THE UNIVERSITY OF WARWICK

June 2012 No.89

Absorptive Capacity and the Growth and Investment Effects of Regional Transfers:

Regression Discontinuity Design with Heterogeneous Treatment Effects

Sascha O. Becker, Peter H. Egger and Maximilian von Ehrlich

WORKING PAPER SERIES

Centre for Competitive Advantage in the Global Economy

Department of Economics

CAGE
Competitive Advantage in the Global Economy

E·S·R·C
ECONOMIC & SOCIAL RESEARCH COUNCIL

Absorptive Capacity and the Growth and Investment Effects of Regional Transfers

A Regression Discontinuity Design with Heterogeneous Treatment Effects*

Sascha O. Becker† Peter H. Egger‡ and Maximilian von Ehrlich§

March 1, 2012

Abstract

Researchers often estimate *average* treatment effects of programs without investigating heterogeneity across units. Yet, individuals, firms, regions, or countries vary in their ability, e.g., to utilize transfers. We analyze Objective 1 Structural Funds transfers of the European Commission to regions of EU member states below a certain income level by way of a regression discontinuity design with systematically *heterogeneous* treatment effects. Only about 30% and 21% of the regions – those with sufficient human capital and good-enough institutions – are able to turn transfers into faster per-capita income growth and per-capita investment. In general, the variance of the treatment effect is much bigger than its mean.

Key words: Regional transfers; Absorptive capacity; Heterogeneous local average treatment effects; Regression discontinuity design.
JEL classification: C21; O40; H54; R11

**Acknowledgements*: We would like to thank two anonymous referees and the editor in charge (Alan Auerbach). We gratefully acknowledge comments from seminar participants at the Universities of Linz, Lugano, Maastricht, Stirling, Tübingen, Wisconsin-Madison, Zurich, at Royal Holloway London, at the Austrian National Bank, at the European Meeting of the Urban Economics Association, at the German Institute of Economic Research (DIW), and at the 8th Christmas Conference of German Expatriate Economists, Berlin.

†*Affiliation:* U Warwick, CAGE, CEPR, CESifo, Ifo, and IZA. *Address:* Department of Economics, University of Warwick, Coventry CV4 7AL, United Kingdom. s.o.becker@warwick.ac.uk.

‡*Affiliation:* ETH Zürich, CAGE, CEPR, CESifo, Federal Reserve Bank of Dallas, Leverhulme Centre for Research on Globalisation and Economic Policy (GEP) at the University of Nottingham, and Oxford University Centre for Business Taxation (OUCBT). *Address:* ETH Zürich, KOF, Weinbergstrasse 35, WEH E6, 8092 Zürich, Switzerland. egger@kof.ethz.ch.

§*Affiliation:* ETH Zürich. *Address:* ETH Zürich, KOF, Weinbergstrasse 35, WEH F9, 8092 Zürich, Switzerland. vonEhrlich@kof.ethz.ch.

1 Introduction

All developed countries use transfers to individuals, firms, or regions to stimulate a change in their behavior or their performance. A prominent example of such transfers is aid in a very broad sense. Transfers may be used as a mere means of redistribution at the level of individuals (possibly but not necessarily with a change in behavior in mind).[1] At the level of firms, regions, or countries, transfers tend to be used to stimulate recipient units' performance (employment, investment, and growth), at least in the medium run.

A significant literature in applied econometrics is concerned with the ex-post estimation of the causal effects of such transfers or programs with non-random assignment. Related work tends to focus on data where treatment assignment is captured by a binary indicator variable and the single parameter of interest measures an *average treatment effect* (ATE). We argue that units and their responsiveness to transfer treatment vary systematically so that a focus on the ATE alone conceals important information to both the econometrician and the policy maker.

With regions or countries, transfers are often aimed at fostering investment and, ultimately, economic growth. Examples of such transfers are un-tied aid programs administered by the World Bank through the International Bank for Regional Development, the European Bank for Regional Development at the international level, or the European Union's (EU) transfers across regions of the Union's member countries. While the justification for and the effectiveness of aid programs is hotly debated,[2] the literature seems to agree that key factors that undermine the goal of aid transfers are low levels of education and bad institutions (such as corrupt politicians, bad administrations, etc.).[3] In broad terms, we refer to human capital and quality of institutions as two dimensions of *absorptive capacity* with regard to rendering transfers effective.[4] Hence, our concept of *absorptive capacity* refers on the one hand to

[1] E.g., a negative income tax may be used as a transfer of the former kind and a training program as a transfer of the latter kind.

[2] For instance, see Dalgaard, Hansen, and Tarp (2004) who argue in favor of an aid-growth link, and Easterly (2003) who questions the effectiveness of aid with respect to economic growth.

[3] Already early work on the effects of aid transfers pointed to the role of education and skills for the responsiveness of countries to such transfers (Rosenstein-Rodan, 1961; Chenery and Strout, 1966; see Chauvet and Guillaumont, 2004, for a more recent argument along those lines). The direct link between political institutions, aid transfers, and economic growth is addressed in Burnside and Dollar (2000, 2004).

[4] In general, human capital (see Mankiw, Romer, and Weil, 1992; Benhabib and Spiegel, 1994; and Becker, Hornung, and Woessmann, 2011) and good institutions (see Mauro, 1995; and Acemoglu, Johnson, and Robinson, 2005) have been found to stimulate economic growth and catch-up of lagging-behind countries or regions *per se*. The arguments made and evidence provided by Rosenstein-Rodan (1961), Chenery and Strout (1966), Burnside and Dollar (2000, 2004), and Chauvet and Guillaumont (2004) suggests that absorptive capacity is also important for transfers

the ability of employing transfers in the most productive way by having a sufficient level of human capital and on the other hand to the ability of allocating funds in an efficient way, which is facilitated by a high quality of institutions.

While most of the arguments along those lines are made in the context of international aid, they also apply for sub-national aid flows among relatively homogeneous countries within the European Union: bad institutions are often mentioned as one reason for why regional transfers are not as effective as they could be,[5] and through capital-skill complementarity a lack of skilled workers in some recipient regions should be considered an important source of lower returns on investment in such regions (see Duffy, Papageorgiou, and Perez-Sebastian, 2004). However, while there is no debate about the qualitative importance of institutions and education not only for economic prosperity but also for returns on investment and transfers, little is known about the quantitative relevance of these factors for the response of investment and economic growth to transfers. Yet, for policy makers the latter is crucial with aid programs and transfers that operate under budget constraints and with tax payers calling their effectiveness into question.

The notion that the responses to fiscal policy at large may be heterogeneous is at the heart of a recent literature at the interface of macroeconomics and public finance. In broad terms, two strands of such research may be distinguished. One line of research on heterogeneous responses to fiscal stimuli supports the view that effects on the same recipients of identical stimuli *vary over time*. For instance, Auerbach and Gorodnichenko (2010, 2011) provide evidence of state-dependent magnitudes of fiscal multipliers. While there is evidence of a positive effect of fiscal multipliers over the long run (see Gemmell, Kneller, and Sanz, 2011; Ramey, 2011), Auerbach and Gorodnichenko's (2010, 2011) findings suggest that effects can vary significantly over the business cycle. Another line of research on heterogeneous responses to fiscal stimuli – which is more closely related to this paper's agenda – supports the view that effects in the same period and of identical stimuli *vary across recipients*. For instance, Suárez Serrato and Wingender (2011) exploit variation *across US counties* in receipt of US federal grants that depend on local population levels to estimate local fiscal multipliers. They find heterogeneity of the impact of government spending with a higher impact in low-growth regions in comparison to high-growth regions. Shoag (2011) uses variation in portfolio returns of defined-benefit pension plans – for which state governments bear the investment risk – *across US states* to identify the effect of state government spending on in-state income and employment. He detects heterogeneity in that the effect is stronger in non-tradable industries (and, consequently, in states which are relatively specialized on such industries) and when

to be used productively by recipients.

[5]For instance, bad institutions in Greece are deemed responsible for the lack of exhaustion of budgeted transfers (Pisani-Ferry, Marzinotto, and Wolff, 2011).

and where economic activity is low.[6]

This paper is devoted to a quantitative analysis of the *heterogeneity* of the treatment effect of the EU's main regional transfer program. The European Commission takes a number of initiatives to foster per-capita income growth and convergence. Such initiatives are subsumed under two major funding programmes: the Structural Funds – which are composed of the European Regional Development Fund (ERDF) and the European Social Fund – and the Cohesion Fund. We focus on the biggest among all initiatives, referred to as Objective 1, whose goal is to provide transfers to the poorest regions of the EU to foster their catch-up towards the EU average. Eligibility for Objective 1 transfers is associated with a discontinuity about GDP per-capita: only regions whose per-capita income (in purchasing power parity) falls short of 75% of the EU average prior to a programming (or budgetary) period are eligible for such transfers during that period.[7] The regional entities which may claim eligibility are so-called NUTS2 regions.[8] The goal of this study is to identify the magnitude at which the effect of Objective 1 treatment on investment and economic growth varies with the quality of institutions and the level of education across regions by using the discontinuity in per-capita income for treatment eligibility as an identification strategy in a regression discontinuity design (RDD). Commonly used RDDs ignore treatment effect heterogeneity in the neighborhood of the threshold and focus on a single local *average* treatment effect (LATE). By way of contrast, we extend that concept to the case of *heterogeneous* local average treatment effects (HLATE). In particular, we wish to infer how the treatment effect of Objective 1 regional transfers varies with the quality of government and the level of human capital – which we interpret as measures of *absorptive capacity* – of targeted regions. The latter allows us to inform policy makers about the extent of the variability of the treatment effect, about reasonable options of redistribution with minimal costs in regions from which and maximal effects in regions whereto funds are redistributed, and about the magnitude of foregone growth stimuli associated with the currently applied scheme.

The paper contributes to two literatures. First, it formulates a flexible RDD which is applicable with fixed but arbitrary numbers of forcing variables (determin-

[6]While Suárez Serrato and Wingender (2011) as well as Shoag (2011) rely on instrumental variable approaches, we use a regression discontinuity design, since the European Union's transfer scheme gives rise to a discontinuity of transfer eligibility depending on regional per-capita GDP, as will become clear later.

[7]Programming periods in EU jargon last for 5 to 7 years. The three most recent Programming periods were 1989-1993, 1994-1999 and 2000-2006. Eligibility is determined in pre-specified years prior to a programming period.

[8]NUTS is the acronym for *N*omenclature des *U*nités *T*erritoriales *S*tatistiques coined by Eurostat, the Statistical Office of the EU, which refers to regional aggregates. NUTS2 regions correspond to groups of counties and unitary authorities with a population of 0.8-3 million inhabitants.

ing treatment eligibility) and fixed but arbitrary numbers of variables the treatment effect varies with. For such designs, we formulate an RDD for the HLATE and illustrate that nonparametric estimators work comparatively well relative to parametric estimators of the multivariate control function, even in small to medium-sized samples. Obviously, with regional per-capita income levels prior to a programming period as one forcing (threshold) variable and Objective 1 treatment interaction with institutional quality and the level of education, the application of interest here is a special case of that general design. Second, with regard to the literature on effects of transfer treatment – such as national or regional aid, of which EU Objective 1 transfers are a prominent example – we shed light on the quantitative importance of institutions and education for the treatment effect heterogeneity on economic outcomes such as per-capita investment and per-capita income growth.

The empirical application to the EU's Objective 1 transfer treatment reveals a great variability of the effect with institutional quality and the level of education. Not only regions but also countries vary substantially with regard to the institutional quality and education level of their NUTS2 regions. This leads to a significant variability of the magnitude of treatment responses not only across regions within countries but also across member countries of the EU. We find that the HLATE is positive at a confidence bound of at least 90% for not more than 21% of the regions with per-capita investment and for 30% of the regions with per-capita income growth. Hence, at that level of statistical significance, politicians could redistribute funds from 140 and 157 regions, respectively, and benefit some regions' investment and growth without harming other regions. Hence, we find that a minimal level of institutional quality and education is necessary for recipient regions to absorb transfers effectively.

The remainder of the paper is organized as follows. In the next section, we outline the econometric model with RDDs for the HLATE in general terms. An Appendix provides evidence on the small sample performance in terms of bias and root mean squared error for identification of the HLATE in the distribution of treatment effects with nonparametric versus parametric control functions. In Section 3, we summarize features of the EU's Objective 1 transfer program. In Section 4, we estimate the HLATE based on an RDD with data on all NUTS2 regions of 25 EU countries and evaluate the role of institutional quality and the level of education for the effectiveness of Objective 1 transfers for regional investment and economic growth. Section 5 concludes with a summary of our main findings.

2 RDD for heterogeneous treatment effects

Our focus is on identification of heterogeneous treatment effects with an RDD where the heterogeneity of treatment effects pertains to interactions with exogenous observable variables.[9] A standard fuzzy RDD, which contains the sharp RDD as a limiting case, exploits discontinuities in the probability of treatment conditional on *one* forcing variable. The result is a research design where the rule giving rise to the discontinuity becomes an instrumental variable for the actual treatment status. In a fuzzy RDD, one can identify a local average treatment effect (LATE) in the sense of Imbens and Angrist (1994). The LATE is the average treatment effect for *compliers*, i.e., those treated who take the treatment only when eligible, but do not get treated when ineligible. Our aim is to employ estimators, where the estimated treatment effect is not (only) local in the sense of being a LATE, but local and heterogeneous in the sense that it is allowed to vary with a fixed but arbitrary number of observables. Accordingly, we refer to this as a heterogeneous local average treatment effect (HLATE). We will allow heterogeneity of treatment effects to vary with variables that do or do not influence treatment status. Moreover, for the sake of generality, we will consider the case of a fixed but arbitrary number of forcing variables (and, hence, discontinuities at potentially more than one treatment threshold) as well as a fixed but arbitrary number of exogenous variables interacting with the treatment effect. Angrist and Fernandez-Val (2010) analyze heterogeneity in local average treatment effects that is caused by instrument-specific compliant subpopulations. They show that differences in observables can be used to explain the heterogeneity of LATE when using different instruments. In Angrist and Fernandez-Val (2010) different instruments yield different experiments the treatment effect is identified from, whereas our approach rests on one instrument and accordingly one set of compliers. An implicit assumption of our estimation strategy is that the distribution of compliers is not affected by the interaction variables. This seems justified since, in our application, we find no systematic correlation between the misclassification of observations according to the treatment rule and their measures of absorptive capacity.

In the following, we outline parametric as well as nonparametric identification strategies for the most general case with many forcing variables and many variables affecting the treatment effects. Building on this, we compare the performance of

[9]The proposed estimators should not be interpreted as to support an unmotivated search for significant treatment effects in subsamples of the data with different characteristics. Such a strategy would involve well-known problems with multiple testing, resulting in an over-rejection of the null hypothesis (of no treatment effect or no heterogeneity of the treatment effect; see Kling, Liebman, and Katz, 2007). But rather, we propose estimators that are applicable where economic theory suggests heterogeneity of a certain kind to materialize in the data.

the derived estimators in Monte Carlo studies (see Appendix C) where we focus on designs which permit graphical illustration.[10]

2.1 Definition of heterogeneous local average treatment effects (HLATE)

Let us use the following notation. First, use T_i to denote a treatment indicator which is equal to one if treatment is received by unit i and zero otherwise. Outcome y_i is a function of the treatment, of the $1 \times K$ vector $\mathbf{x_i}$ of forcing variables, and of the $1 \times L$ vector $\mathbf{z_i}$ of interaction variables that render treatment more or less effective but do not affect treatment assignment. We seek to estimate the heterogeneous local average treatment effect

$$HLATE(\mathbf{x_i} = \mathbf{x_0}, \mathbf{z_i}) = HLATE(\mathbf{x_0}, \mathbf{z_i}) = E[y_{1i}|\mathbf{x_0}, \mathbf{z_i}] - E[y_{0i}|\mathbf{x_0}, \mathbf{z_i}], \qquad (1)$$

where y_{1i} denotes the outcome with treatment and y_{0i} the outcome without treatment, and $\mathbf{x_0}$ denotes the $1 \times K$ vector of threshold values x_{0k} for the K forcing variables with $k = 1, ..., K$.

The challenge for treatment evaluation arises because we observe each unit i only in one of two mutually exclusive states of the world, either with or without treatment, and treatment assignment is not random but depends on the information in $\mathbf{x_i}$. In contrast to the commonly identified local average treatment effect (LATE), the HLATE above allows for variation in the dimensions of $\mathbf{z}_i$. This flexibility is particularly valuable as in many cases where the LATE is not different from zero, the HLATE may vary substantially around the LATE.

In the RDD, the treatment probability is a discontinuous function of the forcing variables

$$P(T_i = 1|\mathbf{x_i}) = \begin{cases} \mathbf{g_1}(\mathbf{x_i}) & \text{if } x_{ik} \geq x_{0k} \ \forall \ k \in K \\ \mathbf{g_0}(\mathbf{x_i}) & \text{otherwise.} \end{cases} \qquad (2)$$

The literature distinguishes two types of RDD: the sharp design where $\mathbf{g_1}(\mathbf{x_0}) - \mathbf{g_0}(\mathbf{x_0}) = 1$ and the fuzzy design where $0 < \mathbf{g_1}(\mathbf{x_0}) - \mathbf{g_0}(\mathbf{x_0}) < 1$. Accordingly, in the sharp design, the treatment probability jumps from zero to one once all K treatment rules are satisfied while the probability jump is less than one in the fuzzy design where treatment assignment is noisy due to exemptions from the rules.

Regardless of whether a sharp or a fuzzy design prevails, the HLATE can be estimated parametrically or nonparametrically under the following assumptions:

[10] Obviously, with more than two variables enforcing treatment status or co-determining treatment effects, graphical illustration becomes difficult. Specifically, we will illustrate two scenarios in the Appendix: a *1-way threshold* scenario where the forcing variable is independent of the variable that interacts with the treatment effect, and a *2-way threshold* scenario with two forcing variables, one of which is allowed to simultaneously affect the magnitude of treatment effects.

Assumption 1 *(Continuity of counterfactual outcomes at threshold vector.)* *$E[y_0]$ and $E[y_1]$ are continuous at $\mathbf{x_0}$.*

This is the standard identifying assumption in an RDD. It states that there should not be any jump in other observable determinants of outcome beyond the forcing variable.

Assumption 2 *(Continuity of interaction variables at threshold vector.)* *The interaction variables $\mathbf{z_i}$ are continuous at $\mathbf{x_0}$.*

This assumption is important for the HLATE to pick up genuine variation in the interaction variables. In our application, we check this assumption by plotting graphs for human capital and quality of government to see whether these measures of absorptive capacity are discontinuous about the forcing variable at the threshold or not (see Figure 2 below).

Assumption 3 *(Random assignment of interaction variables $\mathbf{z_i}$ conditional on $\mathbf{x_i}$.)* *The interaction variables $\mathbf{z_i}$ are uncorrelated with the error term in the outcome equation, conditional on $\mathbf{x_i}$.*

In the context of our application, this assumption states that, conditional on GDP per capita (the forcing variable), regions with different human capital endowments and quality of governance do not differ in unobserved dimensions which are relevant for investment or per-capita income growth. Take the example of two regions with the same pre-treatment level of GDP per capita that differ in their human capital endowment. The fact that, despite different human capital endowment, they achieved the same pre-treatment GDP per capita indicates that there were other factors which, in the past, led the two regions to achieve the same pre-treatment level of GDP per capita. For instance, regions in former communist countries with high human capital endowments might have achieved the same (low) per-capita income as some Western European regions with low human capital endowments. The omitted factor would be the past experience of a communist system in place. Assumption 3 states that such other factors do not systematically contribute contemporaneously to investment or economic growth. We address this particular concern in several ways: first, we run fixed effects regressions (amongst others) which wipe out time-constant factors such as past communist political system experience. Furthermore, we take the absorptive capacity interaction variables as time-constant variables, so that the HLATE picks up factors that facilitate or hinder the effective use of EU transfers over longer horizons. Both human capital endowment and quality of government are factors which hardly vary over time and are thus relatively stable attributes of regions.

In the following, we outline the estimation approaches formally, where the sharp RDD can be understood as a special case of the fuzzy RDD with treatment assignment being a deterministic function of the forcing variables while the fuzzy design allows for some randomness in treatment assignment.

