The effect of Low Emission Zones in the local economy: A synthetic control approach to German cities.

Candidate Nº 41783 - August 15, 2020 - 8604 words

Abstract

Low Emission Zones (LEZ) have been widely adopted in Europe and proven successful in reducing air pollution levels, a growing public concern due to its effects on health and mortality. While LEZ are criticised for "hurting the economy", especially by the retail and transport sector, recent literature suggests LEZ could improve economic performance through the reduction of pollution itself. This makes LEZs a policy that can both harm and boost economic growth. To solve this contradiction and provide the first causal estimates on the effect of LEZ in the economy, I use the Generalized Synthetic Control Method to study the heterogeneous effects of German LEZ in regional GDP per capita. To verify the claims against LEZ I study how this policy affected the share of Gross Value Added (GVA) originated in local trade. Finally, I describe a new methodology to study potential treatment effects on non-treated units and apply it to cities that did not introduce a LEZ. I find (1) great variability on the effects of LEZ in the overall GDP per capita with negative and significant ATT of 2-3%, (2) heterogeneous, although mostly negative, effects in the share of Regional GVA from local trade and (3) negative potential treatment effects for all 3 non-treated cities where the method could give reliable estimates. These results suggest cities' characteristics highly influence the magnitude and sign of the effect of LEZ in the local economy and recommend caution when applying LEZ and similar measures.

***Keywords* —** Low Emission Zones, Economy, Generalized Synthetic Control, Germany, Economic Geography

Acknowledgements

I am grateful for the valuable remarks from my thesis supervisor, Professor Alasdair Jones and professors Kenneth Benoit and Dr. Raúl Aldaz-Peña. I also acknowledge funding and support from the Spanish Central Bank (Banco de España), especially from Laura Hospido. I am thankful for the inputs from participants of the seminar presented at the Spanish Central Bank research department. The views expressed are entirely my own and not necessarily those of my supervisors, LSE or Banco de España.

1 Introduction

Recent research estimates that 92% of the world's population lives in areas where levels of air pollution exceed the World Health Organization's guidelines, with 3 million deaths a year being attributed to ambient air pollution worldwide.[1] Its effect in mortality is complemented by a large body of research on the negative effects of air pollution in human health, such as an increased incidence of respiratory and cardiovascular diseases (Shaddick et al., 2018) but also bronchus, asthma, and lung cancers, with people older than 65 years and children (less than 17-year-old) being especially affected (EPA, 2004). Furthermore, a mature body of research has shown air pollution creates large economic costs (Dechezleprêtre et al., 2019), especially through limiting human capital (Graff Zivin and Neidell, 2013).

A more general view of all the impacts of air pollution is given by the European Commission's "Impact Assessment" of their 2005 plan to reduce the major air pollutants between a 50-85% for 2030 (European Comission, 2013). It estimates 63,600 less premature deaths from long-term exposure to air pollutants, 84 million less sick days, 146,000 additional km^2 of ecosystems protected from eutrophication and 23,000 additional km2 of forest ecosystems protected from acidification. It expects € 4,7 billion of compliance costs and direct economic benefits of € 3,2 billion[2] per year. Finally, they expect no net GDP impact.

To minimise the costs derived from air pollution, substantial and diverse policy initiatives to reduce air pollution have been implemented around the world. Most of them have been based in changes of urban mobility by trying to reduce the number of

[1] In the EU-28 the number of premature deaths per year caused by ambient air pollution is around 446.000 deaths a year (European Environmental Agency, 2019).

[2] From reduced workdays lost, healthcare cost savings, improved crop yields and reduced damage to the built environment.

vehicles or the emissions per vehicle.[3]. As one of these alternatives, Low Emission Zones (LEZ) have been widely implemented across Europe with Germany and Italy being at the forefront of the application of this policy.

Their implementation has delivered on the promise of reducing air pollution (Gehrsitz, 2017; Wolff, 2014) and respiratory and cardiovascular diseases (Pestel and Wozny, 2019) but has found resistance around European cities, from the general German population and local retail businesses and local studies - linked to retail interests - have estimated large negative economic effects of LEZ. With this contradiction in mind, and to develop a fair cost-benefit analysis of LEZ, it is crucial to know how the cities' economy has been affected by LEZ and how local trade reacts after their implementation.

1.1 Low Emission Zones

Low Emissions Zones are geographical areas where the entry of highly pollutant vehicles has been banned based on their level of emissions. The specific geographical limits vary from city to city, from covering the city's historical centre to all the city's geography.

The inauguration of hundreds of LEZ in European cities was motivated by the need to improve air quality and comply with the increasingly strict EU regulations (Holman et al., 2015) with at least 273 LEZ either implemented or planned for the near future. LEZ have been successful as a very significant proportion of cities' air pollution comes from traffic, being especially harmful to humans given they are emitted close to the

[3]Some examples include pedestrian areas, parking schemes, Low Emission Zones, congestion pricing schemes or limitations for certain vehicles at certain times of the day, dedicated lanes for massive transport vehicles, retrofitting of taxis and public buses or subsidies for electric or hybrid cars

ground.[4] Furthermore, there is strong causal evidence that the application of German LEZ has significantly reduced particle pollution in German cities between 2 and 9% (Gehrsitz, 2017; Wolff, 2014) while also reducing respiratory and cardiovascular diseases (Pestel and Wozny, 2019).

As noted before, the large scale application of LEZ around Europe has been heavily shaped by the European legislation on air pollution levels. Most regional and national governments have been either threatened by fines or strongly incentivised to pursue decisive policy action by European regulation. From 2005 to 2007, two-thirds of German cities with more than 100.000 persons were violating European limits and were forced to develop Clean Action Plans (in German: *Luftreinhaltepläne*), defining a set of measures to attain compliance with EU standards. Most of them had a LEZ as the principal measure (Gehrsitz, 2017). As a result of this European legislation, infringement procedures have been opened against 16 Member States[5] and the EU Court of Justice has already handed down judgements in Bulgaria and Poland.[6]

The case of Germany is especially relevant as most German cities over 10.000 inhabitants applied one LEZ from 2008 to 2013 (Gehrsitz, 2017). This was possible because German LEZ have been supported and standardised by the federal government: all of them accept the same "emissions windshield stickers" to be able to enter, with the colour of the sicker signalling their emission standard. Stickers, in increasing pollution emissions, go from green to yellow to red. The strictest LEZ only allow cars with green stickers and the most polluting cars have no sticker at all. With time, German LEZ have become increasingly restrictive and more cars have been banned from entering them (from March 2018, all German LEZ but one only allow

[4]In EU cities, around 35-55% of PM_{10} particles is originated in vehicles (Viana et al., 2008).

[5]Belgium, Bulgaria, the Czech Republic, Germany, Greece, Spain, France, Hungary, Italy, Latvia, Portugal, Poland, Romania, Sweden, Slovakia and Slovenia

[6]Communication from the EC, "A Europe that protects: Clean air for all", accessible here: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0330&from=GA

green stickers).[7] Entering a LEZ without authorisation is fined with 80€ and implies a 1 demerit point in the central traffic registry (Gehrsitz, 2017).

1.2 Economic critiques of LEZ

According to classic economic theory, we would expect policies like LEZ to reduce a city's GDP as they impose a set of costs in the local economy. Old cars are prohibited, some cars have to be retrofitted to attain the new emissions standards and all cars should be checked and labelled by their level of emissions, creating sizeable administrative costs. This can lead to claims of a "Jobs versus the Environment" trade-off as introduced by Morgenstern et al. (2002) in which regulators have to weight both objectives and make compromises. The effect of LEZ can be especially damaging to individuals and business that depend on the transport of commercial goods given a set of costs are being imposed in their supply chains. Furthermore, the commerce that relies on tourists could also be severely affected if clients don't have a windshield sticker to enter the city's commercial areas such as the historic centre. These short-term costs (or lack of earnings) can end up reducing savings and investment, having an impact in the medium- or long-run economic growth.

In an online survey from 2009, over 91% of Germans rejected LEZ as being "too bureaucratic and likely having little effect" while store owners complained that LEZs lead to declining sales (Wolff and Perry, 2010). Furthermore, the German pro-business "Institute for Retail Research" estimated a 7% decrease in customers for stores located within a city's centre after the introduction of a LEZ (Lindstaedt, 2009). An example is the city of Freiburg, often visited by neighbouring French and Swiss tourists, being estimated to lose close to 100M€ per year in revenue by their local retail association (Badische Zeitung, 2009).

[7] For a graphical representation of these stages for all German LEZ see figure 10 in the Appendix.

The effects on the freight transport sector are, in principle, ambiguous. While a LEZ can impose strong direct costs to transport companies they can also benefit from an improvement of their fleet, a higher degree of industry consolidation or a reduction of congestion. Direct costs can come from multiple forms such as the retrofitting of old vehicles, the costs of new ones, or the reduction in the residual value of vehicles that become non-compliant with the new LEZ. On this subject, the study from Browne et al. (2005) of London's LEZ explains that the effect is strongly dependent on the LEZ restrictions and the rate of vehicle replacement of companies. They continue by saying that companies with specialised vehicles, small companies and self-employed owner-drivers end up bearing most of the costs. Dablanc and Montenon (2015) complement these findings by saying that the application of LEZ reduces the number of firms that perform deliveries. They mark this effect as socially desirable as it is expected to promote efficiency by expelling small and less-profitable firms that "find it difficult to maintain a sufficient level of business activity without breaking the freight sector's labour laws and safety standards".

From this literature, we would expect that the costs of the implementation of a LEZ are relatively higher in retail and transport activities with the whole "local trade" sector being more affected than other sectors. If no other sector of the economy benefits in equal or greater amount, the overall growth of the city in question could be reduced.

1.3 Potential economic benefits of LEZ

Even if there are indications that a LEZ might impose some costs in the local economy, it also has the potential to have positive consequences as the reduction in air pollution and congestion might offset other economic costs. This hypothesis is based on recent research by Wolff (2014), Gehrsitz (2017), showing the significant effect of German LEZ on the reduction of pollution and health costs. Specifically, Gehrsitz (2017) estimates that the application of German LEZ reduced average pollution from 0.67 to -1.3μg/m^3. Furthermore, LEZ have also been shown to reduce traffic congestion (Cesaroni et al., 2012), and their reduction of air pollution has significantly reduced chronic cardiovascular and respiratory diseases in nearby hospitals after their implementation (Pestel and Wozny, 2019).

Additionally, recent studies suggest that air pollution reduces aggregate economic output. Dechezleprêtre et al. (2019) show how air pollution causes economy-wide reductions in productivity and economic activity.[8] He concludes that a 1 μg/m^3 ($\approx 10\%$) increase in $PM_{2.5}$'s yearly average concentration causes a 1.2% decrease in GDP and, strikingly, that air pollution abatement costs would be two orders of magnitude smaller than the associated economy-wide reductions in productivity.[9] Hao et al. (2018) also find a significantly negative impact of air pollution on GDP per capita, with an increase of 5 μg/m^3 ($\approx 10\%$) in $PM_{2.5}$'s yearly average concentration causing a reduction of 6.1% of GDP per capita.

This negative effect on the economy can come from multiple causal mechanisms. Although the most documented is the direct effect on human mortality and morbidity

[8] They use wind direction and thermal inversions as instruments in an Instrumental Variable specification. In it, they show how exogenous variation in pollution affects the GDP per capita of NUTS 3 regions in the EU.

[9] Between 2000 and 2015, $PM_{2.5}$ across the European Union declined by 20%. Dechezleprêtre et al. (2019) estimates imply that this increased EU GDP by 2.4%, a 15% of all GDP growth in Europe over this period.

already mentioned, recent research has focused on how pollution directly affects labour as an input of production. As air pollution increases the morbidity of multiple sicknesses, it has also been found to increase absenteeism. Earlier studies focused on school absenteeism (Ransom and Pope, 1992) with more recent research also focusing on absenteeism from work (Aragón et al., 2017; Hanna and Oliva, 2015; Holub et al., 2006).

Air pollution has also been recently found to decrease worker productivity, as it also impairs cognitive and physical activities. After pollution particles are inhaled, they can enter the lungs and pass to the bloodstream, finally affecting multiple organs such as the heart and the brain (Calderón-Garcidueñas et al., 2014; Du et al., 2016; Ranft et al., 2009). This effect has already been documented both for physical work and high-skill intellectual work. Graff-Zivin and Neidell (2012), Adhvaryu et al. (2014) and Chang et al. (2016) investigate the effect of air pollution in workers' productivity in the agricultural sector in California, garment factories in India, and call-centres in China, respectively. All find significant negative effects of pollution on workers' productivity. Intellectual performance is also reduced with higher levels of pollution. Ebenstein et al. (2016) look at standardised high-school examinations in Israel, Roth (2015) examines university examinations in the UK and Zhang et al. (2018) study the effects on a cognitive test score from a nationally representative survey of Chinese families and individuals. All of them coincide that air pollution causes a sizeable decrease of cognitive performance.

1.4 Research Questions

The research questions of this paper are the following:

Main question:

- What was the effect of the application of Low Emission Zones (LEZ) on German cities' local economy?

Complementary questions:

- As its critics suggest, is local trade especially affected by the application of a LEZ?
- What would be the effect of a LEZ in other cities that have not applied it?

The remainder of this paper is structured as follows: Section 2 describes the different data sources and variables of the study. Section 3 is centred in the data analysis strategy, explains the selection of treated and control units and presents the Generalised Synthetic Control Method, its notation, assumptions and robustness tests. Furthermore, it describes how the analysis of each outcome is implemented. Finally, it presents the "Synthetic Treated Method", its justification, uses, and underlying assumptions. Section 4 presents the results for all analysis: the overall effect of LEZ on (1) local GDP per capita, (2) the proportion of local GVA coming from local trade and (3) the potential treatment effect the GDP per capita of non-treated cities. Section 5 concludes and presents ideas for further research.

2 Data sources

The data consisted in two main parts: (1) Sufficient pre- and post-intervention aggregate outputs and economic characteristics of regions, both to test the validity of the identification strategy and be able to see the short- and mid-term effects of the intervention. (2) A detailed description of the application of policies for all cities considered to correctly assign treatment and create a carefully selected control pool. Units were classified as treated regions (German cities that applied a LEZ) and control regions (German or other European cities that have not applied any similar measure). The data was gathered from different sources and aggregated by geographical location.

