

Lock-in Effects of EU R&D Spending on Regional Growth. A Non-parametric and Semi-parametric Conditional Quantile Regressions Approach.

Antonio Acconcia^a, Marta Espasa^b, Leone Leonida^{c,d} and Daniel Montolio^{b,d,}*

Abstract. The purpose of this paper is twofold. First, we study the allocation of European Union (EU) expenditure in Research and Development (R&D) across European regions. Second, we focus on the effects of this variable on regional per capita GDP levels, and on regional growth rates. Using non-parametric and semi-parametric conditional quantiles, we found empirical evidence in favour of different effects of R&D expenditure among conditional quantiles of the per capita income distribution, and of the growth rates distribution. Moreover, we find a “lock-in effect” of R&D spending. A positive relation between growth rates and this component of the EU expenditure is estimated for regions with higher growth rates, with these regions tending to have a higher and common growth rate as R&D expenditure increases. Furthermore, slow growth regions seem to approach to a common but lower growth rate. The estimates relative to the relationship between the per capita regional GDP and the R&D spending confirm these findings.

JEL: C14, O40, O52

Key Word: EU Budget, R&D Expenditure, Growth Rates, Conditional Quantiles.

^a Università Federico II di Napoli (Italy)

^b Universitat de Barcelona (Spain), and Institut d’Economia de Barcelona (Spain)

^c Università Degli Studi Della Calabria (Italy)

^d University of York (England)

* Corresponding author. E-mail: dm127@york.ac.uk

1. Introduction

During the last decade, one of the most hotly debated subjects in economics has been the redistributive capacity of the European Union (EU, hereafter) budget. Among the numerous arguments in favour of an active intervention, the following stands out: its power to mitigate horizontal equity problems due to the intervention of national governments (Davezies-Nicot-Prud'Homme, 1996); its capacity to reduce territorial income disparities, or disparities derived from the integration process¹; its ability to guarantee the existence of the EU itself (Cremer and Pestieau, 1996); and its power to mitigate the negative effects originated by possible asymmetric shocks generated by the European Monetary Union^{2,3}.

This last argument is, indeed, one of the most important since the costs and benefits of economic and monetary integration may not be equally distributed across European regions; it is possible that less developed regions will receive fewer benefits from the integration process. Therefore, it is of crucial importance to design European policies directed to reduce such disparities, and to promote equality of opportunities in the territory. Indeed, if this is not achieved, the process of economic and monetary integration itself can be at risk⁴.

Different studies dealing with the territorial cohesion in Europe highlight that income inequalities inside the EU are very pronounced, especially across regions. Moreover, European income disparities, in terms of per capita Gross Domestic Product (GDP), are more accentuated at a regional level than at a country level. If per capita income across countries seems to have converged (especially across poor countries), the same pattern is not observed across European regions, either if the EU is taken as a whole or inside each of the country members of the Union. It seems that poor European regions do not tend to completely converge with rich regions (see Terrasi, 2000).

The main instrument of the EU to lead to an equal territorial distribution of income is the European budget. Redistributive policies can be implemented either by means of specific instruments, or by means of the overall budget, given that any public intervention (revenues or expenditure) has distributive effects although the interventions are not explicitly planned to fulfil this aim. This is the case of the Research and Development (R&D, hereafter) expenditure, which

¹ Among the numerous studies in favour of this argument, the most important are the Cecchini Report (1988), Padoa-Schioppa Report (1987), and Emerson et al. (1992).

² The Monetary Union supposes not only the transfer of monetary policy to the European Union, but also the existence of substantial limits on fiscal policy established through Stability programs, which reduce national fiscal autonomy.

³ See, among others, Sala-i-Martin and Sachs (1992), Krugman (1993), and Goodhard and Smith (1993).

⁴ The Maastricht Treaty signed in 1992 included the economic and social cohesion as a one of the most important aims to be fulfilled. Moreover, Article 130 of the Maastricht Treaty specifies, as a function of the EU, the reduction of the disparities across regions, and the development of the more depressed areas.

does not have a direct redistributive purpose, but does have important distributional effects. Moreover, these effects can be enlarged by the direct link between this kind of spending and economic growth.

This is the context of the present paper. Our aim is twofold. First, we study the allocation of EU expenditure in R&D across European regions. Second, we focus on the effects of this variable on the regional per capita GDP level and on regional growth rates.

Using a well-known model (Segerström, 2000) and reporting the main stream of the current debate on the effect of R&D spending on the growth rate, we show that the theoretical results are controversial; this calls for an empirical investigation, possibly free from any theory-induced constraints.

Next, we assign the EU budget to European regions, analysing its main features. We found that it has been mainly used to develop depressed areas: regions with a lower level of per capita income receive more European funds. However, the distribution of the R&D spending follows a different pattern: rich regions receive more R&D spending than poor regions.

Finally, using a semi-parametric conditional quantiles regression approach to uncover the effect of the R&D spending on both the per capita GDP level and on its growth rate, we found empirical evidence in favour of different effects among conditional quantiles of the per capita income distribution, and of the growth rates distribution. Moreover, we find a “lock-in effect” of R&D spending. A positive relation between growth rates and this component of the EU expenditure is estimated for regions with higher growth rates, with these regions tending to have a higher and common growth rate as R&D expenditure increases. Furthermore, slow growth regions tend to approach to a common but lower growth rate. These findings are partially in contrast with previous studies (Boldrin and Canova, 1997).

The estimates relative to the relationship between the per capita GDP and the R&D spending confirm these findings. In general, the “pure” effect of the R&D spending on the per capita GDP level seems to be positive. Moreover, for high levels of this type of EU expenditure, the variance of the European regional per capita GDP distribution would be reduced. However, the distribution would show polarisation of income.

The rest of paper is organised as follows. Section 2 briefly describes previous literature dealing with the theoretical and empirical evidence of the effect of the R&D expenditure on economic growth, showing the reason why the answer is essentially empirical. Section 3 describes the regional imputation process of the EU expenditure in R&D, and the main characteristics of its allocation. Section 4 reports on both the methodology employed, and on the estimated effects of the R&D spending on the per capita GDP level and on the growth rates of European regions. Finally section 5, summarises the main findings of this work.

2. The Effect of R&D Spending on Economic Growth. Theoretical Aspects.

The economy represented by R&D-driven endogenous growth models is usually based on three common features. A (final) sector where perfectly competitive firms produce consumption goods by means of intermediate products; an intermediate sector composed of a large number of monopolistic industries where firms produce the intermediate products; research laboratories producing vertical and/or horizontal innovations.

In earlier models, growth arises either by increasing the number of intermediate products, that is through horizontal innovations (Romer, 1990), or by increasing the quality of a given set of intermediate products, that is through vertical innovations (Grossman and Helpman, 1991; Aghion and Howitt, 1992). In these models, the government can play an active role through growth enhancing R&D expenditure, as the growth rate of an economy is directly related to the (relative) amount of resources devoted to R&D. Moreover, R&D subsidies, by encouraging firms to devote more resources to R&D activities, tend to increase the long-run rate of economic growth. Recent papers allow for both types of innovations at the same time, in order to eliminate the so-called “scale-effect” which was a feature of previous models (Howitt, 1999; Segerström, 2000). The main conclusion is that the positive effect of subsidies on growth of earlier models is not robust; in fact, R&D subsidies can either promote or retard long-run economic growth. The latter calls for an empirical investigation.

