Journal of Environmental Economics and Management 83 (2017) 121–144

Contents lists available at ScienceDirect

Journal of
Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

The effect of low emission zones on air pollution and infant health ☆

Markus Gehrsitz

University of Strathclyde, Department of Economics, Strathclyde Business School, Sir William Duncan Building, 199 Catherdral Street, Glasgow G4 0QU, UK

ARTICLE INFO

Article history:
Received 15 March 2016
Available online 9 February 2017

JEL classification:
I18
Q52
Q53
Q58

Keywords:
Air pollution
Particulate matter
Birth weight
Infant health
Low emission zones
Policy evaluation

ABSTRACT

This paper investigates the effect of low emission zones on air quality and birth outcomes in Germany. The staggered introduction of the policy measure creates a credible natural experiment and a natural control group for births and air pollution measurements in cities that enact low emission zones. I show that the introduction of the most restrictive type of low emission zone decreases average levels of fine particulate matter by about 4 percent and by up to 8 percent at a city's highest-polluting monitor. Low emission zones also reduce the number of days per year on which legal pollution limits are exceeded by three. However, these reductions are too small to translate into substantial improvements in infant health. My results are not driven by changes in maternal or city specific characteristics, and are robust to variations in specification and to the choice of control group.

Introduction

Substantial policy initiatives are aimed at improving air quality in order to avoid the adverse effects associated with exposure to air pollution. Legislation such as the Clean Air Act of 1970 (Chay and Greenstone, 2003a) sets limits for pollutants. Cities have experimented with introducing driving restrictions (Davis, 2008), congestions charges (Leape, 2006), and electronic toll collections systems (Currie and Walker, 2011) in order to cut pollution levels. In Germany, a number of cities have recently enacted an even more aggressive policy measure to curb air pollution stemming from motor vehicle exhaust, namely Low Emission Zones (LEZs). LEZs impose an outright ban of vehicles that fail to meet pollution standards from a city's streets.

Yet, surprisingly little is known about the effectiveness of this policy measure and LEZs remain controversial. The General German Automobile Club has conducted a study of the impact of LEZs on air quality for three metropolitan areas and found no effect on air quality (Laberer and Niedermeier, 2009). The results, methodology, and somewhat arbitrary selection of those three treatment cities have been attacked by environmental groups as "flawed" and an "apples to orange comparison" (Reh, 2009). Several cities have commissioned evaluations of their own LEZs. These analyses usually follow a simple and

☆ I thank the Research Data Center of the German Statistical Offices of the Laender for making their birth records available to me through remote execution. I am grateful to Dr. Stefan Weil and Michelle Wittler for their help in facilitating the analysis of data. I benefited from the helpful comments and suggestions of Michael Grossman and David Jaeger.

E-mail address: markus.gehrsitz@strath.ac.uk

http://dx.doi.org/10.1016/j.jeem.2017.02.003
0095-0696/

potentially inadequate first-difference approach and often lead to conflicting findings (Morfeld et al., 2011; Cyrys et al., 2014, among others). The most reliable evaluation to date comes from Wolff (2014), who analyzed the effect of LEZs on pollution levels in 9 (out of today 44) cities which have introduced an LEZ, using a difference-in-differences approach. His study finds that pollution levels were indeed reduced by the introduction of LEZs.

This study builds on Wolff (2014) and provides a comprehensive analysis of the effectiveness of LEZs. I use a similar identification strategy, but extend the data by including more treatment cities over a longer period of time. This allows for an evaluation of different types of LEZs which come in different stages and typically become more restrictive over time. This is an important innovation as the toughening of restrictions over time turns out to be driving much of the pollution reductions associated with LEZs. What is more, a longer evaluation period also allows me to restrict my sample to cities that ultimately adopt an LEZ which lends additional support to my identifying assumptions. Moreover, it allows for a thorough investigation of spillovers and spatial leakage. I investigate whether residents of cities that have not (yet) introduced a zone benefit from LEZs in nearby cities, for example because commuters from these nearby cities have switched to cleaner vehicles. In the same vein, this study investigates spillovers within cities by explicitly distinguishing between pollution monitors that are located within or outside a city's LEZs. Furthermore, my analysis is novel in that it extends beyond an evaluation of average fine particulate concentrations in the air and includes an analysis of the number of days on which daily pollution limits were be met. This is of particular interest to policy makers as it is usually this metric that determines (non)-compliance with regulatory limits that trigger legal, political and financial penalties. Finally, this paper explicitly models health effects. LEZs were primarily introduced as a tool to promote public health rather than to cut pollution per se. This is the first study to look at the health effects of LEZs by leveraging a database that contains detailed information on the universe of births in Germany between 2005 and 2012. This unique data source allows me to investigate whether potential pollution reductions achieved by LEZs result in improved infant health.

An evaluation of LEZs is interesting for several reasons. First, this study can help settle the fierce debate about the success or failure of LEZs and whether this policy measure can be a role model for other cities and countries. This paper is, therefore, of substantial importance to policy makers. Second, the staggered introduction of LEZs throughout Germany provides an attractive natural experiment that overcomes problems of observational studies in estimating causal effects of policy interventions on pollution reduction and health improvement. I focus on fine particulate pollution since measurements for this type of pollution are widely available and the legislation underpinning LEZs was explicitly aimed at these pollutants. In terms of health outcomes, this study focuses on infant health, in particular birth outcomes for which changes in pollution manifest themselves with little delay.

This study finds little in the way of improvements in infant health due to the introduction of an LEZ. Neither average birthweight nor the prevalence of low-weight births appear to be significantly affected by the policy, although I find small reductions in the incidence of stillbirths. This is not to say that LEZs are ineffective in terms of reducing pollution. I find that LEZs reduce PM_{10} concentrations by about 2.5 percent. Reductions are about twice this size at monitors that are located close to major roads. These drops are not offset by pollution increases in areas not covered by an LEZ. On the contrary, small pollution reductions - albeit often statistically insignificant - can also be observed at background stations outside a city's center. I also find that more restrictive zones that only allow the cleanest vehicles into a city's center are driving much of these results and are much more effective than the less restrictive version that was introduced initially. These restrictive zones are associated with reductions in fine particulate pollution of about six percent at a city's highest-polluting monitoring stations. More importantly from a regulatory perspective, the average LEZ reduces the number of days on which pollution limits are exceeded at any of a cities' monitoring stations by three days per year. Since these daily limits are harder to adhere to than limits on average pollution, LEZs appear to make a considerable contribution to cities' compliance with regulatory standards.

The remainder of this paper is structured as follows. Section "Background" provides background information on the timing and nature of LEZ introductions throughout Germany. Section "Data and descriptive evidence" describes my data and Section "Methods" introduces my method and identification strategy. Section "Results" presents my results, provides a thorough discussion of these results, and conducts several robustness checks. Section "Conclusion" concludes.

Background

A key feature of this study is that it exploits a policy or regulatory change that plausibly creates exogenous variation in pollution. Such natural experiments are rare, although economists have been successful in identifying and analyzing the few that exist (Graff Zivin and Neidell, 2013, provide a comprehensive overview). For instance, a series of papers evaluates the effects of the closure and subsequent reopening of a Utah steel Mill in 1987/88 (Pope, 1989; Pope et al., 1992). Chay and Greenstone (2003b) exploit a sharp drop in pollution due to the 1981-1982 recession to establish a link between the concentration of total suspended particulates (TSPs) in the air and infant mortality. Luechinger (2014) investigates the impact of mandated desulfurization at power plants in Germany, which resulted in lower sulfur dioxide concentrations, on infant mortality. Currie and Walker (2011) assess the effect of the introduction of E-ZPass, an electronic toll collections system in New Jersey and Pennsylvania, on birth outcomes and pollution levels. The literature's focus on newborns' health is partly due to the fact that the link between in utero exposure to pollution and adverse health is immediate (Glinianaia et al., 2004; Wang et al., 1997). Currie and Neidell (2005) make a compelling case for this link using a rich data set from California.

Currie et al. (2009) also show that high concentrations in carbon monoxide have adverse effects on infant health in New Jersey. Ha et al. (2001) use data from Seoul, South Korea, to confirm that in utero exposure to pollutants is strongly correlated with low birth weight. Birth weight, in turn, is an important determinant of long-run outcomes such as adult height, intelligence, earnings and educational attainment (Behrman and Rosenzweig, 2004; Black et al., 2007; Currie, 2011). In the short run, low birth weight is strongly associated with neonatal mortality and low health (Almond et al., 2005; Oreopoulos et al., 2008). Newborns are also a very vulnerable part of the population and therefore require special protection.

Much of recent US and EU regulation of air pollution has been targeting fine particulate matter, especially PM_{10} and $PM_{2.5}$. These are fine particles, solid or liquid, which have a diameter of 10 micrometers (μm) or less, and 2.5 μm or less, respectively. Fine particulates are often tainted with traces of heavy metal or other carcinogenic substances. The smaller a particle, the easier it can enter the human lung and blood stream. The European Commission estimates that in the EU, PM_{10} is causing around 348,000 premature deaths every year, making it by far Europe's deadliest air pollutant (Watkiss et al., 2005). It is also the main culprit in hundreds of thousands of respiratory or cardiac hospital admissions. The main source of fine particulate matter are motor vehicles. They account for thirty to fifty percent of PM_{10} emissions and are also a major source of carbon monoxide (CO) and nitrogen dioxide (NO_2) emissions (Friedrich, 2006).

Several EU council directives set numeric limits for the permissible concentration of pollutants in the air. Council Directive 2000/69/EC prohibits EU cities from exceeding a daily eight hour mean of 10 milligrams per cubic meter (10 mg/m^3) of carbon monoxide (CO) from 2005 on. Council Directive 1999/30/EC sets a limit of 40 μg/m^3 for the average annual NO_2 concentration, and an hourly limit of 200 μg/m^3. Both NO_2 limits were effective as of 2010 and the hourly limit was only allowed to be exceeded 18 times per calendar year. The same directive also regulates sulphur dioxide (SO_2) emissions and establishes an hourly limit of 350 μg/m^3 and a daily limit of 125 μg/m^3.[1] Council Directive 1999/30/EC also enforced limits for fine particulate pollution in two phases. In phase 1 (January 2005 through December 2009), measurements of PM_{10} levels were not allowed to exceed a yearly average of 40 μg/m^3 at any of a city's monitoring stations. In addition, the daily average at any of a city's monitors was not allowed to be higher than 50 μg/m^3 on more than 35 days. In phase 2, starting in January 2010, these limits were supposed to be raised to a annual average of 20 μg/m^3 and only 7 exceedance days per year. In particular the latter limit proved to be very hard to attain. In 2007, 70 percent of EU cities with populations of more than 250,000 violated even the less stringent phase 1 limits (Wolff and Perry, 2010). In response, Council Directive 2008/50/*EC* abandoned phase 2 but in return required higher reporting standards and introduced stiffer penalties for non-compliance or non-action.

Between 2005 and 2007, more than two thirds of Germany's cities with a population of more than 100,000 were violating the 35-day limit in at least one year. *All* of them were en route to non-attainment of the much stingier phase 2 limits and would have violated the 7-day limit. As a result, they were forced to develop so called “clean air action plans”. The center piece of these plans, in many cases, was the introduction of a low emission zone. The federal government has broadly supported these action plans and categorized vehicles into 4 classes based on the European emission standards, Euro 1 through 4. On one end of the spectrum are “dirty” Euro 1 diesel-powered vehicles without a particle filter. At the other end of the spectrum are gasoline (or even electric) vehicles with a catalytic converter. Drivers who intend to enter an LEZ have to obtain a sticker that has to be placed on the wind shield and indicates the pollution level associated with their vehicle. The cleanest vehicles are issued a green sticker, somewhat dirtier vehicles receive a yellow or red sticker and the dirtiest vehicles are not issued a sticker at all. LEZs tend to be introduced in different stages, where in stage 1 only vehicles without stickers are banned from entering a zone, in stage 2 vehicles with red stickers are also banned, and in stage 3 only vehicles with a green sticker are allowed.[2]

The size of an LEZ varies from city to city, the average zone covers about 44.5% of the city population.[3] LEZs virtually always cover the city center. Since this is where most of the commercial and cultural activities take place, bypassing an LEZ is usually not an option for drivers of dirty vehicles. For example, Fig. 1 shows Berlin's LEZ. It covers about a third of Berlin's 3.5 million inhabitants including most places of political, cultural, and economic significance. Entering a zone without an adequate sticker is sanctioned with a €40 fine and 1 demerit point in the central traffic registry (€80 since May 2014. The central traffic registry is a federal database containing an account with a unique ID for every traffic offender who has committed a major traffic transgression. Demerit points for different offenses accumulate and eventually trigger further sanctions such as driving license suspensions. In order to avoid repercussions, vehicle owners can, of course, retrofit their vehicles, for example by building in filters, so that they fall in a higher class.

