE models are numerical and empirical applications of Walrasian general equilibrium models in real world circumstances (Shoven and Whalley 1992). These models build on usual assumptions in microeconomics (i.e., utility and profit maximization/cost minimization, perfect competition and most recently monopolistic competition). CGE models are especially well suited to simulate the short- and long run impacts of shocks to the system. A particularly attracting feature of these models is that they do not need as many observations and details in the data as more traditional econometric techniques do.

Spatial CGE modeling is a very recent development in empirical research. Areas of application in regional analysis range from transport modeling to environmental analysis (Donaghy 2009). A particular class of SCGE models follows the tradition of the new economic geography. A couple of such examples include the CGE Europe (Bröcker 1998), Venables and Gasiorek (1999) and the RAEM model (Oosterhaven et. al 2001, Thissen 2003, Ivanova et al 2007). These models are empirical counterparts of new economic geography systems. Resulting from the policy shock each region finds its equilibrium quantities and prices of inputs and outputs in the short run. This does not mean that the whole spatial system is in equilibrium at that stage. This happens only in the long run when inclinations for firms or households to relocate disappear as real incomes across regions equilibrate resulting from previous migrations.

The particular SCGE model integrated into our framework is the modified version of the RAEM model. This model is especially suitable in situations when regional data are only scarcely available for several variables necessary in RAEM. This section draws on the descriptions presented in Varga (2007) and Járosi, Koike, Thissen and Varga (2009).

### **2.3,1 Main model assumptions**

a. The model considers 144 European regions of the EURO zone;

b. The model distinguishes between short run (i.e., a period of one year with the assumption that equilibrium at each region is reached at both goods and factor markets) and long run (several years through which the system is attracted towards a spatial equilibrium as a result of factor movements across regions);

c. The total number of households is assumed fixed;

d. Total housing supply is fixed or exogenously determined in each region;

e. Capital and labor are used in production;

f. Iceberg-type transportation cost (i.e., transportation cost is measured as a portion of the good needed to transport the commodity for a given distance);

g. Capital stock is owned by households (national dividend);

h. The model considers both centripetal and centrifugal forces that form the geographical structure of the economy. Centrifugal forces weaken spatial concentration while centripetal forces work towards further agglomeration. In the model the centrifugal forces are transportation costs and congestion. The level of congestion is measured by per capita housing. As indicated above housing supply is considered fixed in the model consequently increasing population decreases per-capita housing which works against agglomeration. The centripetal force in the model is a positive agglomeration economy measured by the level of Total Factor Productivity in the region in accordance with Equation 7. Increasing concentration of economic activities (measured by the level of employment in the model) increases the probability of interactions among the actors of innovation in the region that results in a higher technological level. Thus increasing concentration works towards further agglomeration. The actual balance between centripetal and centrifugal forces in the model determines the migration of labor and capital. As such the spatial distribution of production, TFP and inputs are all determined by the interplay of centrifugal and centripetal forces.

**2.3.2 Main model equations**

Production is determined by a C-D technology:

(8) $$y_{i,t} = TFP_{i,t} \mathrm{L}_{i,t}^{a_i} \mathrm{K}_{i,t}^{1-a_i}$$

where i stands for region and t is for time, K, L and TFP representing capital, labor and total factor productivity, *a* is production elasticity of labor.

The F.O.B. prices[6] of region *i*

(9) $$\mathrm{q}_{\mathrm{i,t}} = \frac{\mathrm{w}_{\mathrm{i,t}}^{\mathrm{a_i}} \mathrm{r}^{1-\mathrm{a_i}}}{\mathrm{TFP}_{\mathrm{i,t}} \mathrm{a}_{\mathrm{i}}^{\mathrm{a_i}} (1-\mathrm{a}_{\mathrm{i}})^{1-\mathrm{a_i}}}$$

where w is the wage rate, r is capital rent.

The input factor demand functions:

[6] F.O.B. = „free on board"

(10) $$\mathrm{L}_{i,t} = \frac{a_i}{w_{i,t}} q_{i,t} y_{i,t}$$

(11) $$\mathrm{K}_{i,t} = \frac{1-a_i}{r} q_{i,t} y_{i,t}$$

where w and r are the prices of labor and capital.

The utility function of the households:

(12) $$\ln(u_{i,t}) = \alpha_H \ln\left[\frac{H_i}{L_{i,t}}\right] + \beta \ \ln\left[x_{i,t}\right]$$

where H is housing stock, x is final goods.

The budget constraint of the households:

(13) $$w_{i,t}\frac{L_{i,t}}{N_i} + r\frac{\sum_{i=1}^{I} K_{i,t}}{\sum_{i=1}^{I} N_i} = p_{i,t} x_{i,t}$$

where N is population p is price of goods including transportation costs (C.I.F price).

Utility maximization results in the following demand function:

(14) $$x_{i,m,t} = \frac{\beta}{1-\alpha_H}\frac{1}{p_{i,t}}\left(w_{i,t}\frac{L_{i,t}}{N_i} + r\frac{\sum_{i=1}^{I} K_{i,t}}{\sum_{i=1}^{I} N_i}\right)$$

The probability of buying goods in region *i* when living in region *j* is defined as follows:

(15) $$\mathrm{s}_{\mathrm{ij,t}} = \gamma_{\mathrm{ji}}\left[\frac{(1+\tau_{\mathrm{ij}})\mathrm{q}_{\mathrm{i,t}}}{\mathrm{p}_{\mathrm{j,t}}}\right]^{-\mu}$$

where $\tau$ represents the „iceberg transportation cost principle”, that is the quantity of a good that accounts for transportation costs while the good is transported from i to j, $\mu$ and $\gamma$ are constant parameters.

Thus interregional trade volume gets the following form:

(16) $$z_{ij,t} = N_j x_{j,t} s_{ij,t}$$

Aggregate demand in region j gets calculated as follows:

(17) $$N_j x_{j,t} = \sum_{i=1}^{I} z_{ij,t}$$

The cost of transportation is also included in the C.I.F. price:

(18) $$p_{j,t} = \sum_{i=1}^{I} s_{ij,t} q_{i,t} \left(1 + \tau_{ij}\right)$$

Considering Equation (15) this always equals to the following CES form:

(19) $$p_{j,t} = \left\{ \sum_{i=1}^{I} \gamma_{ji} \left[ \left(1 + \tau_{ij}\right) q_{i,t} \right]^{1-\mu} \right\}^{\frac{1}{1-\mu}}$$

**2.3.3 Short run market equilibrium conditions**

- labor market:

(20) $L_{i,t}^{(dem)} = L_{i,t}^{(sup)}$ in every region, $\forall\ i = 1..I$

- capital market:

(21) $$r\left( \left( \sum_{i=1}^{I} K_{i,t}^{(dem)} \right) - K_{TOTAL,t}^{(sup)} \right) = 0$$

- goods market:

(22) $$y_{j,t} = \sum_{i=1}^{I} \left(1 + \tau_{ij}\right) z_{ij,t}$$

**2.3.4 The long run equilibrating mechanism through migration**

Utility differences across regions determine migration:

(23) $$LMIGR_{i,t} = \Phi\left( e^{\Theta \cdot (u_{i,t})} - e^{\Theta \cdot AVG(u_{i,t})} \right) L_{i,t}$$

where $L_{i,t}$ is labor of region i in year t, while $\Phi$ and $\Theta$ are parameters.