First, Eurostat's collection of regional statistics available for NUTS 3 regions[10] (from 2000 to 2019 and covering the whole of the EU) provides yearly statistics to construct the main dependent variables such as GDP, GVA per sector[11] and population. Additionally, Eurostat provides the geographical representation of NUTS 3 regions, allowing to identify how well a given NUTS 3 region represents a given city and to calculate distances between any point and the limits of a NUTS 3 region.

Secondly, the implementation of LEZ in Germany is well documented by the German Environment Agency (UBA, *Umweltbundesamt* in German) with dates for the application of each "stage" of a given LEZ. On the other hand, the announcement dates of each LEZ are not documented on a unique database and were searched individually on historical documents such as local news and each city's "Environmental Plan"

[10]NUTS, or Nomenclature of Territorial Units for Statistics, is a geographical code to reference the subdivisions of countries. NUTS 3 is the most detailed geographic definition with granularity varying by country. Germany has especially detailed NUTS 3 regions and they tend to follow natural city boundaries. An illustration of the NUTS regions in Europe can be found in Figure 9 in the Appendix.

[11]GVA per sector is divided into 6 aggregate sectors of economic activity according to the NACE Rev.2 codes, in parenthesis: **(A)**: Agriculture, forestry and fishing, **(B-E)**: Industry - except construction -, **(F)**: Construction, **(G-J)**: Local trade or "Wholesale and retail trade; transport; accommodation and food service activities; information and communication", **(K-N)**: Financial; real estate; professional, scientific; technical; administrative and support service activities, and **(O-U)**: Public administration and defence; compulsory social security; education; human health and social work activities; arts, entertainment, repair of household goods and other services.

published in their official websites. Figure 10 summaries the treatment status of all LEZ in Germany from 2006 to 2019.

Additionally, data on the implementation of LEZ and similar policies for all European cities has been gathered from specialised databases such as *UrbanAccessRegulations.eu* (financed by the European Commission) and the "Green Zones" mobile application that informs professional carriers and drivers on the state of urban access regulations and LEZ in Europe.[12] Information on treatment status is freely available on their web page and public Android mobile application. The data was recollected and re-structured to a geographic format for further analysis.[13]

[12]UrbanAccessRegulations.eu has been already used by Holman et al. (2015) as a source of data to review the efficacy of European low emission zones to improve urban air quality.

[13]Neither of them have restrictions to web scraping in their web pages.

3 Data Analysis strategy

I use the announcement of German LEZ during the period from September 2009 to 2018 (illustrated in Figure 10) as the treatment indicator. This gives a range of 9-18 years of pre-treatment data and 1-8 years after the treatment to evaluate the fit of the model and estimate the short- and medium-term effects of the policy.

The set of control regions (or control pool) is constructed from other German regions or regions from a selected group of countries. It is constructed for each outcome separately, to maximise the robustness of its estimates. The control cities should have similar characteristics as the treated cities and have not applied similar measures to be considered as controls.

I use this data to estimate a synthetic control for each treated unit and calculate the effect of LEZ in the local economy. I estimate both the heterogeneity of treatment effects and the average treatment effect on the treated. Here follows a careful description of how this method is applied, its bases, assumptions and limitations.

3.1 Synthetic Control and Generalised Synthetic Control

To answer the main question on the effect of LEZ in the city's various economic outputs I use the Generalised Synthetic Control Method (Gsynth) as introduced by Xu (2017). It consists of a modification to the classical Synthetic Control Method (SCM) that, although different, continues to be based on the same theoretical foundations. The SCM was introduced in Abadie and Gardeazabal (2003) to study the effect of terrorism in the Basque Country's macroeconomic performance and is extensively described and expanded in Abadie et al. (2010), Abadie et al. (2015) and Abadie (2019). Specifically, Gsynth differs from the classical SCM given it does not create synthetic controls (counterfactuals) from a weighted average of control regions but as the result of a linear factor model (LFM) and thus from a combination of latent factors and factor weights.[14] Gsynth is especially appealing for the questions and data characteristics of this research given it behaves well with multiple treated units. Furthermore, its estimation procedure with latent factors and factor loadings allows us to interpret them as economic shocks and responses.

Even if the SCM, or any of its variants, has not yet been applied to cities' LEZ, previous research has used very similar techniques. For example, Wolff (2014) and Gehrsitz (2017) applied differences-in-differences (DID) to estimate the causal effect of German LEZ in local air quality and health outcomes.

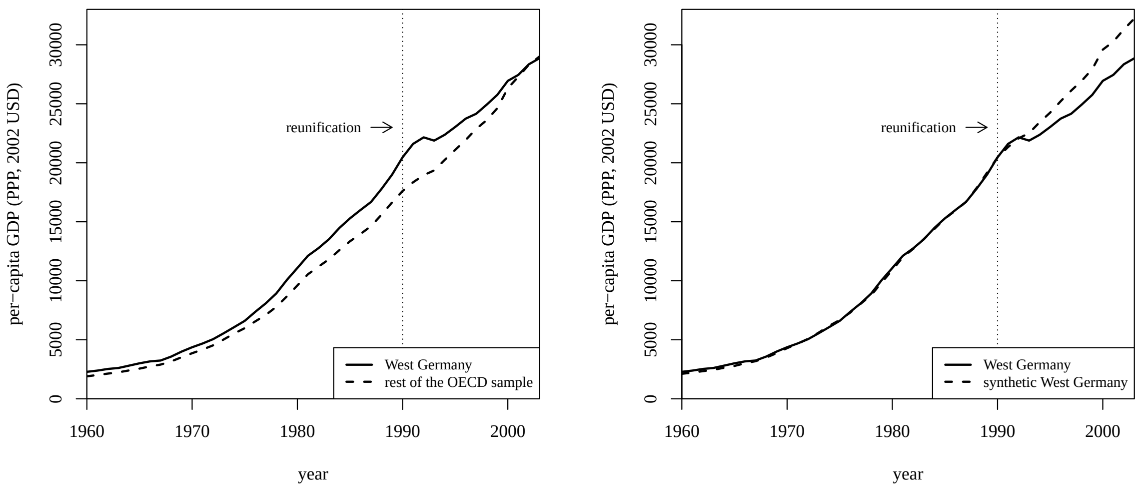
DID and the SCM share multiple similarities, both lay their main identification requirement on having similar trends before treatment and avoid the need of assuming an exogenous treatment assignment. In contrast, although DID allows for the presence of constant unobserved confounders (given it is constructed by taking unit and time differences), the SCM also allows the effects of confounding unobserved characteristics to vary with time (Abadie et al., 2010). Gsynth is based on a linear factor model where these confounding unobserved characteristics (factors) are explicitly estimated to produce a counterfactual for each treated unit.

Relative to DID, the SCM and Gsynth try to avoid ambiguity on how comparison units are chosen and account for the uncertainty that they will indeed reproduce "the outcome trajectory that the affected units would have experienced in the absence of the intervention or event of interest" (Abadie et al., 2010). They do it by creating a

[14]Multiple methodologies build on the SCM and could also in this research, some examples are Abadie and L'Hour (2019) where they include a penalisation term for pairwise differences on characteristics and Bayesian structural time-series models (Brodersen et al., 2015) developed in Google which uses Bayesian Machine Learning procedures to estimate the counterfactual and its credible interval.

synthetic control from non-treated units such that it best follows the pre-treatment path of the treated unit.

The study from Abadie et al. (2015) provides a very understanding example of the SCM. In their paper, they quantify the effects of Germany's reunification in West Germany's GDP by creating a "Synthetic West Germany" from a set of controls from all other OECD countries. At Figure 1(a), we can see how West Germany has a different pre-treatment path to the rest of the OECD sample mean, and thus this average does not constitute a good comparison unit. Figure 1(b) shows how the SCM produces a synthetic control from the set of OECD countries that follows very closely the pre-treatment period of real West Germany creating a credible comparison unit and estimates of the effect of Germany's reunification in West Germany GDP per capita.



(a) West Germany and the mean of the OECD (b) West Germany and its synthetic control

Figure 1: We can see how the SCM creates a "synthetic West Germany" that closely follows the pre-intervention path of West Germany and deviates from if after the intervention.
Source: Abadie et al. (2015).

To briefly explain the Generalised Synthetic Control Method, I first need to set some bases on potential outcomes notation. For a given time t, unit i, a binary treatment $D \in \{0, 1\}$ and outcome $Y_{it}(D)$, the realised outcome of a given unit can be expressed as:

$$\overbrace{Y_{it}(D_{it})}^{\text{Realised Outcome}} \quad = \quad \overbrace{Y_{it}(0)}^{\text{Counterfactual}} \quad + \tau_{it} D_{it}$$

with τ_{it} representing the treatment effect at time t for unit i. This is common to all methods that use comparable non-treated units as counterfactuals such as DID and other synthetic control methodologies.

As mentioned before, Gsynth doesn't estimate weights from control units as the SCM but characterises the path of units in absence of treatment (counterfactual) as a set of common latent factors and factor loadings:

$$Y_{it}(0) = \boldsymbol{x}_{it}^\top \boldsymbol{\theta} + \boldsymbol{\lambda}_i^\top \boldsymbol{f}_t + \gamma_t + \varepsilon_{it}$$

Where $\boldsymbol{x}_{it}^\top \boldsymbol{\theta}$ is the linear product of a set of parameters and predictors, $\boldsymbol{f}_t = (f_{1t}, ..., f_{Jt})^\top$ are J time-varying factors (usually interpreted in economics as "shocks"), and $\boldsymbol{\lambda}_i = (\lambda_{i1}, ..., \lambda_{iJ})^\top$ are J unit-specific factor loadings (interpreted as "responses" to each shock), γ_t represent time fixed effects and ε_{it} are the residuals.

To estimate the parameters, latent factors, factor loadings and treatment effects for each treated unit I use the three-step estimation procedure proposed by Xu (2017):

1. Control units are used to estimate $\boldsymbol{\theta}$, $f_1, ..., f_T$, $\lambda_1, ..., \lambda_{n1}$ and γ_t by minimising the MSE of outcome and counterfactual, $\left(Y_{it} - \widehat{Y_{it}}(0)\right)^2$. For them, there should be no difference given they have no treatment effect.

2. Obtain estimated factor loadings $\hat{\boldsymbol{\lambda}}_i$ for the treated units, conditional of the first step by minimising the MSE between Y_{it} and $\widehat{Y_{it}}(0)$ in the pre-intervention period.

3. Use the estimated values of $\widehat{\boldsymbol{\theta}}$, $\widehat{\boldsymbol{\lambda}_i}$, $\widehat{\boldsymbol{f}_t}$ and $\widehat{\gamma_t}$ to calculate the synthetic control of treated units and calculate the intervention effects as $Y_{it} - \widehat{Y_{it}}(0)$ in the post-treatment period for the treated units.

Gsynth is especially appealing for comparative case studies that have multiple treated units for its transparency and theoretical foundations: (1) We can clearly see the differences in pre-treatment paths between treated and synthetic control making possible to assess its internal validity and choose which treated units have valid synthetic controls. (2) This allows us to make all decisions regarding the set of controls by looking at the pre-treatment fit, safeguarded against specification searches and p-hacking. Gsynth takes further advantage of this and allows to choose the number of factors J by a cross-validation procedure, minimising out-of-sample bias. (3) It's possible to check all estimations are done within the support of the data (no extrapolation) by looking at the factor loadings (λ_i) of treated and control units. (4) Given the construction of the synthetic control is based on a linear factor model, it allows to interpret factors in economic terms giving insight on the economic shocks present on the sample period and the responses of individual units. (5) Xu (2017) describes the implementation of a bootstrap procedure to obtain confidence intervals for the treatment effects and p-values.

3.2 Assumptions and identifying requirements

As DID, the SCM and its variants have very important identifying requirements and assumptions to be able to assess their results as causal estimates. Following technical and practical guidance from Abadie (2019) and Xu (2017), careful thought has been given to attain the various requirements and ensure the assumptions are credible. They are discussed below.

The first requirement is to have a valid set of control regions (or control pool). The treatment and control pool should have similar characteristics and no external shock should affect one and not the other. The most important source of possible bias in this study is the 2009 German scrappage program, a policy specifically thought to stimulate the economy after the 2007 financial crisis. The German scrappage program was the largest in the world by offering a lump-sum subsidy of 2.500€ for buying a new car when the buyer scrapped their old one. From the 14^{th} of January to the 2^{nd} of September, 2009, two million car sells were subsidised, implying the substitution of 2 million cars older than 9 years old for new cars with better emissions standards (at least "Euro 4" compliant) (Kaul et al., 2012).

This creates a potentially significant source of bias for cities that had a LEZ either implemented or announced before the end of this program. Their population had stronger incentives to change their old cars for new ones given they were (or would) not be able to use them inside the LEZ. This implies that these cities would potentially receive larger sums of fiscal stimulus from the central government, modifying their GDP per capita. For that reason, I exclude all cities that had their LEZ either implemented or announced before that date. Figure 10 shows clearly which cities were excluded from the sample of treated cities with almost all major German cities being excluded from the "treated" sample. On the other hand, controls were also restricted to be comparable with treated units by only including NUTS 3 regions that are similar in terms of the outcome variable (ex. GDP per capita) and have similar levels of population density (population/km^2).

Secondly, there must be no interference between units, in other words, that control cities are not influenced by the treatment of nearby treated cities. Wolff (2014) shows that the application of a LEZ correlates with a change towards cleaner vehicles and cities close enough also seem to experience these changes. For that reason, only zones that are at least 60km away from a LEZ are included as controls.[15]

Thirdly, it is important there is no anticipation to the policy as any such effect would bias the selection of the synthetic control. This can be an issue given LEZ in general, and also in Germany, are usually publicised before being enacted to incentivise the public to upgrade their vehicles before its application. As an example, Wolff (2014) shows how the city of Regensburg had a very strong relative increase in green-labelled cars after the LEZ announcement but before its application. This is accounted by creating an "announced" period, and setting it as the start-of-treatment date in the SCM as recommended in Abadie (2019). For example if the start-of-treatment date was set to the implementation and a LEZ harmed GDP per capita before that date, the synthetic control would mask the pre-implementation effect and would give biased results. Although this looks like a potentially important source of bias, previous research on LEZ tends not to control for it. Wolff (2014), Gehrsitz (2017), Pestel and Wozny (2019), Morfeld et al. (2014) and Browne et al. (2005) are some examples.