My identification strategy is based on the fact that cities have introduced LEZs at different points in time, and have upgraded to stages 2 and 3 at different times. This allows me to disentangle the effect of LEZs from seasonal influences and economy wide trends. What is more, the staggered introduction supplies a natural control group for treatment cities, namely communities which have yet to implement their LEZs. The cities of Berlin, Cologne, and Hannover were the first cities to introduce an LEZ in January 2008; Mannheim, Reutlingen, and Stuttgart followed suit in March 2008; and many

[1] Around the same time, limits and target values for lead (Pb), arsenic (As), cadmium (Cd), and nickel cadmium (Ni) in ambient air, as well as benzene, ozone, and polycyclic aromatic hydrocarbons were also introduced.

[2] There are a few exceptions, for instance for ambulances or military vehicles.

[3] This is a rough estimate as city authorities tend to only provide the geographic size (in km^2) of a zone. In cases in which no information on the covered population was obtainable, the percentage of city lands covered by an LEZ was used for this calculation. 44.5% is therefore likely to be a lower bound. Wolff (2014) provides a comprehensive list of current and future LEZs, their size, and - if available - the number of inhabitants living in a zone.

Fig. 1. Low emission zone in Berlin.
Source: City of Berlin Website: http://www.stadtentwicklung.berlin.de/umwelt/luftqualitaet/umweltzone/de/gebiet.shtml.

more introduced LEZs subsequently. Fig. 2 shows all 82 German cities with 100,000 or more inhabitants (as of 2005); cities which at a given point in time introduced a zone are marked with a triangle. 44 out of 82 cities had introduced an LEZ by the beginning of 2013, others plan to do so in the future.

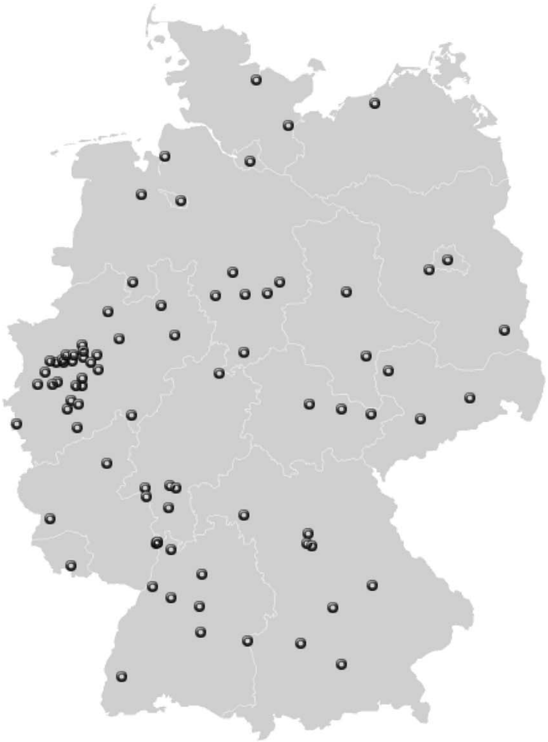
Data and descriptive evidence

Data sources

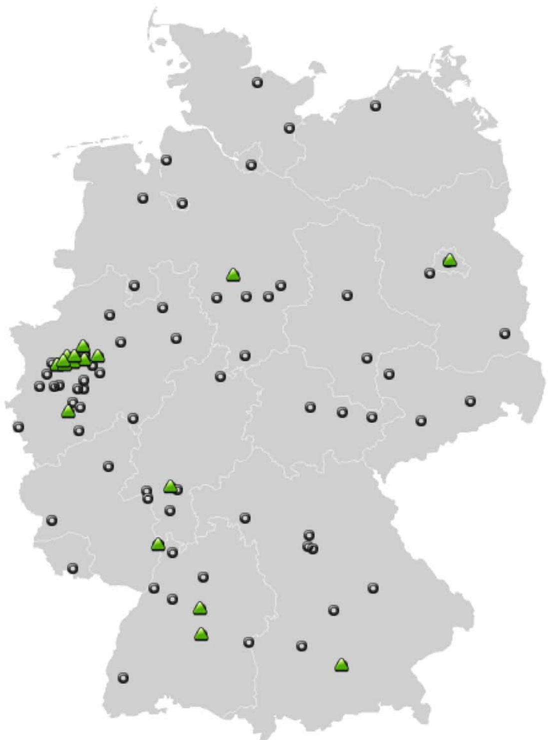
Two main sources of data are used for this study: birth records and pollution measurements. Birth records are available for 2005–2012 and were accessed via remote execution through the Research Data Center of the (German) Statistical Offices of the Laender. Ultimately, the data are based on information that has to be provided by a newborn's parents in order to obtain a birth certificate. By law, births need to be registered with the local civil registry office within 7 days. In order to be issued a birth certificate, parents need to provide a written medical certificate issued by either the delivering hospital or (in the rare case of a home birth) by a certified midwife who was present during the birth of their child. The certificate must contain information about the child's name, the day and time of the birth, the child's birth weight and height, and whether it was a stillbirth or live birth. Parents also need to fill out an application form containing, among other things, their place of residence, dates of birth, nationalities, marital status, the mother's labor force status, the number of previous births in this marriage, the sex of the child, whether it was a twin or multiple birth, and (if applicable) the birth order.

All of this information is entered into a federal registry database. Periodically, the Research Data Center of the German Federal Statistical Office uses these data to create an anonymized and edited data file. Names and exact addresses are erased from the data although the mother's place of residence is not. I use this information on the mother's place of residence to determine whether a newborn child was carried to full term in a city that had an active LEZ. By law every child born in Germany must have a birth certificate, which is also required to be issued a passport and/or ID-card, to obtain health insurance, to request parental allowance, or any kind of government assistance. It is therefore reasonable to assume that the birth data contain information on *all* births in Germany. I focus on three main outcomes: a continuous measure of birth weight in grams, a dummy variable indicating low birth weight ($< 2,500$ g), and a dummy variable indicating a stillbirth. Since small towns and rural areas are unlikely to introduce an LEZ, I focus on births that occurred in cities with populations of at least 100,000 which account for about one third of all births in Germany.

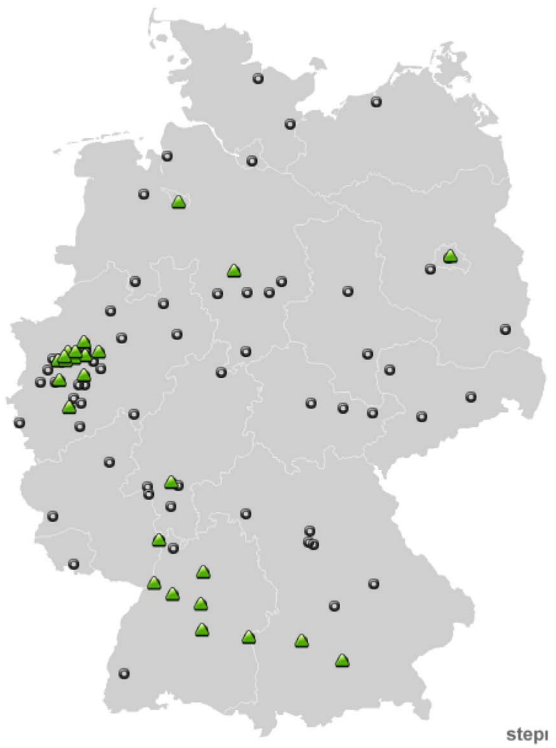
I also obtained daily data on pollution levels in German cities for 2005 to 2012 from the German Federal Environmental Agency. Most communities in my sample have air quality monitors within the city limits that measure PM_{10} concentrations



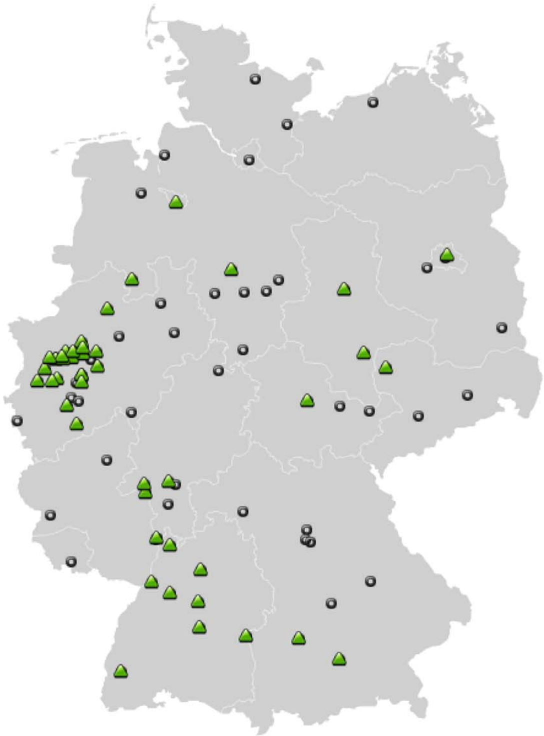
(a) All Cities



(b) Cities with LEZs - 2008



(c) Cities with LEZs - 2009



(d) Cities with LEZs - 2013

Fig. 2. Cities with LEZs over time.
Notes: Maps show all 82 German cities with a population of at least 100,000. Triangles indicate that at the time a low emission zone was in place.

in the air. A different set of monitors also measures the levels of nitrogen dioxide (NO_2), carbon monoxide (CO), and sulfur dioxide (SO_2). Measurement stations differ in terms of their location within a city: a little less than half the stations are "traffic stations", i.e. they are located next to major traffic routes. Obviously pollution measures at traffic stations are likely to be the most affected by LEZs. Therefore, in my analysis I will distinguish between traffic and non-traffic "background" stations. These background stations are located away from busy streets, usually in residential areas and make up about half the stations in my sample. There are also a few "industry" stations in commercial districts which in the analysis were grouped together with background stations.[4] I only use stations that are located within a 30 minute drive of a major (>100, 000 inhabitants) city's center and distinguish between stations based within an LEZ and stations outside an LEZ. The distinction between traffic and background stations and between stations within and outside a zone is useful in shedding light on the roles of spatial heterogeneity and traffic displacement. For example, if drivers took detours and drove around LEZs, the resulting spillovers might defeat the purpose of an LEZ. However, LEZs always cover the city center where most of the economic and cultural life takes place. Thus, one would hope that a zone has negative spillovers, i.e. leads to reductions in pollution not only at stations that are located in the city center but also at those outside of an LEZ. For example, drivers who used to commute into a city center in their dirty cars may have been prompted by an LEZ to either buy a cleaner vehicles or to use public transportation instead. Vehicle registration numbers presented by Wolff (2014) support such a narrative. We will see that the separate analysis of traffic and background stations, and stations within and outside of zones also point to negative spillovers. Unfortunately, spatial spillovers cannot be analyzed in detail for birth outcomes as the mothers' exact addresses are not contained in the birth data.

Pollution measurements are affected by weather conditions. To account for these confounding factors, I have collected data from the National Weather Service ("Deutscher Wetterdienst") and matched the data to pollution monitors. Weather data are collected at stations that are distinct from pollution monitoring stations and also rarer. Hence, some weather stations' measurements are matched to several pollution monitoring stations, at times in different cities. The following control variables are constructed: average daily temperature and its square, daily maximum temperature, daily minimum temperature, one-day lags for average daily temperature and daily maximum temperature, daily average humidity and its square and one-day lag, an interaction between average daily temperature and average daily humidity, daily average precipitation and its square and interaction with the average temperature, daily average wind speed, its lag and an interaction with whether it rained, daily average vapour pressure, and daily average air pressure.[5] The units of measurement and variable means are listed in Tables 1 and 2.

Descriptive statistics

Selection into treatment is an obvious concern with this natural experiment. In order to investigate such selection issues, I have categorized the 82 cities in this study into two groups: "ever adopters" which introduce an LEZ at some point in the 2008 to 2012 time window, and "never adopters" which have not introduced a zone during this time interval. This categorization is further refined by distinguishing between "early adopters" which introduce an LEZ before the end of 2009, and "late adopters" which introduce an LEZ between 2010 and 2012. Table 1 shows that average PM_{10} concentrations in 2005 were slightly higher in cities that eventually introduced an LEZ than in those who refrained from doing so. Ever-adopters also violated the daily PM_{10} limit of 50 $\mu g/m^3$ more frequently than never-adopting cities. Similar differences exist between early and late-adopters (see Table 2). However, these differences in levels do not jeopardize my identification strategy. The validity of a difference-in-differences estimator rests on the common-time-trend assumption, i.e. the assumption that pollution and birth outcome *trends* would have been the same in both treatment and control cities in the absence of treatment.

