RFF CMCC European Institute on Economics and the Environment

Effectiveness, Spillovers, and Well-Being Effects of Driving Restriction Policies

Luis Sarmiento, Nicole Wägner, and Aleksandar Zaklan

Working Paper 21-13
May 2021

Effectiveness, Spillovers, and Well-Being Effects of Driving Restriction Policies

Luis Sarmiento* Nicole Wägner † Aleksandar Zaklan‡

May 21, 2021

We study the effectiveness, spillovers, and well-being effects of low emission zones in Germany, an emission-intensity-based driving restriction rapidly growing in popularity. Using regression discontinuity and group-time difference-in-differences designs, we show that previous estimates of the policy's impact on traffic-related air pollution significantly underestimate its effectiveness. We provide evidence of beneficial and harmful policy spillovers to neighboring areas, and increases in ozone due to changes in the chemical balance with precursor contaminants. Policy effects are heterogeneous by season, with greater decreases in traffic pollutants during winter and increases in ozone during spring and summer. Using individual-level data from the German Socio-Economic Panel, we further find that the policy decreases subjective well-being despite clear evidence of health benefits. The decline in well-being is especially pronounced in the first year after policy implementation and is transitory.

JEL codes: Q53, Q58, I31, I18

Keywords: Low emission zones, air pollution, well-being, health, group-time difference-in-differences, regression discontinuity

*RFF-CMCC European Institute on Economics and the Environment (EIEE), Centro Euro-Mediterraneo sui Cambiamenti Climatici, and German Institute for Economic Research (DIW Berlin). E-Mail: luis.sarmiento@eiee.org. Ph: +39 342 613 3782.

†German Institute for Economic Research (DIW Berlin), Berlin School of Economics, and Technical University Berlin. E-Mail: nwaegner@diw.de. Ph: +49 30 89789 575 (corresponding author).

‡German Institute for Economic Research (DIW Berlin), and Berlin School of Economics. E-Mail: azaklan@diw.de. Ph: +49 30 89789 515.

We thank Jana Hamdan and Shushanik Margaryan for helpful discussions, as well as participants at the Verein für Socialpolitik Annual Conference, EAERE Annual Conference, and seminar members in Amsterdam, Berlin, and Milan for valuable comments. All remaining errors are our own.

1. Introduction

Air pollution is a well-known cause of welfare losses, mainly through its impact on health (e.g. Jayachandran, 2009; Knittel et al., 2016) and productivity (e.g. Zivin and Neidell, 2012; Chang et al., 2019; Sarmiento, 2020). Pollution levels are especially high in urban agglomerations, with tailpipe emissions from motorized vehicles being one of the primary sources (Davis, 2008; Gallego et al., 2013). Given the adverse impacts of road traffic and associated pollution externalities, policymakers in many countries have responded through driving restriction programs like alternate-day travel, congestion pricing, or low emission zones (LEZs).[1] LEZs restrict vehicles from entering specific geographical areas based on their emission intensity. Given the significant potential benefits of driving restriction policies and the associated political, economic, and social costs of limiting mobility, a comprehensive analysis of their effects on pollution and well-being is of great economic and social interest.

There is evidence that LEZs reduce air pollution (Wolff, 2014; Gehrsitz, 2017) and improve health outcomes inside their borders (Margaryan, 2021; Pestel and Wozny, 2019). However, to the best of our knowledge, there is no study focusing on the spatial spillovers arising from the introduction of LEZs, no comprehensive analysis of other pollutants except for coarse particulate matter (PM_{10}) and nitrogen dioxide (NO_2), no estimates of heterogeneous policy effects by season, and no evidence on the policy's impact on the well-being of affected individuals.

We fill these literature gaps in several ways. First, we provide a detailed characterization of LEZs' spatial spillovers to neighboring areas. Next, we examine the impact of the policy on carbon monoxide (CO), a previously ignored contaminant in the LEZs literature that acts as a good proxy for traffic pollution (Gallego et al., 2013). Furthermore, we analyze if the effectiveness of LEZs at reducing NO_2 (Pestel and Wozny, 2019; Gehrsitz, 2017) leads to an unintended increase in ground-level ozone (O_3) because of the chemical interaction between these two particles in the lower-atmosphere; henceforth, we refer to changes in O_3 because of

[1] Congestion pricing charges vehicles to enter specific geographical areas, while alternate-day travel policies restrict a share of the vehicle fleet from circulating for particular periods (hours or days) based on license plate numbering, car's vintage, emissions levels, or other characteristics. Although congestion pricing is typically motivated as a measure against congestion and not pollution (e.g. Leape, 2006), there is evidence that it has associated environmental and health benefits (Simeonova et al., 2019).

the policy's effects on primary pollutants as the policy's chemical spillovers.[2] Additionally, we are also the first study to analyze heterogeneous policy effects by season. Concerning the effects of the policy on individual-level outcomes, we evaluate the effect on self-reported measures of life satisfaction for individuals living inside the LEZ, while directly addressing the health channel previously identified in the literature.

To identify the effect of the policy, we need to account for the triple challenge of the staggered implementation of LEZs, time-varying treatment effects, and spatial spillovers. For this, we pursue a two-step quasi-experimental design using regression discontinuity (RD) and group-time difference-in-differences (gtDD). First, the RD model estimates the local average treatment effect (LATE) for a narrow time window around the implementation date at monitoring stations inside and outside the zones borders. Knowing the extent to which LEZs affect pollution outside their limits allows us to restrict the gtDD's control group to comply with the stable unit treatment assumption (SUTVA). The gtDD's treatment group are all stations within the zone's border, the excluded group all outer stations with significant LATE estimates, and control units all stations further away from the excluded area. Additionally, gtDD circumvents two-way fixed effects difference-in-differences (TWFE-DD)'s vulnerability to the staggered implementation of LEZs and their potentially time-varying impacts on pollution levels (Goodman-Bacon, 2018; De Chaisemartin and d'Haultfoeuille, 2020). Moreover, using gtDD also allows us to analyze the potential bias due to the use of TWFE-DD in the existing literature.

Our empirical design relies on two main panel data sets covering the period between 2005 and 2018. The first set contains daily and yearly average measurements of CO, NO_2, O_3, and PM_{10} at 659 monitoring stations scattered across Germany, as well as information on the location and implementation date of LEZs. The information on LEZs and pollution concentrations comes from the German Environment Agency (UBA). The second data set contains individual-level outcomes from the German Socio-Economic Panel (SOEP), a representative longitudinal survey of individuals living in Germany. As individual-level outcomes,

[2]For further information on the interaction of ground-level ozone and its precursors, we refer the reader to Section A.2 in the appendix.

we use an eleven-point Likert scale of life satisfaction as a proxy for individual well-being, the reported number of yearly doctor visits, and a dummy variable indicating hypertension.

The RD model provides evidence of a 12.1%, 10.7% and 15.0% decrease in CO, NO_2, and PM_{10} at pollution monitors inside the LEZs after the introduction of the policy. In general, RD estimates are more pronounced than the literature's TWFE-DD coefficients, implying possible behavioral adaptation or dynamic treatment effects. Concerning spillovers, we provide evidence that LEZs affect pollution not only within the zones' limit but also in adjacent areas. At pollution stations outside the LEZs, PM_{10}, NO_2 and O_3 increase for stations within 500 meters from LEZs borders by 18%, 22%, and 43%, suggesting traffic displacement to major ring roads. In contrast, PM_{10}, NO_2 and CO decrease between 500 m and 25 km from the zones' border, indicating generalized decreases in air pollution within the zones' catchment area. We detect no spillovers further away than 25 km.

The results from the gtDD design show that the introduction of LEZs reduces PM_{10} and NO_2 by 1.97 and 3.79 micrograms per cubic meter ($\mu g/m^3$) while increasing the concentrations of O_3 by 1.15 $\mu g/m^3$. The increase in O_3 implies changes in the chemical balance between O_3 and its precursors, while the PM_{10} and NO_2 results suggest that previous estimates based on TWFE-DD significantly underestimate the effectiveness of LEZs, with coefficient magnitudes in the gtDD design being around 100% larger; we confirm this bias by using the the Goodman-Bacon decomposition (Goodman-Bacon, 2018). Additionally, we show that the efficacy of LEZs is heterogeneous by season. LEZs are especially effective at decreasing traffic-related pollutants during the winter season when vehicle engines tend to be less efficient and driving restrictions have larger marginal effects on traffic-related pollution (Suarez-Bertoa and Astorga, 2018). In contrast, the effect of LEZs on O_3 levels increases during the spring and summer when lower levels of NO_2 and sunnier days increase its formation. Concerning spillovers, we confirm previous findings on the lack of significant spillovers for PM_{10} (Wolff, 2014; Gehrsitz, 2017) and NO_2 (Pestel and Wozny, 2019), while being the first to provide evidence on the absence of significant spillovers for CO. However, we consistently find that O_3 increases at distances up to 25 km away from LEZs.

Regarding treated persons' life satisfaction, our results show that introducing LEZs significantly decreases individuals' self-reported satisfaction with their lives, especially for subpopulations most affected by the restriction like diesel car owners. On average, the life satisfaction of an individual dwelling inside a LEZ decreases by 0.19 points after policy adoption. This LEZ effect is quite substantive relative to other literature estimates, for instance, amounting to more than 20% of the detrimental life satisfaction effect of becoming unemployed. We show that this reduction in life satisfaction is driven by the first years after LEZ implementation with effects becoming insignificant for later years. Furthermore, we provide evidence of improvements in objective health outcomes by estimating a significant decrease in hypertension that accrues mostly to people aged 60 or older, as well as decreases in the number of annual doctor visits among middle-aged adults and older individuals.

This paper confirms that LEZs are an effective air pollution mitigation policy, at least within the zones' coverage area, while also pointing towards the importance of considering its impact on bordering areas and secondary contaminants like O_3. Furthermore, evidence of a decrease in average life satisfaction due to the zones' introduction suggests that such policies generate adverse well-being effects despite health benefits. Our results indicate that, as long as the local context is similar to Germany, LEZs will likely reduce traffic-related air pollution and are preferred to alternate-day driving policies. One advantageous feature of LEZs is that restricting traffic based on pollution intensity prevents the substitution of restricted vehicles with more emission-intensive ones, shutting down an adaptation channel with adverse pollution effects observed in some studies of alternate-day travel (Davis, 2008; Gallego et al., 2013; Barahona et al., 2020). Moreover, the heterogeneity in pollution spillover patterns indicates that policymakers may want to consider distributional impacts when implementing LEZs. Finally, the negative well-being effect suggests that greater effort should be directed at mitigating adverse well-being effects and securing policy acceptance by the population, e.g., through effectively communicating the health benefits of LEZs.

This paper contributes to the broad literature on the efficacy and consequences of driving restriction policies. Current studies on LEZs indicate that their implementation has been largely beneficial regarding air pollution. Wolff (2014) is an important earlier contribution estimating the effect of German LEZs on concentrations of PM_{10} with TWFE-DD. The study finds a significant decrease in PM_{10} levels and no evidence of spatial spillovers. Malina and Scheffler (2015) confirm this result using fixed-effects panel regressions, while Gehrsitz (2017) and Pestel and Wozny (2019) update the estimates for PM_{10} and provide additional evidence of a negative effect on NO_2. Moreover, Ellison et al. (2013) find similar effects for the London LEZ regarding PM_{10}, a result also confirmed for later years by Zhai and Wolff (2021). Our contributions to the literature on the effectiveness of LEZs are estimating spatial spillovers, extending the analysis to additional major pollutants O_3 and CO, computing the LATE using RD, and robustify the TWFE-DD model using the gtDD estimator.

A vital subset of this literature focuses on changes in health outcomes. Margaryan (2021) uses detailed register data on outpatient and inpatient health care to show that LEZs have clear health benefits, especially for the older population. Pestel and Wozny (2019) show that the policy affects not only broad health measures but also extreme health outcomes such as hospitalizations due to cardiovascular and respiratory conditions. In contrast, Gehrsitz (2017) concludes that pollution reductions induced by LEZs are too small to affect infant health, while Rohlf et al. (2020) provides evidence that the health benefits of LEZs translate into lower health expenditures. Our main contribution to the literature on socio-economic outcomes is examining the impact of LEZs on self-rated measures of life satisfaction. We find that, despite clear health benefits, the effect of LEZs on well-being is negative and driven by the first years after policy adoption.

A further related stream of the literature analyzes the effectiveness of the second dominant type of driving restriction, alternate-day travel policies. In contrast to the LEZ literature, the empirical evidence on alternate-day travel policies is mixed. Using RD designs Davis (2008, 2017) analyzes the effect of driving restrictions under the "Hoy No Circula" (HNC) program in Mexico City and concludes that the program was ineffective concerning pollution and that it even increased both the share of high emitting vehicles in the car fleet. Gallego et al. (2013) suggest that the HNC program in Mexico City and similar transport reforms in Santiago de Chile also increased traffic and pollution. The evidence on similar instruments in China suggests greater effectiveness. Chen et al. (2013) show that Beijing temporarily improved air quality with alternate-day travel restrictions on the eve of the 2008 Olympic games. In the same vein, and using RD models, Viard and Fu (2015) show that alternate-day driving restrictions in Beijing are effective at reducing pollution despite evidence of limited compliance (Wang et al., 2014).

This paper increases the comparability of the literature on LEZs and alternate-day travel by estimating the LATE of LEZs with RD models more often used in alternate-day travel studies. Moreover, we further examine the effect of the policy on CO, a contaminant regularly studied in alternate-day-travel articles because of its ability to approximate traffic-related pollution, that until now, was not considered in the LEZs literature. Our results suggest that LEZs are more effective than alternate-day restrictions at reducing traffic-related air contaminants.

2. Data

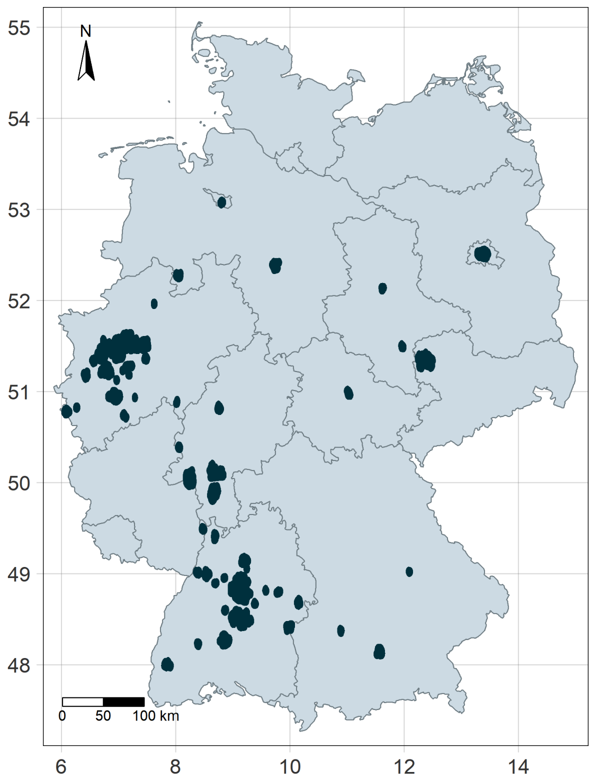
2.1. Air pollution and weather data

Air pollution data comes from 659 monitoring stations scattered throughout Germany and administered by the German Environment Agency (UBA). The dataset contains daily averages of CO, NO_2, O_3, and PM_{10}, between 2005 and 2018.[3] Information on the location and implementation date of each zone also comes from UBA. By 2018, there were 58 active zones in the country with new ones introduced every year between 2008 and 2018, except for 2014.[4] Figure 1 shows the spatial distribution of LEZs and air pollution monitors. LEZs concentrate in the country's most populated urban areas around Berlin, Munich, Stuttgart, Frankfurt, Cologne, and the Ruhrgebiet region. In contrast, monitoring stations are more scattered throughout Germany.

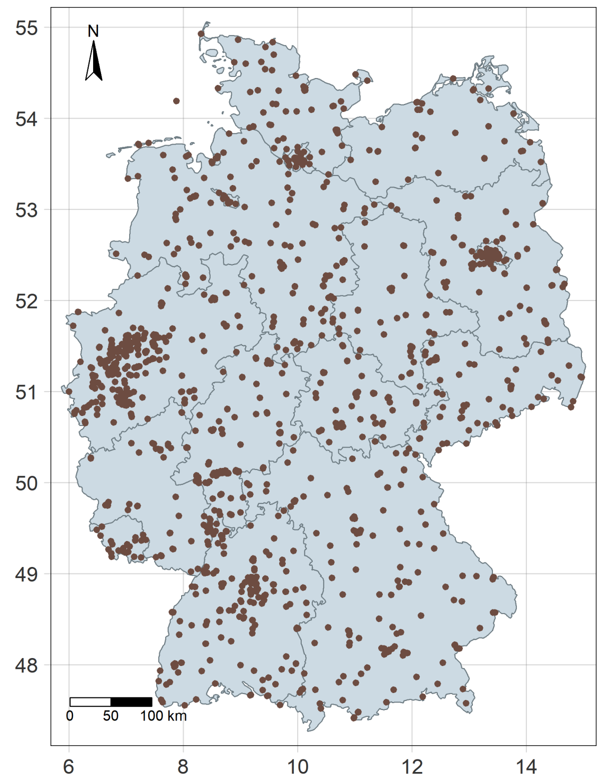
Table 1 lists descriptive statistics of all measured pollutants across monitoring stations inside and outside LEZs. Unsurprisingly, the concentration of traffic-related contaminants is higher for stations inside the LEZs because the zones typically cover city centers. However, O_3 is more eminent at outside stations because of the inverse relationship between ozone and traffic-related contaminants.

