Social capital and economic growth in Europe: nonlinear trends and heterogeneous regional effects

Jesús Peiró-Palomino
Universitat Jaume I

Abstract

After two decades of academic debate on the social capital-growth nexus, discussion still remains open. Most of the literature so far, however, has followed the one-size-fits-all approach, neglecting that the great disparities across geographical units might have implications in this relationship. This article analyzes the role of two social capital indicators on the growth of 237 European regions in the period 1995–2007 by implementing a set of both parametric and nonparametric regressions. Whereas the former impose a linear functional form for the parameters, the latter relax this assumption providing a flexible frame in which the functional form is given by the data. The technique also permits introducing parameter heterogeneity in the analysis by estimating individual regional effects. The results from the parametric analysis show that the sign and the magnitude of the effects hinge upon the indicator considered. In contrast, results from the nonparametric regressions suggest that, while both indicators are significant growth predictors, the effect departs from linearity. Moreover, not all regions benefit from social capital with the same intensity. The most notable difference lies in regions from Central and Eastern Europe countries, where social capital is mostly negative. Other regional conditions such as initial income levels, investment rates or the stock of human capital show a more limited influence.

Keywords: Economic growth; European regions; nonparametric regression; social capital

JEL classification: C14; R11; Z13

Communications to: Jesús Peiró-Palomino, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain. Tel.: +34 964728615, e-mail: `peiroj@uji.es`

1. Introduction

The theory of social capital has received a great deal of attention in the economic growth literature in relatively recent times (see, for instance Putnam, 1993; Knack and Keefer, 1997; Zak and Knack, 2001; Akçomak and Ter Weel, 2009; Bjørnskov, 2012). Many definitions of social capital have been proposed, most of them closely related and sharing the same intuitions. In a broad sense, however, social capital might be understood as a set of informal forms of institutions and organizations based on social relationships, networks and associations that create shared knowledge, mutual trust, social norms and unwritten rules (Durlauf and Fafchamps, 2005). Therefore, social capital is not an unitary concept, but multiple facets coexist, capturing different aspects (see Bjørnskov, 2006). Consequently, the indicators used as a proxy variables for social capital are manifold, whose selection is heavily conditioned by the availability of data. Two of the most used proxies in empirical studies are a measure of *social trust* and levels of *associational life* in society.

Theory suggests that social capital stimulates economic growth throughout a variety of channels. For instance, Putnam (1993) argued that social capital facilitates coordination and cooperation for mutual benefit and it helps in solving problems of collective action. Knack and Keefer (1997) held that it reduces monitoring costs, which translates into a reduction of transaction costs in economic operations. Social capital also facilitates complex agreements by mitigating information asymmetries between negotiating parties (Dearmon and Grier, 2009), as well as it eases knowledge diffusion and innovation processes (Akçomak and Ter Weel, 2009). The environment of mutual trust, participation and cooperation created where social capital is abundant also stimulates other activities which, in turn, are related to higher growth. These are, to name few on which empirical evidence is relatively recent, financial development (Guiso et al., 2004), human capital (Bjørnskov, 2009; Dearmon and Grier, 2011), investment (Zak and Knack, 2001; Dearmon and Grier, 2011; Peiró-Palomino and Tortosa-Ausina, 2013b) or trade (Guiso et al., 2009). In sum, the general idea we might draw from this brief review of the literature is that countries and regions where social capital is abundant perform better in economic terms.

However, for some geographical contexts such as the European regions evidence is still relatively scant and results are more mixed. Schneider et al. (2000) found no effects from social trust to growth, which is in contrast with most of the literature at the country level. Beugelsdijk and Van Schaik (2005) considered 54 (NUTS 1)[1] regions and two indicators of

[1]NUTS stands for Nomenclature of Territorial Units for Statistics. It is a hierarchical system for dividing up

social capital, namely social trust and associational activity. They found that only the latter are related to higher growth. Akçomak and Ter Weel (2009) analyzed the effect of social trust on innovation in 102 regions, finding a positive and significant link. However, they suggested that trust and growth are not directly related. Another common feature is that they are based on samples conformed only with Western Europe regions, whereas evidence for Eastern and Central European regions (ECE hereafter) is still comparatively scant. This is a result of the data shortcomings for these regions, where reliable data on both economic and social variables have been made available only recently.

However, considering these regions is important for some reasons. On the one hand, social values in ECE regions differ from their Western peers. Some authors such as Rose (2000), Paldam and Svendsen (2001) and Żukowski (2007) suggest this is a consequence of the long communist experience, which modified social patterns and negatively affected social capital, which today lags behind the European average. Authors such as Fidrmuc and Gërxhani (2008) defend that this backwardness is a result of poor institutions and low economic development in transition economies and therefore it can be expected that social capital increases at the time development comes. However, others such as Żukowski (2007) are more sceptical and manifest that cultural values in ECE regions are well installed in society and have remained remarkably stable in time.[2] In addition, social capital in these transition economies may adopt negative forms, as suggested by Paldam and Svendsen (2001) and Rose (2000). Against the impossibility of the state and public institutions to cover people's basic necessities, corruption and negative forms of organizations created to fulfill their own interests proliferate, generating harmful effects for the general interest.

On the other hand, there are important disparities in the evolution of income per capita between ECE and Western European regions in recent years. In particular, the former have grown faster than the latter (see Crespo-Cuaresma et al., 2012) in recent times, thus accomplishing the predictions of the neoclassical growth theory, which suggests that economies

the economic territory of the EU for the collection, development and harmonisation of EU regional statistics as well as their socioeconomic analysis.

[2]The stability of social values has become a recent focus of debate. Contributions such as Uslaner (2008) and Nunn and Wantchekon (2011) have shown that social capital (in particular social trust) is inherited from generation to generation. This stability is supported by the data. For instance, while social capital is low in ECE countries, it is historically high in Nordic countries. This becomes an important argument when establishing causality between social capital and economic growth. While the former has remained stable at leat from the World War II, the latter has experienced remarkable changes. Therefore, the idea that social capital is a result of a given economic scenario is difficult to reconcile (Bjørnskov and Méon, 2013).

departing from worse initial conditions grow faster than the richer.[3] The above arguments suggest that growth determinants in both ECE and Western regions might differ, at least in the intensity of their impact. More in particular, the role of social capital might be heterogeneous across groups of regions.

In this regard, the recent study by Peiró-Palomino et al. (2014) analyzed the effect on growth of some indicators of social capital in 85 (NUTS 1) regions of the enlarged European Union (therefore considering ECE regions). Following Bayesian techniques, they found support for positive influences of social capital for the recent period 1995–2008. However, no mention of parameter heterogeneity is made and, in the light of the previous arguments, generalizing social capital effects in a heterogeneous setting such as the one represented by the European regions might be risky.

In order to be consistent with previous literature, this paper also considers social trust and associational activity as proxy variables for the social capital theory. Regarding the geographical and temporal settings, it considers 237 (NUTS 2) European regions in the period (1995–2007)[4]. As the reader might note, this is a similar framework to that considered in Peiró-Palomino et al. (2014). Yet the paper attempts to contribute to the existent literature in some particular respects where evidence is still scant or even nonexistent.

For instance, the level of territorial disaggregation used has important implications when giving policy prescriptions, since NUTS 2 have policy competencies, whereas NUTS 1 correspond to broad macroeconomic regions with no decision power. Accordingly, NUTS 2 regions are the objective of the policies in the frame of the European Regional Policy. These objectives are notably linked to the focus of the paper, since they are mostly addressed to both promoting economic growth and achieving regional income convergence. In addition, by considering NUTS 2 the sample size remarkably increases, achieving (to my knowledge) one of the biggest in the social capital literature.[5] However, the more innovative contribution comes from the methodological side. In particular, both parameter heterogeneity and potential nonlinearities in the estimated parameters are taken into ac-

[3]That result implies that GDP per capita in ECE regions is converging to the Western's regions level. However, the existence of a global trend of convergence at regional and country level does not exclude the possibility that differences inside ECE countries increases (Monastiriotis, 2014).

[4]Year 2008 is excluded because of the arrival to most of the European regions of the economic recession, which started in 2007 in EE.UU.

[5]Data constraints have made samples in the social capital literature to be traditionally smaller than those in studies dealing with economic growth from a more general perspective. Some examples (at both country and regional level) are Knack and Keefer (1997)(29 countries), Beugelsdijk and Van Schaik (2005) (54 regions), Zak and Knack (2001) (51 countries), Akçomak and Ter Weel (2009) (102 regions) or Peiró-Palomino et al. (2014) (85 regions).

count by implementing some of the most recent developments in growth empirics, namely nonparametric regression.

The shortcomings faced ordinary least squares (OLS) estimations to provide robust estimates have lead authors in the growth literature to adopt alternative methods, mainly from Bayesian statistics (Fernandez et al., 2001; Durlauf et al., 2008; Crespo-Cuaresma et al., 2012) (among others) and from the nonparametric field (for instance, Maasoumi et al., 2007; Henderson et al., 2012, 2013). Whereas the former are more addressed to solve the problem of model uncertainty (mostly applying BMA techniques) and still assume linearity in the parameters, the latter are better indicated to study both potential nonlinearities and parameter heterogeneity. In the social capital literature Bayesian methods are still incipient, but we already find contributions such as Horvath (2012) and Peiró-Palomino et al. (2014). However, nonparametric methods are still yet to come to this particular context, where parametric linear specifications are the most common approach.

Consequently, should social capital affects growth nonlinearly we are misinterpreting its true effects. Analogously, if social capital effects are heterogenous across regions, adopting the usual estimation methods (mostly OLS) may lead us to underestimate (overestimate) the effects for some regions. Bayesian methods as those applied in Peiró-Palomino et al. (2014) offer some advantages in this regard by computing posterior distributions for the parameters—i.e. they provide the complete range of values the parameters may take in terms of probability. While this represents some improvement, we still cannot identify social capital effects in particular groups of regions. In addition, despite estimating from a Bayesian point of view provides more flexibility than the early discussed classical estimation methods, linearity in the parameters is still imposed by the analyst.

