Transportation Research Part D 92 (2021) 102731

Contents lists available at ScienceDirect

Transportation Research Part D

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

Do light rail systems reduce traffic externalities? Empirical evidence from mid-size european cities☆

Xavier Fageda *

University of Barcelona & GIM-IREA, Spain

ARTICLE INFO

Keywords:
Mid-size cities
Congestion
Pollution
Light rails
Europe

ABSTRACT

This paper examines the impact of urban light rail systems on congestion, travel time and pollution. Drawing on data from mid-size European cities, I estimate the impact of supply changes for the entire sample and applied a differences-in-differences analysis to a sample of cities that did not have rail systems in the initial year of the considered period. I find evidence that an increase in the supply of rail transport leads to less congestion, less travel time and less pollution. Furthermore, cities with a new rail system have on average 7% less congestion, 1% less travel time and 3% less pollution than cities with no rail systems. The results suggest that light rail systems have been successful in containing the negative externalities associated with car traffic in mid-size European cities.

1. Introduction

Investments in public transportation are generally seen as an essential part of urban strategies to reduce the negative externalities generated by cars. Two of the most damaging externalities for the economy and society are congestion and pollution. Urban congestion results in traffic jams that affect drivers and pedestrians, who have to put up with increasing levels of gridlock, noise and pollution. The economic costs of road congestion are huge, due to loss of time. For example, a recent study by the European Commission (2019) revealed that congestion due to road transport in all European Union countries costs €271 billion. Furthermore, polluting emissions are the main cause of death of 3.3 million people a year in the world (more than AIDS, malaria and the flu together) and car traffic is one of the main causes (Lelieveld et al., 2015).

In recent years, many mid-size cities in Europe have invested in light rail systems, including light metros, light rails and trams. Light rail systems have some advantages over metro systems and buses (Knowles and Ferbrache, 2016). Light rails can operate on steeper gradients and tighter curves than heavy rails, and they can be run at lower costs. They can carry up to three times more passengers than buses and provide faster and more reliable services. However, light rails are more costly than buses including bus rapid transit systems.

In this paper, I examine the impact of light rail systems on congestion, travel time and pollution. I use panel data from a sample of mid-size cities in Europe for the period 2008 to 2019 for congestion and travel time data and 2008 to 2016 for pollution data. I exploit the great variability in supply in recent years, including the expansion of existing systems and the launch of new systems. I estimate the impact of supply changes for the entire sample and applied a differences-in-differences (DiD) analysis to a sample of cities that did not have rail systems in the initial year of the period (baseline year). The DiD analysis examines the impact of light rails over time. The

☆ I acknowledge financial support from Generalitat de Catalunya (2017SGR644), and RecerCaixa (2017ACUP00276).
* Corresponding author at: Department of Applied Economics, University of Barcelona, Av. Diagonal 690, 08034 Barcelona, Spain.
E-mail address: xfageda@ub.edu.

https://doi.org/10.1016/j.trd.2021.102731

Available online 12 February 2021

indicator of congestion is based on the excess travel time in relation to free flow conditions, while the variable of travel time refers to the average travel time for a 20-minute trip. Finally, the indicator of pollution is based on estimates of particles with a diameter of 2.5 µm or less (PM 2.5). According to the World Health Organization, PM 2.5 is one of the most damaging pollutants to health.

Aside from impacts on modal share, light rail investments may have relevant effects on employment and population growth, the relocation of activities and property values. The cost of rail projects usually amounts to hundreds of millions of euros. Bearing this in mind, an analysis of the contribution of light rails to the reduction of negative externalities generated by car traffic is a key component for evaluating their social desirability. In fact, the primary purpose of light rails is usually to reduce such negative externalities.

We may expect a substantial reduction in the negative externalities associated with car traffic if changes in light rail system supply leads to a relevant shift from cars to trains. Such a shift will be dependent on the extent of transit time reductions and lower costs for travelers due to the new/expanded system. Note that light rails may not be as effective in reducing the negative externalities associated with car traffic if the modal shift that they imply is from buses to rails. In an analysis for Australia, Mulley et al. (2017) show that rails are not more environmentally friendly than buses in all contexts. Furthermore, rail investments may spur employment and population growth in the cities that benefit from the investments, which could compensate at least partially the substitution effect from cars to trains.

