Urban Transport Initiatives and Pollution Patterns: Comparing Low Emission Zones and Metro Expansions via Spatial-Temporal Difference-in-Differencing

Jacob L. Macdonald
University of Liverpool
Geographic Data Science Lab
Consumer Data Research Centre

September 20, 2019

Abstract

This research makes use of a high-dimensional neighbourhood-scale database to estimate the pollution mitigating impacts of urban public transportation infrastructure. Short and long-run localized pollution reductions are estimated and attributed to the expansion of the underground metro system of Lisbon, Portugal, and the introduction of a targeted low emission zone in the city centre.

Geostatistical methods are used to process high-frequency and high-resolution open source pollution measures to generate a longitudinal neighbourhood level monthly pollution database spanning since 2000. A large focus compares how fixed-point measures can be interpolated and aggregated across space and time by evaluating robust diagnostics of Kriging and Inverse Distance Weight interpolations. Generalized sets of diagnostics and algorithms are used to select the best, most consistent, model for each of six pollutants.

Long-run effects are estimated under a spatial-temporal difference-in-difference strategy to obtain the average treatment effect of a transit intervention on neighbourhoods in key areas of the city. Results indicate that the expansion of new metro stations have resulted in a decrease in pollution primarily in the city centre and in locations around the newly opened metro stations. Short run localized reductions of PM_{10} immediately following the opening of a metro range up to 2% with longer run reductions of 0.18%. Metro openings had a particularly large impact on decreasing nitrogen-based combustion emission along the riverfront with short-run reductions of 20%, dissipating over time and space.

The LEZ by comparison decreased pollution in and around its boundary and expanded into the city centre with immediate reductions in the subsequent months by up to 4% of PM_{10} and long-run impacts of 0.43%. Some evidence suggests however that the introduction of the LEZ may have shifted pollution patterns elsewhere with increases just outside the boundary and along the Tagus river. The LEZ zone in particular had significant reduction impacts on SO_2 concentration where metro openings did not, capturing the policy's aim of reducing the heaviest polluting vehicles, often running on diesel fuel.

1. Introduction

Broad and efficient transportation networks and infrastructure are core characteristics of an attractive and liveable city. Public transit initiatives, such as the opening of new underground metro stations or the introduction of zonal traffic restrictions, not only influence the daily movement of residents, workers and visitors, but indirectly have important spill-over effects by impacting spatial and temporal patterns of urban air pollutants. In 2010, transportation accounted for almost a quarter of all emissions generated across the globe, of which 40% was from urban transportation specifically (Sims et al. 2014). Local urban transport policies and best practices can therefore yield important contributions to larger scale pollution mitigation and abatement.

Estimating how the introduction of various transit initiatives influence an area's pollution dynamics enables planners and local authorities to evaluate non-monetary environmental benefits and further enact best practices. Challenges exists however in studying such spatially and temporally granular urban dynamics using available open source data.

This work uses geostatistical methods to process high-frequency and high-resolution open source pollution measures to value various transit initiatives in terms of their contribution to the reduction of airborne pollutants. A large focus is on how fixed-point measures can be interpolated and aggregated across space and time. Under hyperparameter optimization, enhanced by the inclusion of temporal lags of predicted pollution, Kriging and Inverse Distance Weight (IDW) families of interpolation are used to generate neighbourhood level monthly pollution concentrations. A variety of parameter, variable and specification choices for each model are compared to ensure the strongest prediction, and generalized sets of diagnostics and algorithms are used to select the best, most consistent, model for each.

The spatial and temporal nature of pollution monitoring is used to generate a neighbourhood level monthly longitudinal database to explore how neighbourhood pollution

concentration has been affected by transportation policy over the past two decades. This high-dimensional database is used to estimate the pollution abating impacts of urban public transportation infrastructure. In particular, the short and long-run localized pollution reductions surrounding the expansion of the underground metro system of Lisbon, Portugal, and the introduction, of a series of traffic based configurations and a targeted low emission zone (LEZ) in the city centre aimed at limiting congested flows of high polluting vehicles.

Long-run effects are estimated under a spatial-temporal difference-in-difference strategy to obtain the average treatment effect of a transit intervention on neighbourhoods in key areas of the city. With limited observations in shorter-run time spans surrounding an intervention, bootstrapping is conducted to provide valid difference-in-difference estimates for immediate effects. This allows for the estimation of month-to-month decaying impacts following the introduction of transit initiatives and further highlights this behaviour over space.

Results indicate that the expansion of new metro stations have decreased pollution primarily in the city centre and around newly opened stations. Short run localized reductions of PM_{10} immediately following the opening range up to 2% with longer run reductions of 0.18%. Metro openings had a particularly large impact on decreasing nitrogen-based combustion emission along the riverfront with short-run reductions of 20%, dissipating over time and space.

Interventions surrounding the LEZ by comparison decreased pollution in and around its boundary and into the city centre with immediate reductions in subsequent months up to 4% of PM_{10} and long-run impacts of 0.43%. Some evidence suggests however that the introduction of the LEZ may have shifted pollution patterns with increases just outside its boundary and along the Tagus river. The LEZ had significant reductions on SO_2 where metro openings did not, capturing the policy's aim of reducing the heaviest polluting vehicles, often running on diesel fuel.

Granular data, from a spatial and temporal dimension, will increasingly continue to shape the valuation of detailed urban-environmental processes. This work develops a set of generalizable criteria and diagnostic selections from which sophisticated geostatistical and temporal methods are used to generate measures of the urban environment. Increasingly detailed data yields increasingly detailed applications, and this work highlights the benefit, in the resolution of spatial and temporal impact evaluations, that can be had by leveraging the use of geostatistical methods in the urban context.

2. Literature Review

Underground metro systems, and regional public transit in general, have a wide range of impacts and spill-over effects as accessibility increases. Not only do network expansions have positive impacts on ridership (Baum-Snow and Kahn 2000, Baum-Snow and Kahn 2005, Goetzke 2008, Zhang et al. 2017) and congestion (Anderson 2014, Adler and van Ommeren 2016), but additional influences are often felt in other markets. This could include impacts to local property prices (Bowes and Ihlanfeldt 2001, Martinez and Viegas 2009, Mohammad et al. 2013, Mulley and Tsai 2016, Li 2018, Mulley et al. 2018), land use and spatial distributions (Cervero and Kang 2011, Roukouni et al. 2012, Gonzalez-Navarro and Turner 2018), or even local labour markets (Sanchez 1999, Kawabata and Shen 2007).

From an environmental perspective, one of the most important spill-over impacts of public transit accessibility is change to local pollution. The evaluation of these impacts at a detailed spatial and temporal resolution however is challenged by the limitations of accurate and timely data. Many works have examined this relationship at varying scales ranging from in-situ sampling (Meinardi et al. 2008, Pereira et al. 2013) or more model-based framework (Anselin and Gallo 2006, Chen and Whalley 2012, Bertazzon et al. 2015).

A common theme across the literature is the identification of shocks to the transit system, often in the form of station openings, in order to estimate the resulting change in

pollution as a proxy for the contribution of public transit. Gendron-Carrier et al. (2018) exploit cross-city variation in subway systems to estimate the impacts on particulate matter one year and a half before and after the opening of respective stations. Across cities, results indicate average reductions of 4% extending 10-kilometres from the city centre. In Granada, the expansion and restructuring of the public transportation network reduced PM_{10} concentrations by up to 33% (Titos et al. 2015). In terms of opening of new stations in an urban setting, Zheng et al. (2019) estimate a difference-in-difference reduction effect on carbon monoxide in the areas surrounding the new subway lines.

One of the most common municipal transport policies used to mitigate pollution is the introduction of zonal based traffic restrictions in key congested areas of a city. Currently across Europe there are total of 264 LEZ's varying in scope and extent (Santos et al. 2019). These zones are geographically delineated areas with targeted enforcement focused on restricting heavy polluting vehicles. Different versions of this type of policy have been enacted in different contexts and could vary based the scope (e.g. restriction based on time of day, year or type of car, or licence plate number) or based on the manner of enforcement (e.g. ticket citations, camera detection, seriousness of fine) (Wolff and Perry 2010, Holman et al. 2015, Zhang et al. 2017).

In general, the introduction of LEZ's in different contexts has led to reductions in local pollution levels. However, the outcome and mechanisms through which pollution is potentially changed is not straightforward. While we would expect direct pollution abatement due to the restriction of vehicles in these zones, the introduction of regulations could have unintended consequences altering commuter or broader transport networks and resulting in behavioural change which can be difficult to capture with available data.

Many studies focus on comparing monitored values inside and outside of LEZ boundaries to estimate differences in pollution (Nunes da Silva et al. 2014). Following the five

years after the implementation of the LEZ in London, Ellison et al. (2013) estimate average reductions from 2.46% to 3.07% using point estimates from four monitoring stations comparing those inside and out of the boundary. Results suggest that effects may be temporary with concentrations reverting towards original levels after some period.

Complementing many studies of LEZ is in-situ measurement of vehicle fleet data to better link underlying pollution reductions to specific mechanism (Ellison et al. 2013, Ferreira et al. 2012). In the Lisbon context, Ferreira et al. (2012) estimate the impact of the LEZ on PM_{10} and NO_2 between 2011 and 2013 by comparing observed effects from the *Avenida da Liberdade* (inside), *Entrecampos* (boundary) and *Olivais* (outside) stations. The reduction in pollution is linked to observed vehicle distributions and ages at different points in time in or around the LEZ. The primary mechanism through which pollution is influenced is more from changes to the traffic composition, removing old fuel-inefficient cars, rather than reductions in traffic volume.

Further behaviour style changes have been noted after the introduction of a LEZ which can be linked to urban pollution. Across Germany, Wolff (2014) estimate an average decline in pollution of around 9% in urban areas, primarily attributed to shifting to greener and less polluting transport modes. Xu et al. (2015) meanwhile highlight the substitution behaviour as people may switch to public transit following the introduction of private driving restrictions.

At a more macro-scale, Boogaard et al. (2012) estimate the average impact of LEZ's introduced in five Dutch cities targeting heavy-duty trucks by comparing urban and suburban monitoring station values pre and post introduction. While estimated effects show a general overall decline in pollution, considering specifically the urban and suburban stations showed no significant differences. Viard and Fu (2015) use a multivariate regression to model aggregate city-level pollution in Beijing controlling for temporal breaks, weekends, holidays,

weather patterns and different transport policies introduced. The authors estimate a decrease in average pollution around 21% from the introduction of a one-day-a-week driving restriction.

Santos et al. (2019) look at relative reductions in pollution levels in Lisbon following the introduction of the LEZ via a multi-dimensional factor design considering a temporal and a spatial dynamic to compare areas before and after the introduction. Estimates indicate that pollution levels from *Avenida da Liberdade* and *Entrecampos* experienced significantly larger declines in ambient pollution between 2009 to 2012. This was estimated via a treatment interaction on zone and time effect, with estimated impacts ranging between a reduction of 22% to 25% for PM_{10} concentration, yet no discernible impacts for NO_2 or NO_X levels.

While many studies provide estimates of pollution impacts from mass public transit and traffic restricting zones, the majority are based on city-level averages or simple mean differences between fixed-point stations. These methods are thus unable to estimate average treatment effects as they may vary across locations in an urban area, and further any differences at the neighbourhood level over time. This lack of heterogeneity can be addressed by making use of more geostatistical interpolation and spatial-temporal modelling.

Anselin and Gallo (2006) incorporate spatial heterogeneity and autocorrelation into a hedonic model of housing prices as influenced by local levels of pollution. The authors highlight the importance of robust interpolation and diagnostics of pollution data prior to any modelling and, in the context of Southern California, conclude that Kriging interpolation techniques provides the best fit for eventual spatial econometric modelling. Significant bias can be introduced into econometric specifications by not using the most appropriate spatial interpolation to generate data (Anselin and Lozano-Garcia 2008).

In an urban setting, limited data availability and detail make comparable transit intervention analyses difficult. It is in this context that this work estimates the within-city spatial and temporal decaying patterns of urban air pollution. Often, the spatial detail comes at

the expense of the temporal detail but using high-dimensional data and differencing estimation this work contributes to better understanding the high-resolution and high-frequency dynamics and trajectories of pollution levels following transit interventions.

The analysis merges robust geostatistical frameworks and diagnostics on interpolated air pollution with spatially detailed urban transportation policy. This allows for a focus on more than just city-wide impacts of various policies and evaluates any potential neighbourhood level disparities and environmental inequalities which may occur following changes to public transit or traffic limiting features.

Further, the discussion is built entirely on open sourced georeferenced data. With current computational abilities it is feasible to do spatially and temporally detailed, robust urban policy analyses. As municipalities are at the forefront of climate change and pollution mitigation, detailed evaluations of transportation interventions can be used to implement best practices to maximize positive spill-over benefits.

3. Pollution Monitoring and Transport Infrastructure

As the capital city and economic hub of the country, the city of Lisbon is densely populated and busy with people, traffic and commerce. In 2017 there were 384,535 firms in the capital, representing around 30% of all those in the country. These businesses attract many workers and visitors with daily increases to the city population of around 70% coming from those commuting into the city centre. The economy is service oriented and heavily driven by accommodation and transport, retail and trade, technology and communication (*Câmara Muncipal de Lisboa* 2018a).

Local economic activity and further, environment patterns, are driven by the city's important location along the Tagus river, with many ports facilitating the trade and transports of goods and people. While the city had a history of manufacturing, agricultural and industrial practices, particularly along the riverfront, in modern times the bulk of these firms have all but

moved out of the urban area and have been replaced by increasingly service and technology-oriented industries. The clustered density around economic hubs of the city, in terms of commerce, population, buildings and traffic, mean that key areas of the city can at times become highly congested leading to high levels of airborne pollutants.