[15] This is the distance from the centre of a treated city to the closest point of a "control" NUTS 3 zone (and thus a conservative measure). A very similar control pool is generated with a distance of 80km between the centroids of treated and control regions.

3.3 Robustness and diagnosis checks

As detailed in Abadie (2019), the main tests needed to assess the robustness of the results from a SCM are regarding the quality of the estimated synthetic control and the significance of the estimates. To test the quality of the synthetic control I perform in-time placebos: fake treatments that are set a number of years before the announcement of the LEZ to see if there is any effect between the fake treatment date and the actual date of the announcement. A robust synthetic control should follow the treated unit outcome path before and during the fake intervention period as no real intervention has happened. I perform in-time placebos 4 years before the announcement of LEZ for all outcomes presented and restrict my results to those cities where there was no significant effect during this fake treatment period.

This is an especially restrictive procedure as I only have 9 pre-treatment periods for some regions. Doing 4-year in-time placebo effectively test if the method can create robust synthetic controls with 20-45% fewer data. For a set of cities, the information of these last 4 years could be crucial to construct the correct synthetic control.

To test for significance I use the parametric bootstrap procedure from Xu (2017) which is further improved by using the Expectation-Maximization algorithm as shown in Gobillon and Magnac (2016)[16]. These standard errors assume $\varepsilon_{1t}, \dots, \varepsilon_{nt}$ are independent and homoscedastic at each time t after controlling for unobserved confounders and time fixed effects.

[16] I use Xu's `gsynth` R package to calculate the resulting synthetic controls, treatment effects and their standard errors.

3.4 The effect on GDP per capita and local trade

To investigate the effect of the application (or announcement) of a LEZ in the local economy I calculate its effect on regional GDP per capita as the most used measure of aggregate value creation for a given economy. It has to be said that it is not a perfect measure of economic well being or value creation. Some effects of reducing pollution, especially the ones related to health and reduced mortality are systematically undervalued by a measure like GDP as the value of non-commercial goods such as health and lifespan is not usually accounted for in economic transactions. Moreover, positive economic effects can be translated into reductions in GDP. For example, a decrease in various respiratory illnesses caused by a reduction in air pollution can reduce the need for costly health interventions and thus cause a decrease in local GDP. With these limitations in mind, the use of GDP per capita as the main outcome is based on its availability in a granular data and its widespread use by economists, policymakers and the general population.

Additionally, I estimate its effect on the share of GVA that originates in activities from local trade, a sector that is thought to be strongly affected by the policy and where most of the critiques to LEZ come from, especially from retail and transport activities.[17] GVA is closely related to GDP, with the only difference being that it does not includes taxes and subsidies and thus only accounts for the value created that is present in economic transactions. This analysis is specifically intended to test if critics from retail are well-founded and up to which point.

[17] I define the "local trade" sector as the equivalent to codes G-J from the NACE Rev.2 classification and includes wholesale and retail trade, transport, accommodation and food service activities, information and communication.

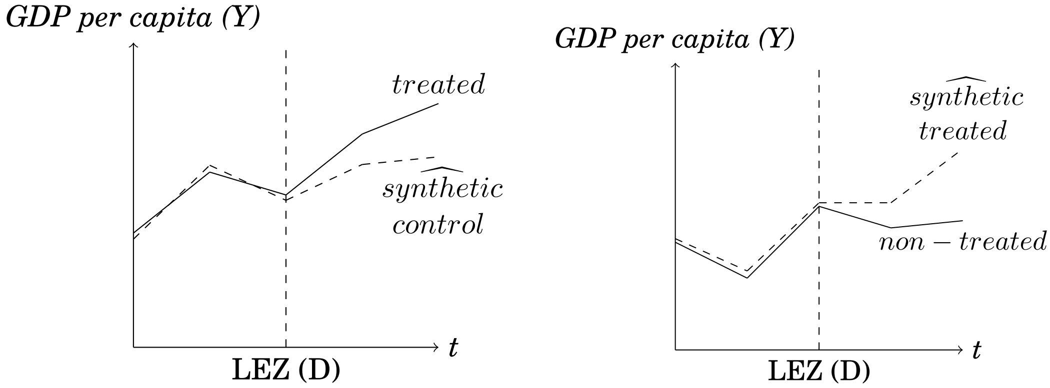
3.5 Estimating potential treatment effects:

As a methodological contribution to make my results more useful for policy decisions, I introduce a new method to calculate potential treatment effects on non-treated units. This is accomplished by reversing the mechanism of the SCM and use the large set of treated cities to construct "synthetic treated" units. This method is referred as the *"Synthetic Treated Method"* on the rest of this paper.

The basis of this method comes from the general intuition that, the same way treatment effects are derived from the difference between the realised outcome of treated units and their non-treated counterfactual $Y_{it}(D = 1) - \widehat{Y_{it}}(D = 0)$, the potential treatment effect can be defined as the difference between the 'treated' counterfactual of control units and their realised outcomes $\widehat{Y_{it}}(D = 1) - Y_{it}(D = 0)$. The methods usually used to estimate $\widehat{Y_{it}}(D = 0)$ can be used to estimate $\widehat{Y_{it}}(D = 1)$ just by changing the relative position of treated and control units.

In this specific case, I use the large set of treated units as controls and the control units as treated to estimate the potential treatment effect of LEZ on the GDP per capita of non-treated cities.

This general reinterpretation of potential outcomes is not restricted to Synthetic Controls and applies to all methods that are based on creating a counterfactual from units that have a different treatment status. Such methods include Difference-in-Differences, Latent Factor Models, Synthetic Controls, and Matching among others. The SCM is especially appealing to this new application given it tries to follow the path of each unit, providing individual estimates of treatment effects, and has clear limits of when it can be applied, providing formal tests for the internal validity of its inferences. After extensive research on the SCM and similar methods, I have not found this application in the literature. No further assumptions are needed and the properties of the classical SCM and Gsynth apply. A graphical illustration is shown in Figure 2.



(a) Classic SCM

(b) "Synthetic Treated Method"

Figure 2: Illustration of SCM and "Synthetic Treated Method" with positive treatment effects and positive potential treatment effects.

As in any other method of causal inference, it is important how the units that constitute the counterfactual are selected and what they represent. For this reason, this study is not the best setting for a demonstration of the "Synthetic Control Method" for two main reasons. First, there is no unique treatment on German LEZ. Although they are standardised in many dimensions, the "synthetic treated" counterfactual comes from cities of varying relative size of LEZ and different implementation calendars. Moreover, the study is done in a context where there are strong temporal trends and thus only cities that announced a LEZ in a given year can be used to construct the counterfactual of what would have happened if a LEZ was announced on that year in a control city. This strongly reduces the number of treated cities I can use as counterfactual. Although these two conditions obscure interpretation of the results and limit the number of treated cities to construct counterfactuals, it is still possible to calculate potential treatment effects and have an intuition of the sign and magnitude of potential treatment effects for some non-treated cities.

4 Results

The results are divided into 3 main parts: the effect of LEZ in aggregate GDP per capita, its effect on the proportion of GVA from local trade and its the potential treatment effect on non-treated cities GDP per capita. As it is only possible to have robust estimates of causal effects from cities that pass the different robustness tests, only those results are presented. 10 cities had robust synthetic controls for local GDP per capita and only 5 for the proportion of local trade in local GVA. Finally, only 3 control cities were suitable for estimating potential treatment effects.

Although the original control pool included NUTS 3 regions from all EU countries[18], no control pool that included regions outside of Germany resulted in robust synthetic controls, with multiple combinations of them being tested.[19] This is likely due to the existence of multiple exogenous shocks during the study period that affected other countries differently. Only by restricting the control pool to non-treated German NUTS 3 regions some robust and credible synthetic controls were possible. The complete set of control regions are mapped in Figure 11 together with the location of treated cities. Finally, and for each outcome of interest, this control set was further restricted to zones with low discrepancies in the outcome or population density to reduce interpolation biases.

Summary statistics are presented for each sample used by comparing treated and control regions and results are presented individually or with aggregates such as the average treatment effect on the treated (ATT).

[18]An illustration of this sample can be found in Figure 9.

[19]This procedure of selection of controls was done exclusively looking at pre-intervention periods and in-time placebos to avoid p-hacking and the manipulation of results.

It is important to note the external validity of these estimates is limited to similar cities (100.000 - 300.000 inhabitants). Furthermore, given the LEZ included in the study were announced later than other (usually bigger) LEZ, it is possible that the cities studied had fewer problems of air pollution than the early adopters and thus their economic outcomes do not benefit as much from the reduction in pollution generated by a LEZ.

A very important note is that all these results are valid only if there are no large idiosyncratic shocks on individual treated or control cities after the treatment date. An example where this does not apply is the city of Magdeburg that suffered from large floods in 2013 and thus no estimate of its treatment effect was estimated after that date.[20]

4.1 The effect of LEZ in cities' GDP per capita

A robust synthetic control could be constructed for the cities of Balingen, Erfurt, Heidenheim, Langenfeld, Liepzig, Magdeburg, Marburg, Mönchengladbach, Remschield, and Siegen. A summary of the outcomes and population density of the treated and control pool is as follows:

[20]Although I looked for similar shocks in all treated cities that were used in this analysis it is possible some were missed, especially for the large control pool.

Variable	Status	n	Min	Mean	Max	Std. dev.
GDP per capita (€)	Control	67	15650	31220	70794	11630
	Treated	10	27338	29929	33000	2074
Population per km^2	Control	67	101	588	2678	600
	Treated	10	197	878	1745	626
Share of GVA (%) from:						
Industry	Control	67	8.7	24.9	56.4	9.3
	Treated	10	11.0	28.4	43.4	13.0
Construction	Control	67	1.6	4.7	8.7	1.8
	Treated	10	3.1	4.6	6.8	1.2
Local trade	Control	67	8.6	19.2	35.4	5.3
	Treated	10	11.8	19.3	28.0	5.2
Agriculture, forestry and fishing	Control	67	0.0	1.0	4.4	1.0
	Treated	10	0.1	0.3	0.9	0.3
Public sector, education, health and arts	Control	67	12.3	25.7	48.0	7.9
	Treated	10	15.0	24.1	37.3	7.8
Financial, real state, professional, scientific and technical activities	Control	67	15.9	24.6	40.2	4.2
	Treated	10	20.0	23.2	29.0	2.7

Table 1: Summary statistics for effect on GDP/capita

From Table 1, we can conclude the treatment and control pool are balanced in the most relevant variables such as GDP per capita, population density and the share of the major sectors in the economy. Sample means are comparable and the extreme values of the treated pool fall into the common support of control regions.[21]

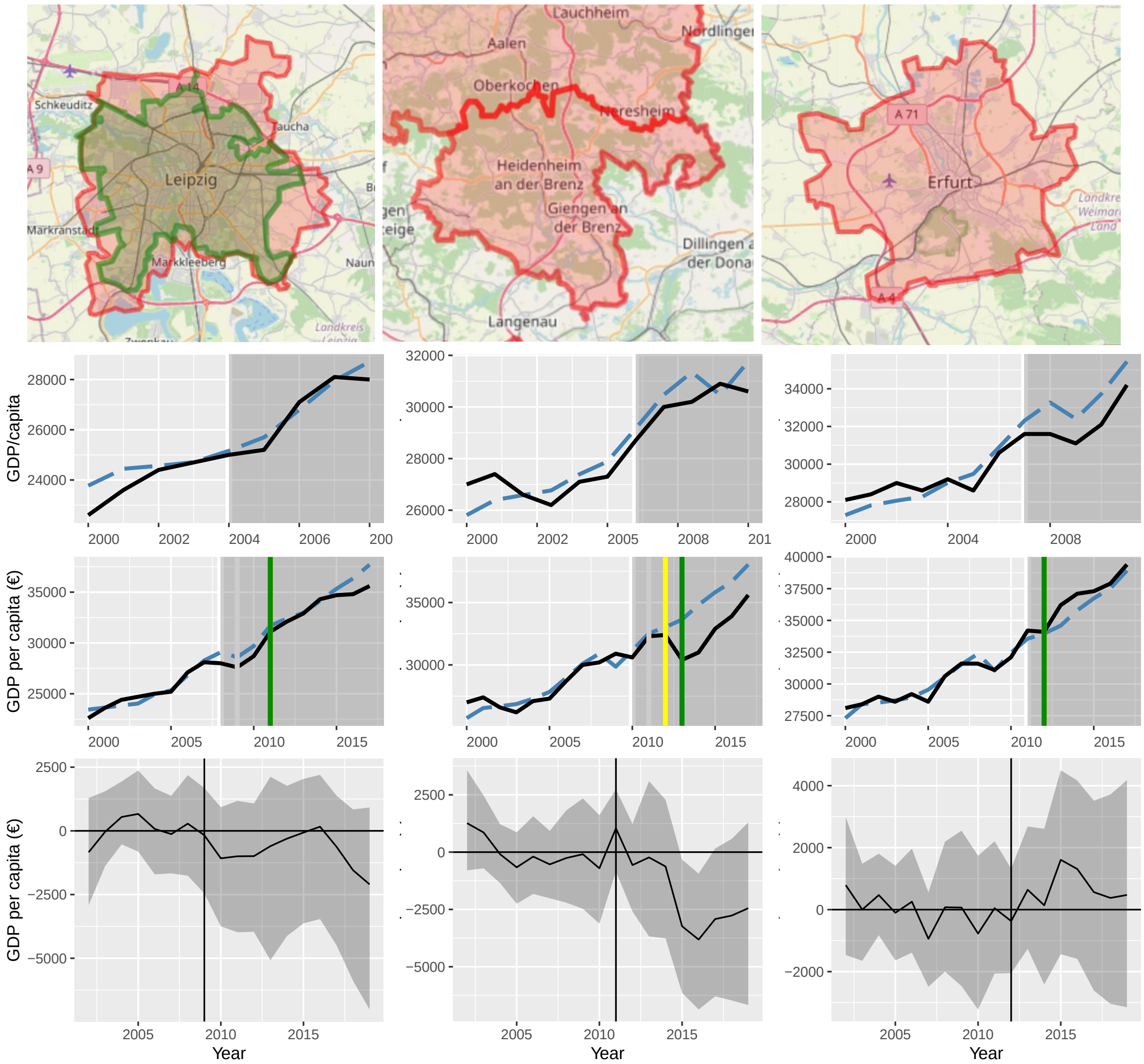
Although the estimated effects for these 10 cities are mostly heterogeneous, the Average Treatment Effect on the Treated (ATT) becomes significantly negative 4 years after its announcement, estimating a reduction in GDP per capita of 1240€ ($\approx$ 2-3%). The results show a large heterogeneity with most cities having no significant effect and only two showing significant and large negative effects: Magdeburg and Heidenheim[22]. Furthermore, no significant effect was found between the announcement and the implementation of the LEZ neither for individual cities or for the aggregate ATT. The resulting ATT and results for 3 example cities are shown in Figure 3 while individual results for other cities are shown in Figures 12-18 in the Appendix.

