$u^b$

b
UNIVERSITÄT
BERN

CRED
CENTER FOR REGIONAL
ECONOMIC DEVELOPMENT

# The persistent effects of place-based policy:
# Evidence from the West-German Zonenrandgebiet

## CRED Research Paper No. 9

Maximilian von Ehrlich
University of Bern,
CRED

Tobias Seidel
University of Duisburg-Essen

April 2016

## Abstract

We show that temporary place-based subsidies generate persistent effects on economic density. As our design allows us to control for agglomeration economies, we attribute an important role to policy-induced locational advantage (e.g. capital structures) in explaining persistent spatial patterns of economic activity. With regard to distributional implications, we show that subsidies have capitalized in land rents, so pre-treatment land owners have benefitted predominantly from the program.

**Center for Regional Economic Development (CRED)**

University of Bern

Schanzeneckstrasse 1

P.O.Box 8573

CH-3001 Bern

Telephone: +41 31 631 37 11

Fax: +41 31 631 34 15

E-Mail: info@cred.unibe.ch

Website: http://www.cred.unibe.ch

The Center for Regional Economic Development (CRED) is an interdisciplinary hub for the scientific analysis of questions of regional economic development. The Center encompasses an association of scientists dedicated to examining regional development from an economic, geographic and business perspective.

**Contact of the authors:**

Maximilian von Ehrlich

University of Bern

Schanzeneckstrasse 1

P.O.Box 8573

CH-3001 Bern

Telephone: +41 31 631 80 75

Email: maximilian.vonehrlich@vwi.unibe.ch

Tobias Seidel

University of Duisburg-Essen

Lotharstraße 65

D-47057 Duisburg

Telephone: +49 (203) 379 - 2625

Email: tobias.seidel@uni-due.de

# 1 Introduction

When supporting underdeveloped regions, policy makers often hope that temporary transfers establish self-sustaining long-run economic development. The effort is substantial. For example, the EU dedicates about one third of its overall budget 2014-2020 to regional policy amounting to more than 350 billion euros (European Commission, 2011). The US does not have a unified regional policy, but annual spending on regional development programs is estimated at about 95 billion US dollars per year (Government Accountability Office, 2012). Also China has installed regional policies that resemble those in the EU in terms of instruments and magnitude (European Commission, 2010).

Despite these efforts, little is known about the long-term consequences of these programs and their underlying mechanisms (Neumark and Simpson, 2015).[1] Using a natural experiment from Germany allows us to make progress in this direction. In 1971, the West German government started a large scale transfer program to stimulate economic development in a well-defined geographical area adjacent to the Iron Curtain. All districts that accommodated either 50 percent of their area or population within a distance of 40 kilometers to the inner-German and Czechoslovakian border on January 1, 1971 became part of the *Zonenrandgebiet* (ZRG).[2] As shown in Figure 1, it stretched from the Danish border in the North to the Austrian border in the South running through four states (Bavaria, Hesse, Lower-Saxony, and Schleswig-Holstein). A major reason for this privileged treatment was to compensate firms and households close to the eastern border for being cut off adjacent markets on the other side of the Iron Curtain. Policy makers were afraid that the remoteness could cause substantial out-migration to the western parts of the country.[3] The program was not intended for a fixed number of years and its termination came as unexpectedly as German reunification. As transfers were redirected towards East Germany after 1990, the place-based policy was phased out until 1994. We are therefore able to study both the contemporaneous and persistent effects of the policy.

---

[1]The literature on placed-based policies has mostly looked into the effects of transfers *during* programs, e.g. Busso, Gregory, and Kline (2013) evaluate the federal empowerment zones program in the US, Glaeser and Gottlieb (2008) examine the place-based policy Appalachian Regional Commission, Gobillon, Magnac, and Selod (2012) study the French enterprise zone program, Becker, Egger, and Ehrlich (2010) focus on income and employment effects of EU Structural Funds.

[2]See Deutscher Bundestag (1970), Drucksache VI/796 and Ziegler (1992, p.9). Zonenrandgebiet literally means area adjacent to the (Soviet occupation) zone, that became the German Democratic Republic. It was common in West Germany to refer to the German Democratic Republic as the "Zone".

[3]See Ziegler (1992) for a more detailed exposition.

Figure 1: THE GERMAN ZONENRANDGEBIET (ZRG), 1971-1994
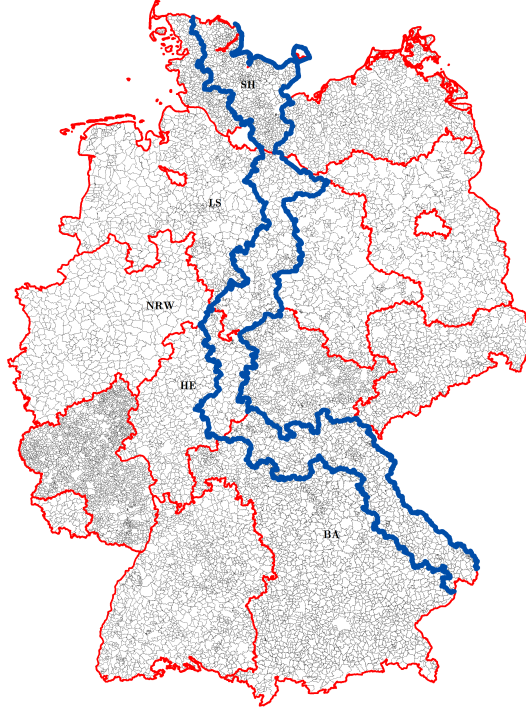


*Note*: The blue lines mark the western border of the ZRG and the Iron Curtain. The grey lines represent the municipalities according to the 1997 classification while the red lines define state borders. The border of the ZRG follows the administrative districts according to the 1971 classification which was modified substantially in the mid-1970s. In most of our analysis we consider the states Schleswig-Holstein (SH), Lower Saxony (LS), North Rhine-Westphalia (NRW), Hesse (HE), and Bavaria (BA).

The institutional setting of the ZRG gives rise to two types of discontinuities that we can use for identification of causal effects. First, we apply a spatial regression discontinuity design (RDD) based on municipalities and grid cells in a close neighborhood on either side of the treatment border. If other relevant factors vary continuously at the ZRG-border, a discontinuity in economic activity at this border can be interpreted as the causal effect of the place-based policy. As the treatment border does not separate areas with different institutions, many concerns of other discontinuities that are important at country borders can be ruled out. Nevertheless, administrative borders are unlikely to be drawn randomly. To establish more credibility of our results, we also exploit the political rule which governed the location of the treatment border. As the treatment probability of districts jumps at a distance of 40 kilometers from the Iron Curtain, we apply a classical

regression discontinuity design. The advantage is that the 40-kilometer rule does not coincide with any administrative boundary nor with geographic features that may cause discontinuities in relevant determinants for outcome. Depending on parametric or non-parametric estimation and the choice of the control function, we find that regional transfers led to higher income per square kilometer in the treatment area by about 30-50 percent in 1986. Undertaking the same exercise for 2010, that is 16 years after the program was eventually stopped, there is no indication that the estimated effects have diminished. We provide evidence that a substantial part of this discontinuity can be explained by local relocation of economic activity. Further, we demonstrate that higher income per square kilometer was driven by higher private and public capital intensity and higher population and employment – both in the mid-1980s and in 2010. However, we find no evidence that the educational composition of the workforce was affected. Identifying causal persistent effects of a place-based policy and their underlying channels establishes a first contribution of this paper.[4]

There is a lively debate in the literature about the *determinants* of persistence in the spatial distribution of economic activity in the long run.[5] For example, Bleakley and Lin (2012) show that historical portage sites in the US serve as a good predictor for the location of cities today. As they do not find substantial differences in the capital stock or capital intensity between portage and non-portage sites, they interpret their findings as evidence in favor of agglomeration economies rather than capital structures. Kline and Moretti (2014) analyze the effects of the Tennessee Valley Authority program showing that manufacturing employment continued to grow after transfers into this region were stopped in 1960. They argue that initial capital investments before 1960 are unlikely to be responsible for this long-run effect as the capital stock would have depreciated several decades later.[6]

It is in the nature of spatial RDD that all determinants of outcome that are continuous at the border are implicitly controlled for. Based on recent work by Turner,

---

[4]Kline and Moretti (2014) have found persistent effects of a place-based policy in the US while Ahlfeldt, Maennig, and Richter (2016) find no persistence of urban renewal policies in Berlin.

[5]For syntheses of the theoretical literature on agglomeration economies see Duranton and Puga (2004).

[6]Schumann (2014) documents persistent effects of different levels of local population due to different settlement policies for refugees in the American and French occupation zones in Germany after World War II. Redding, Sturm, and Wolf (2011) regard the persistent relocation of the main German airport from Berlin to Frankfurt (initiated during the Cold War) as evidence for multiple spatial equilibria. Michaels and Rauch (2014) study the impact of the Roman Empire for the evolution of urban structures in France and the UK.

Haughwout, and van der Klaauw (2014), if externalities (e.g. agglomeration economies like labor-market pooling, technology spillovers, or the home-market effect, among others) dissipate continuously with distance, these determinants cannot be part of the local average treatment effect. However, they certainly do matter for explaining differences in economic density of locations that are more distant from the treatment border. We exploit information from satellite data (capital structures and radiance) at a very fine scale of up to $100m \times 100m$, study subsamples of municipality pairs that are well-connected by transport networks, not separated by undeveloped land (e.g. forests), or characterized by a polycentric structure to verify the plausibility of this assumption. As discontinuities remain prevalent in these exercises, we argue that capital structures play an important role in explaining persistence in spatial outcomes. This reasoning is in line with our findings that the policy has led to an expansion of capital structures (e.g. industrial parks, roads, power networks, and sewage systems) in the Zonenrandgebiet. These structures are likely to generate a persistent effect because the associated planning process has a long-term value. It is easier to maintain established structures than planning new ones on the green field. Moreover, higher economic density generates higher tax revenues that can be reinvested to maintain established structures or local governments have learned to manage commercial activities better during the times of treatment. In essence, these channels are able to raise local productivity of firms and thereby maintain higher investment levels across a historic treatment border.[7] A last explanation for persistence that we explore is potential interactions of shocks with economic density. Studying German reunification in 1990 and EU enlargement in 2004 shows, however, that these arguably important shocks contribute only 2 to 6 percent to the overall treatment effect. Isolating policy-generated locational advantage as an additional explanation for persistent spatial outcomes is a second main contribution of the paper.

Third, we explore the distributional implications of the place-based policy. Policy makers often initiate place-based policies to raise wages and employment, in particular of poor households.[8] However, there is concern that regional transfers eventually capitalize in higher land rents (Glaeser and Gottlieb, 2008) such that the beneficiaries of the policy are those households that owned property *before* the program. If land supply is not infinitely elastic, an increase in economic activity eventually leads to higher land prices.

---

[7]In this regard, our paper links to Davis and Weinstein (2002, 2008) attributing a key role to local structures in explaining the long-term spatial pattern of economic activity.

[8]European Commission (2014).

Our results confirm these concerns. We find that ZRG-transfers raised land rents by around 30 percent both in 1988 and 2010 which approximately offset the nominal per-capita income gain in the recipient regions. In a Rosen-Roback framework, this points to local persistent production amenities that – according to our empirical approach – have to be discontinuous at the treatment border (cf. Rosen, 1979; Roback, 1982). In sum, we show that the place-based policy did affect the spatial economic structure persistently, but the beneficiaries of the subsidies program are (pre-treatment) land owners.

The paper is organized as follows. In the next section, we provide an overview of the historical and institutional background of the transfer program. Sections 3 and 4 introduce the data and identification strategies we use. We present results in section 5 and analyze the underlying mechanisms of persistence. We study the distributional implications of the policy in section 6 before offering concluding remarks.

## 2 Historical background

As Germany's surrender in the Second World War became more likely, the Allied Forces started negotiations about the borders of post-war Germany and the division among the US, the UK, France and the Soviet Union in 1943. Different political ideologies caused growing tensions between the Western Allies and the Soviet Union and eventually led to the division of the country into the Federal Republic of Germany (West Germany) and the German Democratic Republic (East Germany). When the government in East Germany began to install fences and even a death strip at the inner-German border in 1952, passage of goods and people became impossible. Regular transit was only allowed between East and West Berlin until the erection of the Berlin Wall on August 13, 1961 finally closed this last loop hole for nearly 30 years.

While regional transfers in the 1950s targeted primarily former industrial centers that were heavily bombed during the war, politicians in West Germany also responded to the new situation of a divided state.[9] Districts at the inner-German border received support to prevent outmigration of residents and firms. This was a serious concern as the Iron Curtain deteriorated the living conditions for both psychological and economic reasons. At this point, West German policy makers widely regarded the division of Germany as a temporary phenomenon such that transfers were justified to preserve the economic position

---

[9]See Karl (2008) for a more detailed review of regional policy in West Germany.

of the geographical center of pre-war Germany for the time after reunification.[10] Hence, politicians recognized the potentially long-lasting consequence of an event which was then still considered temporary. A further motivation for privileged treatment of the ZRG was geopolitical. An economically strong border region was expected to provide a better buffer against a potential attack of Warsaw Pact troops (Ziegler, 1992).

However, there was no clear rule yet for the allocation of resources. It was not until the late 1960s that the Federal Ministry of Economics suggested a better coordination of regional policy.[11] While a politically-installed committee decided about the eligibility of regions to receive transfers, the Zonenrandgebiet was guaranteed privileged support by law (Zonenrandförderungsgesetz, 1971) within this framework. The federal law of 1971 provided a transparent definition of the ZRG that was never modified until ZRG treatment was eventually stopped in 1994: All districts that accommodated at least 50 percent of their area or population within 40 kilometers to the inner-German or Czechoslovakian border on January 1, 1971 became part of the Zonenrandgebiet.[12] Its area accounted for 18.6 percent of the West German territory accommodating 12.3 percent of the population (see Table A1). It is remarkable that the ZRG boundaries were never modified despite substantial changes in district and municipality borders, especially in the mid-1970s. The ZRG program lost its status in 1994 when Germany was reunified and the focus of regional policy abruptly shifted to the development of the 'New Länder'.

The ZRG transfer scheme comprised a menu of measures. A major focus was laid on subsidies for firm investment. Firms inside the Zonenrandgebiet could apply for investment subsidies of up to 25 percent. For initial investment, the total value of direct subsidies and tax deductions could even reach about 50 percent of the investment volume.[13] Further, firms were eligible for superior credit conditions of the public bank KfW (Kreditanstalt

---

[10]Bundesministerium für innerdeutsche Beziehungen (1987).

[11]This initiative led to two important laws in 1969: (i) the Joint Task "Improvement of the Regional Economic Structure" (Gemeinschaftsaufgabe Verbesserung der regionalen Wirtschaftsstruktur, GRW) (see Eckey (2008) for a historical overview of the Joint Task) and (ii) the Investment Premium Law (Investitionszulagengesetz).

[12]See Deutscher Bundestag (1970), Drucksache VI/796 and Ziegler (1992, p.9). According to a statement by state secretary Sauerborn, the 40-kilometer rule also included less needy regions, but was appealing for practical reasons in the first place (see Protocol of the 39th session of the cabinet committee of economics). It was recognized in the parliamentary debate on June 17, 1971 that the treatment border must remain fixed over time in order to rule out strategic modifications of local district borders (see Protocol of the 128th session of the Bundestag).

[13]See Ziegler (1992) and Zonenrandförderungsgesetz (ZRFG), 1971, available at http://dipbt.bundestag.de/dip21/btd/11/050/1105099.pdf (Anhang 3).

für Wiederaufbau), capital allowances were more generous, and there was a large program of public debt guarantees. Moreover, companies located in the ZRG were treated with priority in public tendering. Beyond firm subsidies, a substantial share of the budget was dedicated towards public infrastructure projects and transfers could also be used for renovation of houses, investments in social housing, day care centers, education and cultural activities. This heterogeneity of measures makes it impossible to report a single money value of the ZRG program.