[3]Throughout the study, CO values come in milligrams per cubic meter (mg/m^3), and NO_2, O_3, and PM_{10} in micrograms per cubic meter ($\mu g/m^3$).

[4]None of the LEZs in Germany is de-activated or annulled. Some LEZs expand their coverage area over time or merge with neighboring LEZs, e.g., in the densely populated Ruhrgebiet. For further information, refer to Table 17 in the appendix, which contains the name, introduction date, and current stringency level of each LEZ.



(a) Low Emission Zones



(b) Air pollution measuring stations

Figure 1: Spatial distribution of low emission zones and pollution monitoring stations in Germany

Notes: The left-hand panel depicts all LEZs introduced between 2008 and 2018. The right-hand side panel shows all pollution monitors in Germany that are active during the study period between 2005 and 2018.

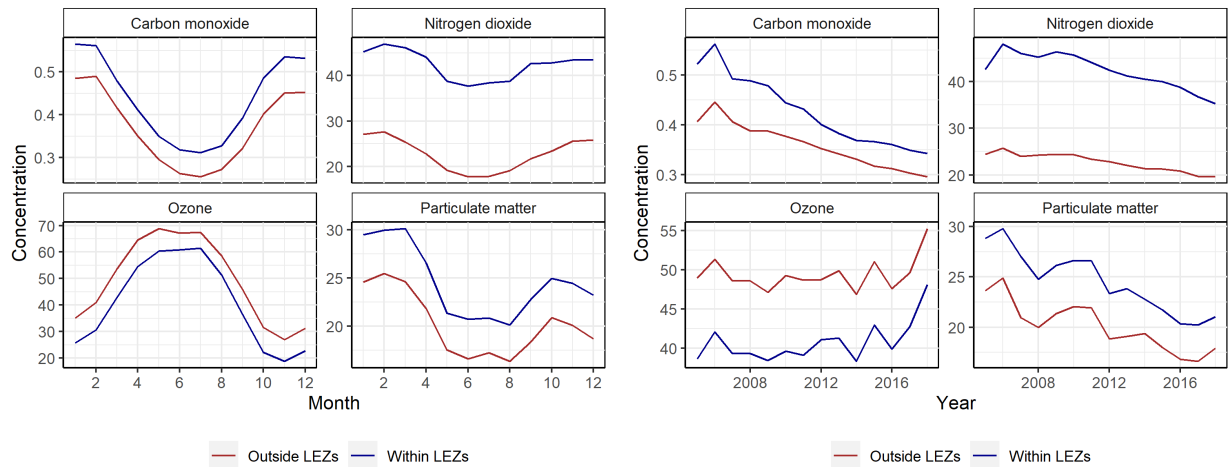
Table 1: Descriptive statistics of air pollution for stations inside and outside low emission zones

	Stations Inside the LEZs				Stations Outside the LEZs			
	Mean	SD	Min	Max	Mean	SD	Min	Max
CO	0.439	0.253	-0.391	3.248	0.371	0.215	-0.148	4.312
NO_2	42.341	22.604	-0.127	246.357	22.752	15.975	-7.694	436.467
O_3	40.745	23.350	-1.922	154.943	49.380	24.408	-2.285	199.896
PM_{10}	24.519	14.716	-2.875	563.966	20.170	13.358	-7.826	1866.604

Notes: This table shows descriptive statistics on daily Carbon monoxide (CO), Nitrogen dioxide (NO_2), Ozone (O_3), and Coarse particulate matter (PM_{10}) measurements from all pollution monitors located inside and outside LEZs, averaged over our sample period from 2005 to 2018.

Figure 2 shows the average monthly and yearly values of all four pollutants. O_3 is higher during spring and summer because of higher solar radiation, while CO, NO_2, and PM_{10} are more prominent during the winter because of residential heating and the lower efficiency of internal combustion engines at low temperatures. Concerning the yearly time trend, CO,

NO_2, and PM_{10} show a steady and downward trend in their concentration across the study period, particularly for stations inside the LEZs. In contrast, O_3 exhibits a stable trend, with noticeable increases at the end of our sample period.



(a) Monthly variation

(b) Yearly variation

Figure 2: Temporal evolution of air pollution across stations inside and outside LEZs

Notes: These figures show the monthly and yearly average concentrations of daily carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and coarse particulate matter (PM_{10}) measurements from all pollution monitors located inside and outside LEZs, averaged over our sample period from 2005 to 2018.

Weather data comes from the German Weather Agency (DWD). DWD provides daily values of wind speed, precipitation, sunshine, temperature, and atmospheric pressure at the station level. We merge the weather and pollution data through inverse distance weighting.[5]

2.2. Individual-level data

To assess the effect of LEZs on individual-level outcomes, we use data from the German Socio-Economic Panel (SOEP) between 2005 and 2018. The SOEP is a representative longitudinal

[5] We interpolate daily weather values at the location of air pollution monitors via inverse distance weighting, using measurements from weather monitors in the vicinity of pollution monitors, as follows:

$$V(pol_{jt}) = \begin{Bmatrix} \frac{\Sigma_i^N \omega(dist_i) * y_{it}}{\Sigma_i^N \omega(dist_{ij})} \\ y_{it} \rightarrow dist_{ij} = 0 \end{Bmatrix} \Longrightarrow \omega(dist_{ij}) = \frac{1}{distance(x_i, x_j)^p}$$

$V(y_{jt})$ is the weighted value of weather at point j and time t, y_{it} refers to the value of weather measured at station i at time t, and $dist_{ij}$ is the distance between pollution monitor j and weather monitor i. The power factor p modifies the heaviness of the weighing load for each monitoring station. The higher p, the larger the weight of closer stations. This paper uses a weight of two and limits the maximum distance to 100 km.

survey that collects information on persons and households in Germany. The data contains our primary outcome of interest, self-rated individual life satisfaction measured on an 11-point Likert scale, plus a wide range of socio-demographic characteristics. Furthermore, we observe the geographic location of households at the street-block level and the exact interview date of individuals, allowing us to determine whether individuals live within the LEZs' boundaries (treatment group) and whether the SOEP interviewed them before or after the zone became active. For individuals residing outside of LEZs, we observe the distance to the closest LEZ and the corresponding implementation date. Since the SOEP incorporated several enlargements and refreshment samples in recent years, we exclude the always treated because their pre-treatment outcomes are unobserved and do not contribute to identification in the group-time difference-in-differences (gtDD).[6] Similarly, we exclude individuals outside of LEZs if the SOEP interviewed them after the closest LEZ came into effect. Additionally, we exclude individuals who change their residence during our study period to avoid potentially confounding effects of residential sorting.

Table 2 depicts the number of persons residing inside LEZs by treatment group. We define treatment groups as the first year in which individuals are observed post-treatment, i.e., the year of the first SOEP interview after LEZ implementation. The fact that new persons are treated every year between 2008 and 2016 illustrates the variation in treatment timing induced by the staggered adoption of LEZs. Early LEZs treat around 42% of all treated persons, with large cities like Berlin and Munich introducing these early LEZs between 2008 and 2009. About 36% of treated cases occur in 2012 and 2013, with the introduction of LEZs in mid-size cities.[7]

Table 3 lists descriptive statistics on individuals dwelling inside and outside LEZs. Both groups exhibit similar average values for life satisfaction, but they differ in other dimensions because LEZs mainly cover city centers. Individuals residing inside LEZs are, on average, more educated, have a higher income, and fewer children relative to individuals outside

[6]By dropping these cases, we avoid unnecessary compositional changes in our treatment group that hinder causal interpretation of our estimates. For similar reasons, we drop individuals surveyed for one single year.

[7]We exclude LEZs introduced in 2017 and 2018 because they were all implemented in relatively small towns, where we observe fewer than 20 treated individuals per treatment group. The small number of units in these treatment cohorts prohibits reliable estimation of group-time ATTs.

Table 2: Number of persons by treatment year

Treatment group	2008	2009	2010	2011	2012	2013	2014	2015	2016	Sum
Number individuals	250	355	65	72	327	189	48	50	80	1436

Notes: This table shows the number of persons per treatment group between 2008 and 2016. Treatment groups are defined as the first year in which individuals are observed post treatment, such that each treatment group contains all persons inside a LEZ with their first (post-treatment) interview after LEZ implementation in the respective calendar year.

Table 3: Descriptive statistics on SOEP individuals

	Inside LEZ		Outside LEZ	
	Mean	SD	Mean	SD
SOEP data				
Life satisfaction [0-10]	7.12	1.73	7.10	1.76
Age [years]	54.30	16.73	54.13	17.03
Is female [%]	0.53	0.50	0.52	0.50
Is employed [%]	0.56	0.50	0.56	0.50
Income [Thsd Euro]	45.13	35.19	42.19	33.80
Education [years]	12.90	3.06	12.26	2.63
Number children	0.44	0.89	0.49	0.92
Owns motor vehicle [%]	0.81	0.39	0.90	0.30
Number motor vehicles	1.26	0.95	1.58	1.06
Owns diesel car [%]	0.32	0.47	0.33	0.47
Number doctor visits	11.06	15.16	10.08	15.09
Has or had hypertension [%]	0.31	0.46	0.32	0.47
Has or had cancer [%]	0.06	0.24	0.06	0.24
Eurostat data				
GDP p.c. [Thsd Euro]	44.34	18.73	30.47	13.24
Population [Thsd]	722.95	796.12	352.53	536.81
Population density [per km^2]	2069.52	1018.83	638.51	882.49
Number individuals	1461		19556	
Number observations	12717		141328	

Notes: This table shows descriptive statistics on all individuals dwelling inside and outside LEZs, averaged over the sample period 2005 to 2018.

the LEZs. Moreover, residents of the city center have more opportunities in choosing their mode of transport, so fewer households inside a zone own a motor vehicle. Regarding health variables, we observe on average one additional doctor visit per year among subjects residing inside LEZs, while the share of hypertension and cancer cases is very similar in both groups. We complement our SOEP sample with supplementary county-level data from Eurostat. Not surprisingly, counties with a LEZ are wealthier in terms of GDP per capita, have a larger population, and are more densely populated.[8]

[8]We use the information on GDP per capita and population density to select a subset of SOEP individuals located in counties with similar population density and GDP per capita.

Figure 3 depicts the annual average life satisfaction of SOEP individuals inside and outside LEZs. Before implementing the first LEZs in 2008, both groups' average life satisfaction developed in parallel, with individuals inside LEZs being slightly more satisfied than people outside. Both groups' life satisfaction dropped in 2009 during the financial crisis, while it experienced an increasing trend in subsequent years.

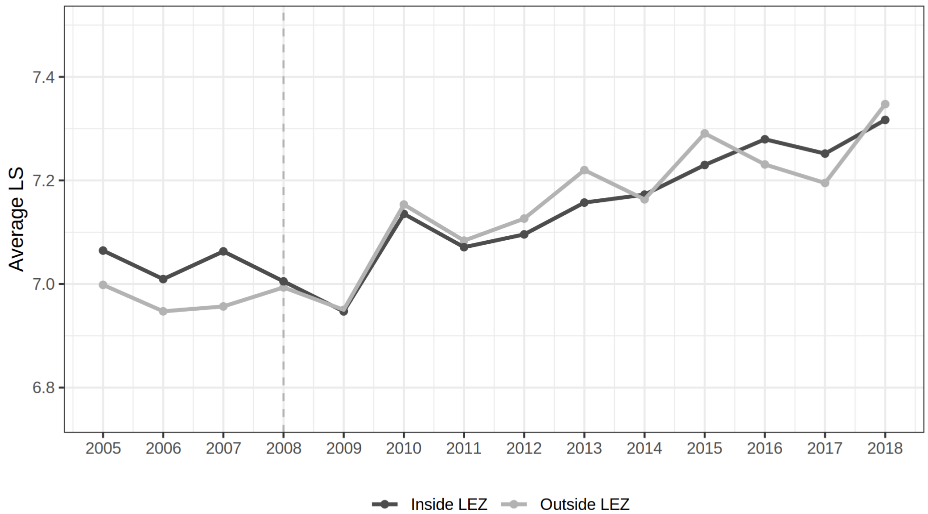


Figure 3: Average life satisfaction over time

Notes: This figure depicts the evolution of average life satisfaction (measured on a 0 to 10 point scale) over time for individuals inside and outside LEZs. The vertical dashed line marks the year 2008 when the first German LEZs were introduced.

3. Research design

We propose a two-step quasi-experimental design to account for the threefold challenge arising from the staggered implementation of LEZs, potentially time-varying treatment effects and spillovers, which would threaten the validity of a standard difference-in-differences (DD) approach. Spillovers violate the stable unit treatment values assumption (SUTVA) necessary for recovering causal estimates in TWFE-DD and gtDD models. Spillovers are relevant for the analysis of LEZs because behavioral adaptations and chemical interactions between pollutants can lead to spatial spillovers. Moreover, the staggered implementation and time-varying treatment effects can assign negative weights to estimates in TWFE-DD models, biasing coefficients, and in some extreme scenarios, even reversing the sign of the effect (Goodman-

Bacon, 2018; Callaway and Sant'Anna, 2020; De Chaisemartin and d'Haultfoeuille, 2020). Time-varying treatment effects arise from the staggered adoption of LEZs, for instance, if the composition of the vehicle fleet changes over time.

First, we use RD models to estimate the LATE of LEZs on daily air pollution within a narrow time window around their implementation date for monitoring stations inside and outside the restriction area.[9] The LATE for stations outside the restriction area determine the influence region of the policy around the implementation date (spatial spillovers). Second, we use gtDD to estimate the average treatment effect (ATE) of LEZs on air pollution and individual-level outcomes. gtDD avoids the bias stemming from negative weights in the weighted computation of the ATE for TWFE-DD models while also letting us compute event-time estimates suitable to examine time-varying treatment effects (Callaway and Sant'Anna, 2020). Moreover, we use the the information on spatial spillovers from the RD design to construct an exclusion region between the zone's border and the estimated outer influence area of the LEZs, and then formally restrict the control group of the gtDD model accordingly.[10]

3.1. Regression discontinuity design

The RD design's main idea is to exploit the sharp discontinuity occurring on the introduction date of LEzs to estimate the LATE. We can determine the policy's impact on treated stations as well as spatial spillovers by looking for significant changes in the pollution levels of outside stations. To the best of our knowledge, this paper is the first to calculate the LATE of LEZs. The LATE may differ from the ATT because of time-varying treatment effects.

[9]Unlike a DD strategy, RD has the advantage that it does not require the correct selection of comparable control groups.

[10]We estimate not only local average treatment effects and average treatment effects for treated units, i.e., stations inside LEZs, but also for "untreated" units, i.e., stations located outside LEZs. These outside stations cannot be considered "untreated" in the usual sense, given that LEZs can affect the stations' readings even though they are not located inside the zone. Throughout the paper, we refer to average treatment effects on the treated (ATT) for stations both inside and outside LEZs for simplicity. However, a more precise term with respect to the treatment effect on outside stations would be the average treatment effect on the unintentionally treated.

Equation 1 shows the RD equation:

$$y_{it} = \beta_{RD} D_{it}(X_{it} \geq 0) + \tilde{\mu}_{-} f(X_{it}) + \tilde{\mu}_{+} f(X_{it} \times D_{it}) + \Omega_t + W_{it} + \varepsilon_{it}\,, \tag{1}$$

where y_{it} are daily average log values of either CO, NO_2, O_3 or PM_{10} for station i at date t. β_{RD} is the coefficient of interest. It captures the LATE around LEZs' introduction date. X_{it} is the running variable, i.e., days until the driving restriction is implemented. D_{it} is a dummy variable equal to one if $X_{it} \geq 0$ and zero otherwise. $\tilde{\mu}_{-} f(X_t)$ and $\tilde{\mu}_{+} f(X_t \times D_{it})$ are linear fits before and after the implementation date. Ω_t contains time fixed effects, while W_{it} is a matrix of weather covariates. We cluster standard errors at the station level and determine the optimal bandwidth with the mean square error expansions plug-in rules described by Cattaneo et al. (2019). Moreover, we report results using robust-bias-corrected confidence intervals. These confidence intervals are different from conventional OLS confidence intervals because they consider the bias stemming from the non-parametric approximation of the local polynomial in determining point estimates and standard errors at the discontinuity (Cattaneo et al., 2019).

When using time as a running variable, a particular worry is that time-correlated unobservables, with discontinuous effects at the date of implementation, may bias the point estimates. For instance, policymakers may introduce LEZs on days with low or high pollution levels, e.g., New Years Day. We account for these potentially discontinuous impacts using a number of alternative fixed effects specifications that control for the day of the week, holidays, school days, and different specifications of weather covariates.