[21] The balance is not that clear for the "Agriculture, forestry and fishing" sector where the control pool has a higher proportion of activity. The sample was not restricted in this dimension given the absolute values of that sector are already low and thus it is hard its differences have a significant effect on the results.

[22] Magdeburg announced its LEZ in May 2011 and implemented it in October of the same year, only allowing yellow and green stickers. Nevertheless, it suffered large floods in 2013 so its results from 2013 and any subsequent years are not included in the calculation of the ATT.

Finally, all of these estimates are relative to other German cities where GDP per capita could have benefited from an increase in cars sells. If we imagine that these benefits are equally distributed in Germany, the estimates would constitute a lower bound of the real effect or the effect of LEZ "clean" of the economic stimulus it can create to the national car industry or any other effect that applies to Germany as a whole.

(a) Liepzig (b) Heidenheim (c) Erfurt



(d) ATT and number of treated regions at each period

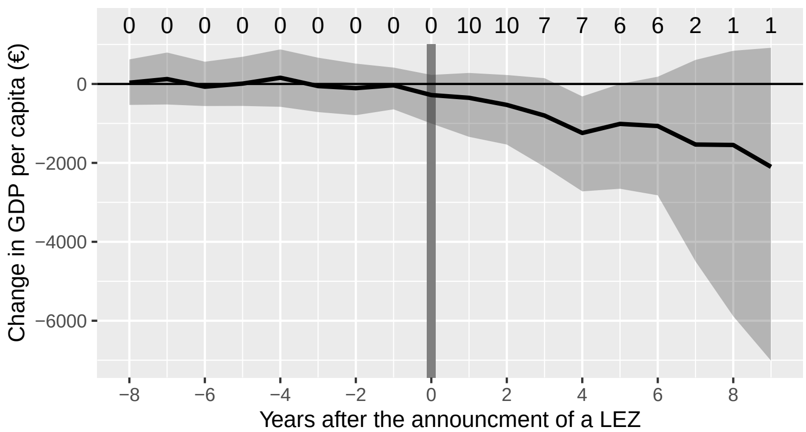
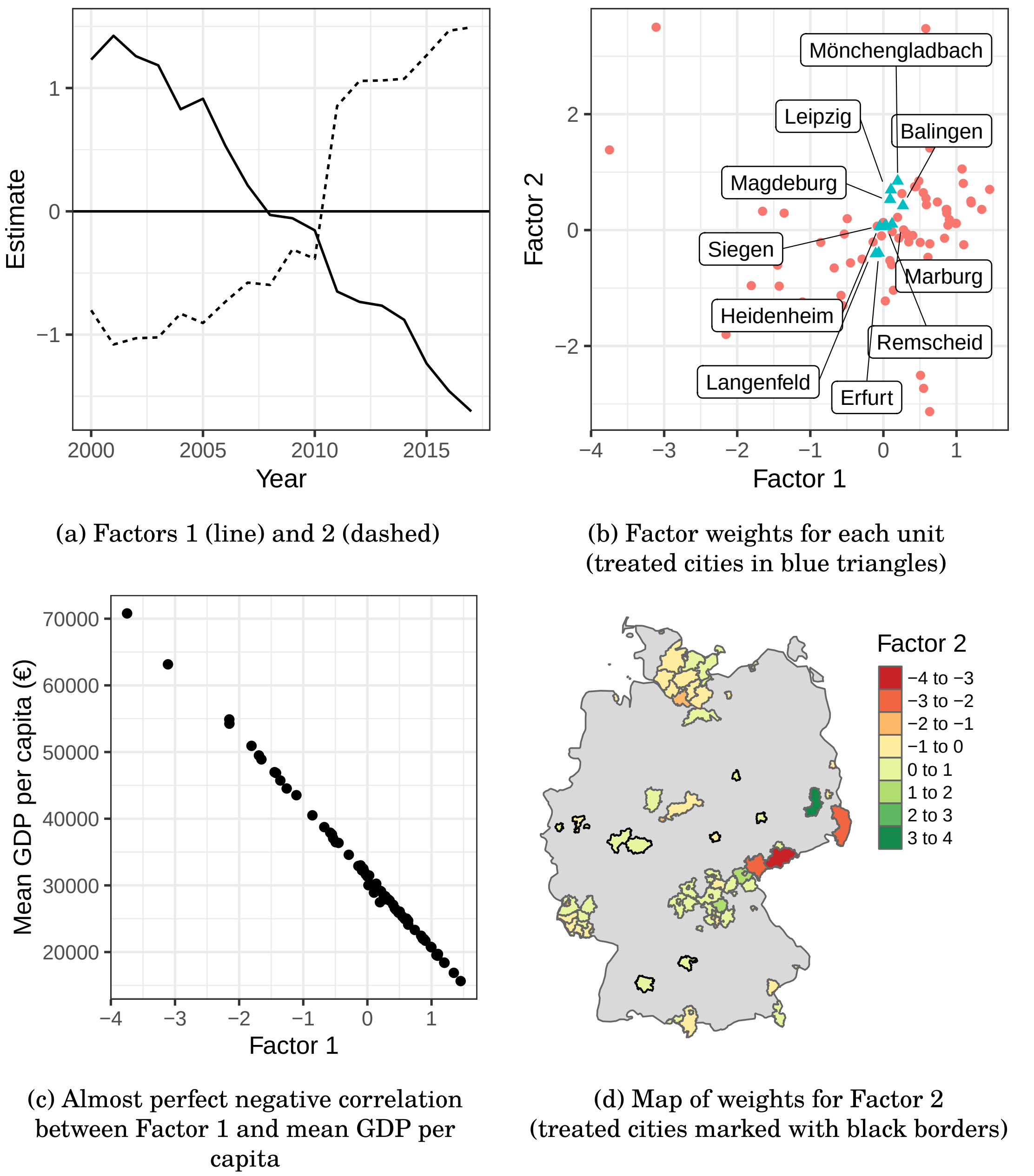


Figure 3: Results for the effect of GDP per capita for Liepzig, Heidenheim and Erfurt. We first see a map of each city with the NUTS 3 region (and the LEZ for Liepzig). The second row shows 4-year pre-intervention placebos to look at the robustness of the synthetic controls. Row 3 shows the aggregate paths of treated units and their synthetic controls with the dates and strength of the implementation maker with colours. Finally the overall effects are shown for each city and the ATT for all cities that had robust synthetic controls. SE are calculated with a 1000-iteration parametric bootstrap.

All diagnosis tests from the estimation of these effects are shown in Figure 4, where the evolution of factors and factor loadings are plotted. The first reassuring result, shown in Figure 4(b), is that estimated factor loadings of the treated units lie in the convex hull of the controls, showing that the estimated counterfactuals are in the common support of the control pool and thus are a result of interpolations, instead of less-reliable extrapolations.

Furthermore, even if factors are not usually interpretable, we can clearly see the first one controls for the differences in absolute GDP per capita growth based on their average level: The first factor has an almost perfect negative correlation with of GDP per capita (Figure 4c), and it is associated with lower growth during the study period (Figure 4a). This makes complete sense as cities usually grow in similar relative terms and thus bigger cities would have higher absolute growth of GDP per capita. Finally, Factor 2, which has to be orthogonal to the first one, is harder to interpret and it describes a strong positive shock in GDP per capita in 2010. From Figure 4(d) it is possible to see some geographical patterns with the zone of Nuremberg being positively affected by this shock and the zone of Hamburg being negatively affected.



(a) Factors 1 (line) and 2 (dashed)

(b) Factor weights for each unit (treated cities in blue triangles)

(c) Almost perfect negative correlation between Factor 1 and mean GDP per capita

(d) Map of weights for Factor 2 (treated cities marked with black borders)

Figure 4: Diagnosis tests for the effect in GDP per capita. All factors are normalised.

4.2 The effect of LEZ on Local Trade

To gather more insight into the effect of LEZ in local economies, I use the same methodology to look at the effects on the share of GVA of local trade (as defined previously). The summary statistics for the outcome and population density in both treated and control pool are as shown in Table 2. In it, we can see how the treated cities lie in the common support of the controls and have similar values for GDP per capita, population density and the share of different sectors in the economy.

Variable	Status	n	Min	Mean	Max	Std. dev.
GDP per capita (€)	Control	63	15650	30370	79200	11743
	Treated	5	25011	28674	32088	2662
Population per km^2	Control	63	101	562	2678	576
	Treated	5	211	980	1754	715
Share of GVA (%) from:						
Industry	Control	63	8.7	24.9	54.8	8.9
	Treated	5	9.5	24.9	43.4	15.6
Construction	Control	63	1.6	4.8	8.7	1.8
	Treated	5	3.7	5.0	6.5	1.3
Local trade	Control	63	10.7	18.7	29.8	4.2
	Treated	5	14.0	18.4	22.9	3.5
Agriculture, forestry and fishing	Control	63	0.0	1.1	4.4	1.0
	Treated	5	0.0	0.3	0.9	0.4
Public sector, education, health and arts	Control	63	13.2	26.1	48.0	7.9
	Treated	5	16.4	28.1	42.0	11.2
Financial, real state, professional, scientific and technical activities	Control	63	13.4	24.5	40.2	4.2
	Treated	5	20.0	23.3	26.2	2.5

Table 2: Summary statistics for effect on local trade

The effects of LEZ on the share of GVA from economic activities related to local trade are heterogeneous. Of the 5 cities studied, selected based on the quality of their pre-intervention placebos and their relevance, Heidenheim and Halle (Salle) had strong negative effects, Magdeburg and Siegen had neutral effects and the city of Mönchengladbach had strong positive effects. An illustration of some these effects is shown in Figure 5, other results can be found in Figures 19 and 20 in the Appendix.

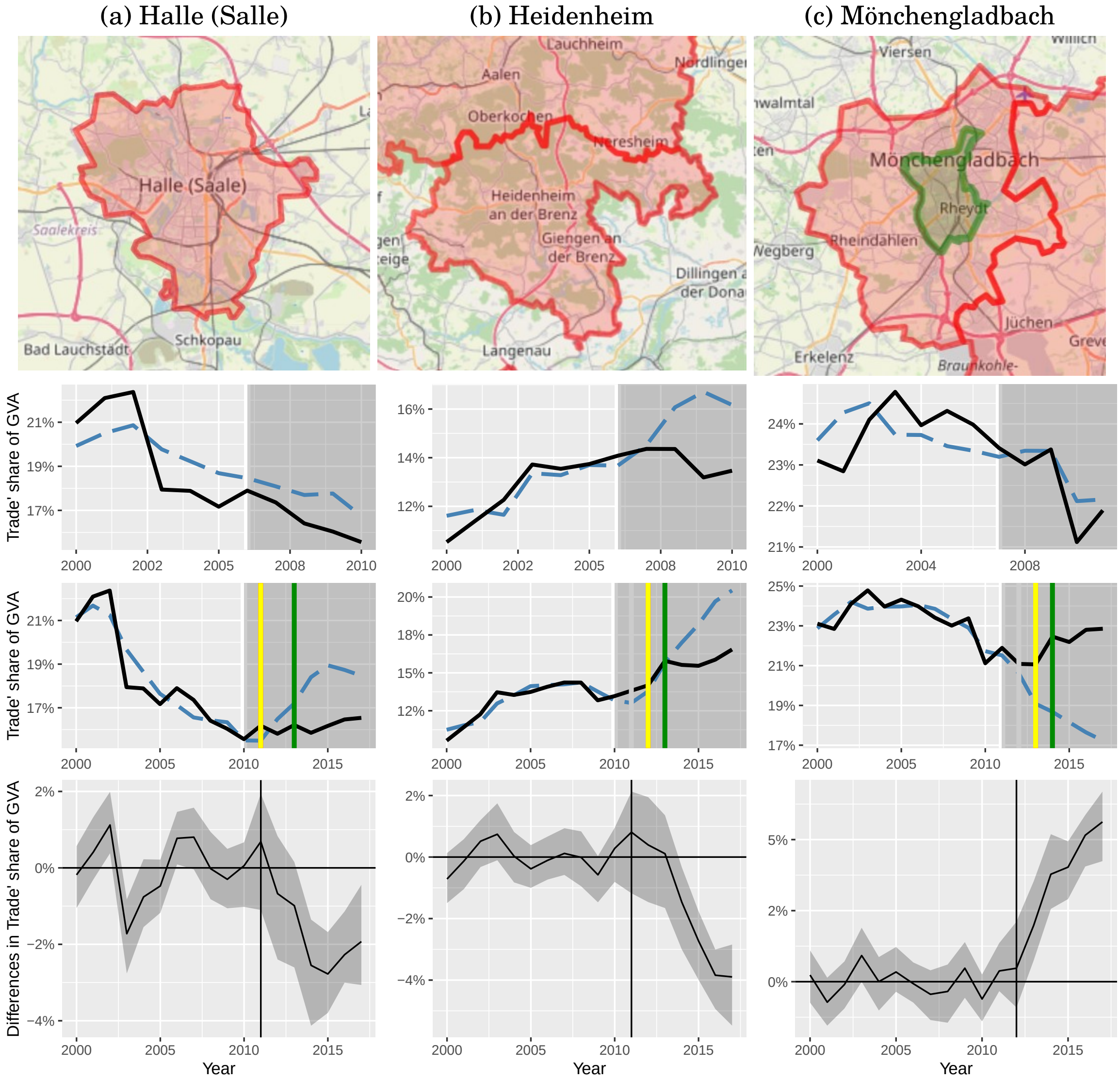


Figure 5: Results for the effect on Local trade share of GDP for Halle, Heidenheim and Mönchengladbach. We first see a map of each city with the NUTS 3 region (and the LEZ for Mönchengladbach). The second row shows 4-year pre-intervention placebos to look at the robustness of the synthetic controls. Row 3 shows the aggregate paths of treated units and their synthetic controls with the dates and strength of the implementation maker with colours. Finally, the overall effects are shown for each city. SE are calculated with a 1000-iteration parametric bootstrap.