Nonparametric kernel regressions as those proposed in Li and Racine (2007) can deal with these issues and, consequently, applications in assessing different growth theories are crescent (see Henderson et al., 2012, 2013).[6] These techniques provide a more flexible framework where the functional form of the parameters is given by the data and not imposed *a priori*. Previous evidence provides good reasons to consider that social capital (proxied by social trust) effects might be nonlinear. For instance, Roth (2009)

[6]The number of theories attempting to explain GDP per capita growth is numerous (Brock and Durlauf, 2001; Durlauf et al., 2008). Apart from the *Solow variables*, which are considered as the baseline model in most of growth studies, other theories are *demography*, *policy*, *geography*, *fractionalisation*, *institutions* and *financial development*, to name few. A careful assessment of (some) of theses theories can be found in former studies such as Sala-i-Martin (1997), and more recently in Durlauf et al. (2008) or Badunenko and Romero-Ávila (2013). More in particular, an evaluation using the techniques employed in this paper is provided in Henderson et al. (2012, 2013).

and Dearmon and Grier (2011) included quadratic terms and conclude that the positive effects of trust on growth and investment, respectively, decline after a given threshold is exceeded. However, the quadratic term is included in the variable trust and not in its associated parameter, which is ultimately the focus of nonparametric regression. Peiró-Palomino and Tortosa-Ausina (2013a) considered a family of quantile regressions and showed that trust effects are sensitive to the level of development of the economies. In particular, they suggest that trust has no effect on GDP per capita in the poorest countries. Yet in the latter case the functional form is linear and what is is obtained is a set of β parameters for different quantiles of the dependent variable—i.e. GDP per capita. Therefore, these studies give some insights on the likely heterogeneity of the effects of social capital but the functional form is imposed.

A complete understanding of the social capital-growth nexus seems pertinent, given the interest of the European policymakers in achieving higher levels of economic integration, both economic and social. In this context it should be evaluated not only which forms of social capital are more beneficial but also analyze how they truly behave, i.e. letting the data define by themselves the true and complete underlying functional form in the social capital-growth relationship and permitting parameter heterogeneity across regions. The methods used in this paper aim precisely to this objective.

The reminder of the paper is structured as follows. Section 2 contains some details on the nonparametric methods used in the paper. Section 3 describes the models to be estimated together with an explanation of the variables and the data sources. Section 4 provides the results and, finally, Section 5 concludes.

2. Empirical methodology

In order to assess the social capital theory as a growth predictor in the European regional context I will proceed in two steps. First, I will perform parametric (OLS) regressions and, second, nonparametric kernel regressions. OLS estimations provide an average β parameter for the effect of the right-hand-side variables on the dependent variable, namely GDP per capita growth. A common empirical strategy in the growth literature consists of estimating a parametric model such as:

$$Y_i = \beta_0 + \sum_{j=1}^{V} \beta_j Z_{ji} + \epsilon_i, i = 1, 2, ...n, \tag{1}$$

where Y_i is the growth rate of GDP per capita for region i, Z_i is a vector of V regressors, either continuous or categorical, and β_0 and β_j are the parameters to be estimated. Finally ϵ_i is a mean zero additive error, i represents a particular region and n is the sample size.

Parametric regressions offer a valuable point of departure, yet it is well known that they face several downfalls derived from the strong assumptions that have to be assumed in order to obtain consistent estimators. For instance, regressors must be uncorrelated with the error term and the functional form must be correctly specified. However, in a model such as Equation (1), linearity is assumed to be the true functional form and other likely forms the parameters might adopt are completely disregarded.

In contrast, an alternative and more flexible framework is provided by nonparametric kernel regressions, which relax many of the assumptions of the parametric framework. One of them is precisely the linearity assumption. Lets consider the nonparametric counterpart to Equation (1), given by:

$$Y_i = m(Z_i) + \epsilon_i, i = 1, 2, ...n, \tag{2}$$

where all components are common to Equation (1) with the exception of $m(.)$, which is an unknown smooth function capturing the conditional relationship between the left-hand-side and the right-hand-side variables in the model.[7]

Among the different alternatives to estimate Equation (2) the methods proposed by Li and Racine (2004) and Racine and Li (2004) are adopted. In particular, both local-constant least-squares (LCLS) and local-linear least-squares (LLLS) estimators.[8] These are methods based on "generalized product kernels", valid for the analysis of mixed data frames with both continuous and categorical (unordered and ordered) variables. As stated in Li and Racine (2007), they do not require neither a predefined functional form nor a distribution of the error term. Both of them are used because they provide different outcomes. Whereas LCLS allows us to identify irrelevant regressors, LLLS is better addressed to detect nonlinearities in the estimated parameters.

LCLS estimates the unknown smooth function $m(.)$ by calculating a local weighted average of the dependent variable Y_i considering the observations with similar values of

[7]Note that the linear specification represented by Equation (1) is also considered by the nonparametric regression and, should the regressors affect the dependent linearly, the model will also provide consistent estimates.

[8]Most of the calculations in the paper were computed using the `np package` implemented in `R`. See Hayfield and Racine (2008) for additional details.

the independent variables Z_i. The degree of similarity is given by the bandwidths, which determine the quantity of averaged observations around each point z_i. Finally, the different point estimates over the range of z are connected to obtain $m(.)$. The conditional mean of this estimator is given by the expression:

$$\hat{m}(z) = \frac{\sum_{i=1}^{n} y_i \prod_{s=1}^{q} K\left(\frac{z_{si}-z_s}{h_s}\right)}{\sum_{i=1}^{n} \prod_{s=1}^{q} K\left(\frac{z_{si}-z_s}{h_s}\right)} \tag{3}$$

where $\prod_{s=1}^{q} K((z_{si} - z_s)/h_s)$ is the product of the kernel functions and h_s is the corresponding bandwidth associated to regressor z_s.

In contrast, the LLLS estimator computes a weighted least-squares regression around every point z_i. The weights are established by a kernel function and a bandwidth vector such that those observations closer to z_i receive more weight. Then, analogously to LCLS, the estimated points are connected, revealing the form of $m(.)$. Lets consider again Equation (2). Adopting a first-order Taylor expansion for the continuous variables, denoted as z^c, we obtain:

$$Y_i \approx m(z) + (z_i^c - z^c)\beta(z^c) + \epsilon_i \tag{4}$$

where $\beta(z^c)$ represents the partial derivative of $m(z)$ with respect to z^c. The LLLS estimator of $\delta(z) \equiv [m(z), \beta(z^c)]'$ is given by

$$\hat{\delta}(z) = [Z'K(z)Z]^{-1}Z'K(z)y \tag{5}$$

where Z is a n x $(q_c + 1)$ matrix with i row $(1, (z_i^c - z^c))$ and $K(z)$ is a n diagonal matrix of product kernel weighting functions for both continuous and categorical data. A second-order Gaussian kernel is selected for continuous variables whereas for categorical variables the choice is the Aitchison and Aitken (1976) kernel. See Li and Racine (2007) for further details.

Independently of the approach followed, either LCLS or LLLS, the choice of the kernel is not a fundamental issue. However, the selection of the bandwidths is an essential point in nonparametric statistics (Hayfield and Racine, 2008). If the bandwidths are too large too many observations are considered in the smoothing process and it yields estimates with low variance and high bias (undersmoothing). In contrast, if the bandwidths are too small it results in estimates with high variance and low bias (oversmoothing). Therefore, band-

widths must be selected carefully. Among the different automated bandwidth selection criteria, the choice is least-squares cross-validation (LSCV), popularized by recent applications (see, for instance Arribas et al., 2010; Henderson et al., 2012, 2013)

The bandwidths not only determine the amount of observations smoothed, but also provide useful information on the relevance and the linearity of the regressors. As stated in Hall et al. (2007), when using the LCLS estimator and the bandwidth associated to one regressor is larger than its upper bound (UB) the regressor is essentially smoothed out. Similarly, Hall et al. (2007) also show that when using LLLS and the bandwidth of a continuous regressor reaches its upper bound, that variable enters linearly, which leads to analogous results than OLS estimation. Because I am first interested in determining if the theory of social capital (proxied by two different indicators) is affecting regional growth I will perform, in a first step, LCLS. Once I have studied relevancy, LLLS will be conducted in order to detect potential nonlinearities in the parameters. For the continuous variables the upper bound is defined as two times their standard deviation whereas for unordered categorical variables it is given by $(q_s - 1)/q_s$, where q_s is the number of values that the categorical variable can take.[9] As explained in Li and Racine (2007), these methods provide a partial estimate for each observation and each regressor, which allows for a more detailed analysis than parametric and semiparametric methods. In this context, this is especially useful since, besides studying the existence of nonlinearities, parameter heterogeneity can be also evaluated—i.e. it can be explored wheter the regressors for particular groups of regions affect GDP growth in a different fashion.

Additionally, with the objective of comparing parametric and nonparametric models and establishing which one is preferable given the data, the Hsiao et al. (2007) test is performed. Under the null hypothesis $(H0 : Pr[E(x|z) = f(z, \beta)] = 1)$ the parametric model is correctly specified. Under the alternative, $(H1 : Pr[E(x|z) = f(z, \beta)] < 1)$, it is not, and then the nonparametric specification is preferred.

3. Models and data

In this section are introduced the models to be estimated as well as the data and their sources. In order to assess the impact of the two indicators of social capital in my sample of European regions five different models are proposed, where these indicators are added

[9]The upper bounds are defined according to the selection of the kernel. For instance, in a Gaussian kernel a bandwidth larger than two times the standard error of the variable assigns equal weight to all observations.

to a baseline model sequentially. The baseline model (*Model 1*) is a standard neoclassical growth regression (see Solow, 1957). The dependent variable is the growth of real GDP per capita and the list of regressors includes the GDP per capita in the first year of the period analyzed (in logs), population growth and investment in physical capital. The model is usually augmented with human capital after Mankiw et al.'s (1992) contribution, who highlighted the importance of education for growth processes.[10]

Data for the variables in the baseline model are all provided by Eurostat. GDP per capita for the sample of 237 regions (NUTS 2) during the period (1995–2007)[11] is measured in real terms (euros of 1999). As usual in growth studies, this relatively long period is subdivided in two shorter subperiods, namely (1995–2001) and (2002–2007). Then, the dependent variable is calculated as the average annual growth of real GDP per capita in these subperiods $(GGDP)$. The real GDP per capita at the beginning of the periods is denoted by (GDP_0) and $(GPOP)$ is the average annual population growth in each subperiod.[12] For the investment variable, gross fixed capital formation $(GFCF)$ is considered (share of GDP). Finally, the stock of human capital, measured as the percentage of population aged between 25 and 65 with tertiary studies (ISCED 5 and 6 in European terminology) is denoted by (HC).