$\sigma$ represents the share of savings in total output that is (1- $\sigma$) part of outputs are consumed by the households and $\sigma$x is investment[7]. Thus the utility function in equation (23) has been changed to the following form:

(24) $$\ln(u_{i,t}) = \alpha \ln\left[ \frac{H_i}{L_{i,t}} \right] + \beta\ \ln\left[ (1 - \sigma\ ) x_{i,t} \right]$$

Equation (23) well exemplifies, that if value of ($u_{i,i}$) in the given region is exactly the average of the ($u_{i,t}$) values of the all regions, then there is no migration in the given region.

[7] The actual value of σ is calculated and taken from the MACRO block.

2.3.5 Parameter values

Some of the parameters are taken from earlier studies/experiences, some of them are estimated econometrically and some of them are calibrated. As a result variables of the system without shocks replicate the observed values of the same variables.

### *2.4 The MACRO block*

The SCGE model accounts for the spatial dynamic effects of policy interventions. The spatial dynamics is driven by the actual balance of centrifugal (transportation costs, congestion) and centripetal (regional TFP) forces and result in the migration of production factors until the full spatial equilibrium is attained. This model is static in the temporal sense. Temporal and spatial changes in TFP resulting from policy shocks are calculated in the TFP block. Temporal changes in capital and labor caused by policy interventions are calculated by the macroeconomic model where temporal adjustments are in focus (but spatial effects are completely missing). In an ideal system spatial and temporal dynamics are integrated right at the regional level. There are already some attempts to integrate the two dimensions in regional models (Ivanova et al 2007, Bröcker and Korzhenevych 2008). However still further efforts are needed to attain a full-fledged solution for a complete theoretical and empirical integration of the temporal dynamisms of policy induced changes in technology, labor and capital in a spatial equilibrium setting.

The applied macroeconomic model in the current GMR-system is Quest III a dynamic stochastic general equilibrium models (DSGE models from here on) macromodel for the Euro zone. DSGE models became the workhorse of modern macroeconomics in the last one and a half decade. These models are called dynamic, as they represent the dynamic aspects of economic activity explicitly capturing the dynamic behavior of agents: they operate with forward-looking decisions of households and firms. They are stochastic, as stochastic shocks to different structural relationships are considered. And finally, these models are general equilibrium models as they work with equilibrium conditions in all markets.

In contrast to more traditional macroeconometric models DSGE models bear the advantage of explicit microfoundations: these models are based on rational optimizing behavior of economic agents. This feature makes them theoretically very coherent on one hand, but creates some difficulties with regards to empirical fit on the other: these models do not capture the data generating process underlying the observed economic time series thus making it especially difficult to bring them to data. However, considerable efforts have been made to increase the empirical fit of DSGE models. For example Smets and Wouters (2003) show that a New Keynesian DSGE model is able to track and forecast time series as well as, if not better than, a vector autoregressive model estimated with Bayesian techniques (BVAR).

In this paper we use the QUEST III model as the macro part of our integrated system. The QUEST III model was developed by the economists of the European Commission' Directorate General Economy and Finance for the Euro zone. This model reflects all basic features of contemporary DSGE models highlighted in the previous section. The QUEST III model is a New Keynesian open-economy DSGE model with active fiscal and monetary policy, forward-looking households and firms, real and nominal rigidities. The model also works with several numbers of exogenous shocks both at the demand and at the supply sides. The model is estimated using Bayesian techniques in order to fit to Euro area data and then used for the evaluation of policy interventions and different kinds of external shocks. For a detailed discussion of the QUEST model and relevant analysis, please consult the paper of Ratto et al. (2009).

### *2.5 Model integration*

Figure 1 describes the way the different sub-models are interrelated in the complex system. Following this figure the current section explains the model structure in details. Without interventions TFP in both the macro model and in the regional models grow with a constant rate. This growth rate (0.974 percent each year) is estimated in the Quest III model.

Step 1 When a TFP-related policy shock happens (in the forms of R&D support resulting in an increase in patent stock or human capital and social capital development) it induces changes in the value of A in Equation 7. The baseline value of TFP ($\mathrm{TFP}_{i,t=0}(1+\mathrm{TFPGROWTH})^{t}$) is then multiplied by the ratio of the value of $\mathrm{A}_{i,t}$ with interventions ($\mathrm{A}_{i,t}^{SH}$) and without interventions ($\mathrm{A}_{i,t}^{0}$). Equations 7a – 31 below show this in details.

(7.a) $$A_{i,t}^{SH} = \omega_i e^{SPATMULT_i \alpha_0} HUMCAP_i^{SPATMULT_i \alpha_1 SOCKAP_i} PATSTCK_{i,t}^{SPATMULT_i \alpha_2 \ln(L_{i,t}/AREA_i)}$$

(30) $$\mathrm{TFP}_{i,t=0} = \mathrm{A}_{i,t=0}^{0}$$

(31) $$TFP_{i,t} = TFP_{i,t=0}(1+TFPGROWTH)^{t}(A_{i,t}^{SH} / A_{i,t}^{0}) \qquad \text{if } t>0$$

Step 2 In the next step $\mathrm{TFP}_{i,t}$ enter the SCGE model where equilibrium values of capital, labor, output, consumption, wages, capital rents and final good prices are calculated for each region and for each year. Differences in utility levels induce factor migration. As a result of this process the equilibrium value of $\mathrm{A}_{i,t}^{SH}$ might not remain the same: changing spatial distribution of labor induces changes in labor density in the power of PATSTCK (Equation 7.a) altering the value of $\mathrm{A}_{i,t}^{SH}$ in the region.

Step 3 In the following step regional TFP values are weighted averaged[8] for each year to get the macro level aggregate in TFP. These annual values enter the MACRO model as a shock in Equation (27) where equilibrium macroeconomic values are estimated for several variables.

Step 4 Equilibrium aggregate values of investment and change in labor calculated in the MACRO model are distributed across regions following the patterns of policy induced changes in TFP:

(32) $$\frac{\Delta L_i}{L'_{i,t}} = E_{t+1/t} \frac{\Delta TFP_i}{TFP_{i,t}}$$

[8] With the following weights: $\mathrm{L}_{i,t}^{a_i}\mathrm{K}_{i,t}^{1-a_i}$

where

(33) $$L'_{i,t} = L_{i,t} + LMIGR_{i,t}$$ [9]

and

(34) $$E_{t+1/t} = \frac{\Delta L_{TOTAL}}{L_{TOTAL,t}} : \frac{\Delta TFP_{AVG}}{TFP_{AVG,t}}$$

with

$$\sum_i L_{i,t} = L_{TOTAL,t}\,;\, \Delta L_{TOTAL} = L_{TOTAL,t+1} - L_{TOTAL,t}\,;\, \Delta L_i = L_{i,t+1} - L'_{i,t}\,;$$

and

$\Delta TFP_{AVG} = TFP_{AVG,t+1} - TFP_{AVG,t}$ and $\Delta TFP_i = TFP_{i,t+1} - TFP_{i,t}$

The resulting change in regional labor is calculated as follows:

(35) $$L_{i,t+1} = L'_{i,t} + E_{t+1/t} \frac{TFP_{i,t+1} - TFP_{i,t}}{TFP_{i,t}} L'_{i,t}$$

Investment increases total capital:

(36) $$K_{TOTAL,t+1} = (1-\delta)K_{TOTAL,t} + INV_{TOTAL,t}$$

Where $\delta$ is the average depreciation rate according to the corresponding value in the MACRO model. Investment shares in output for each year in the MACRO model is taken to the SCGE model as $\sigma$ in Equation 24.