Panel A1a of Fig. A1 in the appendix plots the daily PM_{10} levels over time for "early-adopters", "late-adopters", and "never adopters". Each point represents a daily average across all stations where dark points refer to stations in early-adopting cities and lighter points to stations in late- and never-adopting cities. The vertical line indicates the day on which the first three cities introduced their LEZs, 1 January 2008. To illustrate underlying trends, I overlay the raw data with a smoothed LOWESS lines, again separately for early-, late-, and never-adopters. Clearly there is a fair amount of seasonality with pollution levels increasing substantially during the winter months.[6] All three groups exhibit strikingly similar time trends. In the pre-treatment period, pollution levels in all three groups show a virtually identical pattern for average PM_{10} (see Panel A1a) and the percentage of stations that exceeded the daily limit (see Panel A1b).

Of course, Fig. A1 masks a lot of heterogeneity across cities. Nonetheless, it provides some evidence that the main identifying, common time trend assumption is reasonable.[7] One might be equally concerned about differential time trends in birth outcomes (birth weight, low birth weight, and stillbirth). Table 1 suggests that compared with non-adopting cities, the average birthweight is about six grams lower in adopting cities. This is a very small difference to begin with. More

[4] Dropping industry stations does not considerably affect the results.

[5] I follow Wolff (2014) in this respect in order to make the results comparable.

[6] Some of this seasonality is due to road gritting. In the main analysis I include temperature controls which should pick up most of this effect. I also drop pollutions measurements from new year's days on which the biggest spikes occur due to fireworks.

[7] In Section "Results" , I also show event study graphs that plot the coefficients for leads and lags of the treatment variable. This more direct test lends further support to the common time trend assumption.

Table 1
Table of means: Ever- vs. Never-adopters.

	Year 2005		Year 2012	
	Never-adopters	Ever-adopters	Never-adopters	Ever-adopters
Pollution outcomes				
Daily PM_{10} level (μg/m^3)	25.56	27.14	20.45	22.50
Percentage exceedance Days	0.061	0.075	0.029	0.046
Daily NO_2 level (μg/m^3)	31.99	35.92	29.46	35.19
Daily SO_2 level (μg/m^3)	4.67	5.18	3.11	2.88
Birth outcomes				
Birthweight (in g)	3334	3328	3330	3320
Birthweight <2500 g	0.072	0.071	0.072	0.072
Stillbirth	0.0035	0.0039	0.0035	0.0039
Weather controls				
Daily average temperature (in °C)	9.83	10.34	9.80	10.32
Daily max temperature (in °C)	14.17	14.5	14.15	14.62
Daily minimum temperature (in °C)	5.69	6.28	5.51	5.96
Daily average vapour pressure (in hpa)	9.99	10.17	9.90	10.19
Daily average cloud cover (in eight)	5.32	5.31	5.53	5.47
Daily average air pressure (in hpa)	994.9	995.82	995.43	995.21
Daily average relative humidity (in %)	77.62	76.93	77.55	77.13
Daily average wind speed (in m/s)	3.34	3.14	3.71	3.34
Daily max wind speed (in m/s)	9.84	9.6	10.52	10.07
Daily average precipitation (in mm)	1.92	2.12	1.86	1.96
Mother and child characteristics				
Mother married	0.668	0.682	0.613	0.630
Mother working	0.423	0.394	0.406	0.401
Mother's age	29.47	29.74	30.27	30.55
Father's age	33.27	33.68	33.91	34.38
Under-age mother	0.010	0.010	0.006	0.006
German mother	0.770	0.720	0.789	0.724
Female	0.488	0.485	0.485	0.486
Twin	0.030	0.030	0.032	0.035
Multi-birth	0.001	0.001	0.001	0.001
City characteristics				
Unemployment rate	13.56	13.39	8.11	8.78
Average age	42.82	42.68	42.82	42.68
% Non-German	0.09	0.12	0.09	0.12
GDP per capita (in €1000)	44.23	42.02	44.23	42.02
% College/vocational degree	0.72	0.72	0.72	0.72
N_1	39,164	57,622	37,641	59,376
N_2	73,387	151,566	78,036	164,276

N_1=Number of station-days with valid PM_{10} measurement; N_2=Births with valid birth weight.
Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender for birth statistics, and German Federal Environmental Agency for pollution measurements. Deutscher Wetterdienst for Weather Controls. Federal Employment Agency for unemployment data, all other city characteristics are from the Zensus 2011.

importantly, Fig. A2a in the appendix suggests that the gap neither widens nor narrows much in the pre-treatment period. In fact, trends in birthweight are virtually identical across all types of cities. For example, average birthweight is significantly lower in December than in other months for all three groups, although the smoothness of the lowess lines somewhat obscures this interesting pattern.[8] By and large birthweight is subject to the same seasonalities and follows the same time trend in early-, late-, and never-adopting cities. This remains true if we examine trends in the incidence of low birthweight (< 2.500 g) and stillbirths as is demonstrated by Figs. A2b and c.

While Figs. A1 and A2 illustrate seasonal patterns and variation in the data, the means of all outcomes and control variables are shown in Tables 1 and 2. Both tables compare the means across cities for the earliest available pre-treatment year, 2005, and the latest available post-treatment year, 2012. There are virtually no economically significant differences across city categories with respect to birth outcomes. For instance, in 2005 the average birthweight in ever-adopting cities is 3328 while the average birthweight in never-adopting cities is 3334. What is more, both mother and infant characteristics

[8] This is consistent with Buckles and Hungerman (2013) who find that average birthweight in the US is about 20 grams lower in the winter months than in spring. The same appears to be true in Germany, although interestingly December rather than January is the month with the lowest average birthweight.

Table 2
Table of means: Early- vs. Late-adopters.

	Year 2005		Year 2012	
	Late-adopters	Early-adopters	Late-adopters	Early-adopters
Pollution outcomes				
Daily PM_{10} level (μg/m^3)	25.83	27.81	22.00	22.79
Percentage exceedance days	0.064	0.080	0.046	0.047
Daily NO_2 level (μg/m^3)	32.74	37.55	33.08	36.34
Daily SO_2 level (μg/m^3)	4.26	5.62	1.89	3.37
Birth outcomes				
Birthweight (in g)	3336	3325	3328	3317
Birthweight <2500 g	0.071	0.071	0.072	0.072
Stillbirth	0.0038	0.0039	0.0036	0.0040
Weather controls				
Daily average temperature (in °C)	10.44	10.28	9.95	10.33
Daily max temperature (in °C)	14.66	14.41	14.29	14.63
Daily minimum temperature (in °C)	6.29	6.28	5.62	6.02
Daily average vapour pressure (in hpa)	10.18	10.16	9.97	10.23
Daily average cloud cover (in eight)	5.28	5.33	5.52	5.45
Daily average air pressure (in hpa)	1002.41	991.66	997.62	989.85
Daily average relative humidity (in %)	76.70	77.07	77.32	77.34
Daily average wind speed (in m/s)	3.07	3.19	3.54	3.47
Daily max wind speed (in m/s)	9.55	9.62	10.28	10.31
Daily average precipitation (in mm)	1.92	2.24	1.84	2.07
Mother and child characteristics				
Mother married	0.645	0.693	0.587	0.642
Mother working	0.485	0.367	0.496	0.374
Mother's age	29.44	29.83	30.14	30.67
Father's age	33.46	33.74	34.02	34.48
Under-age mother	0.012	0.009	0.008	0.006
German mother	0.789	0.700	0.794	0.704
Female	0.481	0.486	0.488	0.486
Twin	0.031	0.030	0.033	0.036
Multi-birth	0.001	0.001	0.001	0.001
City characteristics				
Unemployment rate	13.35	13.41	8.31	8.78
Average age	42.51	42.78	42.51	42.78
% Non-German	0.10	0.14	0.10	0.14
GDP per capita (in €1000)	40.05	43.20	40.05	43.20
% College/vocational degree	0.73	0.70	0.73	0.70
N_1	19,460	38,162	22,309	37,067
N_2	34,095	117,471	36,378	127,898

N_1=Number of station-days with valid PM_{10} measurement; N_2=Births with valid birth weight.
Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender for birth statistics, and German Federal Environmental Agency for pollution measurements. Deutscher Wetterdienst for Weather Controls. Federal Employment Agency for unemployment data, all other city characteristics are from the Zensus 2011.

are fairly balanced across cities. For example, the proportion of under-aged mothers is virtually identical in adopting, non-adopting, and early-adopting and late-adopting cities.[9] This indicates that it is unlikely that less resourceful mothers tend to be located in cities which never introduce an LEZ. Again, this supports the identification strategy at hand. For instance, one might be concerned that other policies, such as spending on public health, might be correlated with both mother characteristics and the propensity to introduce an LEZ, so that the measured effect of an LEZ on infant health would be biased. The similarity in levels and even more importantly in time trends of adopting and non-adopting cities, mitigates such concerns. For example, Tables 1 and 2 also show that out of wedlock births have increased over time. This secular trend can be observed nationwide and is equally pronounced in all types of cities.

Weather indicators also show a striking resemblance across different types of cities. For example, in 2005, average daily temperatures were 10.34, 9.83, 10.28, and 10.44 °C in ever-, never-, early- and late-adopting cities respectively. As would be expected given the balance in weather covariates, including them into the regression analysis has little effect on the point estimates, which again supports the notion that treatment and control cities are quite similar. Panel 3a of Fig. 3 plots the

[9] Due to the large sample size, small means differences often show up as statistically significant when tested explicitly. So, while there are no economically meaningful differences in levels, the pre-treatment characteristics are not always perfectly balanced from a purely statistical point of view.

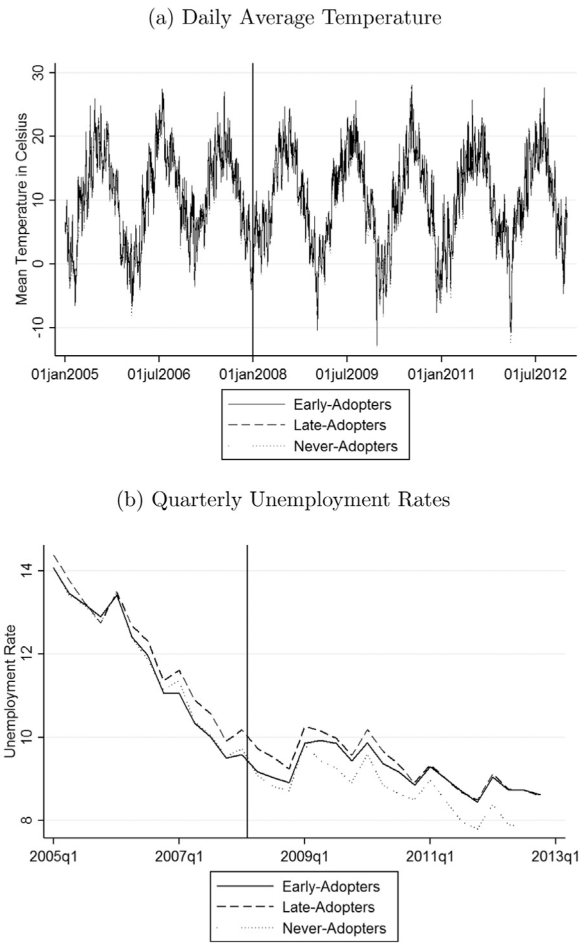


Fig. 3. Covariates over time.
Notes: Raw data based on city averages. The vertical line indicates the first month in which a city introduced an LEZ, January 2008. *Data Source*: Deutscher Wetterdienst and Federal Employment Agency.

daily average temperature over time for different city types. Temperatures follow identical time trends in all three types of city. While this apparent lack of differences is no definite proof for identical time-trends in unobservable characteristics, it is pointing that way. Similarly, there is little evidence that better-off or fast-growing cities were more likely to introduce LEZs (or introduced them earlier). Panel 3b of Fig. 3 plots quarterly city unemployment rates separately by city status.[10] Again, no substantial differences in either levels or time trends stand out. Tables 1 and 2 also compare other city characteristics, such as the average age of the population, percentage of the population without a German passport, GDP per capita, and the share of skilled workers. These indicators stem from the Zensus 2011 and therefore have no time variation. Yet again, the similarity in levels is striking and if the unemployment trends are any indication, little difference in trends can be expected.