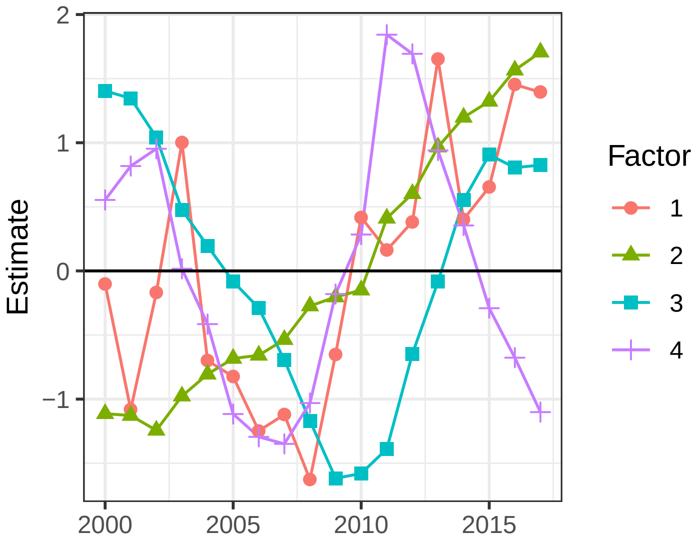
Results for Heidenheim are presented even if it does not perform well in the pre-intervention placebo given it is of special interest for its results in GDP per capita and its fit is credible after the 4 last years are used in the estimation of the synthetic control. ATT is not displayed given the high heterogeneity of estimated treatment effects.

As pointed out previously, the implementation of a LEZ can have negative effects on local trade, mostly though the transport and retail sectors if small freight companies have to incur in costs or if tourism or access to the centre is limited. I estimate that Halle's local trade sector suffers a reduction of 2% of its GVA, or 10% of its overall weight. Halle is a city with a population of 200.000 that implemented a LEZ at the same time that the neighbouring Leipzig. Heidenheim, who showed strong negative effects GDP per capita of their LEZ implementation shows consistent results with its local trade sector being reduced from a 20% to 16% of its the city's GVA.

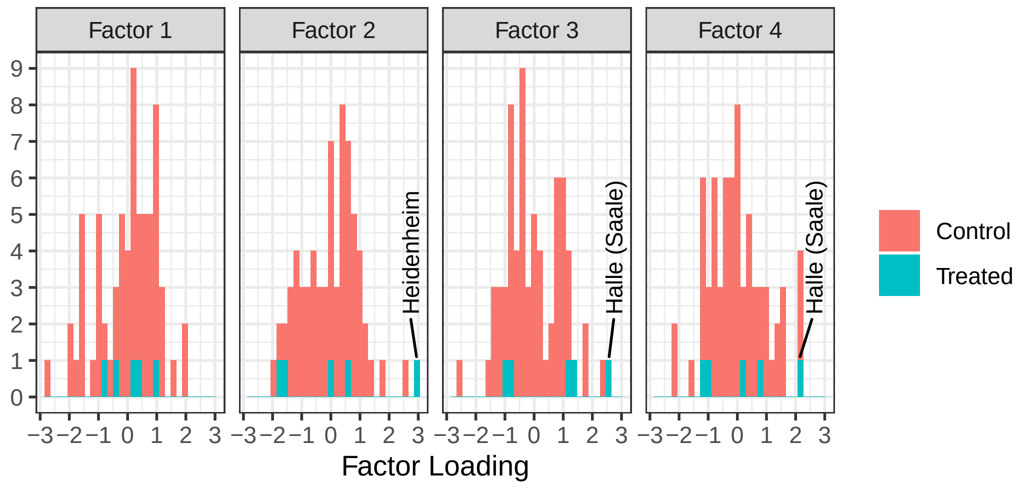
Finally, the effect on Mönchengladbach's LEZ on its local trade is strongly positive, this is probably due to its closeness with the LEZ of Düsseldorf, Krefeld, Neuss and the one for the Ruhr area (the largest LEZ in Germany). All of them were already active when the LEZ for Mönchengladbach (number 43 in Figures 10 and 11) was announced and thus might have absorbed the initial costs of a LEZ while only leaving the benefits of it.

In the estimation of synthetic controls, the number of factors that minimised out-of-sample mean squared error was 4, implying there is large heterogeneity between cities' outcome paths. The evolution of each of the estimated factors is shown in Figure 6a. Regarding factor loadings (Figure 6b) we can see that, although most treated units lie on the common support of control units, the cities of Heidenheim and Halle are sometimes in the edges of the distribution. This implies their synthetic control estimates are constructed with some degree of extrapolation and thus are not as reliable as others.

Although these factors are not usually interpretable, Factor 1 has an almost perfect negative correlation with the mean percentage of GDP represented by local trade. This can be interpreted by looking at Figure 6a and seeing that these cities where local trade has a low weight in the economy saw this value further reduced from 2004 to 2009, in absence of the treatment. Other factors, necessarily octagonal to previous ones, are harder to interpret. Maps of the spatial distribution of all 4 factors can be found in Figure 21, in the Appendix.



(a) Latent factor's evolution during the study period



(b) Histograms of factor loadings (λ_i) for each factor by treatment status.

Figure 6: Factors and factor loadings the effect on local trade (all normalised).

4.3 Potential treatment effect

Various limitations arise when estimating the potential treatment effect of a LEZ on non-treated cities. Although LEZ are standardised in Germany they were announced at different times and implemented with different strengths and calendars. Furthermore, the size of each LEZ relative to the NUTS 3 region (where the GDP per capita is reported) changes from one city to the next. These two dimensions of treatment variability cause that the estimates of potential treatment effects strongly depend on the characteristics of the treated units from where the counterfactual is constructed.

Finally, as LEZ were announced at different years, only treated cities that were announced the same year can be included in the treated pool. This effectively sets the inference of our estimates to what "would have happened" if a LEZ was implemented in these non-treated regions on that specific year.

To simplify the analysis and its results only the year 2011 was chosen to calculate the potential treatment effect, as it was the one when more LEZ were announced in the study period (7). Only 4 of them were well represented by their NUTS 3 regions and though suitable to form the treated pool: Halle (Saale), Magdeburg, Hagen, and Heidenheim.

To reduce uncertainty and dependence of the units in the treated pool the potential treatment effect was estimated for only a set of all non-treated cities. First only cities that had a representative NUTS 3 region and were more than 60km from any actual LEZ were considered. Finally, and given the small set of treated units, the set of non-treated had to be restricted to those that had similar characteristics. The summary statistics of both groups are shown in Table 3.

From it we can see that, although small, the group of treated regions covers most the range of values of non-treated regions.[23] In characteristics such as GDP per capita and the share of GDP of several sectors, the maximum value for non-treated regions falls outside the range of treated ones and both groups have different distributions of the share of industry and local trade in GVA. These differences suggest some caution when interpreting the results as they could create some bias if they are correlated with post-intervention shocks.

Variable	**Status**	**n**	**Min**	**Mean**	**Max**	**Std. dev.**
GDP per capita (€)	Non-treated	3	27911	29211.1	31344	1862
	Treated	4	25011	27863	30027	2203
Population per km^2	Non-treated	3	632	919	1119	255
	Treated	4	211	1081	1754	640
Share of GVA (%) from:						
Industry	Non-treated	3	8.7	15.6	19.7	6.0
	Treated	4	9.5	23.3	43.4	16.2
Construction	Non-treated	3	3.8	4.8	6.6	1.6
	Treated	4	2.9	4.9	6.8	1.9
Local trade	Non-treated	3	19.7	21.4	24.8	2.9
	Treated	4	14.0	18.6	21.7	3.6
Agriculture, forestry and fishing	Non-treated	3	0.1	0.1	0.1	0.0
	Treated	4	0.0	0.3	0.9	0.4
Public sector, education, health and arts	Non-treated	3	28.0	33.7	43.7	8.7
	Treated	4	16.4	30.3	42.0	11.6
Financial, real state, professional, scientific and technical activities	Non-treated	3	23.6	24.5	25.8	1.1
	Treated	4	20.9	22.6	24.7	1.7

Table 3: Summary statistics for potential effect on GDP per capita

The same model as described previously was used to perform the "Synthetic Treated Method" and calculate "synthetic treated" paths for each of the selected non-treated units. Its results are shown on Figure 7.

The results show a consistent and sizeable negative effect of LEZ in GDP per capita of 6-12%. As mentioned before, these results are highly dependent on the set of treated cities where the counterfactual is being constructed and in this case most treated cities

[23]The small sample size of treated regions (who act as controls) is not an issue as the precision of the estimation of potential treatment effects depends on the number of pre-treatment periods and not the size of the control pool.

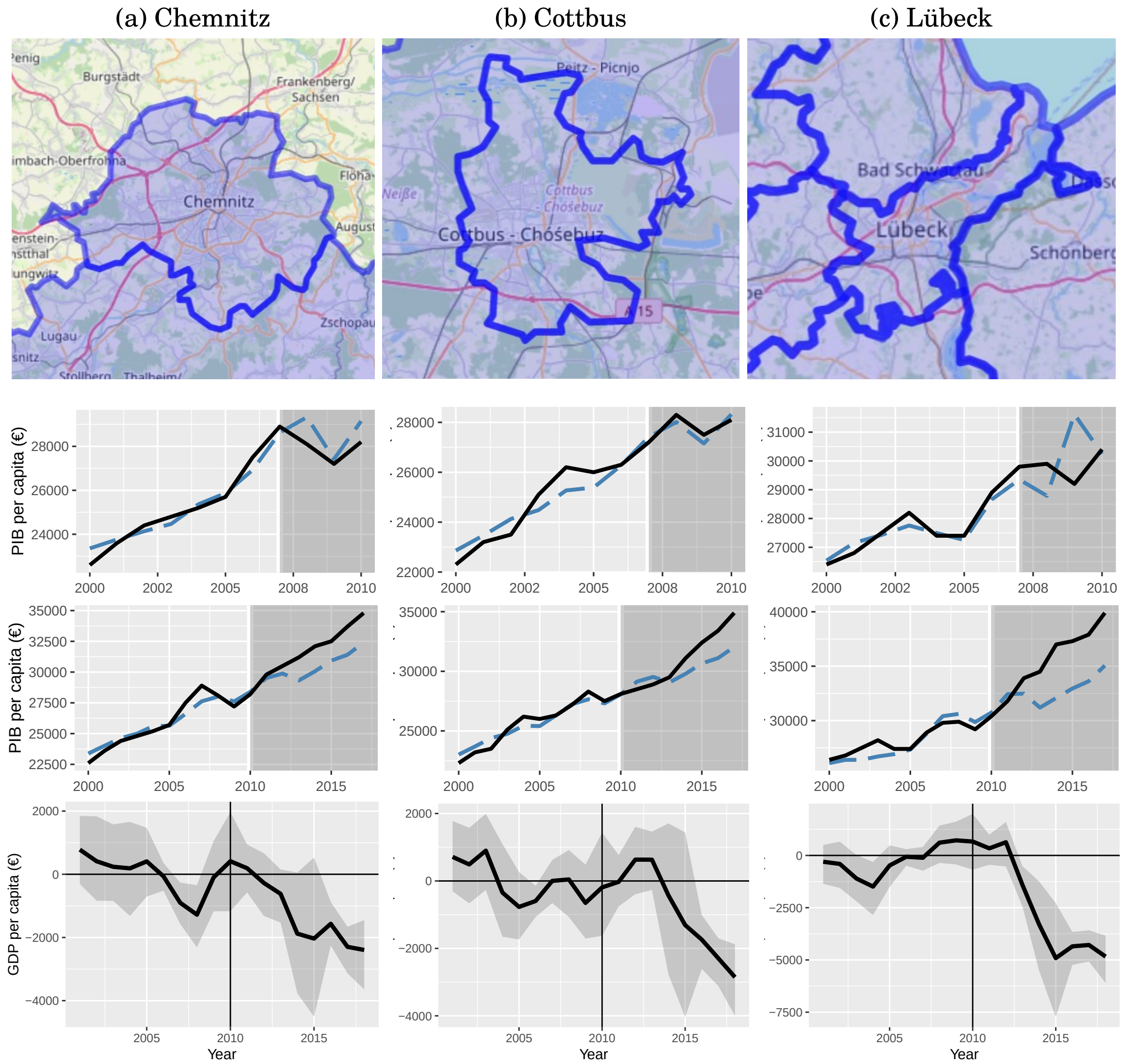


Figure 7: Results for the potential effect of a LEZ smilar to the one in the selected treated cities on Chemnitz, Cottbus and Lübeck. The first row shows a map of each city with its NUTS 3 region and some neighbour control regions (in blue). The second row shows 3-year pre-intervention placebos to look at the robustness of the synthetic controls. Row 3 shows the aggregate paths of the non-treated cities and their "treated controls". Finally the overall effects are presented for each city in the last row. Differences in relative size of LEZ and implementation calendars are not taken into account for the standard errors which are estimated by a boostrap procedure of 1000 iterations.