Empirical literature about the impact of public transportation on congestion and travel time is relatively scarce and the evidence is not conclusive. Some studies identified relevant effects and others did not. The results were not significant in studies that exploited variability across cities in the United States (Baum-Snow and Kahn, 2005; Beaudoin and Lawell, 2018) and the United Kingdom (Lee and Senior, 2013). In contrast, studies that exploited variation within an urban area generally found clearer evidence that public transportation reduces road congestion or car use. This was the case of studies by Baum-Snow and Kahn (2000) on selected US cities, and Zhang et al. (2017) and Yang et al. (2018) on the metro of Beijing, although the results were less positive for the metro of Copenhagen (Vuk, 2005). Furthermore, some studies exploited labor strikes in public transportation and found a substantial increase in road congestion in Los Angeles (Anderson, 2014), Rotterdam (Adler & Van Ommeren, 2016) and German cities (Bauernschuster et al., 2017) in the period in which labor strikes stopped the supply of public transport.

Mixed results have been found in studies of how public transportation affects pollution that exploit the variation between cities. Beaudoin and Lawell (2017) did not find significant effects for the United States, while a study by Gendron-Carrier et al. (2018) that focused on metro openings around the world revealed significant drops in particulate concentrations. Gu et al. (2019) determined a negative statistical relationship between indicators of the quality and quantity of rail systems and the quality of air in large cities in China, but did not find a similar relationship for buses.

Mixed results were also found in studies of how public transportation affects pollution that exploit variation within a city, depending on the type of pollutant considered. Chen and Whalley (2012) found that opening one metro line in Taipei significantly reduced carbon monoxide and nitrogen oxides, while little effect was found for ozone or particular matters. Li et al. (2019) revealed an improvement in air quality in Beijing after metro expansions. Finally, Zheng et al. (2019) examined the effects of opening a new metro system in Changsha (China) and identified a substantial reduction in carbon monoxide pollution but no relevant effects for particulate matter and ozone.

In a review of literature on the role of public transport investments in reducing congestion and pollution, Beaudoin et al. (2015) concluded that extrapolating the effects of one city's transit investment has little capacity to forecast the impact of a potential transit investment in another city. Indeed, studies on the United States generally do not find significant effects, particularly when they focus on the intensive margin (marginal supply increases). The results of studies outside the United States are generally more positive, particularly when they focus on the extensive margin (big supply increases). This difference could be associated with data provided by Mayer and Trevien (2017) showing higher public transport use in Europe than in the United States.

In any case, it seems clear that the estimated impacts of public transportation are specific to the context in which they are analyzed. Hence, it is interesting to use large samples of cities that are themselves a relevant case study.

The main contribution of this paper in relation to previous literature is to examine the impacts of light rail systems (light metros, light rails and tram) on congestion, travel time and pollution in a large sample of mid-size European cities. Previous studies have focused on either congestion or pollution, while I consider these two factors together. Furthermore, I focus on the impacts of light rails while previous studies usually consider different types of public transportation or focus on heavy metro systems that are not a viable alternative for mid-size cities. Finally, with the exception of the study by Gendron-Carrier et al. (2018), all previous studies have based their analyses on a single country or city.

I find evidence that an increase in the supply of rails leads to less congestion, less travel time and less pollution. Furthermore, I find that cities with a new rail system have 7% less congestion, 1% less travel time and 3% less pollution than cities with no rail systems. I find no evidence of different pre-trends of treated and control cities. The impact of rails on congestion, travel time and pollution is gradual, and increases over time.

The rest of the paper is organized as follows. The next section explains the empirical equation that I estimate and discusses the identification strategy. The following section provides relevant information about the data employed in the empirical analysis. This is followed by a section on the results of the econometric estimates. The last section is devoted to a discussion and concluding remarks.