The variables from the social capital theory are selected following Beugelsdijk and Van Schaik (2005), Bjørnskov (2006) and Peiró-Palomino et al. (2014). As already mentioned in the Introduction, two variables are considered, namely social trust $(TRUST)$ and active participation in associations $(ACTIVE)$. By "active" I refer to people who is not merely a member in the group, but they voluntarily participate in the group activities.[13] These two variables are constructed by taking and aggregating

[10]The great power of this small set of regressors for predicting growth in different geographical contexts at both country and regional level has made using the Solow's framework as a starting point an extended strategy when evaluating other theories in growth empirics (see, for instance Durlauf et al., 2008; Henderson et al., 2012, 2013). This is even more common in the literature on social capital and growth, where virtually all the contributions augment the Solow's model by adding a variety of social capital indicators (see Knack and Keefer, 1997; Zak and Knack, 2001; Beugelsdijk and Van Schaik, 2005; Peiró-Palomino and Tortosa-Ausina, 2013a), to name few.

[11]Although more recent data were available at the time of writing, the great and uneven impact of the economic recession across European regions makes it appropriate considering only pre-crisis data. Additionally, the homogenizing process of national accounts in the former URSS countries has made comparable data previous to 1995 unavailable for ECE regions.

[12]Following Mankiw et al. (1992), is commnonplace in the literature adding a fixed coefficient equal to 0.05, which captures depreciation and technological change.

[13]Here the literature has considered both merely participation and active involvement. However, following Bjørnskov (2006) only the latter might be beneficial since interaction inside the group is needed to enjoy the benefits that social capital provides. In line with this argument, Beugelsdijk and Van Schaik (2005) reported a

individual data from the European Values Survey (EVS), wave of 1999. Although the EVS provides surveys for different years (1981, 1990, 1999 and 2008) the data were taken from a year as close as possible to the beginning of the period in order prevent endogeneity, even though the latter is not a particularly worrying issue in social capital studies due to the stability of social values (consequently neither it is considering data exclusively from one wave). In addition, when dealing with social capital some important data constraints arise. Surveys for 1981 and 1990 do not provide regional data. Even in the 1999 wave data at NUTS 2 level— the level of measure for the rest of the variables in the analysis—are not available for most of the regions. As an alternative, social capital data are aggregated at NUTS 1 level, and thus all regions belonging to the same NUTS 1 are assumed to be equal in terms of social capital. This in not especially difficult to reconcile, since social values are expected to be similar in relatively vast geographical areas (such as the ones represented by NUTS 1).[14]

The social trust indicator is based on the widely used question considering the percentage of people who declared trusting others when asked: *"Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?"*. Active participation in groups is measured as the percentage of respondents who mentioned doing unpaid work for a variety of associations, namely a) welfare organization; b) religious organization; c) cultural activities; d) trade unions and political parties; e) local community action; f) development/human rights; g) environment, ecology; h) professional associations; i) youth work; j) sports/recreation; k) women groups; l) peace movement; m) voluntary health; and n) other groups.

The empirical strategy consists of estimating first *Model 1*. *Model 2* and *Model 3* add $TRUST$ and $ACTIVE$, respectively, which are included separately. *Model 4* includes the baseline Solow model and both social capital variables. The most comprehensive model, *Model 5*, further augments *Model 4* by introducing the dummy variable $CAPITAL$, which is equal to one for those regions where the capital city of the country is located and zero otherwise and time effects, controlling for the time period. After merging and averaging all the available data in the two considered subperiods, these models are estimated for a sample of 404 observations. Before proceeding with the results, some descriptive statistics are provided in Table 1. Since one of the purposes is examining possible differences be-

more clear positive effect on growth when considering active participation instead of simple membership.

[14]Mixing different levels of disaggregation is common practice in studies focused on the European regions because of data limitations (see, for instance Akçomak and Ter Weel, 2009).

tween ECE and Western regions (hereafter non ECE regions) regions, the table provides the descriptives separately, distinguishing between the two groups.

In particular, growth rates are higher for ECE regions than for non ECE regions in both periods. This difference is more notable in the second one. GDP per capita at the beginning of the periods (1995 and 2002) was about five times greater in non ECE regions, which shows the great disparities in terms of wealth between both groups. However, population growth ($GPOP$) and gross fixed capital formation ($GFCF$) are similar and quite stable in time. Human capital (HC) was remarkably greater in non ECE regions in the first period, although the difference substantially decreases in the second. Regarding social capital,[15] the most notable differences are in terms of $TRUST$, almost double in non ECE regions. The $ACTIVE$ indicator is relatively low in both groups of regions, although the score is higher in non ECE regions. These descriptives have shown that differences in terms of GDP per capita, growth trends and social capital levels between both groups of regions are substantial. Therefore, in order to fully understand the drivers of growth in the European regions and more specifically the role of social capital, the application of the nonparametric methods introduced in Section 2 might be appealing.

4. Results

4.1. Parametric regressions

This section presents the results for the parametric OLS estimations. Table 2 shows the results for these regressions for the five models introduced in Section 3. *Model 1* only includes the Solow variables. As predicted by the neoclassical theory, poorer regions grow faster as the negative and significant coefficient for GDP_0 indicates. Population growth ($GPOP$) and physical investment ($GFCR$) are positive and negative, respectively. While these signs are contrary to the neoclassical theory, both are nonsignificant. However, human capital (HC) is positive and significant at the 1%, what suggests that growth in recent times in Europe is driven by education rather than physical investment or demographic factors. The results for the Solow variables are stable and consistent across the different specifications, only with the exception of ($GPOP$), which is significant in *Model 2*.

Regarding the variables of the social capital theory, *Model 2* includes $TRUST$, which exhibits a positive sign, although it is nonsignificant. However, $ACTIVE$, considered in *Model*

[15] Note that the values correspond to 1999 in both subperiods.

3, is positive and significant at the 1%. When both $TRUST$ and $ACTIVE$ are considered together in *Model 4*, $TRUST$ becomes negative (although it is only significant at the 10% level) whereas $ACTIVE$ remains positive and significant at the 1%. Finally, the most comprehensive *Model 5* introduces the dummy variable $CAPITAL$ and time effects, the latter not reported. $CAPITAL$ exhibits a positive and significant effect and the rest of variables remain stable in terms of sign and significance only with the exception of $TRUST$, which turns again nonsignificant.

Comparing these results with previous contributions in the European context is difficult because of differences in the sample, both the regions included and its size (it includes ECE regions and its size is remarkably higher). Overlooking these essential differences, the results found for $TRUST$ (especially those in *Models 4* and *Model 5*) are difficult to reconcile with the social capital theory, which predicts a positive effect on growth. However, they would be relatively consistent for the European context, where Beugelsdijk and Van Schaik (2005) found no effects for trust and Schneider et al. (2000) obtained a negative effect. However, the effect found in Peiró-Palomino et al. (2014) for the same period is mostly positive (the Bayesian perspective adopted in this paper does not allow us to talk about significance). The precedents for active participation in associations are again Beugelsdijk and Van Schaik (2005) and Peiró-Palomino et al. (2014), and the results are in line with those found in these studies. Therefore, previous findings for social capital theory in the European context partially hold when considering more disaggregated samples.

In order to assess the validity of the above estimations we perform Hsiao et al. (2007) tests, introduced in Section 2. Results are reported in Table 3. For all models (1–5), the null hypothesis of correct specification form is rejected. These results suggest that the relationship between the dependent (GDP per capita growth) and the regressors included is no longer linear, but it obeys to other functional forms. Therefore, adopting a flexible framework where no functional form is specified becomes an interesting contribution to previous research.

4.2. Nonparametric regressions

In this section results for the nonparametric counterparts of the five models considered in the preceding section are provided. I will proceed in three steps. First, I will compute the bandwidths via Least Squares Cross Validation (LSCV), estimated using both Local-

Constant Least Squares (LCLS) and Local-Linear Least Squares (LLLS) estimators. Second, I will calculate the estimates for each regressor across the different model specifications. And third, I will focus the analysis more tightly on the social capital variables by considering potential heterogeneity in its effects across particular groups of regions.

4.2.1. Bandwidths

Table 4 provides the upper bounds for the regressors as well as the bandwidths for the five models using both LCLS and LLLS estimators. In the table, those bandwidths that hit the upper bound of the corresponding variable are in bold. The bandwidths associated to GDP_0 (in logs) indicate that the regressor is relevant and enters nonlinearly across all models. However, bandwidths for the other variables in the baseline Solow model indicate that neither the relevance nor the functional form remain stable across specifications. For instance, according to the bandwidths obtained applying LCLS in *Model 2*, $GPOP$, $GFCF$ and HC would be irrelevant and bandwidths provided by LLLS show that HC enters the model linearly (although it is irrelevant). Results differ for *Model 3*, where all variables are relevant and affects growth nonlinearly. Results for *Model 4* are similar to those in *Model 2*, while in the most comprehensive *Model 5* $GFCF$ and HC are irrelevant and $GPOP$ enters the model linearly. The dummy variable $CAPITAL$ is irrelevant, since the bandwidth is very close to the upper bound and enters the model nonlinearly. The two social capital variables, $TRUST$ and $ACTIVE$ are relevant and enters the model nonlinearly in all five specifications, since any of the computed bandwidths exceed the corresponding upper bounds.

Some of the results obtained from the nonparametric regressions are somehow in conflict with some of those from the parametric estimations. In particular, focusing on *Model 5*, there are notable differences. For instance, HC is significant in the parametric specification but it is not in the nonparametric counterpart. The contrary holds for $GPOP$ and $TRUST$, which are nonsignificant in the parametric *Model 5* and significant in the nonparametric estimation. For the other variables significance remains unaltered. According to the bandwidths provided using LLLS, $GPOP$ enters the model linearly, while the rest of variables enter nonlinearly. Considering these disparities in the results it seems natural testing for the correct specification form, either parametric or nonparametric. Recall that in the preceding section, the Hsiao et al. (2007) test was performed (see Table 3) and in all cases the appropriateness of the parametric specification was rejected. Therefore, despite the initial

set of parametric regressions are an interesting first approach, conclusions drawn from them might be incomplete and unreliable in the light of the tests and the results derived from the more flexible nonparametric framework.

4.2.2. Nonparametric estimates

This section provides the estimates for the different regressors obtained using the LLLS estimator since, as stated in Li and Racine (2007), the latter has more preferable properties than LCLS. Figure 1 displays graphically the nonparametric results for the continuous regressors in *Model 5*. The different plots in the figure present the estimated nonparametric function $m(.)$ for each regressor in the model, holding the rest of variables constant at their median values. As stated in Maasoumi et al. (2007), this approach permits direct comparison between the parametric and the nonparametric results. In order to save space, only the results for the variables representing the theory on social capital will be explained in detail. I first examine $TRUST$, which shows a fuzzy trend. This relatively unstable pattern would be consistent with the nonsignificant coefficient obtained in the parametric models. The maximum effect of $TRUST$ on growth takes place for medium values. Then the effect progressively decreases but abruptly reaches a peak for values above 0.5 before definitely adopting a downward trend for the highest values.