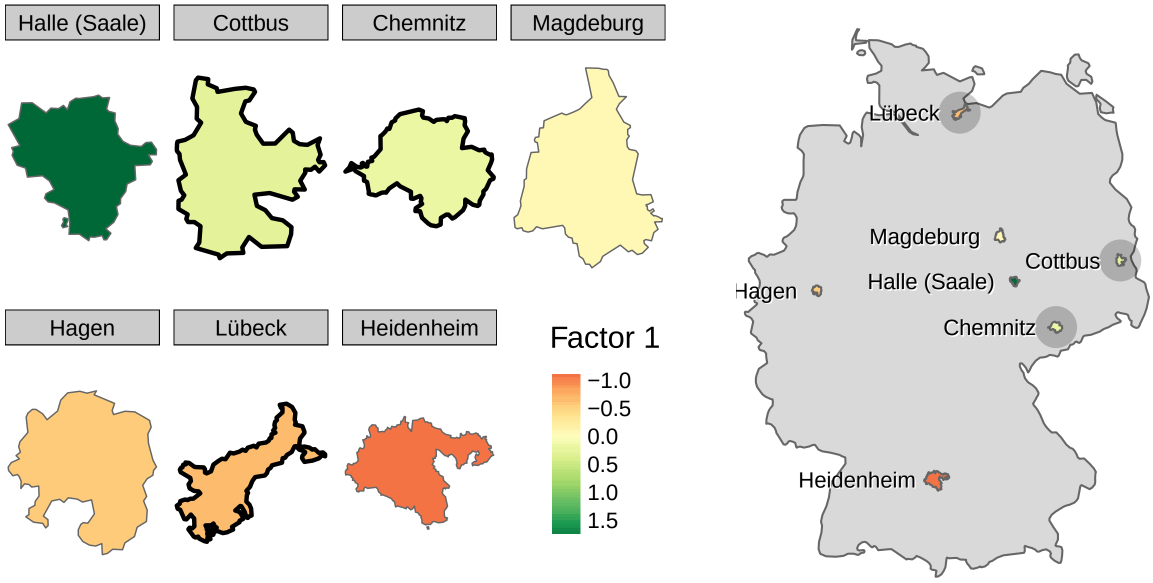
had sharp reductions in GDP per capita after the anouncment of a LEZ, it is the case for Halle, Heidenheim and Magdeburg[24].

These estimates suggest a much stronger negative effect than the ATT estimated for over 10 treated cities (2-3% reductions in GDP per capita). This strong differences could be explained if the assignment mechanism into treatment is sophisticated enough such that the policy is mostly implemented in cities that would have smaller negative effects from its implementation. This is in line with the case that these non-treated cities probably had less air pollution problems when the policy was considered and warns against interpreting the resulting ATT from the previous analysis as a general average treatment effect (ATE).

To illustrate how the potential outcome is constructed the factor loadings of the unique factor identified are shown in Figure 8.[25] From it we can see how non-treated cities are in the common support of treated units, implying that their counterfactual is estimated by interpolation. Furthermore, we can see it perfectly separates cities from West Germany and East Germany and thus gives more "weight" to cities that have similar economic trends. Although there is evidence of convergence in GDP per capita, East and West Germany continue to have major economic differences and thus different paths (Juessen, 2008). This factor controls for such differences and avoids the higher growth of East-German cities bias the potential treatment effect estimates.

[24]Magdeburg suffered from large floods in 2013 and thus all its outcomes after that event are not considered in the calculation. Hagen hadn't a synthetic control robust enough to calculate its treatment effects.

[25]The cross-validation procedure selected only one factor as the alternative that had a smaller out-of-sample error.



(a) Factor loadings (λ_i), East Germany on top

(b) Location of cities

Figure 8: Factor loadings for treated and non-treated cities (marked)

5 Conclusions

Due to a large body of research pointing out the adverse effects of air pollution and a growing political concern since the late 1970s, increasingly strict legislation on air pollution levels was implemented in multiple countries. On this context, LEZ have been widely adopted in European countries with some concern from various economic sectors on its economic impacts, especially the retail and transport sectors.

The positive results of LEZ reducing pollution and improving the overall health outcomes of the population within them have already been documented by previous studies with careful identification strategies of causality. Finally, a recent body of research documents the negative effects of pollution on the economy, especially those through the labour market, indicates that a positive effect of LEZ in the economy was also possible.

Given the aggregate economic effects of LEZ had not yet been studied, this paper tries to estimate them with a Synthetic Control Method strategy based on the *Gsynth* methodology from Xu (2017). The cases used are a set of mid-sized German cities that applied a LEZ from 2009 to 2018.

From the results, individual counterfactuals for 10 treated cities show heterogeneous effects with a significantly negative ATT that implies a reduction of 1240€ ($\approx$ 2.5-3%) in GDP per capita, 4 years after the announcement of a LEZ. Focusing on the effect of LEZ in local trade[26] an estimate on the how LEZ affected its share of regional GVA was also calculated. The results, for 5 different German cities, show important heterogeneity with both significantly positive and negative results. Cities that were surrounded by already-implemented LEZ had positive effects, while

[26] as defined by codes G-J from the NACE Rev.2 classification and including the Retail and Transport sector.

those that had negative effects over GDP per capita also had strong reductions on the share of regional GVA coming from local trade, in line with the sector's concerns over LEZ. Furthermore, I found no significant effect on the period between the announcement and actual application of the policy, neither for the GDP per capita or the proportion of local GVA represented by local trade.

Finally, the calculation of potential treatment effects involved using the treated units to construct "synthetic treated" counterfactuals for non-treated units and estimate what would have happened if the policy was implemented. The results, only robust for 3 small cities, show negative potential treatment effects of LEZ, causing a reduction of $\approx$ 3.000-5.000€ in GDP per capita ($\approx$ 6-12%), 8 years after its announcement. These estimates warn against interpreting the ATT estimate as an ATE and suggest there might be a mechanism of selection into treatment where cities with fewer potential costs of applying a LEZ have a higher probability of implementing this policy.

Overall, the results show that, for the cities studied, the announcement and application of LEZ had either no impact or a sizeable negative impact in the economy. Furthermore, LEZ tended to decrease the weight local commerce had on a city's GVA, in line with critiques from the retail and transport sector. Finally, the application of LEZ in some non-treated cities could have caused sizeable negative effects in their local economies, 2 to 4 times bigger than in those there the policy was actually implemented.

Although the benefits of LEZ reducing air pollution and improving overall health have been already studied by the literature, my conclusions invite to carefully evaluate the possible economic impacts of these policies and work to avoid its negative effects, especially on local trade.

While some cities appear to experience strong negative effects of LEZ, it is also true that others do not. There is even the case of the city of Mönchengladbach in which the LEZ had significant positive effects in the weight of local trade in its economy. Further analysis is needed to complement my results and understand the reasons for these sharp differences. This would help to better understand how different economies react to LEZ and inform future policy decisions on the matter.

Limitations of the results are derived from assumptions of the model, its functional form and the data available. The finite sample properties and bias bounds of the model are studied in Xu (2017), Samartsidis et al. (2019) and Abadie et al. (2010). This last paper shows that a combination of few pre-treatment periods, large transitory shocks and a large and unfiltered control pool will result in biases.

It seems correct to limit the external validity of my results to countries that have strong economic similarities with Germany such as Belgium, The Netherlands and Austria. Other European countries seem to have important differences in economic trends and characteristics. The fact that no robust synthetic controls could be constructed if areas from outside these countries were included further validates this point. To suggest that the effects estimated here apply to other non-studied cities the city-specific characteristics and the LEZ application should be similar to one of the cities we have results of. Finally, it is not recommended to extrapolate the results to developing countries where air pollution levels are usually much higher, cars tend to be older and more pollutant and individuals have less disposable income to buy a new vehicle or modify its emissions standard.

The judgement of external validity, or the assessment of how the results for treated units can help to understand the potential treatment effects of non-treated units, is usually a hard and ambiguous task. The idea that potential treatment effects can be estimated with the same methods of causal inference already used in the literature by modifying the position of treated and control units provides a new set of tools for researchers to formalise the extent of external validity. The introduction of the "Synthetic Treated Method" is an example of such new tools. A wide range of studies could be performed with this technique, especially evaluations of non-random treatment assignments, where the expected causal effect on the treated should differ from the effect on non-treated units as the treatment assignment is probably correlated with its expected effect.

I believe further work is needed to validate and complement my results. Given the positive effects in the economy are expected to be mediated by the reduction of pollution levels, the use of detailed air pollution data and a technique that could separate the causal path of reduced air pollution would be especially insightful. It would allow researchers to better understand the reasons for the effects documented in this paper. Imai and Keele (2010) give a general introduction on mediation causal methods to guide further research. Specifically, the "Mediation Analysis Synthetic Control" methodology introduced in Mellace and Pasquini (2019) could be of special interest to perform this kind of analysis if detailed pollution data could be used.

References

Abadie, A. (2019). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, Forthcoming.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1):113–132.

Abadie, A. and L'Hour, J. (2019). A Penalized Synthetic Control Estimator for Disaggregated Data. *Working Paper,*, (May):1–35.

Adhvaryu, A., Kala, N., and Nyshadham, A. (2014). Management and shocks to worker productivity: evidence from air pollution exposure in an Indian garment factory. *Unpublished Working Paper, University of Michigan*.

Aragón, F. M., Miranda, J. J., and Oliva, P. (2017). Particulate matter and labor supply: The role of caregiving and non-linearities. *Journal of Environmental Economics and Management*, 86:295 – 309. Special issue on environmental economics in developing countries.

Badische Zeitung (2009). Wird die Feinstaubplakette zum Umsatzkiller? *Badische Zeitung*. `https://web.archive.org/web/20100315222338/https://www.badische-zeitung.de/wird-die-feinstaubplakette-zum-umsatzkiller--22330565.html`, accessed on 22-07-2020.

Brodersen, K. H., Gallusser, F., Koehler, J., Remy, N., and Scott, S. L. (2015). Inferring causal impact using bayesian structural time-series models. *Annals of Applied Statistics*, 9:247–274.

Browne, M., Allen, J., and Anderson, S. (2005). Low emission zones: the likely effects on the freight transport sector. *International Journal of Logistics Research and Applications*, 8(4):269–281.

Calderón-Garcidueñas, L., Torres-Jardón, R., Kulesza, R. J., Park, S.-B., and D'Angiulli, A. (2014). Air pollution and detrimental effects on children's brain. The need for a multidisciplinary approach to the issue complexity and challenges. 8.

Cesaroni, G., Boogaard, H., Jonkers, S., Porta, D., Badaloni, C., Cattani, G., Forastiere, F., and Hoek, G. (2012). Health benefits of traffic-related air pollution reduction in different socioeconomic groups: the effect of low-emission zoning in Rome. *Occupational and Environmental Medicine*, 69(2):133–139.

Chang, T., Zivin, J. G., Gross, T., and Neidell, M. (2016). The effect of pollution on worker productivity: Evidence from call-center workers in China. Working Paper 22328, National Bureau of Economic Research.

Dablanc, L. and Montenon, A. (2015). Impacts of environmental access restrictions on freight delivery activities: Example of low emissions zones in europe. *Transportation Research Record*, 2478(1):12–18.

Dechezleprêtre, A., Rivers, N., and Stadler, B. (2019). The economic cost of air pollution: Evidence from Europe. *"OECD Economics Department Working Papers"*.

Du, Y., Xu, X., Chu, M., Guo, Y., and Wang, J. (2016). Air particulate matter and cardiovascular disease: the epidemiological, biomedical and clinical evidence. *Journal of thoracic disease*, 8(1):E8—E19.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8(4):36–65.

EPA (2004). Air quality criteria for particulate matter. volume ii. Washington D.C, `https://cfpub.epa.gov/si/si_public_record_report.cfm` (Accessed 01/06/2020).

European Comission (2013). Impact Assessment. *European Comission*. `https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013SC0531&from=EN`,

accessed on 05-08-2020.

European Environmental Agency (2019). Air quality in europe - 2019 report. EEA Report No 10/2019.

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

Gobillon, L. and Magnac, T. (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. 98(3):535–551.

Graff-Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–73.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. *Journal of Economic Literature*, 51(3):689–730.

Hanna, R. and Oliva, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in Mexico city. *Journal of Public Economics*, 122:68 – 79.

Hao, Y., Peng, H., Temulun, T., Liu, L.-Q., Mao, J., Lu, Z.-N., and Chen, H. (2018). How harmful is air pollution to economic development? new evidence from pm2.5 concentrations of Chinese cities. *Journal of Cleaner Production*, 172:743 – 757.

Holman, C., Harrison, R., and Querol, X. (2015). Review of the efficacy of low emission zones to improve urban air quality in European cities. *Atmospheric Environment*, 111:161 – 169.

Holub, F., Hospido, L., and Wagner, U. J. (2006). Air pollution and labor supply: Evidence from social security data. *Unpublished Working Paper*.

Imai, K. and Keele, L. (2010). A general approach to causal mediation analysis. *Psychol Methods*, 15(4):309–34.

Juessen, F. (2008). A distribution dynamics approach to regional gdp convergence in unified germany. *Empirical Economics*, 37(3):1435–8921.

Kaul, A., Pfeifer, G., and Witte, S. (2012). The incidence of cash for clunkers: an analysis of the 2009 car scrappage scheme in germany. *International Tax and*

Public Finance, 23.

Lindstaedt, B. (2009). Einzelhandel, kundenverkehr und umweltzonen. Institut fuer Handelsforschung (IfH) Seminar Presentation.

Liu, L. and Xu, Y. (2018). panelview: an r package of visualizing panel data. Available at http://yiqingxu.org/software/panelView/panelView.html.

Mellace, G. and Pasquini, A. (2019). Identify more, observe less: Mediation analysis synthetic control.

Morfeld, P., Groneberg, D. A., and Spallek, M. F. (2014). Effectiveness of Low Emission Zones: Large Scale Analysis of Changes in Environmental NO2, NO and NOx Concentrations in 17 German Cities. *PLoS ONE*, 9(8):e102999.

Morgenstern, R. D., Pizer, W. A., and Shih, J.-S. (2002). Jobs versus the environment: An industry-level perspective. *Journal of Environmental Economics and Management*, 43(3):412 – 436.