2. Methodology

The empirical analysis is of mid-size cities in several countries of the European Union (EU) and United Kingdom for which congestion and pollution data is available. The sample includes information on 98 cities in 13 countries, with a population ranging from 100,000 to 1 million inhabitants. Congestion data is for the period 2008 to 2019 and pollution data is for 2008 to 2016 (see below for more details). Fifty-five cities had an urban rail system before 2008, although 11 of them had expanded considerably in the study period. Eleven cities launched a new urban rail service during the study period and 32 cities did not have an urban rail system in the period. Overall, the entire sample contains 1174 observations in regressions with congestion or travel time as a dependent variable and 873 observations in regressions with pollution as a dependent variable. I also consider a subsample based on cities that did not have an urban rail system in the initial year of the period. This subsample is based on 43 cities and contain 514 observations in regressions with congestion or travel time as a dependent variable and 378 observations in regressions with pollution as a dependent variable. The population of cities in this subsample ranges from 100,000 to 700,000 inhabitants.

I estimated the following set of equations for city i in year t:

$$\mathrm{Y}_{it} = \alpha + \delta Rail\_length_{it} + \beta' X_{it} + \mu'_t + \lambda'_i + \varepsilon_{it} \tag{1}$$

$$\mathrm{Y}_{it} = \alpha + \delta D_{it}^{Rail} + \beta' X_{it} + \mu'_t + \lambda'_i + \varepsilon_{it} \tag{2}$$

where all the continuous variables without zero values are transformed using logarithms, so that the influence of outliers is reduced and parameter estimates can be interpreted as elasticities.

Three dependent variables (Y_{it}) are considered. The first is an indicator of congestion that measures the percentage of additional time that a vehicle needs for any trip in the city compared to a situation characterized by free traffic flow conditions. This is the primary indicator offered by TomTom, which is the source from which I extracted the data. The second is an indicator of average travel time using the procedure suggested by TomTom to turn the congestion values into average travel time values. In this regard, I build a variable of average travel time for a 20-minute trip.[1] The third is an indicator of pollution based on annual mean estimates of particulate matters with a diameter of 2.5 µm or less (PM 2.5) with the method outlined in van Donkelaar et al. (2019). I focus on PM 2.5 because data for other pollutants are not available.

In all equations, I consider three attributes of cities as control variables (X_{it}): population, density and income. Population is considered at urban level as a proper measure of the real size of the city. The density variable is the number of inhabitants per square kilometer at city level. The income of the city is given by the GDP per capita at NUTS-3 level.[2]

In the pollution equation, two variables are considered that account for the need for cooling and heating: the cooling degree days index (CDD) and the heating degree days index (HDD). The reason for considering these variables is that a relevant source of emissions is associated with households' electricity consumption, especially when it is generated from coal. Data for these two variables are provided at NUTS-2 level. Furthermore, a variable measuring the amount of precipitation is considered.

The appendix provides more details of the congestion, pollution, HDD, CDD and rain variables.

I estimate a city fixed-effects model that identifies changes from one year to another, as it seemed the most appropriate method to evaluate the effect of rails. Consequently, I included city (λ'_i) and year fixed effects (μ'_t). The model is based on the within transformation of the variables as deviations from the average. Thus, the model allows a comparison of changes in outcomes between cities that have new or expanded rail lines and cities that do not have them. Furthermore, the city-fixed effects model controls for omitted and time-invariant variables correlated with the variables of interest. The year fixed effects control for yearly effects common to all urban areas.

The main variable of the analysis is that of rail systems. Note that I only consider rail lines within the city, not rail lines that link the city center with surrounding cities. The assumption here is that the major shock that can take place in terms of public transportation is associated with rail investments. While bus rapid transit deserves to be considered a competitive alternative to rail systems, it is difficult to analyze in the context of this research, given the lack of consensus about its definition and the lack of data on the network length.

I consider two approaches to identify the impact of rail systems on the negative externalities generated by cars. In equation (1), I use the entire sample of cities and the rails variable is based on the network length of the rail system measured in line kilometers.

However, my main identification strategy relies on the subsample of cities that did not have urban rail systems in the initial year of the period. With this subsample, I identify the impact of the shock of the change from not having to having an urban rail transport system.