A simple "before" and "after" comparison based on the raw data of Tables 1 and 2 is also revealing in that it can yield a rough estimate of the expected treatment effects. An examination of birth outcomes shows little sign of an effect of the policy on birth. Neither does the incidence of stillbirths and light births change much over time. All groups show some reduction in PM_{10} pollution levels and the probability that on a given day an average PM_{10} concentration of 50 μg/m^3 was exceeded. However, there is little indication that these reductions are more pronounced in adopting cities than in non-adopting cities. Of course, the level of aggregation might mask heterogeneity. By and large, all of the above graphical comparisons and the comparison of pre-treatment means are comforting in that they suggest that never adopters are indeed a viable control group for ever adopters, and differences for early adopters can be compared to either differences in late adopters or never adopters.

[10] Technically, unemployment is measured at the county level ("Landkreise und Kreisfreie Staedte"). In the vast majority of cases city and county boundaries are identical. In cases in which they do not exactly overlap, the corresponding county unemployment rate was used.

Methods

A difference-in-differences design is applied to assess the effect of the introduction of an LEZ on pollution and birth outcomes. In order to implement this estimator, the following equation is estimated using OLS:

$$y_{ijt} = \alpha + \beta\, Active_{jt} + \gamma X_{it} + \sum_{i=1}^{814} \theta_i\, Station_i + \sum_{t=1}^{95} \tau_t Time_t + \varepsilon_{ijt} \quad (1)$$

where y_{ijt} is the daily average PM_{10} level or a dummy indicator for a violation of the 50 μg/m^3 daily PM_{10} limit on day t at station i, located in city j. In the evaluation of birth outcomes y_{ijt} denotes birth weight (in g) and indicators for low birth weight ($< 2,500$ g) and stillbirth for child i born at time t in city j. $Active_{jt}$ is a dummy indicator that indicates whether at time t an LEZ was active in city j. For the evaluation of birth outcomes, this main explanatory variable is shifted by the typical gestational time of 9 month, i.e. it takes on a value of 1 if in the 9 months prior to a birth, a low emission zone has been active in the mother's city of residence. X_{it} denotes a set of control variables. In the case of pollution outcomes, these are the weather covariates described in the previous section. In the case of birth outcomes, X_{it} denotes the age of the mother (in years), and three dummies that equal one if the mother is married, in employment, and a German citizen, respectively; I add 3 dummies indicating the sex of the child, whether the child has a twin, and whether the child has multiple twins. I also include station fixed effects in the pollution analysis and city fixed effects in the analysis of birth outcomes, as well as year-month time fixed effects for each (but one) of the 96 months between January 2005 and December 2012. The main coefficient of interest is β which yields the effect of LEZ adoption on the outcomes of interest. This main specification is estimated in a couple of different variations. For instance, I replace the $Active_{jt}$-Dummy by a set of dummies for stage 1, 2, and 3 zones. Many cities were about to upgrade to stage 3 on January 2013, but I only have data available until the end of 2012. Therefore, the sample of stage 3 cities is very small. For that reason, the dummies for stage 2 and 3 were aggregated into a single dummy for most specifications. In order to account for anticipation and/or belated compliance by drivers as well as gestational time, the $Active_{jt}$-Dummy will be replaced with a set of leads and lags in some specifications.[11]

Lastly, inference in a difference-in-differences setting with grouped data is subject to the so called Moulton problem (Moulton, 1986). Errors of observations stemming from the same city might conceivably be correlated. Such intra-group correlation will not bias the coefficient of interest, β, but will lead to inconsistent standard errors. I counter this potential problem by adjusting all standard errors for clustering at the city level (Liang and Zeger, 1986). Bertrand et al. (2004) also note that serial correlation in the error terms constitute another potential threat to the validity of standard errors in a difference-in-differences setting. At the same time, Hansen (2007) points out that with a sufficient number of groups, serial correlation ceases to pose a problem. With 82 clusters, I am well above Angrist and Pischke's (2008) rule of thumb of at least 42 clusters to achieve reliable inference.

Results

LEZs and fine particulate pollution

Table 3 provides the results of the main regression as described by Eq. (1) with station fixed effects and year-month fixed effects included in all specifications. Panel A shows that the introduction of an LEZ reduced daily PM_{10} levels by about 0.67 μg/m^3. This result is robust to the inclusion of controls for local weather conditions. Given a daily average PM_{10} concentration of about 26.5 μg/m^3 in German cities, this translates into roughly a 2.5 percent reduction. Column (3) repeats the analysis using only data from each city's pollution monitor which - on a given day - reported the highest pollution levels in a city. This is the same method that is used to determine whether city's are in compliance with the daily and annual limits set in Council Directive 1999/30/EC. While the focus on each city's highest pollution measurement considerably reduces the sample size and thus inflates standard errors, the point estimate is almost identical in magnitude.

Columns (4) through (6) repeat the analysis but only use stations that are located in close proximity to major traffic links. LEZs have had a more pronounced effect at these stations, reducing PM_{10} levels by about 1.3 μg/m^3. This result is quite robust to the inclusion of weather covariates and a focus on cities' traffic stations reporting the highest pollution levels. Of course, PM_{10} levels tended to be higher at traffic stations to begin with. Still, the effect corresponds to about a 4 percent reduction in fine particulate pollution. For background stations, on the other hand, no statistically significant effect of LEZs can be detected, although the negative sign of the coefficient might be interpreted as weak evidence for reductions in PM_{10} away from major roads. Panel B of Table 3 analyzes the effect of an LEZ on a measure of daily compliance with pollution limits. The binary outcome here is whether on a given day, an average PM_{10} concentration of 50 μg/m^3 was exceeded. My results indicate that an LEZ reduces the probability of such a violation by about 1.1 percentage points. Given a baseline probability of

[11] I also experimented with the inclusion of linear city-specific year-month time trends. While the results for birth outcomes are robust to the inclusion of such trends, my coefficients for daily pollution tend to shrink substantially. However, as is evident from Fig. A1, assuming linearity in time trends is unlikely to do justice to the pattern of underlying time trends and might do more harm than good. For a thorough discussion of the pros and cons of including city-specific time trends, please see Wolfers (2006).

Table 3
Effect of LEZs on air pollution: Diff-in-Diff estimates - All Cities.

	All stations			Traffic stations			Background stations		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Daily PM_{10} levels									
2005 Station Mean		26.510			31.516			23.468	
(SD)		(14.575)			(15.806)			(12.855)	
Active	−0.665***	−0.546**	−0.688	−1.299***	−1.137***	−1.090*	−0.320	−0.216	0.226
	(0.240)	(0.229)	(0.642)	(0.375)	(0.357)	(0.646)	(0.236)	(0.254)	(0.325)
Observations	813,641	813,641	199,865	341,779	341,779	160,848	471,862	471,862	166,885
R-squared	0.279	0.445	0.389	0.253	0.439	0.434	0.238	0.404	0.382
Panel B: Percentage exceedance days									
2005 Station mean		0.069			0.114			0.042	
(SD)		(0.254)			(0.318)			(0.200)	
Active	−0.011***	−0.009***	−0.015	−0.018***	−0.015***	−0.017*	−0.007***	−0.006**	−0.005
	(0.003)	(0.003)	(0.009)	(0.005)	(0.005)	(0.010)	(0.002)	(0.002)	(0.004)
Observations	813,641	813,641	199,865	341,779	341,779	160,848	471,862	471,862	166,885
R-squared	0.129	0.190	0.212	0.144	0.216	0.224	0.108	0.164	0.173
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Max measurement only	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1). Dummy for Active Zone is equal to one if at the time of measurement, an LEZs has been active. Dependent variables are daily PM_{10} levels and a dummy for whether on a given day a PM_{10} concentration of 50 μg/m^3 was exceeded.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

7 percent, this translates into about a 16 percent reduction. This result is robust to the inclusion of weather controls, more pronounced at traffic stations and can to a lesser degree also be observed at background stations.

Despite the evidence presented in Section "Data and descriptive evidence", one might still be concerned about including observations from cities that never introduce a zone ("never-adopters") in the control group. For this reason, Table 4 repeats the analysis using only pollution measurements from stations that are located in cities which eventually adopt an LEZ ("ever-adopters"). As shown in Section "Data and descriptive evidence", there is little indication that never-adopters constitute an inappropriate control group. Nonetheless, it is reassuring that limiting the sample to ever-adopters does not considerably change the results. Column (1) of Panel A suggests that LEZs reduced PM_{10} levels by about 0.72 μg/m^3. Only using data from a city's highest PM_{10} polluting station does not considerably change the point estimate. Columns (4) through (9) again indicate that LEZs led to larger pollution reductions close to busy roads than in residential areas, but that nonetheless borderline significant reductions in fine particulate pollution were achieved at background stations. Panel B shows that the focus on ever-adopting cities does not change the point estimates for exceedance days very much, although the standard errors are slightly larger.

Table 5 illustrates that the stage of a zone matters quite a bit. The higher "treatment intensity" of stage 2 or 3 zones which ban not only cars without sticker but also those with just a red sticker are associated with PM_{10} reductions of about 1 μg/m^3 while stage 1 zones are associated with reductions of about half that size. The effects of stage 2 or 3 zones are even larger when I focus on the cities' highest-polluting stations or on traffic stations. In the most extreme case of column (6), a restrictive zone reduces PM_{10} levels at the cities highest-polluting traffic station by about 1.8 μg/m^3, about a 5.7 percent reduction. A regular LEZ that only bans vehicles without stickers reduces pollution levels by a statistically insignificant 0.68 μg/m^3 relative to not having a zone. However, Panel B of Table 5 indicates that more restrictive zones do not further reduce propensities for exceeding daily pollution limits, except for the highest-polluting stations. In other words, for most stations going from a stage 1 zone to a stage 2 or 3 zone has little effect on the probability that a daily limit is exceeded, but

Table 4
Effect of LEZs on air pollution: Diff-in-Diff estimates - ever-adopters only.

	All stations			Traffic stations			Background stations		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Daily PM_{10} levels									
2005 Station mean		27.148			32.907			23.783	
(SD)		(14.774)			(16.193)			(12.723)	
Active	−0.715**	−0.427	−0.778	−1.016**	−0.819	−1.114	−0.624**	−0.415	−0.320
	(0.289)	(0.330)	(0.720)	(0.477)	(0.379)	(0.581)	(0.280)	(0.361)	(0.403)
Observations	492,311	492,311	106,366	203,651	203,651	91,696	288,660	288,660	96,837
R-squared	0.289	0.489	0.435	0.250	0.475	0.472	0.250	0.454	0.432
Panel B: Percentage exceedance days									
2005 Station mean		0.075			0.132			0.041	
(SD)		(0.263)			(0.338)			(0.199)	
Active	−0.010***	−0.006	−0.015	−0.014**	−0.008	−0.016	−0.008***	−0.007*	−0.010**
	(0.003)	(0.004)	(0.010)	(0.006)	(0.006)	(0.011)	(0.003)	(0.004)	(0.005)
Observations	492,311	492,311	106,366	203,651	203,651	91,696	288,660	288,660	96,837
R-squared	0.131	0.204	0.238	0.147	0.237	0.251	0.109	0.174	0.181
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Max measurement only	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1). Dummy for Active Zone is equal to one if at the time of measurement, an LEZs has been active. Dependent variables are daily PM_{10} levels and a dummy for whether on a given day a PM_{10} concentration of 50 μg/m^3 was exceeded.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

more restrictive zones are effective in reducing peak pollution. For example, column (1) in Panel B indicates that a stage 1 LEZ reduces the probability of exceeding the daily limit of 50 μg/m^3 by about 1.1 percentage points; a stage 2 or 3 zone has virtually the same effect and reduces this probability by 1.2 percentage points. Column (3), on the other hand, shows that in terms of reducing the probability that *any* station in a city exceeds the daily limits, a stage 2 or 3 zone is much more effective than a stage 1 zone. Finally, a focus on ever-adopting cities does not considerably change the results (see Table 6) except that the results for background stations are slightly more suggestive of negative spill-over effects.

I further investigate spatial heterogeneity by categorizing stations into those that are located strictly within an LEZ and those stations that are located outside of an LEZ. I discard measurements from stations within a current or future LEZ and those from cities that never introduce an LEZ and re-run Eq. (1). This boils down to a comparison of stations outside an active LEZ with stations outside an LEZ that has not been activated yet. The difference-in-difference estimate in column (1) of Panel B in Table 7 indicates that a stage 1 zone reduces PM_{10} levels in the outskirts of a city by about 0.42 μg/m^3 while a stage 2 or 3 zone is associated with reductions of about −1.4 μg/m^3. The magnitude of coefficients is similar those in column (7) of Table 6, as it should be since most stations outside LEZs are in fact background stations. These estimates are, however, not particularly stable and decrease in size and statistical significance once weather controls are included. The same is true when the binary indicator of a violation of the daily limit on average fine particulate pollution is used as an outcome. It is, however, comforting that all point estimates remain negative. Hence, while these results may not constitute conclusive evidence of negative spill-overs, at the very least positive spill-overs are highly unlikely. In other words, there is little indication that drivers circumvent zones by driving LEZ-ineligible vehicles through outer boroughs, thus more than off-setting gains in air quality.