While the overall figure is unavailable, we do have data on certain parts of the ZRG program (e.g. subsidies from the Investment Premium Law). This allows us to document that the ZRG received the lion's share of the transfer budget. Between 1984 and 1987 about 60 and 85 percent of total public transfers in the states we consider was directed to the ZRG.[14] Note that data on tax deductions, the value of public tenders, and other monetary advantages that applied specifically to the ZRG are not available such that the treatment intensity of the ZRG was even higher than these numbers suggest. To get an idea about the overall size of the program, estimates range between 1.3-2.5 billion euros (at 2010 prices) per year in the 1980s which amounts to about 194-373 euros per capita.[15] This makes it comparable to the size of current EU Structural Funds amounting to annual transfers of about 230 euros per capita in regions with the highest transfer intensity (Becker, Egger, and Ehrlich, 2010).

## 3  Data

The basis of our empirical work is geographical and administrative data from municipalities and the exact location of the Zonenrandgebiet border. According to the precise definition of the ZRG, we georeference a map of West German districts in 1971 to identify the exact location of both the Iron Curtain (inner-German and Czechoslovakian border) and the ZRG border that separates the treatment from the control area.[16] To georeference municipality data, we use digital maps (shape files) from the Bundesamt für Kartographie und Geodäsie.

---

[14]Documentation of the Joint Task, Rahmenplan No. 13, available at www.bundestag.de.

[15]See Ziegler (1992) and Wirtschaftswoche, 1990 Nr. 99/4. About 6.7 million individuals were living in the ZRG in 1961.

[16]The map we use is provided by the former Bundesforschungsanstalt für Landeskunde und Raumforschung at a scale of 1:1,000,000.

Table 1: OBSERVATIONAL UNITS

| | No. municipalities | | | | No. districts | |
| | Total | | Boundary sample | | | |
| | 1986 | 2010 | 1986 | 2010 | Total | Boundary sample |
|---|---|---|---|---|---|---|
| Non-ZRG | 3,367 | 3,391 | 2,298 | 2,305 | 396 | 202 |
| ZRG | 1,573 | 1,576 | 1,572 | 1,576 | 106 | 107 |
| Total | 4,940 | 4,967 | 3,870 | 3,881 | 502 | 309 |

*Notes*: We consider the states (Länder) Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Hesse, and Bavaria. These five states comprise in total 4,991 and 5,018 populated municipalities in 1986 and 2010, respectively. We lose 51 municipalities due to partial treatment (i.e. ZRG border crosses the municipality) and imprecise assignment to municipal boundaries in the digital maps (see Appendix A for details). The boundary sample on the municipality level contains all municipalities with a distance to the ZRG border of less than 100km; the boundary sample on the district level includes all districts with $M_d \leq 150$ (see section 6). Districts are based on the 1969 classification, municipalities on the 1997 and 2010 classifications.

This georeferencing provides us with relevant distance measures for each municipality and coordinates that we use as controls in several econometric specifications. We compile a unique data set on municipality characteristics between 1984 and 2012 and merge it with the information on location and district affiliation in 1971. In most cases and depending on data availability, we refer to the year 1986 for *contemporaneous* effects of the policy, and estimate the *persistent* effects in 2010.[17] We use (taxable) nominal income per square kilometer as our main proxy of overall economic activity. We further use data broadly categorized in measures of local labor and capital intensity which we introduce below. Details about the data and data sources are provided in Appendix A.

We use two different samples based on municipalities and districts, respectively (Table 1). This is required by the econometric approaches we introduce below. We consider the five states (Länder) that include or border on the treated region: Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Hesse, and Bavaria comprising in *total* 4,991 and 5,018 populated municipalities in 1986 and 2010, respectively. The *boundary sample* of municipalities used in most estimations contains all jurisdictions with a distance to the ZRG border of less than 100 kilometers. This includes all municipalities in the treated region and about 68 percent of the municipalities in the five states west of the ZRG border. For the boundary sample at the district level, we limit the observations to jurisdictions

---

[17]Note that our results are robust to using 2013 data where available.

that are sufficiently close to the threshold determining transfer eligibility which will be described in detail below. This includes again all treated observations and about 50 percent of the districts outside the treated area and in the five states. Note that all our analyses are based on the 1971 district classification such that the number of districts remains constant over time.

## 4 Identification

Regional policy usually targets very specific groups of recipients. For instance, these can be regions lagging behind in terms of economic performance, cities being confronted with a high degree of poverty and emigration, or firms lacking private funds. Hence, public subsidies are not distributed randomly impeding a causal evaluation of such programs. This holds also true for the regional subsidies we analyze. Simple t-tests about the equivalence of the averages in the groups of transfer recipients and controls suggest significant differences for many variables across groups. For instance, income per square kilometer and population density are higher by about 10 and 27 percent in the group of non-subsidized municipalities than in the treatment group and these differences turn out significant at conventional levels. This points to the expectable selection issue and implies that an unconditional comparison may lead to false conclusions.

Yet, the transfer program we study gives rise to two types of discontinuities that generate quasi-random variation and are the basis of most of our econometric exercises. First, we examine observations in a close neighborhood on either side of the treatment border. Provided that other regional characteristics vary smoothly in space, a discontinuous jump in the outcomes of interest at the ZRG border can be attributed to the place-based policy. This approach is referred to as *Spatial Discontinuity Design* or Boundary Discontinuity Design.[18] Second, we exploit a discontinuity in the political rule that governed the treatment eligibility of regions and allows for local randomization of transfer recipience.

---

[18]Other applications of spatial RDD include Black (1999) focussing on school district boundaries to quantify the willingness to pay for a more educated neighborhood, Lalive (2008) identifying the effects of unemployment benefits on the duration of unemployment, Dell (2010) documenting the long-run impact of historical labor market institutions.

## 4.1 Spatial RDD

We denote by $Y_{i0}$ and $Y_{i1}$ the potential outcomes of a municipality $i$ in the situations with and without transfers, respectively. Our aim is to identify the effect of a transfer $T_i$ which corresponds to $\tau = Y_{i0} - Y_{i1}$. As counterfactual situations for individual units are unobservable, we aim at estimating an average treatment effect $E[\tau_i]$ for a group of comparable treated and control units. Our outset represents a special case of a two-dimensional RDD where the location of each municipality relative to the threshold is described by latitude and longitude, $\mathbf{L}_i = (L_{ix}, L_{iy})$. Similarly, the boundary between the treatment area $\mathcal{A}^+$ and the control area $\mathcal{A}^-$ consists of an infinite number of border points $\mathbf{b} = (b_x, b_y) \in \mathbf{B}$.

Due to the geographic nature of the policy measure, assignment to treatment is a discontinuous function of location, $T = 1\{\mathbf{L}_i \in \mathcal{A}^+\}$, where units east of $\mathbf{B}$ receive treatment while those to the west do not. In the spatial discontinuity design, location acts as the so-called forcing variable and we focus on the discontinuity of expected outcome at the geographical border:

$$\tau(\mathbf{b}) \equiv E[Y_{i1} - Y_{i0}|\mathbf{l} = \mathbf{b}] = \lim_{\mathbf{l}^+ \to \mathbf{b}} E[Y_i|\mathbf{L_i} = \mathbf{l}^+] - \lim_{\mathbf{l}^- \to \mathbf{b}} E[Y_i|\mathbf{L}_i = \mathbf{l}^-], \qquad (1)$$

where $\mathbf{l}^+ \in \mathcal{A}^+$ and $\mathbf{l}^- \in \mathcal{A}^-$ refer to locations in treated and control areas, respectively. Accordingly, $\tau(\mathbf{b})$ identifies the average treatment effect at the border point $\mathbf{b}$. In contrast to a one-dimensional regression discontinuity design, our approach yields a function of treatment effects evaluated at each border point $\mathbf{b} \in \mathbf{B}$. For most of our analysis we consider the average treatment effect along the whole border while we explore variations in the treatment effects across locations for sensitivity checks and to analyze the role of agglomeration externalities for persistence.[19]

The identification strategy of a regression discontinuity rests on two comparably weak assumptions (see Hahn, Todd, and van der Klaauw, 2001). First, counterfactual outcomes $E[Y_{i0}|\mathbf{L}_i]$ and $E[Y_{i1}|\mathbf{L}_i]$ have to be continuous at the border, that is all relevant variables besides treatment must change smoothly. Second, selective sorting at the border must be ruled out to ensure that treatment is "as good as" randomly assigned (Lee and Lemieux,

---

[19]See Papay, Willett, and Murnane (2011) for treatment effect heterogeneity in a two-dimensional RDD and non-geographic context. Importantly, this design allows us to limit the estimation to border segments where frictions at municipality borders are likely to be absent.

2010). Hence, municipalities must not be able to (precisely) manipulate their location relative to the treatment border. Since the treatment effect in the geographic discontinuity design is identified for units converging to the boundary, we also pursue the analysis using information on capital structures and luminosity that vary at a very fine spatial scale (e.g. grid cells of $100m \times 100m$) around the border.

The first assumption is fulfilled if the ZRG border was drawn randomly. However, there is reason to argue that administrative boundaries are usually not set at random, but follow some specific features such as rivers, mountains or cultural borders which may lead to discontinuities in other characteristics that matter for outcome. Common ways to address this issue include testing for discontinuities in relevant covariates (Dell, 2010) and removing border segments from the sample that seem to follow a problematic pattern (Black, 1999). While we pursue both robustness checks, we emphasize they are naturally limited in the sense that only a selection of covariates can be checked. Following this path, we thus cannot rule out a discontinuity in another relevant factor with certainty. We use two institutional features in our specific context to rebut these concerns. First, the ZRG border separates a set of 75 individual district pairs over a distance of 1,737 kilometers. These pairs may be divided according to historical routes, but there is no reason to expect that the ones in the treated area had systematically superior or inferior characteristics than the ones in the control area across *all* 75 pairs. Second, the district borders were modified substantially only a few years after the start of the ZRG-treatment whereas the ZRG border remained fully unchanged. Hence, the largest part of the ZRG border did not coincide with the relevant administrative district borders during the time we study.[20] To further improve confidence in our results, we will contrast the discontinuity at the threshold prior to the start of the program with the contemporaneous effects such that time invariant confounding discontinuities will cancel. Finally and most importantly, we will exploit the 40-kilometer rule that determined the actual treatment border, but did not coincide with any administrative or geographical boundary.

The second identifying assumption requires that districts or municipalities cannot (or only imprecisely) select themselves into treatment. In practice this means that municipalities in the control area must not be able to receive transfers by merging with municipalities located inside the originally-defined ZRG or influence the location of the border. As $\mathcal{A}^+$

---

[20]Roughly 57 percent of the 1,737km ZRG border ceased to represent a district border already between 1971 and 1978.

was never changed (despite changes in jurisdictional boundaries), this assumption is justi-fied.[21] Note, however, that individuals and firms may choose their place of residence and thus sort across the border. This potential change in the spatial equilibrium is what we are interested in as it is the consequence of treatment. As in Dell (2010), migration across treated and control regions is one of the channels we study.

We implement the spatial RDD both in a parametric and in a non-parametric way. In the former case we state the conditional expectations in (1) as $E[Y_{i0}|\mathbf{L}_i] = \alpha + f(\mathbf{L}_i) + g_0(D_i)$ and $E[Y_{i1}|\mathbf{L}_i] = \alpha + \tau + f(\mathbf{L_i}) + g_1(D_i)$ where $f(\mathbf{L_i})$ represents flexible polyno-mials of geographic location and $D_i$ refers to the shortest distance from $i$'s centroid to the treatment border ($\mathbf{B}$), i.e. the perpendicular to the closest border point. The inclu-sion of asymmetric distance control functions accounts for the possibility that proximity to the treatment border influences outcomes differently for transfer recipients and non-recipients.[22] Controlling for $\mathbf{L}_i$ may be important as units with the same distance to $\mathbf{B}$ may in fact be quite different if they are located in different parts of Germany (e.g. north versus south or distance to the sea, state/country borders). Thus, the regression model is given by:

$$Y_i = \alpha + g_0(D_i) + f(\mathbf{L_i}) + T_i[\tau + g_1(D_i) - g_0(D_i)] + \varepsilon_i. \tag{2}$$

Since $g_1(D_i) - g_0(D_i)$ converges to zero for observations close to the border, the average treatment effect is captured by $\hat{\tau}$. Since the credibility of the results rest on the correct specification of the control functions, we run alternative regressions with different func-tional forms (e.g. order of the polynomials), with and without coordinate control functions ($f(\mathbf{L_i})$), for different windows around the ZRG border, and we include border-segment fixed effects as well as state fixed effects.[23]

The assumptions about the form of the geographic control functions can be further relaxed by estimating the treatment effect in a non-parametric way. To do so, we employ local linear regressions and estimate the conditional expectations at the border as stated in (1). Notice that we base our estimates for $E[Y_{i0}|\mathbf{L}_i = \mathbf{b}]$ and $E[Y_{i1}|\mathbf{L}_i = \mathbf{b}]$ only on

---

[21]Municipalities that were located outside the ZRG and merged with municipalities in the treatment area could not become eligible for transfers. In such cases the treatment border passes through the munic-ipalities. The jurisdictional boundaries as of January 1, 1971 remained relevant for treatment throughout the duration of the program.

[22]By presuming symmetric functions on both sides of the RDD threshold, a kink may be misinterpreted as a discontinuity (see Lee and Lemieux, 2010).

[23]Limiting the sample to small windows around the threshold can substitute for including a higher order control function but requires a sufficient density of observations.

Table 2: DISTANCE & ASSIGNMENT VARIABLE

|  | ZRG | | | | Non-ZRG | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Mean | Std. | Min | Max | Mean | Std. | Min | Max |
| Distance to $\mathbf{B}$ ($D_i$) | 22,702 | 15,807 | 88 | 97,603 | 40,203 | 28,497 | 183 | 99,953 |
| $M_d$ | 20.337 | 11.322 | 3 | 45 | 88.021 | 29.664 | 42 | 149 |

*Notes*: Distances are in meters and refer to municipality centroids. The assignment variable $M_d$ is defined as the minimum distance (in km) from the Iron Curtain that includes the majority share of the district area. It is determined at the district level according to the 1971 classification. Each municipality is uniquely assigned to a district. Three districts received treatment although not being eligible according to the rule and thus generate fuzzyness. Of those districts being eligible according to the treatment rule all received treatment. We dropped all observations with a distance of more than 150km to the ZRG border and districts with $M_d > 150$.

observations in $\mathcal{A}^+$ and $\mathcal{A}^-$, respectively. As in the parametric approach, we condition on the forcing variable $D_{i\mathbf{b}}$ and estimate univariate local linear regressions for a set of 20 border points $\mathbf{b}^1, ..., \mathbf{b}^{20}$ which are allocated at equal distances along the border.[24] The corresponding results crucially depend on the choice of bandwidth. We derive the optimal bandwidth $h^*$ according to the criterion suggested by Imbens and Kalyanaraman (2012) and use a triangular kernel (see Fan and Gijbels, 1996, and Imbens and Lemieux, 2008).[25] Table 2 reports descriptive statistics on the distance of observations from $\mathbf{B}$. Although the treated area corresponds mostly to a narrow band of 40 kilometers there are treated observations in the north-east (in particular on the island Fehmarn) located at a distance of up to 100 kilometers from the ZRG border. The closest municipal centroids lie at about 88 and 183 meters from $\mathbf{B}$ for the treatment and control groups, respectively.[26] Due to the nature of the transfer program the distance to the ZRG border is positively correlated with the distances to the Iron Curtain. However, as the location of the ZRG border is determined by the districts' shape, size, and location the correlation between distance to $\mathbf{B}$

---

[24]We check the sensitivity of our results with 10 and 30 border points. As an alternative approach we followed Papay, Willett, and Murnane (2011) using a bivariate non-parametric regression with the arguments $L_{ix}$ and $L_{iy}$. Due to the "curse of dimensionality" bivariate local linear regressions require a much higher density of data. For this reason we favor the univariate non-parametric approach. However, all our results are robust to the bivariate non-parametric regression approach. See Appendix, Figure B1 for a more detailed description of the non-parametric specifications.

[25]Alternatively, we use cross-validation procedures and vary the bandwidth manually.