At any date t , firms under perfect competition produce consumption goods and R&D services by a continuum of intermediate products, exploiting the same production function. Specifically the total output of the economy is

$$Y_t = C_t + V_t + H_t = L_{y,t}^{1-\alpha} \int_0^{N_t} A_{i,t} X_{i,t}^\alpha di \quad [1]$$

where Y_t is gross output, C_t is consumption, V_t is vertical R&D expenditures, H_t is horizontal R&D expenditures, $L_{y,t}$ is the input of labour employed to producing output, N_t denotes of how many different intermediate industries exist, $X_{i,t}$ is the flow of intermediate product i , and $A_{i,t}$ is a productivity parameter relative to the latest version of the intermediate product i . Each intermediate product is produced exploiting the linear production function $X_{i,t} = L_{i,t}$, where $L_{i,t}$ is the amount of labour in industry i . Other things equal, total output equation implies that growth can be driven by increases in N_t and/or $A_{i,t}$.

In the simplest horizontal R&D driven endogenous growth model, a deterministic R&D process is assumed such that \dot{N}_t is proportional to the amount of output devoted to R&D, $\dot{N}_t = (1/\mu)H_t$. It follows that $\dot{Y}/Y = \dot{N}/N = H/\mu N$. A subsidy to research, that is a policy such that the government absorbs part of the cost of research for a potential inventor, raises the growth rate.

In the second class of R&D driven endogenous growth models instead, the number of industries is constant while the productivity parameters, $A_{i,t}$, increase with innovations. A research sector for each product i with Poisson arrival rates of innovations is a main feature. Moreover, intersectoral spillover applies: at any date t the economy is indexed by the leading-edge technology which is determined by all sector specific innovations (Aghion and Howitt; 1998). Each vertical innovation in sector i permit firms to produce in sector i using the leading-edge technology, which grows at a rate proportional to R&D expenditure. This implies that the growth rate is proportional to aggregate R&D expenditure; again, the effect of R&D subsidy on growth is positive.

Jones (1995) challenges previous models noting that the large increase in the number of scientist and engineers engaged in R&D since 1950 in advanced countries did not induce a trend in growth rates. Thus, the main conclusion that the long run growth rate is related to the amount of resources devoted to R&D would not adequately fit the data. Furthermore, he finds that in an R&D based model of endogenous growth without the scale effect, R&D subsidies do not have long run growth effects.

Recently, R&D driven endogenous growth models with both horizontal and vertical innovations have been proposed. In the version outlined by Howitt (1999), the scale effect property disappears while other implications of the model are the same as in the earlier models. In particular, R&D subsidies promote long-run economic growth. Howitt (1999) assumes that vertical innovations, that is improvements in the productivity or quality parameter relative to sector i , arrive stochastically following a Poisson process with an arrival rate linear in the amount of R&D expenditure in sector i (as in Aghion and Howitt, 1998). The outcome of research aimed at vertical innovations is composed of a sector specific component and a public good component. The latter creates intersectoral spillover effects inducing growth in the leading-edge parameter. At the same time, horizontal innovations are produced under decreasing returns to R&D expenditure. Each horizontal innovation permits a given firm to produce a new intermediate product with a productivity parameter, which is drawn randomly from the distribution of existing productivity parameters. Increments in intermediate industries destroy the scale effect of previous models retaining the propriety that R&D subsidies have a positive effect on growth.

In a generalised version of Howitt's (1999) model, however, the positive effect of subsidies on growth is not robust. In particular, with more general assumptions about the returns to horizontal

and vertical R&D activities and about how the returns to both activities change over time, Segerström (2000) shows that R&D subsidies can either promote or retard long-run economic growth.

Following the basic Segerström model features, firms engage in both vertical and horizontal R&D activities. Let $A_t \equiv \max\{A_{i,t}; i \in [0, N_t]\}$ denote the leading-edge productivity parameter at time t . Each vertical innovation in sector i permits monopolistic firm to produce using the leading-edge technology. Vertical innovations arrive at Poisson rate, which is directly related to both R&D expenditure flow and firm specific knowledge that is useful for vertical innovation, and inversely related to the leading-edge productivity parameter. As the reward for innovating in a given sector is proportional to A_t the Poisson arrival rates of innovations in different sectors are independent of each other. In particular,

$$\dot{A}_t / A_t = \frac{\lambda_v (V_t)^\delta (Y_t)^{1-\delta}}{N_t (A_t)^d} = \lambda_v (V_t / Y_t)^\delta (Y_t / N_t A_t^d) \quad [2]$$

The parameter $d > 1$ implies that as the leading-edge parameter increases over time research problems become more harder to solve and this depresses the growth rate; $\delta < 1$ measures the degree of diminishing returns to vertical R&D expenditures; division by N_t means that as the number of intermediate industries increase each vertical innovation has a smaller spillover effect on aggregate economy; finally the term Y_t captures the idea that, apart from R&D expenditure, innovation depends on firm specific knowledge that is useful for vertical innovations, the latter depending on the level of output; $\lambda_v > 0$. Horizontal innovations result from R&D aimed at creating new products. At any time t each innovation results in a new intermediate variety whose productivity parameter is drawn randomly from the existing distribution of $A_{i,t}$ across industries. The main implication of this assumption is that the distribution of $A_{i,t} / A_t$ converges monotonically to an invariant distribution. The growth rate of horizontal innovations is governed by an equation similar to [2]:

$$\dot{N}_t / N_t = \frac{\lambda_h (H_t)^\gamma (Y_t)^{1-\gamma}}{N_t (A_t)^d} = \lambda_h (H_t / Y_t)^\gamma (Y_t / N_t A_t^d) \quad [3]$$

Profit maximisation by competitive firms producing Y implies $w_t = (1 - \alpha) \int_0^{N_t} A_{i,t} (X_{i,t} / L_{y,t})^\alpha di$ and $p_{i,t} = \alpha A_{i,t} (L_{y,t} / X_{i,t})^{1-\alpha}$. The latter denotes a constant elasticity (inverse) demand for each intermediate product that, in turn, implies that each incumbent monopolist in industry i charge the standard monopoly mark-up over marginal cost, $p_{i,t} = w / \alpha$.

Substituting the three first order conditions in the output equation yields $Y_t = \alpha^{2\alpha} (1-\alpha)^{-1} (1-\alpha+\sigma)^{1-\alpha} L_{y,t} A_t N_t^{1-\alpha}$. Taking logs of both sides and differentiating with respect to time yields:

$$\dot{Y}_t / Y_t = \dot{A}_t / A_t + (1-\alpha) \dot{N}_t / N_t \quad [4]$$

The economy growth rate depends both on the growth rate of intermediate varieties and on the growth rate of the leading-edge productivity parameter.

Both horizontal and vertical R&D expenditures are subsidised at the proportional rate s . The level of R&D (both horizontal and vertical) is such that the marginal cost of R&D expenditures $1-s$ equals the expected marginal benefit. By assuming that vertical R&D races are perfectly competitive and symmetric R&D firms, first order conditions imply

$$\frac{\lambda_v \delta \Pi_{v,t}}{(A_t)^d} (V_t / Y_t)^{\delta-1} = 1-s \quad \text{and} \quad \frac{\lambda_h \gamma \Pi_{h,t}}{(A_t)^d} (H_t / Y_t)^{\gamma-1} = 1-s \quad [5]$$

where $\Pi_{v,t}$ ($\Pi_{h,t}$) is the expected value of vertical (horizontal) innovation. The assumption of decreasing returns to R&D determines downward sloping marginal benefit curves. Moreover, as the marginal cost of R&D is constant over time in equilibrium the marginal benefit must be constant too.