By and large, Tables 3 through 7 indicate that LEZs moderately reduced pollution levels. Spillovers from parts of a city that have busy roads into residential areas are non-positive and an active LEZ appears to slightly reduces the propensity for

Table 5
Effect of LEZs on air pollution: Diff-in-Diff estimates - all cities.

	All stations			Traffic stations			Background stations		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Daily PM_{10} levels									
2005 Station mean		26.510			31.516			23.468	
(SD)		(14.575)			(15.806)			(12.855)	
Stage 2 or 3 zone	−0.994***	−0.797**	−1.744**	−1.478***	−1.255**	−1.803**	−0.597*	−0.415	0.199
	(0.306)	(0.331)	(0.781)	(0.459)	(0.486)	(0.831)	(0.338)	(0.353)	(0.535)
Stage 1 zone	−0.481**	−0.404*	−0.097	−1.169***	−1.051***	−0.679	−0.188	−0.122	0.242
	(0.240)	(0.229)	(0.673)	(0.380)	(0.340)	(0.640)	(0.240)	(0.273)	(0.327)
Observations	813,641	813,641	199,865	341,779	341,779	160,848	471,862	471,862	166,885
R-squared	0.279	0.445	0.389	0.253	0.439	0.435	0.238	0.404	0.382
Panel B: Percentage exceedance days									
2005 Station Mean		0.069			0.114			0.042	
(SD)		(0.254)			(0.318)			(0.200)	
Stage 2 or 3 zone	−0.012**	−0.010**	−0.028**	−0.017**	−0.014**	−0.026**	−0.008*	−0.006*	−0.005
	(0.004)	(0.004)	(0.011)	(0.007)	(0.007)	(0.012)	(0.004)	(0.003)	(0.003)
Stage 1 zone	−0.011***	−0.009***	−0.007	−0.018***	−0.016***	−0.012	−0.007***	−0.006*	−0.005
	(0.003)	(0.003)	(0.009)	(0.005)	(0.005)	(0.009)	(0.003)	(0.003)	(0.005)
Observations	813,641	813,641	199,865	341,779	341,779	160,848	471,862	471,862	166,885
R-squared	0.129	0.190	0.212	0.144	0.216	0.224	0.108	0.164	0.173
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Max measurement only	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1). Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker was active at the time of measurement. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker was active at the time of measurement. Dependent variables are daily PM_{10} levels and a dummy for whether on a given day a PM_{10} concentration of 50 $\mu g/m^3$ was exceeded.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

violating daily limits for fine particulate pollution even at background stations. The effect size for traffic stations is similar to Wolff's (2014) short-term analysis of LEZs. Wolff (2014) runs several specifications, in some instances excluding cities with very low or very high baseline pollution. His specification that uses all cities with 100,000 without restrictions on their baseline PM_{10} levels most resembles my setup. He finds statistically insignificant reductions in fine particulate pollution at traffic stations of 3.2 percent. The point estimate in this paper (see Table 3) imply effects of in a very similar range. But, due to a sample that covers more cities over a longer time period, my estimates are more precise. What is more, my dataset allows for an analysis of more restrictive stage 2 and 3 LEZs which turn out to be even more effective in reducing pollution levels by almost 8 percent at cities' highest polluting stations.

LEZs and birth outcomes

A natural next step is a "second stage" analysis of infant health, where the effect of a zone introduction on fine particulate pollution constitutes the first stage, and the effect of PM_{10} reductions on birth outcomes can be considered the second stage. However, LEZs may very well have also reduced other forms of pollution.[12] Hence, I evaluate birth outcomes in a reduced

[12] Section "Discussion and robustness checks" will show that NO_2 concentrations were also affected by LEZs. What is more, a study by the Leibniz Institute suggests that the LEZ in the city of Leipzig has reduced the amount of carbon black and ultra-particulate matter (between 0.06 and 700

Table 6
Effect of LEZs on air pollution: Diff-in-Diff estimates - ever-adopters only.

	All stations			Traffic stations			Background stations		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Daily PM_{10} levels									
2005 Station mean		27.148			32.907			23.783	
(SD)		(14.774)			(16.193)			(12.723)	
Stage 2 or 3 zone	−1.068**	−0.783	−2.042*	−1.040	−1.121	−2.223*	−0.971**	−0.274	−0.160
	(0.406)	(0.512)	(1.120)	(0.628)	(0.911)	(1.099)	(0.643)	(0.688)	(0.623)
Stage 1 zone	−0.624**	−0.338	−0.428	−1.007**	−0.729	−1.044	−0.553*	−0.443	−0.370
	(0.286)	(0.261)	(0.729)	(0.456)	(0.641)	(0.805)	(0.728)	(0.370)	(0.430)
Observations	492,311	492,311	106,366	203,651	203,651	91,696	288,660	288,660	96,837
R-squared	0.289	0.489	0.435	0.250	0.475	0.473	0.250	0.454	0.432
Panel B: Percentage exceedance days									
2005 Station mean		0.075			0.132			0.041	
(SD)		(0.263)			(0.338)			(0.199)	
Stage 2 or 3 zone	−0.010*	−0.004	−0.032**	−0.011	−0.001	−0.027*	−0.008*	−0.005	−0.009*
	(0.005)	(0.006)	(0.014)	(0.008)	(0.010)	(0.015)	(0.005)	(0.005)	(0.005)
Stage 1 zone	−0.010**	−0.007	−0.010	−0.014**	−0.010	−0.013	−0.009**	−0.007	−0.010*
	(0.004)	(0.004)	(0.010)	(0.006)	(0.006)	(0.010)	(0.003)	(0.004)	(0.005)
Observations	492,311	492,311	106,366	203,651	203,651	91,696	288,660	288,660	96,837
R-squared	0.131	0.204	0.238	0.147	0.237	0.251	0.109	0.174	0.181
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Max measurement only	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1). Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker was active at the time of measurement. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker was active at the time of measurement. Dependent variables are daily PM_{10} levels and a dummy for whether on a given day a PM_{10} concentration of 50 $\mu g/m^3$ was exceeded.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
***Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

form setting, i.e. I regress birth outcomes on LEZ adoption as described by Eq. (1). The resulting coefficient thus reflects *all* infant health benefits associated with LEZs, not just those operating through PM_{10} reductions.

Column (1) of Table 8 suggests that the introduction of an LEZ has virtually no effect on birthweight. This remains true when I add controls for maternal and child characteristics in column (2). The corresponding point estimate suggests that an LEZ is associated with an increase in birthweight of 0.26 grams. While not statistically significant, this is a relatively precise estimate so that large undetected effects can be ruled out. This remains true when we focus on births to mother who live in ever-adopting cities. Column (3) indicates that an LEZ raises the average birth weight in such a city by less than 3 grams, although the null hypothesis of no effect can again not be rejected at any reasonable level of significance. By way of comparison, a recent study based on birth registry data for Swedish siblings finds that light smoking during pregnancy is associated with a decrease in birthweight by 162 g (Juarez and Merlo, 2013). My findings hold up regardless of whether maternal or child characteristics are controlled for. Neither do effects on birthweight kick in when LEZs become more restrictive as Panel B shows. There is also little evidence for any association between the prevalence of low birthweight ($<2,500$ g) and the presence of an LEZ.[13]

(*footnote continued*)
nanometers) by a third (Rasch et al., 2013).

[13] I also repeated my analysis for the incidence of very low birthweight ($<1,500$ g) and found no effect. Neither is there much in the way of an effect on birthweight, low birthweight and stillbirths when days or trimesters of "exposure" to an LEZ are used as explanatory variables. Results for both

Table 7
Effect of LEZs on air pollution: Diff-in-Diff estimates for outside-zone stations.

	PM_{10}		Exceedance	
	(1)	(2)	(3)	(4)
2005 Station mean	23.09		0.049	
(SD)	(14.76)		(0.217)	
Panel A: Effect of any zone				
Active	−0.616*	−0.268	−0.005	−0.001
	(0.317)	(0.310)	(0.003)	(0.004)
Panel B: Effect by zone restrictiveness				
Stage 2 or 3 Zone				−1.395**
−0.539	−0.008			−0.002
	(0.512)	(0.733)	(0.006)	(0.006)
Stage 1 zone	−0.418	−0.199	−0.004	−0.001
	(0.321)	(0.289)	(0.004)	(0.004)
Observations	221,872	221,872	221,872	221,872
R-squared	0.259	0.454	0.113	0.177
City fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Weather controls	No	Yes	No	Yes
Ever-adopters only	Yes	Yes	Yes	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1) but only uses stations located outside the area covered by a current or future LEZ. Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker was active at the time of measurement. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker was active at the time of measurement. Dependent variables are daily PM_{10} levels and a dummy for whether on a given day a PM_{10} concentration of 50 μg/m^3 was exceeded.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

Although the coefficients in Table 8 are statistically insignificant, they are very precisely estimated. Currie and Walker (2011), for example, find that an electronic toll collection system in New Jersey and Pennsylvania, E-ZPass, substantially reduced the incidence of low birth weight. Their analysis is based on standard errors about 3 times the size of those in columns (4)–(6) of Table 8. Obviously, an electronic toll collection system is a very different intervention than an LEZ. Moreover, mine are average effects that cannot distinguish between infants who were carried to term within an LEZ and those outside. The results for pollution suggest that health benefits might be more pronounced in city centers. The smaller the share of children covered by the zones, the more is the treatment effect biased towards zero. A tendency of young families to move to suburbs could exacerbate this problem. Nonetheless, the comparative size and precision of the estimates in this paper suggests that large effects on birth weight are unlikely to go undetected. For instance, the results in column (6) indicate that the upper bound of a 95% confidence interval would correspond to about a 0.23 percentage point reduction in the incidence of low birth weight, which would roughly translate into a 3.2 percent decrease. Even this upper bound pales in comparison to Currie and Walker's (2011) finding of a 12 percent reduction due to E-ZPass. On the other hand, my study is consistent with Currie and Neidell (2005) who in their study on California do not find statistically significant reductions in pollution to translate into statistically significant gains in birth weight even though both the initial level and subsequent reductions in fine particulate pollution in California in the 1990s were much higher than in Germany in the 2000s. Hence, a likely explanation for the lack of a statistically significant effect of LEZs on birth weight might be that at current pollution levels the marginal return of additional pollution reductions induced by LEZs might be too small to make a dent in average birth weight that is large enough to be detected, even in a sample of over 1.8 millions births.[14]

I do, however, find a statistically significant relationship at the extensive margin, i.e. between the prevalence of stillbirths and the introduction of an LEZ. Stillbirths account for about 0.37 percent of births, so the point estimate of −0.0006 (see column (9)) translates into roughly a 16 percent reduction in the incidence of stillbirths. With about 160,000 births per year

(*footnote continued*)
robustness checks are available from the author upon request.

[14] It should also be noted that much of the existing evidence of the effect of health-enhancing policy interventions stems from the US which differs from Germany in terms of (universal) health care provisions to expecting mothers.

Table 8
Effect of LEZs on Birthoutcomes: Diff-in-Diff estimates.

	Birthweight			Low birthweight			Stillbirth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
2005 Mean		3329.95			0.071			0.0037	
(SD)		(581.44)			(0.252)			(0.061)	
Panel A: Effect of any zone									
Active	−1.4359	0.2575	3.3998	0.0010	0.0003	−0.0007	−0.0000	−0.0000	−0.0006***
	(2.2115)	(1.8943)	(2.5385)	(0.0008)	(0.0008)	(0.0008)	(0.0002)	(0.0002)	(0.0002)
Panel B: Effect by zone restrictiveness									
Stage 2 or 3 zone	−0.0143	−0.2647	1.8066	−0.0005	−0.0005	−0.0013*	0.0001	−0.0003	−0.0000
	(2.2963)	(1.9757)	(2.3781)	(0.0007)	(0.0007)	(0.0007)	(0.0002)	(0.0002)	(0.0002)
Stage 1 zone	−0.7943	−0.1800	1.3036	0.0007	0.0003	−0.0003	−0.0004**	0.0001	−0.0005**
	(2.1095)	(1.7819)	(1.7963)	(0.0010)	(0.0008)	(0.0008)	(0.0002)	(0.0002)	(0.0002)
Observations	1,854,106	1,852,420	1,251,782	1,854,106	1,852,420	1,251782	1,867,733	1,865,948	1,261,329
R-squared	0.0025	0.1289	0.1307	0.0008	0.1481	0.1512	0.0001	0.0006	0.0007
City fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Ever-adopters only	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1) where the outcomes are birthweight (in g) and two dummy indicators for low birth weight ($< 2,500$ g) and stillbirths. Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker has been active for at least 9 months prior to a birth. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker has been active for at least 9 months prior to a birth. Controls include mother's age, labor market status, nationality, and marital status. Controls for the child's sex, and indicators for twin or multiple births are also included.
Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

in cities that have adopted LEZs, this would translate into an annual gain of 96 infant lives. While not negligible, this is a considerably smaller gain than what was achieved, for example, by the desulfurization of power plants in Germany in the late 80s/early 90s. Luechinger (2014) estimates that this policy measure saved between 826 and 1460 infant lives per year.[15] However, my result on stillbirths should be taken with a grain of salt. I only detect this effect when focusing on ever-adopting cities and Panel B counter-intuitively suggests that less restrictive zones are slightly more effective in avoiding stillbirths than more restrictive zones.