Step 5 In the next step the SCGE model is run again to calculate the equilibrium quantities and prices for each region. At this stage the calculations will result in the regional distribution of quantities and prices that bear the impacts of both spatial and temporal dynamisms.

Step 6 In most of the cases the aggregate values of regional output, capital, labor and consumption closely correspond to the respective values in the MACRO model. However, if this close correspondence is not attained Steps 2 to 5 are re-run until this happens.

[9] LMIGR is explained in Equation 23.

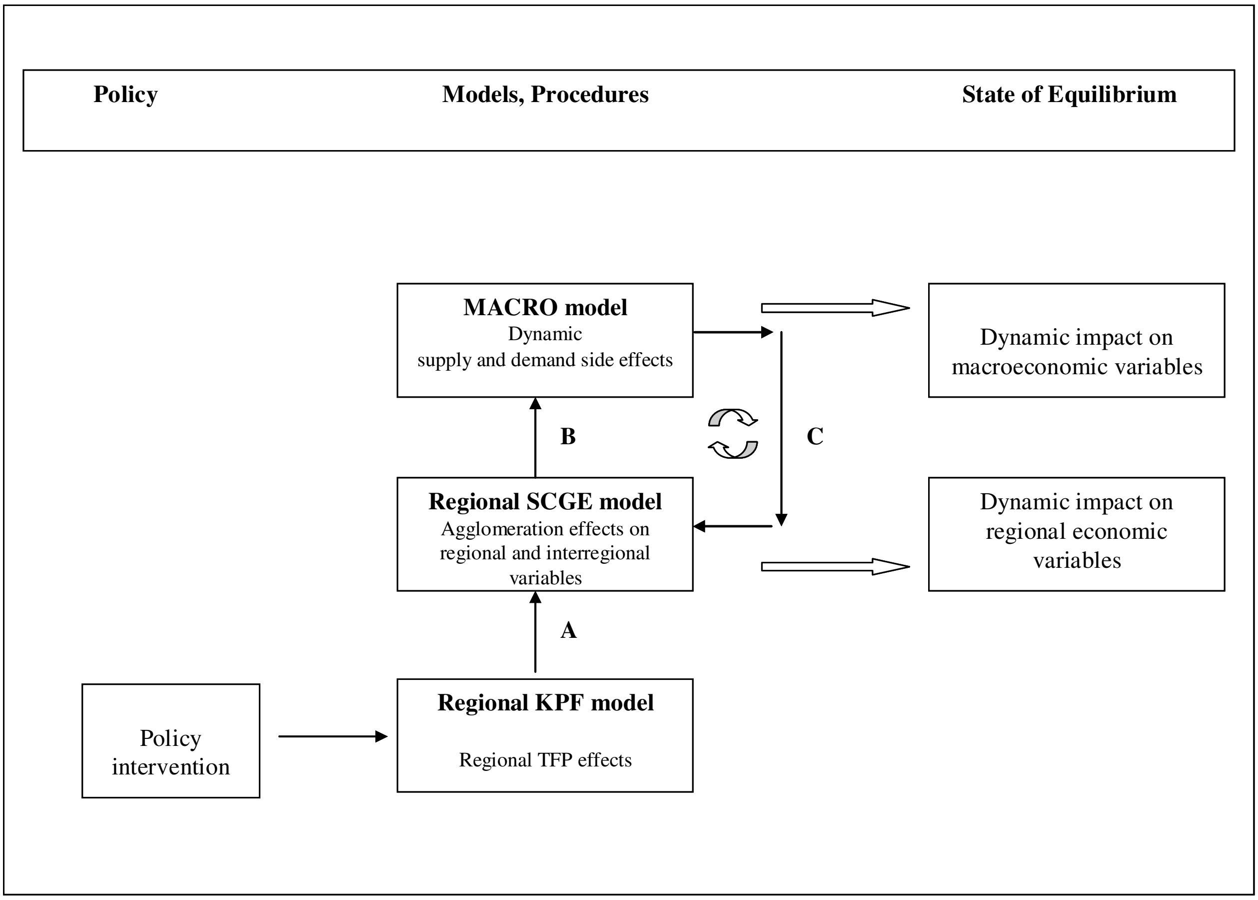


**Figure 1**. The mechanism of the effects of TFP-related policy interventions

## 4. Policy impact analysis

### *4.1 Regional and macro-level impacts of EU FP6 research contributions*

EU Framework programs are designed with the aims of serving the purposes of both scientific progress and technological development. Impact analysis of the FP programs have usually been based on surveys of participants (e.g., Polt, Vonortas, Fischer et al. 2008) which can provide good information at the level of participating institutions or firms, but not at the level of regions where participants located not to mention the level of the European Union. With the help of the complex geographic macro and regional model described in this paper the impacts of EU R&D contributions within the 6$^{th}$ Framework program can be estimated. Main results of the impact analysis are presented in this section.

The Institute for Prospective Technological Studies of the European Commission collected data on FP6 EU R&D contributions and provided the regional and temporal distribution of them for the period of 2003-2007. The monetary values correspond to the information on the projects in the Fall of 2008. Figure 2 exhibits the spatial distribution of funds for the whole period in the Euro zone.