[26]As an alternative to the centroids' distances from $\mathbf{B}$ – which can be very small with narrow municipalities – we approximate the location of a municipality by the average over a sufficiently large number of grid cells within the municipal boundaries. All our results are robust to this alternative. In this case we split each municipality into $100m \times 100m$ grids, determine longitude, latitude as well as distances from $\mathbf{B}$ for each grid cell and take the municipal averages across grid cells to obtain $g(D)$ and $f(\mathbf{L})$.

and distance to the Iron Curtain is only about 0.6 for the boundary sample and reduces to less than 0.05 when we limit the sample to a 20-kilometer window from $\mathbf{B}$. This points to an important advantage of our setting, namely the clear geographic criterion that defined the Zonenrandgebiet.

## 4.2 Classical RDD: Exploiting the political treatment rule

Recall that those districts that accommodated either 50 percent of its area or population within a band of 40 kilometers to the Iron Curtain at the beginning of 1971 became part of the ZRG. The blue-shaded area in Panel A of Figure 2 illustrates the 40-kilometer buffer. It is evident that the ZRG border roughly follows the buffer, but we observe pixel and municipalities at the same distance from the Iron Curtain featuring a different treatment status. The political rule allows us to generate an assignment variable, denoted by $M_d$, indicating a district's minimum distance from the Iron Curtain that includes the majority share of the district's area. Hence, this assignment criterion does not only depend on a municipality's distance from the Iron Curtain but also on the shape of the superordinate district it belongs to. At $M_0 = 40$, we should expect a discontinuity in the probability of receiving treatment which we can exploit as exogenous variation to identify the causal effect of transfers on economic outcomes. As the 40-kilometer buffer has no natural relevance and does not correspond to administrative borders, it is uncritical to presume that there are no discontinuities in other relevant factors at $M_0$.

We compute isodistance-curves from the Iron Curtain using GIS software as illustrated in Panel B of Figure 2. This allows us to compute the area share of each district for each distance to the Iron Curtain. Finally, we determine for each district the minimum distance buffer where the area share exceeds 50 percent. Table 2 reports descriptive statistics of $M_d$ for the treatment and control groups.[27] Apparently, none of the control observations was eligible for treatment and all exceptions belong to the treatment group. If these exemptions from the 40-kilometer rule were not too frequent, we should observe a jump

---

[27]An alternative translation of the treatment rule would be to compute the area share of a district within the 40-kilometer buffer $S_d$. We did this as a robustness check and find a pronounced discontinuity at $S_d = 0.5$ as suggested by the rule. Yet, this assignment variable has the drawback of clustering at $S_d = 0$ and accordingly is less powerful.

Figure 2: ASSIGNMENT VARIABLE

Panel A.                                 Panel B.
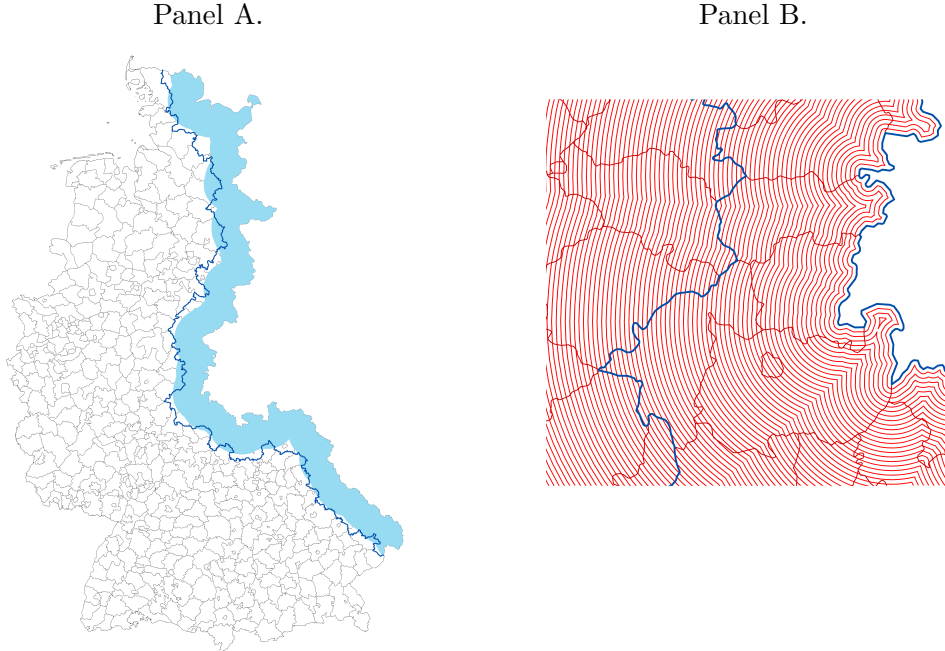


*Notes*: The above maps show district borders according to the 1971 classification. The light blue area in the left hand map marks the 40-kilometer distance from the Iron Curtain, the dark blue line refers to the ZRG border. The right hand map illustrates the buffer lines (in red) drawn in 1km intervals from the Iron Curtain.

in the probability of treatment at the threshold $M_0 = 40$:

$$P(T_{id}|\widetilde{M}_d) = \begin{cases} h_1(\widetilde{M}_d) & \text{if } \widetilde{M}_d \leq 0 \\ h_0(\widetilde{M}_d) & \text{if } \widetilde{M}_d > 0, \end{cases} \tag{3}$$

where $\widetilde{M}_d = M_d - M_0$ denotes the centered version of the assignment variable.

Figure 3 depicts the treatment indicator $T_{id}$ against the assignment variable $M_d$. The discontinuity at 40 kilometers is evident, but the design is fuzzy because a few districts with $M_d > M_0$ still receive ZRG treatment. Overall, non-compliance is not a big issue because only three districts were "mis-assigned". This is most likely driven by the second criterion of the political rule concerning population share, that is the non-compliers are those districts that did not accommodate 50 percent of the area within 40 kilometers to the eastern border, but 50 percent of the population.[28] Although we cannot account for

[28]We lack data about the population distribution within districts such that the second part of the rule may not be considered. Importantly, the rule requires only one of the criteria to be satisfied such that

Figure 3: TREATMENT PROBABILITY
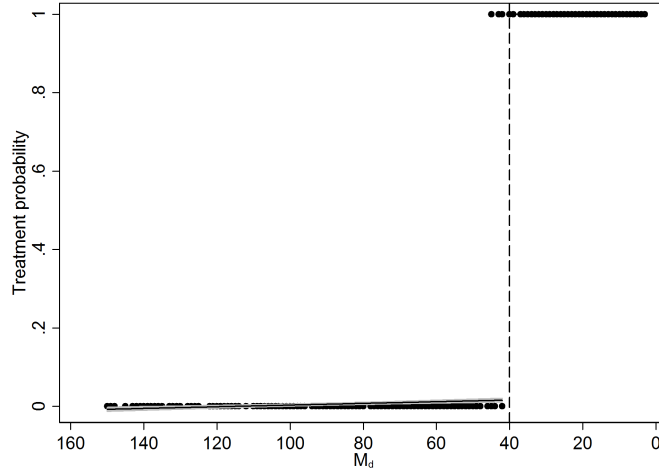


*Notes*: The assignment variable is measured at the district level. We consider only districts overlapping with a 150km buffer from the Iron Curtain. All districts further to the west are dropped from the sample.

this second criterion due to data limitations, we can obtain consistent estimators of the treatment effect by exploiting the discontinuity in the probability. The average treatment effect in this case is given by the ratio between the jump in the outcome and the jump in the treatment probability at $M_0$ (see Lee and Lemieux, 2010, for details).

We estimate the fuzzy RDD in a parametric as well as in a non-parametric fashion. In the latter approach we estimate the conditional expectations of outcome and treatment probability by means of local linear regressions separately for observations with $\widetilde{M_d} > 0$ and those with $\widetilde{M_d} \leqslant 0$. We employ an edge kernel and follow Imbens and Kalyanaraman (2012) in choosing an optimal bandwidth $h^*$ that minimizes the mean squared error of the average treatment effect.[29] The parametric approach follows a 2SLS approach where the

---

$M_d$ suffices as an assignment variable in the spirit of a fuzzy RDD. A precise measure of population distribution within districts was not even available at the time of treatment assignment and all but three districts (Schlüchtern, Einbeck, and Peine) were assigned strictly according to the first part of the rule. Hence, we may also drop those three districts and proceed in the spirit of a sharp RDD which yields almost identical results and even smaller standard errors.

[29]As noted by Imbens and Kalyanaraman (2012) this procedure often leads to bandwidth choices that are similar to those based on the optimal bandwidth for estimation of only the differences in expected outcomes (and applying the same bandwidth to the expectations of treatment probabilities). This holds also true in our case.

regression equations are given by:

$$Y_{id} = \alpha + f_0(\widetilde{M}_d) + T_{id}[\tau + f_1(\widetilde{M}_d) - f_0(\widetilde{M}_d)] + \varepsilon_{id}, \tag{4}$$

$$T_{id} = \gamma + h_0(\widetilde{M}_d) + R_d[\delta + h_1(\widetilde{M}_d) - h_0(\widetilde{M}_d)] + \nu_{id},$$

where $R_d = 1[M_d \leqslant M_0]$ indicates eligibility.[30] Since the political rule is applied at the district level $d$, we correct the estimated variance-covariance matrix for clustering at the level of districts and for heteroskedasticity of arbitrary form. We limit the sample to observations belonging to districts characterized by $M_d \leq 150$.

# 5 Effects on local economic activity

## 5.1 Income per km$^2$

Before turning to regressions, a graphical illustration of the data at the ZRG border is instructive. In Figure 4, we plot our main measure of economic activity (log income per km$^2$) for the years 1986 and 2010 as a function of distance to the ZRG border. In addition the plots include local mean of the outcome within equally-sized bins. The number of bins is chosen according to a data-driven method introduced by Calonico, Cattaneo, and Titiunik (2015). Panels in the left and right column use different windows and different control functions, but both reveal marked discontinuities at the ZRG border, both contemporaneously and persistently. Note that we assign positive and negative distances to the treatment and control region, respectively. As shortcoming of the graphical analysis is that by collapsing the two-dimensional location to a scalar measure of distance from the treatment border we cannot ensure that observations to the left and right of the threshold are de facto located in a short distance from each other.[31] We follow Calonico, Cattaneo, and Titiunik (2015) and apply an optimal data-driven choice of the number of equally sized bins. As a second observation, it seems that transfers have shifted economic activity from the western (non-treated) side of the ZRG border to the eastern (treated) side. We will examine this potential externality more closely in section 5.4 below. While such

---

[30]In what follows, we will generally use linear probability models in the first stage, but the results are very similar to those obtained with a nonlinear probability model in the first stage.

[31]As the ZRG border runs more or less straight from the south to the north this shortcoming of the graphical analysis can be mitigated by controlling for a linear trend of latitude (see earlier version of the manuscript Ehrlich and Seidel, 2015). However, this is not consistent with the data-driven bin choice.

Figure 4: DISCONTINUITIES IN ECONOMIC ACTIVITY

**Contemporaneous effect: 1986**



**Persistent effect: 2010**
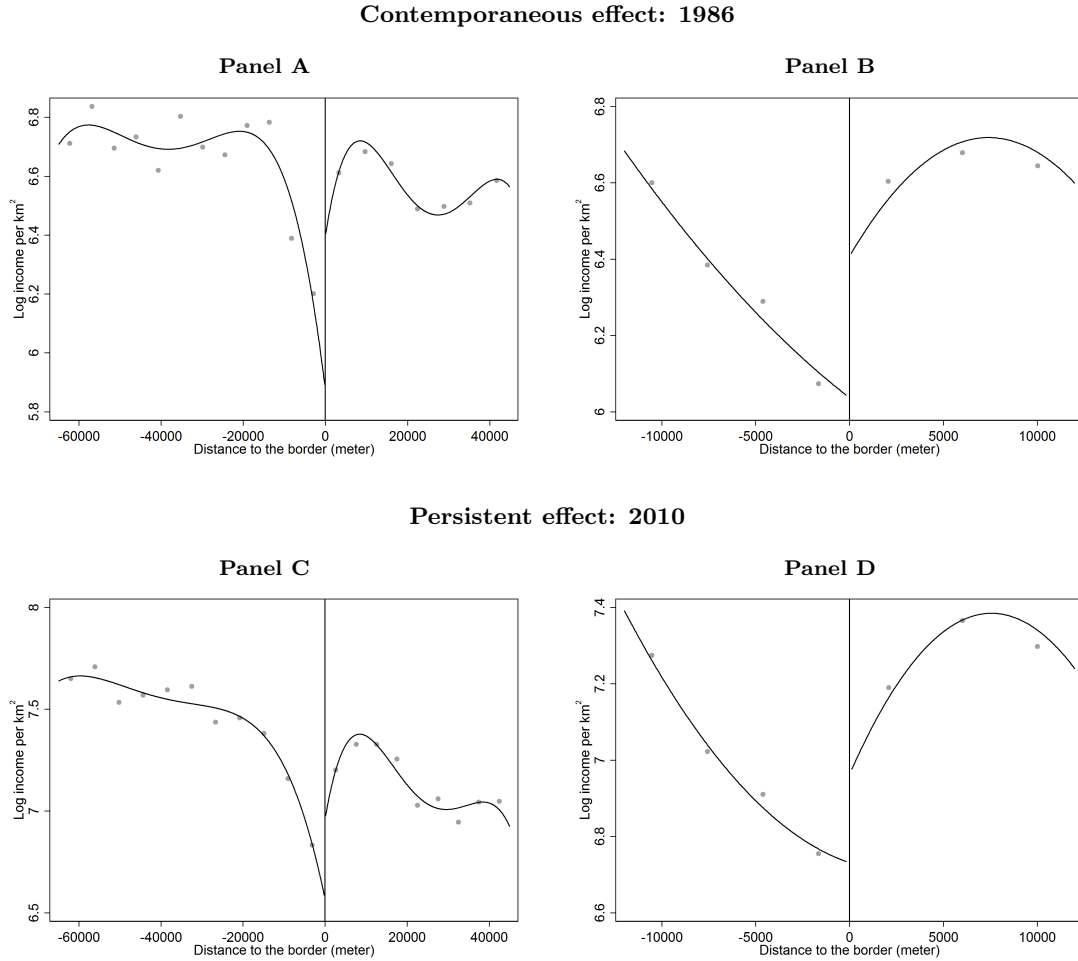


*Notes*: We run separate regressions on each side of the threshold. The plots represent local sample means using nonoverlapping evenly spaced bins on each side of the threshold following the data-driven method for optimal choice of the number of bins described in Calonico, Cattaneo, and Titiunik (2015). The lines represent a 4th-order polynomial distance control function for the 60km window (Panels A and C) and a quadratic control function for the 10km window (Panels B and D).

graphical analyses provide a transparent first assessment of whether a discontinuity exists, they provide only limited information about statistical significance and the magnitude of the effects. We thus turn to regression analysis.