In a balanced growth equilibrium the fractions of GDP allocated to R&D, that is H/Y and V/Y , and Y/NA^d are constant over time. The latter and the labour market clearing condition imply

$$\dot{L} / L = (d-1) \dot{A} / A + \alpha \dot{N} / N \quad [6]$$

where \dot{L} / L is the exogenous growth rate of labour. Equation [6] determines an inverse relationship between \dot{A} / A and \dot{N} / N , given the growth rate of labour. Moreover, substituting in equation [4] at the steady state from equation [6], it follows $\partial(\dot{Y}/Y) / \partial(\dot{N}/N) > 0$ when $(1-\alpha)d > 1$ and $\partial(\dot{Y}/Y) / \partial(\dot{N}/N) < 0$ when $(1-\alpha)d < 1$. Thus, an increase in the steady state growth rate of intermediate varieties, that is decrease in the steady state growth rate of leading-edge technology, can induce an increase or a decrease in the economy growth rate.

This is one of the main results in Segerström (2000). Moreover, Segerström (2000) show that a permanent increase in the R&D subsidy rate s can either promote or retard long run economic growth. According to equations [5], the initial effect of an increase in s is to induce more resources to both vertical and horizontal R&D, determining an increase in innovation rates. When $\gamma < \delta$ the

R&D subsidy increase determines a raise in the share of GDP devoted to vertical R&D to a greater extent than in the share of GDP devoted to horizontal R&D, and the leading-edge productivity parameter growth rate jumps up more than the variety growth rate.

However, after the initial increase the negative effect of increasing complexity of R&D problems over time at a faster rate than usual depresses innovations rates. When $(1-\alpha)d < 1$, the subsidy rate increase raises \dot{A}/A , lowers \dot{N}/N , and it leads the economy towards a new long-run balanced growth equilibrium with a higher growth rate of total output \dot{Y}/Y ; on the contrary, when $(1-\alpha)d > 1$ the subsidy rate increase raises \dot{A}/A , lowers \dot{N}/N , and it leads the economy towards a lower growth rate of total output \dot{Y}/Y .

This brief overview of the present debate should not be intended as exhaustive of all alternative models and ideas about this important issue. The point we want to raise here is that the effect of R&D expenditures and subsidies on growth rates is controversial and far from being clear on theoretical grounds; this is especially true when subsidies come from a public institution, as in the case we are going to analyse. This crucial point calls both for an empirical investigation of the issue, and gives credit to the particular approach, free from theory-induced constraints, that we are going to use in the remaining parts of the paper.

3. Territorial Assignment of the EU R&D Expenditure: Methodology and Description

The Annual Reports of the Court of Auditors provides information concerning expenditure in the member states. However, regional information is only available for expenditure on Structural Actions. Therefore, to shed light into the empirical effect of EU expenditure in R&D on regional income and growth rates, the first step is to establish some hypotheses to determine the assignment of the remaining regional expenditures.

In this section, we will briefly outline the criteria used to assign European R&D spending in 1995. However, in this study, we make use of the regional assignment of other types of European expenditure (Structural Actions, European Agricultural Guidance and Guarantee Fund (EAGGF-Guar), Exterior Actions, and Interior Actions), for more details about the definition of each type of spending and the assignment criteria, see Espasa (2000).

The data provided by the European Court of Auditors in 1999 shows that the EU spent 2.574,9 million Euros in R&D policies (3.21% of the overall European budget). For 1995, our year of interest, the R&D spending was 3.065,5 million Ecus (3,8% of the EU budget). Although the R&D spending represents a relatively small percentage in the overall EU budget, it is important to notice that the research and technological development has been one of the fields that has received

more attention in recent years, specially after the European Act and the Maastricht Treaty. The main purpose of the EU is to reinforce the scientific and technological bases of the European industries to make them more competitive.

Among the projects financed by the EU budget in R&D, there is the investment in the Common Centre for Research (CCR)⁵, a centre that belongs to the EU and is used to undertake research, and to finance the Research Framework Programmes in R&D.

Depending on the execution and financing entities of the projects, we can distinguish between three types of European actions in R&D:

- Direct Actions: research undertaken in the Common Centre of Research (CCR).
- Indirect or Shared Actions: The most common modality of research projects. The EU co-finance research projects in which research groups from different member states participate, and mainly universities, research centres or firms form these groups.
- Coordinated or Concerted Actions: specific programs of research develop by research groups or universities of the member states; the EU only compensates the coordination expenditures of these programs.

10% of the European budget in R&D in 1995 was devoted to the CCR, the rest to the projects inside the Research Framework Programme in Research and Development⁶.

The R&D spending has been assigned to the different member states by the Annual Reports of the Court of Auditors. However, there is a fraction of this expenditure not assigned to any country. Therefore, the percentage of spending assigned is 82,3% in 1995. Looking at the assignation by countries, the country that received the most funds in R&D was United Kingdom with the 20%, followed by France (17%) and Germany (16%). In a second group, there are countries such as Belgium (9%), Netherlands (7%) and Spain (6%). The remaining European countries received less than the 3%.

The regional assignation of R&D spending has been elaborated from the country specific assignation of this expenditure⁷. Moreover, inside each country the regional assignation has

⁵ The CCR is composed of four centres, located in Ispra (Italy), Kalsruhe (Germany), Petten (Netherlands) and Geel (Belgium).

⁶ Our year of interest, 1995, belonged to the Fourth Research Framework Programme in Research and Development (1994-1998).

⁷ The results have been elaborated under the monetary flow approach, rather than the benefit approach.

followed the specific regional spending in R&D (by private and public sector, and universities). This has been the criterion of assignation for France, Italy, Belgium, United Kingdom, Greece, Spain, Finland and Austria⁸.

In Sweden and Germany, the previous criterion is not available, therefore we have used, as a proxy variable the regional distribution of the research personnel (in the private and public sectors, and in universities).

For the assignation of the regional EU R&D spending in Netherlands and Portugal, countries for which we could not dispose of any of the previous variables, we have followed the regional GDP, due to the positive relation between the level of development of a region (indicated by the GDP) and the capacity to undertake projects of research and development.

Finally, the imputation of the expenditure not assigned to any country has followed the criteria previously explained.

The next sub-section will be devoted to the study of the main characteristics of the European regional spending for all the types of expenditure used in this paper.

3.1. Data description

In this sub-section, we describe and analyse the main results of the assignation of the EU budget to European regions. A non-parametric approach is used to study the shape of the distribution for each type of expenditure, focusing in particular on the main conclusions that can be drawn for the analysis of the main variable of interest in this paper: the European R&D spending.

The variable used to represent the income level of European regions is the per capita Gross Domestic Product (GDP) for 1995, for the 121 European regions⁹. This variable is one of the more comparable indexes across different economies and time. Moreover, per capita GDP is a common measure of wealth for an area. The average growth rate has been calculated between 1995 and 1998. We use the country inflation to deflate the series for the regions¹⁰.

Figure 1 reports the estimates of the distributions¹¹ for the different types of expenditure, both for the whole population of regions, and for two sub-samples of them. For instance, the first panel of Figure 1 reports the kernel density estimate of the total EU expenditure (full line), the distribution for the same variable for regions with a per capita income below the European average

⁸ Denmark, Ireland and Luxembourg have been considered as a whole.

⁹ We exclude from the data set the regions for which the European R&D expenditure variable was zero: Aland (Finland), Corse and Departements Doutre-Mer (France), and Ceuta and Melilla (Spain).

¹⁰ The main data source is the "Statistics in Focus" from Eurostat for various years.