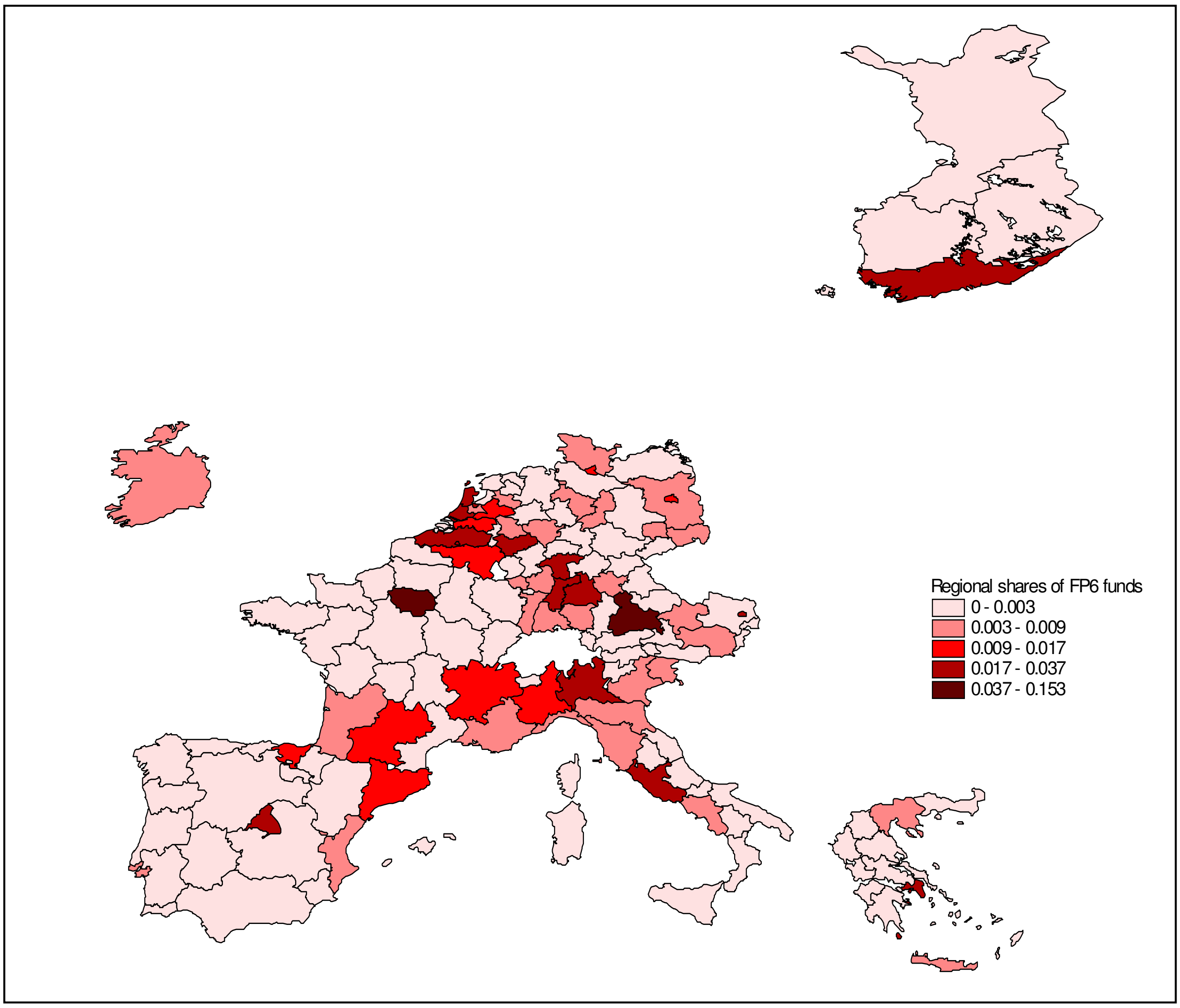


**Figure 2**. Regional distribution of FP6 funds in the Euro-zone, 2003-2007

Euro zone regions are classified according to their level of agglomeration given by the values of the agglomeration index (Equation 3). Regions with values of the index of more than one standard deviation above the mean belong to the first tier. Second tier regions exhibit agglomeration values between the mean and the mean plus one standard deviation. Third tier regions are half standard deviation value below the mean whereas the rest of the regions belong to the fourth tier. Average impacts on GDP in regions belonging to these four tiers are shown in Figure 3.

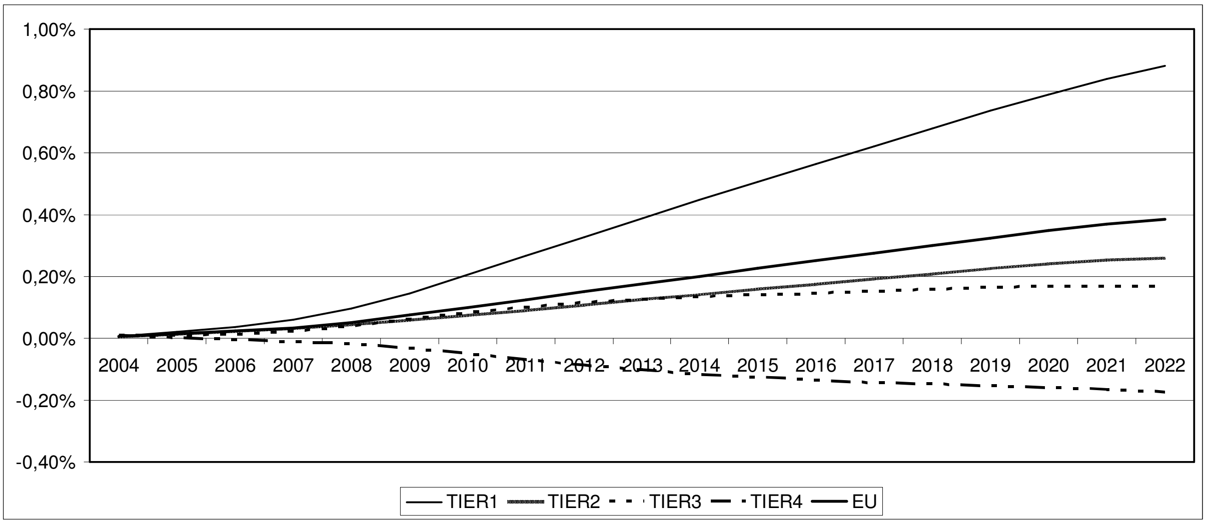


**Figure 3**. Average FP6 impacts on GDP in regions belonging to different agglomeration tiers: percentage differences between scenario and baseline values

As it is clear from the figure the estimated impacts are not dramatic. However one cannot expect large impacts from EU R&D contributions accounting for about 4 percent of regional R&D expenditures on average. More than 60 percent of the funds are won by regions belonging to the first tier. Thus it would not be a surprise if the largest impacts are found in these regions. According to the expectations the relative impacts are highest in first tier regions (by the end of the examination period GDP exceeds its no intervention level by about 0.88 percent) whereas market loss and negative net migration result in a slight decline in average GDP in fourth tier regions. (These regions won less than about 4.5 percent of all the FP6 funds during the period of the program.)

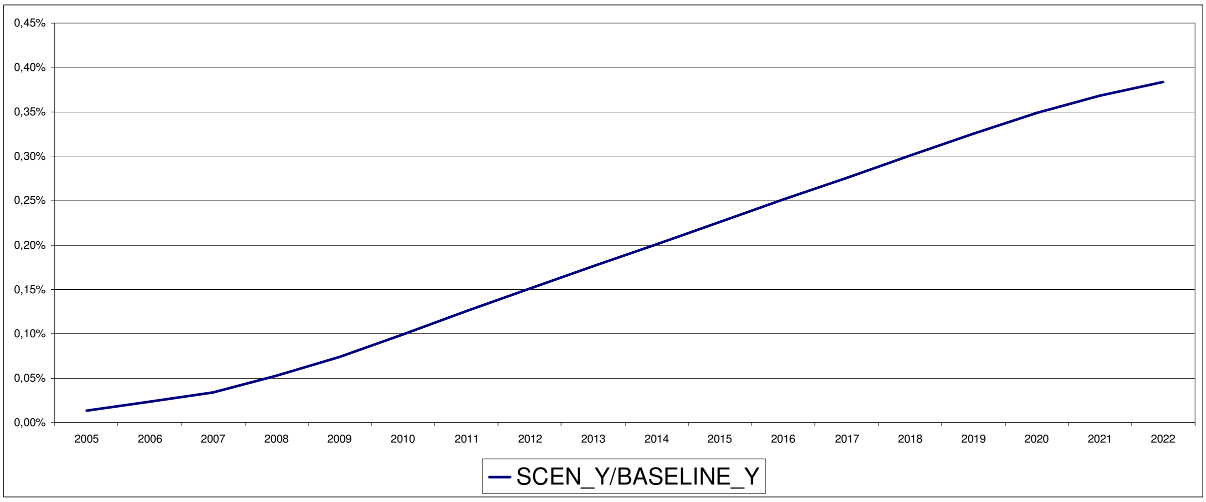


**Figure 4.** Impacts of FP6 funds on EU GDP, Euro-zone, period 2003-2022: percentage differences between scenario and baseline values