2.2 Parametric control function for identification of the HLATE

Assuming that $E[y_i|\mathbf{x_i}, \mathbf{z_i}]$ follows an additive process based on polynomial functions of the columns of $\mathbf{x_i}$ and $\mathbf{z_i}$ we can write the conditional expected outcomes in the counterfactual situations of treatment and non-treatment as follows:

$$E[y_{0i}|\mathbf{x_i}, \mathbf{z_i}] = \alpha + \mathbf{f_0}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{h_0}(\overline{\mathbf{z}}_\mathbf{i}) \tag{3}$$

$$E[y_{1i}|\mathbf{x_i}, \mathbf{z_i}] = E[y_{0i}|\mathbf{x_i}, \mathbf{z_i}] + \beta + \mathbf{f_1^*}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{h_1^*}(\overline{\mathbf{z}}_\mathbf{i}), \tag{4}$$

where $\mathbf{f_0}(\widetilde{\mathbf{x}}_\mathbf{i})$, $\mathbf{h_0}(\overline{\mathbf{z}}_\mathbf{i})$, $\mathbf{f_1^*}(\widetilde{\mathbf{x}}_\mathbf{i})$, and $\mathbf{h_1^*}(\overline{\mathbf{z}}_\mathbf{i})$ are sufficiently smooth polynomial functions of the columns of $\mathbf{x_i}$ and $\mathbf{z_i}$.[11] In order to simplify the interpretation of the coefficients, we define the parametric functions $\mathbf{f_0}(\cdot)$ and $\mathbf{f_1^*}(\cdot)$ in terms of deviations of x_{ik} from the thresholds x_{0k} and $\mathbf{h_0}(\cdot)$ and $\mathbf{h_1^*}(\cdot)$ in terms of deviations of z_{il} from the sample means $E[\mathbf{z}_l]$. Accordingly, $\widetilde{x}_{ik} = x_{ik} - x_{0k}$ and $\overline{z}_{il} = z_{il} - E[\mathbf{z}_l]$. Overall, we can then write

$$E[y_i|\mathbf{x_i}, \mathbf{z_i}] = E[y_{0i}|\mathbf{x_i}, \mathbf{z_i}] + T_i[\beta + \mathbf{f_1^*}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{h_1^*}(\overline{\mathbf{z}}_\mathbf{i})]. \tag{5}$$

Using this notation, the local average treatment effect at the multidimensional threshold level of the forcing variables, $\mathbf{x_0}$, is given by β. The heterogeneous treatment effect for deviations from the sample means in the z-dimensions is measured by $HLATE(\mathbf{x_0}, \mathbf{z_i}) = \beta + \mathbf{h_1^*}(\overline{\mathbf{z}}_\mathbf{i})$.

In the sharp RDD, where the treatment is a deterministic function of the set of forcing variables, we can estimate the treatment effects by the following regression:

$$y_i = \alpha + \mathbf{f_0}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{h_0}(\overline{\mathbf{z}}_\mathbf{i}) + T_i\left[\beta + \mathbf{f_1^*}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{h_1^*}(\overline{\mathbf{z}}_\mathbf{i})\right] + \epsilon_i \tag{6}$$

$$where\ T_i = 1(x_{ik} \geq x_{0k}\ \forall\ k \in K). \tag{7}$$

In the fuzzy RDD, even though the treatment probability jumps when crossing the multidimensional threshold $\mathbf{x_0}$, as indicated in (2), T_i is no longer a deterministic function of $\mathbf{x_0}$. Hence, the identifying assumption of the sharp RDD in (7) is

[11] We use a notation where $\mathbf{f_1^*}(\cdot) \equiv \mathbf{f_1}(\cdot) - \mathbf{f_0}(\cdot)$ and $\mathbf{h_1^*}(\cdot) \equiv \mathbf{h_1}(\cdot) - \mathbf{h_0}(\cdot)$, and where $\mathbf{f_1}(\cdot)$ and $\mathbf{h_1}(\cdot)$ in the treatment state are defined analogously to $\mathbf{f_0}(\cdot)$ and $\mathbf{h_0}(\cdot)$ in the no-treatment state. More generally, one can also allow for interaction terms between columns of $\mathbf{x_i}$ and $\mathbf{z_i}$ and add those interaction terms as additional elements in a (new) $\mathbf{z_i}$ with larger column space. Hence, we will not specifically address this issue. But we note that, in our application, all results are robust to the introduction of interaction terms between (polynomials of) the forcing variable and (polynomials of) the measures of absorptive capacity.

violated. This requires us to specify some functional form for the conditional treatment probability $P(T_i = 1|\mathbf{x_i})$. Let us define a scalar $R_i = 1(x_{ik} \geq x_{0k} \ \forall \ k \in K)$ indicating whether *all rules* underlying the treatment status are fulfilled or not. When $\mathbf{g_1}(\mathbf{x_i})$ and $\mathbf{g_0}(\mathbf{x_i})$ in (2) can be approximated sufficiently well, R_i may serve as an instrument for $P(T_i = 1|\mathbf{x_i})$ conditional on $\mathbf{g_1}(\mathbf{x_i})$ and $\mathbf{g_0}(\mathbf{x_i})$. Using analogous notation as for the outcome, we may determine $\mathbf{g_0}(\widetilde{\mathbf{x}}_\mathbf{i})$, $\mathbf{g_1^*}(\widetilde{\mathbf{x}}_\mathbf{i}) \equiv \mathbf{g_1}(\widetilde{\mathbf{x}}_\mathbf{i}) - \mathbf{g_0}(\widetilde{\mathbf{x}}_\mathbf{i})$, $\mathbf{l_0}(\overline{\mathbf{z}}_\mathbf{i})$, $\mathbf{l_1^*}(\overline{\mathbf{z}}_\mathbf{i}) \equiv \mathbf{l_1}(\overline{\mathbf{z}}_\mathbf{i}) - \mathbf{l_0}(\overline{\mathbf{z}}_\mathbf{i})$. In the first stage of the 2SLS implementation of the fuzzy RDD we estimate:[12]

$$T_i = \mathbf{g_0}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{l_0}(\overline{\mathbf{z}}_\mathbf{i}) + R_i[\delta + \mathbf{g_1^*}(\widetilde{\mathbf{x}}_\mathbf{i}) + \mathbf{l_1^*}(\overline{\mathbf{z}}_\mathbf{i})] + \nu_i. \tag{8}$$

The forcing variables are again measured in terms of deviations from the respective thresholds. Substituting (8) for the treatment indicator T_i in (6) we obtain the reduced form for the fuzzy RDD. Equations (6) and (8) together constitute the IV estimator of the $HLATE(\mathbf{x_0}, \mathbf{z_i})$.

2.3 Nonparametric control function for identification of the HLATE

The parametric estimates of the treatment effects rely on the validity of the approximations $\mathbf{f_0}(\cdot)$, $\mathbf{f_1^*}(\cdot)$, $\mathbf{h_0}(\cdot)$, $\mathbf{h_1^*}(\cdot)$, $\mathbf{g_0}(\cdot)$, $\mathbf{g_1^*}(\cdot)$, $\mathbf{l_0}(\cdot)$, and $\mathbf{l_1^*}(\cdot)$. As has been shown by Hahn, Todd, and van der Klaauw (2001), average treatment effects can be identified nonparametrically under much weaker assumptions (basically continuity restrictions only). This section introduces the nonparametric approach to the estimation of the HLATE.

In a standard RDD with one forcing variable, where $\mathbf{x_i}$ is a scalar and $\mathbf{z_i}$ is absent from the model, identification and consistent estimation of the LATE hinges upon estimation of $E[y_i|x_i]$. In the more general design analyzed here, we have to estimate $E[y_i|\mathbf{x_i}, \mathbf{z_i}]$ in the neighborhood of the multidimensional discontinuity. It can be shown that the HLATE at the multidimensional threshold is given by (see Appendix A for a proof):

$$HLATE(\mathbf{x_0}, \mathbf{z_i}) = \lim_{\mathbf{\Delta} \to \mathbf{0}} \frac{E[y_i|\mathbf{0} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{\Delta}, \mathbf{z_i}] - E[y_i| - \mathbf{\Delta} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{0}, \mathbf{z_i}]}{E[T_i = 1|\mathbf{0} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{\Delta}, \mathbf{z_i}] - E[T_i = 1| - \mathbf{\Delta} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{0}, \mathbf{z_i}]}, \tag{9}$$

where $\mathbf{\Delta}$ denotes a vector of some small, positive deviations from zero. In the sharp RDD the denominator in (9) is simply unity whereas it ranges between zero and one in the fuzzy RDD.

[12] Alternatively, the first stage may be estimated by a nonlinear model. In our application, the results remain unaffected by the choice of a linear or nonlinear first stage.

As pointed out by Hahn, Todd, and van der Klaauw (2001), standard kernel estimators for the above conditional expectations to the left and the right of the threshold yield biased estimates for the treatment effects due to their adverse boundary properties. At boundary points the kernel estimators have a slower rate of convergence than at interior points. Therefore, Hahn, Todd, and van der Klaauw (2001) propose using local linear regressions instead of standard kernel estimates. In our case with multiple interaction and forcing variables we resort to multivariate local polynomial regressions as introduced by Ruppert and Wand (1994).

Let us collect all columns in $\widetilde{\mathbf{x}}_\mathbf{i}$ and in $\overline{\mathbf{z}}_\mathbf{i}$ in the vector $\boldsymbol{\xi}_i$. The first K columns of $\boldsymbol{\xi}_i$ belong to the columns of $\widetilde{\mathbf{x}}_\mathbf{i}$ and the second L columns belong to $\overline{\mathbf{z}}_\mathbf{i}$. We aim at estimating the expectations of y_i in the neighborhood of the multidimensional threshold for given values of $\overline{\mathbf{z}}_\mathbf{i}$. Hence, we fit a polynomial in the neighborhood of a vector $\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{0}$. The local linear estimator for $\lim_{\boldsymbol{\Delta}\to 0} E[y_i|\mathbf{0} < \widetilde{\mathbf{x}}_\mathbf{i} < \boldsymbol{\Delta}, \overline{\mathbf{z}}_\mathbf{i}]$ is given by:

$$\min_{b_0,\mathbf{b}_1} \sum_{i=1}^{N} \{y_i - b_0 - \mathbf{b}_1^T \boldsymbol{\xi}_i\}^2 \mathbf{K}_\mathbf{H}(\boldsymbol{\xi}_i) * 1(\widetilde{\mathbf{x}}_\mathbf{i} > \mathbf{0}), \tag{10}$$

where $\mathbf{K}_\mathbf{H}$ represents a kernel function with bandwidth matrix $\mathbf{H}$. In our applications, we generally use a uniform kernel. For further details on the use of local polynomial regressions we refer to Härdle, Müller, Sperlich, and Werwatz (2004). $HLATE(\mathbf{x_0}, \mathbf{z_i})$ is asymptotically normally distributed as shown in Appendix B.

3 The EU Objective 1 transfer program

Objective 1 funds are one of the biggest expenditure items in the EU budget. They are of particular interest to understand the role of regional transfers for regional investment and growth in Europe, because their explicit aim is to foster investment and per-capita income growth in regions that are lagging behind the EU average and of promoting aggregate growth in the EU (European Commission, 2001). Objective 1 funds are part of the EU's Structural Funds Programme which in turn is the second-largest budget item alongside agricultural expenditure. Within the Structural Funds Programme budget, Objective 1 transfers are by far the largest item in all of the last three so-called programming periods: accounting for 70% of the budget in the 1989-93 period, for 68% in the 1994-99 period, and for 72% in the 2000-06 period (see European Commission, 1997, p. 154 f., and European Commission, 2007, p. 202).

Eligibility for Objective 1 transfers is associated with a discontinuity in real GDP per-capita levels: only regions whose per-capita income (in purchasing power parity) is below 75% of the EU average prior to a programming period are eligible for such transfers. Eligibility is determined in pre-specified years prior to a programming period. The regional entities which may claim eligibility are NUTS2 regions as described above. If the 75%-rule were strictly followed by the EU authorities, there should be perfect compliance, giving rise to a sharp regression discontinuity design. However, some regions that are not formally eligible end up receiving Objective 1 funds and some regions that are eligible do not receive Objective 1 funds. Whereas the reasons for the former might have to do with EU power politics, i.e., some regions negotiating exceptions from the rule, the latter is surprising at first. Measurement error about per-capita income is the main explanation for that: the regional GDP figures that national statistical agencies report to Eurostat may be inaccurate (or even not available at all) at the time eligibility for funding is determined and might later be revised by Eurostat with no effect on foregone funding. In our data, which are revised GDP figures available from Cambridge Econometrics, some regions are below the 75% threshold ex post, but they were not at the time the EU Commission had to decide on eligibility. For instance, the United Kingdom did not deliver GDP data at the required NUTS2-level to the European Commission at the time Objective 1 status was determined in the programming period 1989-93. Only ex post, when the data became available, it turned out that some British NUTS2 regions should have been eligible for Objective 1 funds.

Non-compliance is an issue in 7% of observations, i.e., the design is not quite sharp but fuzzy. In the fuzzy design, the 75%-rule serves as an instrument for actual Objective 1 treatment. In previous work, we have analyzed the LATE of Objective 1 transfers on regional economic growth and found that they had a positive average effect (Becker, Egger, von Ehrlich 2010). The present paper goes substantially beyond earlier work (including ours) on aggregate treatment effects of aid programs in general and European regional redistribution programmes in specific. First, we provide evidence that the main effect of the Objective 1 programme on economic growth runs through investment (as intended). Second, we provide evidence of substantial *heterogeneity* of the LATE in terms of education of the population and institutional quality (our two measures of absorptive capacity). The latter seems particularly important if programmes operate under budget constraints and funding reductions have to be implemented. The method proposed and the evidence provided may assist policy makers in redistributing funds in a way so that detrimental effects on investment and economic growth are minimized while keeping budgetary discipline.

4 The HLATE of EU Objective 1 transfers depending on absorptive capacity

In this section, we provide parametric and nonparametric estimates of the HLATE of Objective 1 treatment on regional per-capita investment and regional growth of per-capita income in the EU. The heterogeneity of the response to Objective 1 treatment is modeled as a function of human capital endowments and/or the quality of regional government as two measures of absorptive capacity.

4.1 Data and descriptive evidence

We use data on NUTS2 regions for the last three completed EU programming periods: 1989-93, 1994-99, and 2000-06. Due to enlargements of the EU during the observation period, the number of NUTS2 regions covered varies between 186 and 251 per period. Hence, a regional unit may be observed in the data once, twice, or thrice. Of course, repeated observation of cross-sectional units should be respected in estimation either by clustering of standard errors or alternative treatment of fixed region-specific effects. For instance, the standard errors of HLATE should be block-bootstrapped (across all years; see Fitzenberger, 1998; and Becker and Egger, 2011, for an application).[13]

For the question of interest, we utilize four types of data from different sources. First, information on NUTS2 regional per-capita GDP at purchasing power parity (PPP) is available from the Regional Database compiled by Cambridge Econometrics. The corresponding data can be utilized to calculate the level of regional average per-capita income in the years specified by the European Commission prior to each programming period – the forcing variable for Objective 1 treatment eligibility. NUTS2 regions whose per-capita GDP fell short of 75% of the EU average were eligible to receive Objective 1 funds from the EU. The same regional GDP data can be employed to determine average annual growth of per-capita income in PPP terms during a programming period.

Second, information about actual Objective 1 treatment is available directly from the European Commission, from various Council Regulations, in particular the Regulations numbered 2052/88, 2082/93, and 502/1999, and in editions of the Official Journal. The data show that there is a discrepancy between the rule and actual treatment, which establishes a fuzziness: about 7% of the data points represent non-compliers with the 75% assignment rule.

[13] In general, this strategy is consistent with the proposal of Anderson (2008, p. 1483f.) to use sampled standard errors instead of Huber-White standard errors in contexts such as ours.

Third, we use two conceptual measures of absorptive capacity at the regional level, human capital endowments and quality of regional government. We employ data on the level of education of the workforce in a region from the European Union Labour Force Survey. More specifically, we utilize data on the share of workers with at least secondary education and allow the response to Objective 1 treatment to vary with it.[14] In a sensitivity analysis, we employ data from the European Values Study (initiated by the European Values Systems Study Group, EVSSG) for the years 1981, 1990, and 1999, to obtain an alternative, complete panel data-set of human capital endowment through interpolation. Appendix E provides details on the construction of time-variant measures of absorptive capacity.

Data on the regional quality of government (QoG) come from various sources. On the one hand, we employ time-invariant data from Charron, Lapuente, and Dykstra (2011). They use a perception-based indicator of QoG based on a 34,000-respondents survey. Their data-set is available for download and contains information at the national level for all 27 EU countries and, at the sub-national level, for 172 NUTS 1 and NUTS 2 regions in the European Union for the year 2009.[15] The variable is standardized within the EU (mean of 0 and standard deviation of 1), such that higher scores equal higher levels of QoG. The QoG index is based on 16 separate survey questions pertaining to three key public services – education, health care, and law enforcement. The respondents were asked to rate their public services with respect to three related concepts of QoG – the quality, impartiality, and level of corruption of the above-mentioned services. Even though the institutional framework is determined to a large extent on the national level, our survey data shows that the implementation of the legal, the education or the health care system varies substantially across regions within countries. Prominent examples for large within country differences in government and administration are Northern and Southern Italy or Flanders and Wallonia in Belgium. Charron, Lapuente, and Dykstra (2011) point

[14] Eurostat delivered NUTS2-level data on education of the workforce for the years 1999 through 2008. Education is measured in three categories, based on UNESCO's International Standard Classification of Education (ISCED): low education refers to ISCED categories 0-2. Medium education refers to ISCED categories 3 and 4, and high education to ISCED categories 5 and 6. Our measure of (at least) upper-secondary education includes ISCED categories 3 to 6. In our sample of NUTS2 EU regions, the correlation coefficient between the share of the work force with at least upper-secondary education in 1999 and in 2008 is 0.91, which shows the stability of human capital endowment over time and makes it an interesting stable measure of the absorptive capacity of a region. The data on human capital represent an (in time) unequally spaced panel. For the baseline regressions, we averaged human capital for each region across the years.

[15] Countries with NUTS 1 level information are Belgium, Germany, Greece, Hungary, Netherlands, Sweden, and United Kingdom. NUTS2 level information is available for Austria, Bulgaria, Czech Republic, Denmark, France, Italy, Poland, Portugal, Slovak Republic, Romania, and Spain. We assume that quality of government is at least as time invariant as human capital endowment.

out that the difference in QoG is more pronounced between the two Italian regions Bolzano and Campania than it is between the national averages of Denmark and Portugal. For a sensitivity analysis we extend our study to a time-variant measure of QoG which we construct from Eurobarometer Survey information.[16,17]

Summary statistics for all variables used in our application are provided in Table 1. As suggested in Section 2, we measure absorptive capacity variables – human capital (HC) and quality of government (QoG) – as deviations from the sample mean. The forcing variable corresponds to average GDP per capita in the threshold years that were crucial to assigning eligibility for Objective 1 transfers. Table 1 reports per-capita income in the threshold years in absolute terms and as a fraction of average EU per-capita income. The Objective 1 binary treatment variable indicates transfer recipience. GDP per capita growth is measured in nominal terms in the average year of the budgetary period and represents together with the per-capita investment levels – measured in logarithmic terms – our outcome of interest.

– Table 1 –

In terms of specification, our model corresponds to the parametric or nonparametric case of a 1-way treatment threshold (in the forcing variable GDP per capita relative to the EU average) in Section 2 and an interaction with one or two regressors. In fact, we present results separately for three cases: (a) HC as the only indicator of absorptive capacity; (b) QoG as the only indicator of absorptive capacity; (c) both HC and QoG as indicators of absorptive capacity which matter simultaneously. A difference with respect to the Monte Carlo analysis in Appendix C lies in the use of repeated observations of cross-sectional units which we allow for in order to exploit variability in the data (taking account of repeated observations in the computation of standard errors throughout).

[16]For this, in particular, we use data from the Mannheim Panel 1970-2002 of the Eurobarometer Survey for the EU15 NUTS2 regions and the original Eurobarometer Survey data for the year 2003 and the New Member Countries' NUTS2 regions of the so-called Eastern Enlargement of the EU in 2004. The variables (scores) used from this survey are *Satisfaction with Democracy* and *Satisfaction with Rule of Law*. In order to gain a unique measure of quality of government, we regress the time-invariant measure by Charron, Lapuente, and Dykstra (2011) on these scores and predict (extrapolate) it for all years covered. See Appendix E for more details.

[17]As an alternative to QoG we considered the impact of social capital (SC), based on data about the voter turnout regarding European Parliamentary elections at the NUTS2 level (from various sources). We find a similar qualitative impact of SC on the effectiveness of Objective 1 transfers as of QoG. However, we suppress these results here for two reasons, namely for the sake of brevity and because QoG appears to be a more direct measure of the quality of institutions than voter turnout. In any case, results on the role of SC for the responsiveness to Objective 1 transfers are available from the authors upon request.