The pattern for $ACTIVE$ is more clearly defined. Its associated parameter shows a parabolic functional form, thus indicating that above a given threshold, which is approximately located around 0.065, the effect of $ACTIVE$ on growth begins to decrease, thus showing that $ACTIVE$ impact obeys to the law of diminishing returns. Nevertheless, note that the error bands become wider after the maximum contribution to growth is reached which would be indicating that the variable is nonsignificant for values above this point. One likely explanation could be the little amount of regions which values above this value (only some regions in the UK and the Netherlands have such high rates of active associationism). Consequently, the precision of the estimate falls dramatically. However, for lower values of $ACTIVE$, where the bulk of observations concentrates, the errors bands are closer and $ACTIVE$ shows an upward tendency and positive effects consistent with the parametric *Model 5*.

I focus now on the partial estimated coefficients. As commented on along the paper, nonparametric regression allows for computing individual estimates for each region, which is a powerful advantage over OLS, which only provides a common estimate for all regions,

namely the average. Table 5 contains partial estimates for the different regressors across the five models. In the table are reported the first quartile (Q1), the median (Q2) and the third quartile (Q3) of the vector of estimated partial effects. Each particular estimate indicates the estimated impact of the regressor on the dependent variable assuming that the rest of variables remain constant at their median. Asymptotic standard errors are provided in parenthesis beyond each estimate. The R^2 for each model is available at the bottom of the table. Note that in all cases the goodness of fit for the nonparametric models is notably higher than that for the parametric counterparts. This gives initial support to the adequacy of adopting nonparametric methods, which offer a better performance when fitting the observed data.

The first interesting insight we might draw from Table 5 is the great variability shown by the partial effects. The estimates for GDP_0, despite showing a moderate variation across models show that the difference between the first and third quartiles are more than twice the estimated effect. In all cases, estimates are negative and significant, which corroborates that poorer regions at the beginning of the period grow faster than the richer. $GPOP$ is always positive in models 1 and 2, but the coefficients are nonsignificant for the first quartile. In models 3, 4 and 5, however, the associated partial effect is negative and significant for the first quartile, a result more according with the neoclassical growth theory. For the median and the third quartiles, the effect is again positive and significance varies across models. $GFCF$ is negative for the first and second quartiles and positive for the third. Coefficients are mostly significant across models except for *Model 4*, where only the first quartile is significant. However, recall that the bandwidths for LCLS hit the upper bound, indicating that the variable is irrelevant for predicting growth. This result is common to the parametric specification across models 1–5, where the variable is nonsignificant. HC is mainly positive across quantiles and specifications, with the exception of the first quartile in models 4 and 5. Yet, remember that HC is only a significant growth predictor in models 1 and 3, where the bandwidths associated to LCLS were below the upper bound.

Regarding the variables representing the social capital theory, the partial effects for $TRUST$ are similar for the three models where the variable is included (models 2, 4 and 5). The effects for the first quartile are negative while the median and the third quartile are positive. All coefficients are significant. These results indicate that the effect of $TRUST$ is mixed, and that might explain why we find this variable nonsignificant in the parametric models. It is likely that $TRUST$ would be positive in some regions and negative in others.

For instance, lets consider *Model 5*. The average effect obtained using OLS is equal to -0.016 and the estimate is not significant. The nonparametric counterpart of the model predicts a significant coefficient of 0.079 for the third quartile, which shows that OLS is underestimating the effect in some regions. Table 5 reports only some particular points from the vector of estimates. In order to provide a complete summary of the results, densities for the whole vector of estimated effects for all the variables are plotted by Figure 2. Figure 2 e) is that corresponding to $TRUST$. While most of the effects concentrate in the vicinity of zero, the effect is clearly positive for some regions while the contrary holds for others. This behavior might explain why $TRUST$ is nonsignificant (in average) in the parametric estimation: the positive effect in one region might be compensated by a negative effect in another.

Regarding $ACTIVE$, results indicate that it is mainly positive across quantiles and models. It only exhibits a negative sign in the first quartile of *Model 4*. In addition, differences between quantiles are notable, specially in models 4 and 5. In *Model 5*, the partial effects for the third quartile are more than double than those for the first. Compared to the parametric counterpart, we find that the latter overestimates the effect of $ACTIVE$, since the coefficient of 0.504 represents the third quartile of the vector of estimates in the nonparametric estimation. Focusing on Figure 2 f), we notice that while the bulk of the effects are in the positive side, for some regions the effect of $ACTIVE$ is manifestly negative. The amount of positive impacts is remarkably higher than the negative ones, thus giving some support to those results in the parametric analysis.

The analysis carried out in this section suggest that generalizing the effect of the variables on GDP per capita growth is no longer an appropriate strategy. For both indicators of social capital evidence of nonlinearities and heterogeneous effects is found. As a consequence, the natural next step would be identifying particular effects across different groups of regions sharing common attributes.

4.2.3. Nonparametric estimates for particular groups of regions

This section focuses on the two indicators of social capital, $TRUST$ and $ACTIVE$, and tries to identify whether the estimated effects vary across different groups of regions. In doing so, I exclusively focus on the most comprehensive *Model 5* and split the vector of the partial effects by considering regional levels of some of the variables included as regressors. Specifically, I analyze the effects of $TRUST$ and $ACTIVE$ for those regions below and

above median values of GDP_0, $GFCF$, HC, $TRUST$ and $ACTIVE$. Note that the obtained estimates are not a result of nonparametric regressions run for each particular group, but a subgroup of estimates from the whole vector of estimates selected given the existent level (below or above the median) of the above variables. In addition, the vector of estimates is also split according to geographical criteria, differentiating between ECE regions and non ECE regions.

Table 6 reports quartile coefficients for the different splits while Figures 3 and 4 offer more complete results by plotting the densities for the whole vector of estimated effects for each subgroup. For all comparisons, equality of the two densities is tested by using the Li (1996) test, which assesses the closeness of two given distributions $h(x)$ and $g(x)$. Under the null hypothesis ($H0 : h(x) = g(x)$), the two distributions are equal. Under the alternative ($H1 : h(x) \neq g(x)$), they statistically differ. Results for these tests are contained in Table 7. Except for $ACTIVE$ when comparing below and above median of this indicator, the null hypothesis of equality of distributions is rejected.

Figure 3 shows the results for $TRUST$, whose median effects are higher in regions below the median initial income. This is consistent with some previous findings in the literature suggesting that trust provides an informal framework guaranteing transactions in low-income economies with weak institutions and poor legal frames (Knack and Keefer, 1997). In contrast, Peiró-Palomino and Tortosa-Ausina (2013a) found irrelevant effects of trust for the poorest economies in a cross country sample. Note that for some low-income regions $TRUST$ effects are in the negative side, clearly below those for regions above median initial income, casting some doubts on where the effects are greater. From the comparisons between groups below and above median investment ($GFCF$) and education (HC) it is fairly difficult to extract a clear pattern. The same stands for regions below and above the median level of $ACTIVE$, whose distributions are very similar. However, the comparison between groups with different levels of $TRUST$ shows that whereas the partial estimated effects for those regions above median $TRUST$ mostly lie in the interval (-2, 2), for the group of regions below median $TRUST$ dispersion is higher, with regions with partial effects both below and above the ones for the group above the median.

The most striking result comes from the comparison between ECE and non ECE regions. $TRUST$ is mostly positive for non ECE regions but negative for an important group of ECE regions. Quartile estimates in Table 6 show that estimates for the non ECE group are always higher than those for the ECE regions group. The latter only shows positive partial effects

for the third quartile. This gives empirical support to some arguments in the literature suggesting that social capital is not always positive, but sometimes might adopt negative forms, pernicious for the general interest. It is likely that ECE countries not only show lower levels of social capital, as some authors such as Fidrmuc and Gërxhani (2008) found and the descriptive statistics corroborate, but also that social capital generate negative consequences for growth in these regions, as suggested by Rose (2000) or Paldam and Svendsen (2001). The reasons explaining the negative effect in these regions are difficult to be determined but they might be linked to the long communist experience, characterized by weak and unreliable institutions, which negatively influenced society. Rose (2000) classified ECE's societies within the group of "antimodern societies", characterized by the organizational failure and the corruption of the formal institutions. Following with Rose' (2000) arguments, under this scenario, individuals are encouraged to create networks of diffuse and informal cooperation where bribing official organisms and using connections to break the established rules become commonplace. When these actions generalize, social capital turns negative and the theoretical positive effects on economic outcomes rapidly vanish.

Focusing on $ACTIVE$, Figure 4 shows differences between regions below and above median initial income. Whereas most of the effects for the two groups are on the positive side, for an important group of poor regions the effect is negative, although admittedly disperse. The comparisons between below and above median levels of investment ($GFCF$), education (HC) or even levels of $ACTIVE$ indicate no remarkable differences (although we reject the $H0$ of equality of distributions for the two former). However, note that when considering the effects of $ACTIVE$ according to the level of $TRUST$ the results are completely analogous to those for the initial income comparison. The group above median $TRUST$ enjoys positive effects of $ACTIVE$ and the partial effects are concentrated in the relatively narrow interval (0, 1), while in the group below median $TRUST$, the estimated effects present higher dispersion and take negative values for some regions and highly positive values for a small group. What the general result would be indicating is that regions below and above median GDP_0 and $TRUST$ coincide. Comparing the quartiles for these subgroups in Table 6, we notice that they are very similar.

Finally, I focus on the comparison between ECE and non ECE regions. The results are again revealing. In the group of non ECE regions, the partial effects of $ACTIVE$ are positive with very few exceptions and they lie in the interval (0, 1). Yet, for many regions in the group of ECE regions $ACTIVE$ negatively affects growth. Dispersion is again remarkable,

since for other regions in the latter group effects are similar or even higher than those for some regions in the non ECE group. Quartile coefficients in Table 6 show that the greatest differences between ECE and non ECE regions are in the first quartile (0.18 in non ECE regions in front of -1.33 in ECE regions). The median is also higher in non ECE regions but for the third quartile we find the reversal situation—i.e. $ACTIVE$ effects are greater in ECE regions.