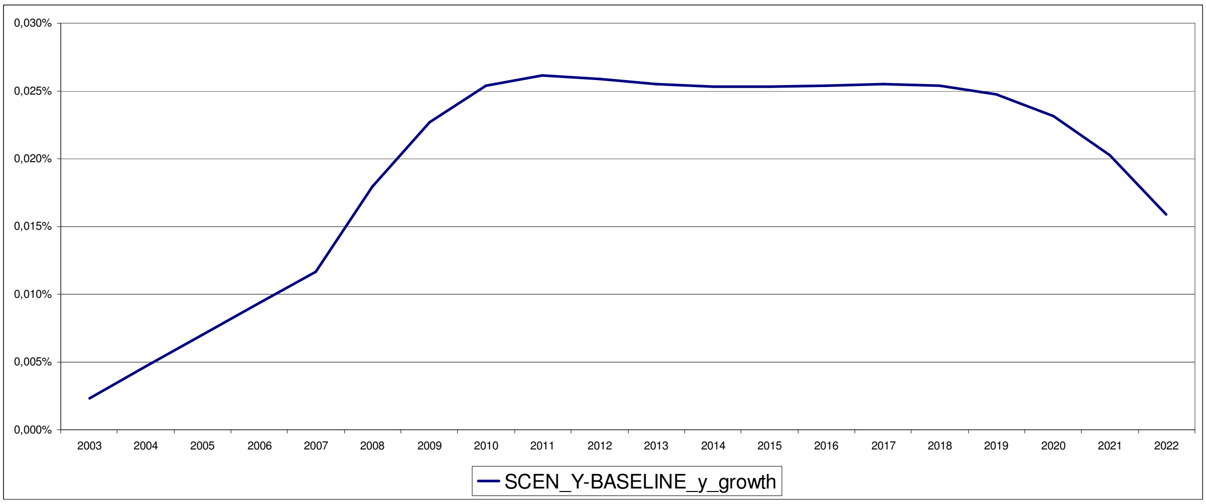


**Figure 5.** The impact of FP6 funds on EU-level GDP growth rates, Euro-zone, 2003-2022: percentage point differences between scenario and baseline values

In Figure 5 the estimated impacts on GDP at the Euro zone level are provided. After a slow increase of the initial impacts from 2008 changes in the differences between the non-intervention (baseline) GDP and the GDP of the FP6 impact start to increase from 2007 (which is caused by the lagged temporal effects as well as the induced agglomeration effects as estimated equations in the TFP block describes it) . By the end of the study period EU level GDP is about 0.38 percent higher than it would be without the 6$^{th}$ Framework program.

Figure 6 shows percentage point differences between EU GDP growth rates with and without the FP6 program. The differences increases until 2010, then slightly declines until 2018 and starts to diminish dramatically after 2019 and are expected to reach the zero difference in later periods (not included in the simulations). This is in accordance with what is expected from temporally positive TFP shocks: they increase GDP levels but not the GDP growth rate in the long run.

### *4.2 R&D specialization and the impact of FP6*

There is an ongoing policy debate among high level decision makers and experts of the European Commission about the necessity and potential impacts of R&D specialization (Pontikakis, Kyriakou, Bavel 2009). Should the European Commission and Member States concentrate R&D resources in technological or geographical areas with high research productivity in the expenses of regions lagging in this respect? What are the potential benefits of such specialization on economic growth and what are (if any) the costs in the sense of increased territorial inequalities in the EU?

Connected to the R&D specialization debate in the European Commission in this policy simulation we are interested if the impact of EU FP6 funds would be different at regional and macro levels if Members States followed a more efficient spatial distribution of their public support on R&D. Assuming that the selection of supported R&D projects in the EU Framework Programs in general follows stricter scientific quality standards than most of the programs of Member States we designed a scenario where 1 percentage of total national R&D expenditures is re-distributed according to the spatial pattern of FP6 funds for each year of the simulation period (2003-2022) and for each country included in the sample. Though the

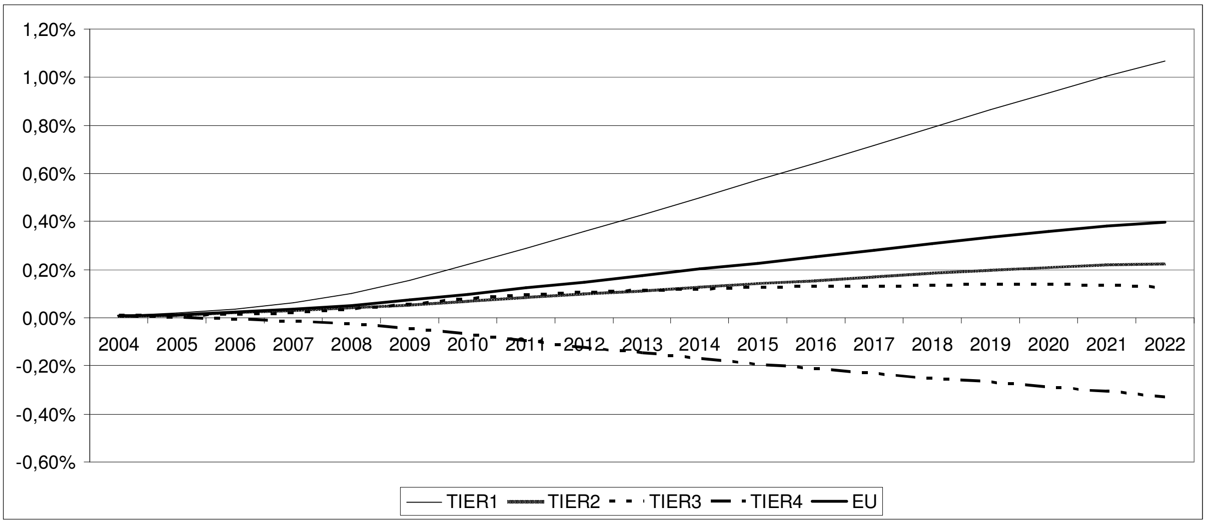


**Figure 6.** The effect of EU FP6 research support augmented with an annual 1 percent quality-oriented redistribution of national R&D expenditures, Euro-zone, 2003-2022: percentage differences between scenario and baseline values

extent of redistribution is purposefully small the simulation is capable of providing information about the trends for regions belonging to different agglomeration tiers as well as for the EU aggregate.

Figure 7 clearly shows that even a 1 percent redistribution of national R&D expenditures would imply significant changes in regional and macro impacts of EU FP6 research support. Tier 1 regions are definite winners of such a quality redistribution. By the end of the simulation period (2022) their GDP would increase by 1.07 percent which is about 20 percent higher than the FP6 impact would be without the quality redistribution of national R&D funds. The impact on Tier 2 and Tier 3 regions is slightly smaller whereas the negative effect on Tier 4 regions would almost double the size of the impact without quality redistribution. There is also a slight positive impact at the aggregate EU level: in 2022 GDP is higher with about 0.46 percent than it would be without the FP6 program.

### *4.2 Compensation for R&D specialization 1: regional human capital support*

The simulation in the previous sub-section clearly indicates that not every region is equally well-prepared for R&D-based development policies. Whereas Tier 1 regions absorb research funds in a more effective manner (due to high agglomeration of technological knowledge and their extensive interregional research collaboration networks) regions belonging to the rest of the tiers might need additional policy measures to catch-up. In this and the next sub-section the potential effects of the support of two intangibles are in the focus: the impacts of human capital development and the support of regional social capital.