– Figure 1 and 2 –

Before turning to regressions, it is useful to have a look at the raw data when pooling them across all three programming periods. Figure 1 depicts the fraction of treated observations against their initial per-capita GDP relative to the EU average – in bins of a width of 1.5 percentage points in the forcing variable – in the years critical for determining Objective 1 eligibility. The discontinuity at 75% is evident, but the design is fuzzy because a number of regions does not comply with the treatment rule.

Such a discontinuity does not appear when plotting equivalent graphs for human capital and quality of government (see Figure 2). Note that this supports Assumption 2 underlying the HLATE, which requires the interaction variables to be continuous at the threshold(s) of the forcing variable(s).[18]

– Figure 3 and Figure 4 –

Unlike RDD plots for homogeneous LATE, the graphs in Figures 3 and 4 are three-dimensional, similar to those shown in the Monte Carlo analysis in Appendix C. They are useful to visualize the interaction between the forcing variable (initial GDP per capita relative to the EU average in a period), the variables relating to absorptive capacity (HC and QoG as deviations from the respective EU average), and the outcome variables (GDP per capita growth and ln of per-capita investments, respectively). Notice that these figures are generated for the subspace of values of HC and QoG where we have relatively good support (see Figure 16 in the Appendix for frequency plots of the data). Since the rule is not applied sharply by the Commission, we expect both treated units (marked by red dots) and untreated units (marked by blue dots) just above and below the threshold of the forcing variable (i.e., at a level of 0.75 or 75%). The surfaces are estimated by using 5th-order polynomial functions in the forcing variable and linear functions of the absorptive capacity variables. These surfaces are estimated separately for both sides of the threshold. The figure clearly points to a continuous impact of the forcing variable on the outcome, and to a discontinuity at the 75% threshold which in turn varies significantly with absorptive capacity. The data indicate a smaller (or even non-existent) treatment effect at the threshold for regions with below-average absorptive capacity and a higher treatment effect for regions with above-average absorptive capacity. The wedges between the two surfaces in the HC and QoG plots indicate heterogeneity of the LATE. Note, however, that the HLATE cannot be directly "inferred" from the wedges in Figures

[18] See also Becker, Egger, and von Ehrlich (2010, Figure 4) for evidence on the absence of discontinuities in other covariates, supporting Assumption 1.

3 and 4. The wedges between the surfaces disregard fuzziness about Objective 1 status, i.e., the true treatment effect needs to take account of the size of the jump in the treatment probability at the 75% threshold.

Hence, we proceed with parametric instrumental variable regression analysis and with nonparametric regression analysis to avoid a bias of the heterogeneous treatment effects accruing to fuzziness.

4.2 Regression results

A first step to scrutinize the heterogeneity of treatment effects displayed in Figures 3 and 4 is to split the sample into observations featuring below- and above-average absorptive capacity and to estimate the LATE for each of these subsamples separately using the fuzzy RDD estimator. Regarding the HC variable employed here, we observe 336 observations with an above-average HC endowment and 310 observations below the average level of HC. The former group exhibits a LATE of about 1.1 percentage points for per-capita income growth which is significant at 5%. In contrast, the LATE for per-capita income growth of the latter group amounts to about 0.2 percentage points and is insignificant. With regard to per-capita investment both groups' LATEs turn out insignificant at the usual levels of significance. Regarding QoG, the 412 observations with an above-average level of QoG feature a LATE of 2.4 percentage points for per-capita income growth – significant at 1% – and one of 34 percent for per-capita investment – significant at 5%. The LATE for the group of below-average QoG (228 obervations) turns out to be insignificant at the usual levels of significance for per-capita income growth and negative for per-capita investment. These first crude results point to a considerable heterogeneity of treatment effects.

Yet, the split of the sample may be arbitrary and we loose substantive information and efficiency by collapsing the two continuous measures of absorptive capacity into binary indicators. A more efficient way of taking into account the heterogeneity of the LATE is to follow the identification strategy for the HLATE as introduced in Section 2. The corresponding regression results are summarized in Tables 2-3 for parametric polynomial instrumental-variable regressions.[19] Each of these tables is organized horizontally in six columns and vertically in four blocs. The six columns refer to 3rd-order, 4th-order, and 5th-order polynomial specifications of the control function of the forcing variable with pooled OLS and fixed NUTS2 region-specific effects each. The four vertical blocs pertain to results for per-capita income growth (or per-capita investment levels) with time-invariant and time-invariant HC and, alternatively, QoG interactions each. In these tables, we report only results for

[19]We present nonparametric regression results only graphically in order to save space.

interactions of the Objective 1 treatment indicator with linear terms of HC and QoG.[20] In general, the main effect of Object 1 in Tables 2-3 roughly corresponds to the LATE, since neither HC nor QoG display a discontinuity at the threshold of the forcing variable (see Figure 2).

– Tables 2-3 –

In almost all specifications of Tables 2-3, both the main effect (LATE) of Objective 1 treatment and the interaction terms with HC and QoG are statistically significant at conventional levels. This means that there is a positive response of outcome – per-capita income growth and per-capita investment levels – to treatment on average, and the response becomes bigger with better endowments of human capital and better institutions.[21] Moreover, we find comfort in the similarity of the pooled OLS and the corresponding fixed effects results on the one hand and in the results based on time-invariant HC or QoG and time-variant measures thereof on the other hand. The former suggests that the RDD is powerful in removing the bias of omitted but possibly confounding unobserved time-invariant variables, while the latter suggests that reliance on time-invariant interacted variables (which are partly measured at the end of the observation period) versus time-variant ones (which are measured at the beginning of each programming period) does not induce a bias. Tables 2-3 suggest two novel insights relative to earlier work.

First, there is evidence of a systematic heterogeneity of the response to Objective 1 treatment with regard to what we call absorptive capacity – measured by HC and QoG. For instance, taking the pooled OLS specification with a 3rd-order polynomial function of the forcing variable and time-invariant HC in Table 2 as the benchmark, a region whose HC is raised by one standard deviation relative to the average receives an Objective 1 HLATE which is $100 \cdot 0.144 \cdot 0.044 \cong 0.63$ percentage points higher than the average HLATE. Notice that such a statement would not be possible with an approach where the sample is split ex ante into a high-HC and a low-HC sub-sample. In comparison, a region whose QoG (time-invariant) is raised by one standard deviation relative to the average receives an Objective 1 HLATE on per-capita income growth which is $100 \cdot 0.815 \cdot 0.005 \cong 0.41$ percentage points higher per annum than the average HLATE. Hence, the responsiveness to Objective 1 treatment is slightly more elastic in HC-space than in QoG-space.

[20]Results for higher polynomial specifications of the interactive variables are available from the authors upon request. However, the more flexible polynomial specifications as well as the nonparametric results reported below indicate that one may safely model the interactions linearly without inducing much bias.

[21]The estimates of the LATE in Table 2 correspond quantitatively to the ones in Becker, Egger, and von Ehrlich (2010).

Second, the positive effect on per-capita income growth goes along with – and is likely intermediated by – an effect on per-capita investment levels. The findings in Table 3 suggest that Objective 1 transfers stimulate investment on net. Hence, more investment is created than is eventually crowded out. Note that no such effects could be detected for employment (or unemployment rates). Accordingly, the Objective 1 transfers stimulate a factor bias in growth: there is new net investment without a short-to-medium-term effect on employment or unemployment.[22] The pooled OLS specification with a 3rd-order polynomial function of the forcing variable and time-invariant HC as the benchmark in Table 3 suggest that a region whose HC is raised by one standard deviation relative to the average receives an Objective 1 HLATE on per-capita investment levels which is $100 \cdot 0.144 \cdot 0.793 \cong 11.42\%$ higher than the average HLATE on ln of investment per capita. In comparison, a region whose QoG (time-invariant) is raised by one standard deviation relative to the average receives an Objective 1 HLATE on investment which is $100 \cdot 0.815 \cdot 0.153 \cong 12.47\%$ higher than the average HLATE.

More detailed insights into the variability of the HLATE of HC or QoG on per-capita income growth and per-capita investment levels can be gained from an inspection of Figures 5-8. The four figures are organized in four panels each. The two panels on the left-hand side of each figure are based on parametric 3rd-order polynomial control functions about real per-capita income, and the two on the left-hand side employ nonparametric control functions.[23] The two panels at the top of each figure are based on time-invariant interaction variables and the ones at the bottom on time-variant counterparts. Figures 5 and 6 address estimates of the HLATE for average annual per-capita income growth, while Figures 7 and 8 refer to average log per-capita investment levels per annum as outcome.

– Figures 5-8 –

In general, the nonparametric estimates in Figures 5-8 display somewhat wider confidence intervals than their parametric counterparts, as we would have expected from the discussion in Section 2. As a consequence, the 90%-confidence intervals of the parametric HLATE estimates are contained in the ones of their nonparametric counterparts. The nonparametric HLATE functions are somewhat steeper in both HC- and QoG-space than their parametric counterparts but not significantly so,

[22]Tables for employment and unemployment corresponding to Tables 2-3 are available from the authors upon request but suppressed here for the sake of brevity. These tables clearly support the aforementioned conclusions.

[23]In general, we follow Ludwig and Miller (2007), Imbens and Lemieux (2008), and Lee and Lemieux (2010, p. 328) in choosing an optimal bandwidth based on the second-stage local linear regression.

and the HLATE functions are steeper in QoG-space than in HC, irrespective of whether per-capita income growth or per-capita investment levels are considered as the outcome. Many but not all of the panels suggest that the average HLATE (i.e., the LATE) is significantly different from zero. In particular, this is true for per-capita investment levels as an outcome in Figures 7 and 8. However, it is important to note that this result may be – and actually is – driven by the omission of the yet other interaction term (HC in the QoG regressions and vice versa): notice that all panels in Figures 5-8 average over a third dimension in the background. In a next step, we will demonstrate that disentangling the role of HC and QoG for outcome is important and reveals a complementary effect of these measures of absorptive capacity.

– Table 4 and Figure 9 –

Table 4 provides parametric 3rd-order and 4th-order polynomial regression results with an integrated RDD model where the HLATE depends on both HC and QoG simultaneously. For reasons of flexibility, we allow the interactive effect Object 1×HC×QoG to be different where both HC and QoG are negative. Since HC and QoG are correlated to some extent (with a correlation coefficient of 0.48), not all parameters of interest can be estimated at high precision. However, the terms involving HG and the ones involving QoG, respectively, are jointly significant at conventional levels according to Wald tests. Indeed, this can be seen from an inspection of Figure 9. This figure displays the HLATE of Objective 1 treatment on annual GDP per capita growth (in the top panel) and on investment per capita levels (in the panel at the bottom) based on a 3rd-order parametric specification of the control function.

Two remarks are in order for an interpretation of Figure 9. First, both panels are three-dimensional plots with HC and QoG on the two horizontal axes and the outcomes on the vertical axes and they involves four colors: red for a positive HLATE and blue for a negative one with light and dark colors for statistically insignificant and statistically significant HLATE point estimates at 10%, respectively.[24] Second, there is a kink at HC$=0$ and QoG$=0$ in the upper panel of the figure which relates to the difference in coefficients for Object 1×HC×QoG (at HC≥ 0 *or* QoG≥ 0) and for Object 1×HC×QoG×I (at HC<0 *and* QoG<0) in Table 4. Since the corresponding coefficients are rather similar for per-capita investment levels as outcome, there is no such discontinuity in the lower panel of the figure.

Figure 9 suggests that there is a minimal level of both HC and QoG necessary for Objective 1 treatment to affect annual per-capita income growth and annual per-capita investment levels positively. At too low a level of one or the other, there is

[24]The confidence bounds are generally block-bootstrapped with 500 replications.

even a danger that Objective 1 treatment reduces the outcome, e.g., through crowding out of investment or of productive economic activity in general. The figure indicates that positive per-capita income growth effects are achieved in a larger HG-QoG-space than this is the case for per-capita investment levels. We may interpret this finding as to suggest some crowding out of investment that would have taken place also in the absence of Objective 1 treatment and of some consumption effects of treatment without any net effects on investment. Notice that the figure suggests that the HG-QoG space with positive (statistically significant or insignificant) HLATEs is marginally larger than the one with negative HLATEs for per-capita income growth while it is much smaller than the one with negative HLATEs for per-capita investment levels. However, the size of the respective HG-QoG subspaces per se is not as interesting without knowing where the data points on EU regions are situated on the surface. The latter information is provided in Table 5 (see also Figure 16 for the distribution of HC and QoG).

– Table 5

Table 5 provides information on the percentage of Objective 1 regions among the EU member countries that received Objective 1 funds and had at the same time sufficiently high levels of human capital and quality of government for realizing positive treatment effects that are significantly different from zero according to the reported confidence bounds. We report results for per-capita income growth in the upper bloc and for per-capita investment levels in the lower bloc of the table and for each country as well as the average economy. The table contains four columns with percentages. The first one is based on a 90% confidence bound and, hence, reflects the percentage of observations which are situated in the dark red areas in Figure 9. The last column is based on the point estimates – disregarding the standard errors – and reflects the percentage of observations which are situated in the dark-red and light-red areas in Figure 9. The columns for 80% and 70% confidence intervals in comparison to the remaining ones provide some evidence as to how close the data points in the light-red area are located to the boundary between the light-red an the dark-red areas in the corresponding panels of Figure 9.

Overall, the figure suggests that HG and QoG is higher in the richer economies of the EU so that Objective 1 treatment is more likely to trigger a positive response in outcome there than elsewhere. However, the intention of the program is to foster cohesion and stimulate investment mainly in regions which lag behind the EU average. Such regions are mainly located in the poorer countries such as the new member states joining the EU in the last programming period covered in this study

(2000-06). Hence, let us compare and focus on these countries[25] and on the cohesion countries in the EU15 area[26] in the discussion of Table 5. The table suggests that, among the considered countries, only Estonia is well-enough endowed with HC *and* QoG as to expect a positive income growth response to treatment even within the 90% confidence bound (the dark-red area in Figure 9) for all recipient regions in the country. Even with a 70% confidence bound (i.e., a significance level of 15%), only Estonia, Hungary, Ireland, Latvia, and Slovenia have a full coverage of their regions among the considered economies with a positive HLATE at that significance level. The last column in Table 5 suggests that a fair share of regions (in the light-blue and dark-blue areas of Figure 9) is estimated to not respond positively to Objective 1 treatment at any significance level.

Of those regions that do experience a significant growth effect from Objective 1 funds, only a subgroup is estimated to respond via investments. At a 90% confidence bound, only Objective 1 regions in Austria, Finland, Germany, the Netherlands, Sweden, and in the United Kingdom display a significant and positive HLATE on per capita investments. The remaining regions' treatment effect seems to operate via stimulated consumption which should generate only temporary income effects. Accordingly, for these regions we are sceptical about whether the transfers will be able to contribute to regional cohesion in the medium to long-run.

4.3 Policy considerations

The results provoke a number of alternative policy conclusions. Figure 9 suggests that significantly positive effects of Objective 1 transfers are only to be had with sufficiently high levels of human capital endowments (HC) and quality of government (QoG). This is the case for only a fraction of the recipients (see Table 5). When using a confidence bound of 90% as in the first column of Table 5, one could say that the European Commission could save money by voiding Objective 1 transfers to about 70% percent of the recipients (in the dark-blue, light-blue, and light-red areas of Figure 9). For the most part, those regions belong in the group of least-developed regions within the EU. By the same token, the Commission could stimulate further growth by reallocating transfers from the just-mentioned 70% of the recipient regions without any positive significant response to the remaining 30% ones (in the dark-red area of Figure 9).[27] Either measure would counteract the very purpose of

[25]The corresponding countries in the sample are: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovenia, Slovak Republic.

[26]These are: Greece, Ireland, Portugal, and Spain.

[27]Even when considering all regions in the light-red or dark-red areas of Figure 9 and, hence, focusing on the point estimates in Table 5, the European Commission could be advised to reallocate

the programme, though, which is reducing per-capita income gaps and stimulating convergence from the tails towards the average within the European Union.

An alternative proposal would be to use funds at the Structural Programme in a more discretionary fashion than at present and to target human capital formation and the development of political as well as administrative institutions (quality of government) in regions which are eligible for transfers. According to our findings, such an approach would be largely complementary to other means of redistribution. In terms of labels of initiatives at the level of the European Commission, this could be seen as an argument in favor of strengthening and broadening efforts around measures taken under the auspices of the *Regional Competitiveness and Employment Objective* (formerly Objective 2) rather than the *Convergence Objective* (formerly Objective 1). Of course, significant changes in the response to transfers induced by such measures should not be expected to happen in the very short run. Both the formation of human capital as well as institutional change take time – most likely about one generation rather than a small number of years. But the returns on those investments in terms of growth effects might be higher than ones on infrastructure and other types of real investments to regions that lack complementary factors such as skilled workers or high-quality institutions to realize the expected growth stimuli.

On a broader scale, the notion that fiscal policy induces heterogeneous responses across recipients is consistent with recent findings on fiscal multiplier effects as mentioned in the introduction, in particular, findings relating to the cross-sectional heterogeneity of such effects. It is interesting to compare the results on transfer response heterogeneity of EU regions with those on the heterogeneity of fiscal multiplier effects in US regions in Shoag (2011) and Suárez Serrato and Wingender (2011). The results of both Shoag (2011) and Suárez Serrato and Wingender (2011) suggest that the responsiveness to fiscal stimuli is better in low-growth (low-economic-activity) regions. Our results complement those findings by suggesting that the availability of human capital endowments and good institutions are crucial for recipients to make productive use of transfers as was argued in earlier work on the effectiveness of foreign aid at the level of countries.

5 Discussion and conclusions

This paper studies the role of absorptive capacity of regions in translating transfers into in per-capita investments and income growth. In particular, we study the importance of absorptive capacity for the treatment effect triggered by regional

transfers from 39% of the regions (in the light-blue and dark-blue areas of Figure 9) to the remaining 61% ones, if per-capita income growth was the Commission's only goal.

transfers under the auspices of Objective 1 under the Structural Funds Programme of the European Commission. A region's initial GDP per capita relative to the EU average determines eligibility of NUTS regions in the European Union to receive transfers out of the Structural Funds budget. Regions whose initial GDP per capita is less than 75% of the EU average are eligible to receive Objective 1 funds. Econometrically, this gives rise to a regression discontinuity design (RDD). To the extent that a region's absorptive capacity systematically influences how efficiently it uses transfers received, we expect heterogeneity in local average treatment effects (LATE) which varies with the recipient region's absorptive capacity. We derive a heterogeneous LATE (HLATE) estimator for the general scenario with multiple thresholds and various interaction variables that affect the treatment effect's magnitude, and we allow for a fuzzy treatment assignment mechanism. In a Monte Carlo simulation, we study the performance of parametric and nonparametric identification strategies for the HLATE and show that both approaches yield consistent estimates.

In our empirical illustration, we show that the heterogeneity of recipient regions with respect to their absorptive capacity matters considerably. Both measures of a region's absorptive capacity, the human capital endowment of the workforce and quality of government, show similar patterns. While the treatment effect is insignificant for regions with a very low level of absorptive capacity it exceeds the average treatment effect for regions with above-average absorptive capacity. We find that only about 30% of the recipient regions display sufficient levels of absorptive capacity to turn the transfers into economic growth. The quality of government and human capital endowments are even more decisive for the investment effects from transfers. Only 20 percentage points of the 30% of regions with significant growth effects realize significant effects on investment. The growth effects of the remaining 10 percentage points are likely consumption effects only which should not be expected to last in the medium to the long run.

Our findings are complementary to recent work on the heterogeneous responses to fiscal stimuli in macroeconomics in the sense that fiscal multipliers may differ dramatically across recipients. We estimate positive responses to stimuli (transfers) to be higher for recipients with higher levels of absorptive capacity measured as an above-average endowment of human capital and an above-average level of quality of government.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2005), Institutions as the fundamental cause of long-run growth, in Philippe Aghion and Stephen Durlauf (eds.), *Handbook of Economic Growth*, North Holland: Elsevier, 386-472.

Anderson, Michael L. (2008), Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and early training projects, *Journal of the American Statistical Association* 103, 1481-1495.

Angrist, Joshua D. and Ivan Fernandez-Val (2010), ExtrapoLATE-ing: External validity and overidentification in the LATE framework, *NBER Working Paper* 16566.

Angrist, Joshua D. and Jörn-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricists's Companion*, Princeton, NJ: Princeton University Press.

Auerbach, Alan J. and Yuriy Gorodnichenko (2010), Measuring the output responses to fiscal policy, *American Economic Journal: Economic Policy*, forthcoming.

Auerbach, Alan J. and Yuriy Gorodnichenko (2011), Fiscal multipliers in recession and expansion, *unpublished manuscript*, University of California at Berkeley.