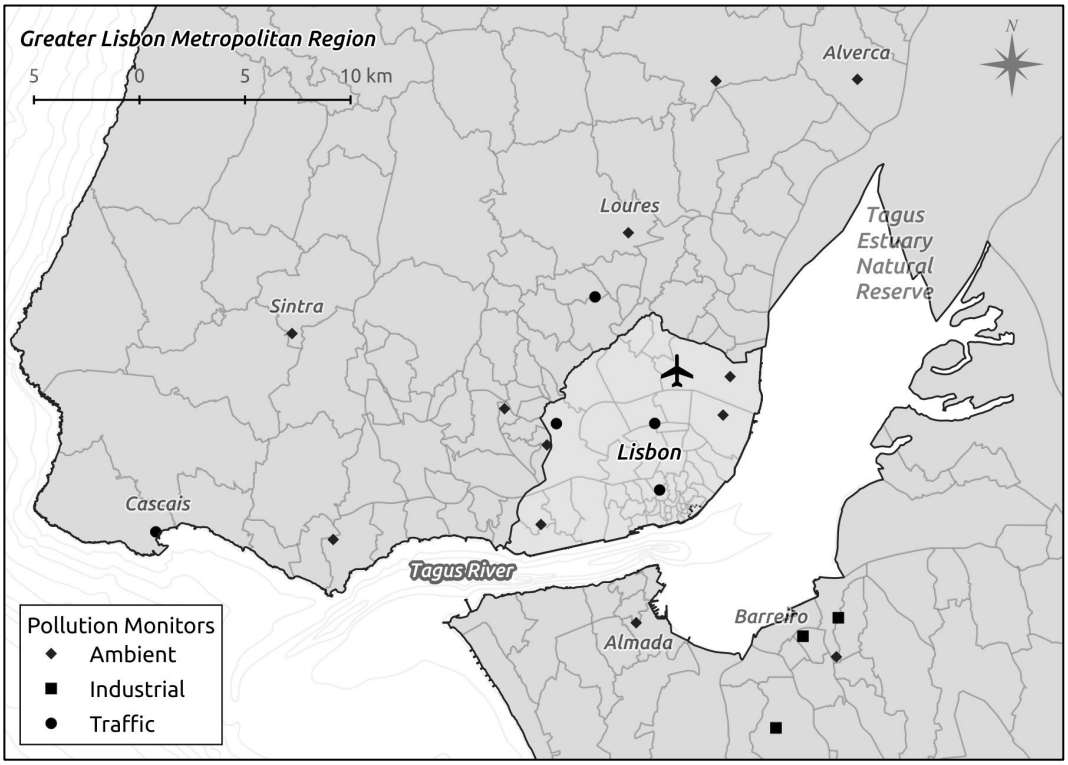
3.1. Local Pollutants and Trends

The intermittent monitoring of various pollutants across the region began in 1995, tracking high-frequency concentrations of common air and ground level pollution. The collection of pollution data in Portugal is managed by *QualAr*, maintained within the *Agência Portuguesa do Ambiente*. There are 68 monitoring stations across the country of which 20 are located in the immediate vicinity of the greater Lisbon region, as seen in figure 1. This grouping of stations represents a density of approximately one station for every 40 km^2 in the greater metro region and one for every 15 km^2 considering only stations located within the municipal boundaries.

There is variation in the timing of when stations began tracking different pollutants, but full coverage daily measures are in general available for six pollutants over 15 years. These include particulate matter (PM_{10}) since 2002, nitrogen emissions (NO, NO_2, NO_X) since 2000, 1999 and 2004 respectively, carbon monoxide (CO) since 2000 and sulfur dioxide (SO_2) since 2001.[1]

[1] Measures of different pollutants may be available for earlier years (starting in 1995) at some stations. To ensure a sufficient base upon which to interpolate, the empirical analysis interpolates only when there is a minimum of six active monitoring stations.

Figure 1. Air Pollution Monitoring Stations in Greater Lisbon, Portugal



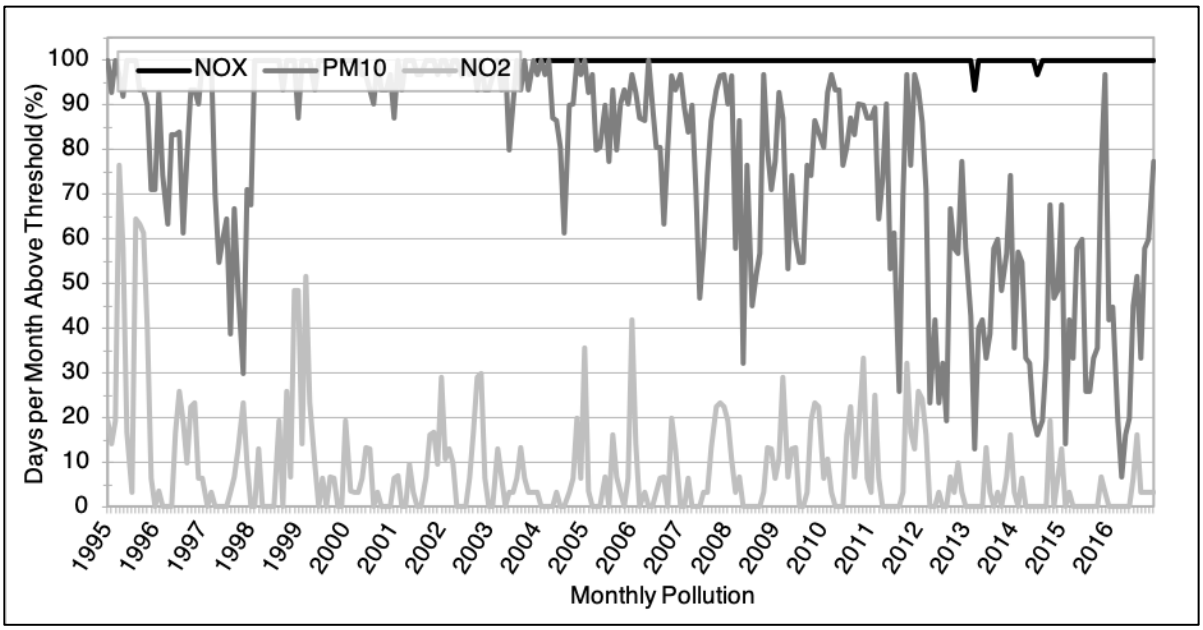
The European Commission has developed a set of air quality standards and regulations which are based on health related research on pollution impacts and are legally binding for member countries (Directive 96/62/EC and subsequent daughter directives). In the case of failing to meet targets, local authorities are then responsible for developing and implementing air quality management plans. Since monitoring began, readings indicated that the Lisbon area consistently exceeded threshold limits for particulate matter and combustion emissions, with high concentrations of PM_{10} across the city and NO_2 particularly to the north. This has led to poor rankings in pollution planning and outcomes relative to other large European cities, however, has also led to a significant investment and focus into local municipal environmental concerns.[2]

Figure 2 shows the proportion of days per month that various pollutants have exceeded their respective regulated maximum threshold limit. We see both particulate matter (PM_{10}) and

[2] European City Ranking 2015: Best practices for clean air in urban transport http://www.sootfreecities.eu/sootfreecities.eu/public/

combustion-based pollution (especially NO_X) consistently fail to meet regulated limits, however significant improvements in particulate matter are seen over time. Decreases in PM_{10} correspond to the development of the 2006 Air Quality Action Plan for Lisbon and Tagus Valley which had a goal of ensuring the compliance of legal limits of air pollution set out by regulatory authorities.

Figure 2. Proportion of Days per Month Exceeding Pollution Threshold Limits



Pollution from PM_{10} is a concern for the region given its serious health implications, especially with regards to respiratory health. While there are no safe levels, a daily maximum threshold of 50 µg/m3 is deemed to represent the limit of harmful concentration not to be exceeded 35 days out of the year. High concentrations of PM_{10} is a common problem, particularly in the main transport corridor leading to the primary business and historic district, *Avenida da Liberdade*.

Particulate pollution however can be highly influenced by regional and larger scale continental trends. Estimated decompositions of local pollution in 2009 indicate that around half of the PM_{10} concentration that year was attributed to external forces. Specifically, in the Lisbon context, particulate levels can be driven by large air masses coming from North African

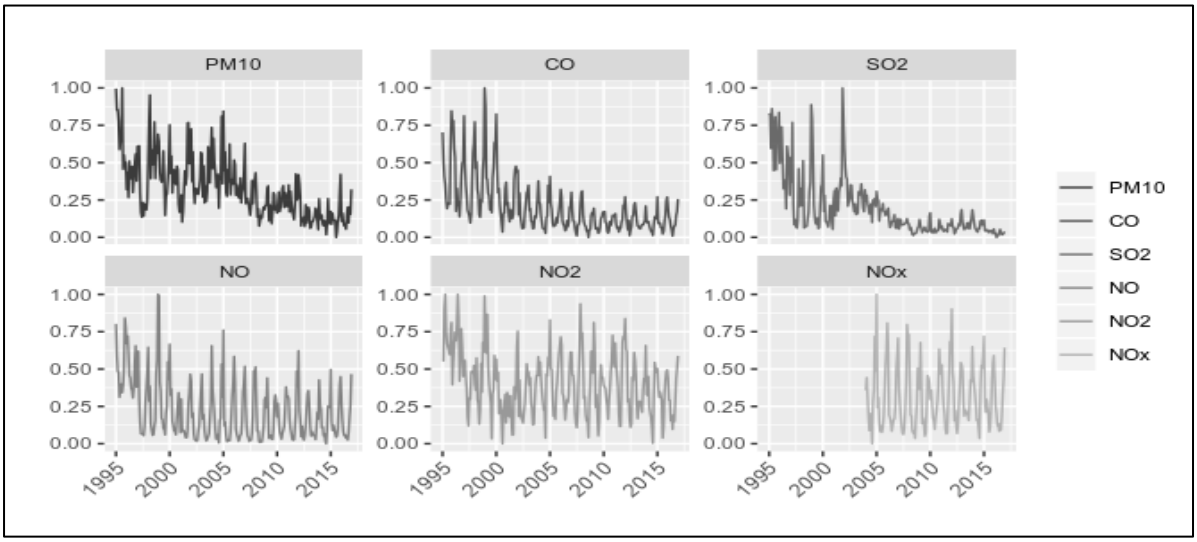
deserts. In 2015, when the region experienced 48 days of atmospheric intrusion by African air masses, there was a significant spike in the monitored values of PM_{10} (*Câmara Muncipal de Lisboa* 2018b).

The other family of pollutants common across Lisbon are nitrogen based and primarily attributed to the combustion of fossil fuels and transportation. These include NO and NO_2 which can be more generally measured as NO_X. While these pollutants could be driven by natural combustion forces, at ground level their concentration is attributed to man-made processes and transportation. There are thresholds set for the level of NO_2 mandating that daily maximum levels not to exceed 200 µg/m3 more than 18 times per year.

As the largest city in the country through which much trade and transport occurs, many heavy vehicles and marine transport pass through Lisbon. These types of transportation often use higher amounts of diesel fuel. Additionally, industrial processes related to manufacturing and trade in the city can be large contributors to SO_2 which is commonly associated with acid rain. Both heavy and light vehicle transportation further contribute to levels of CO emissions in the region.

Figure 3 shows the normalized monthly trends of all pollutants in Lisbon using 2000 as a base year for indexing. In 2016 the average concentration of PM_{10} was 24.96 compared to a low (high) of 23.15 (60.99) in 2014 (1995); NO was 21.76 compared to a low (high) of 18.41 (57.89) in 2013 (1996); NO_2 was 33.48, its lowest value compared to a high of 60.90 (1996); NO_X was 66.25, its lowest compared to a high of 85.08 in 2007; SO_2 was 1.23, its lowest value compared to a high of 14.46 in 1995; and CO was 0.27 compared to a low (high) of 0.25 (0.89) in 2014 (1995).

Figure 3. Standardized Monthly Air Pollution Levels in Lisbon



Partially in response to high levels for some pollutants local environmental quality has become a priority for the municipality. In recent years a significant focus has been dedicated to municipal environmental improvements across many fronts. The city has undertaken ambitious projects in developing urban greenery in the form of tree planting and the provision and maintenance of open spaces, among them the Tagus riverfront running along the South-East edge of the city. Large planned green corridors further aim to link the entire city in an eco-friendly way.

Traffic measures directed specifically at pollution include the introduction of the LEZ around 2012, road restrictions and other residential traffic limitations, the promotion of eco-driving, cycling and public transport. The city has outlined a comprehensive air quality management plan in recent years and based on improvements and planned strategies for the future, Lisbon and the municipal authority, *Câmara Muncipal de Lisboa*, was awarded the European Green Capital 2020.

3.2. Municipal Transportation Infrastructure and Interventions

Two different transportation initiatives are studied in this work, the opening of new metro stations and the introduction of Lisbon's LEZ in the city centre. While the first looks at

the marginal expansion of a public transport network, the second is a traffic limiting policy targeting private ridership and aims to decrease high polluting vehicles in the city centre. Both have the potential to influence the behaviour of drivers in the city by switching to alternative modes of transport or changing their commuting patterns and further, both have the implicit aim of improving transit flows and local pollution levels. Although general comparisons between the resulting pollution reductions can be made, it is important to note how different both types of transit interventions are in terms of their costs, purpose, planning, administration and function.

The Lisbon metro was inaugurated in 1959 with eleven stations running North to South from the historic central business district. Since its inception, stations have been adorned with local art and designed with culturally significant tiles and statues. Different stations, with widely different themes and aesthetics, are known around the world for their uniqueness and attractiveness, and especially for showcasing renowned Portuguese artists and craftspeople.

Construction on the metro continued with nine additional stations opening between 1963 and 1972 after which no new stations were built for almost two decades. Following the political revolution in 1974, the transit system was nationalized in 1975 and has since been run as a public institution. After the nationalization, the metro experienced a revival towards the late 1980's with the construction of new stations and significant expansions to the existing network.

Since 2000, 14 stations have been added to the network, most recently focusing on connecting the international airport and urban peripheries. Currently, the metro consists of 56 stations and 44.5 km of track divided among four lines, with the construction of two additional stations underway for inauguration in the early 2020's. The system services an average of 600,000 riders per day and in 2018 total ridership was 169 million (Grupo ML 2019).

As with most underground systems Lisbon's metro is electric and the carriages themselves do not release any direct combustion emission. Further, as this system is removed from the aboveground road network, congestion is less of an issue. This is compared to other forms of mass public transit like busses which release exhaust as they travel the city, contributing to local pollution levels. Pollution reductions from alterations to the metro network are thus less related to direct changes to public transit exhaust and more related to spill-over and structural changes caused by increasing accessibility and decreasing aboveground private ridership.