Discussion and robustness checks

Overall, the results for fine particulate pollution and birth weight indicate that LEZs have moderately decreased daily PM_{10} levels and even more so the probability that a daily limit is exceeded. However, these reductions appear to be too small to result in statistically significant improvements in birth outcomes. Table 9 shows that the results for fine particulate pollution also hold up if a monthly aggregation rather than daily data is used. A monthly aggregation also has the advantage that the coefficients for exceedance days have a more natural interpretation. For example, column (3) suggests that an LEZ reduces the number of monthly exceedance days by 0.262 or about 3 days a year. Panel B suggests that more restrictive zones are an even better tool in reducing exceedance days. Since the limit on the number of exceedance day is the piece of regulation most cities are struggling with, LEZs appear to be an effective tool in achieving compliance with air quality regulations.

More importantly, one might still be concerned about whether the moderate reductions I find in fine particulate pollution can actually be attributed to the policy at hand or are just random noise. In order to further address such concerns, I evaluate the effect of LEZs on two other pollutants: nitrogen dioxide (NO_2) and sulphur dioxide (SO_2). Motor vehicles are not a major source of the latter. NO_2, on the other hand, is caused by burning fossil fuel and by combustion engines. That is, even though LEZs are primarily designed to reduce particulate pollution they still should reduce NO_2 concentrations but have

[15] It should, however, be noted the Luechinger (2014) evaluates infant mortality (within 1 year of birth) which is distinct from my evaluation of stillbirths.

Table 9
Effect of LEZs on air pollution: Diff-in-Diff Estimates - Monthly Data.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Effect of Any Zone												
	All Stations				Traffic Stations				Background Stations			
	PM_{10}		Exceed. Days		PM_{10}		Exceed. Days		PM_{10}		Exceed. Days	
2005 Station mean	27.426		2.166		33.015		3.750		23.901		1.168	
(SD)	(8.216)		(3.341)		(8.329)		(4.378)		(5.852)		(1.883)	
Active	-0.599^{**}	-0.569	-0.262^{***}	-0.232^{*}	-1.395^{***}	-1.032^{*}	-0.538^{***}	-0.495^{**}	-0.112	-0.371	-0.087	-0.093
	(0.273)	(0.345)	(0.090)	(0.118)	(0.389)	(0.523)	(0.151)	(0.203)	(0.281)	(0.326)	(0.071)	(0.082)
Observations	20,551	11,703	20,551	11,703	8,777	5,108	8,777	5,108	11,774	6,595	11,774	6,595
R-squared	0.847	0.853	0.688	0.694	0.848	0.842	0.742	0.743	0.815	0.830	0.668	0.685
Ever-adopters only	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Panel B: Effect by zone restrictiveness												
	All stations				Traffic stations				Background Stations			
	PM_{10}		Exceed. days		PM_{10}		Exceed. Days		PM_{10}		Exceed. Days	
Stage 2 or 3 zone	-1.069^{***}	-1.284^{**}	-0.307^{**}	-0.339^{*}	-1.860^{***}	-1.495^{*}	-0.587^{***}	-0.588^{*}	-0.484	-1.074^{**}	-0.096	-0.137
	(0.332)	(0.478)	(0.143)	(0.190)	(0.466)	(0.773)	(0.208)	(0.309)	(0.359)	(0.434)	(0.124)	(0.137)
Stage 1 zone	-0.307	-0.401	-0.233^{**}	-0.207^{*}	-1.094^{***}	-0.919^{*}	-0.506^{***}	-0.473^{**}	0.116	-0.208	-0.082	-0.083
	(0.274)	(0.332)	(0.092)	(0.122)	(0.395)	(0.495)	(0.155)	(0.200)	(0.301)	(0.331)	(0.080)	(0.093)
Observations	20,551	11,703	20,551	11,703	8,777	5,108	8,777	5,108	11,774	6,595	11,774	6,595
R-squared	0.847	0.854	0.688	0.694	0.848	0.843	0.742	0.743	0.815	0.831	0.668	0.685
Ever-adopters only	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level.
Regression results correspond to Eq. (1). Dummy for Active Zone is equal to one if at the time of measurement, an LEZs has been active. Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker was active at the time of measurement. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker was active at the time of measurement. Dependent variables are monthly average PM_{10} levels and the number of days in a month on which a PM_{10} concentration of 50 $\mu g/m^3$ was exceeded. All regressions include station and year-month fixed effects and controls for weather conditions.
Data Source: German Federal Environmental Agency, monthly pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

little impact on SO_2 levels. Table 10 suggests that this is indeed the case. For example, column (1) in Panel B suggests that a stage 2 or 3 LEZ reduced NO_2 levels by 1.16 $\mu g/m^3$ in cities that adopted LEZs. This translates into a reduction of about 3.4 percent and is robust to weather controls and to limiting the sample to stations in cities that eventually introduce a zone. On the other hand, there is no consistent evidence for LEZ-induced reductions in SO_2. For the most part columns (5) through (8) of Table 10 suggest a weak negative correlation between SO_2 levels and LEZ adoption, but the adoption of a stage 2 or 3 zone is actually positively associated with SO_2 when only "ever-adopters" are used in the analysis.[16]

My difference-in-differences design relies on the stable-unit-treatment-value assumption (SUTVA). Therefore, a particular concern in the context of this study is spatial leakage. Cities are connected to one another and it is conceivable that a city's air quality is improved by an LEZ in a nearby city that prompts drivers who commute between the two cities to upgrade their cars. Indeed, Wolff (2014) presents a series of figures that show that the adoption of green sticker vehicles is negatively associated with the distance to a nearby LEZ. As a result my analysis might overestimate the effectiveness of LEZs, in particular for late-adopting cities which benefit not only from their own but also from nearby LEZs. However, Table 11 mitigates such concerns. It shows the results of a regression of the form of Eq. (1) in which the presence of an LEZ in a nearby city is the main explanatory variable of fine particulate pollution in a city that has not (yet) launched its own LEZ.[17]

[16] I also analyzed the effect of carbon monoxide (CO) levels, another pollutant produced by motor vehicles and find negative (but statistically insignificant) effects. However, there is only a very limited number of monitors that record CO concentrations. The results are available from the author upon request.

[17] In order to distinguish between spillovers from other cities and own-LEZ effects, a city's observations are dropped from the sample as soon as a city itself introduces a zone.

Table 10
Effect of LEZs on air pollution: other pollutants.

	NO_2				SO_2			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2005 Station mean	34.186		35.916		4.947		5.177	
(SD)	(20.060)		(20.534)		(4.658)		(4.890)	
Panel A: Effect of any zone								
Active	−0.469*	−0.543*	−0.477*	−0.523	−0.378	−0.364	−0.145	−0.133
	(0.274)	(0.282)	(0.259)	(0.360)	(0.275)	(0.278)	(0.271)	(0.275)
Panel B: Effect by zone restrictiveness								
Stage 2 or 3 zone	−1.161*	−1.165	−1.143**	−0.957	0.005	0.011	0.789*	0.792*
	(0.671)	(0.718)	(0.485)	(0.755)	(0.302)	(0.295)	(0.395)	(0.401)
Stage 1 zone	−0.031	−0.148	−0.276	−0.367	−0.553	−0.558	−0.546	−0.529
	(0.290)	(0.360)	(0.360)	(0.419)	(0.631)	(0.611)	(0.356)	(0.362)
Observations	728,835	728,835	400,964	400,964	306,114	306,114	160,092	160,092
R-squared	0.629	0.686	0.630	0.707	0.317	0.355	0.349	0.383
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Ever-adopters only	No	No	Yes	Yes	No	No	Yes	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level. Regression results correspond to Eq. (1) with NO_2 and SO_2 levels (in μg/m^3) as outcomes. Dummy for Active Zone is equal to one if at the time of measurement, an LEZs has been active. Dummy for Stage 1 is equal to one if a zone prohibiting vehicles with no sticker was active at the time of measurement. Dummy for Stage 2 or 3 is equal to one if a zone prohibiting vehicles with no or just a red sticker was active at the time of measurement. All regressions include station and year-month fixed effects and controls for weather conditions. *Data Source*: German Federal Environmental Agency, daily pollution measurements (2005–2012).
***Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

The introduction of an LEZ in a nearby city does not appear to translate into reductions in fine particulate pollution. An active zone within 50 km of a city's center is associated with a small (but insignificant) increase in PM_{10} pollution. Extending the radius to 100 km and focussing on traffic or background stations does not significantly alter the results. In fact none of the coefficients in Panel A of Table 11 are statistically significant. Panel B indicates that air quality improvements due to upgrades to stage 2 or 3 LEZs also do not feed through to nearby cities. Taken together, these results indicate that any violations of the SUTVA due to spillovers are unlikely to induce substantial bias into the main results.[18]

The validity of the difference-in-differences estimator for the analysis of birth outcomes also rests on the assumption that there is no differential sorting that masks an effect of LEZ introduction on newborns' health. This assumption would be violated if mothers with specific characteristics were enticed by the introduction of a zone to move into adopting cities. A more formal test of such sorting and avoidance behavior can be conducted by regressing mother characteristics on LEZ status:

$$MotherChar_{ijt} = \alpha^{MC} + \beta^{MC} Active_{jt} + \sum_{j=1}^{81} \theta_j^{MC} City_j + \sum_{t=1}^{95} \tau_t^{MC} Time_t + \eta_{it} \tag{2}$$

where $MotherChar_{ijt}$ are mother or parent characteristics, namely the mother's age, a dummy indicating that the mother is younger than 18 years, a dummy indicating whether the mother was in employment prior to giving birth, an indicator for whether the mother is a German citizen, an indicator for whether either parent is a German citizen, a dummy for whether the mother is married, and a dummy for recently wedded mothers that is equal to one if the mother got married less than 9 months prior to giving birth. The results for this OLS regression are displayed in Table 12. Despite the large sample size, no significant effect can be found for any of these characteristics. There is no indication that any particular group of mothers sorts into or out of cities that have introduced an LEZ.

Finally, another concern with my identification strategy relates to timing. For instance, it is conceivable that drivers anticipated the widely announced introduction of an LEZ in their city. As a result, they may have retrofitted their cars months before the policy went into place. In the same vein, some drivers may have responded with a delay. Anecdotal evidence suggests that officials in some cities did not impose fines on LEZ violations in the first month of introduction in

[18] I also experimented with a continuous distance measure. Again there was no meaningful relationship between changes in the distance to an LEZ and changes in fine particulate pollution. The results of this specification are available from the author on request.

Table 11
Spatial leakage: effect of an LEZ in a nearby city.