Becker, Sascha O. and Peter H. Egger (2011), Endogenous product versus process innovation and a firm's propensity to export, *Empirical Economics*, forthcoming.

Becker, Sascha O., Peter H. Egger, and Maximilian von Ehrlich (2010), Going NUTS: The effect of EU Structural Funds on regional performance, *Journal of Public Economics* 94, 578-590.

Becker, Sascha O., Erik Hornung, and Ludger Woessmann (2011), Education and catch-up in the industrial revolution, *American Economic Journal: Macroeconomics* 3, 92-126.

Benhabib, Jess and Mark M. Spiegel (1994), The role of human capital in economic development - Evidence from aggregate cross-country data, *Journal of Monetary Economics* 34, 143-173.

Burnside, Craig, and David Dollar (2000), Aid, policies, and growth, *American Economic Review* 90, 847-868.

Burnside, Craig, and David Dollar (2004), Aid, policies, and growth: Revisiting the evidence, *World Bank Policy Research Working Paper* No. 3251.

Charron, Nicholas and Victor Lapuente (2011), Why do some regions in Europe have higher quality of government? *QoG Working Paper Series* 2011:1.

Chauvet, Lisa and Patrick Guillaumont (2004), Aid and growth revisited: Policy, economic vulnerability and political instability, in Bertil Tungodden, Nicholas Herbert Stern, Ivar Kolstad (eds), *Towards Pro-Poor Policies - Aid, Institutions and Globalization*, Washington D.C./Oxford: World Bank/Oxford University Press.

Chenery, Hollis B. and Alan M. Strout (1966), Foreign assistance and economic development, *American Economic Review* 56, 679-733.

Dalgaard, Carl-Johan, Henrik Hansen, and Finn Tarp (2004), On the empirics of foreign aid and growth, *Economic Journal* 114, 191-216.

Duffy, John, Chris Papageorgiou, and Fidel Perez-Sebastian (2004), Capital-skill complementarity? Evidence from a panel of countries, *Review of Economics and Statistics* 86, 327-344.

Easterly, William (2003), Can foreign aid buy growth?, *Journal of Economic Perspectives* 17, 23-48.

Egger, Peter H. and Georg Wamser (2011), The impact of controlled foreign company legislation on real investments abroad: A two-dimensional regression discontinuity design, *CEPR Discussion Paper* 8460.

European Commission (1997), *Regional Development Studies - The Impact of Structural Policies on Economic and Social Cohesion in the Union 1989-99*, Luxembourg: Office for Official Publications of the European Commission.

European Commission (2001), *Unity, Solidarity, Diversity for Europe, its People and its Territory - Second Report on Economic and Social Cohesion*, Luxembourg: Office for Official Publications of the European Commission.

European Commission (2007), *Annex to the 18th Annual Report on Implementation of the Structural Funds (2006)*, Brussels: Commission of the European Communities.

Fitzenberger, Bernd (1998), The moving blocks bootstrap and robust inference for linear least squares and quantile regressions, *Journal of Econometrics* 82, 235-287.

Gemmell, Norman, Richard Kneller, and Ismael Sanz (2011), The timing and persistence of fiscal policy impacts on growth: Evidence from OECD countries, *Economic Journal* 121, F33-F58.

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw (2001), Identification and estimation of treatment effects with a regression-discontinuity design, *Econometrica* 69, 201-209.

Härdle, Wolfgang, Marlene Müller, Stefan Sperlich, and Axel Werwatz (2004), *Nonparametric and Semiparametric Models*, Berlin: Springer.

Imbens, Guido and Joshua D. Angrist (1994), Identification and estimation of local average treatment effects, *Econometrica* 62, 467-475.

Imbens, Guido and Thomas Lemieux (2008), Regression discontinuity designs: A guide to practice, *Journal of Econometrics* 142, 615-635.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz (2007), Experimental analysis of neighborhood effects, *Econometrica* 75, 83-119.

Lee, David S. and Thomas Lemieux (2010), Regression discontinuity designs in economics, *Journal of Economic Literature* 48, 281-355.

Ludwig, Jens and Douglas L. Miller (2007), Does head start improve children's life chances? Evidence from a regression discontinuity design, *Quarterly Journal of Economics* 122, 159-208.

Mankiw, Gregory N., David Romer, and David N. Weil (1992), A contribution to the empirics of economic growth, *Quarterly Journal of Economics* 107, 407-437.

Mauro, Paolo (1995), Corruption and growth, *Quarterly Journal of Economics* 110, 681-712.

Pisani-Ferry, Jean, Benedicta Marzinotto, and Guntram B. Wolff (2011), How European funds can help Greece grow, *Financial Times* Jul 28, 2011.

Ramey, Valerie A. (2011), Identifying government spending shocks: It's all in the timing, *Quarterly Journal of Economics* 126, 1-50.

Rosenstein-Rodan and Paul Narcyz (1961), International aid for underdeveloped countries, *Review of Economic Statics* 43, 107-138.

Ruppert, David and Matthew P. Wand (1994), Multivariate locally weighted least squares regression, *Annals of Statistics* 22, 1346-1370.

Shoag, Daniel (2011), The impact of government spending shocks: Evidence on the multiplier from state pension plan returns, *mimeo*, Harvard University.

Suárez Serrato, Juan Carlos, and Philippe Wingender (2011), Estimating local fiscal multipliers, *mimeo*, University of California, Berkeley.

Wong, Vivian C., Peter M. Steiner, and Thomas D. Cook (2010), Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods, *Institute for Policy Research Northwestern University Working Paper Series* WP-10-02.

Appendix A. Deriving the HLATE

We aim at proving

$$HLATE(\mathbf{x_0}, \mathbf{z_i}) = \lim_{\mathbf{\Delta}\to\mathbf{0}} \frac{E[y_i|\mathbf{0} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{\Delta}, \mathbf{z_i}] - E[y_i| - \mathbf{\Delta} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{0}, \mathbf{z_i}]}{E[T_i = 1|\mathbf{0} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{\Delta}, \mathbf{z_i}] - E[T_i = 1| - \mathbf{\Delta} < \widetilde{\mathbf{x}}_\mathbf{i} < \mathbf{0}, \mathbf{z_i}]}.$$

The outcome difference of observations at the threshold is

$$\begin{aligned} E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] - E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}] = \\ E[T_i\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] - E[T_i\beta|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}] \\ +E[\mathbf{z_i}\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] - E[\mathbf{z_i}\beta|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}] \\ +E[\alpha_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] - E[\alpha_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}]. \end{aligned}$$

We assume that $E[\alpha_i|\mathbf{x_i} = \mathbf{x}]$ is continuous at $\mathbf{x_0}$ such that the last two pairs of terms in the above equation cancel each other as $\mathbf{\Delta}$ moves towards zero. Assuming conditional independence between T_i and β as well as between $\mathbf{z_i}$ and β yields

$$\begin{aligned} &E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] - E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}] = \\ &E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}]E[T_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] + E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}]E[\mathbf{z_i}|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z_i}] \\ &-E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}]E[T_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}] - E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}]E[\mathbf{z_i}|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z_i}]. \end{aligned}$$

Note that conditional independence requires that no selection into treatment on the basis of the expected effect occurs. Assuming that $E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{0}]$ is continuous at $\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{0}$ then delivers

$$\lim_{\mathbf{\Delta}\to 0} E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z}_\mathbf{i}] - \lim_{\mathbf{\Delta}\to 0} E[y_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z}_\mathbf{i}] =$$
$$E[\beta|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{0}, \mathbf{z}_\mathbf{i}] \left(\lim_{\mathbf{\Delta}\to 0} E[T_i|\widetilde{\mathbf{x}}_\mathbf{i} = \mathbf{\Delta}, \mathbf{z}_\mathbf{i}] - \lim_{\mathbf{\Delta}\to 0} E[T_i|\widetilde{\mathbf{x}}_\mathbf{i} = -\mathbf{\Delta}, \mathbf{z}_\mathbf{i}] \right)$$

which can easily be reformulated to obtain $HLATE(\mathbf{z}_\mathbf{i})$ from above.

Appendix B. Standard errors of the HLATE

Under the maintained assumptions in this paper and Assumptions (i)-(vii) in Hahn, Todd, and van der Klaauw (2001), the estimate $\widehat{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i})$ is distributed as

$$n^{2/5}[\widehat{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i}) - HLATE(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i})] \to \mathfrak{N}\,[\mu_{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i}), \Omega_{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i})] \tag{11}$$

where $\mu_{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i})$ approaches zero as $\mathbf{\Delta}$ in (9) approaches zero. $\Omega_{HLATE}(\mathbf{x}_\mathbf{0}, \mathbf{z}_\mathbf{i})$ in (11) is then defined as in Hahn, Todd, and van der Klaauw (2001) conditional on $\mathbf{z}_\mathbf{i}$.

Appendix C. Monte Carlo study

Appendix C.1. Simulation design

In the following we examine the performance of parametric and nonparametric estimators in identifying the HLATE. We consider sharp and fuzzy designs for the HLATE and scenarios where the treatment depends on one (1-way threshold) versus two forcing variables (2-way threshold). In the application in Section 4 only one forcing variable matters for treatment assignment, yet it is useful to consider a more general case other applications rely upon (see Egger and Wamser, 2011). For each case (1-way versus 2-way design), let us consider 3 experiments: a Sharp RDD, a Fuzzy 1 RDD with a low degree of fuzziness, and a Fuzzy 2 RDD with a high degree of fuzziness about treatment assignment (see below). We set the standard deviation of the disturbances ϵ_i to $\sigma_\epsilon = 0.3$ and $\sigma_\epsilon = 0.6$. In any case, ϵ_i is distributed as ϵ_i $i.i.d.N(0, \sigma_\epsilon)$.

We generate the data about x_i and z_i for observation $i = 1, \ldots, N$ based on a grid of $60 \cdot 60$ bins in x-z space. In each dimension, bins take addresses (i.e., values of x_i and z_i) between -2.95 and 2.95 and have a size of 0.1. We, assume that each of the $60^2 = 3,600$ bins hosts 6 observations with identical values of x_i and z_i but an independent draw of ϵ_i. Hence, there is a total number of $21,600$ observations available to the largest data-set possible. This aims at mimicking the empirical situation with RDDs where one allots data points into bins to generate averages of x_i (z_i) and y_i (see Angrist and Pischke, 2009; Lee and Lemieux, 2010). To illustrate the small-sample performance of the nonparametric estimator of the HLATE and compare it with its parametric counterpart, we alternatively consider subsets of that data-set where we consider sub-grids of $40 \cdot 40$ in the support region of $[-1.95, 1.95]$ in x-z space with $40^2 \cdot 6 = 9,600$ observations and $20 \cdot 20$ in the support region of $[-0.95, 0.95]$ in x-z space with $20^2 \cdot 6 = 2,400$ observations.

In each of the experiments, LATE corresponds to the average level of HLATE and is measured by the coefficient on the treatment dummy T_i, i.e., $\beta = 1$.

1-way threshold

With a 1-way threshold rule, the data generating processes can be described as follows.

Sharp RDD:

$$\begin{aligned} y_i =& 1 + T_i + .5T_i z_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_i z_i + \epsilon_i \\ & where\ T_i = 1(x_i \geq 0). \end{aligned}$$

Fuzzy 1 RDD:

$$\begin{aligned} y_i =& 1 + T_i + .5T_i z_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_i z_i + \epsilon_i \\ & where\ P(T_i = 1) = \begin{cases} 1 & \text{if } x_i > b \\ 11/12 & \text{if } 0 \leq x_i \leq b \\ 1/12 & \text{if } -b \leq x_i < 0 \\ 0 & \text{if } x_i < -b \end{cases} \end{aligned}$$

For the simulations, we chose $b = 0.45$ so that the probability of treatment mis-assignment is $1/12$ in the support region of $[-0.45, 0.45]$ in x-space (i.e., in 5 bins to the left and in 5 bins to the right of the 1-way threshold). The maximum of observations in the mis-classification region are $10 \cdot 60 \cdot 6 = 3,600$, $10 \cdot 40 \cdot 6 = 2,400$, and $10 \cdot 20 \cdot 6 = 1,200$, depending on the chosen grid and sample size. Hence, 300, 200, and 100 observations, respectively, are expected to be misclassified. Note that the random process underlying the fuzzyness are drawn for each replication of the Monte Carlo study separately.

Fuzzy 2 RDD:

$$y_i = 1 + T_i + .5T_iz_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_iz_i + \epsilon_i$$

$$where\ P(T_i = 1) = \begin{cases} 1 & \text{if } x_i > b \\ 5/6 & \text{if } 0 \leq x_i \leq b \\ 1/6 & \text{if } -b \leq x_i < 0 \\ 0 & \text{if } x_i < -b \end{cases}$$

As in the Fuzzy 1 design, we chose $b = 0.45$ but we assumed the probability of treatment mis-assignment amounting to 1/6 in the support region of $[-0.45, 0.45]$ in x-space. Hence, depending on the chosen grid and sample size, 600, 400, and 200 observations, respectively, are expected to be misclassified in the Fuzzy 2 design.

The results for the Sharp RDD are illustrated in Figure 10 and the ones for the Fuzzy 1 and Fuzzy 2 RDDs are illustrated in Figure 11. In the 1-way experiments, the treatment is only determined by forcing the variable x_i whereas the outcome is affected by x_i and z_i. The heterogeneous treatment effect appears in the outcome graphs as a wedge between the red (treated) and the blue (untreated) observations. The extent of heterogeneity of LATE is noticeable as the outcome shift between treated and non-treated observations disappears for low values of z_i. In the fuzzy experiments illustrated in Figure 11, the treatment probability (approximated by the fraction of treated observations) jumps at the threshold x_0 by about 0.85 and 0.65 in the Fuzzy 1 and Fuzzy 2 designs, respectively, which reflects the corresponding mis-classification probabilities of 1/12 and 1/6. With a fuzzy design, some of the red observations characterized by $x_i > x_0$ do *not* receive treatment while some of of the blue observations with $x_i < x_0$ *do* receive treatment. This fuzziness blurs the discontinuity in the outcome function and results in a smaller treatment wedge compared to the sharp design. According to equation (9), the treatment effect is measured by the ratio of the outcome wedge and the jump in the treatment probability.

– Figures 10 and 11 –

2-way threshold

With a 2-way threshold, both x_i and z_i serve as forcing variables and LATE also varies with z_i. With respect to x_i, we maintain the threshold value $x_0 = 0$ while now also z_i has to exceed a level of $z_0 = -0.6$ in order to qualify for treatment. For (sharp) treatment assignment we require both rules to be fulfilled at the same time.[28] Distinguishing again between sharp and fuzzy scenarios we consider the

[28]Recent work by Wong, Steiner and Cook (2010) considers multiple threshold rules but requires only one rule to be satisfied for treatment.

following experiments in the 2-way threshold design:

Sharp RDD:

$$\begin{aligned} y_i =& 1 + T_i + .5T_i z_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_i z_i + \epsilon_i \\ & where\ T_i = 1(x_i \geq 0 \wedge z_i \geq -0.6). \end{aligned}$$

Fuzzy 1 RDD:

$$y_i =1 + T_i + .5T_i z_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_i z_i + \epsilon_i$$

$$where\ P(T_i = 1) = \begin{cases} 1 & \text{if } x_i > b \wedge z_i > -.6 + b \\ 11/12 & \text{if } 0 \leq x_i \leq b \wedge -0.6 - b \leq z_i \leq -.6 \\ 1/12 & \text{if } b \leq x_i < 0 \wedge -.6 \leq z_i \leq -.6 + b \\ 0 & \text{if } x_i < b \wedge z_i < -.6 + b \end{cases}$$

As with a 1-way treatment threshold, we chose $b = 0.45$ and the probability of treatment mis-assignment is 1/12 in the chosen support region. However, now treatment mis-classification may vary with both x_i and z_i. Therefore, we chose the support region to be bounded by $[-0.45, 0.45]$ in x-space and by $[-1.05, -0.15]$ in z-space. The maximum of observations in the mis-classification region are $10 \cdot 10 \cdot 6 = 600$, independent of the chosen grid and sample size. Hence, 50 observations are expected to be misclassified in any one of the fuzzy design experiments.

Fuzzy 2 RDD:

$$y_i =1 + T_i + .5T_i z_i + .5x_i + .5z_i + .1x_i^2 + .1z_i^2 + .3x_i z_i + \epsilon_i$$

$$where\ P(T_i = 1) = \begin{cases} 1 & \text{if } x_i > b \wedge z_i > -.6 + b \\ 7/8 & \text{if } 0 \leq x_i \leq b \wedge -0.6 - b \leq z_i \leq -.6 \\ 1/8 & \text{if } b \leq x_i < 0 \wedge -.6 \leq z_i \leq -.6 + b \\ 0 & \text{if } x_i < b \wedge z_i < -.6 + b \end{cases}$$

As in the 2-way threshold Fuzzy 1 design, we chose $b = 0.45$ but we assumed the probability of treatment mis-assignment amounting to 1/6 in the support region of $[-0.45, 0.45]$ in x-space and $[-1.05, -0.15]$ in z-space. Hence, 100 observations are generally expected to be misclassified in the 2-way Fuzzy 2 design. The 2-way Sharp RDD is illustrated in Figure 12 and the corresponding Fuzzy 1 and Fuzzy 2 RDDs are illustrates in Figure 13.

– Figures 12 and 13 –

Notice that, apart from the different design in general, the 2-way (H)LATE estimates are based on a smaller number of cells and observations at the the treatment thresholds. The latter should not have any bearing for the bias but it comes at a loss of precision of the estimates in comparison to the 1-way threshold results.

Appendix C.2. Results

The simulation results for the local average treatment effect (LATE; at $x, z = 0$) are presented in Table 6 for a 1-way threshold design (and in Table 7 for a 2-way threshold design). Those for the heterogeneous local average treatment effect (HLATE; at $x = 0$ across all z) are presented graphically in Figure A.5. Remember that LATE in the sense of the average HLATE corresponds to the coefficient on the treatment dummy T_i, i.e., $\beta = 1$. The bias is measured as a deviation of the estimate $\hat{\beta}$ from the true parameter $\beta = 1$ in percent.

Note that, for the parametric estimates, we use the true functional form, i.e., that of the data-generating process. (Our interest is not in simulating the effect of mis-specification of the control function, but in illustrating the small sample performance of nonparametric relative to parametric estimates of the HLATE.)

The findings can be summarized as follows. First, the estimates of both the nonparametric and the parametric estimates of LATE (β) appear to have a small bias across all experiments considered in the Monte Carlo analysis. In every one of the experiments is the bias of LATE smaller than one percent in absolute value independent of sample size of whether we consider a sharp or a fuzzy RDD (see the panels at the top of Tables 6 and 7). All else equal, the mean squared error tends to be smallest with a sharp design, a smaller value of σ_ϵ, a larger bandwidth considered, parametric rather than nonparametric estimates, and a 1-way instead of a 2-way threshold design. None of that is surprising, since fuzzy designs add noise to the estimation problem by involving a projection of the endogenous treatment status in a first stage; a larger value of σ_ϵ involves more noise at the level of the outcome equation; a smaller bandwidth considered is associated with a smaller number of observations we estimate the HLATE from, thus reducing precision; more flexible nonparametric estimates involve a loss of precision, if the true functional form of the relationship between the forcing variable (x) and also of the variable which interacts with treatment status (z) is a parametric polynomial; and the 2-way threshold design requires more parameters to be estimates – in our case, from a smaller number of observations at which the threshold is observed – which leads to efficiency losses.

– Tables 6 and 7

These insights about LATE also carry over to the estimation of HLATE in Figure 6. Quite obviously, the point estimates are virtually indistinguishable from the true values, but the estimated confidence intervals are smaller for the parametric estimates (which assumes the true functional form) than for the more flexible, local-linear-regression-based nonparametric estimates. Finally, the estimates for the 2-way threshold regressions in Figure 7 have somewhat larger confidence intervals than their counterparts for the 1-way thresholds.

Hence, we may conclude that both the nonparametric and the parametric estimates work well in small to moderately large samples. In empirical circumstances where parametric approximations of unknown functional forms will not work as well as in the Monte Carlo study, where the parametric estimates assumed the correct form of the control function, we expect nonparametric estimates to work quite well. In any case, HLATE can be inferred with very small bias from both nonparametric and parametric control function, irrespective of whether a sharp or a fuzzy design is being considered.

– Figures 14 and 15 –

Appendix D. Frequency of Observations

Figure 16 summarizes the frequency of observations in HC-forcing-variable-space and QoG-forcing-variable-space.