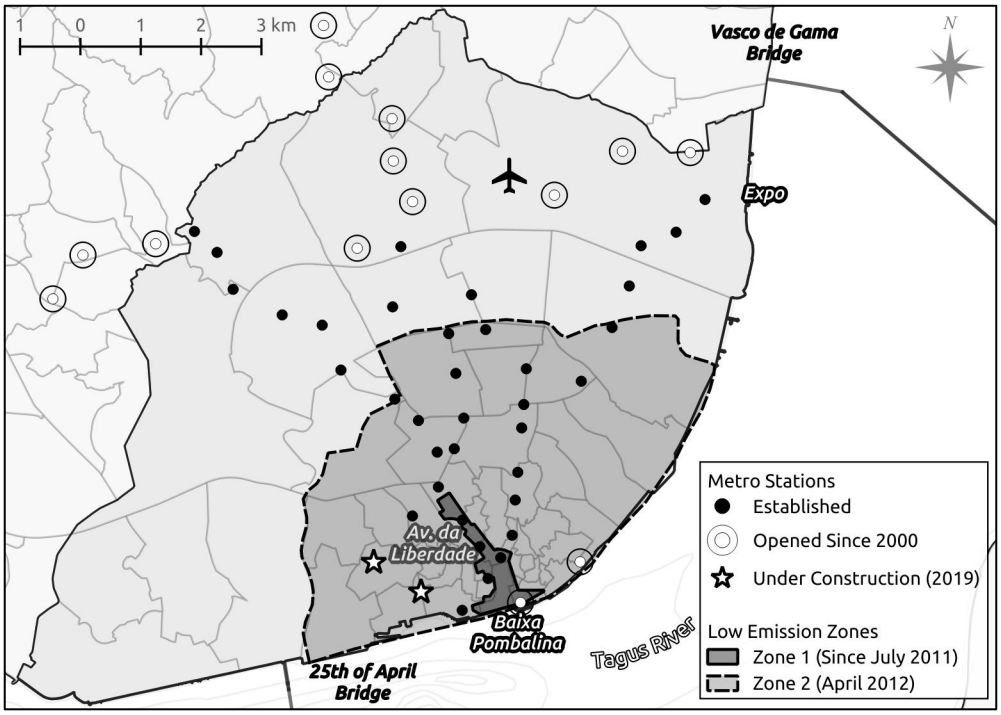
Commuters in the city predominantly use private vehicles for transportation, around 47.7% of residents compared to 19.4% who use busses and 11.63% who use the metro (INE 2011). This contributes significantly to pollution as motor vehicles are primary sources of CO, NO_X, SO_2 and PM_{10}, among other. Vehicle sales in Lisbon are not only much higher, around 71 new vehicle purchases per 1,000 inhabitants compared to the national of 25 per 1,000 in 2017, but further, since 2012, new sales in the capital have risen substantially faster than the national average (INE 2018).

In response to deteriorating pollution levels in the downtown city core, the municipal authority implemented a LEZ in the area with the aim of restricting the worst heavy-polluting transportation fleets. This was part of the larger Air Quality Action Plan introduced in 2006. The current LEZ boundaries cover approximately 30% of the city and are shown in figure 4.

The zone was implemented in three phases starting in 2011. From July of that year to March 2012, during phase one, vehicles from before 1992 were unable to enter an area concentrated around *Avenida da Liberdade*, the primary artery of the city (Zone 1 with a total area of approximately 0.7 km^2). From April 2012, and until January 2015, regulations were strengthened. The original boundary now restricted vehicles manufactured prior to 1996, and an extended area around the city core was introduced, restricting vehicles from before 1992

(Zone 2 with a total area of approximately 25 km^2). Finally, the last alteration came in January 2015, only allowing vehicles from after 2000 in Zone 1 and 1996 in Zone 2.

Figure 4. Lisbon Transit Interventions: Low Emission Zones and Metro Stations



In terms of enforcement, the LEZ restrictions are upheld by the local traffic. The enforcement fine during the first phase amounted to between €25 and €125 for non-compliance. During the first phase, around 20 fines per month were recorded. Given the relatively *ad hoc* manual enforcement policies, plans were made during the last phase of the LEZ implementation to introduce a network of traffic cameras within Zone 2 with license plate reading capabilities in order to increase enforcement and compliance of the regulations (Gonçalves 2014).

Of particular importance for this study is the fact that Lisbon's LEZ was implemented in conjunction with other policy measures. As the LEZ was introduced, so to were other traffic changes, for example altering road axes to remove parking spaces or restructuring the main roundabout (*Marques de Pombal*), to manage traffic volumes and congestion (occurring in September 2016). This implies that isolating the impact directly from the LEZ policy is difficult

since the resulting effect could come from the variety of transport policies introduced around the same time. For this reason, any estimated effects are likely to be representative of the broader collection of transit policies introduced in and around the LEZ and surrounding areas.

4. Spatial Interpolation and Aggregation of Sequential Pollution Monitoring

Any analysis and discussion surrounding potential spill over environmental costs or benefits of an urban transport intervention must necessarily involve accurate, representative and adequately detailed data in terms of the spatial and temporal resolution, often one at the cost of another. While the temporal resolution of pollution data in Portugal is very detailed, in many instances down to the hour, geostatistical interpolation and aggregation is used to expand the spatial dimension for each time interval, allowing for detailed within-urban policy evaluation.

A relatively large body of work has been developed along the lines of statistical interpolation and modelling methods for environmental pollutants over space and time, predicting unknown pollution levels in areas and time where no data is observed. Sophisticated spatial, temporal and, or spatial-temporal interpolations have all been employed, in addition to more deterministic models conditional on external influences such as elevation, the built structure of the city or weather patterns, for example in a geographically weighted regression.

The primary goal of this work however is not the creation of multi-dimensional deterministic pollution models, but rather to employ a battery of diagnostics and selection criteria to those interpolation methods most commonly used in applied urban-scale environmental work. This gives rise to a set of generalizable criteria and algorithms for selecting the best method for the generation of a high-frequency and high-resolution longitudinal urban pollution database. With statistically robust and comprehensive pollution data, it is then possible to evaluate the effects of transit interventions on neighbourhood pollution with sufficient spatial and temporal detail.

4.1. Interpolation Methods and Techniques

In its raw form the air pollution data represents the concentration of a pollutant (PM_{10}, NO, NO_2, NO_X, CO, SO_2) observed at 20 monitors from inside the study region, as well as those in the immediate proximity in all directions (figure 1). The chosen base temporal frequency to which hourly station level observations are first aggregated is at monthly intervals, balancing the granularity of capturing short-term spatially detailed effects while computationally feasible.

Spatial interpolations are conducted at every month rather than using spatial-temporal interpolations or deterministic modelling across multiple dimensions. This is in part due to the application of interest, using temporal cross sections to estimates changes before and after transit interventions. Increased modelling variables or sophisticated temporal dynamics may smooth the data series too much to observe effects, and the final longitudinal database implicitly accounts for the sequential and cross-sectional nature of pollution.

Each sequence of monthly interpolations is estimated varying the underlying parameters, specifications and variables, comparing diagnostics and practical concerns. The methods are compared based on their predictive power and the continuity and consistency of estimates over space and time. Subsequent aggregation to observational units for the econometric specification further yields diagnostic information regarding the interpolation fit for applied modelling.

Two commonly employed families of spatial interpolation methodologies are used to go from the fixed-point static measures of pollution to a continuous distribution over space for each month. Both the Inverse Distance Weight (IDW) and Kriging models are geo-spatial interpolations able to fill in spatially missing values using only the observed concentration of pollution and relative distance between observations. Each family differs however in the underlying assumptions, statistical techniques and customization by the user.

The varying forms of all models are compared according to a range statistics and diagnostics evaluating the predictive power and fit of the interpolated estimates. The main diagnostic to compare models is the root mean square error (RMSE) of prediction obtained via a cross validation approach. More practical concerns focus on the distribution of the interpolated values relative to the observational unit size and further representing adequate heterogeneity across space and time.

Selection of Grid for Interpolation and Neighbourhood for Aggregation

It is important *a priori* to determine the base spatial resolutions onto which the interpolated and aggregated data will be projected. For interpolated data this includes the pixels and their sizes representing the continuity of ground truth. The choice of aggregating unit however is conditional on the research question and the model of interest. In this case, we are interested in evaluating how average neighbourhood level pollution levels have changed, and so the observational units should represent spatial neighbourhoods across the city.

A discrete grid of high-resolution pixels is used as a canvas when interpolating from the monitoring station locations. The choice of grid size must ensure an adequately continuous ground distribution of values which accurately represents marginal changes in pollution moving in any direction. The most important of the criteria is that the size of the interpolation grid is smaller than the observational neighbourhood units of interest, yet large enough for feasible computation. Pixels of 100 meters by 100 meters are chosen, and spatially detailed enough within the study region to capture continuous heterogeneity of pollution.[3]

For this study, 3,623 census enumeration tracts are used to represent the neighbourhood unit of analysis to which interpolated data is then aggregated and modelled. Census tracts in Lisbon are typically delineated along roadways and other natural barriers of the city and thus

[3] Lisbon has a total size of 100 km^2, and a pixel length of 100 m would represent approximately 1% of the city's horizontal or vertical distance.

implicitly separate areas according to the built environment and natural boundaries. These are broadly aligned with the idea of a small-scale city-block neighbourhood capturing the spatial heterogeneity of pollution across the area.

The distribution, size and boundaries of neighbourhoods in an area may be endogenously determined. Any selection of neighbourhood extent however is based on some *a priori* assumptions conditioning the neighbourhood definition on the size of local population or building density, geodemographics, topography, geographic area, or based on the historic and cultural evolution of the city.

Table 1. Size and Density of Lisbon Neighbourhoods

	Min.	Median	Mean	Max.	St. Dev.
Size (km2)	0.00	0.01	0.02	3.70	0.09
Elevation (m)	0.00	70.38	64.51	201.60	31.95
Building Density (per km2)	0.00	1,199	2,052	20,976	2,478
Population Density (per km2)	0.00	12,644	15,214	99,456	13,148
Dist. to Nearest Metro Station (km)	0.02	0.70	1.30	6.70	2.43
Dist. to LEZ Boundary (km)	-0.18	3.08	3.25	7.97	2.22
Dist. to Baixa (km)	0.04	4.15	4.25	9.44	2.43
Dist. to Tagus (km)	0.00	1.75	2.46	8.08	2.06
Dist. to Nearest Freeway (km)	0.00	0.39	0.49	2.78	0.42

Census Tract Level Data (N = 3,623 neighbourhoods)

If neighbourhood units are not structurally dissimilar in terms of their relative size and density over space and time, then model estimates can capture average effects by controlling for these spatial and temporal differences. Table 1 gives some underlying statistics regarding the neighbourhood units chosen for the aggregation of pollution in terms of neighbourhood size, density and locational features.

Inverse Distance Weight Interpolation

The IDW family of interpolations predict pollution values at locations for which no observed measure exists based on weighing observed values from a location proportional to the distance between the locations and some weighting parameter. The general formula for obtaining the interpolated values from the IDW is as follows, where the predicted interpolated value $\mathbb{P}$ at a location x (pixel) for which no data exists is determined according to:

$$\mathbb{P}(x) = \begin{cases} \frac{\sum_{i=1}^{N}[1/d(x,x_i)^{\rho}] \cdot \mathbb{P}(x_i)}{\sum_{i=1}^{N}[1/d(x,x_i)^{\rho}]} & \text{if } \ d(x,x_i) \neq 0 \\ \mathbb{P}(x_i) & \text{if } \ d(x,x_i) = 0 \end{cases} \qquad \text{Eq. 1}$$

If the distance between the pixel location and any of the monitoring stations i is equal to zero ($d(x, x_i) = 0$), then the pixel is exactly at a monitoring location and is assigned that value, $\mathbb{P}(x_i)$. Otherwise, if there is some positive distance between the observed value and the pixel, then the interpolated predicted value is conditional on two parameters: the choice of the number of nearest neighbours, N, and the inverse distance power, ρ. The value of pollution from neighbours (monitoring stations) is weighted according to the respective distance between the observed location and the interpolated pixel of interest.

Given the infinite possible combinations of N and ρ parameters, hyperparameter fine-tuning is done to select among different combinations of realistic parameter sets representing the most commonly used and extreme limit cases. Selecting the best model by varying parameters among a select few within a pre-determined group greatly improves the speed and efficiency in determining the best model. This fine-tuning is done by running all combinations of the interpolation (and respective aggregation) of values and comparing the predictive power among all choices of parameters. How each of the parameter combinations effect the different variants of pollutants over time can help in better understanding the dynamics of pollution in terms of its relative decay and continuity over space from the observed monitored locations.

Four choices for each parameter represent the *de facto* commonly used values and limit cases. The possible values of each parameter are given in table 2. The 16 combinations of all parameters capture some extreme cases, for example, using all monitoring stations N and no decay, $\rho = 0$, attributes to every pixel the average pollution level from across Lisbon. While a combination of parameters like this is acceptable from a geostatistical point of view, this would not provide any spatial variability and would not be of any use for evaluating high-resolution and high-frequency urban policy initiatives.

Table 2. IDW Nearest Neighbour and Weight Parameter Combination

$N = 1$ The singular nearest monitoring station.	$\rho = 0$ A weight of zero indicating no decay in the value moving away from the observed location.
$N = 3$ The three nearest monitoring stations aimed at capturing a broad triangulation around the pixel.	$\rho = 1$ A linear decay proportional to the distance between the observed value and the prediction location.
$N = All$ All monitoring stations in the study region.	$\rho = 2$ Squared inverse distance weight assigning higher weight to closer observed values relative to the linear.
$N = CVOpt.$ A cross-validated determined optimal number of neighbours.	$\rho = CVOpt.$ A cross-validated determined optimal weight parameter.

Given the relatively small sample size of monitoring stations with observed values, the cross-validated (CV) optimal versions of each respective parameter are obtained using a leave one out cross-validation (LOOCV) approach. This method iteratively uses all the data, removing one observed value at a time and running the IDW model sequentially through each removal of an observation. For each sequence, an optimization function determines the values of N and ρ which minimize the RMSE.[4]

This generates a set of optimized values and RMSE criteria for each iteration (removal) of a monitoring station. For each pollution-month interpolation there is a vector of LOOCV RMSE estimates with length equal to the number of monitoring stations. The selected CV optimized set of parameters are the median within-value of all the N and ρ estimates which minimize their respective iteration of the RMSE. The median is chosen so that selected parameters are not sensitive or driven by outliers or any spurious training sample chosen. When comparing results, most frequently the minimum RMSE is attributed to the median parameter values.

Kriging Interpolation

The Kriging family of models are similar to the IDW in that they both interpolate values at areas for which no observations are observed conditional on the locations from which there

[4] The optimization searches for the combination of N and ρ parameters which will minimize the RMSE function according to the Nelder–Mead methodology with initial parameter values of $N = 3$ and $\rho = 2$.

are measurements. Differences arise however in the underlying assumption regarding how to weight the contribution of measurements from different locations. While the IDW bases this weight on the relative distance between a location and its neighbours for which there are observations, the Kriging prediction bases this weight on a Gaussian process. The general format of the Kriging process in in line with the IDW process and can be represented according to equation 4.2.

$$\mathbb{P}(x) = \sum_{i=1}^{N} \omega_i \, \mathbb{P}(x_i) \qquad \text{Eq. 4 – 2}$$

where the inverse distance weights are replaced by a series of optimized weights, ω, minimizing the square deviation between predicted and observed values, much like a regression specification.