	PM_{10}			Exceedance days		
	All (1)	Traffic (2)	Background (3)	All (4)	Traffic (5)	Background (6)
2005 Station Mean	26.510	31.516	23.468	0.069	0.114	0.042
(SD)	(14.575)	(15.806)	(12.855)	(0.254)	(0.318)	(0.200)
Panel A: Effect of any zone						
Active zone within 50 km	0.478	0.189	0.422	0.003	0.189	−0.001
	(0.363)	(0.360)	(0.455)	(0.004)	(0.360)	(0.004)
Active zone within 100 km	0.246	−0.166	0.108	0.003	0.001	−0.003
	(0.311)	(0.381)	(0.344)	(0.004)	(0.005)	(0.003)
Panel B: Effect by zone restrictiveness						
Stage 1 Zone within 50 km	0.512	0.077	0.558	0.004	0.077	0.000
	(0.367)	(0.476)	(0.435)	(0.005)	(0.476)	(0.004)
Stage 2 or 3 zone within 50 km	0.428	0.349	0.222	0.001	0.349	−0.003
	(0.475)	(0.419)	(0.598)	(0.005)	(0.419)	(0.005)
Stage 1 zone within 100 km	0.454	0.057	0.351	0.007*	0.005	−0.001
	(0.298)	(0.403)	(0.327)	(0.004)	(0.006)	(0.003)
Stage 2 or 3 zone within 100 km	−0.039	−0.454	−0.253	0.000	−0.004	−0.007**
	(0.422)	(0.513)	(0.443)	(0.004)	(0.005)	(0.003)
Observations	601,512	257,321	344,191	601,512	257,321	344,191
R-squared	0.441	0.437	0.398	0.191	0.214	0.167
Station fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Standard errors (in parentheses) account for clustering at the city level. Results stem from regressions of pollution levels at different types of pollution monitors on indicators for the presence of an LEZ in close proximity. Dummy for Active Zone is equal to one if at the time of measurement, an LEZ was active within a 50 km (100 km) radius but not the city itself. Dummy for Stage 1 is equal to one if at the time of measurement a zone prohibiting vehicles with no sticker was active within a 50 km (100 km) radius but not the city itself. Dummy for Stage 2 or 3 is equal to one if at the time of measurement a zone prohibiting vehicles with no or just a red sticker was active within a 50 km (100 km) radius but not the city itself. Observations for cities in which at a given time an LEZ was active are not used. All regressions include station and year-month fixed effects and controls for weather conditions.
Data Source: German Federal Environmental Agency, daily pollution measurements (2005–2012).
*** Indicate significance at the 1%-level.
** Indicate significance at the 5%-level.
* Indicate significance at the 10%-level.

Table 12
Design validity - Regression of maternal and child characteristics on zone-dummy.

Variables	(1) Mother's age	(2) Teen birth	(3) Mother working	(4) Mother German	(6) Parents German	(5) Married	(7) Recently married
Active zone	0.019	0.000	0.004	−0.003	0.000	0.002	−0.001
	(0.026)	(0.000)	(0.043)	(0.002)	(0.004)	(0.002)	(0.001)
Observations	1,867,733	1,867,733	1,865,948	1,867,733	1,867,733	1,867,733	1,210,269
R-squared	0.027	0.001	0.073	0.037	0.034	0.063	0.002

Notes: ***/**/* indicate significance at the 1%/5%/10%-level. Standard errors (in parentheses) account for clustering at the city level.
Each coefficients stems from a regression of a mother or child characteristic on an LEZ dummy, a set of city dummies and a set of year-month dummies. Mother's Age is a continuous variable, Teen Birth is a dummy indicating a birth to a mother younger than 18 years, Mother Working is a dummy indicating that the mother is in employment, Mother German is a dummy indicating that the mother is a German citizen, Parents German is a dummy indicating that either the father or the mother is a German citizen, Married is a dummy indicator for a married mother, Recently Married is an indicator for a mother who got married less than 9 months prior to the birth (conditional on the mother being married).
Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender (2005–2012).

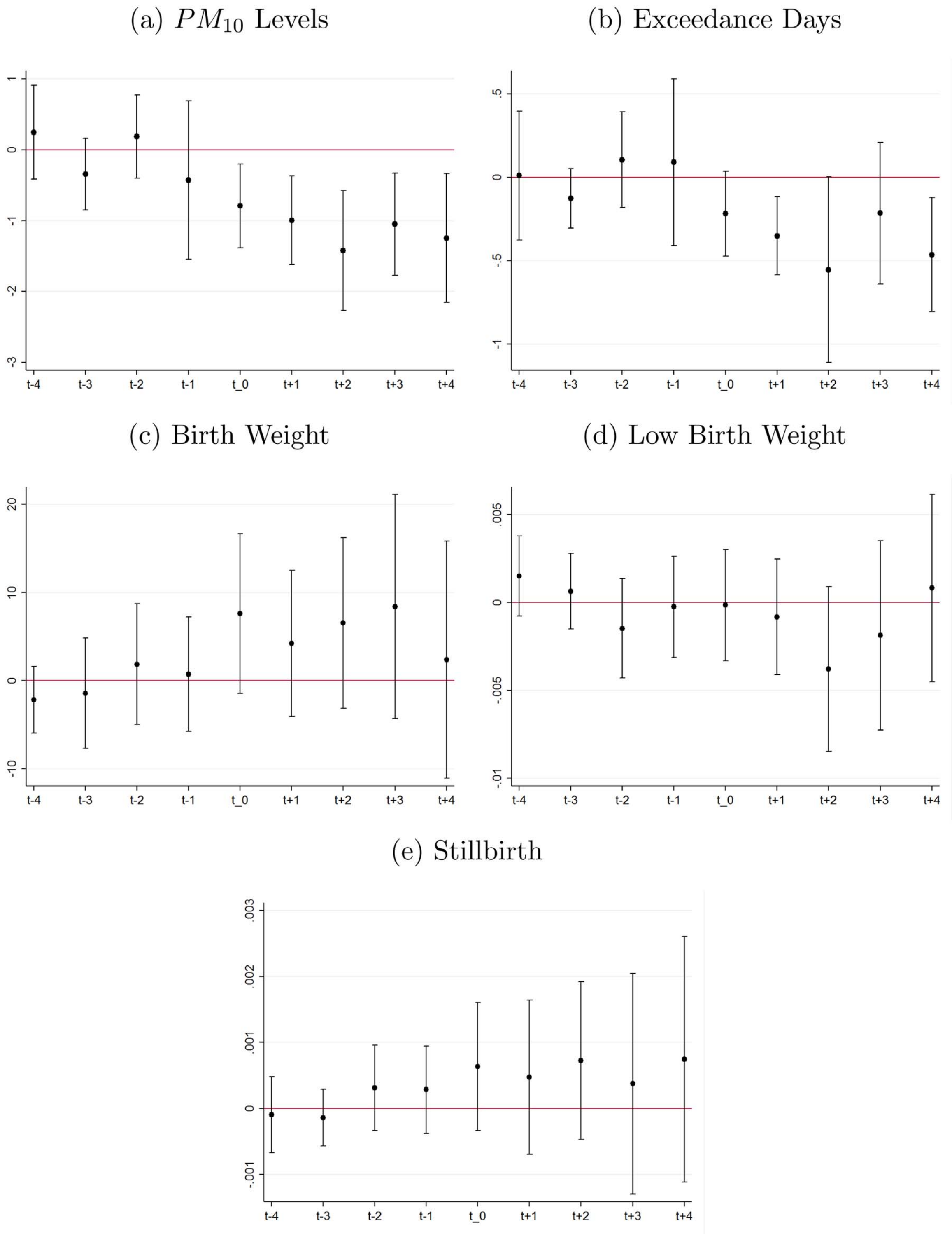


Fig. 4. Leads and lags model.

Notes: Graph displays coefficients and 95% confidence intervals for 6-months leads and lags of the Stage 2/3 - Dummy, as specified in Eq. (3). Outcomes are monthly levels of PM_{10} measured in μg/m^3, days per month on which PM_{10} concentration exceeds 50 μg/m^3, birthweight in grams, the incidence of light births (< 2, 500 g), and the incidence of stillbirths. Reference period is t-5, i.e. more than two years prior to introduction of a Stage 2/3 Zone.

Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender for birth statistics (2005–2012), and German Federal Environmental Agency for pollution measurements (2005–2012).

order to provide drivers with a transition period. I test both hypotheses by applying a model of 6 months leads and lags of the following form:

$$y_{ijt} = \alpha + \sum_{t=-4}^{t=4} \beta_t Active_{jt} + \gamma X_{ijt} + \sum_{j=1}^{81} \theta_j City_j + \sum_{t=1}^{95} \tau_t Time_t + \varepsilon_{ijt} \tag{3}$$

In other words, I run the main regression using 6-months leads and lags of either the $Active_{jt}$ variable or the zone stage indicators, instead of just 1 indicator that switches on and remains on for a given treatment city. The reference period here are time periods more than 24 months prior to the adoption of a zone. Fig. 4 plots the coefficients that correspond to the regression of Eq. (3), along with 95% confidence intervals. There is little evidence for anticipatory effects in either pollution or birth outcomes. In fact, pollution levels in the 6 months leading up to a policy change do not differ much from those in the 6 months following the policy change. Most pollution reductions then appear to materialize within 1-2 years of the policy change after which they level off.[19]

There is remains little support for significant long-run effects on birth outcomes. Average birth weight seems to increase slightly after the introduction of an LEZ. However, these point estimates are accompanied by large confidence intervals. A meta-analysis of the relationship between birth weight and ambient air pollution (Stieb et al., 2012) suggests that on average it takes a decrease in PM_{10} of about 20 μg/m^3 to achieve improvements of 16.8 g which is roughly the upper bound of the confidence intervals in Fig. 4c. LEZs clearly deliver more moderate pollution reductions. Health benefits in line with the higher bounds of the confidence intervals would be implausibly large. Fig. 4, therefore, supports the notion put forward in Section "LEZs and birth outcomes" that the impact of LEZs on infant health might be too small to be precisely estimated even with a sample size of over 1.8 million births.[20]

Conclusion

LEZs are one of the most aggressive policy measures enacted to reduce air pollution and to promote public health. LEZs introduce an outright ban on vehicles that do not meet certain environmental standards. This paper evaluates their effects on air quality and infant health. LEZs are associated with substantial costs, related to enforcement, expenditure on retro-fitting vehicles, earlier than planned replacements of new vehicles, and time lost in complying with new rules. Wolff (2014) estimates that the upgrading of the German vehicle fleet in response to LEZ introductions by itself cost more than 1 billion dollars between 2008 and 2010, although he points out that these upgrades also entail benefits for the owners of these vehicles. What is more, the pro-business Institute for Retail Research estimates that the introduction of an LEZ is associated with a 7 percent decrease in the number of customers to stores located within a city center (Lindstaedt, 2009). Of course, these revenue losses might be offset by revenue gains in other parts of a city.

This study quantifies the benefits of LEZs in terms of improvements in air pollution and infant health. The staggered introduction of LEZs - at different points in time in different cities throughout Germany - provides a series of credible natural experiments to base my results on. These results are robust to changes in the specification and the choice of control group. I find that the adoption of an LEZ is associated with moderate reductions in PM_{10} levels and that the size of these reductions depends on the type of LEZ. A stage 1 zone that only bans the dirtiest vehicles is associated with decreases in PM_{10} levels of 1.5 to 2.5 percent while more restrictive stage 2 or 3 LEZs cut PM_{10} levels by 3 to 4 percent. Reductions are even more pronounced in close proximity to major traffic routes and at stations with particularly high pollution levels. LEZs also help to reduce the number of instances in which an average daily PM_{10} concentration of 50 μg/m^3 is exceeded by about 3 days a year, thus substantially helping cities to comply with the most stringent part of the EU's regulation of particulate matter. My results also suggest that while pollution reductions were more pronounced in areas covered by an LEZ, surrounding areas may also have experienced reduction but to a smaller degree. At the very least, there is no indication that due to drivers circumventing LEZs, pollution levels have increased in surrounding areas thus offsetting improvements in the city centers.

However, these observed improvements in air quality appear to be too small to translate into substantial improvements in infant health outcomes, such as birth weight. Of course, there are other channels through which health and human capital accumulation may have been positively affected by the LEZs' air quality improvements. For instance, reductions in air pollution in general, and lower fine particulate concentrations in particular, improve productivity (Graff Zivin and Neidell, 2012; Chang et al., 2016). Currie et al. (2009) link air pollution (carbon monoxide) to school attendance and Lavy et al. (2015) show that higher ambient air pollution can reduce test scores. A future avenue for research is an evaluation of the effect of LEZs on hospital utilization for respiratory diseases such as asthma attacks. What is more, the estimates in this paper can be considered lower bounds. Control cities that do not introduce LEZs often have alternative clean air action plans in place which are aimed at expanded public transportation, improving traffic flow, etc. Wolff (2014) shows that these measures

[19] The last dummy ($t + 4$) yields the long-run differential effect for after more than 24 months of adoption.

[20] Note that the standard errors in Table 8 are much smaller than those implied by the confidence intervals in Fig. 4. Inflated standard errors are a common feature in lead-lag specifications (see for example Autor (2003)), partly because such a set-up distributes the variation of a single treatment dummy across multiple indicators, leaving in particular the leads with little variation.

have little bearing on pollution levels. Nonetheless, it should be stressed that the estimates in this study reflect the relative merit of LEZs compared with other actions intended to cut emissions, not compared with utter inaction. Hence, while LEZs are certainly no silver bullet and might indeed fall short of policy makers' high expectations, this study suggests that there is some merit to including them into a policy mix. A continued evaluation, e.g. of the now widely discussed introduction of "blue sticker" LEZs that is aimed at nitrogen oxide (NO_x) emitting cars, is therefore in order.