Pestel, N. and Wozny, F. (2019). Low Emission Zones for Better Health: Evidence from German Hospitals. *IZA – Institute of Labor Economics*, (IZA DP No. 12545).

Ranft, U., Schikowski, T., Sugiri, D., Krutmann, J., and Krämer, U. (2009). Long-term exposure to traffic-related particulate matter impairs cognitive function in the elderly. *Environmental Research*, 109(8):1004 – 1011.

Ransom, M. R. and Pope, C. A. (1992). Elementary school absences and pm10 pollution in Utah Valley. *Environmental Research*, 58(1):204 – 219.

Roth, S. (2015). The effect of indoor air pollution on cognitive performance: Evidence from the UK. Unpublished.

Rufibach, K. (2009). reporttools: R functions to generate LaTeX tables of descriptive statistics. *Journal of Statistical Software, Code Snippets*, 31(1).

Samartsidis, P., Seaman, S. R., Presanis, A. M., Hickman, M., and De Angelis, D. (2019). Assessing the causal effect of binary interventions from observational panel data with few treated units. *Statistical Science*, 34(3):486–503.

Shaddick, G., Thomas, M. L., Green, A., Brauer, M., van Donkelaar, A., Burnett, R., Chang, H. H., Cohen, A., Dingenen, R. V., Dora, C., Gumy, S., Liu, Y., Martin, R., Waller, L. A., West, J., Zidek, J. V., and Prüss-Ustün, A. (2018). Data integration model for air quality: a hierarchical approach to the global estimation of exposures to ambient air pollution. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 67(1):231–253.

Tennekes, M. (2018). tmap: Thematic maps in R. *Journal of Statistical Software*, 84(6):1–39.

Viana, M., Kuhlbusch, T., Querol, X., Alastuey, A., Harrison, R., Hopke, P., Winiwarter, W., Vallius, M., Szidat, S., Prévôt, A., Hueglin, C., Bloemen, H., Wåhlin, P., Vecchi, R., Miranda, A., Kasper-Giebl, A., Maenhaut, W., and Hitzenberger, R. (2008). Source apportionment of particulate matter in europe: A review of methods and results. *Journal of Aerosol Science*, 39(10):827 – 849.

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

Wolff, H. and Perry, L. (2010). Policy Monitor: Trends in Clean Air Legislation in Europe: Particulate Matter and Low Emission Zones1. *Review of Environmental Economics and Policy*, 4(2):293–308.

Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1):57–76.

Zhang, X., Chen, X., and Zhang, X. (2018). The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences*, 115(37):9193–9197.

Appendices

A Github repository with the code to replicate my analysis is avaliable at `https://avila-a.github.io/Capstone-Project-ASDS/`. It also includes some data, notes, and interactive maps to explore the set of treated and control regions.

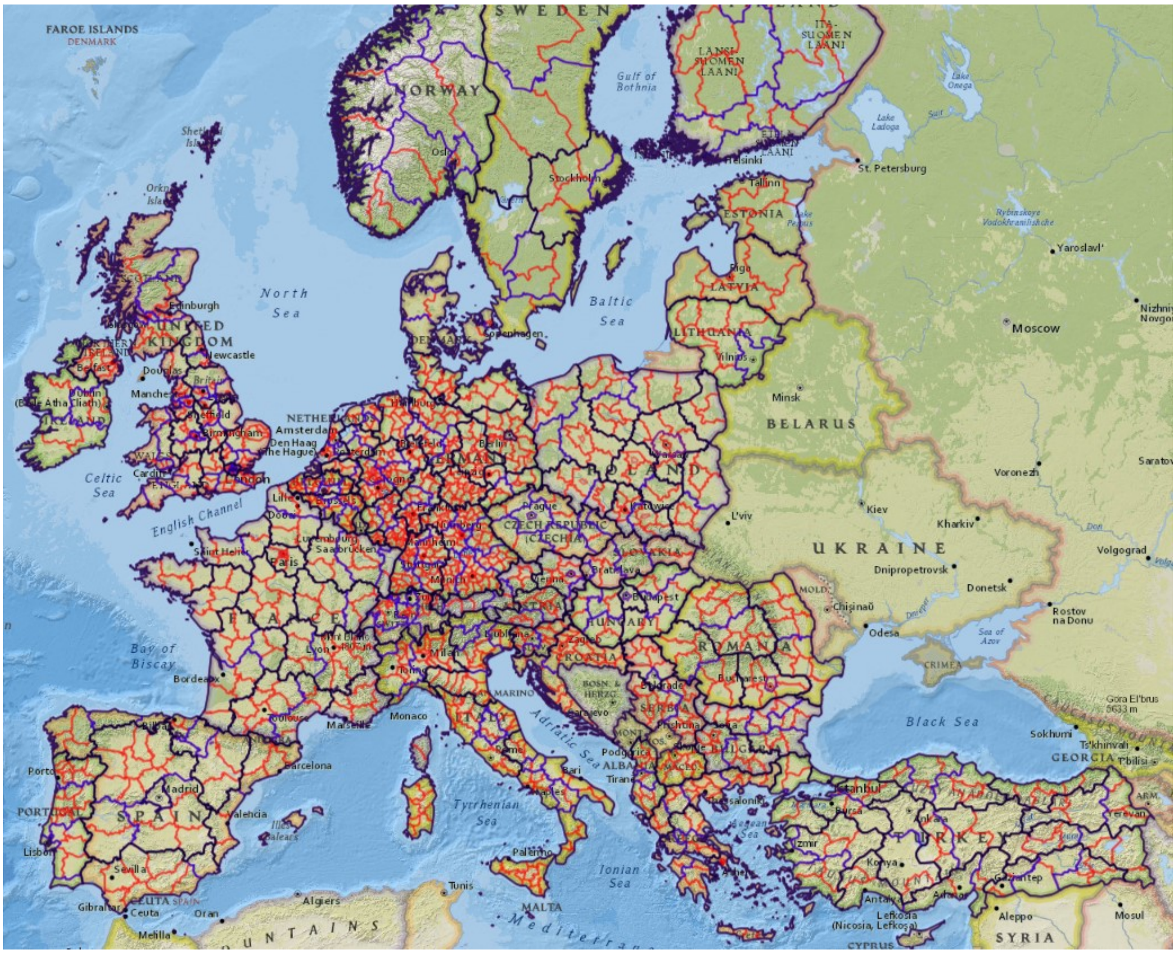


Figure 9: Main coverage of NUTS zones across Europe and Germany, with examples of cities. Borders of NUTS 1, 2 and 3 regions are coloured in black, blue and red, respectively. Some regions are excluded.

Sources: Eurostat, OpenStreetMap and own work. © *EuroGeographics*

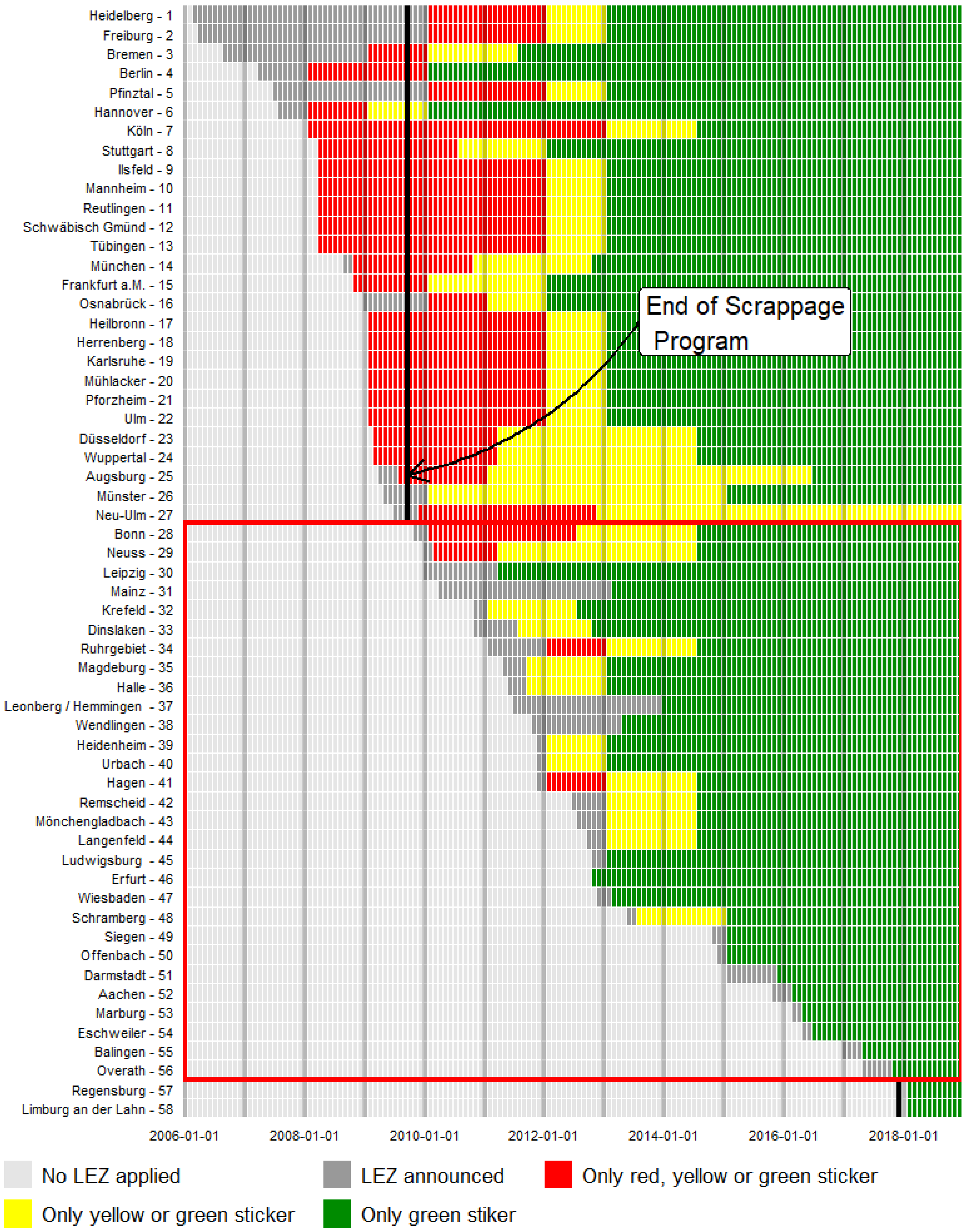


Figure 10: Application of LEZ in Germany during the period studied by announcement date and category of environmental stickers allowed. LEZ included in the sample are marked with a red square and the end of the German scrappage program is marked with a black vertical line. The last 2 cities are not included in the study for lack of post-intervention data.

Sources: Umweltbundesamt, local "Environmental plans" and my own work.

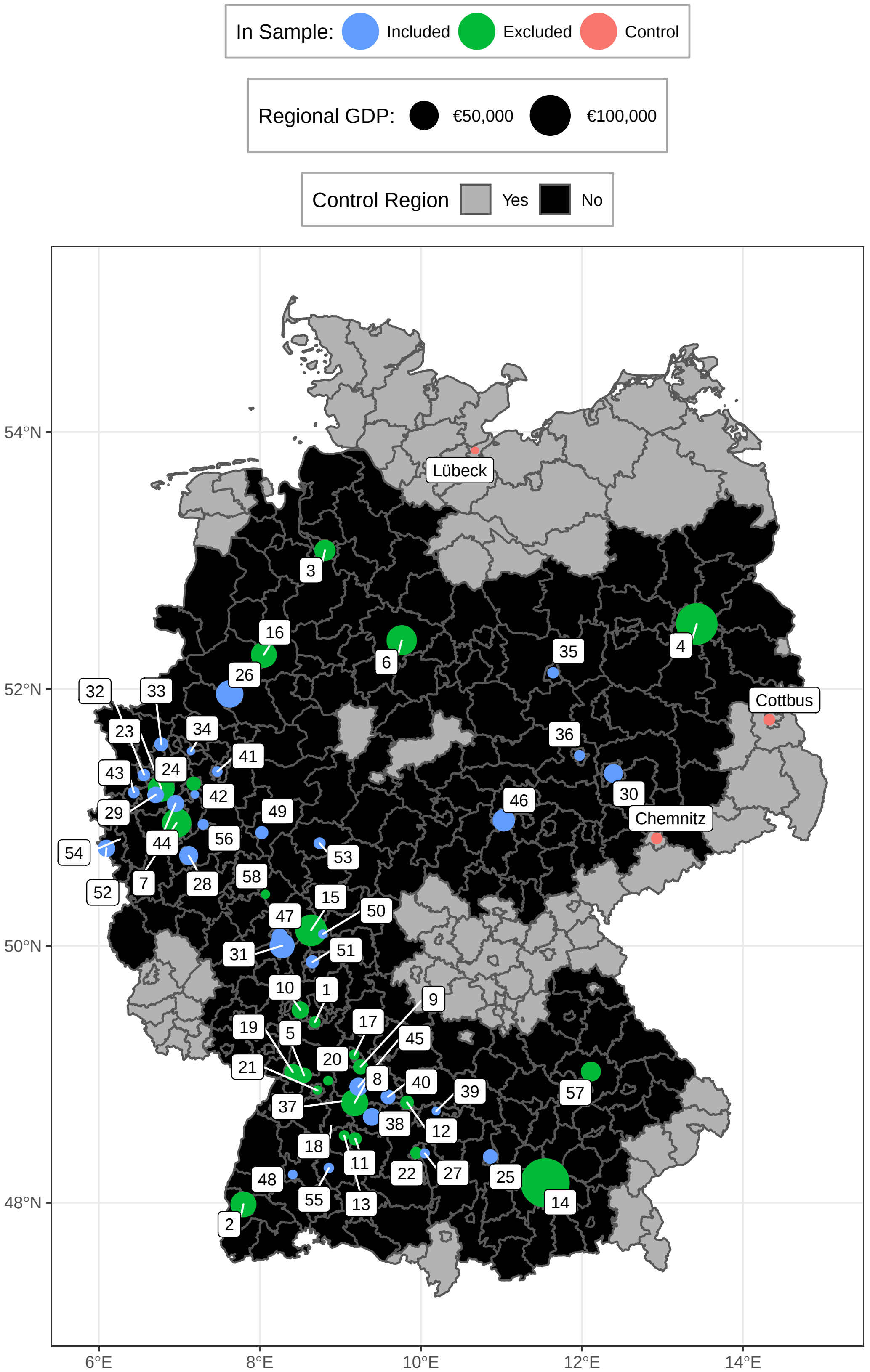
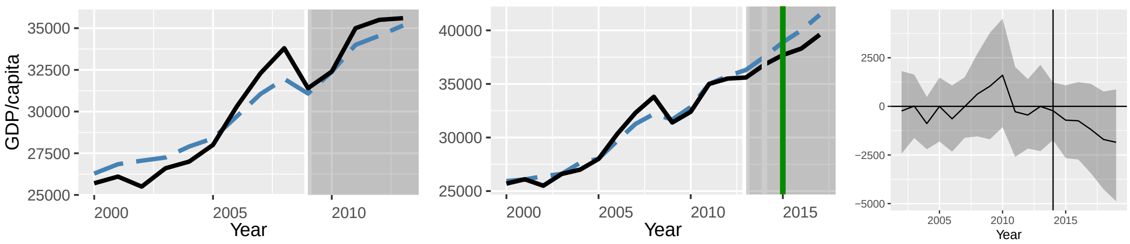


Figure 11: Location of all LEZ and their NUTS 3 regions. Number labels are ordered by announcement date as in Figure 10 and allow to identify each city and if its included or not in the sample of treated regions. Regions are marked in gray if they were not in the UAR or Green Zones databases and were more than 60km away from any LEZ and thus work as controls. The names of 3 control cities where the the "Synthetic Treated Method" is applied are also visible and marked as "Control".