The underlying assumptions related to the structure of the expectation of the observed monitoring station values, $\mathbb{P}(x_i)$, will influence the choice of model specification and complexity. The Ordinary Kriging specification is used if the underlying pollution variable is assumed to come from a random data generating process where the mean is a constant unknown value with random disturbances, $\mathbb{P}(x_i) = \mu + \varepsilon_i$. This would imply that the underlying pollution data follows a spatially stationary process.

If the underlying process varies deterministically, then a trend component can be included. This assumption forms the basis of Universal Kriging specifications, where the expected value of pollution varies deterministically according to other processes, $\mathbb{P}(x_i) = \mu_i + \varepsilon_i$, where the expectation, μ_i, can be expressed in terms of covariates $\mu_i = \sum_k \beta_k z(x_i)$, for k potential predictors in vector z, conditional on location x_i.

The baseline Universal Kriging specifications here is the inclusion of latitude and longitude of the prediction locations as determinants of the predicted value. This would be an appropriate model if we expect some underlying trend where the average pollution varies

deterministically across space – potentially due to some external, unobserved spatial factors such as the variability in wind speed and direction. So, while the set of Ordinary Kriging specifications use an underlying Gaussian process to estimate predicted values, the Universal specifications include a deterministic component. If we expect location to be significant, then the Universal Kriging models should outperform the Ordinary counterparts.

While high-dimensional Kriging specifications can be developed, the goal of this work is not the sophisticated modelling of interpolated air pollution. Additional auxiliary variables can be included to enhance the Universal Kriging model; however, their inclusion is outside the scope of this work. This would require that additional external variables are available at a complete spatial coverage such that each pixel of interest onto which we want to interpolate has underlying data from which to build a model.

There is however one exception which does not require any additional data beyond what is openly available. Given that a sequence of temporal interpolations is conducted at every month, important auxiliary influences can be included in the model in the form of the lagged values of pollution without the burden of having to obtain data on external factors. This could potentially result in significant prediction accuracy increases with minimal additional data processing.

At each time period both baseline Ordinary Kriging and Universal Kriging specifications are enhanced by the inclusion of up to two periods of lags of predicted values. If there is some residual temporal dynamics involved in the prediction of pollution, as would be expected given the continuous nature of these variables over time and space, then the inclusions of past prediction values could improve upon the baseline versions.

Thus, six specifications of Kriging are estimated. The standard Ordinary Kriging specification, which considers that pollution has a random unobserved mean, the Universal Kriging model which considers that pollution is deterministically influenced by the inherent

location across space, and further two additions to each of these base specifications which look at the inclusions of one and two temporal lags.

4.2. Spatial Interpolation Diagnostics and Choice

Spatial interpolation and aggregation are statistical concepts and so the choice of model, sequences and methods to follow can vary widely in different contexts and with different data. There are however a series of diagnostic tools and practical ideas which can be implemented to ensure that the choice of interpolation produces the best and most robust statistical series for each pollutant. This further ensures that any discussions or inferences regarding transit interventions are not driven by the choice of model.

This section outlines the series of criteria and diagnostic inferences guiding the choice of interpolation for generating high-frequency spatial and temporal pollution measures. There is a separate series of interpolations conducted for every pollutant at each month interval since the beginning of their series. This includes 1,307 pollution-month combinations, omitting those with missing values or fewer than six active monitoring stations. Diagnostics are compared between all combinations of methods described, 16 variants of IDW and six variants of Kriging, proving a wealth of information from which to draw conclusions.

One additional point of note is the choice of variable on which to interpolate, namely whether to interpolate directly on the concentration of pollution or whether there are efficiency gains by interpolating directly on the log value of the pollution.[5] With certain interpolation methods, namely the Kriging specifications, relying on an assumption of Gaussian residuals, the pre-transformation of pollution concentrations to log form may provide a better fit and distribution of the observed values to interpolate.

[5] Because the econometric variable of interest is the log of pollution (to estimate percentage change) there is no need to do any post-interpolation transformation and therefore is less alteration to the data. If the final variable of interest is not in log form, then additional cross-validation should be conducted on the post-transformation to check deviations from the original non-transformed ground truth.

Decision Criteria

Several different diagnostics are used to evaluate the overall model performance based on type, parameter selection and variable of interpolation. Given that a temporal sequence of spatially interpolated and aggregated values are generated, no one singular diagnostic value can adequately evaluate the overall model performance across the entire spectrum of interpolations. Diagnostics generated for every month, pollution and model combination will be evaluated using linear regressions to determine how different parameters or model selection influences predictive power.

The criteria for selecting the appropriate interpolation is based on several conditions, however the driving decision should be selecting models with the highest accuracy. The first step is disregarding any observations with inadequate representation, here by removing interpolations with less than six active monitoring stations or missing values.

The predictive accuracy is evaluated using the RMSE.[6] Given the relatively small number of monitoring stations, the overall RMSE for each model is estimated using a LOOCV approach, systematically removing one monitoring station at a time and evaluating the RMSE between the predicted value and the observed value left out. The average of these RMSE across all the sequentially left-out monitoring stations provides the index upon which we compare all model specifications. A low RMSE indicates that the model more accurately predicts the ground truths.

The distribution of the interpolated values is also important to consider, specifically whether the distribution of values obtained correspond to the distribution of values following the aggregation to the neighbourhood (census tract) level units. This is important for considering whether the results from the interpolation have a spatial resolution which is

[6] The general formula for the RMSE evaluates the difference between some value of x and it's expected (or predicted) value: $RMSE = \sqrt{[avg((x - \bar{x})^2)]}$

detailed enough to accurately be aggregated without changing the underlying structure of the pollution values.

A mismatch between these distributions would signal an additional skew being added to the data in moving from the interpolated values to the aggregated values, potentially biasing any results. The is evaluated using the Kolmogorov-Smirnov test statistic to determine if the interpolated and the aggregated series are drawn from the same continuous empirical distribution.

The distributional concern should further be complemented by looking at the continuity and variation of interpolated values. As is particularly the case with IDW models with zero weights, the interpolated value could lack spatial variation if the predictions represent simple unweighted averages. This is evaluated by considering the number of unique values obtained from interpolation and ensuring this is at least as large as the number of observational units so as not to introduce further alterations to the data.

Finally, it is important to have temporal consistency in the choice of specification. Since different pollutants are entirely different series the choice of model can vary depending on pollution but within, there should be a consistent method for generating the data across the time span. Selecting different specifications at different intervals will change the underlying assumptions in the data series. Given the interest is in estimating temporal breaks, it is necessary to reduce any time inconsistencies from different methodologies.

Diagnostics and Final Selection of Interpolated Data

To guide the choice of interpolation model, a set of linear regressions are performed to see how different parameters, specifications, and variables influence the overall prediction accuracy and fit of the interpolated values. For every pollutant-month combination, this

diagnostic information is available for all interpolations and further for comparing each specification using interpolation on the direct concentration and interpolation on log values.[7]

A regression model estimates the impact that model specification, number of active stations, time and distribution of value variability (as measured by the number of unique interpolated values) have on the RMSE of prediction. A pooled model first controls for these characteristics, and further for each pollutant and respective pollutant-specification interaction. Subsequent pollution-specific models are estimated to determine the best specification for each data series. Results on these diagnostics are presented in table 3.[8]

The full specification with control variables significantly determines the variability of the prediction power of models with an R^2 of 82.5%. As the number of monitoring stations increase, the variation in the RMSE decreases (t-value of -28.38), and this is also seen as we consider observations occurring in later years (decrease in t-value from -10.42 in to -39.43 in 2013-2016). While robust and accurate modelling can significantly enhance the prediction, these results show that inherently the power of any model is conditioned by the external availability of data.[9]

[7] This gives $1{,}307 \cdot (16 + 6) \cdot 2 = 57{,}508$ sets of parameter combinations and resulting accuracy measures.
[8] t-values estimate how many standard deviation reductions (or increase) each specification has on the RMSE from the study-wide average value. Direct effects shown only and full values for all diagnostics available upon request.
[9] While not shown here for brevity, similar results are found when considering as a dependent variable the Kolmogorov-Smirnov test statistic. The more monitoring stations and later collection time, for example, relatively reduces the KS statistic, indicating no significant difference between the interpolated and respectively aggregated distribution of values. This captures the idea that the interpolated values must be well suited for the observational units, and a lower KS statistic indicates a closer correspondence between the interpolated and aggregated values.

Table 3. Interpolation Model Diagnostics: Full and Sub Linear Regression t-Values

	Full (Model 1)		PM_{10} (Model 2)		NO (Model 3)		NO_2 (Model 4)		NO_X (Model 5)		CO (Model 6)		SO_2 (Model 7)	
Dep. Variable: RMSE of Prediction	*Level's*	*Log's*	*Level's*	*Log's*	*Level's*	*Log's*	*Level's*	*Log's*	*Level's*	*Log's*	*Level's*	*Log's*	*Level's*	*Log's*
Nearest Neighbour: 1; Weight: 0	0.00	-90.99	17.12	-40.90	7.89	-23.89	8.16	-39.36	10.00	-35.72	1.28	17.09	1.32	-12.62
Nearest Neighbour: 1; Weight: 1	0.00	-90.99	17.12	-40.90	7.89	-23.89	8.16	-39.36	10.00	-35.72	1.28	17.09	1.32	-12.62
Nearest Neighbour: 1; Weight: 2	0.00	-90.99	17.12	-40.90	7.89	-23.89	8.16	-39.36	10.00	-35.72	1.28	17.09	1.32	-12.62
Nearest Neighbour: 1; Weight: CV	0.00	-90.99	17.12	-40.90	7.89	-23.89	8.16	-39.36	10.00	-35.72	1.28	17.09	1.32	-12.62
Nearest Neighbour: 3; Weight: 0	-13.86	-91.24	7.11	-41.33	4.43	-24.31	1.30	-39.63	3.05	-35.83	0.64	14.91	-2.47	-13.16
Nearest Neighbour: 3; Weight: 1	-12.47	-71.08	-0.95	-16.88	0.12	-11.86	-1.09	-18.17	3.01	-4.86	1.69	8.17	-0.97	-6.06
Nearest Neighbour: 3; Weight: 2	-12.38	-71.07	-0.78	-16.87	-0.29	-11.85	-1.31	-18.17	3.02	-4.86	1.70	8.24	-0.55	-6.02
Nearest Neighbour: 3; Weight: CV	-12.28	-71.09	-1.26	-16.89	-0.27	-11.87	-1.40	-18.18	3.04	-4.86	1.68	8.00	-1.12	-6.11
Nearest Neighbour: All; Weight: 0	*Ref.*	-91.30	*Ref.*	-41.54	*Ref.*	-24.48	*Ref.*	-39.61	*Ref.*	-35.88	*Ref.*	12.49	*Ref.*	-13.09
Nearest Neighbour: All; Weight: 1	-15.77	-71.14	-2.83	-16.95	-1.01	-11.92	-1.36	-18.18	2.57	-4.87	1.50	7.30	-0.90	-6.09
Nearest Neighbour: All; Weight: 2	-15.73	-71.13	-2.09	-16.92	-1.13	-11.89	-1.82	-18.19	2.57	-4.87	1.57	7.56	-0.64	-6.07
Nearest Neighbour: All; Weight: CV	-15.41	-71.13	-3.16	-16.96	-1.40	-11.93	-1.36	-18.17	2.62	-4.87	1.54	7.26	-0.61	-6.13
Nearest Neighbour: CV; Weight: 0	-14.80	-91.25	6.15	-41.35	2.94	-24.35	0.38	-39.65	2.58	-35.83	0.51	14.60	-2.52	-13.23
Nearest Neighbour: CV; Weight: 1	-13.07	-71.09	-1.17	-16.89	-0.29	-12.00	-1.39	-18.38	2.93	-4.86	1.66	8.21	-1.00	-6.13
Nearest Neighbour: CV; Weight: 2	-12.82	-71.08	-0.92	-16.88	-0.50	-11.99	-1.50	-18.37	2.96	-4.86	1.68	8.30	-0.56	-6.08
Nearest Neighbour: CV; Weight: CV	-12.81	-71.10	-1.48	-16.90	-0.37	-12.01	-1.61	-18.38	2.96	-4.86	1.65	8.06	-1.21	-6.19
Ordinary Kriging	-17.63	-71.95	-3.32	-16.91	-1.07	-12.27	-1.79	-18.90	2.33	-5.08	1.51	7.52	-0.72	-7.23
Ordinary Kriging: 1 Temporal Lag	-46.68	-71.80	-8.15	-17.11	-7.03	-12.51	-8.88	-18.90	-1.60	-5.00	0.75	4.79	-3.70	-7.94
Ordinary Kriging: 2 Temporal Lags	-46.09	-71.65	-8.18	-17.08	-6.94	-12.46	-8.84	-18.84	-1.52	-4.97	0.81	4.88	-3.57	-7.85
Universal Kriging	-9.92	-71.00	-1.95	-17.94	0.25	-11.81	0.77	-18.06	3.35	-4.85	1.80	8.30	0.56	-5.74
Universal Kriging: 1 Temporal Lag	-45.27	-71.52	-8.11	-17.78	-6.74	-12.26	-7.75	-18.34	-1.40	-4.93	0.87	4.28	-2.58	-6.77
Universal Kriging: 2 Temporal Lags	-43.77	-71.40	-8.03	-17.69	-6.65	-12.24	-6.50	-18.27	-1.20	-4.92	1.00	4.55	-2.24	-6.62
Observations	56,996		7,864		8,656		9,292		6,808		8,300		8,212	
Pollution Controls	Yes		No		No		No		No		No		No	
Pollution-Specification Interactions	Yes		No		No		No		No		No		No	
Adjusted R^2	0.8241		0.8667		0.6449		0.7949		0.8206		0.4301		0.3636	
Residual Std. Error	3.2602		1.3260		4.5601		2.5763		6.4886		0.2555		1.6824	
F Statistic	853.85***		1,044.70***		321.79***		735.75***		636.44***		128.83***		96.72***	

The choice of model specification for each pollutant is done by estimating pollution-specific sub-models with relevant time and active station controls. The reference interpolation category used is all nearest neighbours with a weight of zero – or in other words the Lisbon level average of the respective pollution at any month in time. The negative t-value attributed to each model specification therefore represents the relative decrease in the average RMSE, resulting in a better prediction power for that give specification over the alternative simple average.

Clear gains are seen when using the log version in terms of reduced RMSE. This is particularly the case when comparing the Kriging specification of models which are conditional on an underlying Gaussian distribution. This suggests that, when possible, conducting pre-transformation of the data should be done to create a more standard distributed variable.