Starting with the spatial RDD, Table 3 confirms the first impressions from the plots: Regional transfers to the Zonenrandgebiet exerted a strong and significant effect on eco-

Table 3: Spatial RDD: Income per km$^2$

| | Distance control | | Coordinate control | | Non-parametric | | |
|---|---|---|---|---|---|---|---|
| | 3rd | 5th | 2nd | 3rd | $h^*$ | $0.8 \times h^*$ | $1.2 \times h^*$ |
| **Contemporaneous effect** | | | | | | | |
| ATE | 0.484*** | 0.583*** | 0.296*** | 0.528*** | 0.311*** | 0.552*** | 0.239*** |
| | (0.099) | (0.147) | (0.079) | (0.099) | (0.079) | (0.099) | (0.069) |
| | [0.111] | [0.157] | [0.089] | [0.110] | - | - | - |
| Adj. R$^2$ | 0.16 | 0.17 | 0.14 | 0.16 | - | - | - |
| AIC | 10,750 | 10,732 | 10,869 | 10,741 | - | - | - |
| Obs. | 3,870 | 3,870 | 3,870 | 3,870 | 3,143 | 2,297 | 3,694 |
| **Persistent effect** | | | | | | | |
| ATE | 0.503*** | 0.542*** | 0.296*** | 0.535*** | 0.370*** | 0.518*** | 0.288*** |
| | (0.095) | (0.142) | (0.076) | (0.095) | (0.077) | (0.097) | (0.067) |
| | [0.108] | [0.154] | [0.086] | [0.107] | - | - | - |
| Adj. R$^2$ | 0.21 | 0.22 | 0.20 | 0.22 | - | - | - |
| AIC | 10,454 | 10,438 | 10,541 | 10,404 | - | - | - |
| Obs. | 3,881 | 3,881 | 3,881 | 3,881 | 3,095 | 2,203 | 3,652 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis, Conley (1999) standard errors in squared brackets. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(4) include state indicators, where (1) and (2) include segment fixed effects in addition. Columns (5)-(7) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012).

nomic activity (log income per km$^2$). We run three types of regressions. First, we include a distance control function using asymmetric 3rd- and 5th-order polynomials with segment and state fixed effects. We choose the polynomial orders on the basis of the AIC. Second, we directly control for the location of each municipality by including coordinates in addition to the Euclidean distance. Here we choose 2nd- and 3rd-order polynomials and add state fixed effects.[32] In each case we report robust standard errors as well as standard errors that correct for spatial dependence of unknown form using the method introduced by Conley (1999). Finally, we run non-parametric regressions where the optimal bandwidth $h^*$ is computed according to Imbens and Kalyanaraman (2012) and varied manually for sensitivity analysis in columns (6)-(7).

Among the parametric regressions both the adjusted $R^2$ and the AIC suggest that the

---

[32]The cubic polynomial of latitude and longitude is defined as $L_{ix} + L_{iy} + L_{ix}^2 + L_{iy}^2 + L_{ix}^3 + L_{iy}^3 + L_{ix}L_{iy} + L_{ix}^2 L_{iy} + L_{ix}L_{iy}^2$. Note that we choose lower order polynomials for $f(.)$ than for $g(.)$ because the bivariate control function requires more parameters to be estimated than the corresponding univariate control function.

Table 4: Fuzzy RDD: Income per km$^2$

| | Parametric $M_d$ | | Non-parametric | | |
| | 2nd | 3rd | $h^*$ | $0.8 \times h^*$ | $1.2 \times h^*$ |
|---|---|---|---|---|---|
| **Contemporaneous effect** | | | | | |
| ATE | 0.428** | 0.482** | 0.535*** | 0.476*** | 0.564*** |
| | (0.198) | (0.199) | (0.087) | (0.098) | (0.082) |
| Adj. R$^2$ | 0.077 | 0.083 | - | - | - |
| AIC | 11,110 | 11,088 | - | - | - |
| Obs. | 3,875 | 3,875 | 2,143 | 1,617 | 2,581 |
| **Persistent effect** | | | | | |
| ATE | 0.435** | 0.485** | 0.360*** | 0.255*** | 0.424*** |
| | (0.207) | (0.211) | (0.076) | (0.079) | (0.073) |
| Adj. R$^2$ | 0.134 | 0.139 | - | - | - |
| AIC | 10,793 | 10,773 | - | - | - |
| Obs. | 3,885 | 3,885 | 2,874 | 2,664 | 3,041 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors clustered at the district level in parenthesis. Observations with $M_d > 150$ are dropped from the sample. Columns (1) and (2) refer to fuzzy RDD specifications using a two-stage instrumental variables procedure and include state indicators. Note that the instrument is highly relevant in each of the first stages. Specifications (3)-(5) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012).

5th- and 3rd-order polynomials are preferred in case of the distance and coordinate control functions, respectively. However, the reduction in AIC is only marginal which indicates that there is no further gain to adding higher order terms. Coordinates capture location more precisely than distance from simple segment fixed effects such that we favor the specifications in columns (4) and (5). The latter refers to the non-parametric approach with optimal bandwidth $h^*$ which requires less restrictive functional form assumptions. We find that income per km$^2$ is predicted to be about 30-50 percent higher than in the counterfactual without regional subsidies in 1986, depending on the specification. Moreover, we can reject the zero for all specifications at a confidence level of 99 percent. The lower panel displays the corresponding specifications for the persistent effects of transfers measured in 2010. Notably, all specifications indicate again a positive and highly significant effect. Most importantly, these estimates remain remarkably similar for each type of specification in the two panels.

As we have argued before, we can identify causal effects of regional transfers under even weaker identifying assumptions by exploiting the discontinuity in the probability of receiving transfers at a distance of 40 kilometers from the ZRG-border. It can be virtually

ruled out that the 40-kilometer threshold mattered for economic outcomes in the absence of the ZRG program such that this approach is unaffected by potential confounding factors. However, it comes at the cost of lower efficiency as treatment assignment is carried out on the district level. Columns (1) and (2) in Table 4 show regressions that use 2nd- and 3rd-order polynomials of $M_d$ as control functions while columns (3)-(5) report non-parametric regression outcomes with the optimal bandwidth $h^*$ and manual adjustments. Standard errors are generally clustered on the level of districts and we obtain qualitatively identical results if we estimate the specifications on a sample collapsed by districts. Note that the non-parametric and contemporaneous estimate increases somewhat compared to the spatial RDD, but the overall picture shows very similar results when comparing the estimates to the corresponding coefficients in the spatial RDD in Table 3. This establishes confidence in the consistent estimation of the treatment effect. Notice that all specifications yield highly significant treatment effects at conventional levels.

Talking about economic magnitude, the effects might appear fairly high at first sight, but need to be qualified in at least two respects. First, the predicted average treatment effect in 1986 is the consequence of subsidies since 1971. As we have documented in section 2, transfers to the Zonenrandgebiet have been quite substantial every year. Second, it is quite plausible that these estimates include negative externalities of shifting activity from the control area to the treatment area, so these estimates must not be interpreted as *new* economic activity generated by the place-based policy. However, we argue that the estimates reflect the total causal effect of transfers into the Zonenrandgebiet on the spatial equilibrium. We provide a more detailed analysis of local relocation in section 5.4.

Although we have discussed identifying assumptions and their plausibility in this context in detail, a straightforward placebo test is to check whether there was a discontinuity in economic activity prior to treatment. Unfortunately, income data is unavailable at the municipality level before 1975, so we take GDP data at the more aggregated district level. Using estimates for 1961, it is immediate from Table 5 that none of the specifications reveal higher economic activity in the Zonenrandgebiet that was established only ten years later. The point estimates are positive in the parametric and negative in the non-parametric specifications, but all of the estimates are far from being statistically significant. Further, we use pre-treatment information about population density which is available at the municipality level and confirms that there are no discontinuities at the ZRG border prior to the start of the program (see Figure 5 below).

21

Table 5: PRE-TREATMENT – 1961

| Income per km$^2$ | Parametric $M_d$ | | Non-parametric | | |
|---|---|---|---|---|---|
| | 2nd | 3rd | $h^*$ | $0.8 \times h^*$ | $1.2 \times h^*$ |
| ATE | 0.023 | 0.008 | -0.186 | -0.297 | -0.189 |
| | (0.440) | (0.443) | (0.601) | (0.737) | (0.541) |
| Adj. R$^2$ | 0.350 | 0.352 | - | - | - |
| AIC | 1,040 | 1,041 | - | - | - |
| Obs. | 309 | 309 | 176 | 141 | 193 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. Regressions are based on district level data. Observations with $M_d > 150$ are dropped from the sample. Columns (1) and (2) refer to fuzzy RDD specifications using a two-stage instrumental variables procedure and include state indicators. Note that the instrument is highly relevant in each of the first stages. Specifications (3)-(5) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012).

## 5.2 Economic channels

What are the underlying channels of higher economic activity in the Zonenrandgebiet? Most obviously, as transfers were primarily targeted to subsidize firm investments and public infrastructure, we employ data on private and public capital to explore whether discontinuities prevail. The German Statistical Office provides detailed information about municipalities' land coverage which we use as capital proxies. *public capital* measures the area share of a municipality covered by public infrastructure like streets, railway tracks, airports, seaports, public squares, or public buildings. Similarly, *private capital* represents the area share of a municipality covered by industrial parks, commercial buildings and residential homes. Further focusing on commercial capital (*industrial capital*) allows for insights into the relative importance of business activity versus residences. We also use the *business tax base* as an alternative proxy for the private capital stock.[33] A second reason for higher economic activity per square kilometer could be changes in population and employment. Investment subsidies may also raise labor demand and labor productivity (arguably through higher capital stock) affecting the migration decision of households. Furthermore, the ZRG program also supported renovation of private homes, social housing, and cultural activities rendering living in the treatment area more appealing. Finally, we explore whether the human capital of the workforce differs systematically between the treatment and the control area. We proxy human capital by the share of residents with a

---

[33]Note that the business tax base is defined homogeneously across all municipalities in Germany. *private capital* is the sum of *residential capital* and *industrial capital*.

tertiary degree.

For the sake of brevity, we show results from the spatial RDD stressing that the findings

Table 6: CHANNELS I: CAPITAL

| | Contemporaneous effects | | | Persistent effects | | |
|---|---|---|---|---|---|---|
| | Coordinate control | | Nonparametric | Coordinate control | | Nonparametric |
| | 2nd | 3rd | $h^*$ | 2nd | 3rd | $h^*$ |
| **Business tax base per km$^2$** | | | | | | |
| ATE | 0.366*** | 0.720*** | 0.652*** | 0.463*** | 0.848*** | 0.800*** |
| | (0.124) | (0.157) | (0.148) | (0.114) | (0.144) | (0.142) |
| Adj. R$^2$ | 0.17 | 0.19 | - | 0.18 | 0.20 | - |
| AIC | 12,795 | 12,718 | - | 13,244 | 13,161 | - |
| Obs. | 3,533 | 3,533 | 2,318 | 3,792 | 3,792 | 2,299 |
| **Private capital stock** | | | | | | |
| ATE | 0.197*** | 0.341*** | 0.291*** | 0.193*** | 0.298*** | 0.278*** |
| | (0.062) | (0.078) | (0.070) | (0.058) | (0.073) | (0.066) |
| Adj. R$^2$ | 0.11 | 0.13 | - | 0.07 | 0.08 | - |
| AIC | 8,895 | 8,830 | - | 8,420 | 8,369 | - |
| Obs. | 3,845 | 3,845 | 2,730 | 3,851 | 3,851 | 2,839 |
| **Industrial capital stock** | | | | | | |
| ATE | 0.415** | 0.666*** | 0.525*** | 0.344* | 0.407* | 0.476*** |
| | (0.191) | (0.248) | (0.182) | (0.182) | (0.236) | (0.175) |
| Adj. R$^2$ | 0.11 | 0.13 | - | 0.11 | 0.12 | - |
| AIC | 3,875 | 3,860 | - | 3,668 | 3,656 | - |
| Obs. | 1,259 | 1,259 | 859 | 1,234 | 1,234 | 822 |
| **Public capital stock** | | | | | | |
| ATE | 0.147*** | 0.111*** | 0.153*** | 0.172*** | 0.138*** | 0.207*** |
| | (0.032) | (0.040) | (0.039) | (0.032) | (0.040) | (0.040) |
| Adj. R$^2$ | 0.27 | 0.27 | - | 0.25 | 0.26 | - |
| AIC | 3,885 | 3,851 | - | 3,760 | 3,718 | - |
| Obs. | 3,855 | 3,855 | 2,433 | 3,865 | 3,865 | 2,312 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(2) and (4)-(5) refer to parametric specifications and include state indicators. Columns (3) and (6) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012). *Business tax base per km$^2$* is measured in logarithmic terms. The three measures of capital stocks are bounded between zero and unity and which renders estimating linear models inappropriate. Thus we apply a logit transformation to *public capital*, *private capital* and *industrial capital*. Note that data on *industrial capital* is available for 1988 in only three states. Therefore, we restrict the contemporaneous and persistent estimates on these states. Using data on all states in 2010 yields similar results.

are robust to using fuzzy RDD.[34] Table 6 summarizes contemporaneous and persistent effects of the transfer program on capital. We only report 2nd- and 3rd- order polynomials of the augmented coordinate control specifications and non-parametric regressions based on the optimal bandwidth $h^*$. The estimates suggest that the ZRG treatment has led to a markedly higher stock of both private and public capital. For example, the business tax base per square kilometer is predicted to be around 60-70 percent higher in 1986. Looking at persistence in 2010, we still find a highly significant effect at an even higher level of around 80 percent. Taking the area share covered by plants and residential structures, our estimates suggest that transfers have raised the capital stock by about 30 percent both in 1984 and 2010. Distinguishing between industrial and residential structures, we observe that ZRG treatment led to a higher increase in industrial premises as the effect on *industrial capital* turns out higher than the effect on aggregate *private capital*. The public capital stock is predicted to be about 10-20 percent higher compared to the counterfactual.

Turning to labor as a potential channel for higher economic activity and using the same specifications as above, Table 7 reveals that population density was raised by about 40 percent with no indication of a decline in the long term. Econometrically speaking, commuting is costless at the ZRG-border so the change in population can only be attributed to subsidies for social housing and renovation of private residences. The discontinuity in employment per square kilometer is even more pronounced at about 60-70 percent indicating substantial commuting into the Zonenrandgebiet.[35] However, we find no evidence that the composition of the workforce with respect to skills was affected by treatment. The share of high-skilled employees in the Zonenrandgebiet does not differ from the counterfactual scenario without transfers.

It is informative to take a closer look at how the magnitude of effects has developed over time. As we have argued in the previous subsection, GDP is only available at the district level and at fewer intervals than population data. Since we have found significant and large effects of ZRG transfers on population density, we run the specification with coordinate control functions for a number of years between 1961 and 2010. Figure 5 reveals differences in population densities between the Zonenrandgebiet and the control

---

[34]Results from the fuzzy RDD can be obtained from the authors upon request.

[35]Monte, Redding, and Rossi-Hansberg (2015) show that differences in commuting intensities yield substantial heterogeneity of the local employment elasticity to productivity shocks. By analyzing the effect on variables measured at place of residence (income, population) and place of work (business tax base, employment) we can infer the effect of transfers on commuting intensities.

Table 7: Channels II: Labor

| | Contemporaneous effects | | | Persistent effects | | |
|---|---|---|---|---|---|---|
| | Coordinate control | | Nonparametric | Coordinate control | | Nonparametric |
| | 2nd | 3rd | $h^*$ | 2nd | 3rd | $h^*$ |

**Population per km$^2$**

| | | | | | | |
|---|---|---|---|---|---|---|
| ATE | 0.239*** | 0.434*** | 0.372*** | 0.290*** | 0.473*** | 0.425*** |
| | (0.069) | (0.087) | (0.077) | (0.071) | (0.089) | (0.079) |
| Adj. R$^2$ | 0.19 | 0.21 | - | 0.18 | 0.21 | - |
| AIC | 9,846 | 9,746 | - | 9,988 | 9,876 | - |
| Obs. | 3,870 | 3,870 | 2,745 | 3,881 | 3,881 | 2,717 |

**Employment per km$^2$**

| | | | | | | |
|---|---|---|---|---|---|---|
| ATE | 0.418*** | 0.658*** | 0.692*** | 0.467*** | 0.723*** | 0.741*** |
| | (0.108) | (0.137) | (0.124) | (0.110) | (0.140) | (0.133) |
| Adj. R$^2$ | 0.18 | 0.20 | - | 0.16 | 0.17 | - |
| AIC | 13,120 | 13,052 | - | 12,407 | 12,346 | - |
| Obs. | 3,826 | 3,826 | 2,601 | 3,665 | 3,665 | 2,269 |

**Human capital**

| | | | | | | |
|---|---|---|---|---|---|---|
| ATE | 0.016 | 0.213* | 0.113 | -0.076 | 0.116 | -0.054 |
| | (0.082) | (0.109) | (0.079) | (0.071) | (0.092) | (0.074) |
| R$^2$ | 0.13 | 0.14 | - | 0.12 | 0.14 | - |
| AIC | 3,555 | 3,541 | - | 5,372 | 5,337 | - |
| Obs. | 1,782 | 1,782 | 1,373 | 2,576 | 2,576 | 1,886 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(2) and (4)-(5) refer to parametric specifications and include state indicators. Columns (3) and (6) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012). *Population per km$^2$*, *Employment per km$^2$*, and *Income per capita* are measured in logarithmic terms. *Human capital* is bounded between zero and unity and which renders estimating linear models inappropriate, thus we apply a logit transformation.

area. Note that the points and bars illustrate the point estimates and confidence bands of parametric specifications according to (2). These outcomes shed light on how transfers unfolded their effects over time. It is immediate that the difference in population density developed fairly quickly over the first five to ten years after the introduction of the transfer program and reached a steady level until the end of the 1980s. We observe no evidence for a decline of this difference after the program was stopped. Further there is no significant difference in population density prior to treatment. This finding is in line with the insights from Table 5 that there was no discontinuity in GDP per square kilometer in 1961.