¹¹ In our estimates, we always use a Gaussian Kernel and the Jones' rule for the optimal bandwidth because they are among the most common and less controversial. See, Silverman (1986).

per capita GDP (dotted line), and finally the density relative to the regions with a per capita income above the average (dashed line). Similarly, the first panel of Figure 2 reports the kernel density estimate of the total EU expenditure (full line), the distribution for the same variable for regions with a growth rate below the European average (dotted line), and finally the density relative to regions with a growth rate above the European average (dashed line).

The estimated distribution for the overall EU spending is unimodal (see the first panel in Figure 1). Moreover, the total EU budget has been mainly received by regions with a per capita GDP below the European average, reinforcing the idea that EU expenditure has redistributive functions. This first impression is confirmed by looking at the densities of the single components of the budget. EU intervenes extensively in the poorer part of the continent using the spending in Structural and Exterior actions. The same pattern, even if less pronounced, can be found by looking at the distribution for EAGGF-Guarantee. Interestingly, the only type of expenditure with a different pattern is the R&D spending.

Although the R&D spending distribution is unimodal, rich regions received much more R&D subsidies than poor regions in 1995. Therefore, for the distribution of this variable, the EU follows different criteria than that for the other types of expenditure. This characteristic will have important consequences in the analysis of its effects on the per capita GDP and on the growth rate distributions of European regions.

If we analyse the EU spending in 1995 with respect to the growth rates of the European regions for the period 1995-1998 (Figure 2, panel 1), it seems that the total budget only slightly stimulated growth. The mode representing the fast growth regions is, indeed, on the right with respect to the mode representing the slow growth regions. The analysis relative to the decomposition of the budget does not clarify which part of it had this effect. Therefore, it becomes of crucial importance to isolate the effects of each type of EU spending on the growth rate of regional economies, and this is the purpose of next section, focusing on the effects of the R&D spending.

4. R&D Spending and Regional Economic Growth

This section analyses the effect of the R&D level of spending provided by the EU to the European regions, both on the level of the regional per capita GDP and on regional growth rates.

In order to analyse this relationship, we estimate conditional quantiles. This approach estimates the conditional median (instead of the conditional mean) of the dependent variable together with all the distribution quantiles of the investigated relationship, when the independent variables vary.

Fig. 1. Density Estimates for different Components of the EU Budget and for sub-samples “rich” and “poor” regions

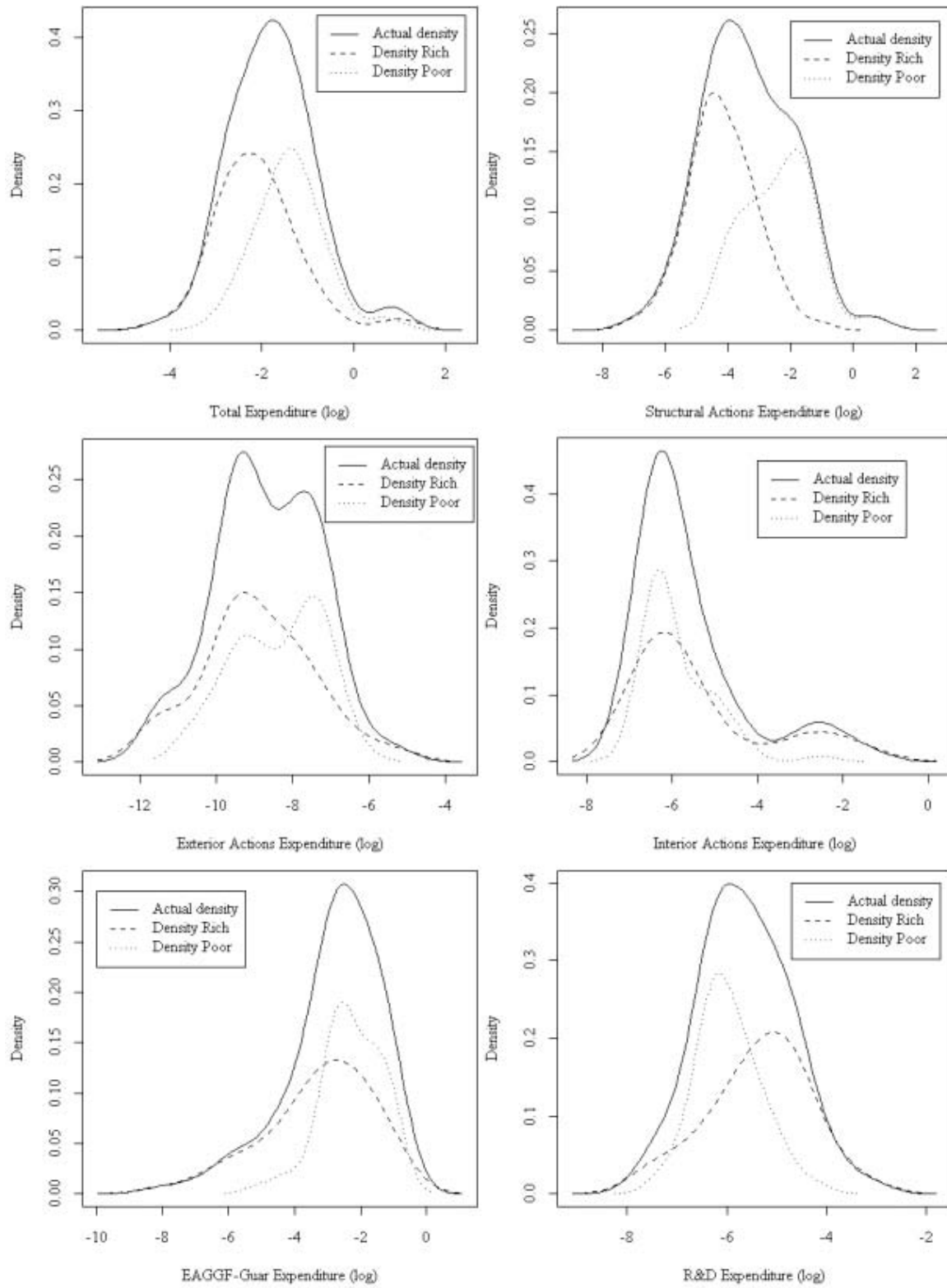
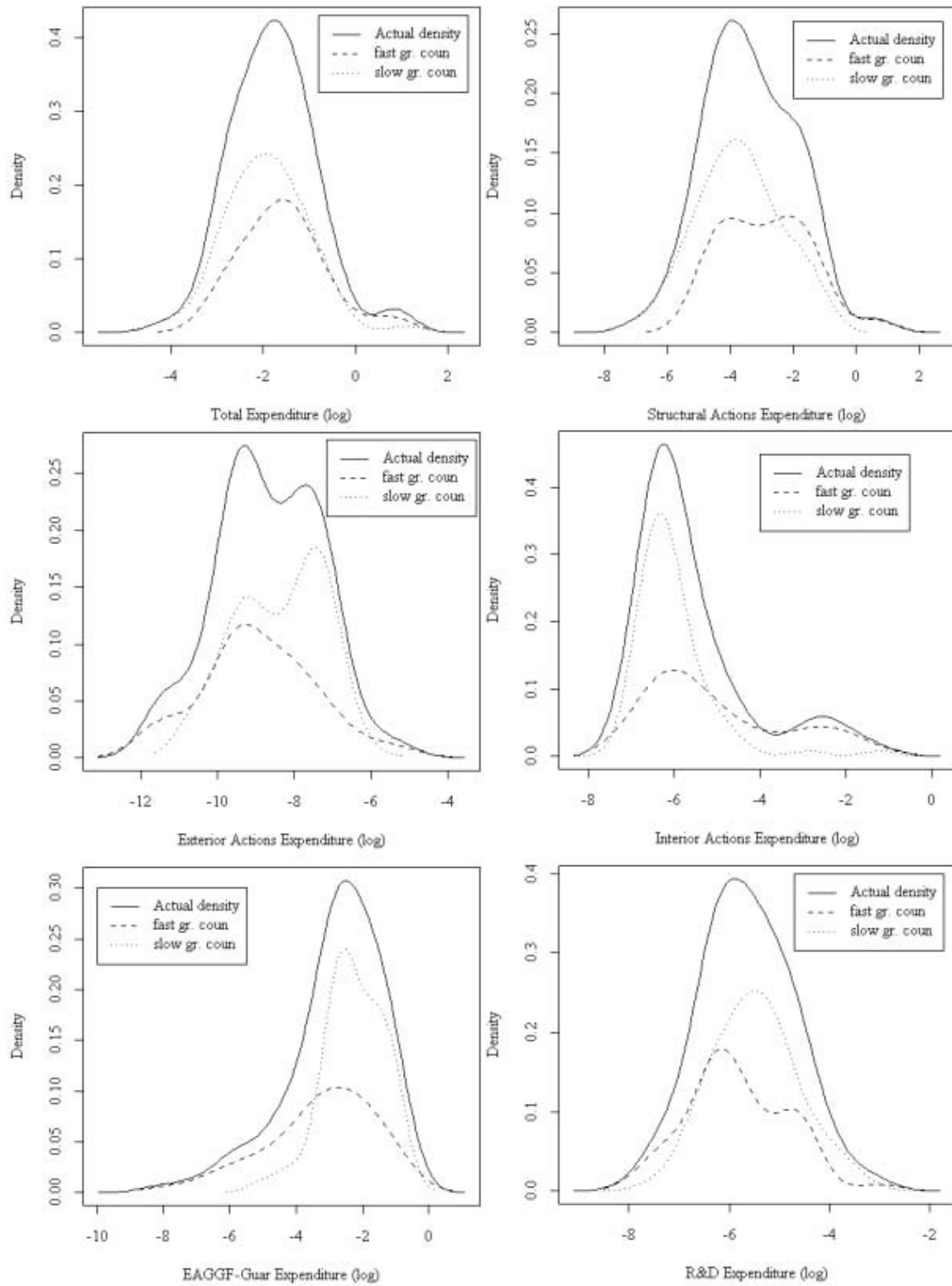