Sources: Umweltbundesamt, Eurostat and own work. *© EuroGeographics for the admin. boundaries.*

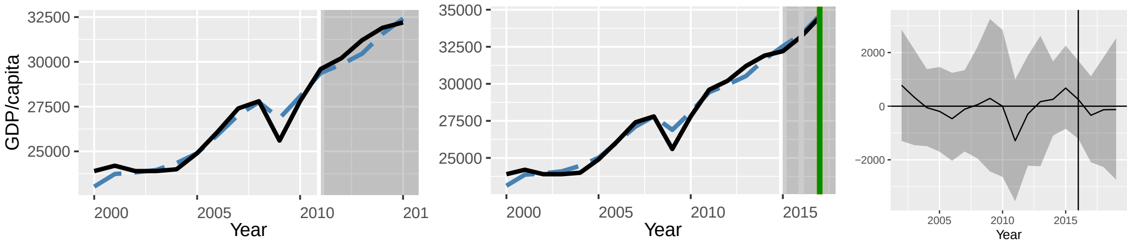
Zoomable figure

Additional results for the effect of LEZ in local GDP per capita:



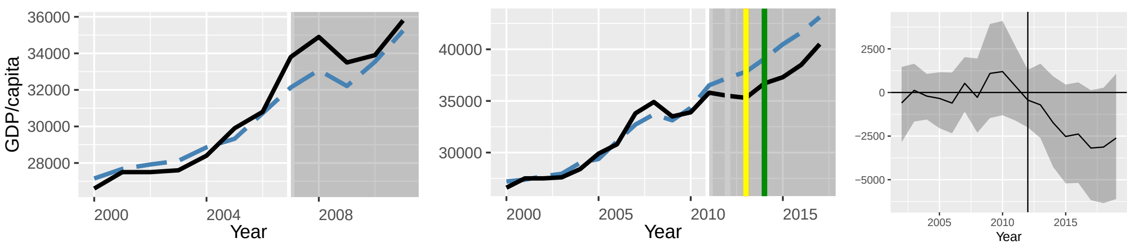
(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 12: Results for the effect of Siegen LEZ in its GDP per capita



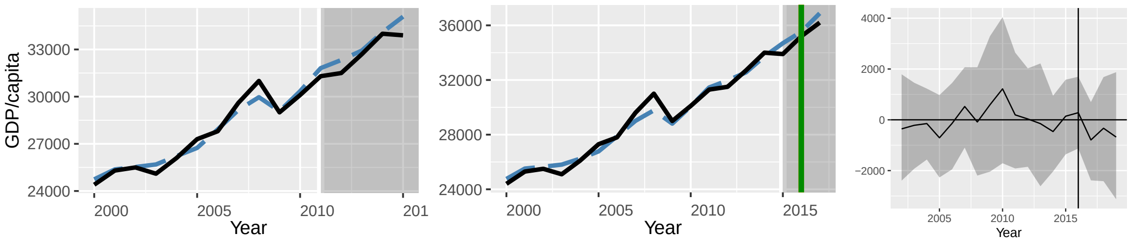
(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 13: Results for the effect of Balingen LEZ in its GDP per capita



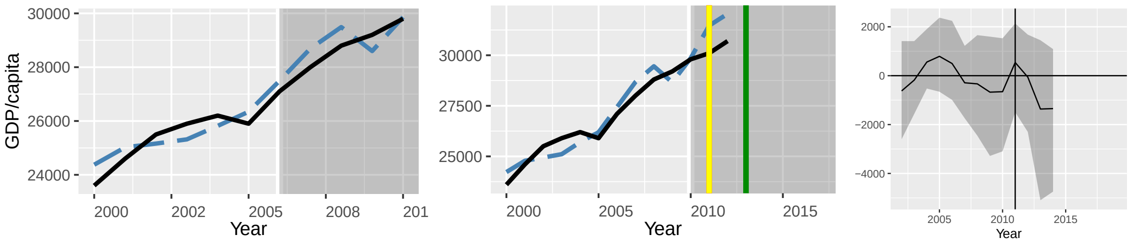
(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 14: Results for the effect of Langenfeldt LEZ in its GDP per capita



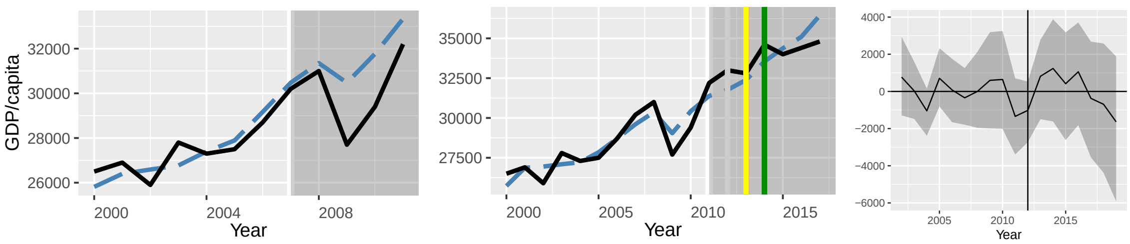
(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 15: Results for the effect of Marburg LEZ in its GDP per capita



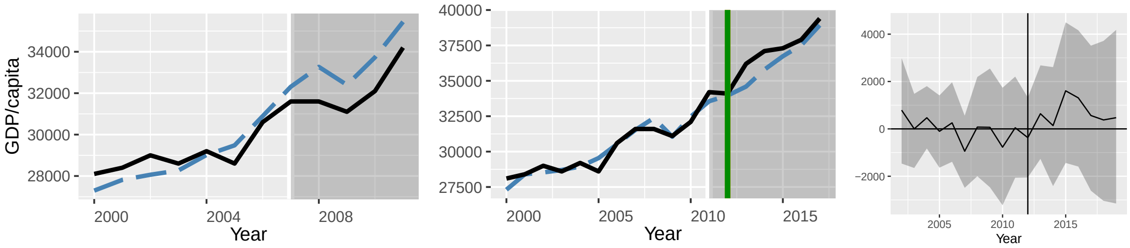
(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 16: Results for the effect of Magdeburg LEZ in its GDP per capita. No effect is estimated after the large floods of 2013.



(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

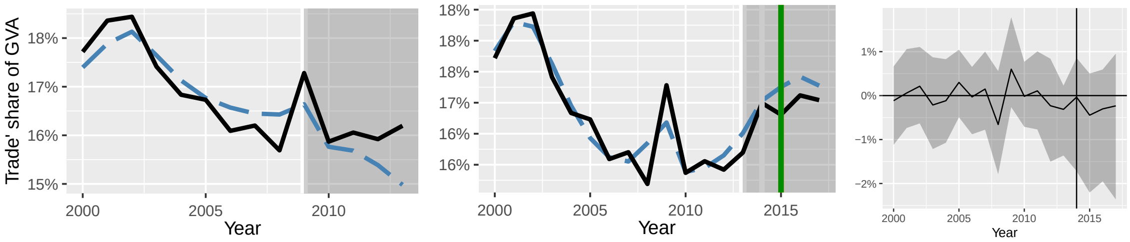
Figure 17: Results for the effect of Remsheid LEZ in its GDP per capita.



(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

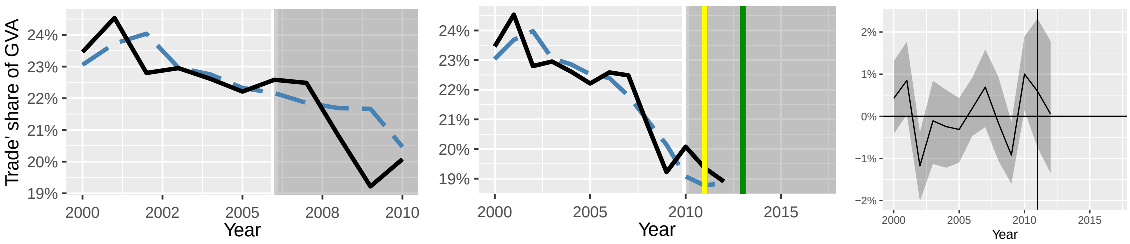
Figure 18: Results for the effect of Erfurt LEZ in its GDP per capita.

Additional results for the effect of LEZ in the share of GVA from local trade:



(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 19: Results for the effect of Siegen LEZ in its share of GVA from local trade.



(a) Pre-intervention placebo (b) Outcome paths (c) Estimated effect

Figure 20: Results for the effect of Magdeburg LEZ in its share of GVA from local trade.

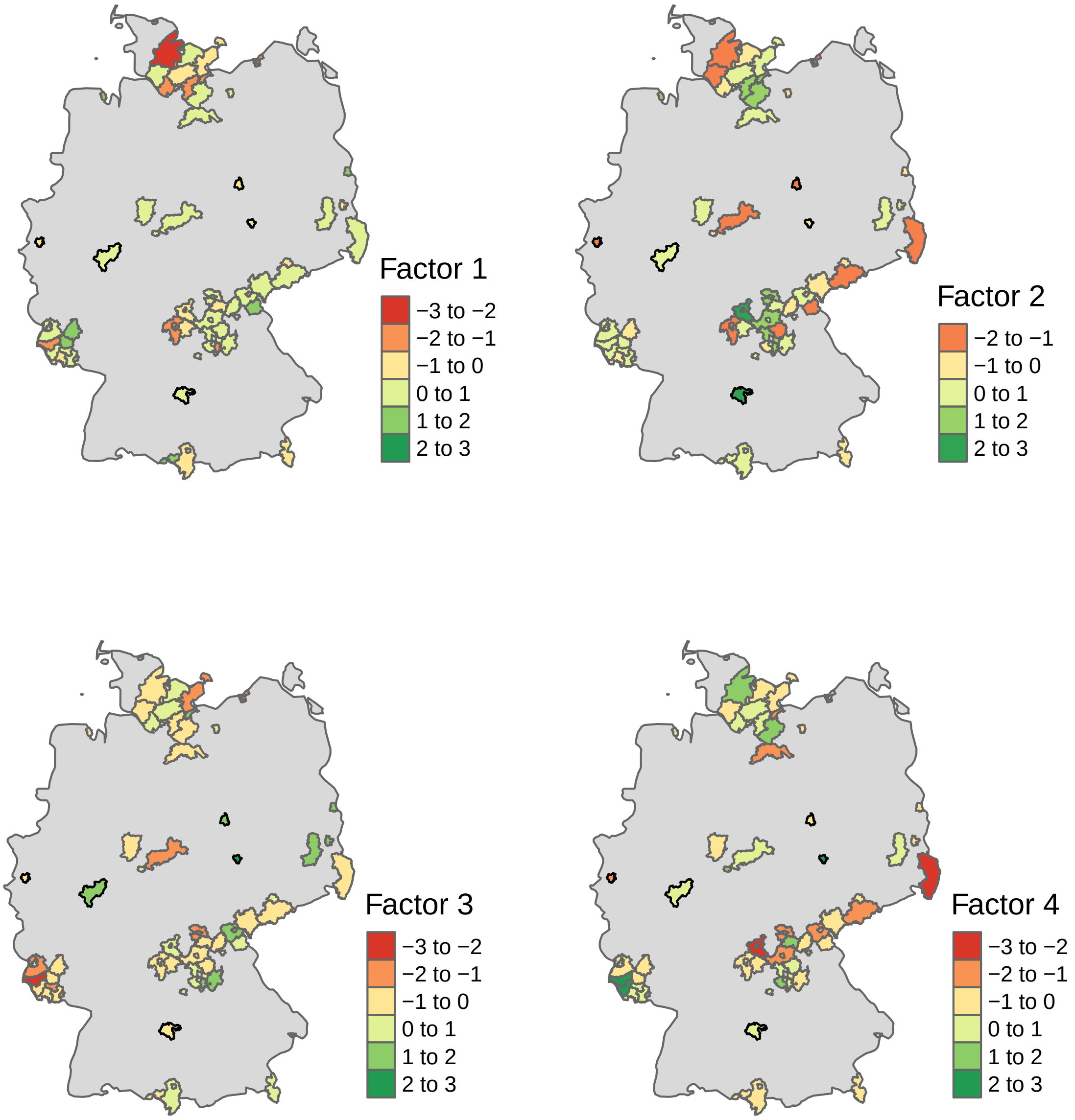


Figure 21: Standarized factor weights for the 4 factors identified to model the share of GVA from local trade. Treated regions are marked with black borders. The appearance of geographical clusters suggests that the underlying unobservables represented by a given factor should respond to local phenomena. ***Zoomable figure***