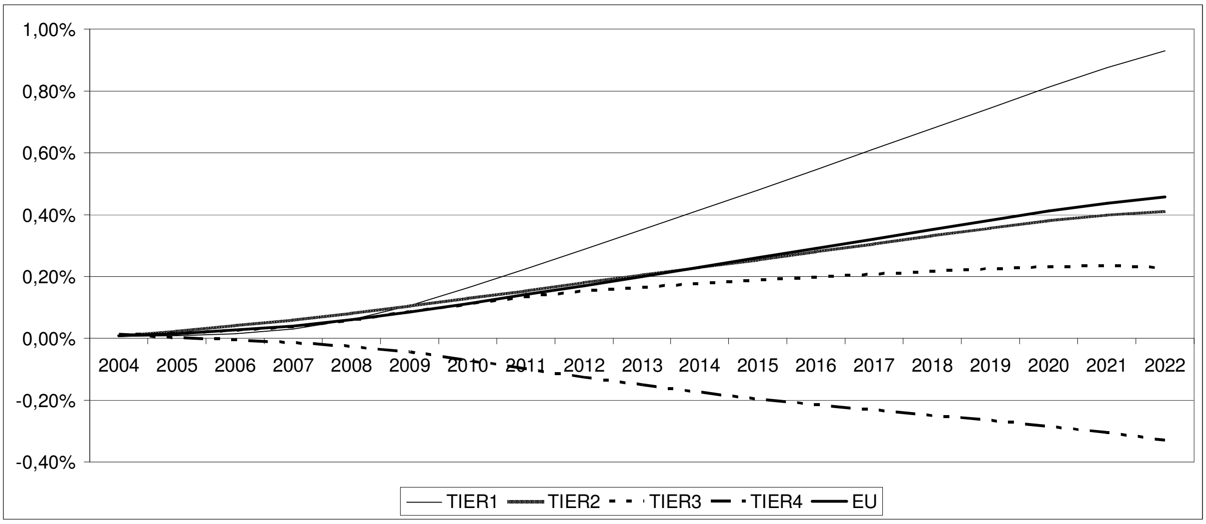


**Figure 7.** The effect of a 0.5 percent annual increase of human capital in Tier 2, 3 and 4 regions to compensate for the impact of the quality-oriented redistribution of national R&D expenditures, Euro-zone, 2003-2022: percentage differences between scenario and baseline values

To what extent regional human capital development is able to compensate the adverse effects of a quality redistribution of national R&D for relatively less developed regions? In this simulation the previously detailed policy mix of EU FP6 research support and a 1 percent quality redistribution of national R&D funds is extended with a 0.5 percent annual increase of human capital (that cumulates to an about 10 percent increase of regional human capital over the simulation period) in Tier 2, Tier 3 and Tier 4 regions. The impacts are depicted in Figure 8. Tier 2 and Tier 3 regions absorb human capital development in a very effective way: by the end of the study period the impact of FP6 is about two times higher in these regions than what it would be without the compensation for the quality redistribution of R&D. However for Tier 4 regions human capital development has a practically zero impact as compared to the results in Figure 7. The impact on GDP in the Euro-zone is about 10 percent higher when the policy mix of FP6 and regional quality distribution of R&D is extended by human capital development.

### *4.3 Compensation for R&D specialization 2: regional social capital development*

Though changing regional culture is perhaps the most challenging policy task it is interesting to speculate about the likely effects of social capital development. Figure 9 shows the impacts of a policy scenario where regional social capital is increased annually by 0.05 percent (which cumulates to an about 1 percent increase in social capital over the whole study period) in Tier 2, Tier 3 and Tier 4 regions. Though the targeted increase in social capital is small the results show that policies aiming at such development can be very powerful. Very similar to the results of the previous scenario Tier 2 and Tier 3 regions absorb social capital development in a very effective way: by the end of the study period the impact of FP6 is again about two times higher in these regions than what it was without the compensation. However for Tier 4 regions social capital development again has a practically zero impact as compared to the results in Figure 7. Similar to the findings of the previous scenario the impact on GDP in the Euro-zone is about 10 percent higher when the policy mix of FP6 and regional quality distribution of R&D is extended by social capital development.

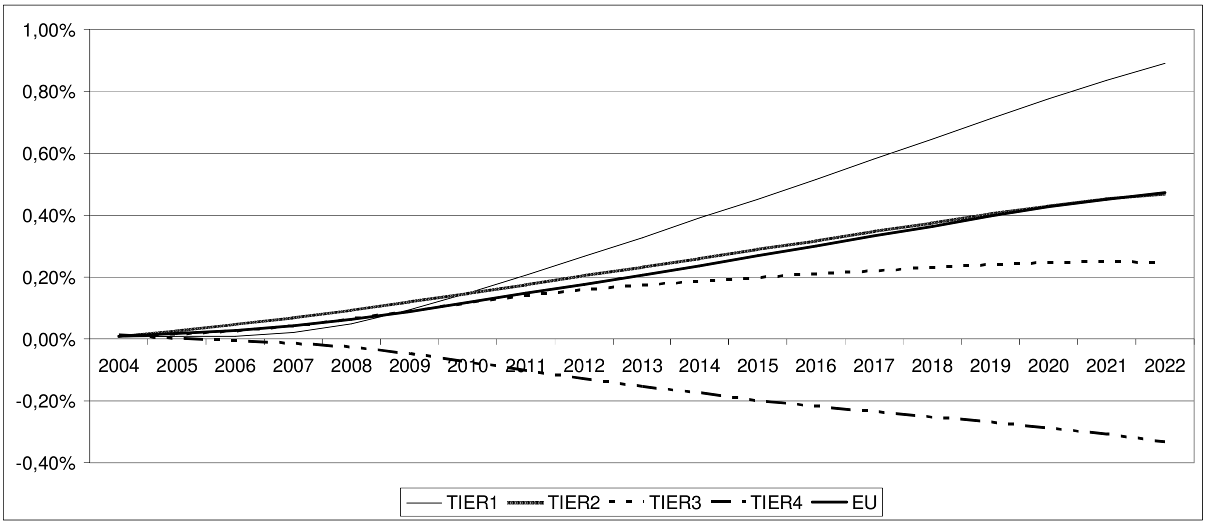


**Figure 8.** The effect of a 0.05 percent annual increase of social capital in Tier 2, 3 and 4 regions to compensate for the impact of the quality-oriented redistribution of national R&D expenditures, Euro-zone, 2003-2022: percentage differences between scenario and baseline values

### *4.4 Policy implications*

Policy analyses in the previous sub-sections lead to the following implications for regional policies aiming at supporting intangible assets in the forms of R&D, human capital and social capital.

- Compared to the relatively small share of EU Framework Program research support in Member States' R&D budgets regional and EU level economic impacts of FP6 expenditures are considerable. It suggests that this policy instrument is an effective tool not only for promoting scientific publication activities but also for supporting regional and macro level productivity and economic development.

- Redistributing R&D funds to regions where research productivity is the highest is a promising economic policy instrument in the hands of Member States. This instrument increases regional GDP in the most agglomerated regions as well as at the level of the European Union. However, as expected there is a small negative effect on regions with average development and a more adverse effect on lagging regions.

- There are policy instruments to compensate for the negative effects of specialization in the form of a spatial quality redistribution of R&D resources. Continuous regional human capital development can successfully overcompensate the adverse effects in regions where technological knowledge is about medium developed. There is also a considerable impact of regional human capital development on GDP at the macro level.

- Compensating for R&D specialization in the form of persistent social capital development is also a powerful tool for Member States to improve economic positions of regions with medium-level agglomeration of technological knowledge. This policy option results in a significant macro level GDP impact as well.

- It is clear from the policy analyses that EU regions where agglomeration of technological knowledge shows the lowest levels are not responsive to compensations in forms of either human capital or social capital development. These regions should be considered separately when local development policies are formed. They are not (yet) able to be the sites of future knowledge-based development. Instead, specific sectoral policies aiming at leisure or tourism would be more effective for those regions.

## References


Acs, Z., Anselin, L. and Varga, A. 2002 Patents and innovation counts as measures of regional production of new knowledge. *Research Policy* 31, 1069-1085

Anselin L 1988 *Spatial econometrics: methods and models*. Boston, Kluwer Academic Publishers

Anselin, L., Varga, A. and Acs, Z. 1997 Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics* 42, 422-448.