In general terms, the results for $ACTIVE$ suggest that the negative effects are for some regions in ECE countries, which are also those with lower levels of $TRUST$ and lower initial income (GDP_0). Recall that in the $TRUST$ analysis I only found clear evidence of different effects in the geographical comparison, but when comparing $TRUST$ and GDP_0 levels, differences were not so clear. Therefore, given that both $TRUST$ and $ACTIVE$ are social capital indicators—although they admittedly measure different aspects—my suggestion is that $ACTIVE$ effects are more subject to geographical location (ECE vs non ECE) than to the existent levels of GDP_0 or $TRUST$. The rationale is that the initial income and trust are precisely lower in ECE regions and that leads to similar results for these comparisons. In any case, the findings suggest that social capital effects are heavily heterogeneous across regions in Europe. Differences between ECE and non ECE regions seem important, thus showing that considering only average effects we are missing important subleties.

4.3. Endogeneity

Endogeneity issues are well known in growth studies. Under this circumstance, the values taken by the independent variables might be given by those of the dependent variable and the analyst has to deal with reverse causality concerns. Accordingly, a common robustness check consists of estimating with instrumental variables.

When evaluating the social capital theory as a growth determinant, however, a collection of theoretical underpinnings lead us to consider social capital variables as exogenous. The main argument lies in the stability of social values over time, especially social trust. As suggested by Bjørnskov (2012) and Bjørnskov and Méon (2013), the latter has remained stable at least from the World War II. Therefore, recent economic history might have not determined social capital levels. Nevertheless, in order to provide the results with further robustness and ensure endogeneity is not driving the results I implement an instrumental variables (IV) approach in the frame of nonparametric regression, where these methods have been developed only recently. Some alternatives on this issue are the proposals by

Newey and Powell (2003), Su and Ullah (2008), Horowitz (2011) or Darolles et al. (2011). I selected the Su and Ullah (2008) procedure, which can deal with endogenous regressors in a kernel setting. As argued by Henderson et al. (2013), empirical applications of this method are virtually nonexistent, despite its interesting properties. Lets consider again Equation 2, but now we have a single endogenous regressor p. Then the model is defined as

$$Y_i = m(Z_i, p_i) + \epsilon_i, i = 1, 2, ...n, \tag{6}$$

and

$$p_i = n(Q_i) + \epsilon_i, i = 1, 2, ...n, \tag{7}$$

where Q_i in Equation 7 is a vector of instrumental variables, $n(.)$ is an unknown smooth function and ϵ_i is the error term. The other components in both Equations 6 and 7 are common to Equation 2.

Perhaps the major problem in (IV) estimations is the choice of appropriate instruments. In the social capital literature some authors such as Akçomak and Ter Weel (2009), Tabellini (2010), Bjørnskov (2010) and Bjørnskov and Méon (2013) have proposed a variety of instruments based on geographical, political or historical variables. However, most contributions are based on cross-country samples and therefore some of the instruments such as language features (pronoun-drop characteristic) or governmental regimes (monarchy vs. republic) (see Bjørnskov and Méon, 2013) are suitable for countries but not for regions, since these variables are common in those regions belonging to the same country. In my sample, conformed by regions from both Western and Eastern Europe, difficulties exacerbate. While for Western regions the amount of available data is still reasonably high, it is far more complicated to find certain data for ECE regions and, as a result, other common instruments in the literature at the regional level such as, for instance, seventeenth century literacy rates (see Akçomak and Ter Weel, 2009) cannot be used.

In order to overcome these difficulties, instruments were chosen by following the nobel strategy introduced by Henderson et al. (2013). Then I focus on Model 5 and allow for the social capital indicators to be endogenous. Remember that the LCLS estimator for Model 4 considered three variables as irrelevant for predicting growth, namely $GPOP$, $GFCF$ and HC. However, it is likely that these variables are related to the social capital indicators.

Then, a first stage consists of running a LCLS estimation where the dependent variable is the potential endogenous regressor and the explanatory variables are those variables both relevant and irrelevant in Model 4.[16] In contrast to Henderson et al. (2013), where only one variable is potentially endogenous, in my model both indicators of social capital are treated as endogenous. Then, two different regressions are performed (one by indicator) and in this first stage the other social capital variable is not included as regressor—i.e. $ACTIVE$ is not included for predicting $TRUST$ and viceversa. Columns 2 and 3 in Table 8 show the computed bandwidths. Note that only with the exception of $GFCF$ in the $TRUST$ equation the bandwidths are below the corresponding upper bounds, thus showing that the variables are good predictors of the dependent variable—i.e. they can be considered suitable instruments.

In the second stage, the Su and Ullah (2008) estimator proposes running the original model (Model 5) including all the regressors in that model, both exogenous and potentially endogenous, as well as the estimated residuals from the first stage. Note that the first stage consisted of two separate regressions and therefore two residuals were generated, namely $\mu_{T\hat{R}UST_i}$ and $\mu_{AC\hat{T}IVE_i}$. Accordingly, these two residuals are included in the second stage. This time the model is estimated using LLLS, since as mentioned in the preceding section it is preferable to LCLS for computing estimates. Bandwidths from this estimation are available in Table 8, column 5. Results corroborate that both indicators of social capital enter the model nonlinearly also when controlling for endogeneity. Finally, Table 9 contains the quartile estimates for the instrumented Model 5. In order to ease comparisons, in columns 2–4 the original results are reported, while the results for the instrumented counterpart are contained in columns 5–7. Note that the results are qualitatively equivalent and quantitative differences are slight. Only for the variable $ACTIVE$ it is observed a higher dispersion in the estimated quartile coefficients. In the instrumented Model 5, the first quartile is negative and the third is almost double than in the original Model 5. The median effect is fairly similar. In any case, the results show that conclusions hold when endogeneity of social capital is accounted for. The results for the rest of control variables also remain virtually unaltered.

[16]See Su and Ullah (2008) for further technical details on this estimator.

5. Concluding remarks

In the last three decades the theory on social capital as a growth determinant has generated an intense academic debate. This paper feeds the discussion providing some new findings on this issue using two of the most accepted indicators of social capital, namely social trust ($TRUST$) and active participation in associations ($ACTIVE$). Two are the main contributions of the paper to the existent literature. First, in the evaluation of growth theories it is essential to consider whether the functional form of the model to be estimated is correctly specified (Henderson et al., 2012). The nonparametric techniques applied permitted taking into account this circumstance and detecting the existence of nonlinearities, not only for the social capital variables, but for the rest of usual regressors in growth models. Second, it considers the likely existence of parameter heterogeneity, thus showing that the effects of social capital widely differ across regions.

The results suggest that the linear specification imposed by the parametric methods is not the true underlaying relationship between the two indicators of social capital and growth. Neither it is for most of the variables included. I considered five different specifications and in all cases linearity was rejected as a valid specification. The nonparametric specifications showed that the effect of $TRUST$ and $ACTIVE$ heavily departs from linearity. The most striking point, however, is that $TRUST$ is not significant in the parametric analysis, which is in line with previous research for the European regions, but it is significant in the nonparametric one. Therefore, considering only the parametric scenario we would conclude that $TRUST$ is not a relevant growth predictor as result of model misspecification. Regarding $ACTIVE$, both parametric and nonparametric analysis coincided in considering this variable as significant.

Nevertheless, the average coefficient provided by the parametric analysis is far from capturing the influence of the regressors in some regions, where regressors' impact might depart remarkably from the mean value. The nonparametric methods applied permitted calculating individual partial effects and evaluating the potential parameter heterogeneity of social capital. Different subgroups of regions were conformed according to the level of initial income, investment and education levels, social capital levels, and, more interestingly, ECE and non ECE regions. Whereas for most of the comparisons the estimated effects of social capital were similar, differences exacerbated when comparing ECE and non ECE regions. In the former, social capital effects are mostly negative, whereas the contrary holds for the latter. Therefore, generalizing the positive effects of social capital is no longer

appropriate.

These results have different implications in the context of the EU Cohesion Policy. I suggest that, despite previous evidence supporting social capital as a growth predictor is abundant, social capital is not the panacea for development and in some cases could be even negative for achieving higher growth. Two aspects should be considered when designing policies addressed to improve social capital levels. First, the existent stock of social capital in each region should be considered, since the desired positive effect would be only achieved for some particular levels of social capital and it varies according to the indicator. Second, these policies should be applied carefully in some regions where we might achieve undesired effects.

Until now, we knew about the existence of negative effects of social capital. Authors such as Rose (2000) or Paldam (2000) suggested that in ECE regions negative social capital might be present. The empirical findings of this paper support these previous considerations. Unfortunately, exploring the reasons why social capital is negative in ECE regions falls out of the scope of this paper. Probably, following the suggestions of the latter authors one might think of poor and corrupted formal institutions and the legacy of communism in ECE's societies as the main responsible factors, but empirical evidence is still yet to come. Undoubtedly, it is a challenging topic for future research initiatives in the context of social capital and growth in Europe.

Acknowledgements

I am grateful to Daniel J. Henderson and Jeffrey S. Racine for kindly providing assistance with technical doubts. I also acknowledge the financial support of the Ministerio de Ciencia e Innovación (ECO2011-27227) as well as the Generalitat Valenciana (VALi+d ACIF 2080/2011). The usual disclaimer applies.

References

Aitchison, J. and Aitken, C. G. G. (1976). Multivariate binary discrimination by the kernel method. *Biometrika*, 63(3):413–420.

Akçomak, İ. and Ter Weel, B. (2009). Social capital, innovation and growth: evidence from Europe. *European Economic Review*, 53(5):544–567.

Arribas, I., Pérez, F., and Tortosa-Ausina, E. (2010). The determinants of international financial integration revisited: The role of networks and geographic neutrality. *Studies in Nonlinear Dynamics and Econometrics*, 15(1):article 4.

Badunenko, O. and Romero-Ávila, D. (2013). Financial development and the sources of growth and convergence. *International Economic Review*, 54(2):629–663.

Beugelsdijk, S. and Van Schaik, T. (2005). Social capital and growth in European regions: an empirical test. *European Journal of Political Economy*, 21(2):301–324.

Bjørnskov, C. (2006). The multiple facets of social capital. *European Journal of Political Economy*, 22(1):22–40.

Bjørnskov, C. (2009). Social trust and the growth of schooling. *Economics of Education Review*, 28(2):249–257.

Bjørnskov, C. (2010). How does social trust lead to better governance? An attempt to separate electoral and bureaucratic mechanisms. *Public Choice*, 144(1-2):323–346.

Bjørnskov, C. (2012). How does social trust affect economic growth? *Southern Economic Journal*, 78(4):1346–1368.

Bjørnskov, C. and Méon, P.-G. (2013). Is trust the missing root of institutions, education, and development? *Public Choice*, 157(3–4):641–669.