– Figure 16 –

Appendix E. Construction of Time-variant Measures of Absorptive Capacity

While the time-invariant measures of HC and QoG can directly be based on information about the NUTS2 level from a single source, namely European Union Labour Force Survey and Charron, Lapuente, and Dykstra (2011), respectively, time-variant measures thereof have to rely on imputation from several levels of aggregation and even several sources. We constructed time-variant measures on HC and QoG by as follows.

Towards a time-variant measure of HC at the NUTS2 level: Most of the required annual data points on time-variant HC for all NUTS2 regions and the same years in which the forcing variable is measured (i.e., prior to the respective programming periods) are available from the European Values Study (EVS) which is conducted by the European Values Systems Study Group, EVSSG. The EVS classifies the surveyed population into three education categories: low, medium, and high. We define $s_{2ht} \in [0, 1]$ as the share of medium plus high skilled in all surveyed individuals at the NUTS2 level and transform it logistically as $\sigma_{2ht} = \ln \frac{s_{2ht}}{1-s_{2ht}} \in (-\infty, \infty)$. We interpolate σ_{2ht} linearly, but are still left with missing data. In a next step, we use the same measure at the next higher level, namely NUTS1, σ_{1ht}

and use it to impute missing cross-sectional data of σ_{2ht} by σ_{1ht}. Finally, for the few remaining missing observations, we impute σ_{2ht} by σ_{0ht}, where the latter is measured at the country level. Then, we use the imputed counterpart to σ_{2ht}, referred to as $\hat{\sigma}_{2ht}$, and retransform it so as to obtain an imputed measure of HC as $\hat{s}_{2ht} = \frac{\hat{\sigma}_{2ht}}{1+\hat{\sigma}_{2ht}}$. The latter is used for some of the results in the main text.

Towards a time-variant measure of QoG at the NUTS2 level: The imputation of the QoG index involves two types of data and two steps. First of all, we utilize Eurobarometer Survey data on two variables, namely citizen's satisfaction with the state of democracy in their region, and their satisfaction with the rule of law in their region. Both variables take on scores between 0 and 100 and they are mapped into the unit space before being transformed akin to the HC share as above. Then, those two measures are interpolated each and retransformed. Let us call those imputed measures at the NUTS2 level that are now bounded between 0 and 1 $\hat{r}_{2ht}$ for satisfaction with rule of law and $\hat{d}_{2ht}$ for satisfaction with the state of democracy.

In a second step, we regress the cross-sectional QoG index on these two measures in linear models in order to impute the missing data on QoG prior to the first programming period so as to obtain (average) values for the same years where the forcing variable is measured. This results in the time-variant, imputed measure of QoG as used for some of the results in the main text.

Tables and Figures

Table 1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Obs.
	(1)	(2)	(3)	(4)	(5)
GDP per capita growth rate	0.043	0.017	-0.008	0.131	646
ln(Investments/Capita)	8.222	0.430	6.483	9.406	646
Objective 1	0.317	0.466	0	1	646
Avg. GDP per capita in threshold years	12,792.870	4,714.408	4,448.148	37,835.190	646
Avg. GDP per capita in threshold years/EU avg.	0.913	0.297	0.300	2.328	646
Human capital (HC) - time-invariant	0	0.144	-0.376	0.318	646
Human capital (HC) - time-variant	0	0.184	-0.367	0.412	646
Quality of government (QoG) - time-invariant	0	0.815	-2.756	1.187	640
Quality of government (QoG) - time-variant	0	3.185	-11.387	5.109	640

Notes: Units of observation are EU NUTS2 regions. GDP, investment, and population data are from Cambridge Econometrics; information about Objective 1 treatment is available directly from the European Commission, from various Council Regulations, in particular the Regulations numbered 2052/88, 2082/93, and 502/1999, and in editions of the Official Journal (see also Becker, Egger, and von Ehrlich, 2010); the human capital (HC) variable measures the share of the workforce with at least upper-secondary education (ISCED categories 3 to 6); the quality of government (QoG) index comes from Charron and Lapuente (2011). HC and QoG variables are normalized to zero, by detracting the EU average; see the main text for more detail. For HC and QoG we construct the corresponding time-variant variables using information from the European Value Survey, Eurobarometer, and the Barro-Lee dataset on educational attainments.
We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS2 regions we calculated GDP per capita growth for the years 1989 and 1990 using information from the GDR's statistic yearbook; the East-German investment per capita in the first programming period is measured as the average over the years 1991,1992, and 1993. The EU QoG index is not available for the Spanish region Ceuta and Melilla.

Table 2: GDP/Capita Growth Rate, Objective 1 Treatment and Absorptive Capacity

	3rd-order polynomial		4th-order polynomial		5th-order polynomial	
	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
			Human Capital			
Time-invariant						
Object1	.007	.007	.004	.004	.010	.012
	(.004)*	(.004)*	(.004)	(.004)	(.004)***	(.004)***
Object1×HC	.044	.046	.046	.044	.048	.045
	(.010)***	(.009)***	(.010)***	(.009)***	(.010)***	(.009)***
HC	.008	-	.006	-	.007	-
	(.008)		(.008)		(.008)	
Const.	.041	.011	.041	.041	.038	.037
	(.002)***	(.011)	(.002)***	(.002)***	(.003)***	(.003)***
Obs.	646	646	646	646	646	646
R^2	.293	.319	.295	.302	.291	.311
Time-variant						
Object1	.009	.009	.005	.006	.010	.011
	(.004)**	(.004)**	(.004)	(.004)♯	(.004)***	(.004)***
Object1×HC	.042	.038	.041	.041	.042	.041
	(.008)***	(.008)***	(.008)***	(.008)***	(.008)***	(.007)***
HC	.003	.003	.003	.004	.004	.005
	(.004)	(.004)	(.004)	(.004)	(.004)	(.004)
Const.	.040	.015	.040	.040	.037	.037
	(.002)***	(.013)	(.002)***	(.002)***	(.003)***	(.003)***
Obs.	646	646	646	646	646	646
R^2	.259	.291	.266	.274	.261	.278
			Quality of Government			
Time-invariant						
Object1	.011	.013	.008	.007	.010	.010
	(.004)***	(.004)***	(.003)**	(.003)**	(.003)***	(.003)***
Object1×QoG	.005	.005	.006	.006	.006	.006
	(.002)**	(.002)***	(.002)***	(.002)***	(.002)***	(.002)***
QoG	.002	-	.001	-	.001	-
	(.002)		(.002)		(.002)	
Const.	.039	.001	.039	.040	.037	.037
	(.002)***	(.012)	(.002)***	(.002)***	(.003)***	(.003)***
Obs.	640	640	640	640	640	640
R^2	.238	.286	.247	.255	.244	.257
Time-variant						
Object1	.010	.012	.006	.006	.009	.009
	(.004)**	(.004)***	(.003)*	(.004)*	(.003)***	(.003)***
Object1×QoG	.001	.001	.001	.001	.001	.001
	(.0005)**	(.0005)***	(.0005)***	(.0005)**	(.0005)***	(.0005)***
QoG	.0005	.0005	.0004	.0004	.0004	.0004
	(.0004)	(.0003)	(.0004)	(.0004)	(.0004)	(.0004)
Const.	.039	.001	.039	.039	.037	.037
	(.002)***	(.012)	(.002)***	(.002)***	(.003)***	(.003)***
Obs.	640	640	640	640	640	640
R^2	.233	.28	.241	.243	.238	.25

Notes: ***, **, *, ♯ denote significance at the 1, 5, 10, and 15% level, respectively. Standard errors are clustered at the NUTS2 level. First-stage regressions are probit models. The polynomial functions are allowed to have different parameters to the left and the right of the threshold. The quality of government variable (QoG) refers to the EU QoG index by Charron and Lapuente (2011); for the construction of the time-variant version see Appendix XX. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only.

Table 3: LN(Investments/Capita), Objective 1 Treatment and Absorptive Capacity

	3rd-order polynomial		4th-order polynomial		5th-order polynomial	
	Pooled OLS	FE	Pooled OLS	FE	Pooled OLS	FE
			Human Capital			
Time-invariant						
Object1	-.075	-.044	.072	.076	.111	.150
	(.076)	(.077)	(.072)	(.070)	(.061)*	(.060)**
Object1×HC	.793	.728	.792	.748	.870	.806
	(.229)***	(.222)***	(.229)***	(.218)***	(.236)***	(.221)***
HC	.315	-	.346	-	.342	-
	(.132)**		(.138)**		(.140)**	
Const.	8.135	8.324	8.120	8.109	8.134	8.106
	(.043)***	(.230)***	(.055)***	(.056)***	(.061)***	(.062)***
Obs.	646	646	646	646	646	646
R^2	.563	.575	.576	.582	.578	.589
Time-variant						
Object1	-.075	-.050	.072	.082	.107	.141
	(.082)	(.083)	(.074)	(.074)	(.062)*	(.061)**
Object1×HC	.638	.694	.624	.624	.677	.662
	(.230)***	(.217)***	(.228)***	(.230)***	(.231)***	(.230)***
HC	.274	.309	.292	.312	.291	.318
	(.081)***	(.081)***	(.081)***	(.075)***	(.081)***	(.074)***
Const.	8.134	8.419	8.111	8.093	8.112	8.080
	(.042)***	(.234)***	(.052)***	(.054)***	(.057)***	(.059)***
Obs.	646	646	646	646	646	646
R^2	.546	.559	.557	.563	.557	.567
			Quality of Government			
Time-invariant						
Object1	.166	.215	.260	.264	.226	.244
	(.086)*	(.091)**	(.087)***	(.083)***	(.072)***	(.072)***
Object1×QoG	.153	.137	.131	.122	.127	.118
	(.051)***	(.048)***	(.051)**	(.048)**	(.051)**	(.048)**
QoG	.097	-	.111	-	.110	-
	(.024)***		(.024)***		(.024)***	
Const.	8.050	7.966	8.033	8.031	8.051	8.040
	(.040)***	(.187)***	(.048)***	(.047)***	(.051)***	(.054)***
Obs.	640	640	640	640	640	640
R^2	.581	.592	.58	.595	.584	.595
Time-variant						
Object1	.141	.190	.234	.234	.207	.225
	(.085)*	(.089)**	(.086)***	(.084)***	(.071)***	(.070)***
Object1×QoG	.037	.034	.031	.029	.030	.028
	(.013)***	(.012)***	(.013)**	(.012)**	(.013)**	(.012)**
QoG	.026	.030	.029	.031	.029	.031
	(.006)***	(.005)***	(.006)***	(.005)***	(.006)***	(.005)***
Const.	8.056	7.954	8.041	8.039	8.058	8.047
	(.040)***	(.188)***	(.047)***	(.047)***	(.051)***	(.053)***
Obs.	640	640	640	640	640	640
R^2	.58	.591	.581	.593	.584	.594

Notes: ***, **, *, ♯ denote significance at the 1, 5, 10, and 15% level, respectively. Standard errors are clustered at the NUTS2 level. First-stage regressions are probit models. The polynomial functions are allowed to have different parameters to the left and the right of the threshold. For the construction of the time-variant version of the human capital variable (HC) see Appendix XX. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only.

Table 4: Objective 1 Treatment, Human Capital, and Quality of Government

	3rd-order polynomial		4th-order polynomial	
	Pooled OLS	FE	Pooled OLS	FE
	Dependent Var.: GDP/Capita Growth Rate			
Object1	.009	.009	.010	.012
	(.003)***	(.003)***	(.003)***	(.003)***
Object1×QoG	-.003	-.002	-.001	-.001
	(.004)	(.003)	(.003)	(.003)
Object1×HC	.026	.029	.027	.028
	(.016)♯	(.014)**	(.016)*	(.014)**
Object1×QoG×HC	.055	.067	.050	.066
	(.038)♯	(.035)*	(.037)	(.035)*
Object1×QoG×HC×I	-.116	-.098	-.099	-.094
	(.062)*	(.053)*	(.060)*	(.052)*
QoG	.005	.005	.005	.004
	(.002)***	(.002)***	(.002)***	(.002)***
HC	.024	.015	.023	.014
	(.015)♯	(.012)	(.015)♯	(.012)
QoG×HC	-.084	-.077	-.084	-.076
	(.034)**	(.026)***	(.033)**	(.025)***
QoG×HC×I	.136	.106	.129	.103
	(.055)**	(.043)**	(.054)**	(.042)**
Const.	.040	.017	.039	.012
	(.002)***	(.011)	(.002)***	(.013)
Obs.	640	640	640	640
R^2	.322	.378	.321	.386
F-Stat.: joint significance HC terms	10.43	9.33	10.64	9.12
F-Stat.: joint significance QoG terms	5.13	4.01	5.61	4.79
	Dependent Var.: Ln(Investments/Capita)			
Object1	.006	.126	.046	.127
	(.073)	(.056)**	(.077)	(.057)**
Object1×QoG	.279	.218	.277	.222
	(.088)***	(.072)***	(.090)***	(.071)***
Object1×HC	.660	.755	.725	.760
	(.232)***	(.231)***	(.246)***	(.236)***
Object1×QoG×HC	1.086	.565	.934	.579
	(.536)**	(.517)	(.579)♯	(.540)
Object1×QoG×HC×I	.736	.560	1.024	.646
	(1.001)	(.912)	(1.041)	(.918)
QoG	.074	.094	.072	.092
	(.039)*	(.038)**	(.039)*	(.037)**
HC	-.015	.257	-.037	.233
	(.160)	(.167)♯	(.153)	(.164)
QoG×HC	.213	-.035	.232	-.022
	(.290)	(.252)	(.288)	(.257)
QoG×HC×I	-.445	.261	-.604	.146
	(.630)	(.674)	(.622)	(.658)
Const.	14.997	15.418	15.009	15.389
	(.041)***	(.212)***	(.051)***	(.237)***
Obs.	640	640	640	640
R^2	.646	.699	.647	.7
F-Stat.: joint significance HC terms	24.86	23.21	23.28	22.44
F-Stat.: joint significance QoG terms	16.74	7.74	13.20	7.22

Notes: ***, **, *, ♯ denote significance at the 1, 5, 10, and 15% level, respectively. Indicator I is unity for observations where QoG as well as HC are negative and zero otherwise. Standard errors are clustered at the NUTS2 level. First-stage regressions are probit models. The polynomial functions are allowed to have different parameters to the left and the right of the threshold. The human capital variable (HC) as well as the quality of government variable (QoG) are time-invariant. The sample consists of the EU12 NUTS2 regions for the first period, the EU15 NUTS2 regions for the second period, and the EU25 NUTS2 regions for the third programming period. We miss information on the four French overseas-départements and the two autonomous Portuguese regions Madeira and Azores for all three periods. For the Dutch region Flevoland we miss information for the first period only. Regarding the East-German NUTS2 regions, we calculated GDP per capita growth for the years 1989 and 1990 using information from the GDR's statistic yearbook.

Table 5: Percentage of Objective 1 regions with significant positive HLATE per country

Country	Confidence Interval 90%	80%	70%	Point Estimate
	GDP/Capita Growth			
Austria	100	100	100	100
Belgium	100	100	100	100
Czech Republic	0	43	57	100
Estonia	100	100	100	100
Finland	100	100	100	100
France	0	0	0	0
Germany	100	100	100	100
Greece	8	8	8	23
Hungary	86	86	100	100
Ireland	40	100	100	100
Italy	0	0	0	0
Latvia	0	0	100	100
Lithuania	0	0	0	100
Malta	0	0	0	0
Netherlands	100	100	100	100
Poland	6	25	44	100
Portugal	0	0	0	0
Slovak Republic	0	0	0	100
Slovenia	0	50	100	100
Spain	0	0	0	68
Sweden	100	100	100	100
United Kingdom	50	75	100	100
Average	30	36	40	61
	Ln(Investments/Capita)			
Austria	100	100	100	100
Belgium	0	0	0	0
Czech Republic	0	0	0	0
Estonia	0	0	0	0
Finland	100	100	100	100
France	0	0	0	0
Germany	100	100	100	100
Greece	0	0	0	0
Hungary	0	0	0	0
Ireland	0	0	40	100
Italy	0	0	0	55
Latvia	0	0	0	0
Lithuania	0	0	0	0
Malta	0	0	0	0
Netherlands	100	100	100	100
Poland	0	0	0	0
Portugal	0	0	0	21
Slovak Republic	0	0	0	0
Slovenia	0	0	0	0
Spain	0	0	0	0
Sweden	100	100	100	100
United Kingdom	13	13	25	88
Average	21	21	23	34

Notes: The percentages in the table are based on the same estimates as Figure 6. Note that Cyprus, Denmark, and Luxembourg did not receive Objective 1 funds during the period under consideration.

Table 6: Local Average Treatment Effect (1-Way Threshold)

Panel A: Bias of Average Treatment Effect

	Sharp RDD		Fuzzy 1 RDD		Fuzzy 2 RDD	
	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$
Parametric						
$N = 60^2 \cdot 6$	0.012	-0.061	0.013	0.015	-0.068	-0.076
$N = 40^2 \cdot 6$	0.004	-0.039	0.005	0.006	-0.046	-0.056
$N = 20^2 \cdot 6$	0.017	0.116	0.022	0.031	0.147	0.203
Nonparametric						
Bandwidth 2/3	-0.017	-0.044	-0.020	-0.026	-0.046	-0.047
Bandwidth 1/3	-0.007	-0.124	-0.015	-0.037	-0.158	-0.244
Bandwidth 1/6	0.079	-0.050	0.101	0.101	-0.047	-0.039

Panel B: RMSE of Average Treatment Effect

	Sharp RDD		Fuzzy 1 RDD		Fuzzy 2 RDD	
	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$
Parametric						
$N = 60^2 \cdot 6$	0.006	0.027	0.008	0.010	0.033	0.042
$N = 40^2 \cdot 6$	0.014	0.062	0.019	0.027	0.082	0.116
$N = 20^2 \cdot 6$	0.061	0.237	0.098	0.183	0.379	0.712
Nonparametric						
Bandwidth 2/3	0.010	0.040	0.013	0.019	0.054	0.079
Bandwidth 1/3	0.027	0.109	0.043	0.089	0.176	0.368
Bandwidth 1/6	0.126	0.488	0.193	0.472	0.741	1.788

Notes: All estimates result from Monte Carlo simulations with 2000 replications. The random error terms in the outcome equation as well as the random process underlying the fuzziness are drawn for each replication separately. The nonparametric estimates result from local linear regressions with uniform kernel. The variance of the error term in the outcome equation is denoted by σ_ϵ. Fuzzy 1 (2) refers to a data generating process with a misassignment probability of 1/12 (1/6) within 5 bins at both sides of the threshold. The largest sample refers to a grid range $[-2.95, 2.95]$ with 0.1 intervals. Accordingly, x and z feature 60 different values each. We observe each $x - z$ combination 6 times. The bias as well as the RMSE of the average treatment effect are measured in percent.