Bayar, A 2007 *Simulation of R&D Investment Scenarios and Calibration of the Impact on a Set of Multi-Country Models*. European Commission DG JRC Institute for Prospective Technological Studies (IPTS).

Bröcker, J. 1998 Operational spatial computable general equilibrium modeling. Annals of Regional Science 32, 367-87.

Bröcker, J. and Korzhenevych A. 2008 *Forward Looking Dynamics in Spatial CGE Modelling*. Unpublished manuscript, pp.17.

Dettori, B., Marrocu, E. and Paci, R. 2009 *Total factor productivity, intangible assets and spatial dependence in the European regions*. IAREG working paper WP5/03, June.

Donaghy, K 2009 CGE Modeling in space. In Capello, R and Nijkamp, P *Handbook of Regional Growth and Development Theories*. Edward Elgar Publishers, 389-422.

ESRI 2002 *An examination of the ex-post macroeconomic impacts of CSF 1994-1999 on Objective 1 countries and regions*. Dublin.

Fujita, M., Krugman, P. and Venables, A. 1999 *The Spatial Economy*. MIT Press.

Griliches Z 1990 Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 28, 1661-1707.

Ivanova, O., Heyndrickx, C., Spitaels, K., Tavasszy, L., Manshanden, W., Snelder, M., and Koops, O. 2007 *RAEM: version 3.0. Final Report*. Transport & Mobility Leuven, pp. 77.

Járosi, P., Koike, A., Thissen, M. and Varga, A. 2009 Regionális fejlesztéspolitikai hatáselemzés térbeli számítható általános egyensúlyi modellel: A GMR-Magyarország SCGE modellje. KRTI working papers 2009/3.

Jones C 2002 *Economic Growth*. W. W. Norton, London.

Krugman, P 1991 Increasing returns and economic geography. *Journal of Political Economy* 99, 483-499.

Oosterhaven J., Knaap T., Rijgrok C. and Tavasszy L. 2001 *On the development of RAEM: The Dutch spatial general equilibrium model and its first application to a new railway link*. Paper presented at the 41th Congress of the European Regional Science Association, Zagreb, August 29.

Polt W., Vonortas N., Fisher R., et al. 2008 Innovation Impact Study Final Report. European Commission DG Research.

Pontitakis D., Kyriakou D. and Bavel R. (Eds.) 2009 *The Question of R&D Specialisation. Perspectives and policy implications*. JRC Scientific and Technical Reports. European Commission Joint Research Centers Institute for Prospective Technological Studies, Seville.

Ratto, M., Roeger, W., Veld, J. 2009 QUEST III: An estimated open-economy DSGE model of the euro area with fiscal and monetary policy. *Economic Modelling*, 26, 222-233.

Romer P 1990 Endogenous technological change. *Journal of Political Economy* 98, S71-S102.

Schalk, H. and Varga, A. 2004 *The economic effects of EU Community Support Framework interventions. An ex-ante impact analysis with EcoRET, a macroeconomic model for Hungary*. Center of Applied Economic Research Münster (CAWM), University of Münster, Münster.

Shoven, J. B. and J. L. Whalley 1992 Applying General Equilibrium. Cambridge University Press, Cambridge.

Smets, F. and Wouters R. 2003 An estimated stochastic dynamic general equilibrium model of the euro area. *Journal of the European Economic Association*, Vol 1, No. 5, pp 1123-1175.

Thissen M 2003 RAEM 2.0 *A regional applied general equilibrium model for the Netherlands*. TNO working papers, pp 19.

Varga, A 2000 Local academic knowledge transfers and the concentration of economic activity**.** *Journal of Regional Science* 40, 289-309.

Varga, A. 2006 The spatial dimension of innovation and growth: Empirical research methodology and policy analysis. *European Planning Studies* 9, 1171-1186.

Varga, A. 2007 *GMR-Hungary: A complex macro-regional model for the analysis of development policy impacts on the Hungarian economy*. PTE KTK KRTI working papers 2007/4.

Varga, A. 2008 From the Geography of Innovation to Development Policy Analysis: The GMR-approach. *Annales d'Economie et de Statistique* 87-88, 83-102.

Varga, A. and Schalk, H. 2004 Knowledge spillovers, agglomeration and macroeconomic growth. An empirical approach. *Regional Studies* 38, 977-989.

Varga, A., Schalk, H., Koike, A, Járosi, P., Tavasszy, L. 2008 *Integrating the effects of geography into EU Cohesion Policy impact modeling: The GMR-approach*. Paper presented at the modeling workshop of the German Section of the European Regional Science Association.

Varga, A., Pontikakis, D., Chorafakis, G. 2010 Aggomeration and network effects in regional R&D productivity. Working Paper IAREG WP5/25.

Veld Jan in't 2007 Model-based ex-ante evaluations of the impact of the Cohesion Policy programme 2007-20015: A comparison of QUEST and HERMIN model results. European Commission DG ECFIN.

Venables, A. J. and Gasiorek, M 1999 *The Socio-Economic Impact of Projects Financed by the Cohesion Fund: AModeling Approach*. Luxembourg: European Commission.

## APPENDIX

### Table A1. Regression Results for Log (Patents) for 189 EU regions, 2000-2002 (N=567)

| Model | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Estimation | OLS | OLS | OLS | OLS | OLS | 2SLS- Spatial Lag (INV2) |
| Constant | -1.6421*** | -0.3107 | -0.5391* | -1.7864*** | -1.7227*** | -2.3006*** |
| | (0.1776) | (0.2316) | (0.2806) | (0.2381) | (0.2372) | (0.2743) |
| W_Log(PAT) | | | | | | 0.2455*** |
| | | | | | | (0.0631) |
| Log(GRD(-2)) | 1.0822*** | 0.8453*** | 0.9585*** | 0.7142*** | 0.6879*** | 0.7088*** |
| | (0.0308) | (0.0407) | (0.0886) | (0.0377) | (0.0384) | (0.0377) |
| Log(GRD(-2))*Log(δ(-2)) | | 0.3242*** | 0.3222*** | 0.2443*** | 0.2136*** | 0.1439*** |
| | | (0.0389) | (0.0389) | (0.0351) | (0.0363) | (0.0396) |
| Log(GRD(-2))*Log(NETRD(-2)) | | | -8.675E-05 | | | |
| | | | (6.03E-05) | | | |
| Log(PSTCK(-2)) | | | | 0.2502*** | 0.2536*** | 0.1804*** |
| | | | | (0.0203) | (0.0202) | (0.0272) |
| PAHTCORE | | | | | 0.4814*** | 0.4614*** |
| | | | | | (0.1568) | (0.1526) |
| $R^2$-adj | 0.69 | 0.72 | 0.72 | 0.78 | 0.78 | |
| Log Likelihood | -885.30 | -852.36 | -851.32 | -784.69 | -779.98 | |
| Sq. Corr. | | | | | | 0.80 |
| Multicollinearity Condition Number | 7 | 10 | 24 | 13 | 13 | |
| F on pooling (time) | 0.9071 | 0.6777 | 0.5644 | 0.8143 | 0.6425 | |
| F on slope homogeneity | 0.4815 | 0.7613 | 0.5836 | 0.6485 | 0.4645 | |
| White test for heteroscedasticity | 0.7529 | 1.0462 | 12.8409 | 3.6634 | 12.1852 | |
| LM-Err | | | | | | |
| Neighb | 111.78*** | 69.36*** | 66.85*** | 26.95*** | 23.46*** | |
| INV1 | 252.17*** | 129.64*** | 117.26*** | 29.87*** | 26.13*** | |
| INV2 | 215.12*** | 121.59*** | 114.45*** | 32.40*** | 29.24*** | |
| LM-Lag | | | | | | |
| Neighb | 142.53*** | 100.88*** | 99.03*** | 24.99*** | 25.89*** | |
| INV1 | 247.03*** | 159.07*** | 153.47*** | 28.16*** | 27.96*** | |
| INV2 | 237.99*** | 148.93*** | 145.48*** | 31.42*** | 30.95*** | |

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; W_Log(PAT) is the spatially lagged dependent variable where W stands for the weights matrix INV2. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$. In model (6) the Durbin-Wu-Hausman test for Log(GRD(-2)) and Log(GRD(-2))* Log(δ(-2)) does not reject exogeneity. The instruments were selected following the 3-group method. For the spatial lag term the instruments are the spatially lagged explanatory variables.