Overall, Kriging specifications are superior for most pollutants and the inclusion of one temporal lag provides the better fit. While the inclusion of a secondary temporal lag in the prediction yields strong results, compared to one temporal lag it appears that these models may be overfitting the data series. In terms of the IDW family of models, it is always best to include all observations in generating the predicted value rather than a localized subset of monitoring stations.

The RMSE of prediction for PM_{10} is well described by the choice of model, with an R^2 value of 86.67%. For the interpolation of PM_{10} values, similar gains in prediction power come from using all the available data in an IDW model and using the Kriging specification. Still, however, the Kriging family of models have a clear gain over the IDW with better fit and significantly larger reductions in the RMSE of prediction. The preferred PM_{10} specification is the Universal Kriging using one temporal lag. This specification is also the best choice for modelling of SO_2, consistently better than any of the IDW models. Suggesting a spatial deterministic component in the distribution of these series.

In terms of combustion-based pollutants, NO and NO_2, the Kriging specifications using log pollution values are superior to the alternatives. In order to choose between potential specifications, the top general two best IDW and Kriging interpolations are estimated, including both the Universal and Ordinary Kriging with one temporal lag and the IDW using an optimally determined weight parameter with three and all nearest neighbours. Across the entire time series, a tabulation of which model has the lowest RMSE at every month out of the potential top candidates shows that for NO and NO_2 the Ordinary Kriging model with a temporal lag has the best prediction 119 times out of 198 and 137 times out of 213 respectively.

No consistent model had any significant improvement in the prediction power of CO pollution relative to using the study-wide average value. The different specifications explain little of the variation in RMSE with an R^2 of 43.01%. This suggests that the interpolation may not accurately be capturing the true spatial dynamic of this series and more complex underlying features may be influencing local CO values.

Even though the average spatial interpolation does not generate robust predictions, this does not mean that the empirical strategy will not be able to capture the temporal dynamics before and after a transit intervention. The Kriging specification of models remain relatively those with the smallest RMSE. This allows us to have some spatial variability compared to the city-wide average yet provides some level of predictive power over the alternatives. The Ordinary Kriging with one temporal lag is the best prediction 119 times out 191 and is chosen as the preferred specification for CO. Care is taken with these results however, and limited inferences are made with discussion focused on general trends and patterns.

Not all pollutants are well predicted by the Kriging models, and results indicate NO_X is best predicted using models with weights of zero. This would seem to suggest that this pollutant is also better predicted by localized averages without any weighting or Gaussian-based interpolation, and thus has less decay over space and inherently less spatial variability.

To balance the need for concentration heterogeneity with choosing the model minimizing the RMSE, the three nearest neighbours are chosen with a weight of zero. This uses the local average of the closest monitoring stations without any spatial decay.

Figure A1 of the appendix shows selected pollution-month interpolated spatial grid and corresponding aggregated neighbourhood observation units for the respectively chosen preferred interpolation model.

5. Empirical Spatial-Temporal Impact Estimates of Urban Transit Initiatives

The mean aggregation of the preferred interpolated data to the neighbourhood level observation units across monthly intervals yields a high-dimensional spatial-temporal longitudinal panel database with the same neighbourhood's pollution concentration repeated at every month. The empirical methodology employed makes use of this structure comparing different neighbourhoods across space before and after the introduction of the transit initiatives.

Under a panel data difference-in-difference econometric strategy the long-run average reduction impacts for various pollutants can be estimated. This is further complemented by a non-parametric bootstrapping procedure to estimate an equivalent short-run difference-in-difference average treatment effect for each impact. This bootstrapping overcomes the limited short-run sampling allowing us to further estimate temporal decays. These estimated spatial-temporal impacts are based on taking the temporal introduction of the transit initiatives and comparing the differences before and after while accounting for the varying spatial orientation of neighbourhoods, namely by considering those that are closer to key areas of the city where we would expect pollution levels to be altered following an intervention .

5.1. Spatial-Temporal Difference-in-Difference Specification

A variety of models are estimated for each of the transit interventions of interest to evaluate the local impact of pollutants, and any spatial decaying effects, from key locations

around the city related to metro stations and the LEZ. Here, the same econometric model is applied to all the pollutants to observe how equivalent interventions impacted the various pollutants differently. The empirical specifications in equations Eq. 4 – 3 and Eq. 4 – 4 control for the heterogeneity of neighbourhoods while estimating the treatment impact of the respective transit intervention by difference-in-difference across space and time.

$$ln(\mathbb{P}_{it}) = \beta + \alpha_i + \beta_1 \text{Location}_i + \beta_2 \text{Metro}_t + \beta_3[\text{Location}_i \times \text{Metro}_t] + \beta_4 \text{Year} + \varepsilon_{it} \quad \text{Eq. 3}$$

$$ln(\mathbb{P}_{it}) = \beta + \alpha_i + \beta_1 \text{Location}_i + \beta_2 \text{LEZ}_t + \beta_3[\text{Location}_i \times \text{LEZ}_t] + \beta_4 \text{Year} + \varepsilon_{it} \quad \text{Eq. 4}$$

Here, $ln(\mathbb{P}_{it})$ represents the log value of the various pollution measures for the $i = 1, \dots, 3{,}623$ census tract neighbourhoods for each monthly interval from $t =$ January 2000 to December 2016. An overall average effect is captured in the intercept, β, while neighbourhood level heterogeneous effects are captured in the fixed effects, the individually varying parameters, α_i. The ε_{it} are classical Gaussian error terms. Year controls are included to capture broad annual trends in concentration trajectories.

The Location_i variable indicates the $\{0, 1\}$ assignment to one of the chosen spatial treatment areas, indicating whether a neighbourhood is in, for example, *Baixa* or near a metro station recently opened. The Metro_t and LEZ_t represent the temporal treatment, assigning pre and post-transit intervention classification.

The specifications above estimate the direct impacts of both the space (location) treatment and the temporal (transit) treatment, as well as the interaction between these two which represents the average treatment effect, β_3, via a difference-in-difference estimation. In this scenario the first difference represents the change in pollution captured before and after a transit initiative, while the second difference captures whether a neighbourhood is proximate or not to five chosen locations across the city where we would expect pollution to be impacted – near the newly opened metro, along the riverfront, in the central business district, near the LEZ or along busy thoroughfares.

If we expect that there are important omitted neighbourhood variables which are not measured or cannot be observed, then a fixed effect estimation can address this omitted heterogeneity. One of the key assumptions in the use of these neighbourhood level fixed effect controls however is that any unobserved and omitted heterogeneous effects among them is time invariant. That is, across the timespan no neighbourhood experienced structural change in their trajectory or dynamics over and above the changes other neighbourhoods experienced.

The fixed effect, or within group, uses the individual group mean to identify the impact of intervention over space and time. This means that the parameter values are estimated using a standard OLS applied to the within-group demeaned values. This estimation is numerically equivalent to including a dummy variable for each census tract neighbourhoods. As the estimation is done on demeaned variables no general intercept value is obtained from the estimation.

Choice of Temporal Treatment

This work looks specifically at two groups of transit related interventions introduced over time. Exploiting the timing of these different initiatives is used to estimate and compare their mitigating impacts on localized airborne pollution across the city. While the mechanisms of these interventions are very different, the environmental aims of reducing air pollution are the same. The goal of this work is not to compare and evaluate these two transit interventions against each other, but rather to explore the patterns observed in the generated longitudinal neighbourhood pollution data against known transportation related changes with spatial and temporal dynamics.

The two temporal treatment variables are introduced in a cumulative way. For estimating the impact from opening metro stations, the variable represents the cumulative number of stations opening over the course of the study period. This controls directly for the existing density of the transportation network in place at any given time. The estimate therefore

represents the marginal change of an additional metro station to the existing network of stations already opened.

In terms of the LEZ, the treatment variable is cumulative in the sense of considering the relative intensity of when the policy entered into effect. Different levels of strictness accompanied the introduction of the different phases of the LEZ, and the cumulative variable captures this increasing intensity and restrictiveness as it was phased in over time.

Choice of Spatial Treatment

In identifying the spatial impacts of interest, it is necessary to determine the extent to where we would expect inter-urban air pollution to be most affected by each transit intervention. This could be highly dependent on the study region of question and conditional on local topographic constraints and the built structure of the city. It is not always clear exactly where and how local transit patterns may respond to various changes to transportation infrastructure, and so spill-over impacts from the creation of a metro station in the urban periphery, for example, could be felt in *Baixa* if periphery residents change their behaviour or commuting patterns into the city centre.

Five different key areas around the city are considered: proximity to the metro station which has opened, proximity to the LEZ boundary, distance to the city centre (*Baixa*), distance to the Tagus riverfront, and distance to the nearest freeway. As metro stations open and public transit becomes more accessible, or with the introduction of a traffic limiting LEZ in the city centre, the alterations to commuting and traffic patterns could impact pollution concentration in any combinations of these areas.

One of the key assumptions when estimating impacts is the stable unit treatment value assumption. This states that the treatment assignment of one observation (in a neighbourhood 'nearby' one of the key areas) does not affect the potential outcome of others (pollution levels in the 'non-nearby' neighbourhoods). This could particularly be the case with a spatial

treatment assignments such as proximity to key areas in the city. Given the continuity of pollution over space, the cut-off treatment threshold between what defines 'nearby' to the area of interest or not can be subjective. Proximate neighbourhoods are necessarily influenced by adjacent pollution, and so the distinction between these spatially treated units and controls can be fuzzy and should be addressed.

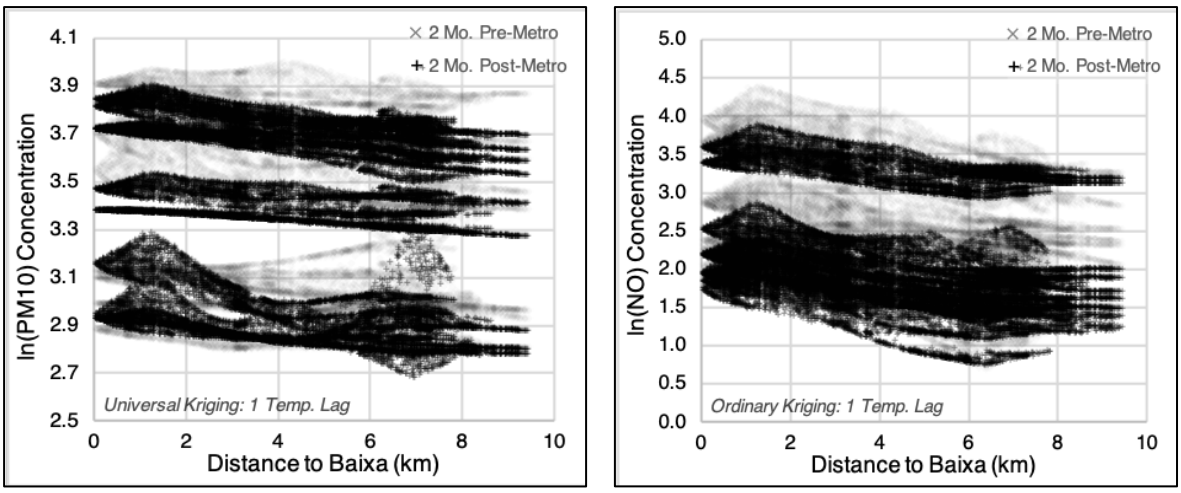
The identification of this cut-off threshold defining the $\{0, 1\}$ spatial treatment needs to represent a positive distance at which point neighbourhoods are considered in proximity to a given area of the city where pollution is likely impacted. These spatially treated units are then compared to those far enough away from these areas so as not to be influenced by the intervention effect (spatial controls). Depending on the context and underlying understanding of the study region, this could range from, for example, considering nearby neighbours as those very local units within 100 meters of downtown versus those within 1,000 meters if we would expect that the transit intervention has a larger impact over space.

To address these possible concerns, a sequence of potential spatial treatment assignments is estimated for all combinations of pollution and interventions. This takes, for example, the distance to the LEZ boundary, and assigns the spatial treatment classification sequentially moving marginally away from the location. This enables a plot of the decaying fuzzy effect that various treatments may have on pollution over space comparing neighbourhoods first at 100 meters away and then at sequentially larger distances.

Each fixed effect model is at every potential spatial treatment assignment in 50 meters intervals between 100 and 500 meters. At larger buffer distances, the impacts are estimated at every 250 meters from 750 meters up to 3 kilometres. The rationale behind estimating the spatial impacts at such a large distance can be seen by looking at an example plot of how PM_{10} or NO concentration decays over space for neighbourhoods pre and post-transit intervention.

Figure 5 shows that decreases in pollution concentration following the introduction of a LEZ was felt up to around 1.25 kilometres away from *Baixa* by comparing the spatial decay two months before and after treatment. Prior to the treatment, pollution effects increased around the city centre locally for both PM_{10} and NO after which point levels begin dropping. After the treatment however, we see consistent spatial decay in levels. This effect is highly conditional on the spatial concept considered, and for example, the spatial decay located around freeways is much more localized and the estimated impacts therefore are only estimated much more locally. Figure A2 of the appendix shows the similar plot for the impact of proximity to freeways.

Figure 5. Spatial Decay of PM_{10} and NO Pre and Post Treatment in Baixa



5.2. Bootstrapped Short-Term Pollution Reduction Impacts

The estimation of the fixed effect parameters in the difference-in-difference specification makes use of the entire time series of data spanning from the start of monitoring to the end of 2016. The estimated impacts are therefore representative of the average effects spread out over this relatively long span. With limited data in the periods directly before and after a given transit intervention, it is difficult to use a panel data structure to estimate the short-term immediate effects occurring directly following a treatment.

To overcome this limitation, a non-parametric bootstrapping process is used to estimate the immediate temporal effects in the subsequent months following an intervention, and the respective decay of these effects both over space and over time. The bootstrapped average treatment effect is based on systematically increasing the estimation by including new observations in one-month intervals before and after the introduction of the treatment.