3.2. Group-time difference-in-differences

The gtDD retrieves the policy's group- and time-specific ATTs on pollution levels and individual-level outcomes, based on comparisons between treated and never-treated units.[11] Equation 2 shows the gtDD's estimation equation.

$$y_{itg} = \beta_{ge} LEZ_{itg} + \lambda_i + \gamma_t + \varepsilon_{it}\,, \tag{2}$$

[11] In this subsection, we refer to relevant units as stations to increase readability, although our units of interest can either be monitoring stations or SOEP individuals.

where y_{itg} is the yearly average pollution level for station i, treated in year g, measured at time t. The treatment group g corresponds to all stations treated in $t = g$, e.g., all stations treated in 2008 are part of the 2008 treatment group.[12] β_{ge} are the point estimates of interest and represent the group-time ATTs of LEZs, i.e. the ATT for stations in group g at time since treatment $e = t - g$. LEZ_{itg} is a dummy variable equal to one if, in period t, station i in group g is treated. Finally, λ_s and γ_t are station and year fixed effects. To estimate the dynamic treatment effect across treatment groups we aggregate β_{ge} according to:

$$\beta_e = \sum_{g \in G} \omega_{gt}^e \beta_{ge} \quad \forall \quad \omega_{gt}^e = 1[g + e \leq T] \times P[G = g | G + e \leq T], \tag{3}$$

where $P[G = g|G + e \leq T]$ is the probability of being first treated in period g, conditional on being observed e periods after treatment, and β_e is the average treatment effect e periods after treatment. This aggregation is comparable to traditional TWFE-DD event study regressions but without the negative weights issues. In this weighted sum, the group-specific weights are always positive and lead to event-time average treatment effects with more weight on larger group sizes (Callaway and Sant'Anna, 2020). Note that because the length of treatment can vary across treatment units, treatment groups' composition can change with e. We provide estimates robust to this potential pitfall by balancing the groups concerning e, i.e., we only aggregate β_{ge} for a subset of stations treated for at least n periods. To determine the ATT across all groups and periods, we calculate

$$\beta = \frac{1}{\kappa} \sum_{g \in G} \sum_{e=1}^{T} \omega_{gt} \beta_{ge} \quad \forall \quad \omega_{gt} = 1[t \geq g] \times P[G = g | G \leq T], \tag{4}$$

where $\kappa = \sum_{g \in G} \sum_{t=2}^{T} 1[t \geq g] \times P[G = g|G \leq T]$ ensure that the weights on β_{ge} sum up to one. β is the weighted sum of β_{ge} with larger weights for larger group sizes. As with β_e, this weighted sum avoids the negative weights of TWFE-DD models.

[12]The control group is restricted to never-treated stations and we assume no anticipation effects prior to treatment.

4. Effectiveness and spillovers with respect to air pollution

4.1. Regression discontinuity design

Table 4 contains the main results of the RD design for PM_{10}, NO_2, CO, and O_3.[13] The first column contains no controls, column two includes year and week fixed effects, and column three adds daily temperature, precipitation, and humidity as linear covariates. Column three is our preferred specification as it controls for possible time-varying unobservables biasing our coefficients like new years week or weather events. However, point estimates are qualitatively similar across designs. To simplify interpretation, we transform the log estimates into percentage changes using $exp(\beta)-1$, i.e., an estimate of 0.1 implies a 10% increase in air pollution.[14] The estimated bandwidths range between 27 and 51 days around the implementation date, hence point estimates can be interpreted as the LATE of LEZs on air pollution within one to two months after policy adoption.

The results show that LEZs are effective at reducing CO, NO_2, and PM_{10}. For CO, point estimates suggest a decrease of 5.2%, 12.1%, and 7.7% for the first, second, and third specifications. However, the estimate for the first column is not statistically different from zero. For NO_2, results show a reduction between 6.1% and 10.7%, O_3 presents no statistically significant result, and PM_{10} decreases between 11.5% and 15%.

Table 5 tests the robustness of the RD design against different combinations of fixed effects and weather covariates. Column (1) contains year, week, and station fixed effects. Column (2) adds to (1) weekday, public-holiday and school-day fixed effects. Column (3) contains year, week, daily temperature, precipitation, and humidity plus additional weather covariates like sunshine, atmospheric pressure, wind speed, and humidity. And column (4) contains year, week, temperature, precipitation, and humidity in linear and quadratic terms.

[13] The available number of observations varies across pollutants because most stations monitor only a subset of pollutants. Also, the number of observations increases with the bandwidth, since a larger bandwidth implies a larger time window around the implementation date underlying the RD estimates.

[14] The results in levels are in table 19 of appendix section A.4. In general, they confirm the results of the log design; the introduction of LEZs reduces the concentration of PM_{10}, and NO_2 by 1.66 and 4.36 $\mu g/m^3$. As in the specification with logs, we find no significant point estimates for O_3 and only a significant decrease of 0.058 milligrams per cubic meter (mg/m^3) for CO in the second specification.

Table 4: Effect of the introduction of LEZs on air pollution

(a) Carbon monoxide (CO)

	(1)	(2)	(3)
LEZ	-0.052	-0.121^{***}	-0.077^{**}
	(0.046)	(0.046)	(0.038)
Optimal Bandwidth	32	32	51
No. observations	2510	2510	3911

(b) Nitrogen dioxide (NO_2)

	(1)	(2)	(3)
LEZ	-0.061^{*}	-0.107^{***}	-0.098^{***}
	(0.037)	(0.037)	(0.035)
Optimal Bandwidth	31	30	33
No. observations	5851	5665	6039

(c) Ozone (O_3)

	(1)	(2)	(3)
LEZ	0.098	0.051	-0.065
	(0.073)	(0.067)	(0.053)
Optimal Bandwidth	39	46	48
No. observations	2750	3247	3463

(d) Coarse particulate matter (PM_{10})

	(1)	(2)	(3)
LEZ	-0.115^{**}	-0.150^{***}	-0.116^{***}
	(0.048)	(0.045)	(0.037)
Optimal Bandwidth	27	28	35
No. observations	4335	4335	5452

Notes: Regression discontinuity (RD) estimates of the impact of LEZs on daily log concentrations of carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and coarse particulate matter (PM_{10}). We transform log estimates with $exp(\beta)-1$ and interpret them as the percentage change in the concentration of each particle because of the implementation of LEZs, i.e., an estimate of 0.1 is equivalent to a 10% increase. The first column has no controls, the second adds year plus week-of-the-year fixed effects, and the third further includes weather covariates (daily temperature, precipitation, and humidity). The RD model selects optimal bandwidths based on mean squared error optimization (Calonico et al., 2020). ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter ($\mu g/m^3$) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Reassuringly, point estimates are robust to alternative fixed effects and weather controls, remaining negative, very similar, and statistical significant for PM_{10}, NO_2, and CO.

Table 5: Effect of the introduction of LEZs on air pollution, robustness checks

(a) Carbon monoxide (CO)

	(1)	(2)	(3)	(4)
	-0.123^{***}	-0.119^{***}	-0.099^{***}	-0.108^{***}
	(0.045)	(0.046)	(0.039)	(0.043)
Bwd	32	32	44	36
N. Obs	2432	2510	3367	2744

(b) Nitrogen dioxide (NO_2)

	(1)	(2)	(3)	(4)
	-0.105^{***}	-0.104^{***}	-0.091^{***}	-0.088^{***}
	(0.037)	(0.036)	(0.033)	(0.035)
Bwd	30	30	34	32
N. Obs	5665	5479	6413	6039

(c) Ozone (O_3)

	(1)	(2)	(3)	(4)
	0.043	0.042	-0.066	-0.102
	(0.067)	(0.067)	(0.049)	(0.056)
Bwd	44	47	43	46
N. Obs	3105	3319	3034	3319

(d) Coarse particulate matter (PM_{10})

	(1)	(2)	(3)	(4)
	-0.156^{***}	-0.150^{***}	-0.087^{**}	-0.118^{***}
	(0.045)	(0.045)	(0.035)	(0.039)
Bwd	28	28	32	28
N. Obs	4491	4335	4968	4491

Notes: Regression discontinuity (RD) estimates of the impact of LEZs on daily log concentrations of carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and coarse particulate matter (PM_{10}). We transform log estimates with $exp(\beta)-1$ and interpret them as the percentage change in the concentration of each particle because of the implementation of LEZs, i.e., an estimate of 0.1 is equivalent to a 10% increase. Column (1) contains year, week of the year, and station fixed effects. Column (2) adds to (1) weekday, holidays, and school-day fixed effects. Column (3) contains year, weekof-the-year, temperature, precipitation, and humidity plus additional weather covariates, i.e, sunshine, atmospheric pressure, wind speed, and humidity. Column (4) contains year, week of the year, temperature, precipitation, and humidity in linear and quadratic terms. The RD model selects optimal bandwidths based on mean squared error optimization (Calonico et al., 2020). ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter ($\mu g/m^3$) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Borrowing from Chen and Whalley (2012), we test if implementing the LEZ affects the concentration of sulfur dioxide (SO_2). The largest share of SO_2 emissions comes from residential heating and heavy industry, whereas car emissions contribute only a minor share. Consequently, finding significant effects of LEZs on SO_2 levels may suggest that we are capturing the impact of other industrial policies. Table 6 shows that for the three main specifications, the effect of introducing LEZs on SO_2 is never statistically different from zero.

Table 6: Effect of the introduction of LEZ on sulfur dioxide

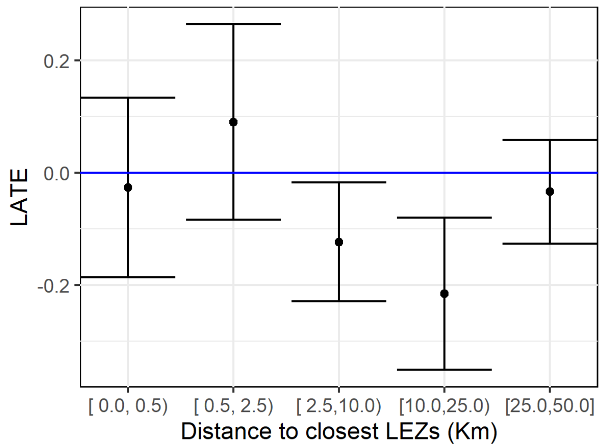
	(1)	(2)	(3)
LEZ	0.057	0.048	0.083
	(0.073)	(0.066)	(0.055)
Bwd	96	89	121
N. Obs	4752	4401	5949

Notes: Regression discontinuity (RD) estimates of the impact of LEZs on daily log concentrations of sulfur dioxide (SO_2). We transform log estimates with $exp(\beta)-1$ and interpret them as the percentage change in the concentration of each particle because of the implementation of LEZs, i.e., an estimate of 0.1 is equivalent to a 10% increase. The first column has no controls, the second adds year plus week-of-the-year fixed effects, and the third further includes weather covariates. The RD model selects optimal bandwidths based on mean squared error optimization (Calonico et al., 2020). sulfur dioxide (SO_2) reported in micrograms per cubic meter ($\mu g/m^3$). Significance levels denoted by $^{***}p<0.01$; $^{**}p<0.05$; $^{*}p<0.1$.

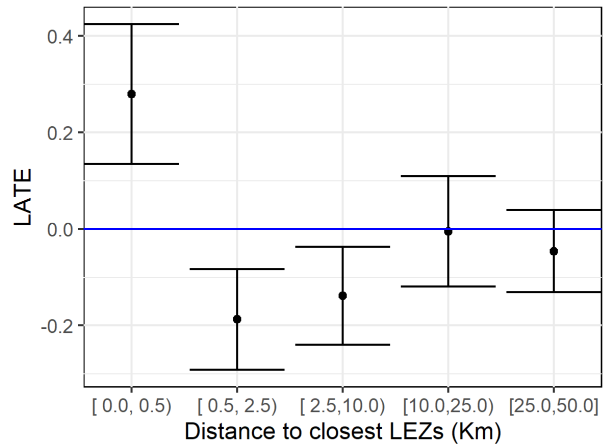
4.2. Spillover effects

This section examines the policy's impact on pollution levels outside the restriction area that can arise from behavioral adaptation or the geographical displacement of pollution. For this, we define five distinct rings based on their distance to the LEZs' borders. The ring between 0 and 500 meters examines the impact in the immediate vicinity of the LEZ while the four rings between 500 meters and 50 km analyze spillover effects at more distant stations. The thickness of rings increases with distance from the LEZ, due to the declining density of measuring stations. Figure 4 depicts point estimates and 95% confidence bands for CO and NO_2 across five outer rings.

For CO, we find no evidence of spillover effects in the rings between 0 and 2.5 km. However, for the samples between 2.5 and 25 km, coefficients are negative and statistically significant. Particularly, there is a 13% reduction in CO for stations between 2.5 and 10 km and a 24% decrease for those between 10 to 25 km. After 25 km, the coefficients are no longer



(a) Carbon monoxide (CO)



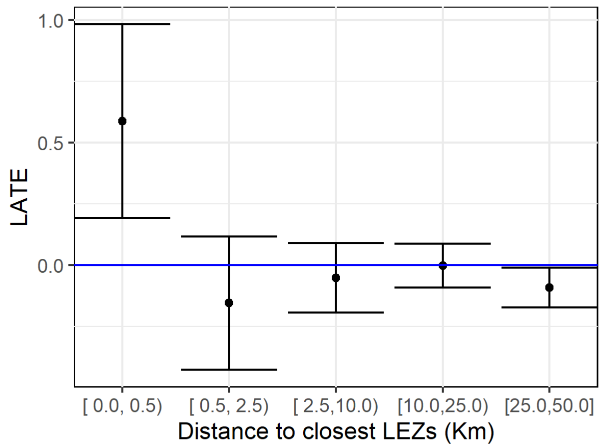
(b) Nitrogen dioxide (NO_2)

Figure 4: Spillover analysis (RD)

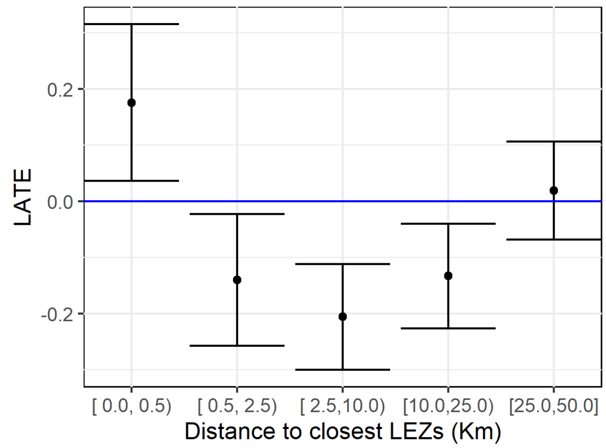
Notes: Regression discontinuity (RD) estimates on the impact of LEZs on daily log concentration of carbon monoxide (CO) and nitrogen dioxide (NO_2) for all stations within the specified distance ring from LEZs, e.g., Coefficients for the [0.0-0.5) buffer refers to the effect of LEZs on all stations 500 m away from the zone's border. We transform log estimates with $exp(\beta)-1$ and interpret them as the percentage increase in each particle because of the implementation of LEZs, i.e., an estimate of 0.1 is equivalent to a 10% increase. All regressions contain weather covariates plus year and year of the week fixed effects. The RD model selects optimal bandwidths based on the mean squared error optimizations (Calonico et al., 2020).

statistically different from zero, suggesting that stations further away than 25 km are not significantly affected by the restriction. Concerning NO_2, point estimates show an increase for the first ring between 0 and 500 meters. Usually, stations between 0 and 500 meters coincide with ring roads surrounding city centers. If the LEZ restricts inner-city traffic, it is likely that the number of vehicles on ring roads would increase because of behavioral adaptations like looking for parking at the outer edge of the zone, driving to the limit and changing transport modes, or driving around the city instead of traveling through it. However, as we move away from the ring road, the effect reverses, revealing significant reductions for areas between 0.5 and 10 km. For all distances beyond 10 km, coefficients are not statistically different from zero. The results for NO_2 suggest that there are two forces at play outside the LEZ. On one hand, fewer trips with high-pollution cars towards the LEZ decrease background NO_2 levels. On the other, the restriction redirects traffic to ring roads increasing pollution values for stations close to the zone's border.

Figure 5 contains the results for O_3 and PM_{10}. For O_3, we observe evidence of harmful spillovers for the stations close to the LEZs border and no significant effects in the other rings. These increases in O_3 may reflect chemical spillovers stemming from declining precursor



(a) Ozone (O_3)



(b) Coarse particulate matter (PM_{10})

Figure 5: Spillover analysis (RD)

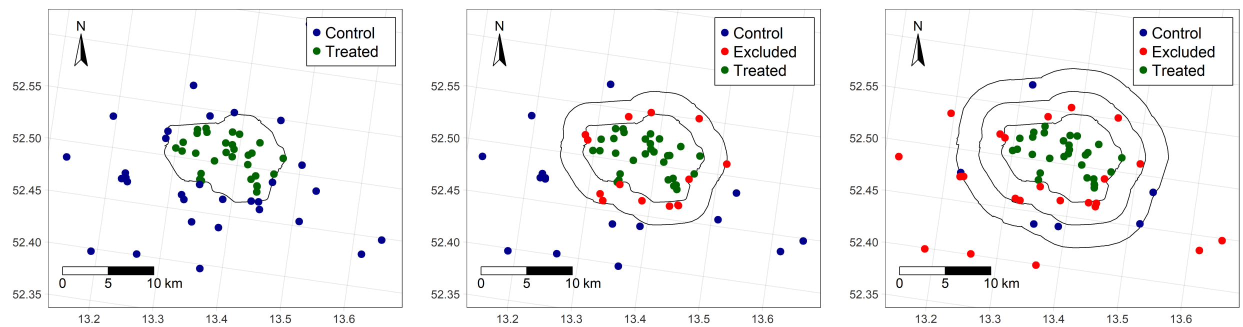
Notes: Regression discontinuity (RD) estimates on the impact of LEZs on daily log concentration of ozone (O_3) and coarse particulate matter (PM_{10}) for all stations within the specified distance buffer, e.g., Coefficients for the [0.0-0.5) buffer refers to the effect of LEZs on all stations 500 m away from the zone's border. We transform log estimates with $exp(\beta)-1$ and interpret them as the percentage increase in each particle because of the implementation of LEZs, i.e., an estimate of 0.1 is equivalent to a 10% increase. All regressions contain weather covariates plus year and year of the week fixed effects. The RD model selects optimal bandwidths based on the mean squared error optimizations (Calonico et al., 2020).

contaminants in the zone's area. Finally, the results concerning PM_{10} are similar to NO_2. We estimate that LEZs cause an increase in PM_{10} levels in the area immediately adjacent to it while reducing its levels in rings up to 25 km away from the zone's limit. As with NO_2 and CO, we find no significant impact at distances of more than 25 km.