Figure 5: Dynamics in population density
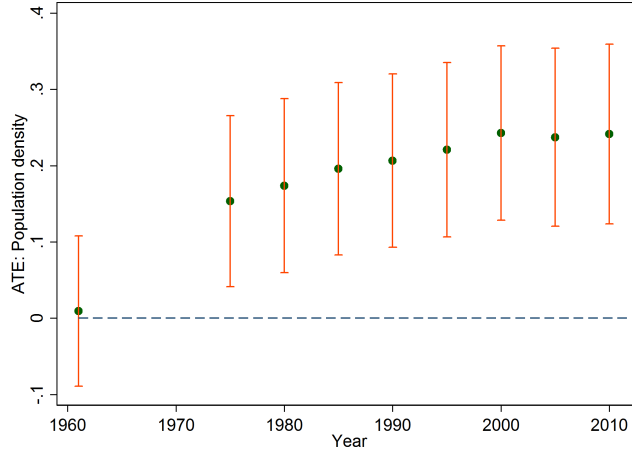
## 5.3 What explains persistence?

To guide ideas, we build on a very simple, but general framework. Consider two initially symmetric regions accommodating locations of unit measure $x$ between the boundaries $-\bar{x}$ and $\bar{x}$.[36] The common (treatment) border is referred to by $x = 0$. One region with locations $x \in [0; \bar{x}]$ receives treatment $T$ (referred to by $+$) while the other region with locations $x \in [-\bar{x}; 0]$ does not (referred to by $-$). A region can be understood as an area composed of municipalities or as a municipality composed of small grid cells – depending on the level of aggregation in our empirical analysis. We define economic activity as output in location $x$ as

$$Q(x) = A(x, I)I(x, A, T). \tag{5}$$

Output is increasing in physical inputs $I(\cdot)$ (for example, labor, capital, public goods or a combination thereof) and in a production amenity $A(\cdot)$. We assume the production amenity to be composed of two parts such that $A(x, I) = A_{ext}(x, I)A_{own}(x, I)$. While the

---

[36]Note that in the empirical analysis we weight outcomes by area and consider equally distributed grid cells.

first part dissipates in space, the second component is confined to the respective location. To ensure that there is always some economic activity in both regions, we impose that $A(\cdot)$ is hump-shaped in $I$ capturing the idea that dispersion forces dominate agglomeration forces at a certain level of economic density or economic frictions (like trade costs). For example, $A_{ext}(\cdot)$ can be understood as agglomeration economies stemming from knowledge spillovers, labor market pooling, supply linkages or home-market effects, among others (Marshall, 1920, and Duranton and Puga, 2004). $A_{own}(\cdot)$, in contrast, captures all effects on productivity that occur only under the condition of being located in that location (e.g. better public facilities, capital structures, tax benefits, etc.). Our specification of economic activity allows for hysteresis as inputs depend on production amenities and vice versa (as e.g. in Krugman, 1991). With respect to the second part in (5), we assume that (i) treatment $T$ raises inputs and (ii) more productive locations have an incentive to invest more, so $\partial I(\cdot)/\partial A(\cdot) > 0$.[37]

To make both the contemporaneous and persistent effects of treatment transparent, we evaluate the model at three distinct points in time: 1. *before* treatment, 2. *during* treatment, and 3. *after* treatment.

**1. Before treatment.** As both regions are initially identical and $T^{+} = T^{-} = 0$, it is immediate that input and production amenity levels are identical. Relating (5) to our empirical specification by taking logs and comparing economic activity at the treatment border $x = 0$ implies: $lnQ(0^{+}) - lnQ(0^{-}) = 0$, where $x = 0^{+}$ and $x = 0^{-}$ refer to locations converging to $x = 0$ from the treatment and the control side, respectively.

**2. During treatment.** Now suppose that one region receives treatment $T^{+} > 0$ while $T^{-} = 0$ (e.g. capital subsidies or provision of public infrastructure). This leads to an increase in $I^{+}$ in a first step while $I^{-}$ remains unchanged. We refer to this as the *direct input effect* of treatment. In a second step, the difference in $I(\cdot)$ raises the production amenity stimulating more input investments in consecutive steps until $\partial A(\cdot)/\partial I(\cdot) = 0$. We call this the *self-reinforcing effect*. Importantly, however, externalities do not stop at local borders so locations close to $x = 0$ benefit from input intensities on both sides of the threshold. We follow recent work by Turner, Haughwout, and van der Klaauw (2014) by assuming that $A_{ext}(x)$ is a weighted average of input intensity of neighboring locations

---

[37]This feature is in line with heterogeneous firms models, e.g. Melitz (2003) or Melitz and Ottaviano (2008).

according to

$$A_{ext}(x) = \frac{1 - \gamma(x)}{2} \int_{-\bar{x}}^{0^-} I(x)dx + \frac{1 + \gamma(x)}{2} \int_{0^+}^{\bar{x}} I(x)dx. \tag{6}$$

We define $\gamma(x)$ to be a weakly increasing continuous function with $\gamma(x) = -1$ if $x \leq -x' \ \wedge \ -x' > -\bar{x}$, $\gamma(0^-) = \gamma(0^+) = 0$ and $\gamma(x) = 1$ if $x \geq x' \ \wedge \ x' < \bar{x}$. Hence, according to (6) locations $x = 0^-$ and $x = 0^+$ experience the same level of externality, that is $A_{ext}(0^+) = A_{ext}(0^-)$. With respect to the location-specific production amenity, we observe a discontinuity at $x = 0$ due to differences in input intensities at this location. Using these insights and taking logs of (5), we obtain at $x = 0$ a discontinuity

$$\Delta Q(0) = \Delta I(0) + \Delta A_{own}(0),$$

where $\Delta$ refers to the difference in the logs of the respective variables. Importantly, the observed discontinuity does not contain $A_{ext}$. Note that agglomeration externalities do raise economic activity, but it is in the nature of spatial RD identification that all continuous effects are not part of the discontinuity. Intuitively, self-reinforcing externalities exert the same incentive for investment at $x = 0^+$ and $x = 0^-$. Figure 6 provides a graphical illustration of this argument.

**3. After treatment.** After the end of the program, locations in the former treatment region no longer receive transfers and $T^+ = T^- = 0$. Hence, for the gap in economic activity to persist, our framework suggests two explanations: either inputs $I$ stay at the same level even in the absence of treatment (e.g. capital does not depreciate) or location-specific productivity $A_{own}$ remains at a higher level in the former treatment area persistently.

While some type of inputs like public infrastructure generate long-run effects, they depreciate as well and we should at least observe a decline in the discontinuity over time. This is especially true if we consider investments that were undertaken at the beginning of the treatment period in the early 1970s. Recall from Figure 5 that the average treatment effect reached its persistent level only within a few years. Moreover, our results do not provide any evidence for a decline in employment or capital stocks. We thus favor the interpretation that the policy has generated a location-specific productivity advantage. This advantage could stem from better governance with regard to managing commercial

Figure 6: Direct effect versus productivity effect
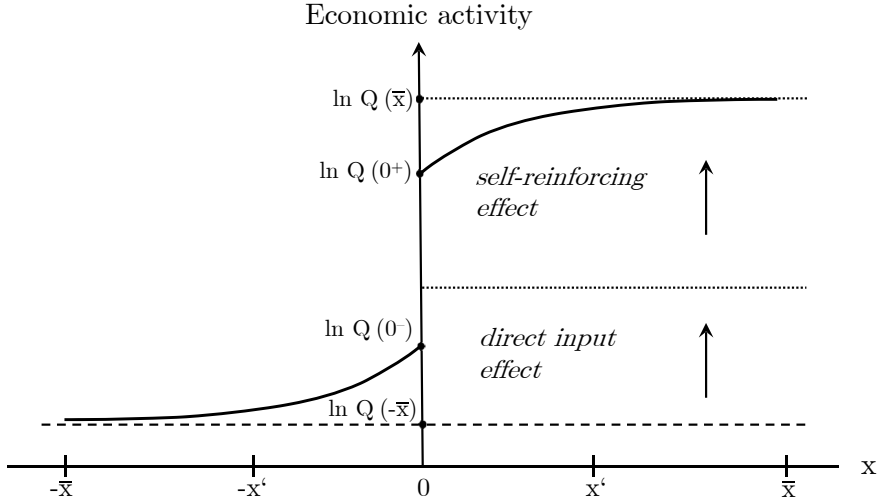


*Notes*: The dashed line refers to the situation without transfers. The solid line refers to the output with transfers and the dotted lines illustrate the "direct input effect" and the "productivity effect". The area left of $x = 0$ belongs to the control group whereas observations right of $x = 0$ receive transfers.

activities that municipalities have developed during the treatment period. Alternatively, we have shown that the policy has raised economic activity and thus the tax base. With higher tax revenues, municipalities have more means to maintain better local infrastructure. Moreover, it might be easier to attract firms due to the sunk factors associated with land-use planning, so "commercial and residential structures may serve as focal points" (Redding, 2010). Further, adjustment costs may prevent firms from relocating after the end of the policy. With fixed costs and imperfect divisibility of investments it may be more profitable for firms to reinvest in existing structures instead of building new facilities on the green field. While the latter argument does not apply for entry of new firms the expansion of an existing commercial zone might generally be less expensive than the development of a new zone.

**Continuity of externalities.** The conclusion that *agglomeration externalities* cancel in the RD design hinges on the assumption that $\gamma(x)$ is a continuous function. So far, we have built our empirical analysis on administrative data at the municipality level such that outcome variables were assigned to the municipalities' centroids. A natural concern would be that there is little or no economic activity at the boundaries of jurisdictions introducing frictions in the diffusion of externalities around the treatment border. In our

29

simple theoretical framework this would imply $\gamma(0^-) = -\varepsilon$ and $\gamma(0^+) = \varepsilon$ with $0 < \varepsilon \leq 1$. If $\varepsilon = 1$, any spillover at the treatment border is ruled out.

The assumption that $\gamma(x)$ is a continuous function at $x = 0$ becomes more credible if we were able to use more disaggregated data around the treatment boundary. In that case, the distance between $x = 0^+$ and $x = 0^-$ effectively converges to zero. Two sources of satellite data prove useful in this context. First, we use information on capital structures at a grid cell level of $100m \times 100m$ provided by the European Environmental Agency's CORINE project. The satellite data contain information on numerous different land cover classes and we set $PrivateCapital = 1$ if a location is covered by private capital structures.[38] Note that the area-weighted sum of $PrivateCapital$ is highly correlated with our municipality-level variable for private capital (correlation coefficient of 0.84). Second, we exploit night-light radiance as a proxy for local GDP (see Henderson, Storeygard, and Weil, 2012) at grid cells of $30 \times 30$ arcseconds ($926m \times 926m$ at the equator).[39]

It is evident from Table 8 that both data sources are associated with positive and highly significant average treatment effects. Note that the estimates for $Prob(PrivateCapital = 1)$ on the grid-cell level are well in line with the estimates for the area share of a municipality covered by private capital as displayed in Table 6. The latter indicated that the logit-transformed area share of private capital, i.e. the odds ratio, increased by 19-34 percent due to transfers. Given that $Prob(PrivateCapital = 1)$ is about 6 percent in our data, a 1.6 percentage points increase in $Prob(PrivateCapital = 1)$ as displayed in columns (1) and (2) of Table 8 corresponds to $ln(\frac{0.076}{1-0.076}\frac{1-0.06}{0.06}) \approx 0.253$, i.e. an increase of about 25.3 percent in the odds ratio. To compare the finding relating to radiance, we first compute a conversion factor of income per $km^2$ and radiance per $km^2$ of 1.045 using West German municipality data. Thus, the radiance estimates suggest an increase in GDP per $km^2$ of 21-31 percent. Most importantly, the effects estimated from grid-cell data show again a high degree of persistence. Hence, the analysis of grid-cell data suggests that there must be something beyond agglomeration externalities that drives persistence.

We pursue an additional exercise to further examine the potential role of frictions at

---

[38]In particular, $PrivateCapital = 1$ if a place is covered by 'Continuous urban fabric', 'Discontinuous urban fabric', 'Industrial or commercial units', or 'Construction sites' and zero otherwise.

[39]This data is provided by the Defence Meteorological Satellites Program – Operational Linescan System (DMSP-OLS). The satellite data report digital integer numbers ranging from 0 to 63. These may be converted to radiance as a measure of night luminosity by using the formula $radiance = digitalnumber^{1.5}$ for a spatial unit which is denoted in terms of Watts/cm$^2$/sr/nm (in words: Watts per squared centimeter per steradian per nanometer of wave length).

Table 8: SPATIAL RDD: GRID CELL DATA

| | Contemporaneous effects | | | Persistent effects | | |
|---|---|---|---|---|---|---|
| | Coordinate control | | Nonparametric | Coordinate control | | Nonparametric |
| | 2nd | 3rd | $h^*$ | 2nd | 3rd | $h^*$ |
| **Prob(PrivateCapital = 1)** | | | | | | |
| ATE | 0.016*** | 0.016*** | 0.013*** | 0.017*** | 0.019*** | 0.031*** |
| | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) |
| Adj. $R^2$ | 0.01 | 0.01 | - | 0.01 | 0.02 | - |
| AIC | -210,193 | -218,436 | - | 1,214,637 | 1,201,129 | - |
| Obs. | 7,786,402 | 7,786,402 | 1,138,278 | 7,786,402 | 7,786,402 | 1,661,047 |
| **Log(Radiance)** | | | | | | |
| ATE | 0.298*** | 0.281*** | 0.252*** | 0.285*** | 0.206*** | 0.198*** |
| | (0.015) | (0.021) | (0.019) | (0.013) | (0.018) | (0.023) |
| Adj. $R^2$ | 0.08 | 0.10 | - | 0.09 | 0.12 | - |
| AIC | 256,617 | 254,503 | - | 290,388 | 286,205 | - |
| Obs. | 107,776 | 107,776 | 49,304 | 125,527 | 125,527 | 24,196 |

*Notes:* ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. We drop all observations outside a 40km window of the ZRG border. The dependent variable in columns (1)-(3) is a binary indicator which is unity if a $100m \times 100m$ grid is covered by private buildings and zero otherwise. In columns (4)-(6) we use log(Radiance) as computed from satellite night-light data for grid cells of 30 arc-seconds (about $926m \times 926m$ at the equator). Information on land coverage and radiance refer to the years 1990 and 1992, respectively. Columns (3) and (6) refer to non-parametric specifications where the bandwidth $h^*$ is computed according the algorithm introduced by Imbens and Kalyanaraman (2012) and standard errors are computed according to Imbens and Lemieux (2008).

the treatment border. The idea is to exploit heterogeneity along the ZRG border and to estimate the treatment effects for radiance and private capital structure based on grid cells that belong to municipalities where potential frictions are presumably less pronounced. First, we restrict our sample to the 10 percent of municipalities that feature the highest (area-weighted) number of roads crossing the treatment border. Second, we compute the amount of undeveloped land within a 1km buffer on both sides of the treatment border. Using this information we restrict the sample to the 10 percent of municipalities featuring the lowest share of undeveloped land in the neighborhood of the boundary. Third, suspecting that unobservable social networks may display frictions at the treatment border, we estimate the effects only for polycentric municipalities arguing that ties within such scattered municipalities are weaker (relative to cross-border networks) than in monocentric municipalities. If frictions play a role, we would expect lower estimates in these exercises compared to those displayed in Table 8. As shown in Appendix D, this is not

unambiguously the case. While some point estimates go up, others are somewhat lower than those in the full sample. Importantly, these estimates based on small subsamples are not different from the corresponding ones in Table 8 at conventional significance levels. In conclusion, it is impossible to entirely rule out agglomeration economies as an explanation for the estimated discontinuities on the basis of observable variation along the border, but the evidence clearly points towards a further important channel to explain spatial persistence.