In considering only observations directly around any of the respective interventions, the sample size of the data decreases, and so too does the statistical power and variation captured by the original model. The non-parametric aspect means that the estimate returns the bootstrapped value of the mean differences across space and time treatments as opposed to being estimated through a deterministic linear model as is done when a longitudinal database format is available. The bootstrap statistic, β_{BS}, which is evaluated for the subsample of temporally proximate observations for all intervals from one month to a year is:

$$\beta_{BS} = Avg.\left[\overline{\mathbb{P}}^{Sp.Treat=1}_{Tm.Treat=1} - \overline{\mathbb{P}}^{Sp.Treat=0}_{Tm.Treat=1}\right] - Avg.\left[\overline{\mathbb{P}}^{Sp.Treat=1}_{Tm.Treat=0} - \overline{\mathbb{P}}^{Sp.Treat=0}_{Tm.Treat=0}\right] \qquad \text{Eq. 5}$$

The same set of spatial treatment effects as in the parametric model are used, and so a bootstrap estimate is obtained for increasing distances from *Baixa*, the Tagus riverfront, opening metro stations, the LEZ boundary and freeways. The temporal treatments, however, are converted to dummy variables representing the month of the intervention. This provides an indicator to identify the specific base reference month when a metro opened, or the LEZ was introduced. From this reference month, estimates for the immediate temporal effects can be calculated by systematically estimating the average impact in the first month post-treatment, the second month post-treatment, and continuing sequentially for up to one year following the transit intervention.[10]

[10] In the case of multiple cumulative treatments, such as opening new metro stations, the mutually exclusive before and after observations are taken. Any observations which happen to fall in the same number of months prior to the opening of a station and equally within the months following the opening of another station, are removed.

For each combination of treatment effect of interest, and now further for each month following the intervention, a bootstrapped resampling is used to estimate the average treatment effect. The idea behind this is to use the limited temporal subsample of observations directly immediate to the transit intervention and resample these values with replacement. The average treatment effect is then estimated for the resampled levels of pollution. A total of 500 iterations for each model are run with estimated impact coming from average effect. Bootstrapped standard errors are also calculated and allow us to evaluate the significance of any estimated impacts.

6. Impact Estimations

This section outlines the results of the estimation for the impact evaluation of the two transit policies of interest: the opening of new metro stations and the introduction of a LEZ. The empirical strategy first looks at the long-run effect of either transit initiative on pollution levels in key areas of the city. This is estimated via difference-in-difference to get the average impact on different neighbourhood locations pre and post intervention. The pollution variables used correspond to the best interpolated model for each respective pollutant as described in section 4.

The difference-in-difference strategy uses the entire time series of data starting after 2000. For every pollutant, this includes up to 216 months and 3,623 neighbourhoods for over 700,000 total observations per model. The average effect across this time span therefore represents a longer-term impact averaged out over almost two decades. Estimates presented are converted directly into the percentage change caused by the interventions.

This model estimation is complemented with the bootstrapped short-run effects which show how pollution levels changed in the months directly after the opening of a new metro station or the timing of the LEZ. The bootstrapped estimates draw from the pollution in neighbourhoods directly surrounding the interventions to compare the pre and post-treatment

effect. It is comparable to the difference-in-difference estimator, calculating the average treatment effect for the subsample of values occurring sequentially in every month before and after the respective treatments up to one year.

Both long-run difference-in-difference estimates and short-run bootstrap estimates for the impact of metros and the LEZ are obtained for every pollutant across the five key areas of the city. The spatial treatment effect estimated varies in distance relative to each of these locations ranging from 100 meters to 3 kilometres away. In order to present all results in a systematic manner, the estimated coefficients from each model are all plotted together to highlight the decaying impact of pollution across space and time.[11]

6.1. Average Short and Long-Run Impacts of Metro Accessibility on Pollution

Particulate Matter

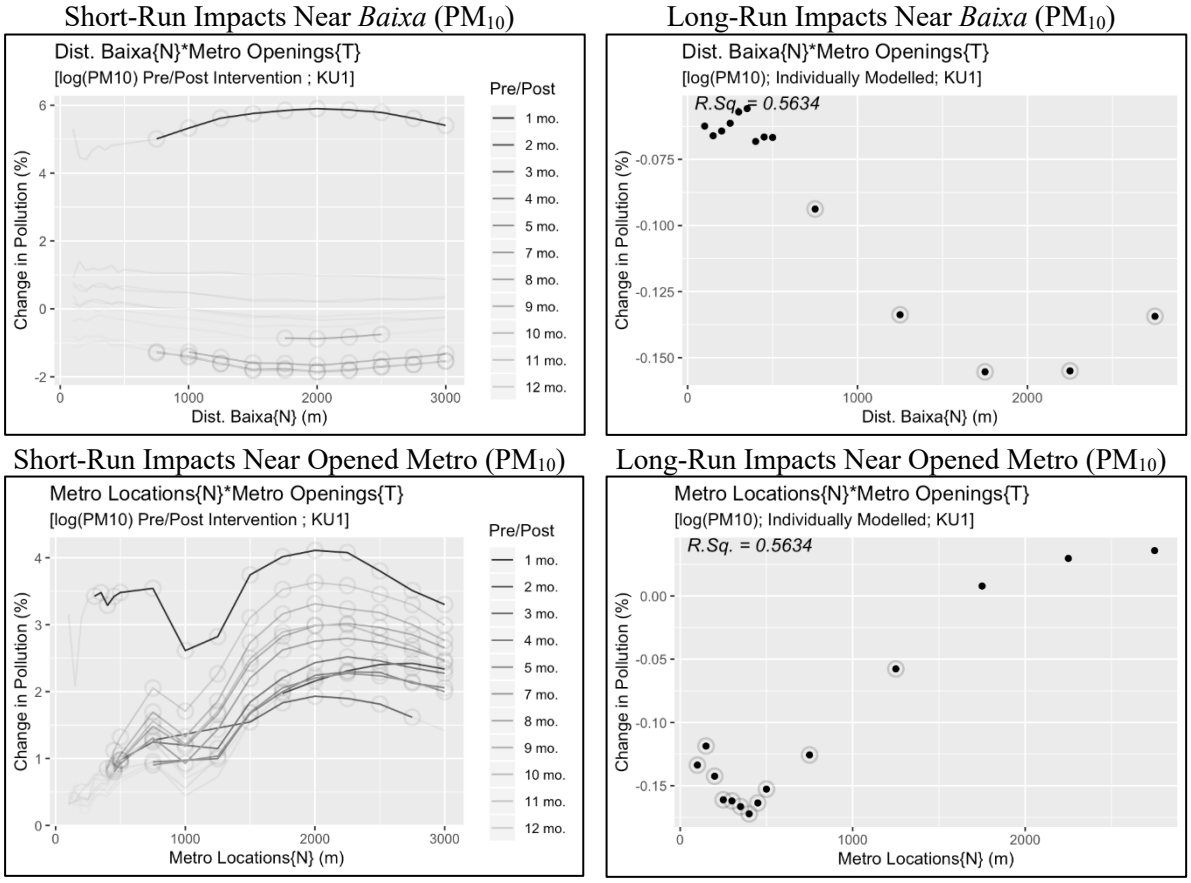
Overall, we observe positive impacts coming from the opening of new metro stations related to the reduction of particulate pollution in the city. Looking first at the direct impact across the entire study region, the average effect of the metro expansion has a median estimated long-run impact of -2.15% across the varying specifications, corresponding to β_2 from equation Eq. 4 – 3.

Focusing on the spatial and temporal decay of effects, estimates show a reduction in PM_{10} following the opening of new metro stations extending to some distance away from the city centre. Figure 6 shows the estimated significant impact of metro openings on sequentially increasing spatial treatments, increasing from locally within 100 meters of the city centre and opening metro stations up to 3 kilometres away. The average value of the R^2 is presented across each of the difference-in-difference models. Consistent with expectations from figure 5, the

[11] For brevity, only the most salient features and results are presented. Estimated values, and respective significance levels, are plotted sequentially in order to highlight spatial trends and robustness in considering marginal increases in the estimated values over time and space.

significant impacts are strongest at the point where post-treatment pollution levels appear to decay away from the city centre.

Figure 6. Short and Long-Run Metro Opening Impacts on PM_{10}

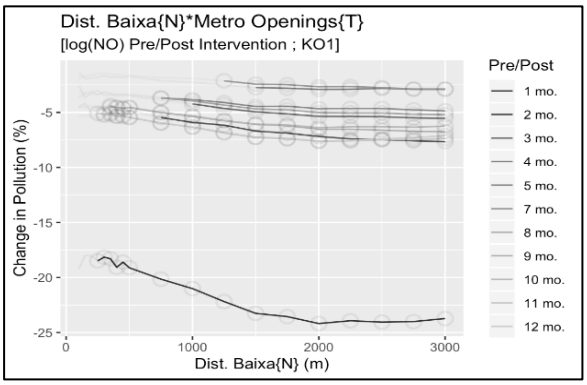


Results indicate that up to approximately 2 kilometres away from the city centre new metro stations across the city has yielded a long-term reduction in local particulate matter around 0.15%. In the short term, the bootstrap plots in the months directly before and after an opening indicate some immediate increase in PM_{10} around *Baixa*, a pattern further observed around the location of the metro station itself. With construction and preparations ongoing up to and including opening day, the immediate jump in pollution could be measuring this effect, compounded if the station opened towards the end of the month. This effect reduces in the subsequent months and immediate short-run local pollution levels reduce by almost 2% up to two and three months following the opening of a new station.

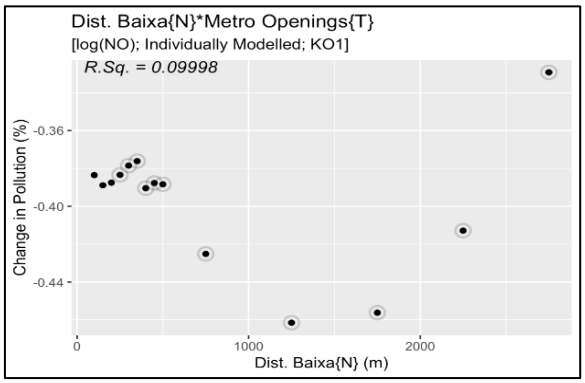
Around the metros, PM_{10} has a long term drop up to 0.175% in the immediate vicinity of the station up to 500 meters away. Weaker yet still negative reductions around 0.05% continue to be experienced in neighbourhoods up to around 1.5 kilometres away from the metro station. Although there appear to be some short-term increase in pollution in the months following an opening, this effect again dissipates over time.

Figure 7a. Metro Opening Impacts on Combustion-Based Emissions Near *Baixa*

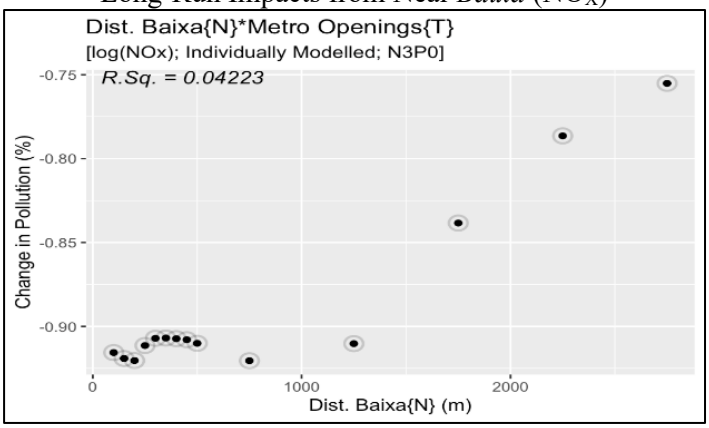
Short-Run Impacts Near *Baixa* (NO)



Long-Run Impacts Near *Baixa* (NO)



Long-Run Impacts from Near *Baixa* (NO_X)



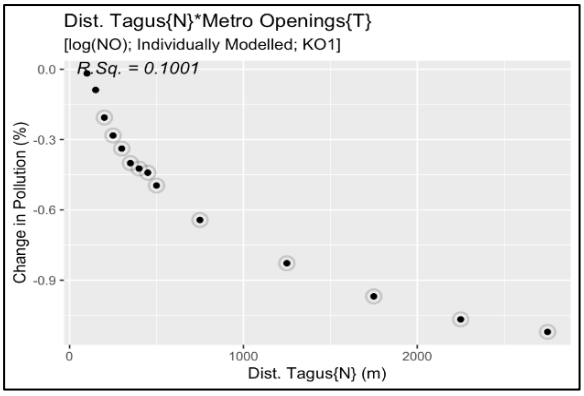
Nitrogen-Based Combustion Emissions

Metro openings decreased NO in *Baixa* by around 0.45% up to around 1.5 kilometres in the long run as seen in figure 7a. Short term effects indicate the largest decrease in the first month, around 20%, and in subsequent months this effect dies off. Up to almost a year later, however, impacts of 2.5% could be felt at large 2 kilometer distances from *Baixa*. Similarly, very localized spatial effects around *Baixa* show decreases in NO_X around 0.9%.

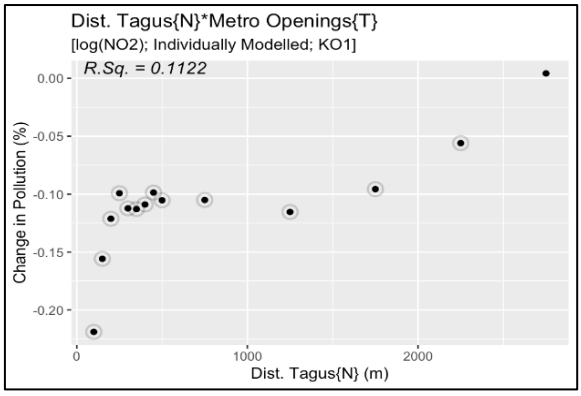
Combustion emissions have the greatest reduction around the riverfront with NO, NO_2 and NO_X all decreasing in this area following the introduction of new metro stations, as seen in figure 7b. The reduction of NO and NO_X is broad, extending away from the Tagus and reaching up to 0.9% and 0.2% up to 3 kilometres away. Around metro stations themselves, there are local reductions of NO_X around 0.3%. This impact dies off over space but decreases NO_X in neighbourhoods up to 2 kilometres away from newly opened stations.

Figure 7b. Metro Opening Impacts on Combustion-Based Emissions Near the Tagus

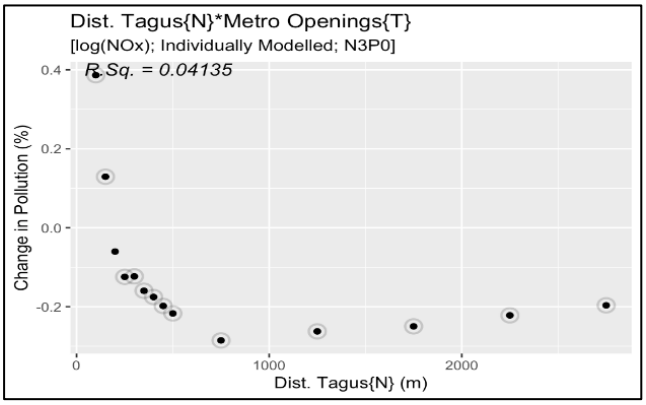
Short-Run Impacts Near the Tagus (NO)



Long-Run Impacts Near the Tagus (NO_2)



Long-Run Impacts from the Tagus (NO_X)



Carbon Monoxide

Although the average predictive power of the interpolated data at each time interval for CO has little improvement, it is still possible that a difference-in-difference specification can capture the temporal impacts from the observed point values. Broadly, the direct estimated impacts caused by the cumulative opening of metro stations, corresponding to β_2, has a median

average reduction of 2.18%. This captures the general temporal impact of these metro treatments on neighbourhood CO concentration.