These results provide evidence of spillover effects stemming from the introduction of LEZs. Interestingly, spillovers can both increase or decrease contaminants in areas outside LEZs. On average, the restriction decreases air pollution within the zone, increases it in the ring road area, and decreases it at further stations. Moreover, we find no evidence of spatial spillovers at stations further away than 25 km, allowing us to formally restrict the gtDD's control group by excluding all stations likely to be affected by the policy (all stations between zero and 25 km). If the control group in the gtDD is not restricted to stations unaffected by the LEZs, the estimation can underestimate the true magnitude of the average treatment effect for contaminants with beneficial spillovers or overestimate it when harmful spillovers predominate.

4.3. Group-time difference-in-differences

We use three main specifications of the control group in the gtDD estimations, a raw, a buffer, and a doughnut design. Figure 6 provides intuition for these designs using the LEZ of Berlin. In the raw specification, treated and control units are stations inside and outside



(a) Raw (b) Buffer (c) Doughnut

Figure 6: Main specifications for the gtDD design, Berlin low emission zone

Notes: These figures illustrate the three different spatial designs of the gtDD design using Berlin's LEZ. For scale purposes, the figure does not correspond to the actual distance for the exclusion areas in the buffer and doughnut specifications.

the LEZs. In the buffer design, we specify a 25 km exclusion area between the treatment and control group to avoid spatial spillovers threatening the validity of SUTVA. Treated units are all stations inside the zone (green), excluded stations are all stations between 0 and 25 km from the zone's border (red), and the control group all stations further away from the 25 km buffer (blue). Recall that we formally derive the 25 km buffer by implementing the RD model on stations located in different distance-rings from the LEZs. Finally, in the doughnut design, we restrict the outer edge of the control group to 75 km to increase the comparability of treatment and control units and avoid a large share of rural stations in the control group.[15] We choose the doughnut design as our preferred specification because it balances the threat of spillovers with a closer geographical match between the treatment and control groups. However, all specifications are consistent with the common trend assumption and deliver similar ATT estimates.

[15]Very similar results hold at other distances, i.e., 100, 150, 200 km. They are available upon request.

4.3.1. Dynamic treatment effects

Figure 7 plots the annual event-time ATTs (β_e) for NO_2 and PM_{10} across all treatment groups in the doughnut design. Each coefficient corresponds to the ATT across all treated groups (g) for each period before and after treatment, where zero denotes the first year after LEZ implementation. The grey shaded areas represent 95% confidence intervals.

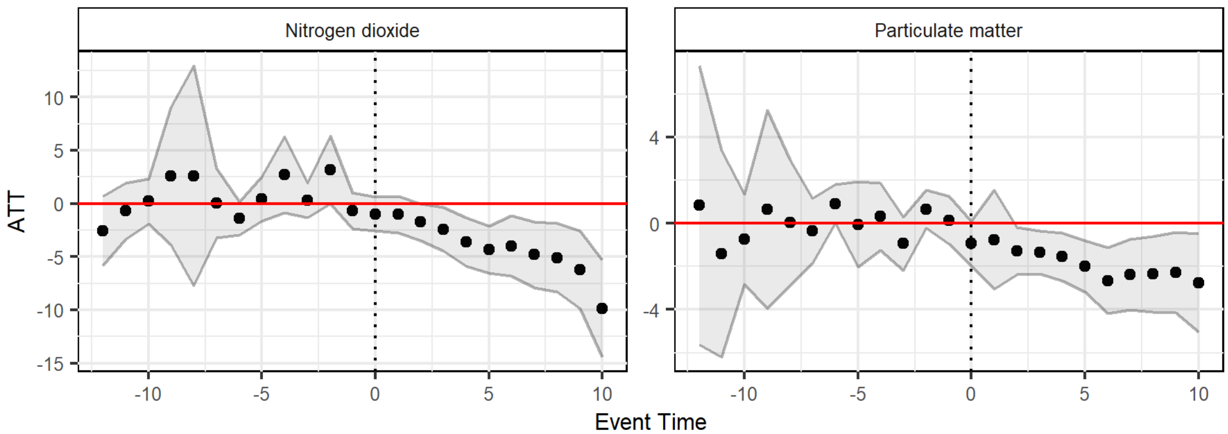


Figure 7: Event-time ATTs on nitrogen dioxide (NO_2) and coarse particulate matter (PM_{10})

Notes: Event-time group-time difference-in-differences (gtDD) estimates (β_e) of the impact of LEZs on yearly air pollution levels for the doughnut design. Event time measured in years before/after LEZ introduction. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. Grey ribbons represent 95% confidence intervals. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. nitrogen dioxide (NO_2) and coarse particulate matter (PM_{10}) reported in $\mu g/m^3$

Reassuringly, results show no significant coefficient for either contaminant before introducing the LEZs at event time zero. This outcome confirms common trends between treatment and control groups. For NO_2, estimates point towards a significant negative effect of LEZs. From the second period onwards, we see statistically significant reductions of NO_2, and in the last time interval, the event-time ATT is as large as -10.3 $\mu g/m^3$. Concerning PM_{10} from the second period onward, the event-time ATT is negative and statistically different from zero. In the last period, the LEZ leads to a reduction in PM_{10} of 3.4 $\mu g/m^3$.

Figure 8 presents analogous estimates for CO and O_3. As for NO_2 and PM_{10}, common trends hold. For CO, we find suggestive although statistically insignificant evidence of negative event-time ATTs after LEZ implementation, with a point estimate in the last period of -0.12 mg/m^3. Concerning O_3, point estimates are positive and even statistically significant

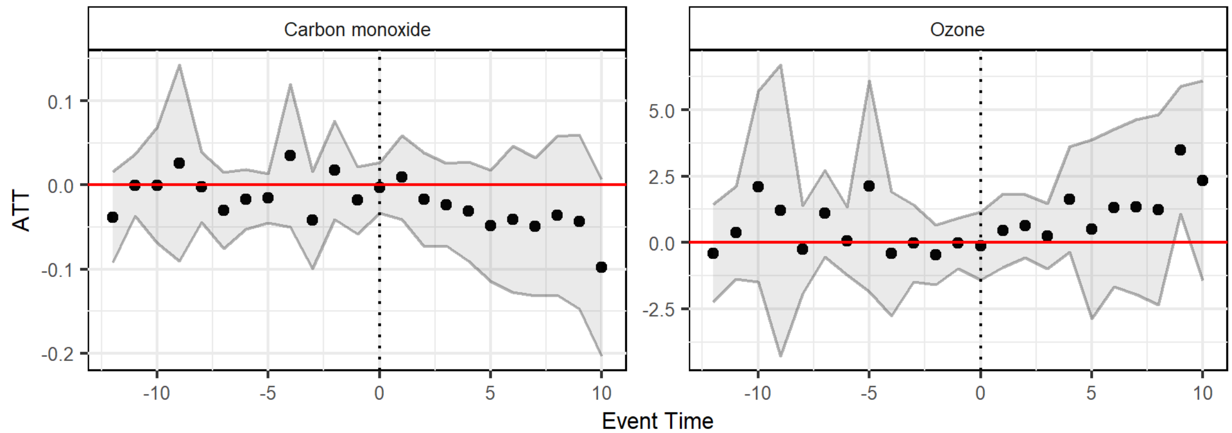


Figure 8: Event-time ATTs on carbon monoxide (CO) and ozone (O_3)

Notes: Event-time gtDD estimates (β_e) of the impact of LEZs on yearly air pollution levels for the doughnut design. Event time measured in years before/after LEZ introduction. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. Grey ribbons represent 95% confidence intervals. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. O_3 reported in $\mu g/m^3$ and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3).

at some intervals. For instance, in the second-to-last period, the LEZ increases O_3 by a significant 4.3 $\mu g/m^3$.

These results show a significant reduction of PM_{10} and NO_2 after introducing LEZs in Germany. Moreover, although statistically insignificant, they also suggest decreases in CO and increases in O_3. Figures 12 and 13 in Appendix A.5 show the same results for the buffer and raw specifications. In the raw design, results are very similar to the doughnut specification. However, although coefficients remain quite similar in the buffer model, standard errors increase, reducing the significance of each event-time ATT.

One caveat of the gtDD estimates is that the composition of treated units changes from period to period. For instance, all stations treated in the last three years of the sample period would only have event-time ATT for $e > 3$, making it difficult to compare estimates across e values. For robustness, we restrict the event-time ATT to avoid compositional changes between periods. We do this by only analyzing the event-time ATT of stations treated for $e \in (1,2,3,4,5)$. Figure 9 shows the results for the balanced design of the doughnut specification. Reassuringly, results confirm that compositional changes do not drive point estimates. The negative coefficients for PM_{10} and NO_2 are very similar, while the statistically insignificant coefficients still hold for CO and O_3.

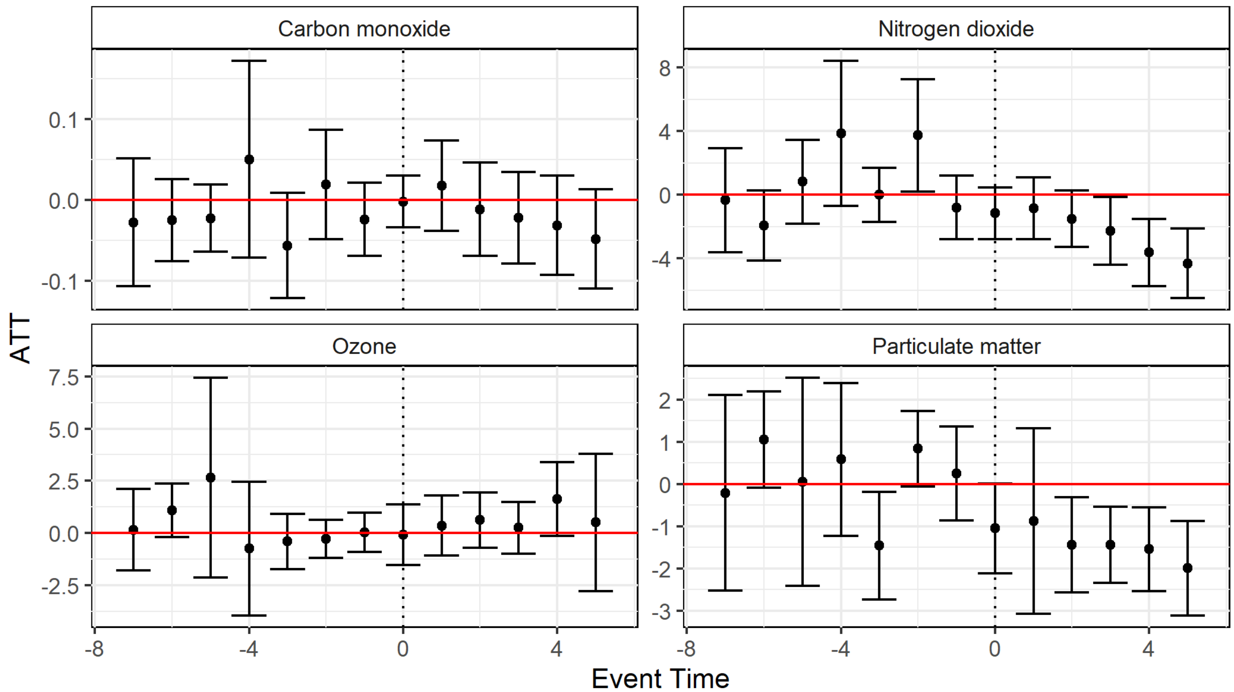


Figure 9: Event-time ATTs for the balanced doughnut specification

Notes: Event-time gtDD estimates (β_e) of the impact of LEZs on yearly air pollution levels for the doughnut design in a balanced panel for treated units at least for $e > 3$. Event time measured in years before/after LEZ introduction. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. Grey ribbons represent 95% confidence intervals. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level.

4.3.2. Average treatment effect on the treated

Next, we calculate the ATT (β) across all groups and periods for the doughnut, buffer, and raw specifications. Doing so allows us to compare our results with estimates from the literature using TWFE-DD models. Moreover, we include an additional specification where only stations in Hamburg serve as the control group. Hamburg is the only large urban area in Germany with no LEZ. Table 7 shows the point estimates for all four designs.

All specifications suggest that LEZs decrease air pollution within their borders. For CO, all coefficients are negative but statistically insignificant. For NO_2, all coefficients are negative, statistically significant at the one percent level, and very similar. In the preferred doughnut specification, LEZs decrease NO_2 concentrations by 3.792 μg/m^3. Our estimates suggest that LEZs are far more effective with respect to NO_2 than findings in the existing literature, where Pestel and Wozny (2019) estimate that LEZs decrease NO_2 levels by about 1.6 μg/m^3, while

Table 7: Effect of the introduction of LEZ on air pollution (gtDD)

(a) Raw

	CO	NO_2	O_3	PM_{10}
	−0.028 (0.020)	-3.399^{***} (0.783)	0.926 (0.597)	-1.688^{***} (0.398)
N.Stations	270	622	387	563
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(b) Buffer

	CO	NO_2	O_3	PM_{10}
	−0.026 (0.022)	-3.477^{***} (0.791)	1.246^{**} (0.568)	-1.862^{***} (0.420)
N.Stations	208	472	286	419
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(c) Doughnut

	CO	NO_2	O_3	PM_{10}
	−0.032 (0.021)	-3.792^{***} (0.858)	1.155^{*} (0.592)	-1.968^{***} (0.424)
N.Stations	141	299	164	274
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(d) Hamburg

	CO	NO_2	O_3	PM_{10}
	−0.053 (0.035)	-3.942^{***} (1.010)	2.001^{***} (0.731)	-1.749^{***} (0.604)
N.Stations	83	163	52	138
N.Groups	9	9	9	9
N.Periods	14	14	14	14

Notes: Group-time difference-in-differences (gtDD) estimates (β) of the impact of LEZs on yearly air pollution levels for four different specifications and air pollutants. In the raw specification, treated and control units are stations inside and outside the LEZs. In the buffer design, we include a 25 km exclusion area between the treatment and control group. In the doughnut design, we restrict the outer edge of the control group to 75 km to increase the comparability of treatment and control units. In the Hamburg specification, we restrict control stations to the city of Hamburg. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter ($\mu g/m^3$) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Gehrsitz (2017) finds a reduction by about 0.5 $\mu g/m^3$. Regarding O_3, we find significant estimates in the buffer, doughnut, and Hamburg designs. In the doughnut model, the zone's introduction increases the concentration of O_3 by 1.155 $\mu g/m^3$. Finally, all specifications point to a significant reduction in PM_{10}. In the preferred specification, the ATT for PM_{10} is a reduction of 1.968 $\mu g/m^3$. This is also in excess of findings in the TWFE-DD literature, where Pestel and Wozny (2019) estimate a reduction by about 1.4 $\mu g/m^3$ and Gehrsitz (2017) finds a reduction by about 0.7 $\mu g/m^3$. Only the estimate by Wolff (2014), which uses log-transformed values of the pollution outcomes, is in line with our estimates. The 9.1% effect of LEZs on PM_{10} levels found by Wolff (2014) translates to a decrease by about 2.1 $\mu g/m^3$.[16]

The discrepancy between our NO_2 and PM_{10} estimates and previous studies' findings is not surprising, given that these studies exclusively apply TWFE-DD models, which are biased due to the staggered adoption of LEZs and their time-varying effects on pollution levels. For a multi-period framework with variation in treatment timing, Goodman-Bacon (2018) shows

[16] We obtain a reduction by 2.1 $\mu g/m^3$ by taking 9.1% of the 2007 average PM_{10} level of 23.1 reported by Wolff (2014) in Table 1.