Brock, W. A. and Durlauf, S. N. (2001). What have we learned from a decade of empirical research on growth? Growth empirics and reality. *World Bank Economic Review*, 15(2):229–272.

Crespo-Cuaresma, J. C., Doppelhofer, G., and Feldkircher, M. (2012). The determinants of economic growth in European regions. *Regional Studies*, 48(1):44–67.

Darolles, S., Fan, Y., Florens, J.-P., and Renault, E. (2011). Nonparametric instrumental regression. *Econometrica*, 79(5):1541–1565.

Dearmon, J. and Grier, K. (2009). Trust and development. *Journal of Economic Behavior & Organization*, 71(2):210–220.

Dearmon, J. and Grier, K. (2011). Trust and the accumulation of physical and human capital. *European Journal of Political Economy*, 27(3):507–519.

Durlauf, S. N. and Fafchamps, M. (2005). Social capital. In Aghion, P. and Durlauf, S. N., editors, *Handbook of Economic Growth*, volume 1B, chapter 26, pages 1639–1699. North-Holland, Amsterdam.

Durlauf, S. N., Kourtellos, A., and Tan, C. M. (2008). Are any growth theories robust? *The Economic Journal*, 118(527):329–346.

Fernandez, C., Ley, E., and Steel, M. F. (2001). Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics*, 16(5):563–576.

Fidrmuc, J. and Gërxhani, K. (2008). Mind the gap! Social capital, East and West. *Journal of Comparative Economics*, 36(2):264–286.

Guiso, L., Sapienza, P., and Zingales, L. (2004). The role of social capital in financial development. *American Economic Review*, 94(3):526–556.

Guiso, L., Sapienza, P., and Zingales, L. (2009). Cultural biases in economic exchange? *Quarterly Journal of Economics*, 124(3):1095–1131.

Hall, P., Li, Q., and Racine, J. S. (2007). Nonparametric estimation of regression functions in the presence of irrelevant regressors. *The Review of Economics and Statistics*, 89(4):784–789.

Hayfield, T. and Racine, J. S. (2008). Nonparametric econometrics: The np package. *Journal of Statistical Software*, 27(5):1–32.

Henderson, D. J., Papageorgiou, C., and Parmeter, C. F. (2012). Growth empirics without parameters. *Economic Journal*, 122(559):125–154.

Henderson, D. J., Papageorgiou, C., and Parmeter, C. F. (2013). Who benefits from financial development? New methods, new evidence. *European Economic Review*, 63:47–67.

Horowitz, J. L. (2011). Applied nonparametric instrumental variables estimation. *Econometrica*, 79(2):347–394.

Horvath, R. (2012). Does trust promote growth? *Journal of Comparative Economics*, 41(3):777–788.

Hsiao, C., Li, Q., and Racine, J. S. (2007). A consistent model specification test with mixed categorical and continuous data. *Journal of Econometrics*, 140(2):802–826.

Knack, S. and Keefer, P. (1997). Does social capital have an economic payoff? A cross country investigation. *Quarterly Journal of Economics*, 112(4):1251–1288.

Li, Q. (1996). Nonparametric testing of closeness between two unknown distribution functions. *Econometric Reviews*, 15:261–274.

Li, Q. and Racine, J. S. (2004). Cross-validated local linear nonparametric regression. *Statistica Sinica*, 14(2):485–512.

Li, Q. and Racine, J. S. (2007). *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton and Oxford.

Maasoumi, E., Racine, J., and Stengos, T. (2007). Growth and convergence: A profile of distribution dynamics and mobility. *Journal of Econometrics*, 136(2):483–508.

Mankiw, N., Romer, D., and Weil, D. (1992). A contribution to the empirics of economic growth. *Quarterly Journal of Economics*, 107(2):407–437.

Monastiriotis, V. (2014). Regional growth and national development: transition in central and eastern europe and the regional kuznets curve in the east and the west. *Spatial Economic Analysis*, 9(2):142–161.

Newey, W. K. and Powell, J. L. (2003). Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578.

Nunn, N. and Wantchekon, L. (2011). The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–3252.

Paldam, M. (2000). Social capital: one or many? Definition and measurement. *Journal of Economic Surveys*, 14(5):629–653.

Paldam, M. and Svendsen, G. T. (2001). Missing social capital and the transition in Eastern Europe. *Journal of Institutional Innovation, Development and Transition*, 5(5):21–34.

Peiró-Palomino, J., Forte, A., and Tortosa-Ausina, E. (2014). Does social capital matter for European regional growth? *BBVA Foundation WP*, 06/2014.

Peiró-Palomino, J. and Tortosa-Ausina, E. (2013a). Can trust effects on development be generalized? A response by quantile. *European Journal of Political Economy*, 32:377–390.

Peiró-Palomino, J. and Tortosa-Ausina, E. (2013b). Social capital, investment and economic growth: some evidence for Spanish provinces. *Spatial Economic Analysis*, forthcoming.

Putnam, R. (1993). *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton Univ Press, New Jersey.

Racine, J. S. and Li, Q. (2004). Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics*, 119(1):99–130.

Rose, R. (2000). Getting things done in an antimodern society: social capital networks in Russia. In Dasgupta, P. and Serageldin, I., editors, *Social Capital: A Multifaceted Perspective*, pages 147–172. World Bank.

Roth, F. (2009). Does too much trust hamper economic growth? *Kyklos*, 62(1):103–128.

Sala-i-Martin, X. (1997). I just ran two million regressions. *American Economic Review*, 87(2):178–183.

Schneider, G., Plumper, T., and Baumann, S. (2000). Bringing Putnam to the European regions. *European Urban and Regional Studies*, 7(4):307–317.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Ser.B,53(3):683–690.

Solow, R. M. (1957). Technical change and the aggregate production function. *Review of Economics and Statistics*, 39(3):312–320.

Su, L. and Ullah, A. (2008). Local polynomial estimation of nonparametric simultaneous equations models. *Journal of Econometrics*, 144(1):193–218.

Tabellini, G. (2010). Culture and institutions: economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716.

Uslaner, E. (2008). Where you stand depends upon where your grandparents sat. *Public Opinion Quarterly*, 72(4):725–740.

Zak, P. J. and Knack, S. (2001). Trust and growth. *Economic Journal*, 111(470):295–321.

Żukowski, R. (2007). *Social Capital and Challenges of Development in Latin America and East Central Europe*. CESLA, Warsaw University.

Table 1: Descriptive statistics

	1995–2001						2002-2007					
	Non ECE regions			ECE regions			Non ECE regions			ECE regions		
Variable	Obs.	Mean	s.d.	Obs.	Mean	s.d.	Obs.	Mean	s.d.	Obs.	Mean	s.d.
$GGDP$	190	0.050	0.031	46	0.102	0.031	192	0.036	0.014	46	0.111	0.045
GDP_0	190	17,736	6,995	46	2,892	1,386	192	24,078	8,784	46	5,558	2,867
$GPOP$	192	0.053	0.005	46	0.048	0.004	192	0.055	0.006	46	0.048	0.003
$GFCF$	161	0.208	0.055	46	0.218	0.071	156	0.213	0.045	46	0.216	0.052
HC	189	0.214	0.083	46	0.136	0.067	192	0.246	0.081	46	0.170	0.070
$TRUST$	192	0.334	0.138	46	0.184	0.055	192	0.334	0.138	46	0.184	0.055
$ACTIVE$	192	0.037	0.022	46	0.022	0.013	192	0.037	0.022	46	0.022	0.013

Notes: ECE regions correspond to regions from Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Romania and the Slovak Republic.
Social capital variables are taken from European Values Survey, wave of 1999. These values are common in both subperiods.

Table 2: Parametric regressions

	Dependent variable: GDP growth ($GGDP$)				
Variables	Model 1	Model 2	Model 3	Model 4	Model 5
$(Intercept)$	0.405***	0.407***	0.419***	0.408***	0.401***
	(0.018)	(0.019)	(0.017)	(0.018)	(0.018)
$log(GDP_0)$	−0.039***	−0.039***	−0.041***	−0.040***	−0.040***
	(0.347)	(0.002)	(0.002)	(0.002)	(0.002)
$GPOP$	0.216	0.219***	0.073	0.049	0.116
	(0.244)	(0.244)	(0.232)	(0.232)	(0.225)
$GFCF$	−0.030	−0.031	−0.014	−0.008	−0.026
	(0.028)	(0.028)	(0.027)	(0.027)	(0.026)
HC	0.099***	0.098***	0.091***	0.094***	0.060***
	(0.017)	(0.017)	(0.016)	(0.016)	(0.017)
$TRUST$		0.003		−0.018*	−0.016
		(0.011)		(0.010)	(0.010)
$ACTIVE$			0.431***	0.463***	0.504***
			(0.064)	(0.067)	(0.066)
$CAPITAL$					0.022***
					(0.004)
N	404	404	404	404	404
$R^2(Adjusted)$	0.531	0.530	0.578	0.580	0.616
F_{STAT}	115.20***	91.94***	111.40***	93.69***	72.20***
Time control	No	No	No	No	Yes

Notes: Standard errors are in parentheses. *, ** and *** denote significance at 10%, 5%, and 1% significance levels, respectively. Invariant time effects control for common features in the two periods of analysis (1995–2001) and (2002–2007).

Table 3: Tests of appropriateness of the parametric models (Hsiao et al., 2007)

	Model 1	Model 2	Model 3	Model 4	Model 5
Jn-statistic	12.828	10.580	5.676	9.820	9.764
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes: The null hypothesis being tested is whether the parametric specification is correct $(H0 : Pr[E(x|z) = f(z, \beta)] = 1)$, against the alternative that it is not $(H1 : Pr[E(x|z) = f(z, \beta)] < 1)$. Boostraped (399 repetitions) p-values are in parenthesis.

Table 4: Nonparametric regressions, bandwidths

	Dependent variable: GDP growth ($GGDP$)										
		Model 1		Model 2		Model 3		Model 4		Model 5	
Variables/method	UB	LCLS	LLLS	LCLS	LLLS	LCLS	LLLS	LCLS	LLLS	LCLS	LLLS
$ln(GDP_0)$	1.622	0.134	0.276	0.154	0.205	0.095	0.242	0.1528	0.261	0.287	0.748
$GPOP$	0.012	0.007	0.008	**1,809**	0.005	0.006	0.007	**22,195**	0.003	0.010	**1,364**
$GFCF$	0.106	0.016	0.057	**149,738**	0.033	0.016	0.042	**383,800**	0.025	**1,149,916**	0.035
HC	0.173	0.019	0.052	**0.270**	**0.421**	0.033	0.066	**0.269**	0.147	**0.640**	0.075
$TRUST$	0.278			2.05e-06	0.059			1.16e-04	0.065	0.005	0.029
$ACTIVE$	0.043					0.007	0.012	0.017	0.027	3.0e-04	0.024
$CAPITAL$	0.500									0.499	0.007
$Time$	0.500									0.007	0.024

Notes: Bandwidths are computed using Least Squares Cross Validation (LSCV). A bandwidth in bold indicates that it exceeds the upper bound (UB). In LCLS estimations this is synonym of irrelevance of the regressor while in LLLS indicates that the regressor enters the model linearly.