Fig. 2. Density Estimates for different Components of the EU Budget and for sub-samples “fast growth” and “slow growth” regions



In other words, it makes it possible to study the relationship between the GDP and the R&D expenditure for different GDP levels, and to study the evolution of this relationship when the R&D expenditure provided to the regions by the EU is allowed to vary.

Moreover, a non-parametric conditional quantile regression approach is used, because it makes possible to avoid a priori assumptions about the specific functional form, while obtaining a graphical idea of the existing relationship between the variables under examination. This approach is preferred when the scope of the analysis is to search particular stylised facts, because its specific characteristics can be independent of any (and possibly partial) theoretical reasoning.

Finally, we show how the results change when a semi-parametric conditional quantile regression approach is adopted to eliminate the effects of other variables (rather than R&D spending) that can affect the estimated relationship.

4.1. Issues on Non-Parametric and Semi-Parametric Conditional Quantile Regressions

In its parametric formulation, the regression relative to conditional quantiles was introduced by Koenker and Basset (1978). The underlying idea is to estimate, instead of the conditional mean, a set of quantiles of the distribution (for instance, the first, the median and the third), in the hypothesis that the mean relationship is not sufficient to detect whether the relationship exists and, more importantly, if it is the same for the whole distribution. In other words, we would have a regression line for every quantile of the relationship we are estimating.

However, in this context, as well as in many others, the assumption that the relationship between the variables is linear may not be appropriate. Therefore, the issue of specifying the functional form of the model emerges; this is a particularly hard choice especially when the impact of the independent on the dependent variable is not clear, and there is not an underlying theoretical model providing equilibrium solutions.

Furthermore, instead of testing for different functional forms, it is possible to use a non-parametric approach. It allows us to avoid a priori assumptions about the specific functional form and to obtain a graphical idea of the existing relationship between the variables under examination.

Different non-parametric approaches have been used in the literature, such as the spline smoothing (Koenker et. al., 1992), and the kernel density estimation (Abberger, 1997). In the following, we use the latter approach.

Let YN be the (\ln of) per capita GDP index and RD (\ln of) the EU R&D expenditure assigned to the regions. The joined cumulative density function $F(rd, yn)$ gives the proportion of the population where $RD \leq rd$ and $YN \leq yn$ hold simultaneously. The existence of the bivariate distribution $f(rd, yn)$

is assumed. In this case, the cumulative density function of the per capita GDP level, conditioned to the R&D expenditure is given by:

$$F(yn|rd) = \int_{-\infty}^{yn} f(x|rd)dx = \int_{-\infty}^{yn} \frac{f(rd, x)}{f(rd)} dx \quad [7]$$

where

$$f(rd) = \int_{-\infty}^{\infty} f(rd, yn)dyn \quad [8]$$

represents the marginal distribution of RD .

The inverse function $F^{-1}(p|rd)$ of [7] gives the p -quantile (parametrically estimated) of YN conditioned to $RD = rd$. It should be noted that this is the same procedure by which the median of the distribution can be estimated (i.e. $p = .5$). By varying p (between 0 and 1, of course) an infinite number of quantiles of the investigated relationship can be easily obtained.

The problem is, at this stage, that the functional form of $f(rd, yn)$ is unknown and must be estimated. Instead of using a priori assumptions about its shape, we estimate it non-parametrically. Further, in this case, this choice is the only option since a specific parametric specification does not exist, see Trede (1998).

Let n be the number of observations for which we have the per capita GDP level and the R&D expenditure. With, rd_i and yn_i the measured variables relative to the i th individual of the population, with $i = 1, \dots, N$.

The unknown densities in [7] are substituted by their non-parametric estimates. Starting from the non-parametric estimation of the bivariate densities:

$$\hat{f}(rd, yn) = \frac{1}{nh_1h_2} \sum_{i=1}^n K\left(\frac{rd - rd_i}{h_1}\right) K\left(\frac{yn - yn_i}{h_2}\right) \quad [9]$$

where h are the bandwidths and $K(\cdot)$ is the kernel¹².

Substituting [9] into [7] and rearranging, we obtain:

$$\hat{F}(yn|rd) = \frac{\sum_{i=1}^n K\left(\frac{rd - rd_i}{h_1}\right) G\left(\frac{yn - yn_i}{h_2}\right)}{\sum_{i=1}^n K\left(\frac{rd - rd_i}{h_1}\right)} \quad [10]$$

¹² In the bivariate case, the choice of the form of the relationship between the two kernels is arbitrary. In any case, we chose the simplest form, which is the multiplicative one. Obviously, this issue does not arise in the univariate case.

where

$$G(z) = \int_{-\infty}^z K(x)\delta x$$

represents the cumulative density function of the kernel function for a given variable x .

By inverting equation [10], we obtain the non-parametric quantile p that we are looking for, which depends on the R&D expenditure level of the EU towards the regions.