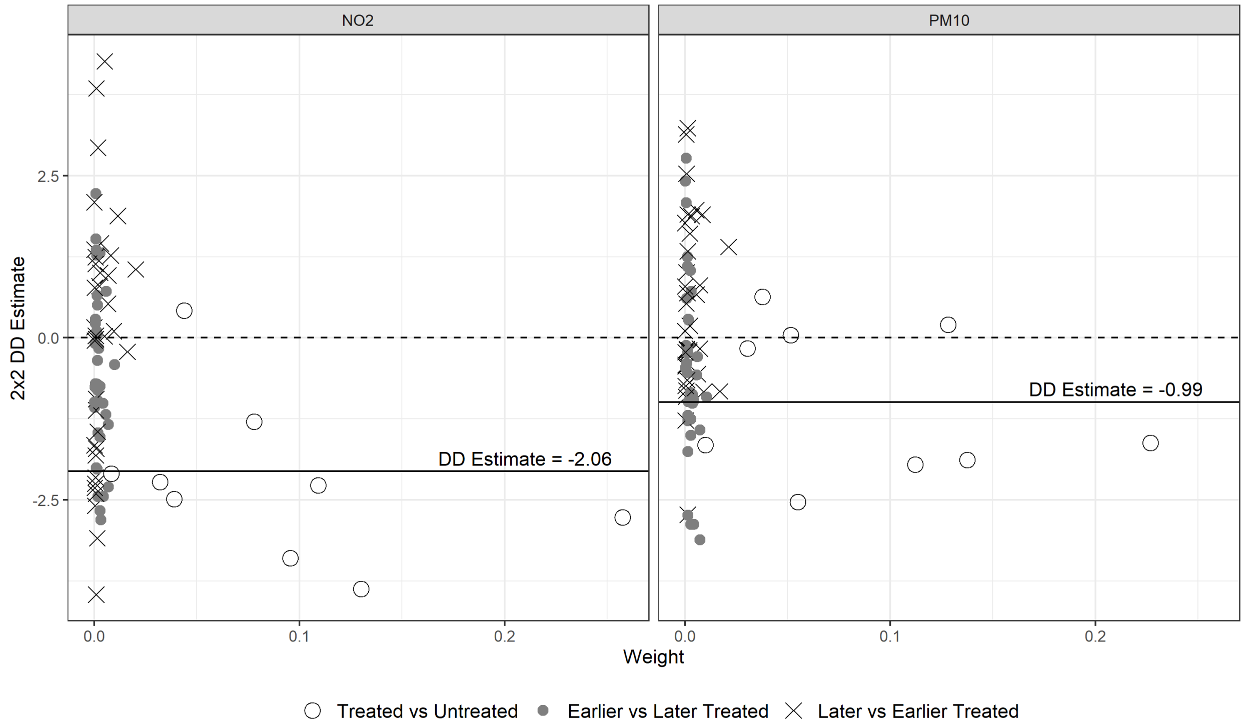


Figure 10: Goodman-Bacon decomposition of TWFE-DD estimates

Notes: The figures plot the 2x2 DD estimates from the Goodman-Bacon (2018) decomposition against their associated weights. Solid horizontal lines represent the TWFE-DD estimates of the LEZ impact on annual levels of NO_2 and PM_{10}, respectively. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. Each estimation sample consists of a balanced panel of stations measuring the respective pollutant between 2005 and 2018.

that the TWFE-DD coefficient is a weighted average of all possible two-group two-period (2x2) DD estimators. Figure 10 depicts the decomposition for NO_2 and PM_{10} proposed by Goodman-Bacon (2018), with the point estimates for each 2x2 DD comparison on the y-axis and the corresponding weights on the x-axis.[17] The symbols represent 2x2 DD estimates based on three different types of comparisons: Treated vs. Untreated, Earlier Treated vs. Later Treated, and Later Treated vs. Earlier Treated. The solid horizontal lines depict the TWFE-DD estimates of the zones' impact on the respective pollutant,[18] which are roughly half the size of their gtDD counterparts in Table 7. The figure illustrates that the TWFE-DD estimates are biased towards zero due to the 2x2 DD comparisons between treated and already-treated stations (Later vs. Earlier Treated), depicted by the cross symbols.

[17] Weights are based on treatment group sizes and treatment timing of the respective 2x2 DD comparisons.
[18] The TWFE-DD is equal to the weighted average of all 2x2 DD estimates.

This timing group frequently generates positive point estimates,[19] because already-treated stations (Earlier Treated) experience strong decreases in pollutant levels years after LEZ implementation, as shown in Figure 7. Using their post-treatment outcomes as a control group for stations treated in later years results in underestimation of the true LEZ impact on the Later Treated. In our setting, the bias is so severe that for most of the Later vs. Earlier Treated comparisons the sign of the coefficient reverses relative other comparisons. Since the gtDD framework circumvents these problems by excluding these timing comparisons, the point estimates are closer to the true parameter, hence more pronounced in magnitude than the TWFE-DD estimates.

4.3.3. Heterogeneous treatment effects

Working with yearly averages may mask the effects of LEZs during specific periods. For instance, ozone is predominantly present during the spring and summer months because of its interaction with solar radiation, while carbon monoxide levels are higher in the winter due to the burning of gas, fuel-oil, and wood for residential heating alongside a lower efficiency of internal combustion engines at low temperatures. To explore heterogeneous effects by season, we calculate the average exposure of each contaminant by season. Table 8 shows seasonal ATT estimates for CO, NO_2, O_3, and PM_{10} using the doughnut specification.

Point estimates differ across seasons. In winter, we estimate significant decreases for CO, NO_2, and PM_{10} by -0.044 mg/m^3, -3.31 μg/m^3, and -1.86 μg/m^3, respectively. The significant and larger effect in winter for CO is likely driven by internal combustion engines' lower efficiency at cold temperatures. These lower efficiencies raise traffic policies' effectiveness because of increases in the winter emissions of older vehicles.[20] In the spring, the ATTs for CO, NO_2, and PM_{10} remain without significant changes, while the coefficient for ozone increases in size and becomes significant and positive, implying that LEZs increase spring

[19] On average, the 2x2 DD estimates for the Later vs. Earlier comparisons amount to 0.7 μg/m^3 for NO_2 and 0.49 μg/m^3 for PM_{10}. These types of comparisons get a weight of 12 percent, so this timing group is relatively influential for the overall TWFE-DD parameter.

[20] Santiago de Chile alternate-day travel program is an example of differentiated traffic restrictions by season. Every year, in the winter months between April and March, Santiago restricts the number of vehicles on the road by limiting the transit of specific plate numbers (Rivera, 2021)

Table 8: Seasonal effects of the introduction of LEZs on air pollution

(a) Winter

	CO	NO_2	O_3	PM_{10}
	-0.044^{*}	-3.318^{***}	0.680	-1.866^{***}
	(0.024)	(0.758)	(0.535)	(0.532)
N.Stations	136	295	159	271
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(b) Spring

	CO	NO_2	O_3	PM_{10}
	-0.043^{*}	-4.385^{***}	1.732^{**}	-1.868^{***}
	(0.024)	(1.021)	(0.828)	(0.444)
N.Stations	136	295	159	271
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(c) Summer

	CO	NO_2	O_3	PM_{10}
	-0.019	-3.605^{***}	1.955^{***}	-2.005^{***}
	(0.019)	(0.984)	(0.746)	(0.552)
N.Stations	136	295	159	271
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(d) Autumn

	CO	NO_2	O_3	PM_{10}
	-0.027	-4.595^{***}	0.749	-2.691^{***}
	(0.025)	(0.900)	(0.470)	(0.486)
N.Stations	136	295	159	271
N.Groups	9	9	9	9
N.Periods	14	14	14	14

Notes: Group-time difference-in-differences (gtDD) estimates (β) of the impact of LEZs on yearly air pollution levels for four different seasons and air pollutants. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. We calculate the average exposure in each season by averaging pollution values for each day of the year within each season. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter (μg/m^3) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

concentrations of O_3 by 1.73 μg/m^3. In the summer, all particles but CO are statistically significant, and in autumn, only NO_2 and PM_{10}.[21]

4.3.4. Spillover effects

Next, we use the group-time difference-in-differences (gtDD) model to analyze low emission zones' impact outside their borders. For this, the model restricts the analysis to the 25 km influence area suggested by the RD design and determines the effect on the pollution levels of stations outside the zone. To check for spillovers in different areas outside the LEZs, we examine all stations' between the zone's edge and one, five, ten, and twenty-five kilometers from it. The treatment group comprises all stations within the buffer zone, and the control group, all stations further away than the 25 km limit, e.g., in the 1 km buffer, we consider all stations between the zone's border and 1000 m as treated and all stations further away than 25 km as control. Table 9 shows the results across all four buffers and air contaminants.

[21] Figure 14 in appendix A.5 explores the seasonal dimension of ATTs by examining the effect of LEZs on monthly pollution levels. CO is only significant between February and April, NO_2 and PM_{10} across the year, and O_3 between the sunny months of April and July.

Table 9: Spillovers across different radii

(a) Spillovers between 0 and 25 km from LEZs

	CO	NO_2	O_3	PM_{10}
	0.018	0.218	1.252**	0.188
	(0.018)	(0.797)	(0.562)	(0.366)
N.Stations	189	452	319	414
N.Groups	9	9	9	9
N.Periods	14	14	14	14

(b) Spillovers between 0 and 10 km from LEZs

	CO	NO_2	O_3	PM_{10}
	0.033	0.313	1.199**	0.599
	(0.021)	(0.701)	(0.591)	(0.417)
N.Stations	169	397	282	362
N.Groups	6	7	7	7
N.Periods	14	14	14	14

(c) Spillovers between 0 and 5 km from LEZs

	CO	NO_2	O_3	PM_{10}
	0.028	0.195	1.001	0.763
	(0.024)	(0.957)	(0.905)	(0.543)
N.Stations	156	364	259	331
N.Groups	4	6	6	6
N.Periods	14	14	14	14

(d) Spillovers between 0 and 1 km from LEZs

	CO	NO_2	O_3	PM_{10}
	0.016	0.715	1.821**	0.717
	(0.021)	(1.243)	(0.893)	(0.718)
N.Stations	142	338	241	307
N.Groups	4	5	4	5
N.Periods	14	14	14	14

Notes: Group-time difference-in-differences (gtDD) estimates (β) of the impact of low emission zones on air pollution for four different buffer rings outside the zone's limit. Treated stations are all those stations inside the outside ring, and control stations all stations further away than 25 km from the zone's border. We calculate the average exposure in each ring by averaging pollution values for each day of the year for all stations within the ring. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter (μg/m^3) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

We do not find evidence of spillovers for CO, NO_2, or PM_{10} across any buffer zone. Although their point estimates are always larger than zero, they are never significant. For PM_{10} and NO_2, this estimated absence of spillovers is in line with the existing literature (Wolff, 2014; Pestel and Wozny, 2019). For O_3, on the other hand, there is evidence of harmful spillovers to individuals living outside of the LEZ. In the 0 to 25 km buffer zone, LEZs increase O_3 by 1.252 μg/m^3. This effect also holds in the 0-10 and the 0-1 km buffer while remaining positive but statistically insignificant in the 0-5 km interval. As with the results for stations inside the LEZs, we explore if the spillovers vary by season. Table 10 shows the point estimates for winter, spring, summer, and autumn.

We find no statistically significant spillovers across any radii or season for CO and NO_2. However, it is worth noting that the majority of coefficients are larger than zero. For O_3, spillovers concentrate in the summer months with increases as large as 5.89 μg/m^3 for the 0-1 km buffer. Furthermore, coefficients also show significance in the smallest and largest radii for the autumn months of September, October, and November. Next, we find harmful spillovers for PM_{10} in the spring and summer months, with statistically significant increases as large as 2.165 μg/m^3.

Table 10: Spillovers across different seasons and radii

(a) Carbon Monoxide

	0-1 km	0-5 km	0-10 km	0-25 km
Winter	−0.007	0.026	0.031	0.026
	(0.024)	(0.028)	(0.024)	(0.021)
Spring	0.018	0.030	0.030	0.021
	(0.019)	(0.027)	(0.020)	(0.018)
Summer	0.012	0.022	0.025	0.010
	(0.020)	(0.025)	(0.017)	(0.017)
Autumn	0.019	0.033	0.033	0.014
	(0.018)	(0.027)	(0.021)	(0.019)

(b) Nitrogen Dioxide

	0-1 km	0-5 km	0-10 km	0-25 km
Winter	0.265	0.176	0.327	0.309
	(1.158)	(0.791)	(0.633)	(0.756)
Spring	0.535	−0.474	0.049	−0.161
	(1.833)	(1.157)	(0.917)	(0.794)
Summer	0.483	1.149	1.060	0.166
	(1.432)	(1.179)	(0.934)	(0.824)
Autumn	1.535	0.178	0.227	−0.240
	(1.219)	(0.935)	(0.746)	(0.796)

(c) Ozone

	0-1 km	0-5 km	0-10 km	0-25 km
Winter	0.415	0.583	0.167	−0.300
	(0.573)	(0.447)	(0.440)	(0.576)
Spring	0.027	0.674	0.489	0.085
	(0.806)	(0.855)	(0.683)	(0.633)
Summer	5.895***	2.863***	3.074***	2.339***
	(1.918)	(0.921)	(0.732)	(0.605)
Autumn	1.194*	0.836	1.125	1.330**
	(0.616)	(0.687)	(0.744)	(0.543)

(d) Particle Matter

	0-1 km	0-5 km	0-10 km	0-25 km
Winter	−0.167	0.599	0.252	−0.208
	(1.155)	(0.684)	(0.628)	(0.545)
Spring	2.165***	1.900***	1.619***	0.913**
	(0.787)	(0.710)	(0.507)	(0.391)
Summer	0.181	1.039	1.131**	0.428
	(0.716)	(0.652)	(0.486)	(0.354)
Autumn	0.544	−0.266	−0.263	−0.296
	(0.878)	(0.607)	(0.442)	(0.383)

Notes: Group-time difference-in-differences (gtDD) estimates (β) of the impact of low emission zones on air pollution for four different seasons buffer rings outside the zone's limit. Treated stations are all those stations inside the outside ring and control stations all stations further away than 25 km from the zone's border. We calculate the average exposure in each ring-season combination by averaging pollution values for each day of the year for all stations within each ring and season. The gtDD model only controls for station and year fixed effects. Standard errors clustered at the station level. ozone (O_3), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}) reported in micrograms per cubic meter ($\mu g/m^3$) and carbon monoxide (CO) in milligrams per cubic meter (mg/m^3). Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

5. Well-being and health effects

5.1. Average treatment effect on the treated

To analyze LEZs' impact on individual-level well-being and health outcomes, we apply our gtDD design to three different samples. First, we include all SOEP individuals (i.e., without restricting the control group in any way). Second, we exclude control persons residing within 25 km distance from a LEZ to account for potential spatial spillovers. Spatial spillovers in life satisfaction might arise either due to changes in pollution levels in close vicinity to the LEZs, or due to individuals outside of LEZs being restricted in their mobility behavior, e.g., because they have to travel into or through the LEZ. Third, we further restrict the control group to individuals living in counties that are similar to LEZ counties in terms of population density and GDP per capita.[22] Restricting the sample based on county characteristics increases the similarity between treatment and control groups by avoiding comparing urban vs. rural populations.

Table 11: Effect of LEZ introduction on life satisfaction

	(1)	(2)	(3)
ATT	-0.162^{***}	-0.190^{***}	-0.139^{**}
	(0.057)	(0.054)	(0.060)
N.Individuals	20963	12588	5972
N.Groups	9	9	9
N.Periods	14	14	14

Notes: Group-time difference-in-differences (gtDD) estimates of the impact of LEZs on life satisfaction of individuals living inside the LEZs. Point estimates represent the simple aggregation across all groups and time periods (β). The first column lists results obtained on the full sample, the second column restricts the control group to individuals in residences further than 25km away from the nearest LEZ, and the third column restricts the control group further to individuals living in counties that are similar to LEZ counties in terms of population density and GDP per capita. Standard errors clustered at the household level. Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

We estimate the ATT for individuals living inside the LEZs and present the simple aggregation across all groups and time periods of the gtDD estimates (β). Table 11 shows the estimated effects of LEZ on life satisfaction for all three samples. All point estimates are negative and statistically significant, indicating that LEZs decrease the life satisfaction of their inhabitants, despite their overall effectiveness at reducing the concentration of NO_2, PM_{10}, and CO, alongside their associated health benefits (Margaryan, 2021; Pestel and Wozny, 2019).

Excluding control individuals within 25 km distance to any LEZ increases the magnitude of point estimates relative to the full sample. These estimates imply that residing in a LEZ decreases life satisfaction by 0.19 points, or 2.7% of the average life satisfaction before implementation. The magnitude of the effect is quite substantial relative to other determinants of life satisfaction identified in the literature. For example, using SOEP data, Kassenboehmer and Haisken-DeNew (2009) find that becoming unemployed decreases life satisfaction by up to 0.9 points, suggesting that the impact of LEZ amounts to more than 20% of the unemployment effect. Further restricting the control group to persons living in similar-sized

[22]Specifically, we limit the control group to individuals in counties that are at least as densely populated as the LEZ county with the lowest population density (195 per km^2 in Marburg) and with greater GDP per capita than the LEZ county with the lowest GDP per capita (18,460 Euro in Bottrop).

counties halves the sample size, attenuates point estimates, and increases standard errors. Nevertheless we still observe a significant 0.14 point decrease in life satisfaction.

For all three samples, we provide evidence of parallel pre-trends.[23] Note that the estimates for the full sample might still suffer from potential spillover impacts, whereas the most restrictive sample is relatively small, which introduces the possibility of too low statistical power. As the sample with the 25 km buffer specification accounts for spillovers, is moderately sized and exhibits parallel pre-trends, in our further presentation we focus on results for this sample.[24]

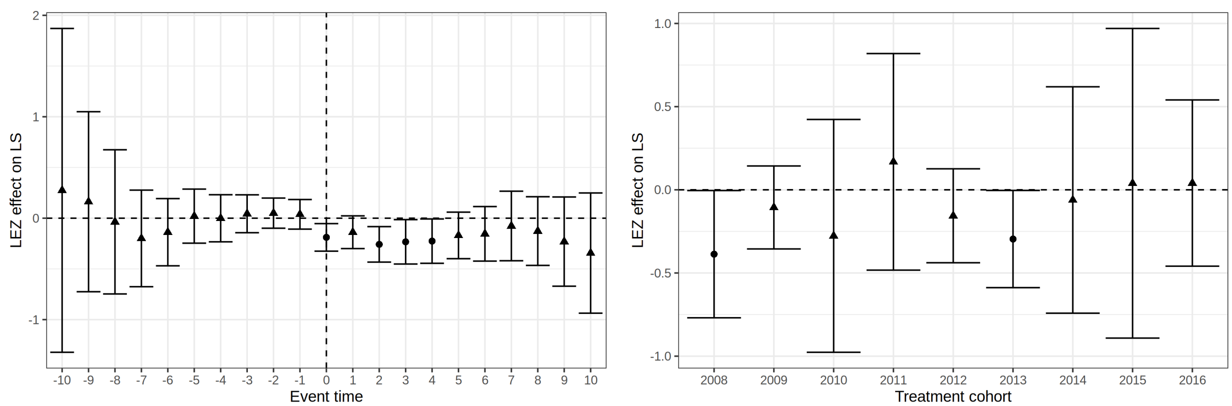


Figure 11: Dynamic and group-specific LEZ effects on life satisfaction

Notes: Dynamic (left panel) and group-specific (right panel) gtDD estimates of the impact of LEZs on life satisfaction of individuals living inside the LEZs. Dynamic effects refer to years before/after LEZ introduction; group-specific effects refer to treatment groups. The sample of control individuals is restricted to residences further than 25km away from the nearest LEZ. Standard errors clustered at the household level.