Table 7: Local Average Treatment Effects (2-Way Threshold)

Panel A: Bias of Average Treatment Effect

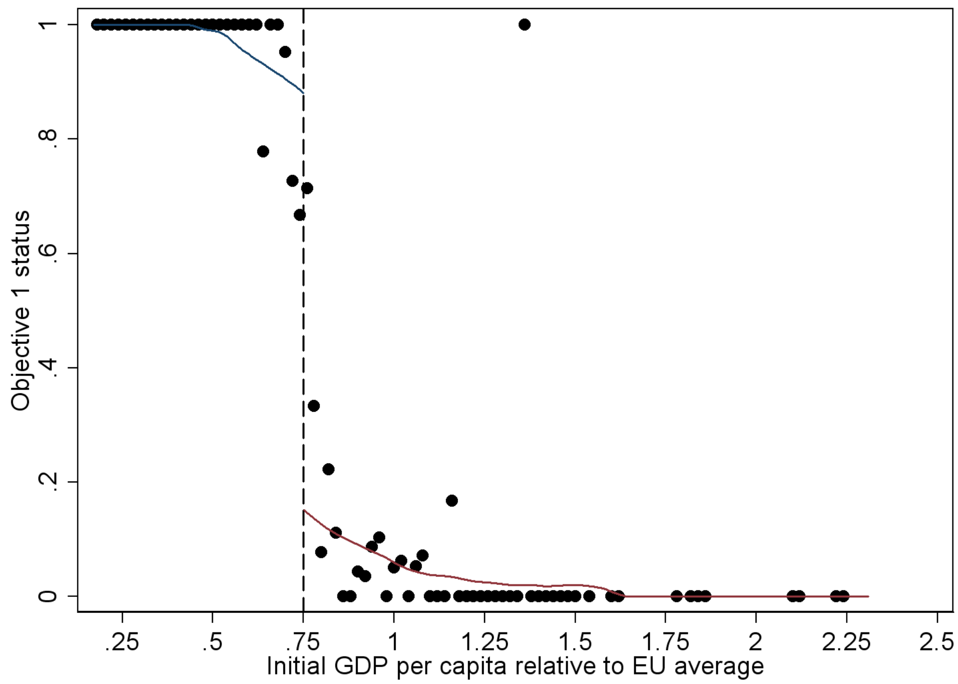
	Sharp RDD		Fuzzy 1 RDD		Fuzzy 2 RDD	
	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$
Parametric						
$N = 60^2 \cdot 6$	0.012	-0.097	0.012	0.013	-0.098	-0.101
$N = 40^2 \cdot 6$	0.047	0.051	0.049	0.051	0.053	0.056
$N = 20^2 \cdot 6$	-0.001	0.175	-0.003	-0.005	0.181	0.188
Nonparametric						
Bandwidth 2/3	-0.004	-0.079	-0.002	0.005	-0.088	-0.099
Bandwidth 1/3	0.026	-0.088	0.032	0.041	-0.087	-0.091
Bandwidth 1/6	0.077	-0.035	0.072	0.053	-0.019	0.049

Panel B: RMSE of Average Treatment Effect

	Sharp RDD		Fuzzy 1 RDD		Fuzzy 2 RDD	
	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$	$\sigma_\epsilon = 0.3$	$\sigma_\epsilon = 0.6$
Parametric						
$N = 60^2 \cdot 6$	0.007	0.027	0.007	0.008	0.028	0.029
$N = 40^2 \cdot 6$	0.014	0.058	0.015	0.016	0.062	0.066
$N = 20^2 \cdot 6$	0.054	0.197	0.059	0.067	0.218	0.247
Nonparametric						
Bandwidth 2/3	0.016	0.066	0.017	0.018	0.069	0.075
Bandwidth 1/3	0.042	0.165	0.047	0.061	0.185	0.243
Bandwidth 1/6	0.159	0.599	0.185	0.290	0.691	1.081

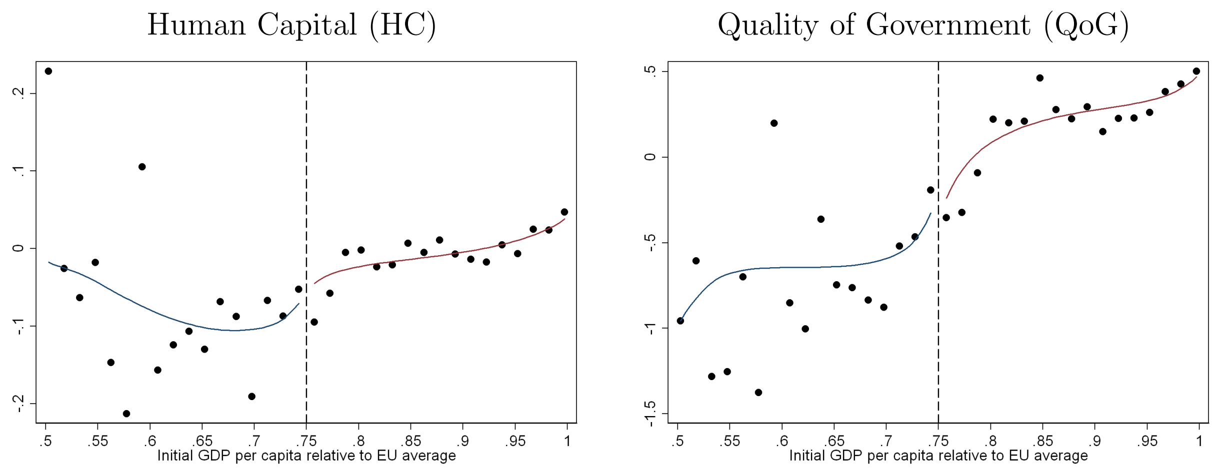
Notes: All estimates result from Monte Carlo simulations with 2000 replications. The random error terms in the outcome equation as well as the random process underlying the fuzziness are drawn for each replication separately. The nonparametric estimates result from local linear regressions with uniform kernel. The variance of the error term in the outcome equation is denoted by σ_ϵ. Fuzzy 1 (2) refers to a data generating process with a misassignment probability of 1/12 (1/6) within 5 bins at both sides of the two dimensional threshold. The largest sample refers to a grid range $[-2.95, 2.95]$ with 0.1 intervals. Accordingly, x and z feature 60 different values each. We observe each $x - z$ combination 6 times. The bias as well as the RMSE of the average treatment effect are measured in percent.

Figure 1: Objective 1 status and the 75% GDP threshold



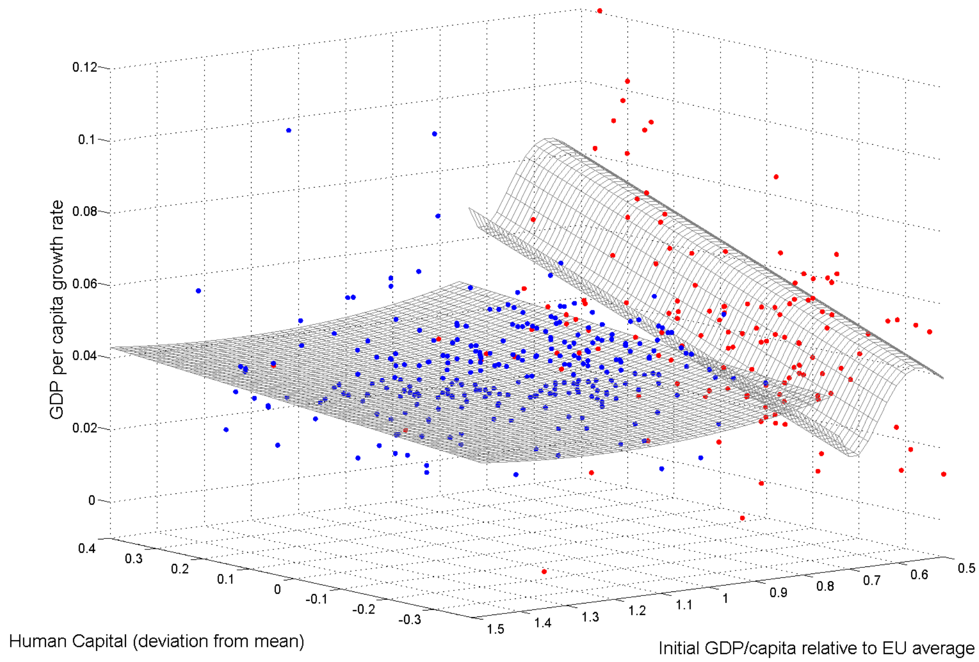
Note: The figure shows average treatment rates in equally-sized bins of 1.5 percentage points which are plotted against the per-capita GDP level that applied in the years relevant for the decision about Objective 1 status. The graph represents a local polynomial smooth based on an Epanechnikov kernel with a rule-of-thumb bandwidth. Note that the outlier at about 1.3 times the EU average which received treatment represents only one observation, namely Berlin in the 1989-1993 programming period. All results are robust to the exclusion of Berlin.

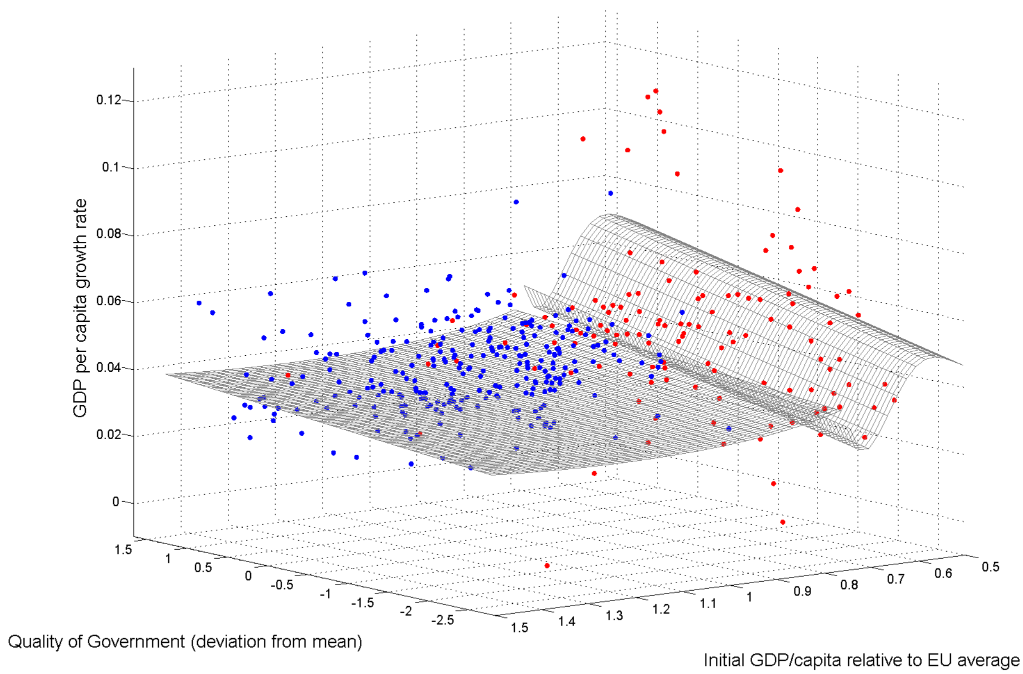
Figure 2: Human Capital, Quality of Government and the 75% GDP threshold



Note: The figures show averages of HC and QoG in equally-sized bins of 1.5 percentage points which are plotted against the per-capita GDP level that applied in the years relevant for the decision about Objective 1 status. The graphs represent a 2nd-order local polynomial function.

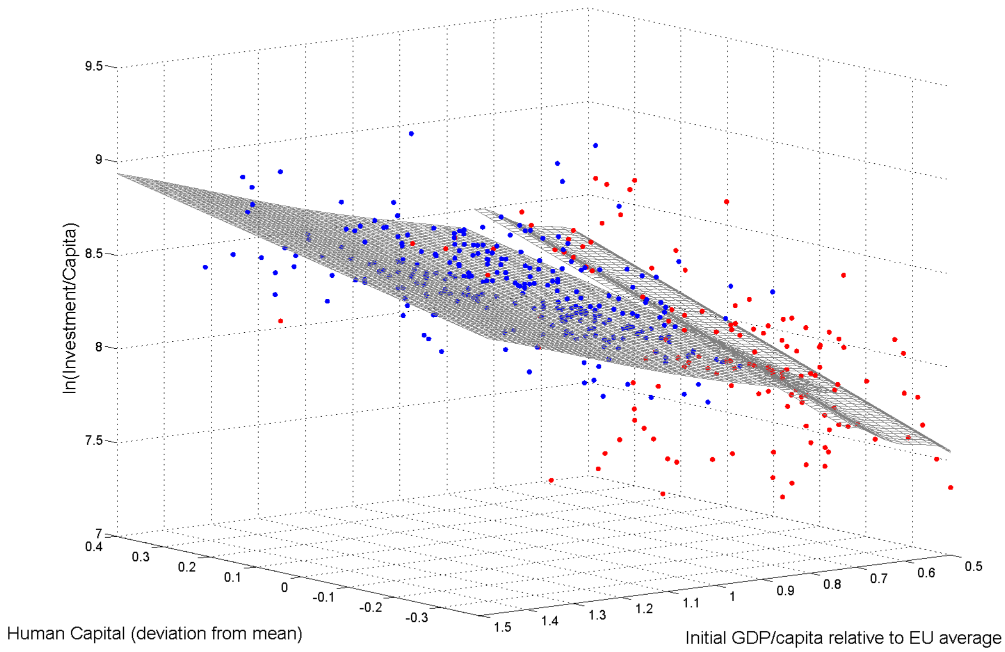
Figure 3: GDP per capita growth rate, Objective 1 Treatment, and absorptive capacity

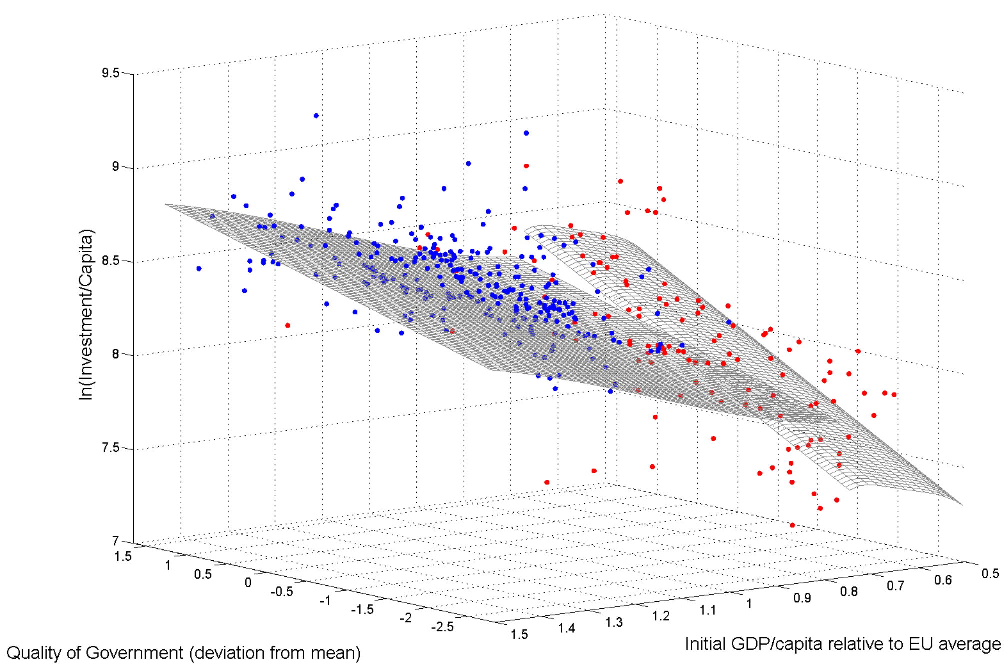




Note: The upper and lower figures illustrate the relationship between the outcome, forcing variable, human capital, and quality of government, respectively. The red (blue) dots indicate observations which received (did not receive) Objective 1 treatment. The surfaces represent 5th-order polynomial functions of per-capita GDP and linear functions of human capital and quality of government, respectively. These functions are estimated on both sides of the 75% threshold separately.

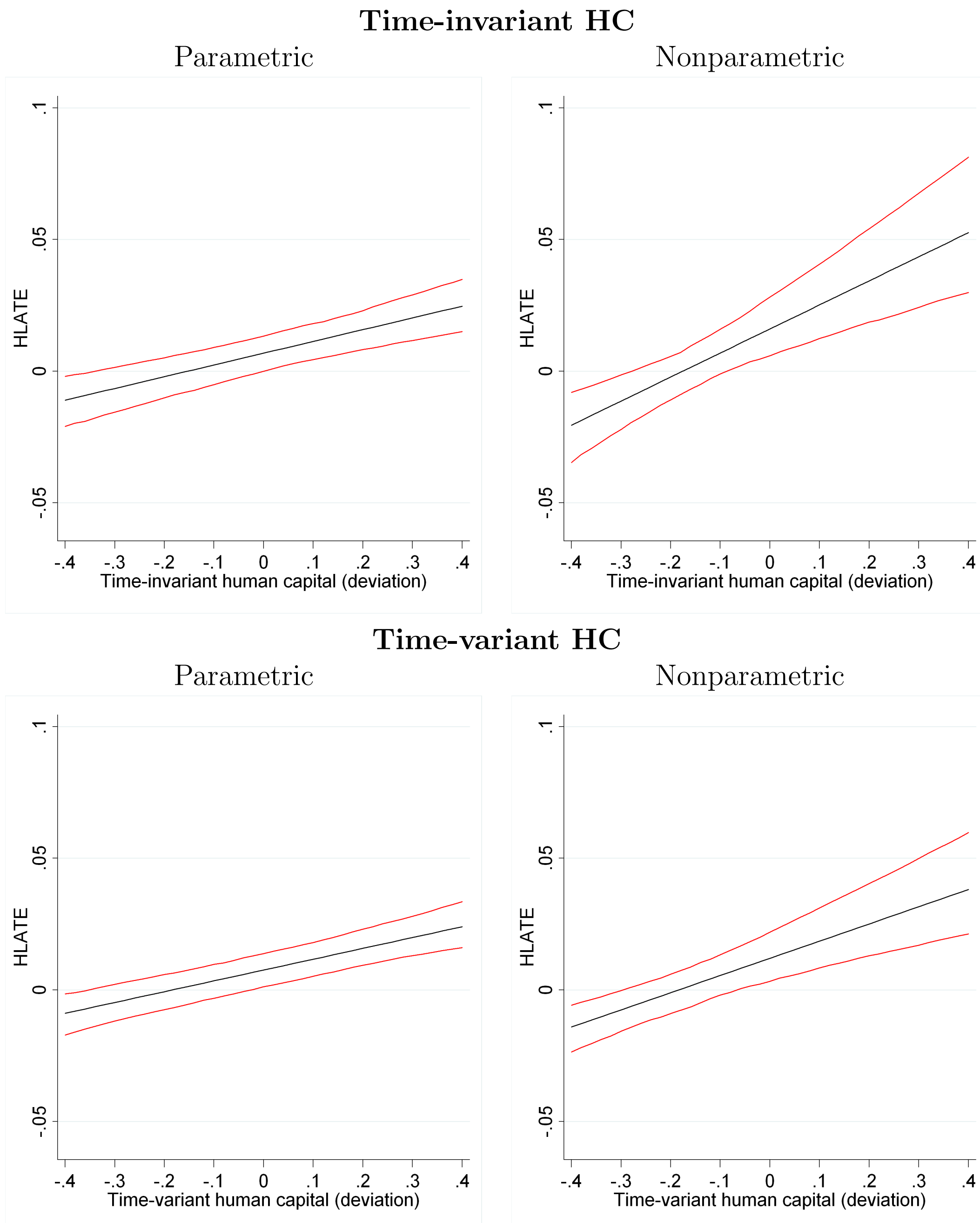
Figure 4: LN(Investments/Capita), Objective 1 Treatment, and absorptive capacity





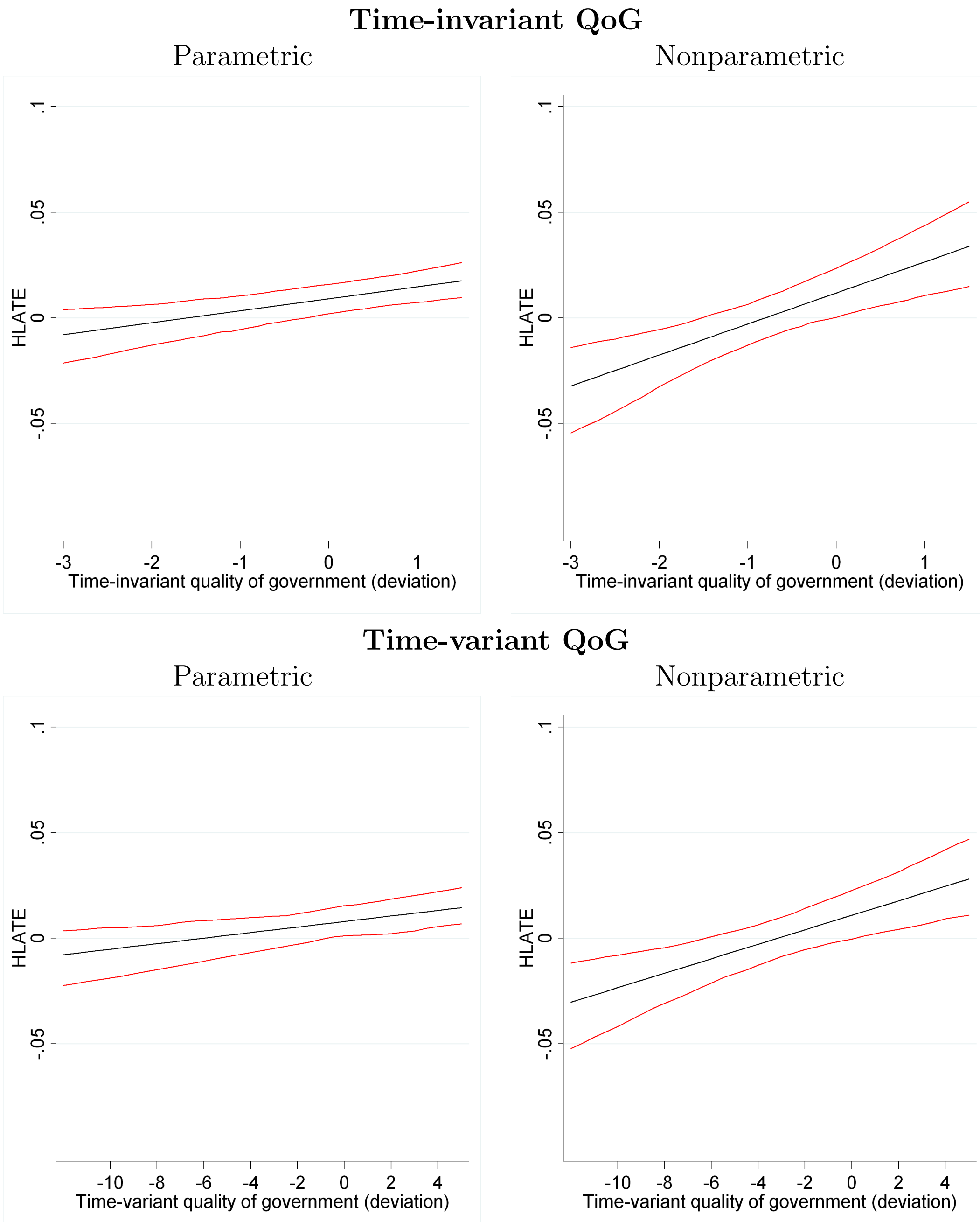
Note: The upper and lower figures illustrate the relationship between the outcome, forcing variable, human capital, and quality of government, respectively. The red (blue) dots indicate observations which received (did not receive) Objective 1 treatment. The surfaces represent 5th-order polynomial functions of per-capita GDP and linear functions of human capital and quality of government, respectively. These functions are estimated on both sides of the 75% threshold separately.