Appendix A

See Figs. A1 and A2.

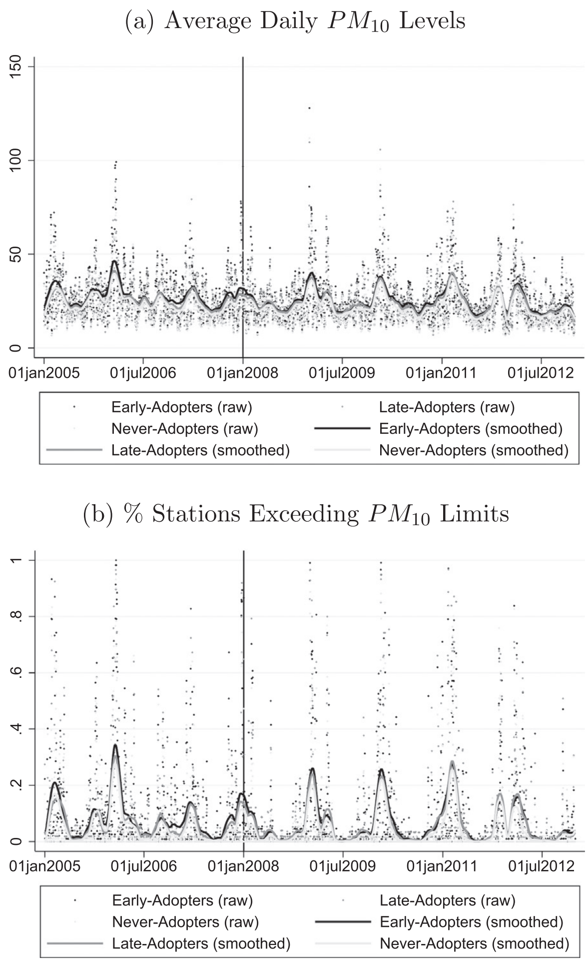


Fig. A1. Pollution over time.
Notes: Each dot shows daily average PM_{10} concentration by city-type. Smoothed LOWESS (locally weighted scatterplot smoothing) regression lines are superimposed to highlight trends and seasonality. For the lines, a tricube weighting function, as developed by Cleveland (1979), was applied with a bandwith of 0.03. The vertical line indicates the first month in which a city introduced an LEZ, January 2008. *Data Source*: German Federal Environmental Agency, pollution measurements (2005–2012).

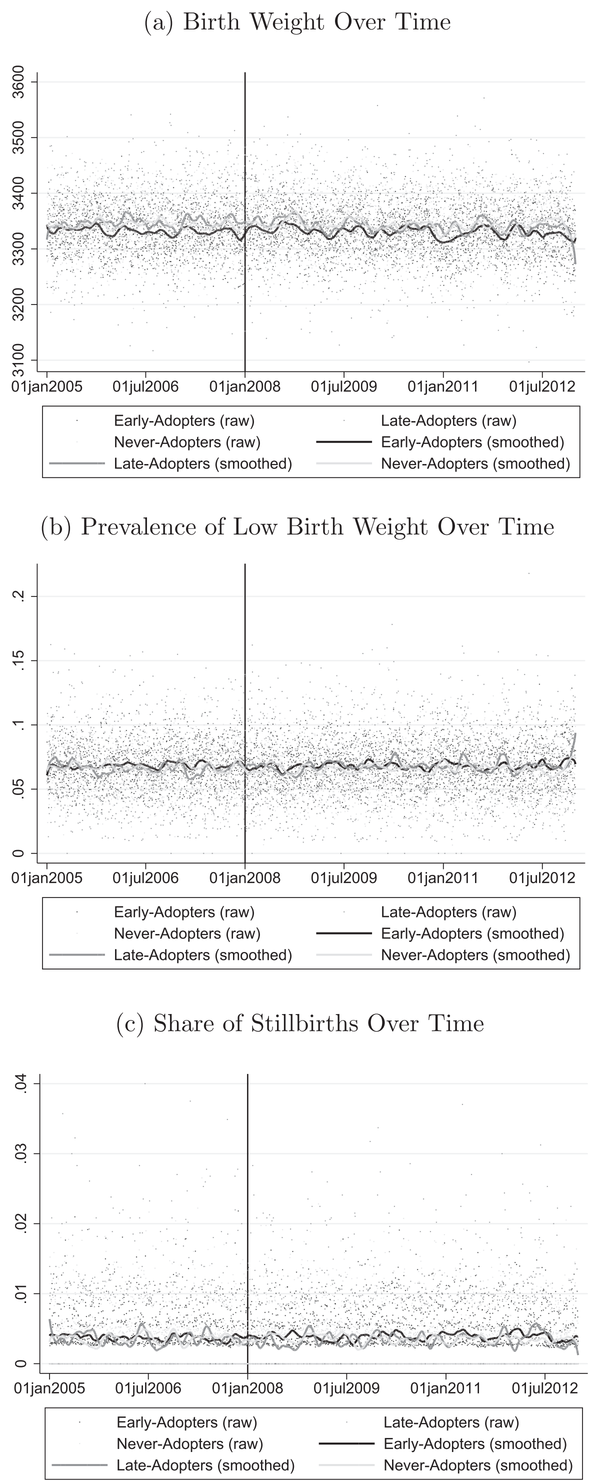


Fig. A2. Birth outcomes over time.
Notes: Each dot shows daily average birth outcomes (i.e. average birthweight, percentage low birthweight, and percentage stillbirths) by city-type. Smoothed LOWESS (locally weighted scatterplot smoothing) regression lines are superimposed to highlight trends and seasonality. For the lines, a tricube weighting function, as developed by Cleveland (1979), was applied with a bandwith of 0.03. The vertical line indicates the first month in which a city introduced an LEZ, January 2008.
Data Source: Research Data Centers of the German Federal Statistical Office and the Statistical Offices of the Laender (2005–2012).

References

Almond, D., Chay, K.Y., Lee, D.S., 2005. The costs of low birth weight. Q. J. Econ. 120, 1031–1083.
Angrist, J.D., Pischke, J.-S., 2008. Mostly Harmless Econometrics: an Empiricist's Companion. Princeton University Press.
Autor, D.H., 2003. Outsourcing at will: the contribution of unjust dismissal doctrine to the growth of employment outsourcing. J. Labor Econ. 21, 1–42.
Behrman, J.R., Rosenzweig, M.R., 2004. Returns to Birthweight. Rev. Econ. Stat. 86, 586–601.
Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? Q. J. Econ. 119, 249–275.
Black, S.E., Devereux, P.J., Salvanes, K.G., 2007. From the cradle to the labor market? The effect of birth weight on adult outcomes. Q. J. Econ. 122, 409–439.
Buckles, K.S., Hungerman, D.M., 2013. Season of birth and later outcomes: old questions, new answers. Rev. Econ. Stat. 95, 711–724.
Chang, T., Graff Zivin, J., Gross, T., Neidell, M., 2016. Particulate pollution and the productivity of pear packers. Am. Econ. J. Econ. Policy 8, 141–169.
Chay, K.Y., Greenstone, M., 2003a. Air quality, infant mortality, and the clean air act of 1970. Working Paper 10053, National Bureau of Economic Research.
Chay, K.Y., Greenstone, M., 2003b. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. Q. J. Econ. 118, 1121–1167.
Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. 74, 829–836.
Currie, J., 2011. Inequality at birth: some causes and consequences. Am. Econ. Rev. 101, 1–22.
Currie, J., Neidell, M., 2005. Air pollution and infant health: What can we learn from California's recent experience?. Q. J. Econ. 120, 1003–1030.
Currie, J., Neidell, M., Schmieder, J.F., 2009. Air pollution and infant health: lessons from New Jersey. J. Health Econ. 28, 688–703.
Currie, J., Walker, R., 2011. Traffic congestion and infant health: evidence from E-ZPass. Am. Econ. J. Appl. Econ. 3, 65–90.
Cyrys, J., Peters, A., Soentgen, J., Wichmann, H.-E., 2014. Low emission zones reduce PM10 mass concentrations and diesel soot in German cities. J. Air Waste Manag. Assoc. 64, 481–487.
Davis, L.W., 2008. The effect of driving restrictions on air quality in Mexico City. J. Political Econ. 116, 38–81.
Friedrich, A., 2006. Barriers and Bottlenecks to Implementing Integration. GermanEnvironmental Agency (Umweltbundesamt) - Institutional Arrangements for Policy Integration.
Glinianaia, S.V., Rankin, J., Bell, R., Pless-Mulloli, T., Howel, D., 2004. Particulate air pollution and fetal health: a systematic review of the epidemiologic evidence. Epidemiology 15, 36–45.
Graff Zivin, J., Neidell, M., 2012. The impact of pollution on worker productivity. Am. Econ. Rev. 102, 3652–3673.
Graff Zivin, J., Neidell, M., 2013. Environment, health, and human capital. J. Econ. Lit. 51, 689–730.
Ha, E.-H., Hong, Y.-C., Lee, B.-E., Woo, B.-H., Schwartz, J., Christiani, D.C., 2001. Is air pollution a risk factor for low birth weight in Seoul? Epidemiology 12, 643–648.
Hansen, C.B., 2007. Asymptotic properties of a robust variance matrix estimator for panel data when is large. J. Econ. 141, 597–620.
Juarez, S.P., Merlo, J., 2013. Revisiting the effect of maternal smoking during pregnancy on offspring birthweight: a quasi-experimental sibling analysis in Sweden. PLoS One 8, e61734.
Laberer, C., Niedermeier, M., 2009. Wirksamkeit von Umweltzonen. ADAC Untersuchung.
Lavy, V., Ebenstein, A., Roth, S., 2015. The long run economic consequences of high-stakes examinations: evidence from transitory variation in pollution. American Economic Journal: Applied Economics, forthcoming.
Leape, J., 2006. The London congestion charge. J. Econ. Perspect. 20, 157–176.
Liang, K.-Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.
Lindstaedt, B., 2009. Einzelhandel, Kundenverkehr und Umweltzonen. Institut fuer Handelsforschung (IfH) Seminar Presentation.
Luechinger, S., 2014. Air pollution and infant mortality: a natural experiment from power plant desulfurization. J. Health Econ. 37, 219–231.
Morfeld, P.D.P., Spallek, M., Groneberg, D., 2011. Zur Wirksamkeit von Umweltzonen: design einer Studie zur Ermittlung der Schadstoffkonzentrationsaenderung fuer Staubpartikel (PM10) und andere Groessen durch Einfuehrung von Umweltzonen in 20 deutschen Staedten. Zentralblatt für Arbeitsmedizin, Arbeitsschutz und Ergon. 61, 148–165.
Moulton, B.R., 1986. Random group effects and the precision of regression estimates. J. Econ. 32, 385–397.
Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L., 2008. Short-, medium-, and long-term consequences of poor infant health an analysis using siblings and twins. J. Human. Resour. 43, 88–138.
Pope, C.A., 1989. Respiratory disease associated with community air pollution and a steel mill, Utah Valley. Am. J. Public Health 79, 623–628.
Pope, C.A., Schwartz, J., Ransom, M.R., 1992. Daily mortality and PM10 pollution in Utah valley. Arch. Environ. Health Int. J. 47, 211–217.
Rasch, F., Birmili, W., Weinhold, K., Nordmann, S., Sonntag, A., Spindler, G., Hermann, H., Wiedensohler, A., Löschau, G., 2013. Significant reduction of ambient black carbon and particle number in Leipzig as a result of the low emission zone. Gefahrstoffe- Reinhalt. der Luf 11–12, 483–489.
Reh, W., 2009. Wirksamkeit von Umweltzonen. BUND Hintergrund.
Stieb, D.M., Chen, L., Eshoul, M., Judek, S., 2012. Ambient air pollution, birth weight and preterm birth: a systematic review and meta-analysis. Environ. Res. 117, 100–111.
Wang, X., Ding, H., Ryan, L., Xu, X., 1997. Association between air pollution and low birth weight: a community-based study. Environ. Health Perspect. 105, 514–520.
Watkiss, P., Pye, S., Holland, M., 2005. CAFE CBA: baseline analysis 2000 to 2020. Report to the European Commission DG Environment.
Wolfers, J., 2006. Did unilateral divorce laws raise divorce rates? A reconciliation and new results. Am. Econ. Rev. 96, 1802–1820.
Wolff, H., 2014. Keep your clunker in the suburb: low-emission zones and adoption of green vehicles. Econ. J. 124, 481–512.
Wolff, H., Perry, L., 2010. Policy monitor trends in clean air legislation in Europe: particulate matter and low emission zones. Rev. Environ. Econ. Policy 4, 293–308.