5.2. Dynamic and heterogeneous treatment effects

Next, we investigate the dynamic and heterogeneous effects by treatment groups. The left panel in Figure 11 plots event-time ATTs for the 25 km buffer sample. Results show a significant decrease in life satisfaction in the first year after the LEZ is activated. The negative impact on life satisfaction is slightly weaker in the second year, but it persists in subsequent years before rebounding back to insignificance. This indicates a transitory effect of LEZs on life satisfaction, potentially related to delayed policy acceptance.

[23]Figure 15 in Appendix A.6 depicts the dynamic event-time ATTs β_e for all samples: Indicative of parallel trends, none of the coefficients of the pre-treatment periods is significantly different from zero.

[24]Appendix A.7 lists the results on the most restrictive control group.

The right panel in Figure 11 depicts the aggregated ATTs for each treatment group. The most pronounced effect on life satisfaction occurs for the 2008 treatment group. The effects for all other groups are smaller and often statistically insignificant, indicating that the first LEZs drive the overall negative treatment effect. This is consistent with the novelty effect of LEZs – residents inside the early LEZs were the first to experience driving restrictions due to environmental policies in Germany. To enter a LEZ, cars need a colored windshield sticker indicating its emission category, such that every car owner had to order a sticker for a small fee – even for cars with the lowest emission intensity. Hence, policy acceptance might have been limited, especially for early zones, resulting in a more substantial impact on residents' well-being. In addition, individuals in early LEZs were more affected by the policy because the emission standards underlying the LEZs restricted a larger share of cars in earlier years, and car owners had to retrofit or replace their vehicles with higher likelihood. New vehicles usually comply with LEZ standards, so a renewal of the vehicle fleet results in fewer restrictions, which could also explain the null effect on life satisfaction from 2014 onwards.[25]

Table 12: Heterogeneous LEZ effects on life satisfaction

(a) By motor vehicle (MV) ownership

	With MV	Without MV
ATT	-0.186^{***} (0.058)	-0.162 (0.169)
N.Individuals	5227	709

(b) By diesel car ownership

	Diesel	Other Fuels
ATT	-0.276^{***} (0.099)	-0.151^{**} (0.071)
N.Individuals	1739	3473

(c) By income quartiles

	Q1	Q2	Q3	Q4
ATT	-0.186 (0.142)	-0.265^{**} (0.119)	-0.204^{*} (0.122)	-0.197^{**} (0.077)
N.Individuals	3145	3145	3146	3144

(d) By age groups

	$\geq$ 65 years old	$<$ 65 years old
ATT	-0.097 (0.112)	-0.240^{***} (0.063)
N.Individuals	4163	9953

Notes: Group-time difference-in-differences (gtDD) estimates of the impact of LEZs on life satisfaction using subsets of individuals living inside the LEZs. The control group consists of individuals in residences further than 25km away from the nearest LEZ. Subsamples are split based on motor vehicle ownership, diesel vehicle ownership, income quartiles and age groups. Point estimates represent the simple aggregation across all groups and time periods (β). Standard errors clustered at the household level. Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

To highlight potential mechanisms through which LEZs affect life satisfaction, we analyze heterogeneous effects for various sub-samples. In panel (a) of Table 12 we split the sample

[25] Results also hold when controlling for compositional changes of the treatment group at different event times ($e \in c(1,2,3,4,5)$).

based on motor vehicle ownership and estimate separate ATTs for persons with and without cars. Note that the SOEP only surveyed mobility-related information in 2015. Consequently, the sample split is based on cross-sectional differences. As expected, motor vehicle owners' life satisfaction declines significantly due to LEZs. The point estimate for the sample without motor vehicles is around the same magnitude as in the full sample, but imprecisely estimated because of the small sample size.

Panel (b) splits the subsample of motor vehicle owners into households with at least one diesel-fueled vehicle and households with vehicles using other fuel types (mostly gasoline, ethanol fuel mixtures, and natural gas). Individuals' life satisfaction decreases in both samples, but the effect on diesel-car owners is almost twice as large as in the other group. This is in line with LEZs targeting mostly diesel vehicles, as standards for diesel vehicles are more refined and stricter than for gasoline vehicles.[26] Consequently, diesel-car owners living inside a LEZ are more likely to be restricted by the policy, and might have to retrofit or replace their vehicle after policy implementation.

Panel (c) shows that the negative effect on life satisfaction persists at similar magnitudes across all income quartiles.[27] Still, the impact is lowest in the first income quartile, which is plausible given that more than a third of its households do not own a motor vehicle; this sample consists of low-paid workers, retirees, and welfare recipients who typically cannot afford a car. The largest impact occurs in the second quartile, where the share of vehicle owners is much higher at over 90%, hence this quartile is more likely to be affected by the policy. Effects on the third and fourth quartiles are lower than in the second, potentially because these individuals can more easily adapt to the driving restrictions due to greater financial resources. Nevertheless, LEZs still have a statistically significant negative effect on life satisfaction, even for high-income persons.

The last panel in Table 12 lists results across age subgroups. We focus on a cutoff of 65 years. People older than 65 are retirees or near retirement - they are less likely to own a car and, on average, own fewer cars relative to working-age individuals. In contrast, people younger than 65 have very different mobility requirements, e.g., due to children in

[26] Table 16 in Appendix A.1 lists the relevant standards for both fuel types.

[27] Quartiles are based on annual net household income (after taxes) averaged over the whole sample period; the cutoffs are [0; 27,000], (27,000; 34,300], (34,300; 48,700] and $> 48{,}700$ Euros.

the household. These differences appear in the point estimates, with strong effects on life satisfaction for individuals younger than 65 and no significant effect with a smaller point estimate for people aged 65 or older.

5.3. Health outcomes

We complement our results by looking at the impact of LEZs on objective health measures and utilization of the health care system as proxied by the number of doctor visits. Since 2009, the SOEP bi-annually surveys several illness categories and the number of doctor visits within the last twelve months prior to the interview. We restrict our sample to individuals and time periods where information on specific illnesses is available.[28] Since the recent empirical literature shows that LEZs decrease cardiovascular diseases (Margaryan, 2021; Pestel and Wozny, 2019), we focus on hypertension as a risk factor for cardiovascular conditions. Moreover, as a falsification test, we analyze the effect of LEZs on cancer. Cancer often develops over long time periods and it is unlikely that the marginal pollution effect would trigger changes in the cancer rate.

Table 13: LEZ effect on health care utilization and health outcomes

	LS	Doctor visits	Hypertension	Cancer
ATT	-0.163^{*}	-1.292	-0.046^{**}	0.010
	(0.087)	(0.938)	(0.022)	(0.013)
N	9218	9208	9218	9218
G	4	4	4	4
T	5	5	5	5

Notes: Group-time difference-in-differences (gtDD) estimates of the impact of LEZs on life satisfaction and objective health outcomes of individuals living inside the LEZs. The control group consists of individuals in residences further than 25km away from the nearest LEZ. Point estimates represent the simple aggregation across all groups and time periods (β). Standard errors clustered at the household level. Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 13 reports results for the LEZ effect on life satisfaction, doctor visits, hypertension, and cancer. The negative LEZ effect on life satisfaction persists in this smaller sample with a magnitude comparable to our baseline results. There is suggestive evidence that LEZs decrease the number of doctor visits, although we cannot establish statistical significance at conventional levels. Concerning objective health outcomes, we estimate a significant decrease

[28]Since health outcomes are only available since 2009, we drop individuals initially treated in 2008 and 2009 because we do not observe their pre-treatment outcomes.

in the likelihood of hypertension.[29] The probability of developing hypertension drops by 4.6 percent after implementation.[30] The point estimate of the LEZ effect on the probability of developing cancer is, as expected, statistically insignificant.

Lastly, we validate our results by analyzing heterogeneous health effects for different age groups in Table 14. The decrease in number of doctor visits is especially pronounced for middle-aged adults and older individuals, with almost two avoided doctor visits per year for individuals in these age groups. The decrease in the probability of developing hypertension is visible across all age groups, with the effect becoming stronger for older individuals. People aged 60 to 80 years benefit the most from LEZs, as their probability for hypertension decreases by 8.2%, which is in line with recent empirical results that health benefits of LEZs accrue mostly to the older population (Margaryan, 2021). The probability of developing cancer is not significantly affected by LEZ implementation in any age group.

Table 14: LEZ effect on health outcomes by age groups

(a) Effect on annual doctor visits

	Age [20, 40)	Age [40, 60)	Age [60, 80)
ATT	−0.081	−1.830*	−1.786
	(1.268)	(1.105)	(2.103)
N.Individuals	2519	4520	3618
N.Groups	4	4	4
N.Periods	5	5	5

(b) Effect on hypertension

	Age [20, 40)	Age [40, 60)	Age [60, 80)
ATT	−0.018	−0.031	−0.082*
	(0.038)	(0.031)	(0.043)
N.Individuals	2054	4256	3613
N.Groups	4	4	4
N.Periods	5	5	5

(c) Effect on cancer

	Age [20, 40)	Age [40, 60)	Age [60, 80)
ATT	−0.005	0.003	0.023
	(0.004)	(0.015)	(0.030)
N.Individuals	2054	4256	3613
N.Groups	4	4	4
N.Periods	5	5	5

Notes: Group-time difference-in-differences (gtDD) estimates of the impact of LEZs on objective health outcomes of individuals living inside the LEZs. The control group consists of individuals in residences further than 25km away from the nearest LEZ. Point estimates represent the simple aggregation across all groups and time periods (β). Subsamples are split based on age groups. Standard errors clustered at the household level. Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

[29] Figure17 in Appendix A.6 shows that the effect on hypertension is immediate, manifesting itself in the first year after LEZ implementation, and point estimates are rather stable in later years, despite increasing standard errors.

[30] Using the pre-treatment hypertension rate of 31 percentage points among LEZ individuals (cp. Table 18) to calculate the potential reduction in hypertension cases yields an estimate of 1.4 percentage points. Given that more than 6.6 million people lived inside a LEZ in 2018, a simple back-of-the-envelope calculation suggests that these driving restrictions avoided at least 93,000 hypertension cases in Germany.

Overall, our results suggest that LEZs decrease subjective well-being of individuals living inside their borders. This negative life satisfaction effect is immediate but transitory as it disappears several years after policy implementation. We show evidence that individuals residing in the early LEZs drive the overall effect, potentially due to a novelty effect of driving restrictions in Germany, and due to compositional changes in the vehicle fleet making the policy less binding in later years. Furthermore, adverse life satisfaction effects are heterogeneous and especially pronounced among diesel car owners and younger individuals, who are more likely to be affected by the policy. In addition, we provide evidence that LEZs decrease the likelihood of developing hypertension, mostly in the older population, though these health benefits do not seem sufficient to counteract the drop in life satisfaction due to LEZ implementation.

6. Conclusion

This paper analyzes the effectiveness and spillover effects of low-emission zones (LEZs) on air pollution, as well as the impact on individual-level outcomes. Concerning pollution, we consider LEZs' influence on the concentration of carbon monoxide (CO), nitrogen dioxide (NO_2), and coarse particulate matter (PM_{10}), both within and outside their borders. Estimating the effect outside the LEZs' borders allows us to examine policy spillovers to neighboring regions. Moreover, we analyze whether the policy's influence on the concentration of traffic contaminants has the unintended result of raising the level ozone (O_3) because of the lower-atmosphere chemical interaction between O_3 and NO_2. This is the first contribution analyzing the effect of the policy on CO, a good proxy for traffic pollution. Moreover, this is the first paper to evaluate heterogeneous effects of the policy by season. Regarding individual-level outcomes, we analyze the impact of LEZs on self-reported life satisfaction and health outcomes as captured by doctor visits and hypertension cases.

To identify the effect of the policy on pollution measuring stations and individuals, we pursue a two-step quasi-experimental design that handles the triple challenge of accounting for the staggered implementation of LEZs, time-varying treatment effects, and policy spillovers. First, using a regression discontinuity (RD) strategy we estimate the local average treatment effect (LATE) of LEZs on daily pollution levels within a narrow time window around its implementation date for stations inside and outside the zones' borders. The LATE for stations outside the zone helps us identify its external influence area. Second, we use the results on spatial spillovers from the RD design to constrain the control group of a group-time difference-in-differences (gtDD) model that robustly identifies the average treatment effect on the treated for pollution, life satisfaction, and health outcomes in our setting with staggered implementations and time-varying treatment effects.

RD estimates show that LEZs reduce CO by 12.1%, NO_2 by 10.7%, and PM_{10} by 15%. Moreover, we find that LEZs affect pollution both within the zones' limit and in adjacent areas. PM_{10}, NO_2 and ozone (O_3) levels increase for stations within 500 meters from LEZs' borders, suggesting some traffic displacement to major roads. In contrast, PM_{10}, NO_2 and CO decrease further away from the LEZ border, up to a distance of 25 km.

Based on gtDD estimates we show that the introduction of LEZs consistently reduces PM_{10} and NO_2 while increasing the concentrations of O_3, by about 1.1 $\mu g/m^3$. Our estimate of the decrease in NO_2 levels by about 3.8 $\mu g/m^3$ is more than 100% larger in magnitude than results in the existing literature, while our estimates for PM_{10} of a reduction by 2 $\mu g/m^3$ are also larger than findings in the majority of the literature. These findings suggest that a two-way fixed effects difference-in-differences (TWFE-DD) strategy significantly underestimates the effectiveness of LEZs due to the problems arising from the staggered implementation and time-varying treatment effects. We confirm this intuition using the Goodman-Bacon decomposition (Goodman-Bacon, 2018).

Additionally, we show that results are heterogeneous by season, with LEZs being particularly effective at decreasing traffic-related pollutants during the winter. At the same time, the effect of LEZs on O_3 focuses on the spring and summer months when lower levels of NO_2 and sunnier days increase O_3 formation. Concerning spillovers based on the gtDD model, we consistently find that O_3 increases at distances up to 25 km away from LEZs. On average, no spillovers are detected for the other contaminants, although with suggestive evidence of seasonal spillovers for PM_{10}.

Our analysis of individual-level outcomes reveals that persons dwelling inside LEZs experience a substantive decrease in their life satisfaction, with average effects ranging from -0.19

to -0.14 points, about 15% to 21% of the unemployment effect on life satisfaction obtained in previous studies based on SOEP data. Owners of diesel cars and younger individuals under 65 years experience even larger decreases. The former are more likely to be affected by the policy that restricts mostly diesel cars, and the latter are more likely to own a car since they have greater mobility needs than the older population. Furthermore, these well-being effects are strongest in the first years of policy implementation and are transitory.

With respect to health outcomes, we observe a significant drop in hypertension cases after LEZ implementation. These health benefits mostly accrue to people aged between 60 and 80 years, while the effects for younger people are less pronounced. Furthermore, we show suggestive evidence that the number of annual doctor visits, as a proxy for health care utilization, decreases due to LEZs. Nevertheless, these health benefits, which should increase life satisfaction, are not sufficient, at least on average, to counteract the negative well-being effects of LEZs. Taken together, our analysis with respect to overall well-being suggests that the impacts of driving restrictions are mainly borne by younger individuals, whose share of vehicle ownership is greater than for older people. The health benefits, on the other hand, accrue more strongly to the older population.

Our results confirm that LEZs are an effective air pollution mitigation policy while also pointing towards the importance of considering its impact on bordering areas and secondary contaminants like O_3. Cities facing ozone problems like Los Angeles, Mexico City, or Delhi should consider the increase in O_3 induced by LEZs before implementing this kind of driving restriction policies. Compared to the mixed evidence on the alternate-day driving literature, LEZs are more effective at reducing traffic-related air pollution. One advantage of LEZs is that restricting traffic based on pollution intensity prevents the substitution of restricted vehicles with more emission-intensive ones. Furthermore, our finding of a decrease in average life satisfaction due to the LEZs' introduction shows that such policies generate adverse well-being effects despite their health benefits.

References

Barahona N., Gallego F. A., and Montero J.-P. (2020) Vintage-specific driving restrictions. The Review of Economic Studies, 87(4): 1646–1682.

Callaway B. and Sant'Anna P. H. (2020) Difference-in-differences with multiple time periods. Journal of Econometrics.

Calonico S., Cattaneo M. D., and Farrell M. H. (2020) Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. The Econometrics Journal, 23(2): 192–210.

Cattaneo M. D., Idrobo N., and Titiunik R. (2019) A practical introduction to regression discontinuity designs: Foundations: Cambridge University Press.

Chang T. Y., Graff Zivin J., Gross T., and Neidell M. (2019) The effect of pollution on worker productivity: evidence from call center workers in China. American Economic Journal: Applied Economics, 11(1): 151–72.

Chen Y. and Whalley A. (2012) Green infrastructure: The effects of urban rail transit on air quality. American Economic Journal: Economic Policy, 4(1): 58–97.

Chen Y., Jin G. Z., Kumar N., and Shi G. (2013) The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality. Journal of Environmental Economics and Management, 66(3): 424–443.

Davis L. W. (2008) The effect of driving restrictions on air quality in Mexico City. Journal of Political Economy, 116(1): 38–81.

——— (2017) Saturday driving restrictions fail to improve air quality in Mexico City. Scientific reports, 7(1): 1–9.

De Chaisemartin C. and d'Haultfoeuille X. (2020) Two-way fixed effects estimators with heterogeneous treatment effects. American Economic Review, 110(9): 2964–96.

Ellison R. B., Greaves S. P., and Hensher D. A. (2013) Five years of London s low emission zone: Effects on vehicle fleet composition and air quality. Transportation Research Part D: Transport and Environment, 23: 25–33.