As one of the primary sources of CO is vehicle emission, we would expect that the opening of new metro stations may reduce emissions from private drivers. Estimates indicate localized reductions near metro stations and broader impacts in *Baixa*, 0.25% up to 750 meters and 0.29% up to 3 kilometres respectively (figure A3 of the appendix).

4 – 6.2. Average Short and Long-Run Impacts of Low Emission Zones on Pollution

Particulate Matter

The before and after difference is generally larger when comparing the interventions surrounding the LEZ and the introduction of new metro stations. This however only evaluates pre and post intervention averages and the complex mechanisms of pollution abatement cannot cleanly be attributed to any specific source. The larger decrease following the LEZ is not surprising given the zones explicit aim of reducing pollution, however should be interpreted with caution given the number of transit initiatives introduced in conjunction with the LEZ, and thus could be capturing the combined effect of this temporal dynamic.

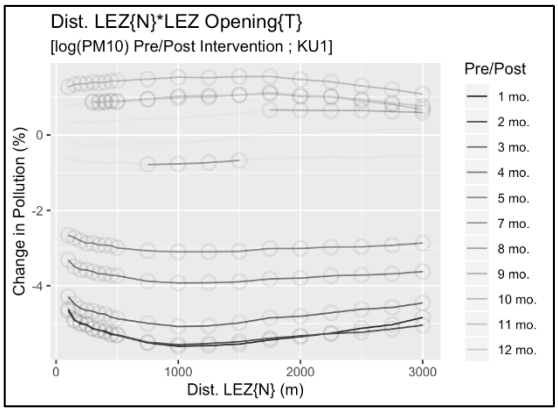
The difference-in-difference estimate can identify the local change in the area directly around the LEZ itself where pollution reduction is more targeted. A long-term reduction up to 0.30% is estimated in the immediate area around the LEZ, with broader impacts extending up to 3 kilometres away from the boundary with a broad area reduction up to 0.43%. Similar effects are confirmed when considering the distance to *Baixa,* where the LEZ is located, with broad effects up to 0.25% felt up to 3 kilometres. The effect of the LEZ has thus had a wide impact on pollution in the city centre over the long term since its introduction.

Large local decreases of PM_{10} are observed in the months directly following the intervention as expected. In the first 5 months after the introduction, the reduction ranged

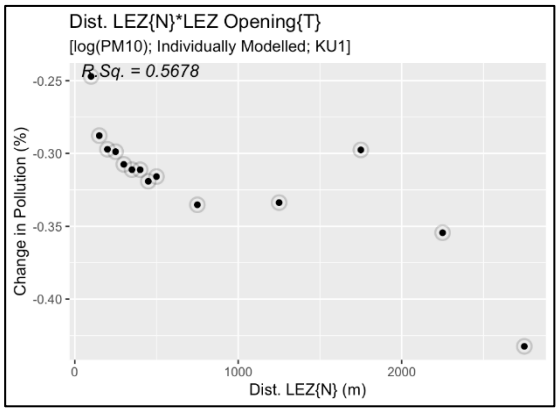
between around 2% to 4%. This effect appears to die off over time and revert, and also observed in the literature as drivers or enforcers may become complacent and regulations laxed.

Figure 8. Short and Long-Run LEZ Impacts on PM_{10} and NO_X

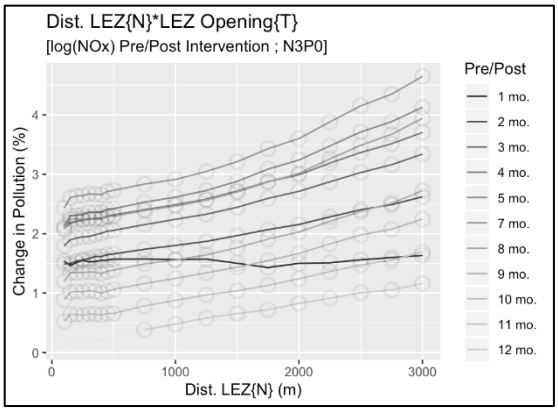
Short-Run Impacts Near LEZ (PM_{10})



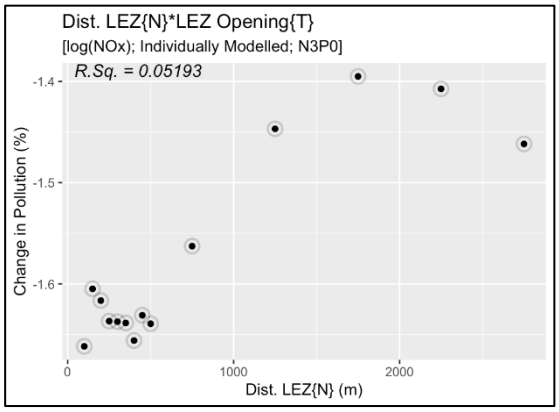
Long-Run Impacts Near LEZ (PM_{10})



Short-Run Impacts Near LEZ (NO_X)



Long-Run Impacts Near LEZ (NO_X)



Interestingly, we see some increases in pollution along the Tagus river over the long run, however small. The estimated impact represents an increase in pollution around 0.2% near the waterfront. While the zone itself has improved in air quality, there is some suggestion that the introduction of the LEZ may have led to shifts in driving patterns where now commuters or heavy vehicles entre into the city centre via alternative routes – many of which are along the river.

Nitrogen-Based Combustion Emissions

In terms of NO_X there are long-run reductions in pollution due to the introduction of the LEZ very localized of around 1.65%. This reduction decays closer to zero over space indicating the strongest reductions of NO_X are closer to the LEZ boundary after its introduction. The short run dynamics reveal some increases in localized pollution. In the month directly following the introduction of the LEZ pollution increase was constant across space, however, in the subsequent months NO_X levels further away from the LEZ began increasing at relatively faster rates.

As LEZ's primarily target traffic patterns, these short-term increases could be a result of peripheral areas surrounding the LEZ having increased spill-over traffic to avoid entering the zone after it was introduced. This can further be seen with slight increases in freeway pollution following the introduction of the LEZ in the magnitude of 0.17% within 200 meters. If traffic patterns are indeed shifted, then we would expect that drivers may take alternative routes to by-pass the LEZ and enter the city centre.

Reductions of NO_2 were broader and included decreases around the Tagus riverfront of around 1.5%. The reductions are very localized and significant for NO_2, and short-term impacts indicate strong decreases very nearby in the months immediately following the introduction which decay both over space and time. Thus, while the LEZ was efficient in reducing the concentration of NO_2 it is a very localized effect.

Similar effects are seen with NO, however, are much more extensive. The pollution reduction of around 3.5% caused by the LEZ and stemming from the riverfront was experienced by neighbourhoods up to 3 kilometres away. Short-run impacts mirror this trend with strong decreases in NO concentration which remain large as we consider a larger zone around the Tagus.

Sulfur Dioxide

The impacts on SO_2 from models looking at metro openings have a poor fit and results that are not meaningful, however estimates are much more robust and intuitive when considering the impacts from the LEZ. This shows the importance of considering the full resulting outcomes of each treatment as they related to different urban transports and ultimately different pollutants. While metro policies are targeted towards light vehicle drivers, incentivizing them to use public transport, LEZ's are targeted towards more heavy polluting vehicles. So, while a metro treatment is not likely to impact SO_2 emissions, given the low proportion coming from light vehicles, a LEZ policy targeting these heavy vehicles directly should.

Similar localized reductions in SO2 are seen around *Baixa* and the LEZ itself from around 5.8% in the immediate proximity with impacts extending out to 2 kilometres from *Baixa* and 1 kilometre from the LEZ. Along the Tagus riverfront, the introduction of the LEZ yielded reductions of over 5% in neighbourhoods within up to 3 kilometres. As we would expect transit patterns to shift, further reductions along the freeways within the city are estimated to reach around 1.5% up to 750 meters away. Plots for auxiliary impacts from sulfur dioxide available in figure A4 of the appendix.

7. Conclusions

Although this work does not aim to evaluate the full costs and benefits associated to metro expansions and LEZ's, it does provide context to the relative impact of each initiative in terms of pollution changes and spatial-temporal patterns. The results indicate that within-urban spatial and temporal decay patterns are significant and do not impact all neighbourhoods equally. Transit interventions therefore have potentially very localized effects which may cause varying impacts conditional on how commuting patterns respond. As pollution concentrations

tend to be spatially concentrated in urban areas, it is important to take these high-resolution dynamics into account when evaluating different transit interventions.

Understanding the local and neighbourhood level impacts of pollution can be used to better target hotspots where residents are at greater risks of negative impacts. Further, non-transit related pollution mitigation strategies, such as the planting of trees or developing green infrastructures (walls, roofs, parks), can be better targeted knowing where pollution is most prevalent. While long run pollution reduction around newly opened metro stations are observed, the largest effect appears to come generally from the reduction in pollution in the city centre or near the riverfront. This is similarly seen with the resulting impacts from the introduction of the LEZ, located in the centre.

Of particular interest from these results, in comparison to general average city-level impacts, there is some evidence to support the idea that the introduction of the LEZ yields alterations to transit patterns and ultimately pollution. Results suggest a post intervention increase in pollution at the external boundaries of the LEZ and along freeways. Further, results highlight that certain interventions are only appropriate for certain goals, for example, the goal of reducing SO_2 emissions should be concentrated on LEZ enforcement of heavy-duty vehicles rather than incentivizing light vehicle drivers to switch to public transit.

The results from the applied study highlight the benefit to be gained in terms of increased spatial and temporal complexity and understanding of the impacts. While in the relative context Lisbon is not a heavily polluted city, there are still clear environmental improvements to be made and best practices for other municipalities with similar infrastructure and conditions. The broader impact of metro stations and LEZ fall in line with estimated impacts from other large cities, particularly when looking at the larger short run effects estimated compared with more traditional point differences in pollution levels at different stations inside and outside of LEZs.

More important considerations should further be given to the relative costs of each transit intervention. LEZ's are very low-cost relative to the expansion of a new metro, which include capital heavy construction projects rather than simple enforcement. From a purely pollution reduction point of view therefore, the results from the LEZ suggest that interventions surrounding this goal had relatively large long run reductions in local pollutants.

However, this does not negate or diminish the benefit of metro expansions on pollution reduction. Since the primary goal of metro expansions is not the reduction of pollution, any observed decrease in pollutants constitute a positive spill-over. This value of pollution reduction is a non-market benefit that is often not taken into account when estimating the costs and benefits of any intervention. Therefore, the value of pollution reduction via metro expansions is crucial for evaluating the true benefits, based on current dynamics.

This work looks at the role of geostatistical methods to generate longitudinal data for detailed urban environmental statistical analysis. The application of these methods highlight the increasing detail with which urban analytics can be used to better understand dynamic processes at a highly refined spatial and temporal detail, often lacking in studies of the urban environment. This highlights the advantage of leveraging currently available geospatial data to estimate and value urban processes and impacts. From a practical point of view these estimates are crucial for better understanding urban dynamics in different contexts and better valuing location spillovers.

These procedures and estimates highlight the quality of open source data for the generation of high-dimensional longitudinal neighbourhood level databases, enabling the study of spatial and temporal dynamics of urban pollution. As data and computational power allows for higher resolution data at a higher frequency, new and relevant urban-scale intervention analyses can be conducted to better guide any discussions and best practices related to transit or other policy areas influencing the urban environment.

References

Adler, Martin W., and Jos N. van Ommeren. 2016. "Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes." *Journal of Urban Economics* 92: 106-119.

Anderson, Michael L. 2014. "Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion." *American Economic Review* 104 (9): 2763-2796.

Anselin, Luc, and Nancy Lozano-Gracia. 2008. "Errors in variables and spatial effects in hedonic house price models of ambient air quality." *Empirical Economics* 34 (1): 5-34.

Anselin, Luc, and Julie Le Gallo. 2006. "Interpolation of Air Quality Measures in Hedonic House Price Models: Spatial Aspects." *Spatial Economic Analysis* 1 (1): 31-52.

Baum-Snow, Nathaniel, and Matthew E. Kahn. 2000. "The effects of new public projects to expand urban rail transit." *Journal of Public Economics* 77: 241-263,

Baum-Snow, Nathaniel, and Matthew E. Kahn. 2005. "Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities, 1970-2000." *Brookings-Wharton Papers on Urban Affairs* 147-206.

Bertazzon, Stefania, Markey Johnson, Kristin Eccles, and Gilaad G. Kaplan. 2015. "Accounting for spatial effects in land use regression for urban air pollution modelling." *Spatial and Spatio-temporal Epidemiology* 14-55: 9-21.

Boogaard. Hanna, Nicole A.H. Janssen, Paul H. Fischer, Gerard P.A. Kos, Ernie P. Weijers,Flemming R. Cassee, Saskia C. van der Zee, Jeroen J. de Hartog, Kees Meliefste, Meng Wang, Bert Brunekreef, and Gerard Hoek. 2012. "Impact of low emission zones and local traffic policies on ambient airpollution concentrations." *Science of the Total Environment* 435-436: 132-140.

Bowes, David R., and Keith R. Ihlanfeldt. 2001. "Identifying the Impacts of Rail Transit Stations on Residential Property Values." *Journal of Urban Economics* 50: 1-25.

Câmara Municipal de Lisboa (CML). 2018a. "Lisbon: The Economy in Figures."

Câmara Municipal de Lisboa (CML). 2018b. *European Green Capital 2020 – Application*.

Cervero, Robert, and Change Deok Kang. 2011. "Bus rapid transit impacts on land uses and land values in Seoul, Korea." *Transport Policy* 18 (1): 102-116.

Chen, Yihsu, and Alexander Whalley. 2012. "Green Infrastructure: The Effects of Urban Rail Transit on Air Quality." *American Economic Journal: Economic Policy* 4 (1): 58-97.

Ellison, Richard B., Stephen P. Greaves, and David A. Hensher. 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.

Ferreira, Francisco, Pedro Gomes, Ana Cristina Carvalho, and Hugo Tente. 2012. "Evaluation of the Implementation of a Low Emission Zone in Lisbon." *Journal of Environmental Protection* 3 (9): 1188-1205.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A. Turner. 2018. "Subways and Urban Air Pollution." *National Bureau of Economic Research* Working Paper 24183.