## Table A2. Regression Results for Log (Publications) for 189 EU regions, 2000-2002 (N=567)

| Model | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Estimation | OLS | OLS | OLS | OLS | OLS | 2SLS Heteroscedasticity Robust |
| Constant | 1.4026*** | 2.3886*** | 2.196*** | 2.3395*** | 2.4568*** | 2.6137*** |
| | (0.1298) | (0.1645) | (0.202) | (0.1711) | (0.1697) | (0.3199) |
| Log(GRD(-2)) | 0.942*** | 0.445*** | 0.480*** | 0.4158*** | 0.4523*** | 0.4317*** |
| | (0.0225) | (0.0597) | (0.633) | (0.066) | (0.0602) | (0.1262) |
| Log(GRD(-2))*Log(δ(-2)) | | | -0.0462 | | | |
| | | | (0.0282) | | | |
| Log(GRD(-2))* WFP5_Log(RD(-2)) | | 0.0004*** | 0.0004*** | 0.0004*** | 0.0004*** | 0.0003*** |
| | | (4.40E-05 ) | (4.40E-05) | (4.56E-05) | (4.68E-05) | (9.26E-05) |
| Log(PSTCK(-2)) | | | | 0.01758 | | |
| | | | | (0.01689) | | |
| PUBCORE | | | | | 0.2247** | 0.3293*** |
| | | | | | (0.1032) | (0.0977) |
| $R^2$-adj | 0.76 | 0.79 | 0.79 | 0.79 | 0.79 | |
| Log Likelihood | -707.30 | -670.05 | -668.70 | -669.51 | -667.89 | |
| Sq. Corr. | | | | | | 0.79 |
| Multicollinearity Condition Number | 7 | 22 | 23 | 27 | 24 | |
| F on pooling (time) | 0.6694 | 0.9269 | 0.6712 | 0.7141 | 0.7055 | |
| F on slope homogeneity | 0.2059 | 0.357 | 0.2752 | 0.2683 | 0.2501 | |
| White test for heteroscedasticity | 44.575*** | 77.378*** | 84.013*** | 92.231*** | 86.884*** | |
| LM-Err | | | | | | |
| Neighb | 0.7199 | 0.7727 | 0.7518 | 0.9808 | 0.5749 | |
| INV1 | 3.3586* | 2.5407 | 1.8767 | 3.4006* | 2.6595 | |
| INV2 | 0.3687 | 0.9367 | 0.8782 | 1.2604 | 1.020 | |
| LM-Lag | | | | | | |
| Neighb | 12.214*** | 3.0067* | 2.4689 | 4.2311** | 3.7861* | |
| INV1 | 1.6479 | 0.0642 | 0.4640 | 0.061 | 0.0069 | |
| INV2 | 5.2928** | 0.6649 | 0.1242 | 1.9522 | 1.1352 | |

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neigh is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$. In Model 5 the Durbin-Wu-Hausman test for Log(GRD(-2)) and Log(GRD(-2))* Log(NETRD(-2)) rejects exogeneity at the level of $p < 0.1$. In Model 6 the instruments were selected following the 3-group method.

**Table A3. Regression Results for (GRD2001-GRD1998) for EU regions (N=189)**

| Model | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Estimation | OLS | OLS | OLS | OLS-Heteroscedasticity Robust (White) |
| Constant | -604.429*** | -735.41*** | -299.107*** | -299.107*** |
| | (90.8252) | (101.405) | (78.3494) | (68.7176) |
| BETAPAT1998 | 1145.6*** | 910.258*** | 351.824*** | 351.824*** |
| | (147.511) | (167.819) | (125.294) | (118.165) |
| BETAPUB1998 | | 364.853*** | 190.322** | 190.322*** |
| | | (131.181) | (93.4943) | (69.8948) |
| RDHCORE | | | 360.98*** | 360.98*** |
| | | | (26.3212) | (47.4151) |
| $R^2$-adj | 0.24 | 0.27 | 0.63 | 0.63 |
| White test for heteroscedasticity | 52.3206*** | 57.8899*** | 42.2263*** | |
| LM-Err | | | | |
| Neighb | 0.1133 | 0.0231 | 0.0674 | |
| INV1 | 0.0092 | 0.1976 | 1.1476 | |
| INV2 | 0.0895 | 1.8205 | 0.9415 | |
| LM-Lag | | | | |
| Neighb | 0.0960 | 0.0434 | 0.1026 | |
| INV1 | 2.6971 | 0.9635 | 1.9972 | |
| INV2 | 0.5956 | 0.5309 | 1.9896 | |

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: Neighb is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix. *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.

**Table A4. Regression Results for (EMPKI2001-EMPKI1998) for EU regions (N=189)**

| Model | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Estimation | OLS | OLS | OLS | ML – Spatial Error (INV2) with Heteroscedasticity weights |
| Constant | 5399.78* | 8821.36*** | 9955.96*** | 11168.3*** |
| | (3032.61) | (3314.62) | (3267.78) | (2879.48) |
| EMPKI1998 | 0.071*** | 0.054*** | 0.032*** | 0.0262** |
| | (0.006) | (0.009) | (0.012) | (0.011) |
| EMPKI1998*GRD1998 | | 3.788E-06** | 5.043E-06*** | 5.624E-06*** |
| | | (1.582E-06) | (1.604E-06) | (1.604E-06) |
| RDCORE | | | 19896.5*** | 21321.1*** |
| | | | (6614.64) | (6366.96) |
| LAMBDA | | | | -0.0181** |
| | | | | (0.009) |
| $R^2$-adj | 0.41 | 0.42 | 0.45 | 0.45 |
| Multicollinearity Condition Number | 2 | 4 | 6 | |
| White test for heteroscedasticity | 27.37*** | 28.182*** | 34.522*** | |
| LM-Err | | | | |
| Neighb | 0.922 | 0.164 | 0.042 | |
| INV1 | 0.052 | 0.023 | 0.28 | |
| INV2 | 1.008 | 3.263* | 5.878** | |
| LM-Lag | | | | |
| Neighb | 2.181 | 1.846 | 1.916 | |
| INV1 | 0.479 | 0.043 | 0.645 | |
| INV2 | 4.000* | 4.574** | 4.316** | |

Notes: Estimated standard errors are in parentheses; spatial weights matrices are row-standardized: LAMBDA is the spatial autoregressive coefficient; Neighb is neighborhood contiguity matrix; INV1 is inverse distance matrix; INV2 is inverse distance squared matrix; *** indicates significance at $p < 0.01$; ** indicates significance at $p < 0.05$; * indicates $p < 0.1$.