EU (1999) Directive 1999/30/EC of 22 April 1999 relating to limit values for sulphur dioxide, nitrogen dioxide and oxides of nitrogen, particulate matter and lead in ambient air.Technical report, European Parliament.

——— (2000) Directive 2000/69/EC of the European Parliament and of the Council of 16 November 2000 relating to limit values for benzene and carbon monoxide in ambient air.Technical report, European Parliament.

——— (2002) Directive 2002/3/EC: of the European Parliament and of the Council of 12 February 2002 relating to ozone in ambient air.Technical report, European Parliament.

——— (2008) Directive 2008/50/EC: of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe.Technical report, European Parliament.

Gallego F., Montero J.-P., and Salas C. (2013) The effect of transport policies on car use: Evidence from Latin American cities. Journal of Public Economics, 107: 47–62.

Gehrsitz M. (2017) The effect of low emission zones on air pollution and infant health. Journal of Environmental Economics and Management, 83: 121–144.

Goodman-Bacon A. (2018) Difference-in-differences with variation in treatment timing.Technical report, National Bureau of Economic Research.

Jayachandran S. (2009) Air quality and early-life mortality evidence from Indonesia's wildfires. Journal of Human Resources, 44(4): 916–954.

Kassenboehmer S. C. and Haisken-DeNew J. P. (2009) You're Fired! The Causal Negative Effect of Entry Unemployment on Life Satisfaction. The Economic Journal, 119(536): 448–462.

Knittel C. R., Miller D. L., and Sanders N. J. (2016) Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. Review of Economics and Statistics, 98(2): 350–366.

Leape J. (2006) The London congestion charge. Journal of Economic Perspectives, 20(4): 157–176.

Malina C. and Scheffler F. (2015) The impact of low emission zones on particulate matter concentration and public health. Transportation Research Part A: Policy and Practice, 77: 372–385.

Margaryan S. (2021) Low emission zones and population health. Journal of Health Economics, 76: 102402.

Monks P., Archibald A., Colette A., Cooper O., Coyle M., Derwent R., Fowler D., Granier C., Law K., Stevenson D. et al. (2015) Tropospheric ozone and its precursors from the urban to the global scale from air quality to short-lived climate forcer.. Atmospheric Chemistry & Physics Discussions, 15: 8889–8973.

Pestel N. and Wozny F. (2019) Low Emission Zones for Better Health: Evidence from German Hospitals. IZA Discussion Paper No. 12545.

Rivera N. (2021) Air quality warnings and temporary driving bans: Evidence from air pollution, car trips, and mass-transit ridership in Santiago. Journal of Environmental Economics and Management (Forthcomming).

Rohlf A., Holub F., Koch N., and Ritter N. (2020) The effect of clean air on pharmaceutical expenditures. Economics Letters, 192: 109221.

Sarmiento L. (2020) Waiting for My Sentence: Air Pollution and the Productivity of Court Rulings. DIW Berlin Discussion Paper.

Simeonova E., Currie J., Nilsson P., and Walker R. (2019) Congestion pricing, air pollution, and children's health. Journal of Human Resources: 0218–9363R2.

Suarez-Bertoa R. and Astorga C. (2018) Impact of cold temperature on Euro 6 passenger car emissions. Environmental pollution, 234: 318–329.

Viard V. B. and Fu S. (2015) The effect of Beijing's driving restrictions on pollution and economic activity. Journal of Public Economics, 125: 98–115.

Wang L., Xu J., and Qin P. (2014) Will a driving restriction policy reduce car trips?—The case study of Beijing, China. Transportation Research Part A: Policy and Practice, 67: 279–290.

Wolff H. (2014) Keep Your Clunker in the Suburb: Low-emission Zones and Adoption of Green Vehicles. The Economic Journal, 124(578): F481–F512.

Zhai M. and Wolff H. (2021) Air pollution and urban road transport: evidence from the world's largest low-emission zone in London. Environmental Economics and Policy Studies: 1–28.

Zivin J. G. and Neidell M. (2012) The impact of pollution on worker productivity. The American Economic Review, 102(7): 3652–3673.

A. Appendix

A.1. Introduction to LEZs

Policymakers in the European Union minimize air pollution's health risks by setting limit values to urban centers.[31] The limit values are legally binding. In case of non-attainment, the member states must propose and implement action plans to reduce the risk or duration of future limit violations; if the member states fail to implement sufficient measures to reduce pollution, repeated non-attainment results in financial penalties. Table 15 portrays current exposure limits in the European Union.

Table 15: EU air quality regulations

Pollutant	Concentration	Avg. period	Legal nature	Permitted exceedance days per year
CO	10 mg/m^3	Max. daily 8 h mean	Limit value as of 1.1.2005	NA
NO_2	200 μg/m^3	1 hour	Limit value as of 1.1.2010	18
NO_2	40 μg/m^3	1 year	Limit value as of 1.1.2010	NA
O_3	120 μg/m^3	Max. daily 8 h mean	Target value as of 1.1.2010	25 days averaged over 3 years
PM_{10}	50μg/m^3	24 hours	Limit value as of 1.1.2005	35
PM_{10}	40μg/m^3	1 year	Limit value as of 1.1.2005	NA

Notes: Source: EU (2008).

Germany implemented the 22nd Ordinance of the Federal Immission Control Act (Bundes-Immissionsschutzgesetz - BImSchG) to comply with EU legislation. This law made EU limit values legally binding as of January 2005. In the following years, many cities could not adhere to the limit values for NO_2 and PM_{10}. Between 2005 and 2007, 89 urban centers violated the daily PM_{10} limit of 50 μg/m^3 on more than 35 days in at least one year,[32] and 54 cities exceeded the annual NO_2 limit of 40 μg/m^3 for at least one year. Consequently, German federal states and local administrations had to draw up action plans for improving air quality. These action plans targeted traffic exhaust-related pollutants and commonly involved introducing low emission zones (LEZ).

[31] Directive 1999/30/EC (EU, 1999) define permissible concentrations for NO_2, SO_2 and PM_{10}, Directive 2000/69/EC (EU, 2000) set limits for carbon monoxide (CO), and Directive 2002/3/EC (EU, 2002) focuses on O_3. These legislations were revised in 2008 and unified into the single Directive 2008/50/EC (EU, 2008) that defines current limit values and detailed measurement procedures for all criteria pollutants.

[32] Among these, 52 were large cities with more than 100,000 inhabitants, which amounts to 65 percent of all large cities in Germany.

The 2007 Immission Control Act (35th BImSchV) provides the legal basis for introducing low emission zones by giving state and local governments the right to restrict access to specific city areas for cars not complying with predefined emission standards. Germany enforces LEZs through colored stickers on car's windshields: Only automobiles with a specifically colored label can enter the LEZ. Red stickers represent the highest emitting vehicles, and green stickers the least emitting ones; Table 16 lists details on the stringency of emission standards and stickers. The policy is enforced by police and municipal authorities and infringement results in fines for the vehicle driver.

Table 16: Relevant emission standards for LEZ sticker categories

	No Sticker	Red	Yellow	Green
Diesel	Euro 1 or older	Euro 2 / Euro 1 with particle filter	Euro 3 / Euro 2 with particle filter	Euro 4 or better / Euro 3 with particle filter
Gasoline	Without catalytic converter	–	–	Euro 1 with catalytic converter or better

Notes: Relevant emission standards for LEZ sticker categories defined in the Ordinance on the marking of vehicles (35th BImSchV). The Euro standards represent the EU emission regulations for new light duty vehicles based on Directive 70/220/EEC and its amendments.

A.2. Why ground-level ozone is special

Ground-level ozone differs from the other criteria pollutants in that it is a secondary pollutant, i.e., it requires precursors as it is created through the interaction of solar radiation with nitrous oxides (NO_x) and volatile organic compounds (VOC), whose concentration in the air is much increased beyond natural levels by the combustion of fossil fuels. Road traffic is one of the major causes of this increase.

The relationship between the concentrations of ozone and its precursors is complex, as there are two sides to the interaction. On the one hand, the interaction of solar radiation with NO_x and VOC forms ozone. On the other, it also degrades it. The balance between the two sides leads to patterns in ozone concentrations that deviate from those of other criteria pollutants: In areas with high levels of precursor pollution, such as urban centers with dense vehicle traffic, ozone concentrations are lower than in suburban areas. The reason is that at high concentrations of precursor pollutants, ozone degrades faster than it is formed, whereas the formation process dominates at lower levels of precursors, e.g., in rural areas. This phenomenon is sometimes referred to as the ”ozone paradox” (Monks et al., 2015).

A.3. Data

Table 17: List of Low Emission Zones in Germany

State	Zone	Level	Level 1	Level 2	Level 3	Expansion
BW	Asperg - Ludwigsburg	3	01.01.13	01.01.13	01.01.13	No
BW	Balingen	3	01.04.17	01.04.17	01.04.17	No
BW	Bietigheim-Bissingen - Ludwigsburg	3	01.01.13	01.01.13	01.01.13	No
BW	Freiberg am Neckar - Ludwigsburg	3	01.01.12	01.01.12	01.01.13	Yes
BW	Freiburg im Breisgau	3	01.01.10	01.01.12	01.01.13	No
BW	Heidelberg	3	01.01.10	01.01.12	01.01.13	No
BW	Heidenheim	3	01.01.12	01.01.12	01.01.13	No
BW	Heilbronn	3	01.01.09	01.01.12	01.01.13	No
BW	Herrenberg	3	01.01.09	01.01.12	01.01.13	No
BW	Ilsfeld	3	01.03.08	01.01.12	01.01.13	No
BW	Ingersheim - Ludwigsburg	3	01.01.12	01.01.12	01.01.13	Yes
BW	Karlsruhe	3	01.01.09	01.01.12	01.01.13	No
BW	Kornwestheim - Ludwigsburg	3	01.01.13	01.01.13	01.01.13	No
BW	Leonberg - Leonberg	3	01.03.08	01.01.12	01.01.13	Yes
BW	Leonberg und Umgebung	3	02.12.13	02.12.13	02.12.13	No
BW	Ludwigsburg - Ludwigsburg	3	01.03.08	01.01.12	01.01.13	Yes
BW	Ludwigsburg und Umgebung	3	01.01.13	01.01.13	01.01.13	No
BW	Mannheim	3	01.03.08	01.01.12	01.01.13	No
BW	Markgroeningen - Ludwigsburg	3	01.07.11	01.07.11	01.01.13	Yes
BW	Moeglingen - Ludwigsburg	3	01.01.13	01.01.13	01.01.13	No
BW	Muehlacker	3	01.01.09	01.01.12	01.01.13	No
BW	Pfinztal	3	01.01.10	01.01.12	01.01.13	No
BW	Pforzheim	3	01.01.09	01.01.12	01.01.13	No
BW	Pleidelsheim - Ludwigsburg	3	01.07.08	01.01.12	01.01.13	Yes
BW	Remseck am Neckar - Ludwigsburg	3	01.01.17	01.01.17	01.01.17	No
BW	Reutlingen 1	3	01.03.08	01.01.12	01.01.13	Yes
BW	Reutlingen 2	3	01.02.15	01.02.15	01.02.15	No
BW	Schramberg	3	01.07.13	01.07.13	01.01.15	No
BW	Schwaebisch Gmuend	3	01.03.08	01.01.12	01.01.13	No
BW	Stuttgart	3	01.03.08	01.07.10	01.01.12	No
BW	Tamm - Ludwigsburg	3	01.01.13	01.01.13	01.01.13	No
BW	Tuebingen 1	3	01.03.08	01.01.12	01.01.13	Yes
BW	Tuebingen 2	3	01.01.15	01.01.15	01.01.15	No
BW	Ulm	3	01.01.09	01.01.12	01.01.13	No
BW	Urbach	3	01.01.12	01.01.12	01.01.13	No
BW	Wendlingen	3	02.04.13	02.04.13	02.04.13	No
BY	Augsburg	3	01.07.09	01.01.11	01.06.16	No
BY	Muenchen	3	01.10.08	01.10.10	01.10.12	No
BY	Neu Ulm	2	01.11.09	05.11.12		No
BY	Regensburg	3	15.01.18	15.01.18	15.01.18	No
BE	Berlin	3	01.01.08	01.01.10	01.01.10	No
BR	Bremen	3	01.01.09	01.01.10	01.07.11	No
HE	Darmstadt	3	01.11.15	01.11.15	01.11.15	No
HE	Frankfurt a.M.	3	01.10.08	01.01.10	01.01.12	No
HE	Limburg an der Lahn	3	31.01.18	31.01.18	31.01.18	No
HE	Marburg	3	01.04.16	01.04.16	01.04.16	No
HE	Offenbach	3	01.01.15	01.01.15	01.01.15	No
HE	Wiesbaden	3	01.02.13	01.02.13	01.02.13	No
NS	Hannover	3	01.01.08	01.01.09	01.01.10	No

Table 17 continued from previous page

NS	Osnabrueck	3	04.01.10	03.01.11	03.01.12	No
NW	Aachen	3	01.02.16	01.02.16	01.02.16	No
NW	Bochum - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Bonn	3	01.01.10	01.07.12	01.07.14	No
NW	Bottrop - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Castrop-Rauxel	3	01.01.12	01.01.13	01.07.14	No
NW	Dinslaken	3	01.07.11	01.07.11	01.10.12	No
NW	Dortmund - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Duisburg - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Duesseldorf	3	15.02.09	01.03.11	01.07.14	No
NW	Eschweiler	3	01.06.16	01.06.16	01.06.16	No
NW	Essen - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Gelsenkirchen - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Gladbeck	3	01.01.12	01.01.13	01.07.14	No
NW	Hagen	3	01.01.12	01.01.13	01.07.14	No
NW	Herne	3	01.01.12	01.01.13	01.07.14	No
NW	Herten	3	01.01.12	01.01.13	01.07.14	No
NW	Koeln 1	3	01.01.08	01.01.13	01.07.14	Yes
NW	Koeln 2	3	01.04.12	01.01.13	01.07.14	
NW	Krefeld	3	01.01.11	01.01.11	01.07.12	No
NW	Langenfeld	3	01.01.13	01.01.13	01.07.14	No
NW	Moenchengladbach	3	01.01.13	01.01.13	01.07.14	No
NW	Muelheim	3	01.10.08	01.01.13	01.07.14	Yes
NW	Muenster	3	01.01.10	01.01.10	01.01.15	No
NW	Neuss	3	15.02.10	01.03.11	01.07.14	No
NW	Oberhausen - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Overath	3	01.10.17	01.10.17	01.10.17	No
NW	Recklinghausen - Ruhrgebiet	3	01.10.08	01.01.13	01.07.14	Yes
NW	Remscheid	3	01.01.13	01.01.13	01.07.14	No
NW	Ruhrgebiet	3	01.01.12	01.01.13	01.07.14	No
NW	Siegen	3	01.01.15	01.01.15	01.01.15	No
NW	Wuppertal	3	15.02.09	01.03.11	01.07.14	No
RP	Mainz	3	01.02.13	01.02.13	01.02.13	No
SN	Leipzig	3	01.03.11	01.03.11	01.03.11	No
ST	Halle	3	01.09.11	01.09.11	01.01.13	No
ST	Magdeburg	3	01.09.11	01.09.11	01.01.13	No
TH	Erfurt	3	01.10.12	01.10.12	01.10.12	No

Table 18: Descriptive statistics on SOEP individuals living inside LEZ

	Total	Before	After	Diff	t-stat	
Life satisfaction [0-10]	7.12	7.09	7.14	0.04	-1.44	
Age [years]	54.30	52.49	55.65	3.15	-10.63	***
Is female [%]	0.53	0.53	0.53	0.00	-0.29	
Is employed [%]	0.56	0.54	0.57	0.02	-2.39	*
Income [Thsd Euro]	45.13	44.24	45.79	1.55	-2.46	*
Education [years]	12.90	12.70	13.04	0.34	-6.16	***
Number children	0.44	0.51	0.39	-0.12	7.30	***
Owns motor vehicle [%]	0.81	0.83	0.80	-0.03	4.14	***
Number motor vehicles	1.26	1.28	1.26	-0.02	0.99	
Owns diesel car [%]	0.32	0.33	0.31	-0.02	2.13	*
Number doctor visits	11.06	11.22	10.94	-0.28	1.02	
Has or had hypertension [%]	0.31	0.31	0.31	-0.00	0.14	
Has or had cancer [%]	0.06	0.05	0.06	0.01	-1.12	
GDP p.c. [Thsd Euro]	44.34	40.66	47.08	6.42	-19.69	***
Population density [per km^2]	2069.52	1950.37	2158.34	207.96	-11.63	***

Notes: This table lists pre- and post-treatment averages of all individuals living inside a LEZ for the sample period 2005 to 2018.