Goetzke, Frank. 2008. "Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York." *Urban Studies* 45 (2): 407-417.

Gonçalves, Rui N. 2014. The Lisbon Low Emission Zone Enforcement Methods and Procedures. Portuguese Environment Agency. Presentation to then Workshop on Low Emission Zones (European Comission TAIEX).

Gonzalez-Navarro, Marco, and Matthew A. Turner. 2018. "Subways and urban growth: Evidence from earth." *Journal of Urban Economics* 108: 85-106.

Grupo Metropolitano de Lisboa (ML). 2019. "Relatórios e Contas Consolidado 2018"

Holman, Claire, Roy Harrison, and Xavier Querol. 2015. "Review of the efficacy of low emission zones to improve urban air quality in European cities." *Atmospheric Environment* 111: 161-169.

Instituto Nacional de Estatistica (INE). 2011. Mode of transport used on commuting (No.), Population and housing census - 2011.

Instituto Nacional de Estatistica (INE). 2018. Veículos novos vendidos por 1000 habitantes (N.º) por Local de residência (NUTS - 2013) e Tipo de veículo; Anual - Conservatórias do Registo Automóvel.

Kawabata, Mizuki, and Qing Shen. 2007. "Commuting Inequality between Cars and Public Transit: The Case of the San Francisco Bay Area, 1990-2000." *Urban Studies* 44 (9): 1759-1780.

Li, Zheng. 2018. "The impact of metro accessibility on residential property values: An empirical analysis." *Research in Transportation Economics* 70: 52-56.

Martinez, L. Miguel, and Jose Manuel Viegas. 2009. "Effects of Transportation Accessibility on Residential Property Values: Hedonic Price Model in the Lisbon, Portugal, Metropolitan Area." *Transportation Research Record: Journal of the Transportation Research Board* No. 2115.

Meinardi, Simone, Paul Nissenson, Barbara Barletta, Donald Dadbub. 2008. "Influence of the public transportation system on the air quality of a major urban centre. A cast study: Milan, Italy." *Atmospheric Environment* 42: 7915-7923.

Mohammad, Sara I., Daniel J. Graham, Patricia C Melo, Richard Anderson. 2013. "A meta-analysis of the impact of rail projects on land and property values." *Transportation Research A* 50: 158-170.

Mulley, Corinne, Chi-Hong (Patrick) Tsai. 2016. “When and how much does new transport infrastructure add to property values? Evidence from the bus rapid transit system in Sydney, Australia.” *Transport Policy* 51: 15-23.

Mulley, Corinne, Chi-Hong (Patrick) Tsai, and Liang Ma. 2018. “Does residential property price benefit from light rail in Sydney?” *Research in Transportation Economics* 67: 3-10.

Nunes da Silva, Fernando, Renata A. Lajas Custodio, and Helena Martins. 2014. “Low Emission Zone: Lisbon’s Experience.” *Journal of Traffic and Logistic Engineering* 2 (2): 133-139.

Pereira, Boscolli Barbosa, Edimar Olegario de Campos Jr., and Euclides Antonio Pereira de Lima. 2013. “Biomonitoring air quality during and after a public transportation strike in the centre of Uberlandia, Minas, Gerais, Brazil by Tradescantia micronucleus bioassay.” *Environmental Science and Pollution Research* 21 (5): 3680-3685.

Roukouni, A, Basbas S, and Kokkalis A. 2012. “Impacts of metro station to the land use and transport system: Thessaloniki Metro case.” *Transport Research Arena – Europe 2012* 48: 1155-1163.

Sims R., R. Schaeffer, F. Creutzig, X. Cruz-Núñez, M. D’Agosto, D. Dimitriu, M. J. Figueroa Meza, L. Fulton, S. Kobayashi, O. Lah, A. McKinnon, P. Newman, M. Ouyang, J. J. Schauer, D. Sperling, and G. Tiwari, 2014: Transport. In: Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Sanchez, Thomas W. 1999. “The Connection Between Public Transit and Employment.” *Journal of the American Planning Association* 65(3): 284-296.

Santos, Francisca M., Alvaro Gomez-Losada, and Jose C. M. Pires. 2019. “Impact of the implementation of Lisbon low emission zone on air quality.” *Journal of Hazardous Materials* 365 (5): 632-641.

Titos, G., H. Lyamani, L. Drinovec, F.J. Olmo, G. Mocnik, and L. Alados-Arboledas. 2015. “Evaluation of the impact of transportation changes on air quality.” *Atmospheric Environment* 114: 19-31.

Viard, V. Brian, and Shihe Fu. 2015. “The effect of Beijing’s driving restrictions on pollution and economic activity.” *Journal of Public Economics* 125: 98-115.

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

Wolff, Hendrik, and Lisa Perry. 2010. “Trends in Clean Air Legislation in Europe: Particulate Matter and Low Emission Zones.” *Review of Environmental Economics and Policy* 4 (2): 293-308.

Xu, Yangfei, Qinghua Zhang, and Siqi Zheng. 2015. “The rising demand for subway after private driving restriction: Evidence from Beijing’s housing market.” *Regional Science and Urban Economics* 54: 28-37.

Zhang, Wei, C-Y Cynthia Lin Lawell, and Victoria I. Umanskaya. 2017. “The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence.” *Journal of Environmental Economics and Management* 82: 181-220.

Zheng, Siqi, Xiaonan Zhang, Weizeng Sun, and Jianghao Wang. 2019. “The effect of a new subway line on local air quality: A case study in Changsha.” *Transportation Research Part D* 68: 26-38.

Appendix

Figure A1. PM_{10} Optimal Interpolation and Aggregation

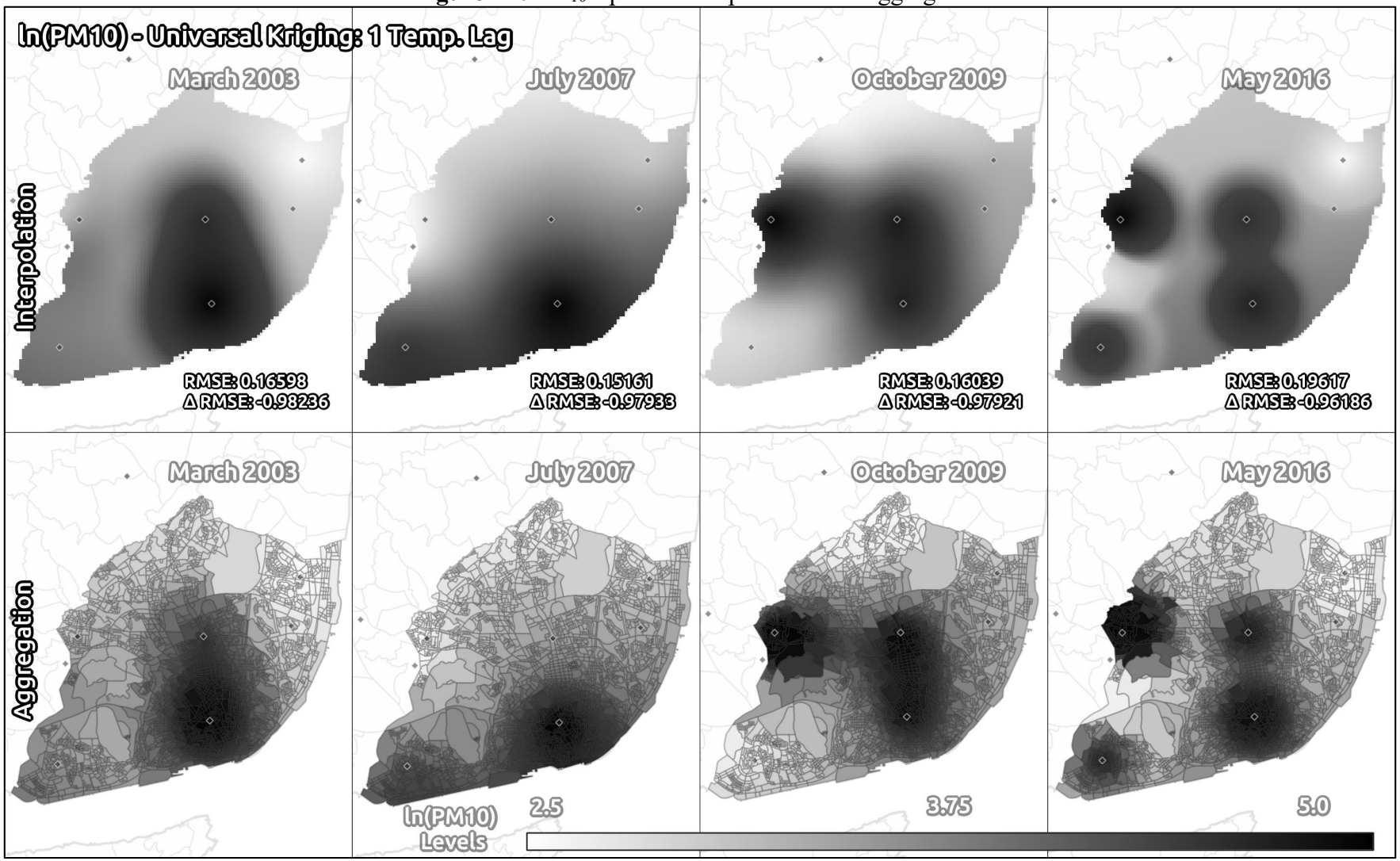


Figure A2. Spatial Decay of PM_{10} and NO Pre and Post Treatment Near Freeway

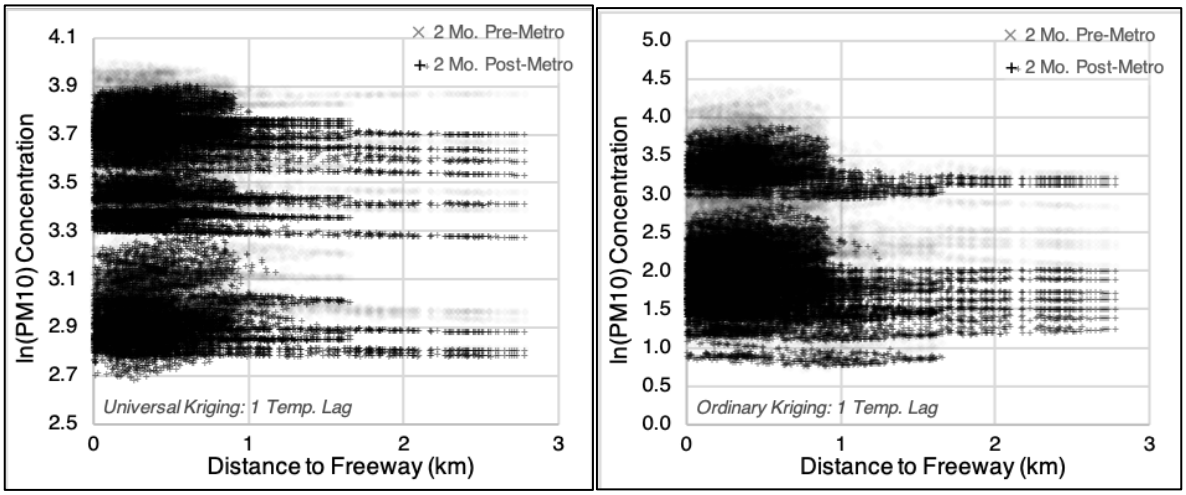


Figure A3. Carbon Monoxide and Auxiliary Plots

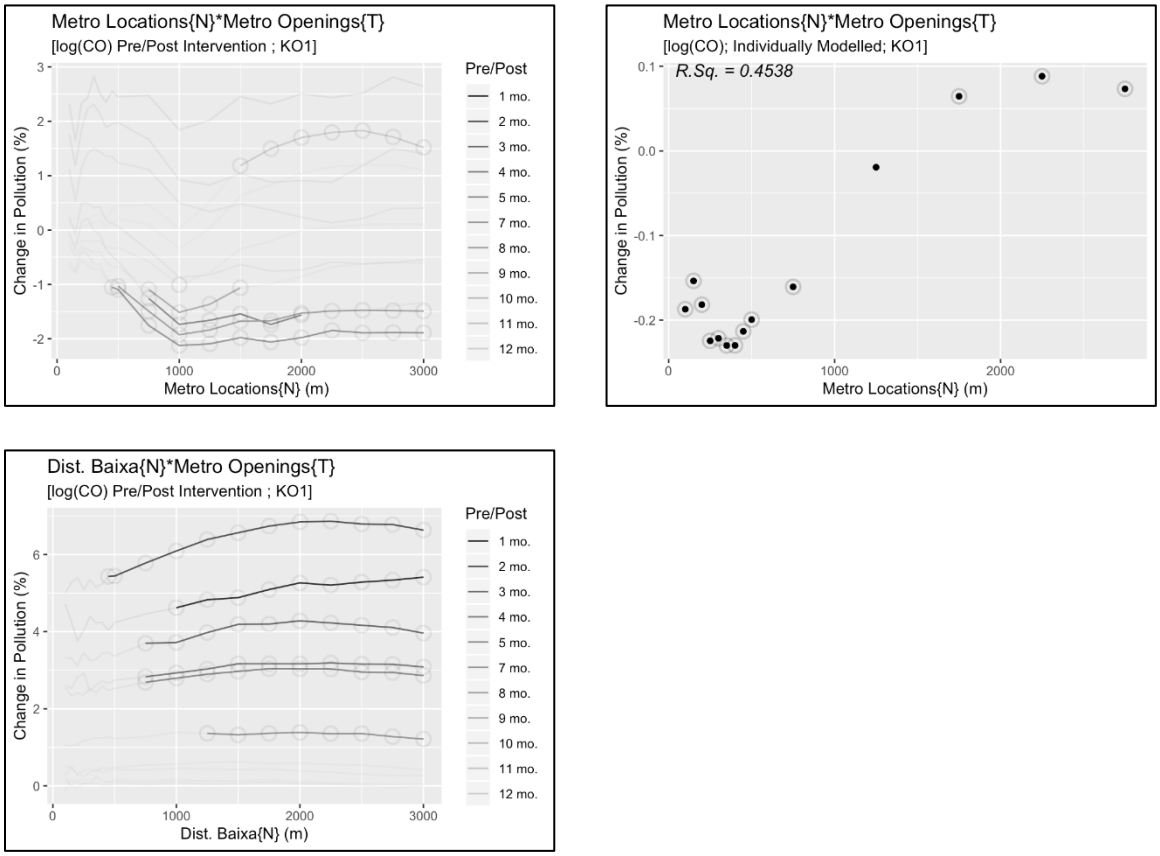


Figure A4. Auxiliary Plots of LEZ Impacts