Figure 5: HLATE and GDP/Capita Growth rate for different levels of Human Capital



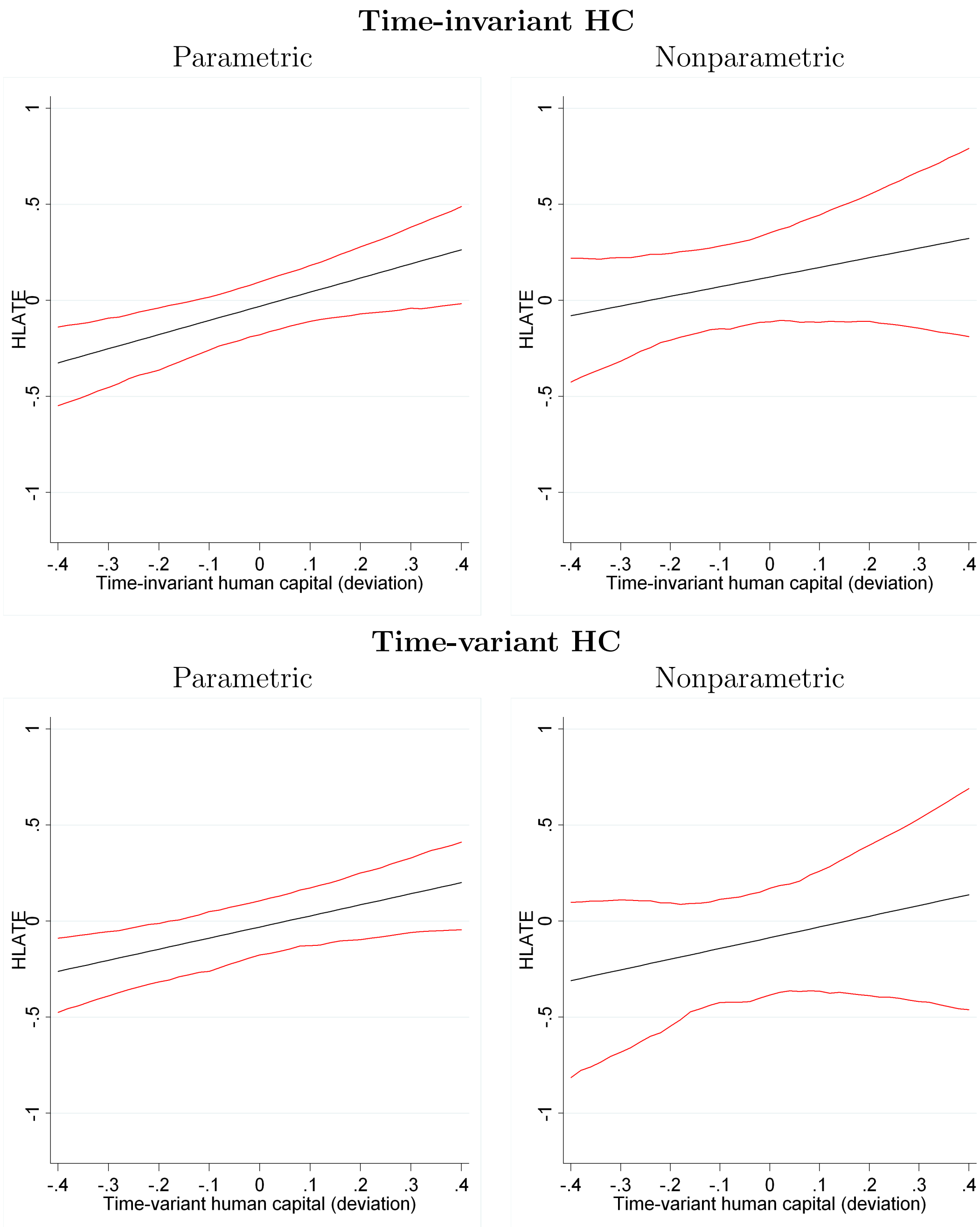
Note: The black line illustrates the point estimates, the red lines represent the 90% confidence intervals. The parametric estimates are derived from a specification with 3rd-order polynomials of initial GDP per-capita and linear human capital. The nonparametric estimates are based on an optimal bandwidth selection procedure following Ludwig and Miller (2007). The confidence intervals are derived from bootstrapped standard errors with 500 replications.

Figure 6: HLATE and GDP/Capita Growth rate for different levels of Quality of Government



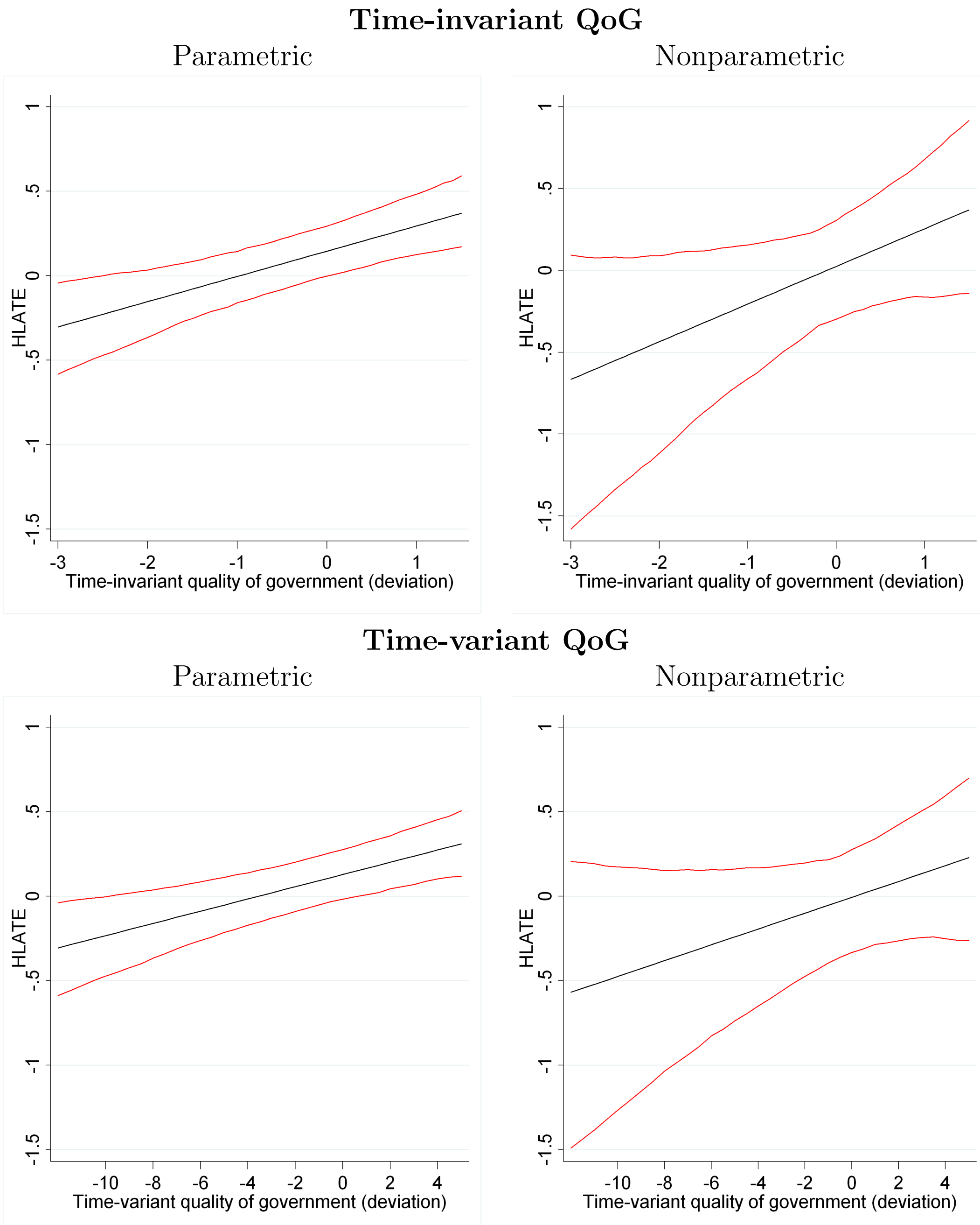
Note: The black line illustrates the point estimates, the red lines represent the 90% confidence intervals. The parametric estimates are derived from a specification with 3rd-order polynomials of initial GDP per capita and linear quality of government. The nonparametric estimates are based on an optimal bandwidth selection procedure following Ludwig and Miller (2007). The confidence intervals are derived from bootstrapped standard errors with 500 replications.

Figure 7: HLATE and LN(Investments/Capita) for different levels of Human Capital



Note: The black line illustrates the point estimates, the red lines represent the 90% confidence intervals. The parametric estimates are derived from a specification with 3rd-order polynomials of initial GDP per-capita and linear human capital. The nonparametric estimates are based on an optimal bandwidth selection procedure following Ludwig and Miller (2007). The confidence intervals are derived from bootstrapped standard errors with 500 replications.

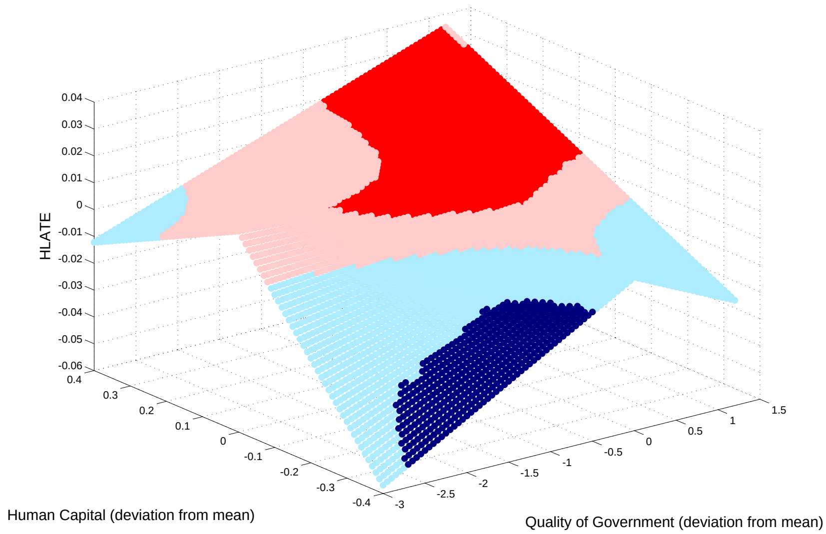
Figure 8: HLATE and LN(Investments/Capita) for different levels of Quality of Government



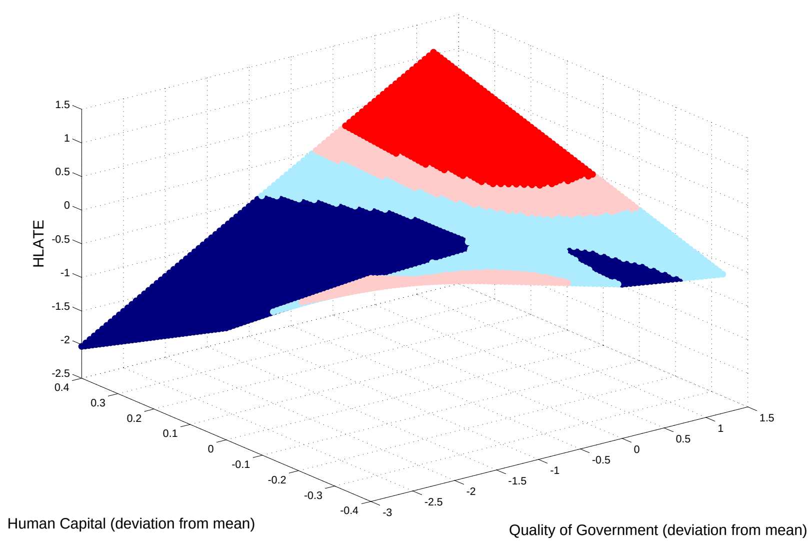
Note: The black line illustrates the point estimates, the red lines represent the 90% confidence intervals. The parametric estimates are derived from a specification with 3rd-order polynomials of initial GDP per capita and linear quality of government. The nonparametric estimates are based on an optimal bandwidth selection procedure following Ludwig and Miller (2007). The confidence intervals are derived from bootstrapped standard errors with 500 replications.

Figure 9: HLATE for different levels of Human Capital (HC) and Quality of Government (QoG)

GDP/Capita Growth Rate



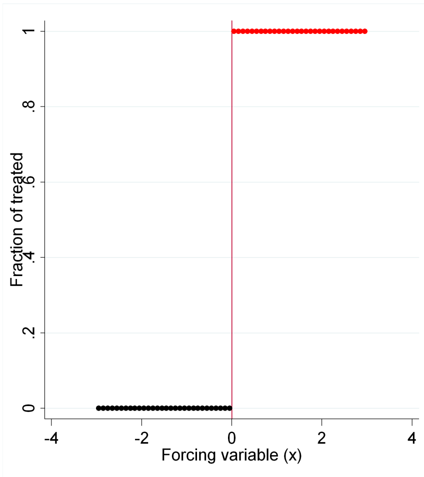
LN(Investments/Capita)



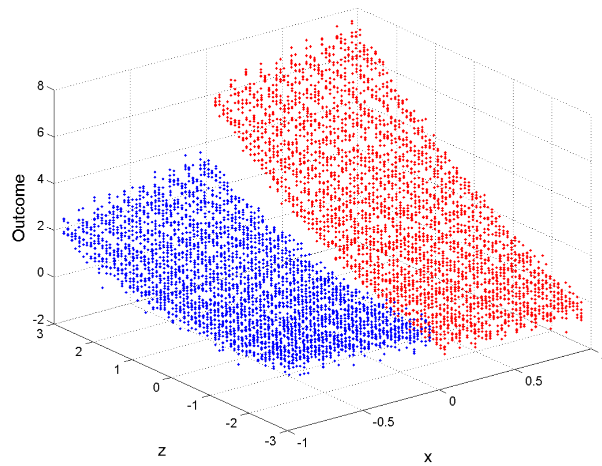
Note: The light-red and light-blue areas refer to insignificant positive and insignificant negative effects, respectively. The dark-red and dark-blue areas indicate significant positive and significant negative effects, respectively. We choose the 90% confidence interval – calculated on the basis of bootstrapped standard errors with 500 replications – to determine significance of the HLATE. The predictions stem from parametric OLS regressions with a 3rd-order polynomial of per-capita GDP and linear HC and QoG.

Figure 10: Sharp RDD (1-Way Threshold)

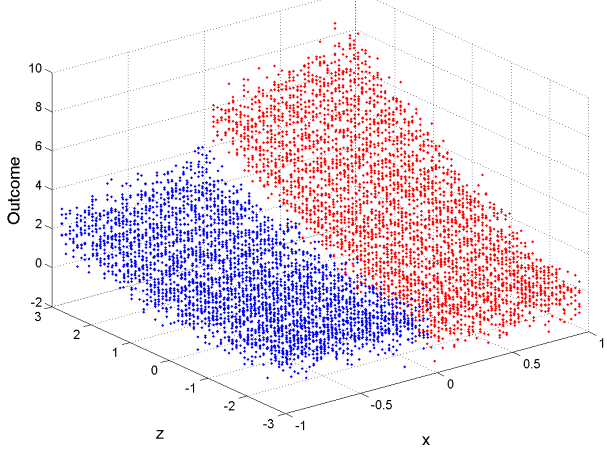
Treatment



Outcome ($\sigma_\epsilon = 0.3$)



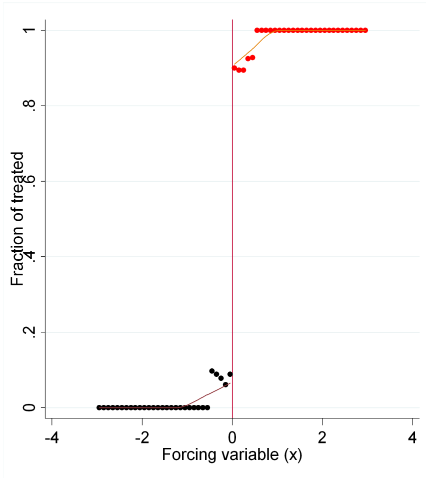
Outcome ($\sigma_\epsilon = 0.6$)



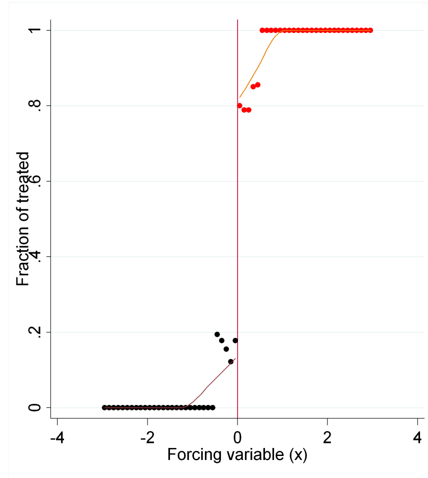
Note: The upper left figure shows average treatment rates in equally-sized bins of 0.1 which are plotted against the forcing variable x. The other two figures show average outcome rates plotted against the forcing variable x and the interaction variable z. Blue (red) dots indicate untreated (treated) observations. For illustration purpose, we focus on the range $x = [-1, 1]$. σ_ϵ refers to the standard deviation of the error term in the outcome function. That is, the greater is σ_ϵ the less precise is the control function of x.