A.4. Regression discontinuity design, robustness checks

Table 19: Effect of the introduction of LEZ on air pollution (levels)

(a) Carbon monoxide

	(1)	(2)	(3)
	-0.009	-0.058^{***}	-0.000
	(0.020)	(0.020)	(0.018)
Bwd	52	44	71
N	4066	3445	5543

(b) Nitrogen dioxide

	(1)	(2)	(3)
	-2.543^{*}	-5.140^{***}	-4.360^{***}
	(1.320)	(1.299)	(1.235)
Bwd	39	36	40
N	7172	6602	7363

(c) Ozone

	(1)	(2)	(3)
	2.495	1.653	-1.059
	(1.725)	(1.593)	(1.163)
Bwd	38	39	61
N	2679	2750	4319

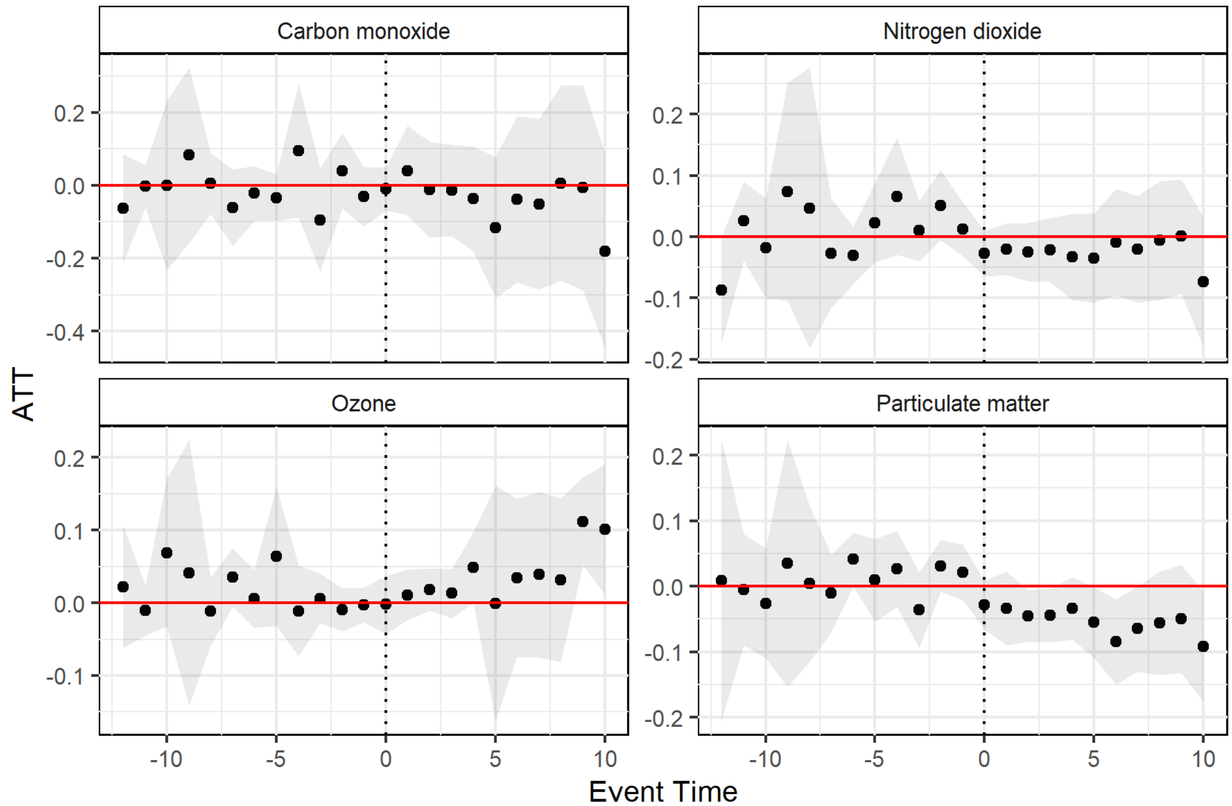
(d) Coarse particulate matter

	(1)	(2)	(3)
	-3.148^{**}	-4.641^{***}	-1.669^{*}
	(1.270)	(1.240)	(0.910)
Bwd	31	29	48
N	4810	4651	7542

Notes: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. RD estimates on the impact of low emission zones on the concentration of coarse particulate matter, nitrogen dioxide, carbon monoxide, and ozone. Optimal bandwidths are selected based on the MSE optimization proposed by Calonico et al. (2014).

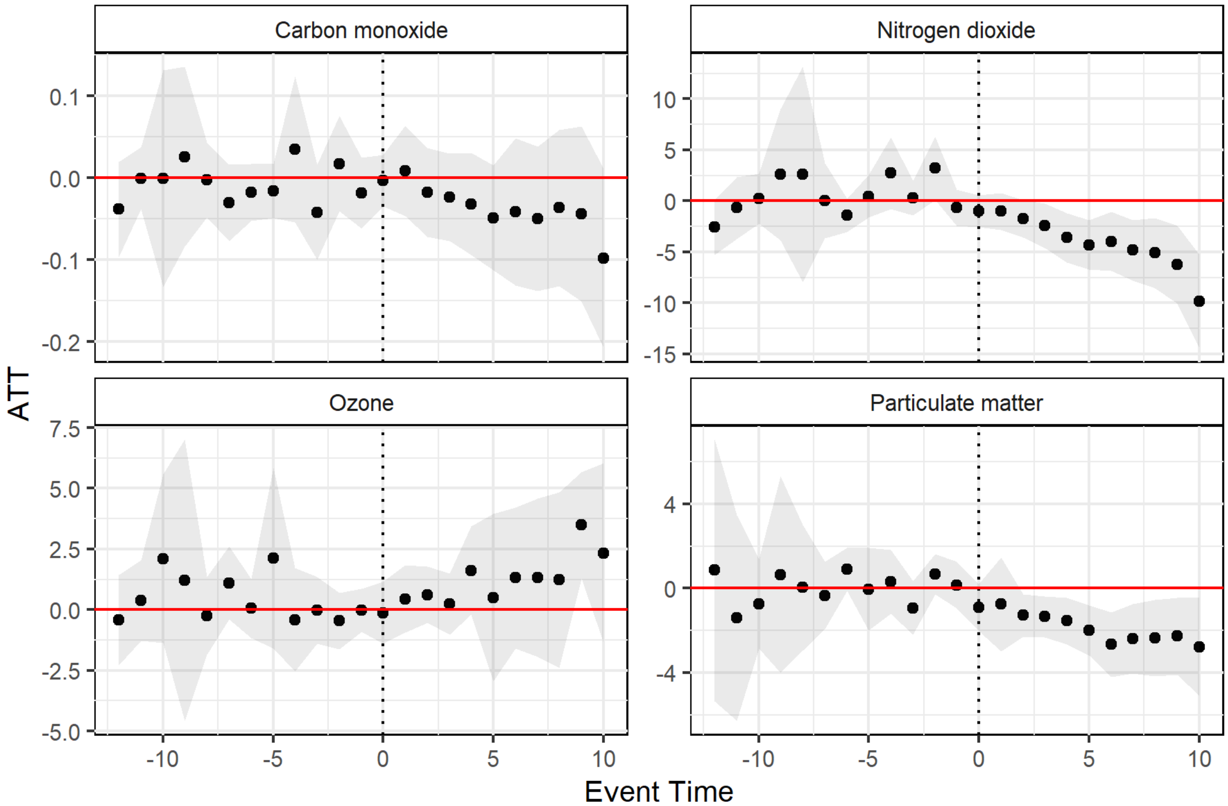
A.5. Group-time difference-in-differences design, robustness checks

Figure 12: Event-time ATTs for the buffer specification



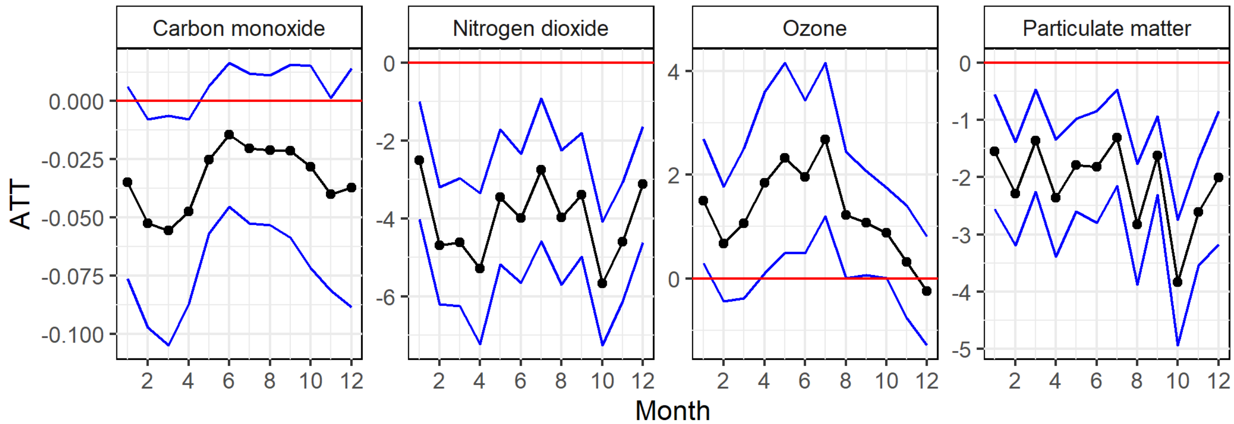
Notes: Event-time group-time difference-in-differences (gtDD) estimates (β_e) of the impact of LEZs on yearly air pollution levels for the buffer design. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border. Grey ribbons represent 95% confidence intervals. The gtDD model controls for station and year fixed effects. Standard errors clustered at the station level.

Figure 13: Event-time ATTs for the raw specification



Notes: Event-time group-time difference-in-differences (gtDD) estimates (β_e) of the impact of LEZs on yearly air pollution levels for the buffer design. Treated stations are all those stations inside the zone, and control stations all stations outside the zone. Grey ribbons represent 95% confidence intervals. The gtDD model controls for station and year fixed effects. Standard errors clustered at the station level.

Figure 14: Event time average treatment effects by month of the year

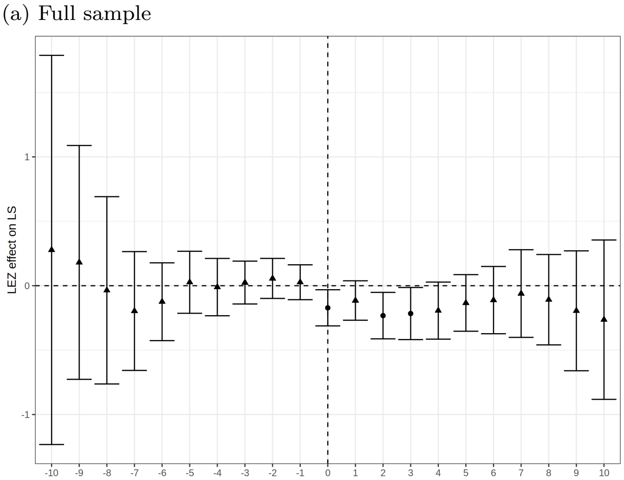


Notes: Group-time difference-in-differences (gtDD) estimates (β) of the impact of LEZs on monthly air pollution levels for the doughnut design. Treated stations are all those stations inside the zone, and control stations all stations further away than 25 km from the zone's border and up until 75 km. Blue lines represent 95% confidence intervals. The gtDD model controls for station and year fixed effects. Standard errors clustered at the station level.

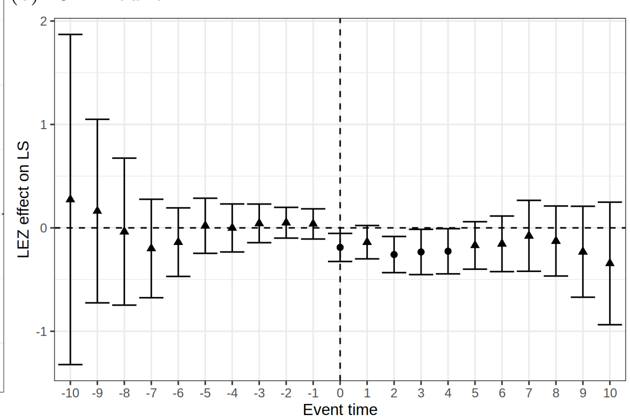
A.6. Effect of LEZs on well-being and health outcomes, additional results and robustness checks

Figure 15: Dynamic effects of LEZs on life satisfaction across samples

(a) Full sample



(b) 25 km buffer



(c) 25 km buffer, similar counties

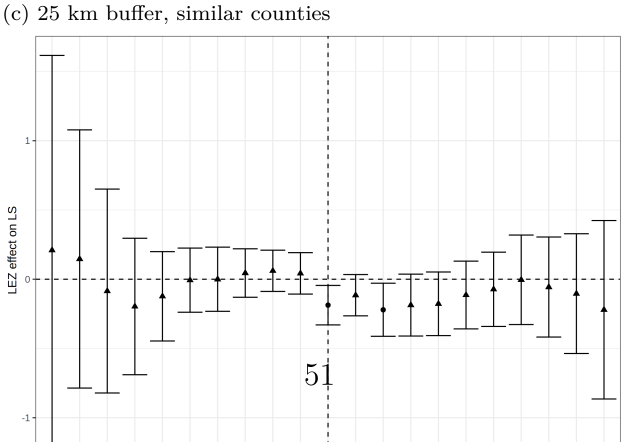
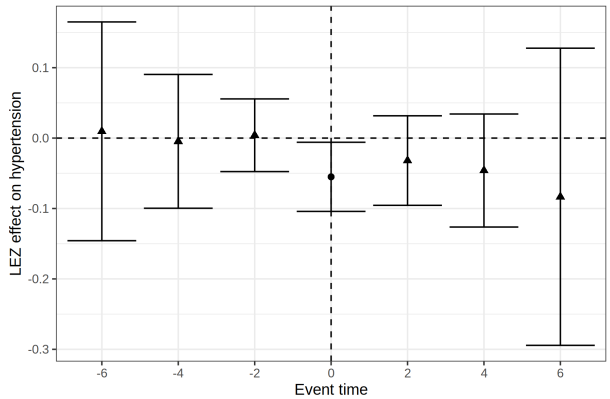


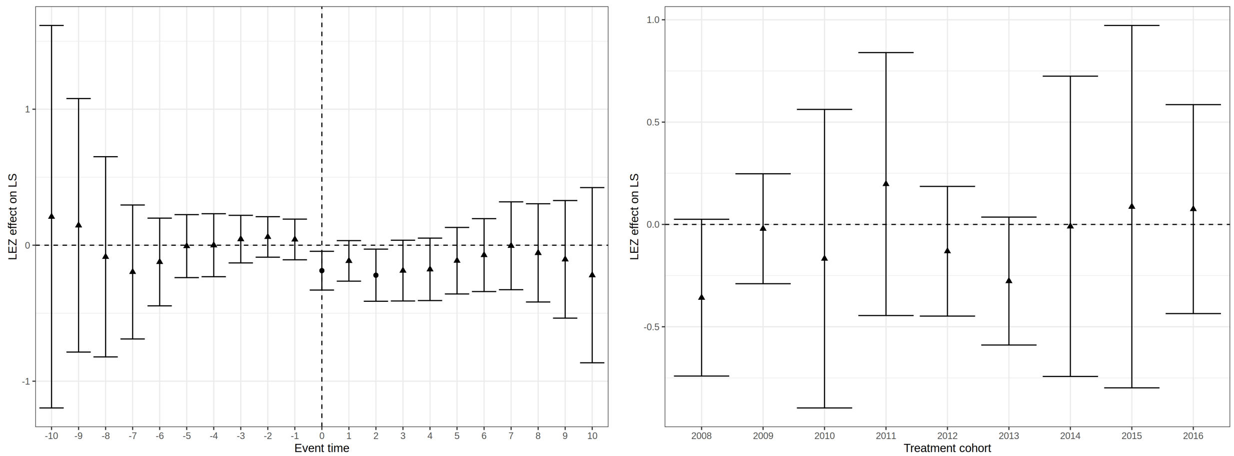
Figure 17: Dynamic LEZ effects on hypertension



Notes: Dynamic gtDD estimates (β_e) of the impact of LEZs on the probability to develop hypertension of individuals living inside the LEZs. Dynamic effects refer to years before/after LEZ introduction.The sample of control individuals is restricted to residences further than 25km away from the nearest LEZ. Standard errors clustered at the household level.

A.7. Effect of LEZs on well-being and health outcomes, additional results using alternative control group

Figure 18: Dynamic and cohort-specific LEZ effects on life satisfaction



Notes: Dynamic (left panel) and group-specific (right panel) gtDD estimates of the impact of LEZs on life satisfaction of individuals living inside the LEZs. The sample of control individuals is restricted to residences further than 25km away from the nearest LEZ in similar counties. Standard errors clustered at the household level.

Table 20: Heterogeneous LEZ effects on life satisfaction

(a) By MV ownership

	Owns MV	Without MV
ATE	−0.110*	−0.059
	(0.063)	(0.166)
N	2354	438

(b) (b) By diesel car ownership

	Diesel	Other fuels
ATE	−0.181	−0.083
	(0.113)	(0.072)
N	760	1588

(c) By income quartiles

	Q1	Q2	Q3	Q4
ATE	−0.177	−0.242*	−0.141	−0.160*
	(0.137)	(0.125)	(0.136)	(0.082)
N	1714	1519	1386	1349

(d) By age groups

	≥ 65y	< 65y
ATE	−0.054	−0.179**
	(0.113)	(0.071)
N	1958	4733

Notes: Group-time difference-in-differences (gtDD) estimates of the impact of LEZs on life satisfaction using subsets of individuals living inside the LEZs. The sample of control individuals is restricted to residences further than 25km away from the nearest LEZ in similar counties. Subsamples are split based on motor vehicle ownership, diesel vehicle ownership, income quartiles and age groups. Point estimates represent the simple aggregation across all groups and time periods (β). Standard errors clustered at the household level.Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Table 21: LEZ effect on health outcomes

	LS	Doctor visits	Hypertension	Cancer
ATE	-0.123	-1.350	-0.038^{*}	0.009
	(0.095)	(0.921)	(0.020)	(0.013)
N	4042	4039	4042	4042
G	4	4	4	4
T	5	5	5	5

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{\cdot}p < 0.1$

Notes: Group-time DID estimates of the impact of LEZs on life satisfaction and objective health outcomes of individuals living inside the LEZs. The sample of control individuals is restricted to residences further than 25km away from the nearest LEZ in similar counties. Standard errors clustered at the household level. Significance levels denoted by $^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.