**Interaction effects.** Finally, the persistence in economic activity could also be driven by an *interaction* between density and economic shocks (e.g. globalization, fiscal policy, labor market reforms etc.). In our simple framework $\frac{\partial Q/\partial A}{\partial I} > 0$ holds such that a common, positive productivity shock benefits formerly subsidized regions more than the control regions as long as input differences remain. To shed light on the importance of interaction effects we exploit two exogenous shocks that were arguably most important for the region, German reunification in 1990 and EU Enlargement in 2004.[40] To identify the effects, we combine the discontinuity approach with time variation to examine whether discontinuities differ *before* and *after* the events. Comparing municipalities in the close neighborhood of the ZRG-border ensures that municipalities are affected similarly by changes in market access. We focus on population density due to superior data coverage. Our differences-in-discontinuities specification can be directly derived from (2), so we have

$$Y_{it} = \alpha + \gamma_t + g_0(D_i) + f(\mathbf{L_i}) + T_i[\tau + g_1(D_i) - g_0(D_i) + \beta S_t] + f(\mathbf{L_i} S_t) + \varepsilon_{it}. \qquad (7)$$

Note that $T_i$ indicates whether a municipality is located in the Zonenrandgebiet and $S_t$ is a dummy variable equal to one after the shock has occurred (post-1990 and post-2004, respectively) and zero otherwise. Based on the insights from Redding and Sturm (2008) that the benefits of market access are declining in distance, we include a term that interacts the shock with location, $f(\mathbf{L_i} S_t)$. Note that accounting for regional subsidies raises the causal effect of market access as identified in Redding and Sturm (2008) by about 45 percent as shown in an earlier version of the manuscript (see Ehrlich and Seidel, 2015). Moreover, time fixed effects $\gamma_t$ absorb all common per-period effects of the shocks. We are interested in the contribution of the respective shocks for the overall discontinuity, that is $\beta/(\beta + \tau)$. While the main effects shown in Table 9 are very similar to the benchmark

---

[40]In a related difference-in-differences approach, Brülhart, Carrère, and Trionfetti (2012) study the employment response in Austrian municipalities due to the Fall of the Iron Curtain.

specifications in Table 7, it is immediate that the interaction effects of German reunification and EU enlargement with ZRG-treatment explain only between 2 and 6 percent of the total discontinuity.[41]

Table 9: REUNIFICATION & EU INTEGRATION

| Population per $km^2$ | Reunification | | EU integration | |
| --- | --- | --- | --- | --- |
| | 2nd. order | 3rd. order | 2nd. order | 3rd. order |
| ZRG | 0.240*** | 0.464*** | 0.305*** | 0.504*** |
| | (0.071) | (0.089) | (0.072) | (0.090) |
| S × ZRG | 0.016*** | 0.008 | 0.009* | 0.011* |
| | (0.005) | (0.007) | (0.005) | (0.006) |
| | includes coordinate controls | | | |
| Adj. $R^2$ | 0.196 | 0.223 | 0.181 | 0.209 |
| AIC | 18,969 | 18,704 | 18,973 | 18,714 |
| Obs. | 7,502 | 7,502 | 7,456 | 7,456 |

*Notes:* ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Standard errors clustered at municipality level in parenthesis. We drop all observations outside a 100km window of the ZRG border. *ZRG* indicates location in the Zonenrandgebiet, *S* is a shock that refers to *Reunification* in columns (1) and (2) and to *EU − Integration* in columns (3) and (4). In the former case the years refer to 1989 and 1995 whereas the years are 2003 and 2010 in the case of *EU − Integration*.

## 5.4 Local relocation

An aspect that we have so far ignored is local relocation at the treatment border. The above results do not allow inference about the net effect of the policy as our estimate of the average treatment effect is likely to contain a relocation externality that does show a discontinuity at the treatment border. Firms that are located in the Non-ZRG close to the treatment border are *not* indifferent between locating one meter to the left and one meter to the right of the threshold. We should thus expect that ZRG-treatment draws economic activity from the control area to the Zonenrandgebiet. Assuming that migration costs are increasing in distance, we should observe particularly strong relocation activities in a close neighborhood of the ZRG-border. Combining this effect with the net agglomeration externality/productivity effect in Figure 6 could result in a function that is represented by the solid curve in Figure 7. The way we have incorporated it implies that

---

[41]Further robustness checks, most importantly substitutive policies after the end of the program, are relegated to Appendix C.

Figure 7: Relocation
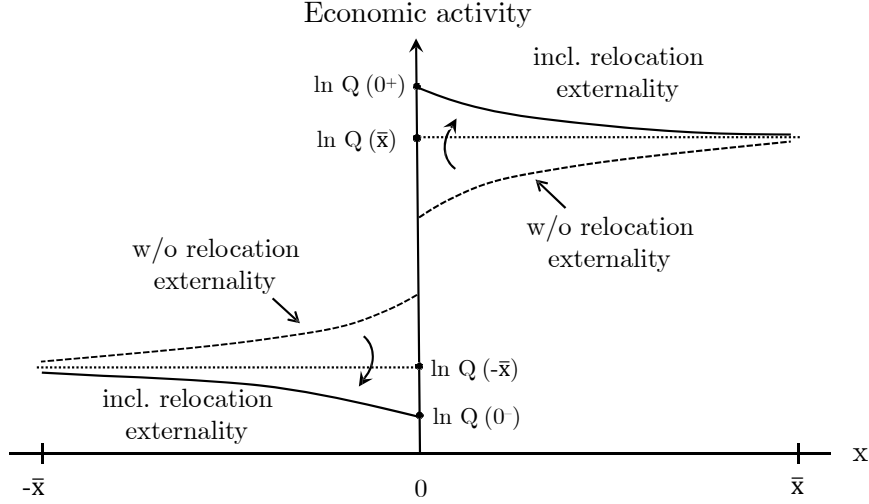
the relocation externality dominates, but we have no priors about the magnitude of this relocation externality. If this shifting of economic activity is relevant in size, our estimates of $lnQ(0^+) - lnQ(0^-)$ capture both the direct effect due to subsidies and the relocation externality due to further inputs moving from the control to the treated area. Note, however, that the continuous agglomeration externalities still cancel at the ZRG border and locational advantage has to play a key role for inputs to relocate and in particular to remain in the ZRG region once transfers were phased out.

In the following we aim to quantify the degree of local relocation included in our estimates. We apply a *spatial exclusion* approach that relies on estimating the effect for treatment and control observations that are located sufficiently distant from each other. The idea is that these municipalities are not affected by (local) externalities of the policy (see Neumark and Kolko, 2010). According to Figure 7 the aim is to estimate $lnQ(\bar{x}) - lnQ(-\bar{x})$. Obviously, this approach contradicts the identification strategy of the spatial RDD which relies on the comparison of outcomes for observations in a close neighborhood. However, we may execute this exercise in the fuzzy RDD. We exploit the fact that each district with $M_d \sim M_0$ accommodates a number of municipalities with

varying distances to the treatment border. Thus, by excluding municipalities in the close neighborhood of **B** we can remove the part of district outcome that is potentially contaminated by spillovers (see Figure E1 in the appendix for details). We estimate (4) for a sub-sample that excludes all municipalities bordering **B** and additionally establish minimum distances between treatment and control units of 10 kilometers and 20 kilometers in two alternative specifications. Yet, the requirement to maintain a high density of districts at the threshold $M_0$ limits the minimum distance we can establish between treatment and control observations. It may therefore be impossible to eliminate relocation externalities completely.[42]

The estimation results for the spatial exclusion approach are highlighted in Table 10. There are three specifications: (i) "border" excludes all municipalities adjacent to the treatment border, (ii) "5,000m" ignores all jurisdictions whose centroid is located within 5 kilometers to the ZRG border, and (iii) "10,000m" is a similar exercise for the 10-kilometer range. Note that all results are based on the contemporaneous sample.

Starting with the baseline results in columns (2) and (3) of Table 4, we observe from specification "border" that the average treatment effect drops from 0.482 to 0.449 and from 0.535 to 0.465 in the parametric and nonparametric specifications, respectively. Further restricting the sample to minimum distances of up to 10,000 meters yields a reduction of the point estimates of about one third compared to the benchmark results. Hence, assuming that the externality dissipates linearly implies that relocation activities must occur within 60km to explain the total effect. Note, however, that this would comprise a substantial part of West Germany as the mean distance to **B** across all non-treated municipalities (in all states) is about 113km. With a minimum distance of 10,000m we loose efficiency and the share of observations displaying a level of $M_d$ in the neighborhood of 40 drops considerably.[43] In summary, there is evidence for substantial relocation externalities. This is consistent with our finding that a shift of population towards the subsidized regions is among the key drivers of the aggregate effect on income per square kilometer. However, we may not completely determine the spatial extent of these negative externalities.

---

[42]An alternative would be to estimate the gradient of the solid line in Figure 7. A negative (positive) gradient in the treated (control) region would reveal that the relocation externality dominates and the size of the absolute magnitude provides information about the degree of relocation. We have implemented this alternative approach in an earlier version of the manuscript (see Ehrlich and Seidel, 2015) and show that the implied degree of relocations turns out qualitatively similar to the spatial exclusion approach.

[43]While $M_d \in [30, 50]$ holds for about 15 percent of the observations in our benchmark specification, this share drops to about 8 percent with the 10km-exclusion window.

Table 10: Relocation externalities

| Log income | 3rd. order polynomial of $M_d$ | | | Nonparametric $h^*$ | | |
|---|---|---|---|---|---|---|
| per km$^2$ | border | 5,000m | 10,000m | border | 5,000m | 10,000m |
| ATE | 0.449** | 0.396* | 0.336 | 0.465** | 0.435** | 0.358 |
| | (0.220) | (0.209) | (0.239) | (0.217) | (0.218) | (0.259) |
| Adj. R$^2$ | 0.09 | 0.09 | 0.10 | - | - | - |
| Obs. | 3,514 | 3,408 | 3,084 | 1,784 | 1,678 | 1,360 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. We exclude all municipalities that are bordering **B** (columns (1), (6)), and are within a distance of 5km and 10km from **B** (columns (2),(3) and (5),(6)). We estimate the fuzzy RDD specifications using a parametric 2SLS approach in columns (1)-(3) and the nonparametric approach in columns (4)-(6). Observations with $M_d > 150$ are dropped from the sample and standard errors are clustered at the district level.

# 6  Incidence: Who benefitted from transfers?

Policy makers often have low-income households in mind when favoring transfers to lagging regions.[44] However, according to spatial equilibrium theory it is unclear who eventually benefits from the place-based policy. If subsidies raise local investments and wages, it is likely that higher incomes translate into immigration, higher demand, and thus higher prices for land and housing. As a consequence, pre-treatment property owners reap the benefits and higher nominal income is eaten up by higher land rents.

We run the same regressions as in the spatial discontinuity approach with income per capita and land prices as outcome variables to shed light on this question. For contemporaneous effects, land prices are only available for a subsample of municipalities (Lower-Saxony), while we have information for all relevant states in 2010. We observe from Table 11 that nominal income per capita has increased by about 4-8 percent both contemporaneously and persistently. However, land prices went up by about 25-35 percent, depending on the specification. As households in Germany spend about 30 percent of their net income on rents,[45] real wages in the Zonenrandgebiet have not increased.

According to the framework established by Rosen (1979) and Roback (1982) increases in consumption amenities should be associated with lower real wages in spatial equilibrium. Differences in nominal wages may however prevail as a result of differences in production amenities. This is consistent with persistent location-specific productivity advantages in

---

[44]See, for example, European Commission (2014).
[45]Source: Federal Statistical Office, 2012, Fachserie 15 Reihe 1, Wirtschaftsrechnungen.

## Table 11: Per-capita income and land prices

| | Contemporaneous effects | | | Persistent effects | | |
|---|---|---|---|---|---|---|
| | Coordinate control | | Nonparametric | Coordinate control | | Nonparametric |
| | 2nd | 3rd | $h^*$ | 2nd | 3rd | $h^*$ |
| **Log income per capita** | | | | | | |
| ATE | 0.028** | 0.083*** | 0.038*** | 0.006 | 0.067*** | 0.044*** |
| | (0.012) | (0.015) | (0.014) | (0.013) | (0.017) | (0.014) |
| Adj. $R^2$ | 0.79 | 0.80 | - | 0.14 | 0.17 | - |
| AIC | -3,545 | -3,749 | - | -2,874 | -3,047 | - |
| Obs. | 3,870 | 3,870 | 3,035 | 3,881 | 3,881 | 2,867 |
| **Log land prices** | | | | | | |
| ATE | 0.121 | 0.392*** | 0.269** | 0.220*** | 0.353*** | 0.260*** |
| | (0.111) | (0.150) | (0.117) | (0.049) | (0.060) | (0.063) |
| Adj. $R^2$ | 0.23 | 0.31 | - | 0.25 | 0.30 | - |
| AIC | 1,759 | 1,648 | - | 6,749 | 6,493 | - |
| Obs. | 982 | 982 | 410 | 3,635 | 3,635 | 2,350 |

*Notes*: Land prices per $m^2$ and income per capita are in logs. ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. In 1988 we have only data for Lower Saxony. Land prices correspond to so-called 'Bodenrichtwerte' which are expert evaluations of the land value net of the structures value. These values exist for land allocated to different usage types (housing, business and industry) of which we take the average. Note that the results are robust to individual usage types. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(2) and (4)-(5) refer to parametric specifications and include state indicators. Columns (3) and (6) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012).

the formerly subsidized regions.

These capitalization effects have not been documented in such a pronounced way in the context of place-based policies (see Neumark and Simpson, 2015).[46] One reason why we do find evidence for such effects could be the long-term time horizon of the policy. Transfers were granted for the time of German division which was unforeseeable to end in 1990. As migration decisions are forward-looking, the indefinite time horizon of the policy could have supported the effectiveness of the program substantially.

Note that local subsidies are likely to exert positive externalities on land prices in the neighboring regions as shown by Rossi-Hansberg, Sarte, and Owens (2010) for an urban revitalization program in Richmond (VA). If this externality is continuous in space, our

---

[46]Busso, Gregory, and Kline (2013) find little evidence of capitalization effects when assessing the federal empowerment zones program in the US.

discontinuity estimates again reflect the direct effect on land prices in the treated area which does not include potential externalities.

## 7 Conclusions

We have shown in this paper that temporary regional transfers are able to affect the spatial pattern of economic activity in the long run. As the policy was connected to German division, households and firms expected subsidies to be paid for a longer period and the volume of transfers was substantial. These circumstances were likely influential for the effectiveness of the policy as migration and location decisions are forward-looking.

If agglomeration externalities are continuous at the treatment border, the regression discontinuity design is shown to control for this mechanism as an explanation for persistence. We have used data at the finest spatial scale and executed several robustness checks to explore the validity of assuming frictions to be absent at the border which assures continuity of agglomeration economies. While agglomeration economies have been shown to be an important mechanism to explain spatial economic outcomes in many contexts, our empirical approach allows us to highlight policy-induced locational advantage as another important channel. For example, the development of land is likely to exert long-term effects due to the sunk-cost nature of urban planning or municipalities take advantage of higher economic activity by reinvesting higher tax revenues into local infrastructure to attract commercial activity. Moreover, persistent locational advantage can be generated by sizable adjustment costs at the firm level. Analyzing interaction effects between density and prominent aggregate shocks shows that the formerly subsidized places benefit disproportionately from increases in market access. However, interactions can only explain a small part of the long-term effects.