It has to be stressed that the non-parametric conditional quantile gives the effect of the R&D spending on the per capita GDP level, together with the effect of all the other variables that can also affect both variables. With these estimates, we cannot isolate, in other words, the estimated effect of R&D spending.

If we are interested in a sort of “pure effect” of the European R&D expenditure, we should wash up the effect of all variables affecting both the GDP level and the R&D expenditure from the previous estimated relationships. In order to estimate this effect, we have to “hold constant” those variables. This is the purpose of the semi-parametric estimation approach.

Essentially, this approach consists in estimating a *partially linear regression model* (Johnston and Di Nardo, 1997). Starting from the following model

$$yn_i = X_i\beta + f(rd_i) + \varepsilon_i \quad [11]$$

where one part of the model is linear and the rest of the model is non-linear. Rearranging it as follows

$$yn_i - X_i\beta = f(rd_i) + \varepsilon_i \quad [12]$$

a residualised yn is obtained, to be estimated non-parametrically against the rd level.

In practice, the linear part of the model is used to wash up the potential effect of X_i on yn_i . Applying the non-parametric approach on the residualised yn for the quantiles involved in the previous estimates “corrects” the estimates themselves from possible omitted variables effects.

Articles by Estes and Honoré (1995) and Yatchew (1997) suggest to first difference adjacent values of rd , in order to remove the non-parametric effect of the non-linear part of the model in the semi-parametric regression. According to these authors, it is possible to treat the parametric portion

of the model as a fixed effect¹³. Following this procedure, in order to have a consistent estimate of the β 's, we sort the data by ascending value of rd . Ordered in this way, provided the first derivative of f is bounded by a constant, we then calculate the “first difference” adjacent values of the regressand and all the regressors in the sorted data base:

$$yn_i - yn_{i-1} = (x_i - x_{i-1})\beta + (f(rd_i) - f(rd_{i-1})) + \varepsilon_i - \varepsilon_{i-1} \quad [13]$$

Running the OLS regression on the first differenced variables:

$$\Delta yn = \Delta X\beta + \Delta\varepsilon_i \quad [14]$$

gives consistent estimates of the coefficient.

According to Estes and Honoré (1995) and Yatchew (1997), if $f(rd)$ is upward bounded, the adjacent $f(rd)$'s values become closer and closer to each other and, provided f is continuous, the first difference $\Delta f(rd)$ approaches zero as the number of observations increases. Under these conditions, the estimates of β are consistent.

Once estimates of these coefficients are obtained, we compute the residuals as follows:

$$\hat{u} = yn_i - X_i\beta = f(rd_i) + \varepsilon_i \quad [15]$$

to run our quantile estimator ([7]-[10]). The same methodology applies to the estimates relative to the effect of the EU R&D spending on the average growth rates of European regions.

4.2. Non-Parametric Conditional Quantile Regressions

The results shown in Figure 3 indicate the effect of the European R&D spending on the regional per capita GDP. In this figure, as well as in all the others in this section, we report seven quantiles of the conditional distribution: the estimates relative to the first 10%, 15%, 25%, 50%, 75%, 85% and 90% of the distribution respectively.

¹³ In the following linear model:

$$y_{it} = X_{it}\beta + \varepsilon_{it}$$

the error structure for disturbance term is:

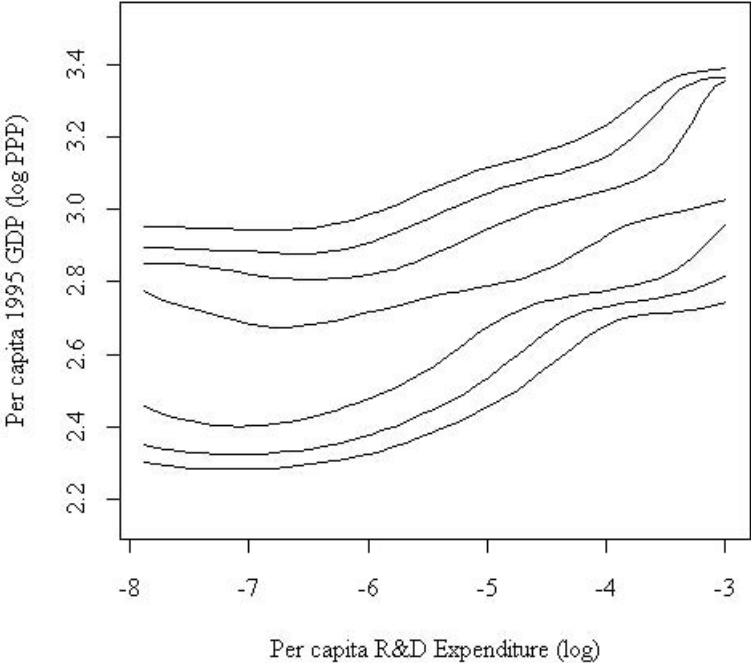
$$\varepsilon_{it} = \alpha_i + \eta_{it}$$

where it is assumed that η_{it} is uncorrelated with X_{it} . The first term in this expression (α_i) is an individual effect. It varies across individuals or the cross section unit, but is constant across time; it may or may not be correlated with the explanatory variables. On the other hand, η_{it} varies unsystematically (i.e. independently) across time and individuals. “This formulation is the simplest way of capturing the notion that two observations from the same individual will be more “like” each other than observations from two different individuals” (Johnston and Di Nardo, 1997). Models where the individual effect α_i is assumed to be correlated with the explanatory variable are called fixed effect models.

Our variable of interest shows, after a locally negative effect for low levels of R&D spending (not very pronounced for the first two quantiles), a positive effect on the regional per capita GDP. This effect is present, even with different strength, on all the quantiles of interest. We can conclude, therefore, that if the R&D spending reaches a certain level for all the European regions, it will have a positive effect on their levels of per capita GDP (the wealth of the area will increase).

Regions starting with higher per capita GDP will experience an increase in per capita level of income. Poorer regions will have a stronger effect and then, once the median position of the distribution is reached, the R&D expenditure will decrease its effect at the level found for the richer regions.

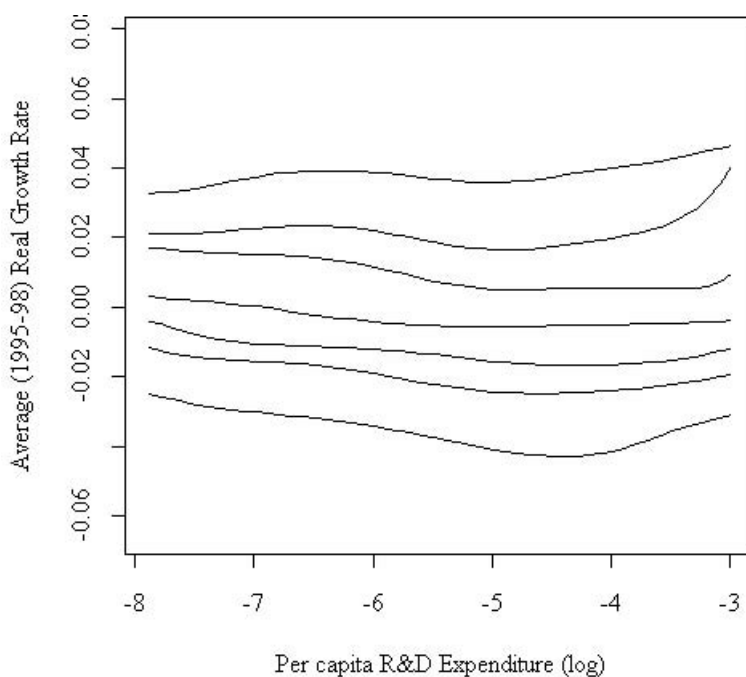
Fig. 3. GDP Level and R&D expenditure. Non-Parametric Conditional Quantiles



Combining these results together, and looking at the distance between the first and the last quantile, the non-parametric conditional quantile regression seems to indicate that the distance among the quantiles of the conditional (to R&D spending) distribution of the per capita GDP of the European regions will decrease, i.e. the R&D expenditure seems to induce convergence in per capita GDP levels. It seems, moreover, that regions set in the same part of the distribution tend to cluster.