Table 5: Nonparametric regressions, LLLS quartile estimates for the continuous regressors

	Dependent variable: GDP growth ($GGDP$)														
	Model 1			Model 2			Model 3			Model 4			Model 5		
Variables	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
$ln(GDP_0)$	-0.069	-0.047	-0.030	-0.071	-0.052	-0.040	-0.057	-0.041	-0.023	-0.052	-0.034	-0.021	-0.057	-0.035	-0.016
	(0.006)	(0.003)	(0.003)	(0.003)	(0.008)	(0.006)	(0.005)	(0.003)	(0.002)	(0.005)	(0.011)	(0.003)	(0.004)	(0.003)	(0.001)
$GPOP$	0.054	0.393	0.719	0.141	0.577	0.973	-0.280	0.116	0.989	-0.538	0.388	0.921	-0.152	0.168	0.663
	(0.213)	(0.097)	(0.210)	(0.338)	(0.139)	(0.202)	(0.114)	(0.214)	(0.277)	(0.364)	(0.872)	(0.347)	(0.000)	(0.000)	(0.000)
$GFCF$	-0.216	-0.091	0.025	-0.289	-0.142	0.027	-0.224	-0.065	0.088	-0.276	-0.104	0.08	-0.354	0.024	0.139
	(0.040)	(0.034)	(0.006)	(0.033)	(0.015)	(0.028)	(0.033)	(0.009)	(0.037)	(0.067)	(0.087)	(0.115)	(0.156)	(0.009)	(0.023)
HC	0.020	0.101	0.143	0.035	0.093	0.129	-0.005	0.040	0.117	0.000	0.051	0.149	-0.012	0.031	0.099
	(0.035)	(0.017)	(0.041)	(0.018)	(0.026)	(0.024)	(0.010)	(0.048)	(0.025)	(0.012)	(0.005)	(0.017)	(0.012)	(0.003)	(0.011)
$TRUST$				-0.008	0.033	0.069				-0.050	0.024	0.083	-0.018	0.014	0.079
				(0.014)	(0.011)	(0.010)				(0.028)	(0.010)	(0.026)	(0.011)	(0.003)	(0.037)
$ACTIVE$							-0.039	0.467	0.744	0.021	0.373	0.788	0.143	0.322	0.504
							(0.127)	(0.163)	(0.224)	(0.120)	(0.050)	(0.128)	(0.041)	(0.094)	(0.086)
N	404			404			404			404			404		
R^2	0.816			0.854			0.916			0.958			0.958		
Time/capital controls	No			No			No			No			Yes		

Notes: Estimates for a particular continuous regressor correspond to first (Q1), median (Q2) and third (Q3) quartile of the vector of partial effects for that regressor. Beyond each estimate the corresponding asymptotic standard error is in parenthesis. In all cases, estimations are performed considering that the rest of variables in the model remain constant at the median.

Table 6: Nonparametric regression, LLLS quartile estimates for the social capital variables in Model 5 across different subgroups

	Dependent variable: GDP growth ($GGDP$)					
	$TRUST$			$ACTIVE$		
Split/variable	Q1	Q2	Q3	Q1	Q2	Q3
Below median $ln(GDP_0)$	-0.081	0.041	0.097	0.055	0.288	0.592
	(0.018)	(0.014)	(0.025)	(0.103)	(0.077)	(0.097)
Above median $ln(GDP_0)$	-0.010	0.006	0.029	0.188	0.348	0.462
	(0.003)	(0.003)	(0.041)	(0.038)	(0.015)	(0.068)
Below median $GFCF$	-0.010	0.010	0.099	0.179	0.323	0.471
	(0.011)	(0.004)	(0.018)	(0.023)	(0.077)	(0.079)
Above median $GFCF$	-0.034	0.018	0.069	0.097	0.348	0.559
	(0.008)	(0.019)	(0.016)	(0.062)	(0.016)	(0.080)
Below median HC	-0.035	0.025	0.090	0.076	0.369	0.545
	(0.014)	(0.017)	(0.011)	(0.050)	(0.045)	(0.079)
Above median HC	-0.013	0.012	0.065	0.175	0.287	0.468
	(0.006)	(0.007)	(0.014)	(0.069)	(0.014)	(0.060)
Below median $TRUST$	-0.035	0.024	0.075	0.086	0.344	0.586
	(0.015)	(0.017)	(0.013)	(0.073)	(0.077)	(0.089)
Above median $TRUST$	-0.011	0.011	0.089	0.166	0.307	0.454
	(0.004)	(0.004)	(0.017)	(0.060)	(0.015)	(0.018)
Below median $ACTIVE$	-0.063	0.006	0.108	0.116	0.297	0.504
	(0.017)	(0.003)	(0.045)	(0.051)	(0.014)	(0.056)
Above median $ACTIVE$	-0.012	0.019	0.043	0.157	0.349	0.498
	(0.012)	(0.006)	(0.052)	(0.031)	(0.013)	(0.071)
ECE regions	-0.150	-0.086	0.045	-1.338	0.212	0.815
	(0.012)	(0.019)	(0.029)	(0.064)	(0.104)	(0.079)
Non ECE regions	-0.004	0.019	0.083	0.187	0.328	0.469
	(0.003)	(0.008)	(0.016)	(0.052)	(0.070)	(0.060)

Notes: Values correspond to the first (Q1), median (Q2) and third (Q3) quartiles of the vector of partial effects for $TRUST$ and $ACTIVE$ in each subgroup. Beyond each estimate the corresponding asymptotic standard error is in parenthesis.

Table 7: Nonparametric comparison (Li, 1996) of the estimated densities for different subgroups in Model 5

		$TRUST$	$ACTIVE$
Below vs. above GDP_0	t-statistic	46.951	13.288
		(0.000)	(0.000)
Below vs. above $GFCF$	t-statistic	17.757	7.163
		(0.000)	(0.000)
Below vs. above HC	t-statistic	12.338	12.150
		(0.000)	(0.000)
Below vs. above $TRUST$	t-statistic	23.646	12.003
		(0.000)	(0.000)
Below vs. above $ACTIVE$	t-statistic	2.768	0.271
		(0.002)	(0.393)
ECE vs. non ECE regions	t-statistic	36.520	59.054
		(0.000)	(0.000)

Notes: p-values are in parenthesis.

Table 8: Nonparametric instrumental variable estimation of Model 5 (Su and Ullah, 2008), bandwidths

	Stage I (LCLS)			Stage II (LLLS)	
	UB	D.V: $TRUST$	D.V: $ACTIVE$	UB	D.V: $GGDP$
$ln(GDP_0)$	1.622	0.111	0.179	1.622	1.181
$GPOP$	0.012	0.002	0.005	0.012	**1,832.48**
$GFCF$	0.106	**0.117**	0.017	0.106	0.028
HC	0.173	0.015	0.024	0.173	0.067
$TRUST$	0.278			0.278	0.071
$ACTIVE$	0.043			0.043	0.021
$CAPITAL$	0.500			0.500	0.001
$Time$	0.500			0.500	0.020
$\mu_{\widehat{TRUST}}$				0.147	0.043
$\mu_{\widehat{ACTIVE}}$				0.024	0.012

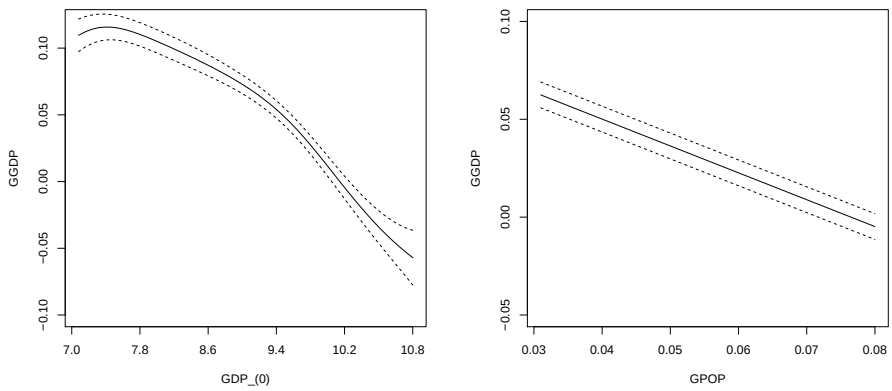
Notes: D.V. refers to "Dependent Variable". Bandwidths are computed using Least Squares Cross Validation (LSCV). A bandwidth in bold indicates that it exceeds the upper bound (UB). In LCLS estimations this is synonym of irrelevance of the regressor while in LLLS indicates that the regressor enters the model linearly. $\mu_{\widehat{TRUST}}$ and $\mu_{\widehat{ACTIVE}}$ are the residuals from Stage I to be included in Stage II.

Table 9: Nonparametric instrumental variables regression, LLLS quartile estimates for the continous variables in the instrumented Model 5

	Dependent variable: GDP growth ($GGDP$)					
	Model 5			IV Model 5		
Variables	Q1	Q2	Q3	Q1	Q2	Q3
$ln(GDP_0)$	-0.057	-0.035	-0.016	-0.050	-0.039	-0.028
	(0.004)	(0.003)	(0.001)	(0.000)	(0.000)	(0.003)
$GPOP$	-0.152	0.168	0.663	-0.257	0.298	1.221
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$GFCF$	-0.354	0.024	0.139	-0.288	0.006	0.105
	(0.011)	(0.004)	(0.018)	(0.020)	(0.025)	(0.022)
HC	-0.012	0.031	0.099	-0.048	0.038	0.112
	(0.008)	(0.019)	(0.016)	(0.062)	(0.016)	(0.080)
$TRUST$	-0.018	0.014	0.079	-0.046	0.034	0.121
	(0.011)	(0.003)	(0.037)	(0.008)	(0.018)	(0.020)
$ACTIVE$	0.143	0.322	0.504	-0.183	0.298	0.976
	(0.041)	(0.094)	(0.086)	(0.063)	(0.095)	(0.017)
N	404			404		
R^2	0.958			0.980		

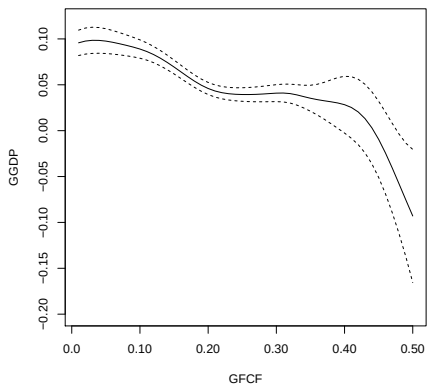
Notes: Estimates for a particular continuous regressor correspond to first (Q1), median (Q2) and third (Q3) quartile of the vector of partial effects for that regressor. Beyond each estimate the corresponding asymptotic standard error is in parenthesis. In all cases, estimations are performed considering that the rest of variables in the model remain constant at the median.