A second conclusion relates to distributional implications of place-based policies. We have identified substantial capitalization effects of transfers such that higher nominal incomes in the Zonenrandgebiet were derogated by higher land rents. As a consequence, transfers primarily benefitted pre-treatment land owners in the Zonenrandgebiet rather than raising real wages.

# References

**Ahlfeldt, Gabriel, Wolfgang Maennig, and Felix Richter.** 2016. "Urban Renewal After the Berlin Wall: A Place-Based Policy Evaluation." *Journal of Economic Geography* 16(2), 1-28.

**Becker, Sascha, Peter Egger, and Maximilian von Ehrlich.** 2010. "Going NUTS: The Effect of EU Structural Funds on Regional Performance." *Journal of Public Economics* 94(9-10): 578-590.

**Black, Sandra.** 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114(2): 577-599.

**Bleakley, Hoyt, and Jeffrey Lin.** 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127(2): 587-644.

**Brülhart, Marius, Céline Carrère, and Federico Trionfetti.** 2012. "How Wages and Employment Adjust to Trade Liberalization: Quasi-Experimental Evidence from Austria." *Journal of International Economics* 68(1): 68-81.

**Bundesministerium für innerdeutsche Beziehungen.** 1987. "Ratgeber Zonenrandförderung." Bonn.

**Busso, Matias, Jesse Gregory, and Patrick Kline.** 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103(2): 897-947.

**Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik** 2015 "Optimal Data-Driven Regression Discontinuity Plots." *Journal of the American Statistical Association* 110(512): 1753-1769.

**Conley, Timothy C.** 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92(1): 1-45.

**Davis, Donald, and David Weinstein.** 2002. "Bones, Bombs, and Break Points: The Geography of Economic Activity." *American Economic Review* 92(5): 1269-1289.

**Davis, Donald, and David Weinstein.** 2008. "A Search for Multiple Equilibria in Urban Industrial Structure." *Journal of Regional Science* 48(1): 29-65.

**Dell, Melissa.** 2010. "The Persistent Effects of Peru's Mining MITA." *Econometrica* 78(1): 1863-1903.

**Duranton, Gilles and Diego Puga.** 2004. "Micro-foundations of urban agglomeration economies." in Henderson, Vernon and Jacques François Thisse (eds.), Handbook of Regional & Urban Economics, Vol. 4., Amsterdam: North-Holland, 2063-2117.

**Eckey, Hans-Friedrich.** 2008. "Verbesserung der regionalen Wirtschaftsstruktur. Gesetz über die Gemeinschaftsaufgabe vom 6. Oktober 1969." in: Eberstein, H.H. and H. Karl (eds.): Handbuch der regionalen Wirtschaftsförderung, Abschnitt V, Köln.

**Ehrlich, M. von and T. Seidel.** 2015. "The Persistent Effects of Regional Policy: Evidence from the West-German Zonenrandgebiet, CESifo Working Paper No. 5373.

**European Commission.** 2010. "EU-China Cooperation Activity on 'Regional Policy' within the Policy Dialogue between DG Regio and NDRC. Final Report of the Chinese Exports." Brussels.

**European Commission.** 2011. "Panorama Inforegio 40 – Cohesion Policy 2014-2020. Investing in Europe's Regions." Brussels.

**European Commission.** 2014. "Investment for jobs and growth. Promoting development and good governance in EU regions and cities. Sixth report on economic, social and territorial cohesion." Brussels.

**Fan, Jianqing, and Irene Gijbels** 1996. "Local Polynomial Modelling and its Applications." Chapman and Hall, London.

**Glaeser, Edvard, and Joshua Gottlieb.** 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity* Spring: 155-253.

**Gobillon, Laurent, Thierry Magnac, and Harris Selod.** 2012. "Do unemployed workers benefit from enterprise zones? The French experience." *Journal of Public Economics* 96(9-10): 881-892.

**Government Accounting Office.** 2012. "Federal Economic Development Grants to Communities with High Rates of Poverty and Unemployment." GOA-04-938R.

**Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw.** 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69(1): 201-209.

**Henderson, Vernon, Adam Storeygard, and David Weil.** 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102(2): 994-1028.

**Imbens, Guido, and Karthik Kalyanaraman.** 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies* 79(3): 933-959.

**Imbens, Guido, and Thomas Lemieux.** 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics* 142(2): 615-635.

**Karl, Helmut.** 2008. "Entwicklung und Ergebnisse regionaler Wirtschaftspolitik in Deutschland." in: Eberstein, H.H. and H. Karl (eds.): Handbuch der regionalen Wirtschaftsförderung, Abschnitt II, Köln.

**Kline, Patrick, and Enrico Moretti.** 2014. "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *Quarterly Journal of Economics* 129(1): 275-331.

**Lalive, Rafael.** 2008. "How do Extended Benefits Affect Unemployment Duration? A Regression Discontinuity Approach." *Journal of Econometrics* 142(2): 785-806.

**Lee, David, and Thomas Lemieux.** 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48(2): 281-335.

**Marshall, Alfred.** 1920. *Principles of Economics.* Macmillan, London.

**Melitz, Marc J.** 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71, 1695-1725.

**Melitz, Marc J. and Gianmarco I.P. Ottaviano.** 2008. "Market Size, Trade, and Productivity." *Review of Economic Studies* 75, 295-316.

**Michaels, Guy, Ferdinand Rauch.** 2014. "Resetting the Urban Network: 117-2012." unpublished manuscript, London School of Economics.

**Monte, Ferdinando, Stephen J. Redding, Esteban Rossi-Hansberg.** 2015. "Commuting, migration and local employment elasticities." NBER Working Paper 21706.

**Neumark, David, and Helen Simpson.** 2015. "Place-Based Policies." in Duranton, Gilles, Vernon Henderson, and William C. Strange (eds): Handbook of Regional & Urban Economics - Volume 5, pp. 1197-1287.

**Neumark, David, and Jed Kolko.** 2010. "Do Enterprise Zones Create Jobs? Evidence from California's Enterprise Zone Program." *Journal of Urban Economics* 68(1): 1-19.

**Papay, John, John Willett, and Richard Murnane.** 2011. "Extending the Regression Discontinuity Approach to Multiple Assignment Variables." *Journal of Econometrics* 161(2): 203-207.

**Redding, Stephen.** 2010. "The Empirics of New Economic Geography." *Journal of Regional Science* 50(1): 297-311.

**Redding, Stephen, and Daniel Sturm.** 2008. "The Costs of Remoteness: Evidence from German Division and Reunification." *American Economic Review* 98(5): 1766-1797.

**Redding, Stephen, Daniel Sturm, and Nikolaus Wolf.** 2011. "History and Industry Location: Evidence from German Airports." *Review of Economics and Statistics* 93(3): 814-831.

**Roback, Jennifer.** 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90(6): 1257-1278.

**Rosen, Sherwin.** 1979. "Wage-Based Indexes of Urban Quality of Life." in P. Miezkowski and M. Strazheim (eds): Current Issues in Urban Economics , pp. 74-104.

**Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Raymond Owens III.** 2010. "Housing Externalities." *Journal of Political Economy* 118(3): 485-535.

**Schumann, Abel.** 2014. "Persistence of Population Shocks: Evidence from the Occupation of West Germany after World War II." *American Economic Journal: Applied Economics* 6(3): 189-205.

**Turner, Matthew, Andrew Haughwout, and Wilbert van der Klaauw.** 2014. "Land Use Regulation and Welfare." *Econometrica* 82(4): 1341-1403.

**Ziegler, Astrid.** 1992. "Regionale Strukturpolitik: Zonenrandförderung – ein Wegweiser?" Köln: Bund-Verlag.

## Appendix

## A  Data

### A.1  Characteristics of the Zonenrandgebiet

Table A1 displays the number of observations by state as well as the area and population shares of the ZRG.

A list of all 1971 districts that belonged to the ZRG is contained in the federal law 'Zonenrandförderungsgesetz', 1971. Data on population by municipality is available from the Federal Statistical Office for the years 1975-2010. For the years 1950-1970 we have acquired the data from the Statistical Offices of the five Länder we consider. Likewise, information on income, business taxes, and employment was provided by the statistical offices of the individual states. Data on municipal area shares covered by private, public, commercial, industrial, and residential capital was provided by the Federal Institute for Research on Building, Urban Affairs and Spatial Development. The data on land prices are provided by the states' expert committees and by $F\&B$ real estate consulting

Table A1: CHARACTERISTICS OF THE ZONENRANDGEBIET

|  | No. districts | | No. municipalities | | Area ZRG | Pop. ZRG |
|---|---|---|---|---|---|---|
|  | Non-ZRG | ZRG | Non-ZRG | ZRG | in % | in % |
| West Germany | 396 | 106 | 6,839 | 1,573 | 18.6 | 12.3 |
| Schleswig-Holstein | 4 | 14 | 414 | 710 | 53.3 | 81.3 |
| Lower Saxony | 44 | 25 | 711 | 282 | 28.6 | 33.4 |
| North-Rhine Westfalia | 86 | 0 | 386 | 0 | 0 | 0 |
| Hesse | 34 | 13 | 323 | 96 | 27.8 | 19.1 |
| Bavaria | 131 | 54 | 1,533 | 485 | 25.8 | 21.9 |
| Other West-German states | 97 | 54 | 3,472 | 0 | 25.8 | 21.9 |

*Notes*: The states (Länder) Schleswig-Holstein, Lower Saxony, Hesse, and Bavaria belonged to the ZRG. We add North Rhine-Westphalia as it borders with the ZRG, but drop the city states of Hamburg and Bremen. The districts correspond to the 1971 classification while we use data from 1986 for the number of municipalities, and the year 1961 for population shares.

**Georeferencing.** To georeference municipality data, we use digital maps (shape files) from the Bundesamt für Kartographie und Geodäsie. As they are only available since 1997, we assign each municipality to a district in 1971 and drop all observations where the municipality cannot be linked to a district with at least 90 percent of its area (20 municipalities or about 0.4 percent of the sample).[47] Moreover, we drop 31 municipalities due to partial treatment (i.e. ZRG-border crosses the municipality based on the 1997 or 2010 classification). The *boundary sample* of municipalities contains all jurisdictions with a distance to the ZRG border of less than 100 kilometers. This includes all municipalities in the treated region and about 68 percent of the municipalities in the five states west of the ZRG border. For the boundary sample at the district level, we limit the observations to jurisdictions that are sufficiently close to the threshold determining transfer eligibility i.e. $M_d \leq 150$. This includes again all treated observations and about 50 percent of the districts outside the treated area and in the five states. Note that all our analyses are based on the 1971 district classification such that the number of districts remains constant over time.

Depending on the municipality classification of each data source we assign it either to the 1997 or 2010 shape file of municipal boundaries. Thereby, we link 1971 districts (and thus treatment status), coordinates, and distances from the ZRG boundary to the outcome variables. Geospatial data processes have been performed and documented in ArcGIS using shape files about administrative boundaries in 1971, 1997, and 2010. Table A2 provides information about the data sources of all variables we use and the respective level of spatial aggregation.

**Administrative & satellite data.** Table A3 displays summary statistics for all variables in the boundary sample used in our analysis. *income* and *business taxbase* are measured in current 1,000 euros and in current terms. The average income per square kilometer was about 1.5 million euros during the transfer program and it increased to about 2.8 million euros in 2010. Note that averages across municipalities deviate from the country average because cities receive the same weight as small municipalities. Municipalities are the smallest administrative units comprising on average about 33 square kilometers with a high standard deviation and a minimum of 0.45 square kilometers while the largest municipality stretches over 359 square kilometers. The average number of inhabitants

---

[47]This may happen due to changes in administrative boundaries that were frequent especially in the 1970s. Note that all our results are robust to the exclusion of *all* municipalities that could not perfectly be assigned to a 1971 district.

increased from about 5,900 in 1986 to 6,500 in 2010. Accordingly, the municipal average of population density reaches about 160 (178 in 2010) individuals per square kilometer which is well below the German average of about 220. Likewise, per capita income and employment density average at relatively low levels of about 25,000 euros and 40 employees per square kilometer. Note also that 19 (7) municipalities exist in 1986 (2010) that have no employment and taxable business income such that they will be dropped when specifying these outcomes in logarithmic terms.

If not stated differently, all variables refer to municipalities. *public capital* and *private capital* measure the area share of a municipality covered by public infrastructure (area used for streets, railway, airports, seaports, public squares, public buildings etc.) and private structures (industry, business, and housing), respectively. *industrial/private capital* refers to the ratio of industry and business structures in total private structures. *human capital* refers to the share of residents with tertiary education. Land prices correspond to current prices per $m^2$ and stem from evaluations of expert committees.

Private capital and radiance are also summarized for grid cells of $100m \times 100m$ and $30 \times 30$ arc-seconds (about $926m \times 926m$ at the equator), respectively. *Radiance* of grids is computed according to digital integer numbers reported by satellite data of the Defence Meteorological Satellites Program – Operational Linescan System (DMSP-OLS).[48] These data measure night-time lights in the year 1992 and are widely used in research (see, e.g., Henderson, Storeygard, and Weil, 2012). The information on $PrivateCapital$ is provided by the European Environmental Agencies CORINE project for the year 1990. The data contains a variable that indicates 44 different land cover classes. We set $PrivateCapital = 1$ if a place is covered by 'Continuous urban fabric', 'Discontinuous urban fabric', 'Industrial or commercial units', or 'Construction sites' and zero otherwise. Note that the area-weighted sum of $PrivateCapital$ is highly correlated with our municipality-level variable for private capital which is described in section 3 (correlation coefficient of 0.84). Moreover, the averages for the area share of private capital measured at the grid-cell and municipality levels are relatively close to each other even though the data sources differ. The deviation is possibly again due to the identical weight attached to municipalities of different size when summarizing the municipality level data.

---

[48]The satellite data report digital integer numbers ranging from 0 to 63. These may be converted to radiance as a measure of night luminosity by using the formula $radiance = digitalnumber^{1.5}$ for a spatial unit which is denoted in terms of Watts/cm$^2$/sr/nm (in words: Watts per squared centimeter per steradian per nanometer of wave length).

Table A2: Data sources

| Variable | Spatial aggregation | Source |
|---|---|---|
| area | municipality/district | Shapefiles from Federal Institute for Research on Building, Urban Affairs and Spatial Development; Data by municipality from Federal Statistical Office |
| income | municipality | Statistical Offices of the Länder |
| population | municipality | Federal Statistical Office |
| employment | municipality | Statistical Offices of the five Länder |
| human capital | municipality | Federal Employment Agency |
| public capital | municipality/grid cell | Federal Office for Building and Regional Planning; European Environmental Agency |
| private capital | municipality/grid cell | Federal Office for Building and Regional Planning; European Environmental Agency |
| business taxbase | municipality | Statistical Offices of the Länder |
| land price | municipality | States' expert committees; F&B real estate consulting |
| radiance | grid cell | NOAA Defence Meteorological Satellites Program |
| roads | municipality/grid cell | Openstreetmap.org |

Table A3: Descriptive statistics of outcome variables

| | Contemporaneous | | | | Persistent | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std.dev. | Obs. | Year | Mean | Std.dev. | Obs. | Year |
| area in km$^2$ | 33.042 | 33.378 | 3,870 | 1986 | 33.167 | 33.572 | 3,881 | 2010 |
| income/km$^2$ | 1522.511 | 2965.65 | 3,870 | 1986 | 2848.013 | 4701.864 | 3,881 | 2010 |
| radiance (grid cell) | 47.928 | 72.063 | 107,776 | 1992 | 112.420 | 107.019 | 125,527 | 2010 |
| population | 5931.77 | 20318.12 | 3,870 | 1986 | 6530.91 | 21848.77 | 3,881 | 2010 |
| population/km$^2$ | 160.083 | 238.896 | 3,870 | 1986 | 178.333 | 258.123 | 3,881 | 2010 |
| employment/km$^2$ | 39.074 | 99.598 | 3,845 | 1986 | 48.798 | 119.153 | 3,672 | 2010 |
| income/capita | 24.716 | 11.010 | 3,870 | 1986 | 31.461 | 7.247 | 3,881 | 2010 |
| human capital | 0.028 | 0.024 | 1,782 | 1985 | 0.061 | 0.048 | 2,576 | 2010 |
| public capital | 0.044 | 0.020 | 3,856 | 1984 | 0.051 | 0.023 | 3,866 | 2010 |
| private capital | 0.048 | 0.048 | 3,856 | 1984 | 0.067 | 0.057 | 3,866 | 2010 |
| private capital (grid cell) | 0.061 | 0.239 | 7,786,402 | 1990 | 0.074 | 0.262 | 7,786,402 | 2012 |
| industrial capital | 0.007 | 0.012 | 1,259 | 1988 | 0.007 | 0.009 | 1,234 | 2010 |
| business taxbase/km$^2$ | 7.606 | 23.929 | 3,709 | 1986 | 18.194 | 77.084 | 3,870 | 2010 |
| business taxbase/employee | 0.201 | 0.403 | 3,667 | 1986 | 0.425 | 0.626 | 3,642 | 2010 |
| land price | 26.197 | 21.016 | 982 | 1988 | 74.188 | 70.601 | 3,635 | 2010 |

*Notes*: We consider the states (Länder) Schleswig-Holstein, Lower Saxony, North Rhine-Westphalia, Hesse, and Bavaria. We restrict the sample to observations located within 100km of the ZRG border. *income* and *business taxbase* are measured in current 1,000 euros, *human capital* refers to the share of residents with tertiary education, *public capital* and *private capital* measure the area share of a municipality covered by public infrastructure (area used for streets, railway, airports, seaports, public squares, public buildings etc.) and private structures (industry, business, and housing), respectively. *industrial capital* refers to the area share of a municipality covered by industry and business structures. Note that the latter is available for only three states. Land prices correspond to current prices per $m^2$.