Figure 4 presents the results regarding the effect of the R&D European spending on the regional growth rates; they indicate that the median effect of our variable on the per capita GDP distribution is small and negative.

Fig. 4. Growth rates and R&D expenditure. Non-Parametric Conditional Quantiles.



Moreover, the fast growing regions grow at higher rates as R&D increases; however, this is not true for the observation set in the low quantiles of the growth rates distribution. The analysis for the estimated quantiles indicates that R&D spending has a slight positive effect, which is only for high levels of the R&D spending. In general, as the variance of the distribution increase, this means that the growth rates tend to diverge.

4.3. R&D Spending and Regional Economic Growth. A Lock-in Effect?

Despite the previous discussion, the non-parametric conditional regression approach analysed so far does account for the effect of other variables that can also affect the level of the regional GDP and the growth rate of the regional economies. In particular, we should wash up from the previous estimates, the effect of the other component of the EU budget both on the GDP level and the growth rates. Therefore, to obtain a sort of “pure” effect of the European spending in R&D on the regional level of per capita GDP and on the growth rate of the European regions, we “control” the previous estimates for variables expected to affect the observed relationship.

The estimations presented in this paragraph control the previous quantile regressions for variables that can affect both the per capita GDP levels and the growth rates. Model [16] is used for the GDP level (following, of course the differentiating procedure shown in part 4.1):

$$yn_i - \left(\sum_{i=1}^4 \beta_i X_{ii} + \sum_{j=1}^4 \beta_j BC_{ji} \right) = f(rd_i) + \varepsilon_i \quad [16]$$

and for the growth rates

$$g_i - \left(\sum_{s=1}^5 \beta_s X_{si} + \sum_{j=1}^4 \beta_j BC_{ji} \right) = f(rd_i) + \varepsilon_i \quad [17]$$

In model [16], the set of variables chosen to control the estimates are:

- Three variables (X_{ii}) indicative of the wealth of an area: the unemployment level, population and density of population.
- Four variables (BC_{ji}) representing the other EU budget components: the per capita EU expenditure in Structural Actions, EAGGF-Guarantee, Internal Actions and External Actions.

In model [17], the level of the regional per capita GDP itself is used as a conditioning variable. All variables are taken in logs¹⁴.

Results of the estimates are presented in Figures 5 and 6. Furthermore, these figures present the comparison between the non-parametric and the semi-parametric conditional quantile regressions. As can be immediately observed, they show a different picture with respect to the first set of estimates. The overall effect of the R&D spending is bigger in the semi-parametric conditional quantile regression than in the non-parametric conditional quantile regression. Quantiles are closer together, and they are all upward sloping.

Although for regions starting from low levels of per capita GDP estimates are still locally negative, the effect on this part of the relation is reduced; more generally, this variable has a positive impact on per capita GDP level. The effect on the median is linear and positive. The turning point completely disappears for the observation set on the high quantile of the estimated relation.

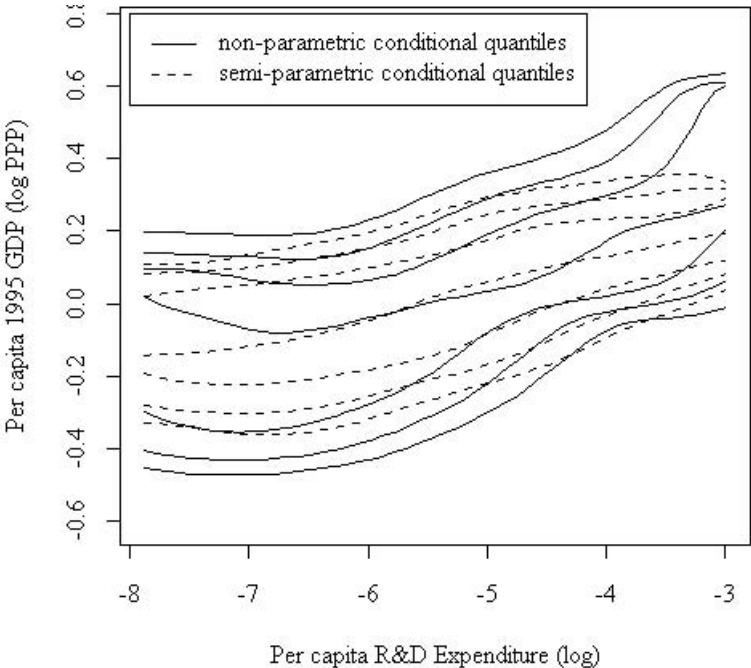
¹⁴ We have to highlight the difficulty of finding data at a European regional level for important variables such as private capital stock or public capital stock, variables that could be introduced in the regression if found. However, the variables available, and hence finally chosen, seem to be a good approximation of the determinants of the per capita GDP and the growth rates of the regional economies.

However, if all regions would receive the same level of R&D spending than the regions that received more, then the effect of this variable would be a positive effect on the per capita income. Looking at the semi-parametric estimate, the quantiles are decisively closer to each other with respect to the previous estimate. This is essentially because, even if all quantiles are upward sloping, the estimates show that for high level of R&D spending, this variable has a bigger impact on the lower quantiles.

Figure 5 also shows that European regions would concentrate in two clusters in the conditional distribution of per capita GDP. For high levels of R&D spending the distribution becomes polarised at two levels: above and below the median of the distribution. In other words, rich regions (receiving high quantities of R&D spending) would experience increases on the per capita level of income that tend to cluster them; a very similar effect holds for regions starting with low initial levels of per capita income. Therefore, this variable induces global convergence but polarisation of per capita GDP.

This process can be interpreted as a lock-in effect of the R&D spending: richer regions, possibly due to their economic structure, can use this type of EU spending in a more productive way than poorer regions.

Fig. 5. GDP Level and R&D expenditure. Semi-Parametric Conditional Quantiles



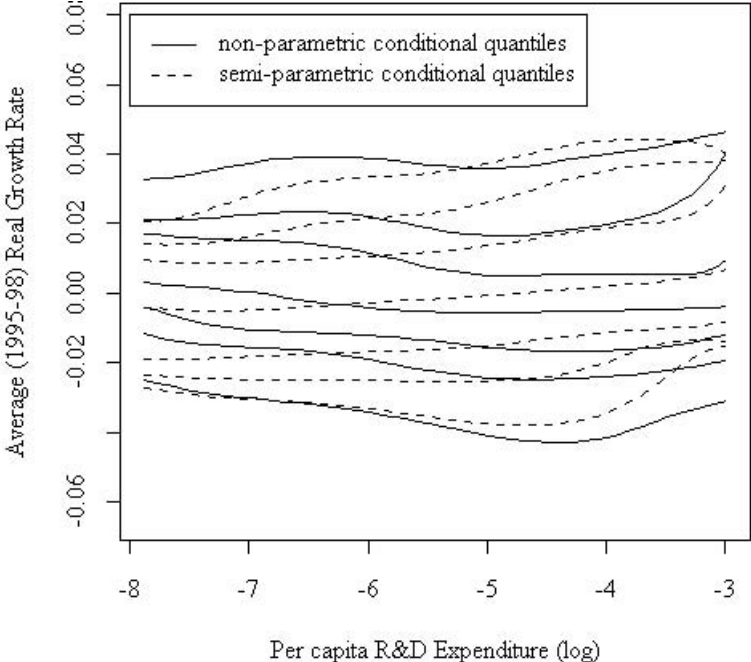
This confirms the idea that the R&D spending is more productive in those regions with a better economic structure. These results are confirmed with the analysis of the effect of the R&D spending on the growth rate of the European regions (Figure 6).