Figure 11: Fuzzy RDD (1-Way Threshold)

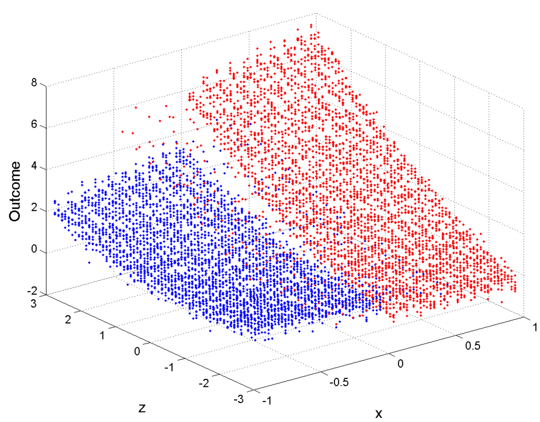
Treatment (Fuzzy 1)



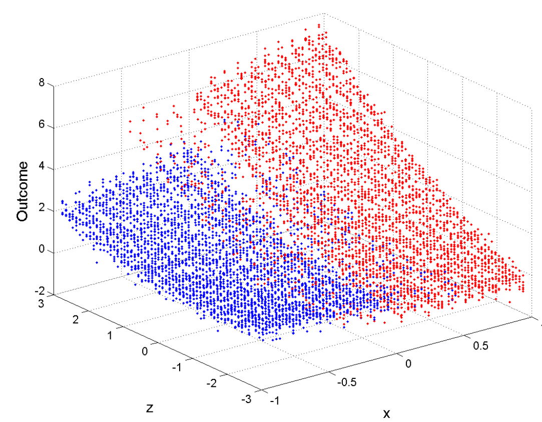
Treatment (Fuzzy 2)



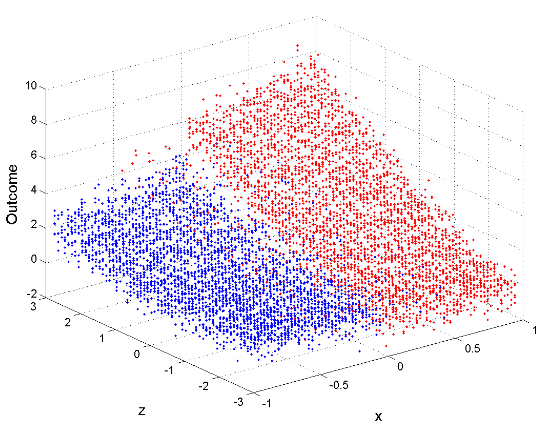
Outcome (Fuzzy 1, $\sigma_\epsilon = 0.3$)



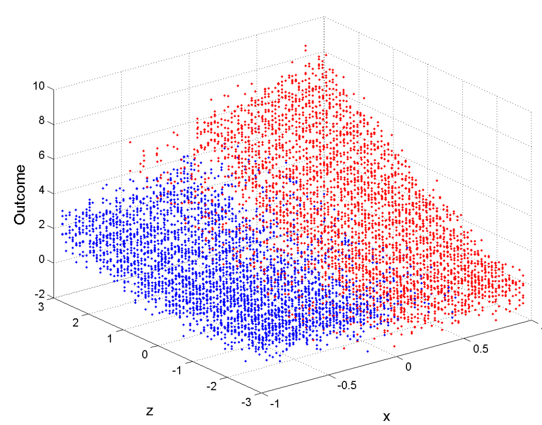
Outcome (Fuzzy 2, $\sigma_\epsilon = 0.3$)



Outcome (Fuzzy 1, $\sigma_\epsilon = 0.6$)

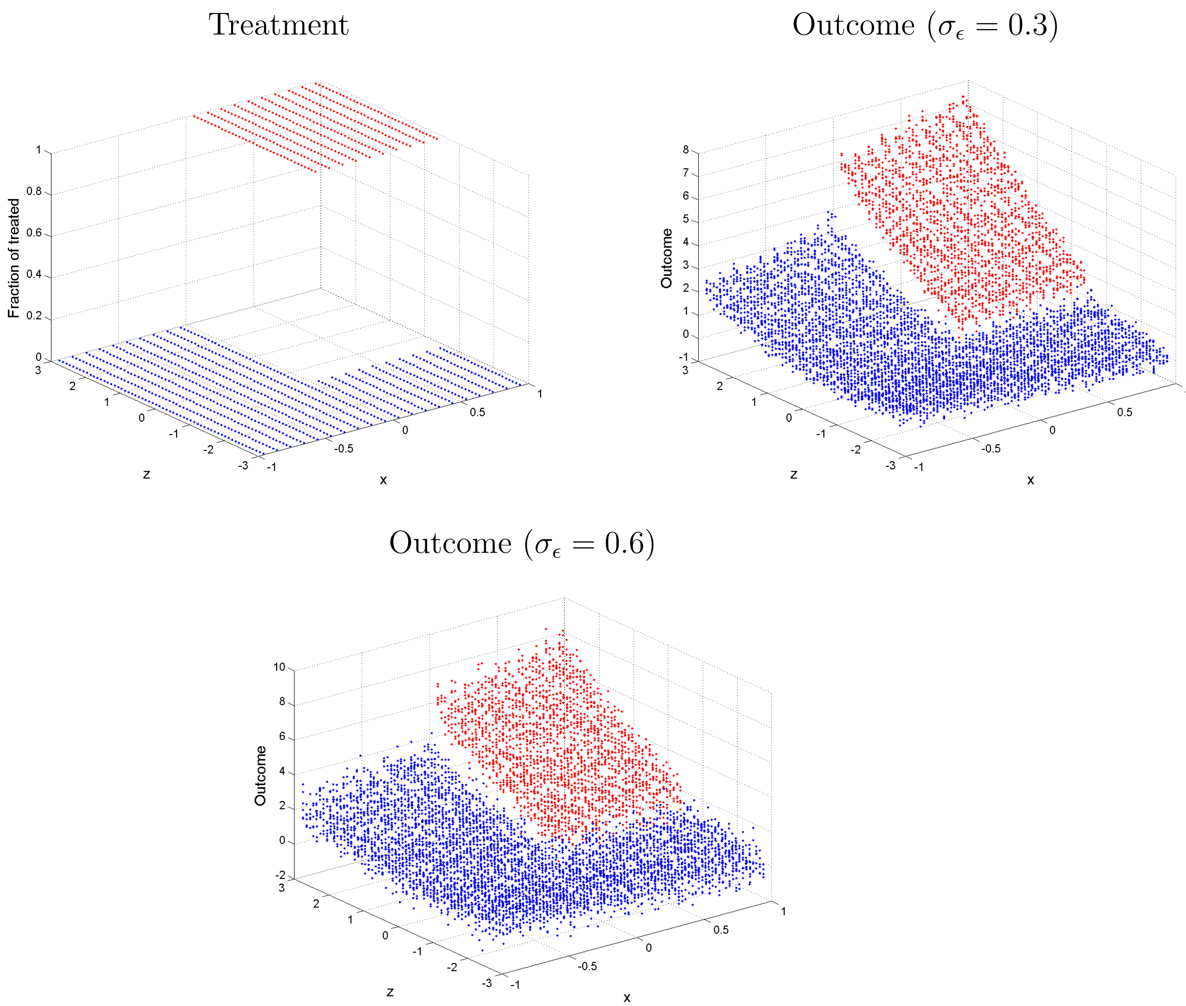


Outcome (Fuzzy 2, $\sigma_\epsilon = 0.6$)



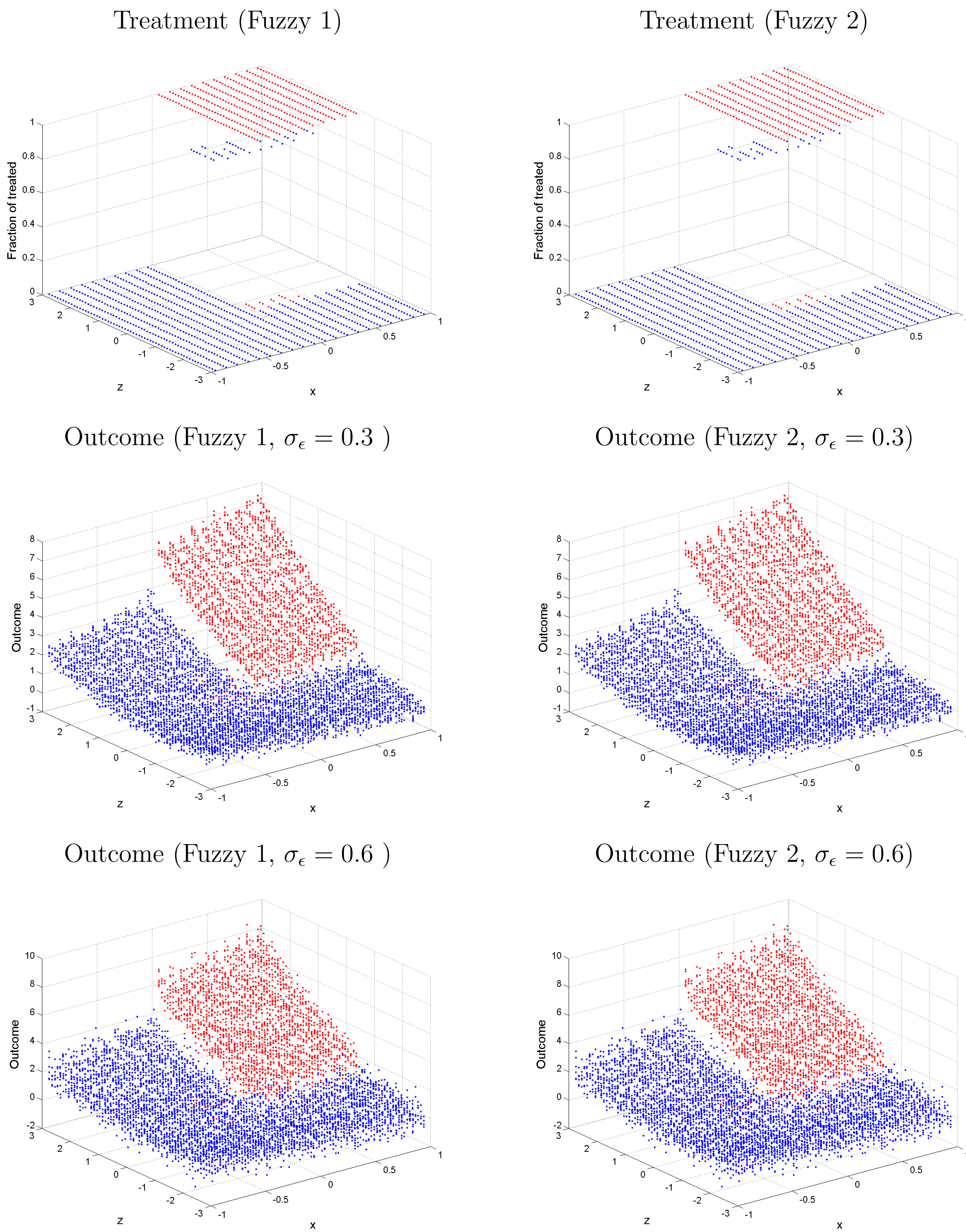
Note: The upper figures show average treatment rates in equally-sized bins of 0.1 which are plotted against the forcing variable x. The figures in the two lower panels show average outcome rates plotted against the forcing variable x and the interaction variable z. Blue (red) dots indicate to untreated (treated) observations. For illustration purpose, we focus on the range $x = [-1, 1]$. σ_ϵ refers to the standard deviation of the error term in the outcome function while Fuzzy 1 (2) indicates a misassignment probability of 1/12 (1/6). Accordingly, the greater is σ_ϵ the less precise is the control function of x, and Fuzzy 2 represents a less precise relationship between the treatment the treatment rule than Fuzzy 1.

Figure 12: Sharp RDD (2-Way Threshold)



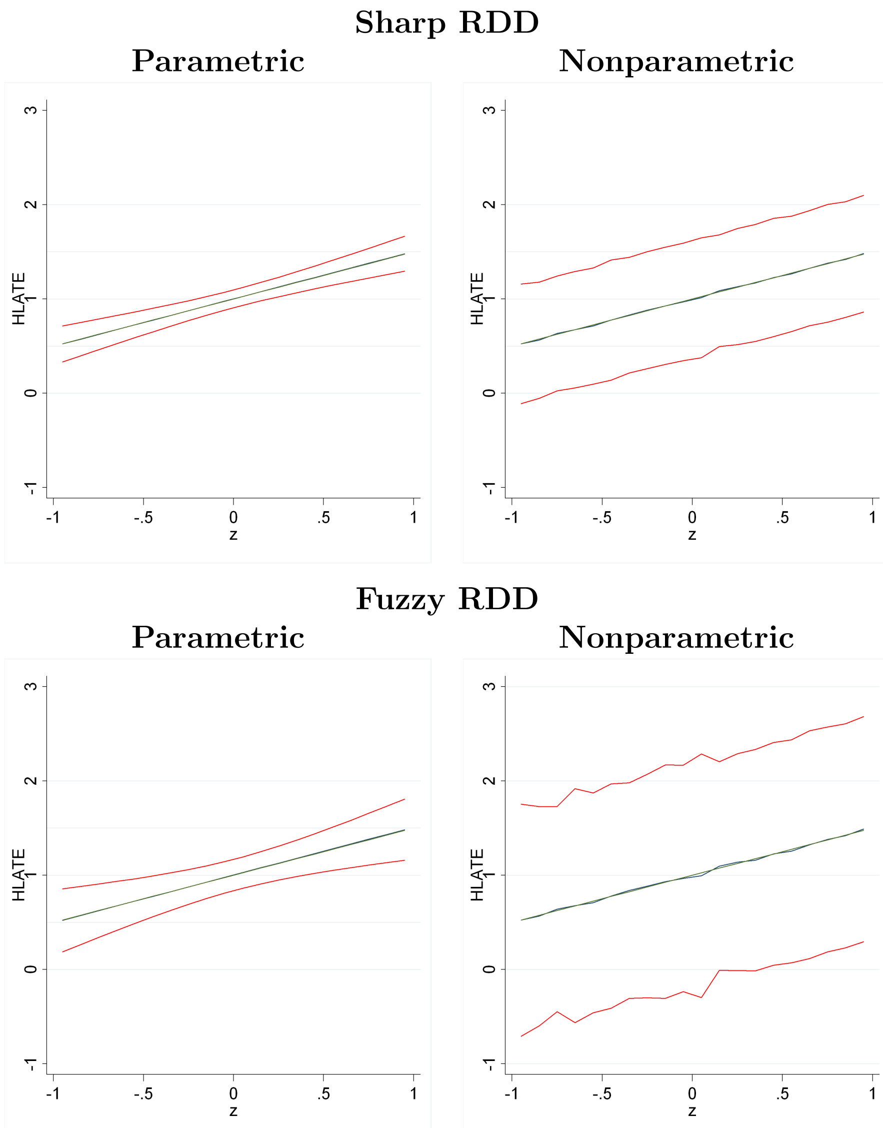
Note: The upper left figure shows average treatment rates in equally-sized bins of 0.1 which are plotted against the two forcing variables x and z. The other two figures show average outcome rates plotted against the forcing variables x and z. In addition to determining the treatment probability, z affects the treatment effect via an interaction term in the outcome equation. σ_ϵ refers to the standard deviation of the error term in the outcome function. That is, the greater is σ_ϵ the less precise is the control function of x.

Figure 13: Fuzzy RDD (2-Way Threshold)



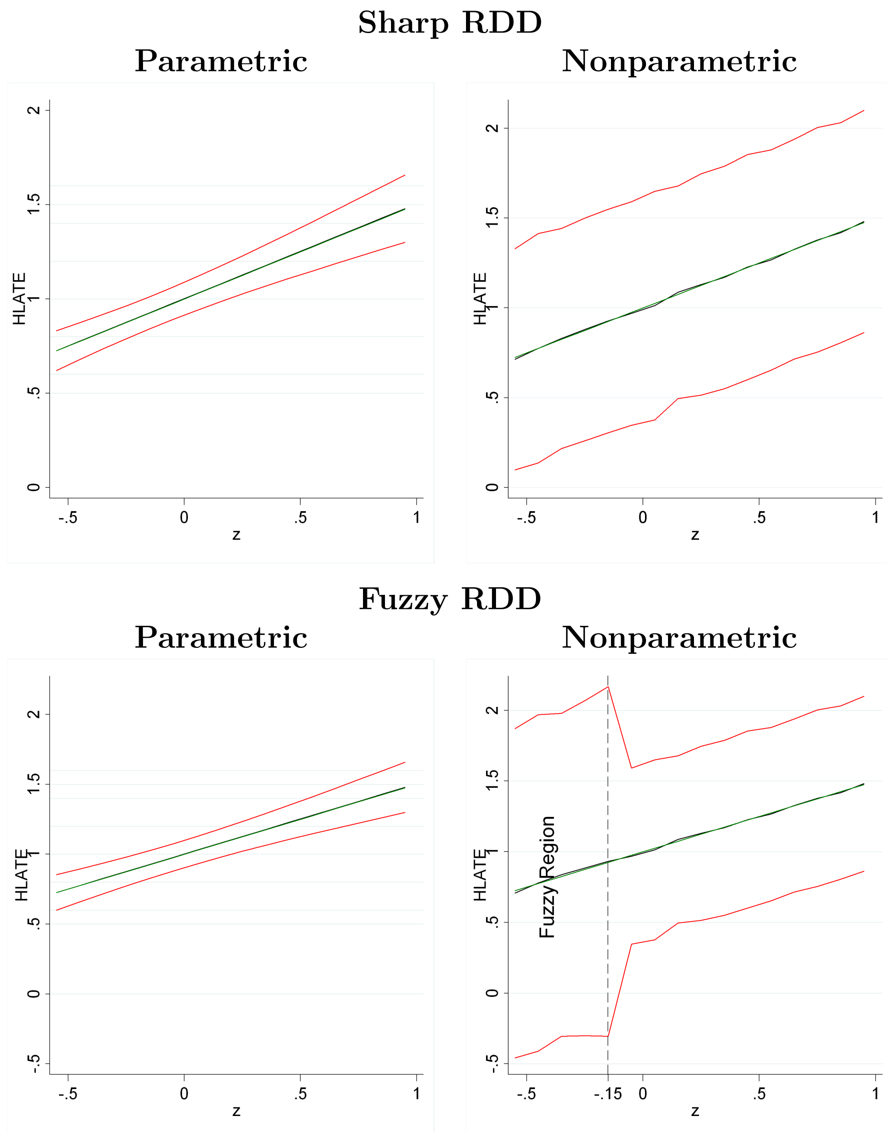
Note: The upper figures show average treatment rates in equally-sized bins of 0.1 which are plotted against the forcing variables x and z where red (blue) dots indicate observations that qualify (do not qualify) for treatment according to the treatment rule. The figures in the two lower panels show average outcome rates plotted against the forcing variables x and z where red (blue) dots indicate observations that did (did not) receive treatment. In addition to determining the treatment probability, z affects the treatment effect via an interaction term in the outcome equation. σ_ϵ refers to the standard deviation of the error term in the outcome function, while Fuzzy 1 (2) indicates a misassignment probability of 1/8 (1/4). Accordingly, the greater is σ_ϵ the less precise is the control function of x, and z, and Fuzzy 2 represents a less precise relationship between the treatment the treatment rule than Fuzzy 1.

Figure 14: Heterogenous local average treatment effects (1-Way Threshold)



Note: The figures show treatment effects at the x_0 threshold (we restrict the sample to one bin on each side of x_0) plotted against the interaction variable z. All figures are based on experiments with $\sigma_\epsilon = 0.6$ where the fuzzy design refers to a data-generating process with a misassignment probability 1/6. The parametric figures are derived from an $N = 20^2 \cdot 6$ sample. For the nonparametric figures, we choose a bandwidth of 1/6. The green line illustrates the true effect, the black line illustrates the point estimates, and the red lines represent the 90% confidence intervals.

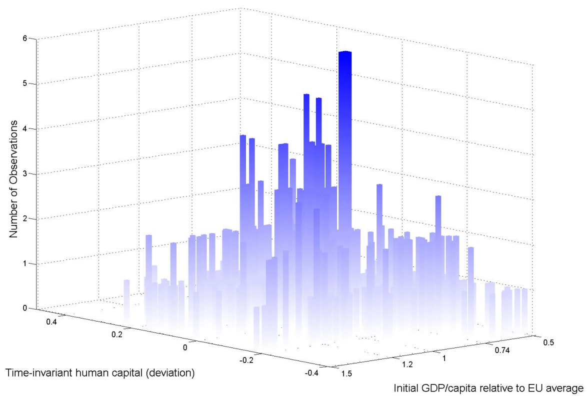
Figure 15: Heterogenous local average treatment effects (2-Way Threshold)



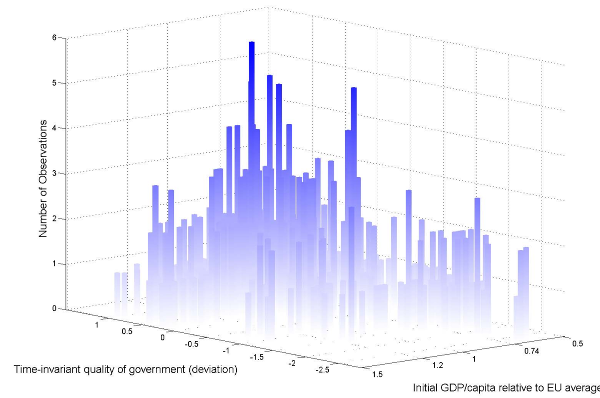
Note: The figures show treatment effects at the x_0 threshold (we restrict the sample to one bin on each side of x_0) plotted against the interaction variable z. All figures are based on experiments with $\sigma_\epsilon = 0.6$ where the fuzzy design refers to a data-generating process with a misassignment probability 1/6. Note that the fuzzyness is bounded in the z dimension by [-1.05, -0.15] which results in a lower degree of precision of the HLATE in this interval. The parametric figures are derived from an $N = 20^2 \cdot 6$ sample. For the nonparametric figures, we choose a bandwidth of 1/6. The green line illustrates the true effect, the black line illustrates the point estimates, and the red lines represent the 90% confidence intervals.

Figure 16: Frequency plots

Human Capital (HC)



Quality of Government (QoG)



Note: The two figures illustrate the number of observations in the human capital/per-capita GDP bins and the quality of government/per-capita GDP bins, respectively. These bins correspond to the ones used in Figure 3.