# B   Non-parametric estimation

For the nonparametric identification strategy we resort to local linear regressions as these are particular well-suited for inference in the RDD (see Fan and Gijbels, 1996; Imbens and Lemieux, 2008). We employ an edge kernel function and choose different bandwidths according to optimality criteria. We compute the distance of each municipality's centroid to 20 (or 30) border points that are allocated at equal distances along **B** as shown in Figure B1. Then, we assign municipalities to the closest border point, add border-point fixed effects, and use the distance to the respective border point in the local linear regressions to estimate $E[Y_{i0}|\mathbf{L}_i = \mathbf{b}]$ and $E[Y_{i1}|\mathbf{L}_i = \mathbf{b}]$. All our results are insensitive to choosing 20 or 30 border points (red dots and blue triangles in Figure B1, respectively).

As a further robustness check, we refrain from allocating border points and estimate bivariate local linear regressions based on Cartesian coordinates. In this approach we use a product kernel $K_{h_x}(L_{ix} - L_{0x})K_{h_y}(L_{iy} - L_{0y})$ and minimize:

$$\sum_{i=1}^{n} \left\{Y_i - \alpha - (L_{ix} - L_{0x})\beta_1 - (L_{iy} - L_{0y})\beta_1\right\}^2 K_{h_x}(L_{ix} - L_{0x})K_{h_y}(L_{iy} - L_{0y}). \quad \text{(B.1)}$$

Again, this is done separately for units west and east of **B** to obtain $\hat{\alpha}$ for both sides. The pair of bandwidths is chosen according to a cross-validation criterion. In practise this approach is less efficient than the univariate approach based on border points because the additional dimension requires disproportionately more observations. Therefore, we present our non-parametric results generally for the border point approach and note that all results are robust to employing bivariate local linear regressions (corresponding tables are available upon request).

# C   Substitutive policies after 1990

We have shown that transfers to the Zonenrandgebiet led to a *persistent* increase in income per square kilometer in the target region, channelled through higher capital stock, population and employment. This outcome can be interpreted as causal if the former ZRG-border does not exhibit any other discontinuity ex post. For example, one can think about policy makers trying to compensate households and firms in the former treatment area in various ways. However, German municipalities do not have control over many pol-

Figure B1: Border points
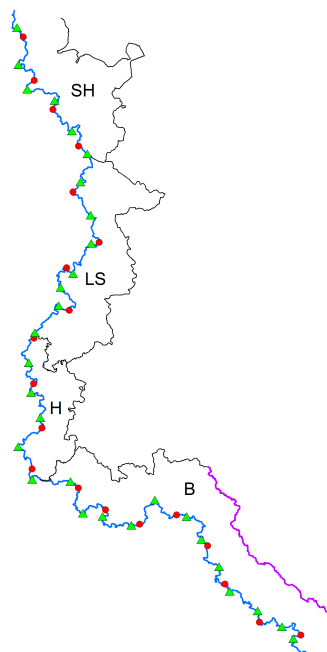


*Notes*: The red dots (green triangles) mark the 20 (30) border points we employ in our analysis. These are allocated at equal distances along the ZRG border. The black lines mark the inner-German border and the state borders. The treated states were Bavaria (B), Hesse (H), Lower Saxony (LS) and Schleswig-Holstein (SH). The Czechoslovakian border is marked in purple. When splitting up the sample into treated units closer to the former GDR or to Czechoslovakia we use the perpendicular distances of municipal centroids or pixels to the respective borders.

icy instruments. Important tax rates like income taxes are chosen at the federal level and have to be approved by the states. And those taxes that municipalities can set themselves are mostly too small to be relevant for location decisions. The business tax rate is an important exception. One could hypothesize that municipalities in the Zonenrandgebiet lowered their business tax rates after 1994 to compensate firms for the loss in subsidies. Based on data from the German Statistical Office, Table C1 shows that there is no discontinuity in business tax rates between 1994 and 2010. Our preferred specifications show insignificant effects. Those that are significant rather point towards higher tax rates in the Zonenrandgebiet.

Further, policy makers could have decided to compensate former ZRG-municipalities by an alternative transfer scheme that substituted the old program, at least to some extent. We use data from the main regional transfer program, the Joint Task, to illustrate that

Table C1: Substitutive policies

| | Local business tax rates (2010) | | | Federal transfers per capita 1994-2010 | | |
|---|---|---|---|---|---|---|
| | Coordinate control | | Nonparametric | Coordinate control | | Nonparametric |
| | 2nd | 3rd | $h^*$ | 2nd | 3rd | $h^*$ |
| ATE | 0.011* | -0.001 | 0.012 | 45.177 | 163.825 | 6.219 |
| | (0.007) | (0.008) | (0.007) | (159.865) | (202.927) | (33.509) |
| Adj. $R^2$ | 0.30 | 0.31 | - | 0.01 | 0.04 | - |
| AIC | -8,433 | -8,450 | - | 69,799 | 69,795 | - |
| Obs. | 3,881 | 3,881 | 2,783 | 3,870 | 3,870 | 1,168 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(2) and (4)-(5) refer to parametric specifications and include state indicators. Columns (3) and (6) refer to non-parametric specifications where $h^*$ denotes the optimal bandwidth computed according to Imbens and Kalyanaraman (2012). We apply a logit transformation to *business tax rates* which are bounded between zero and unity. Federal transfers are aggregated over the years 1994-2010 and divided by the municipal population in 1994. We use the absolute level of *Federal transfers per capita* (instead of logs) in order to account for zeros. These results are robust to a zero-inflated Poisson model that accounts for the relatively high number of zero transfers.

there is no discontinuity in transfer recipience for municipalities at the former ZRG border between 1994 and 2010 (Table C1, columns (4)-(6)). Instead, we observe a discontinuity at the former inner-German border. This result fits into the general picture that regional transfers moved to the new Länder after German reunification (see Figure C1).

Information on federal regional transfers by recipient municipality, year, and transfer type was provided by the Federal Office for Economic Affairs and Export Control. We aggregate over transfer types (infrastructure and subsidies to the private sector) and over years. While it still holds true after 1994 that transfers are channeled to the east of Germany, we do not observe a discontinuity at the former ZRG border as is shown in Figure C1. The left-hand plot corresponds to the results presented in Table C1. Taking into account that the program was terminated in 1994 and that the former ZRG border does not correspond to district borders while transfer intensities are determined by districts, this is not surprising. Moreover, from the right-hand plot in Figure C1 we observe a pronounced discontinuity of transfer intensity at the former inner-German border. In addition to federal transfers we have accounted for resources redistributed within states according to municipal fiscal equalization schemes (in the case of Bavaria). These within state transfers 'Bedarfszuweisungen' are at much lower scale (about 0.3 percent of the federal transfers) and display no discontinuity at **B**.

Figure C1: TRANSFERS 1994-2010
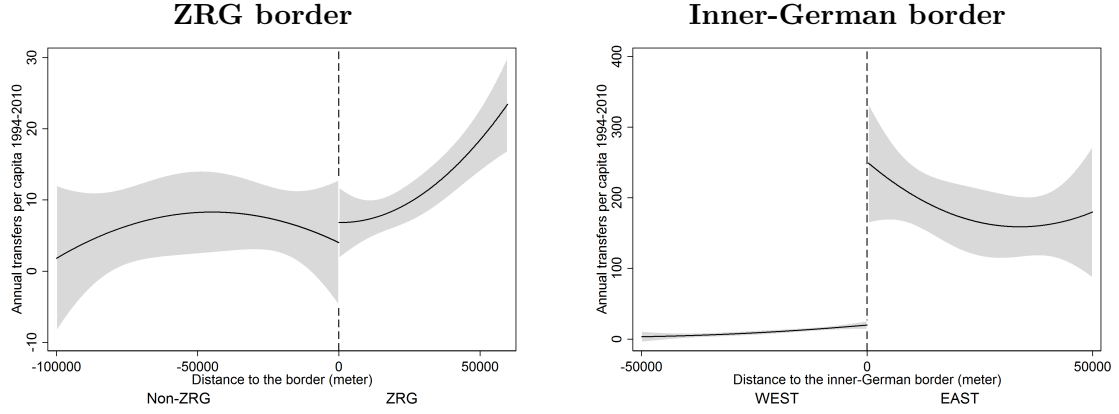
## D  Grid-cell data

For the analysis at the fines spatial scale we use data on *PrivateCapital* and *Radiance* for grids of $100m \times 100m$ and $30 \times 30$ arc-seconds (about $926m \times 926m$ at the equator), respectively (see details about the data in Appendix A). For a sensitivity check about the role of frictions at municipality borders we consider three subsamples that are arguably less prone to inter-municipality frictions. Table D1 reports in columns (1)-(3) the contemporaneous and in columns (4)-(6) the persistent average treatment effects of the ZRG transfers on the probability of a grid cell being covered by private capital and on log radiance.

The first sensitivity check limits the sample to municipality pairs around the treatment border that are highly connected. For this we employ data from OpenStreetMap provided by ThinkGeo which includes all streets and roads in Germany. We drop all small pathways and trails and use GIS techniques to extract roads that cross the ZRG border **B**. Then, we compute for each municipality around **B** the number of roads which cross both the ZRG border as well as the respective municipality. To adjust for municipality size this variable is weighted by municipality area. We keep only grid cells belonging to municipalities in the upper 10th percentile of the distribution of roads crossing **B** and the municipality. Thus the sample size drops by about 90 percent compared to Table 8 while the point estimates (see columns (1) and (4)) remain highly significant and well in line with those presented

in Table 8.

For the second experiment we draw a buffer at 1km distance from both sides of **B**. Employing land-use data from the CORINE project we calculate the share of non-cultivable land within this buffer for all municipalities. We define as non-cultivable land the categories 23-44 ("Forest and and semi natural areas", "Wetlands", "Water bodies"). All municipalities with more then 5 percent of their land within the 1km buffer being noncultivable are dropped from the sample. Thus, we estimate the effect on grid cells belonging to municipality pairs that feature homogeneous characteristics in the neighborhood of **B** as the land could be developed at relatively low costs. Again the estimates in columns (2) and (5) are not increasing substantially compared to the benchmark in Table 8 which supports the assumption that frictions are not decisive for our results.

The third experiment focuses on polycentric municipalities arguing that social networks are more dispersed in scattered neighborhoods. For each grid cell in the data set we calculate the number of grids in a 200, 300, and 400 meter radius that are covered by capital structures. That is we construct a spatial lag for $capital_i^{spatlag} = \mathbf{W} \times capital_j$ for $j \neq i$ and $\mathbf{W}$ referring to a weight matrix with elements being unity if the distance between $i$ and $j$ is less than 200, 300, or 400 meters depending on the specification and zero otherwise. Using the municipality distributions of $capital_i^{spatlag}$ we refer to a point $i$ as a municipality center if it features the highest number of covered grid cells in the surrounding neighborhood i.e. if $capital_i^{spatlag}$ corresponds to the municipality maximum. If this binary definition of a center is unity for multiple adjacent grid cells we aggregate these cells using ArcGIS. In case this definition yields several but *non-adjacent* municipality centers (grid-cells that feature the municipality maximum of $capital_i^{spatlag}$) we refer to the municipality as a polycentric municipality. Independent of the chosen radius, the majority of municipalities turns out monocentric according to the above definition. For the estimates presented in Table D1 we use the 200 meter radius and keep only municipalities with at least two non-adjacent centers. The treatment effects remain highly significant and the magnitude is very similar to the benchmark. Note that this holds also true for using the 300 and 400 meter radii.

## E   Spatial exclusion approach

The spatial exclusion approach is illustrated in Figure E1 where the blue line marks the ZRG border. Blue and yellow shaded areas refer to municipalities in our boundary

Table D1: Robustness: Grid Cell Data

| | Contemporaneous effects | | | Persistent effects | | |
|---|---|---|---|---|---|---|
| | Connectedness | Undeveloped | Polycentric | Connectedness | Undeveloped | Polycentric |
| | | | ***Prob(PrivateCapital=1)*** | | | |
| ATE | 0.024*** | 0.019*** | 0.007*** | 0.023*** | 0.022*** | 0.014*** |
| | (0.001) | (0.002) | (0.001) | (0.001) | (0.002) | (0.002) |
| Adj. $R^2$ | 0.01 | 0.02 | 0.01 | 0.01 | 0.02 | 0.02 |
| Obs. | 776,639 | 147,424 | 758,968 | 776,639 | 147,424 | 758,968 |
| | | | ***Log(Radiance)*** | | | |
| ATE | 0.222*** | 0.376*** | 0.203*** | 0.217*** | 0.241*** | 0.214*** |
| | (0.022) | (0.049) | (0.035) | (0.020) | (0.044) | (0.030) |
| Adj. $R^2$ | 0.12 | 0.17 | 0.11 | 0.10 | 0.24 | 0.09 |
| Obs. | 10,626 | 2,790 | 11,579 | 13,037 | 3,296 | 12,665 |

*Notes*: ***, **, * denote significance at the 1-, 5-, and 10-percent level, respectively. Robust standard errors in parenthesis. We drop all observations outside a 100km window of the ZRG border in the parametric specifications. Columns (1)-(3) and (4)-(6) refer to the contemporaneous and persistent estimates, respectively. All estimates correspond to the parametric specifications and include asymmetric control functions of distance to the ZRG border as well as coordinate controls. Columns (1) and (4) base on municipality pairs at the ZRG border and in the upper 10th percentile of connectedness. Columns (2) and (5) restrict the sample to municipality pairs at the ZRG border that are not separated by undeveloped land such as rocks, water bodies, or forests. Columns (3) and (6) restrict the sample to polycentric municipality pairs at the ZRG border. Observations refer to grid cells of $100m \times 100m$ and $30 \times 30$ arc-seconds for private capital and radiance, respectively.

sample that belong to treated and non-treated areas, respectively. District boundaries (according to the 1971 classification) and municipal boundaries are drawn in red and black, respectively. In the spatial exclusion approach, we drop all treated (non-treated) municipalities within a certain distance from the closest non-treated (treated) municipality. The excluded municipalities are located between the blue and yellow areas (the map corresponds to a 20-kilometer minimum distance between the treatment and the control group). By dropping those municipalities in the close neighborhood of the treatment border, we remove the share of GDP of a district that is potentially affected by positive or negative spillover effects.

Figure E1: RELOCATION EXTERNALITIES – SPATIAL EXCLUSION APPROACH