Figure 1: Nonparametric regression (LLLS), Model 5

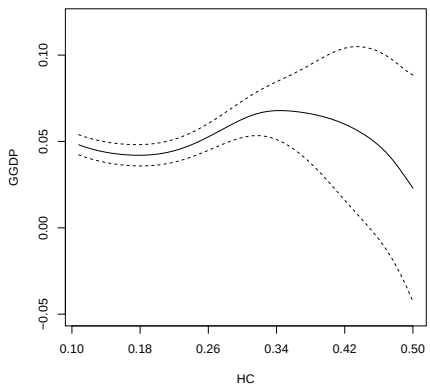


(a) *GDPPCI*

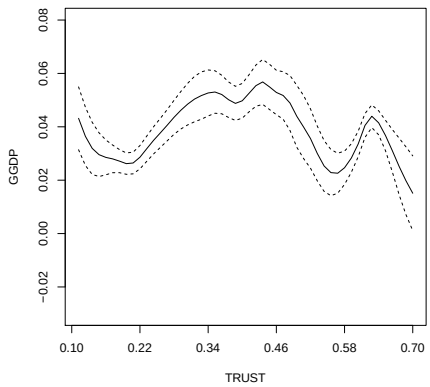
(b) *GPOP*



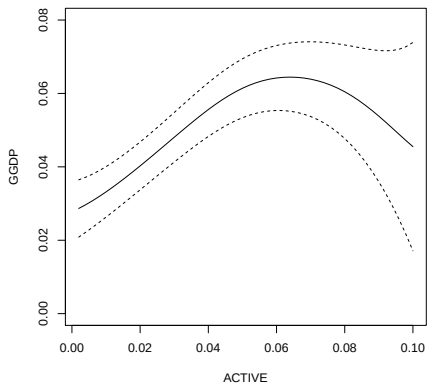
(c) *GFCF*



(d) *HC*



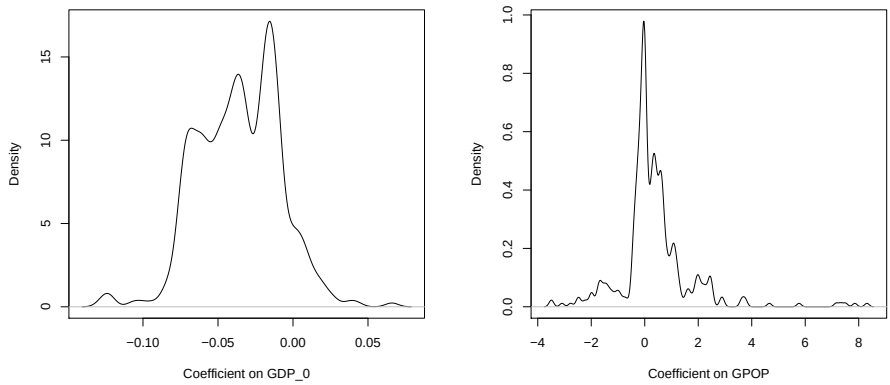
(e) *TRUST*



(f) *ACTIVE*

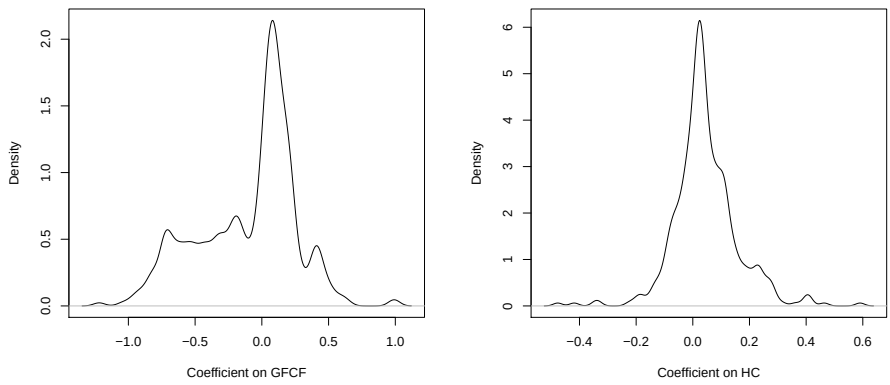
Notes: In the plots, the solid line represents the estimated function for each regressor and the dotted lines are the associated standard errors.

Figure 2: Densities for the LLLS estimated coefficients for the continuous variables in Model 5, Sheather and Jones's (1991) bandwidths



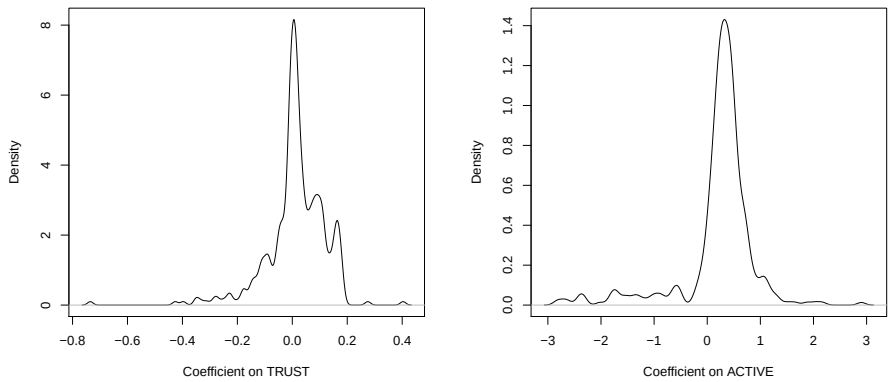
(a) *GDPPCI*

(b) *GPOP*



(c) *GFCF*

(d) *HC*



(e) *TRUST*

(f) *ACTIVE*

Figure 3: LLLS estimated coefficients for TRUST for specific groups in Model 5, Sheather and Jones's (1991) bandwidths

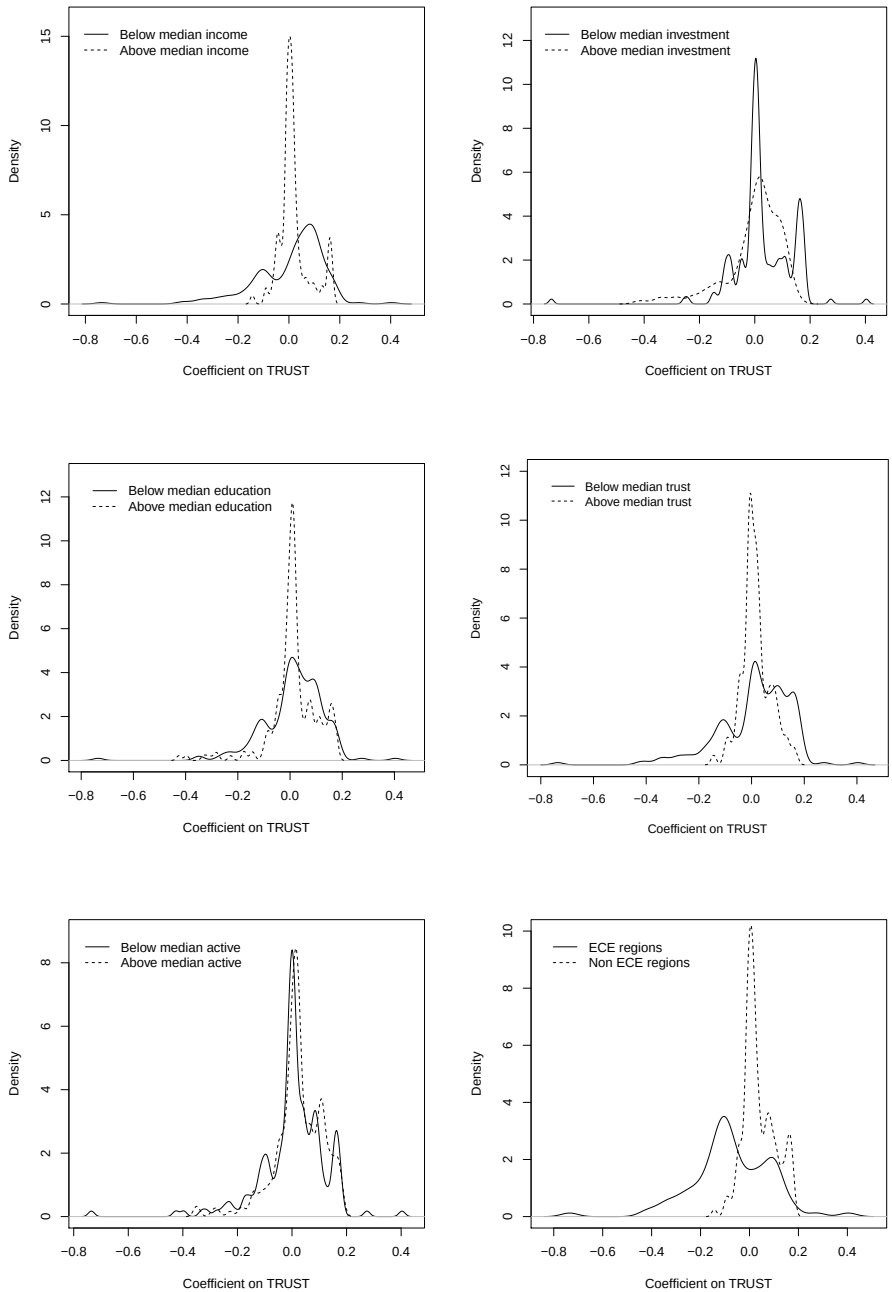


Figure 4: LLLS estimated coefficients for ACTIVE for specific groups in Model 5, Sheather and Jones's (1991) bandwidths