The effect of the R&D spending on the growth rate is small but after conditioning the quantiles, the relation is positive for almost all the quantiles. All of them have a higher slope than in the non-parametric estimates; this effect is very clear at the median position of the growth rates distribution.

This variable shows stronger effects on the growth rates of regions located in higher quantiles than the ones set in the lower quantiles. Regions with higher growth rates will experience a strong increase in the growth rates; regions with low growth rates will experience a smaller increase in their growth rates.

The conditional distribution of the growth rates tends to converge to two levels of growth rates as long as R&D spending increases. This means that R&D spending makes both the regional GDP and growth rates cluster. These results confirm what has already been found on the analysis for per capita income levels: that there is a lock-in effect for this component of the EU budget.

Fig. 6. Growth rates and R&D expenditure. Semi-Parametric Conditional Quantiles.



5. Conclusions

Now, we summarize the main conclusion that can be drawn from the imputation done at a regional level of the EU budget, and from the effects of the R&D spending on per capita GDP level and growth rates of the European regions.

The analysis of the EU budget shows that it has been mainly received by regions with a per capita GDP below the European average, reinforcing the idea that the EU expenditure has redistributive functions. This pattern is clear for expenditure such as Structural Actions, EAGGF-Guarantee, Interior and Exterior Actions. We found that the only type of expenditure with a different pattern is the R&D spending. Moreover, a first look to the relation between the distribution of the European funds in 1995 and the average growth rate for the 1995-1998 period shows a less clear picture. Regions that grew faster received a similar amount of total EU spending than regions that had lower growth rates in the period considered. This allowed us to analyse the effect of the R&D spending on per capita GDP level and growth rates of European regions.

Using non-parametric and semi-parametric conditional quantile regressions, we found a “lock-in effect” of EU spending on R&D at a regional level. The R&D spending has a stronger positive effect on regions with higher growth rates with respect to regions with lower growth rates. Moreover, we find European regions concentrate in two clusters in the conditional distribution of growth rates. For high levels of R&D spending, the distribution becomes polarised at two levels: above and below the median of the distribution, this is what we have called the “lock-in effect”.

Furthermore, the “pure” effect of R&D expenditure on regional per capita GDP levels seems to be positive. However, the conditioned per capita income distribution shows global convergence (less variance) but polarisation of per capita income.

6. References

- Abberger, K. (1997), “Quantile smoothing in financial time series”, *Statistical Papers* 38.
- Aghion P., Howitt P. (1998), *Endogenous Growth Theory*, The MIT Press Cambridge, Massachusetts.
- Annual Reports of the Court of Auditors. Various years.
- Barro, R. and Sala-i-Martin, X. (1992), “Convergence”, *Journal of Political Economy*, 100,2.
- Boldrin, M. and Canova, F. (2001), “Inequality and Convergence in Europe’s Regions: Reconsidering European Regional Policies”, *Economic Policy*, vol 0 (32).
- Cecchini, P. (1988), *Europa 1992: Una apuesta de futuro*, Alianza Editorial. Madrid.
- Cremer, H. and Pestieau, P. (1996), “Distributive implications of European Integration”, *European Economic Review*, vol 40.

- Davezies, L., Nicot, B. H. and Prud'Homme, R. (1996), *Economic and Social Cohesion in the European union: The Impact of Member State's Own Policies*. European Commission
- Emerson, M., Gros, D., Italianer, A., Pisan-Ferry, J. and Reichenbach, H. (1992), *One market, one money. An evaluation of the potential benefits and costs of forming an economic and monetary union*, Oxford University Press, Oxford
- Espasa, M. (2000), "El poder redistributivo del Presupuesto de la Unión Europea. Análisis a través de los flujos fiscales regionales", Doctoral Thesis, Universitat de Barcelona.
- Estes, E. and Honoré, B. (1995), "Partially linear regression using one nearest neighbour", Princeton University, Department of Economics, March 1995, mimeo.
- Eurostat (various years) "Statistics in Focus".
- Goodhard, C. and Smith, S. (1993), "Stabilization", *European Economy*, vol 5, Report and Studies.
- Grossman, G. and Helpmann, M. (1991), *Innovation and Growth in the Global Economy* Cambridge, Mass. And London: MIT Press.
- Howitt P. (1999), Steady Endogenous Growth with Population and R&D Inputs Growing, *Journal of Political Economy*, **107**(4), pp. 715-730.
- Johnston, R. and Di Nardo, R. (1997), *Econometric Methods*, McGraw Hill, IV Edition.
- Jones, C. I. (1995), "Time Series Test of Endogenous Growth Models", *The Quarterly Journal of Economics*. Vol 110 (2).
- Koenker, R., Portnoy, S. and Ng, P. (1992), "Nonparametric estimation of conditional quantile functions", in Dodge, Y. (Ed), *L₁-Statistical Analysis and Related Methods*. North-Holland, New York.
- Koenker, R. and Basset, G. (1978), "Regression Quantiles", *Econometrica*, vol 46, Issue 1.
- Krugman, P. (1993), "Lessons of Massachusetts for EMU", in Torres, F. and Giavazzi, F. (eds): *Adjustment and Growth in the European Monetary Union*, Cambridge University Press, Cambridge.
- Padoa-Schioppa, T. (1987), *Eficacia, Estabilidad y Equidad: Una estrategia para la evolución del sistema económico de la Comunidad Europea*, Alianza Editorial, Madrid.
- Quah, D. T. (1996a), "Empirics for Growth and Convergence", *European Economic Review*, 40.
- Quah, D. T. (1996b), "Regional Convergence Clusters Across Europe", *European Economic Review*, 40.
- Romer P. M. (1990), Endogenous Technological Change, *Journal of Political Economy*, 98(5) pt. 2, pp. 71-102.
- Sala-i-Martin, X. and Sachs, J. (1992), "Fiscal federalism and optimum currency areas: evidence for Europe and United States" in Canzoneri, M., Grilli, V. and Masson, P. (eds): *Establishing a Central Bank: Issues in Europe and Lessons from the US*, Cambridge University Press, Cambridge.
- Segerstrom P. S. (2000), The Long Run Growth Effects of R&D Subsidies, *Journal of Economic Growth*, **5**(3), September, pp. 277-305.
- Silverman, B. W. (1986), *Density Estimation for Statistic and Data Analysis*, New York, Chapman and Hall.
- Solow, R. (1956), "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, 70.

- Terrasi, M. (2000), "National and Spatial Factors in EU Regional Convergence", presented at the 40th European Congress of the European Regional Science Association. Barcelona 2000
- Trede, M. (1998), "Making mobility visible: a graphical device", *Economic Letters*, vol 59.
- Wand, M. P., Marron, J. S., and Ruppert D. (1991), "Transformation in Density Estimation", *Journal of American Statistical Association*, 86.
- Yatchew, A. (1997), "An elementary estimator of the partial linear model", *Economic Letters*, vol 57.