In equation (2), I apply the logic of differences in differences (DiD), which is a common methodology used within the treatment evaluation framework (see Angrist and Picke, 2009; Gertlet et al., 2016 for details). The basic idea of DiD is to isolate the effect of the treatment (i.e. having a rail system) over a particular outcome by comparing changes in the outcomes of a set of cities that received treatment (treated group) with the changes in outcomes of a similar group of cities that were not exposed to treatment (control group). In the DiD analysis, the sample is restricted to cities that did not have a rail system in the initial year of the study period. The rail variable is defined as a dummy variable that takes the value one from the year in which the rail system started operation.

I also reframe equation (2) in two ways. In equations (2a and 2b), I add a variable that differentiate between two types of rail systems. High-capacity systems are light metros that can operate at higher capacity and speed than light rails or trams but less than typical heavy rails like metros. Medium-capacity systems are light rails and trams. Differences between light rails and trams are often indistinct and a given system may combine multiple features. The main difference is that light rails operate on exclusive lanes, but most

[1] For example, if the congestion measure is 29% for Liverpool this turns into an average travel time of 25.8 min (20 + 0.29 X 20).

[2] The NUTS classification (Nomenclature of territorial units for statistics) is a hierarchical system for dividing up the economic territory of the European Union into three different hierarchical levels — NUTS level 1, level 2 and level 3 respectively — detailing larger to smaller territorial units.

Table 1
Descriptive statistics of the variables used in the empirical analysis.

Variable	Source	# Obs.	Mean	Standard deviation	Minimum value	Maximum value
Congestion (excess travel time in % in relation to free flow conditions)	TomTom Traffic Index (2020): https://www.tomtom.com/en_gb/trafficindex	1174	23.6720	7.3063	10	56
Travel time (Average travel time in minutes for a 20-minute trip)	TomTom Traffic Index (2020): https://www.tomtom.com/en_gb/trafficindex	1174	24.7344	1.4612	22	31.2
Pollution (estimates of PM 2.5 - μg/m^3)	Atmospheric Composition Analysis Group (2020): http://fizz.phys.dal.ca/~atmos/martin/?page_id = 140	873	14.9002	5.0575	4.6	38.7
Length of rail (kilometers)	Urban rail (2020): (http://www.urbanrail.net/), World Metro database (2020): (http://mic-ro.com/metro/table.html) and websites of operators	1176	34.9619	45.3133	0	198
D^{rail} (dummy variable)	Urban rail: (http://www.urbanrail.net/), World Metro database: (http://mic-ro.com/metro/table.html) and websites of operators	1176	0.6267	0.4838	0	1
Population (thousands of inhabitants – urban area	Word Urbanization Prospects, United Nations (2020)	1176	541.5023	195.7836	279.38	1107.73
Density (number of inhabitants per square kilometer)	Eurostat (2020)	1176	2824.6140	1751.5600	260	9379
Income (thousands of euros per inhabitant – NUTS 3)	Eurostat (2020)	1176	32.1904	14.6753	7	93
HDD (Heating degree days index – NUTS 2)	Eurostat (2020)	882	2495.9620	768.7718	19.8	5680.7
CDD (Cooling degree days index – NUTS 2)	Eurostat (2020)	882	90.0902	129.4825	0	517.7
Rain (precipitation sum with unit 0.01 mm)	European Climate Assessment dataset (2020): https://www.ecad.eu/	882	72562.6100	24905.500	6760	233,320

trams considered in this study also include grade-separated systems. With this equation, I examine a potential differential impact between high-capacity and medium-capacity rail systems.

In equation (2c), the rail dummy variable is now defined as $\sum_{k=-3}^{3} D_{it}^{Rail}$ so that it equals 1 if the rail system is in operation in the city *k* years from year *t*. This definition implies that k = 0 represents the first year following the launch of the rail system, k = -1 is the year prior to treatment and k = 1 is the year after treatment. With this event study analysis, I examine the timing of rail effects after treatment. I choose three years before and after treatment to capture the timing for most of the treated cities. Importantly, the fundamental identification assumption of the DiD analysis is the existence of parallel trends between groups in the absence of treatment. If the assumption holds, by comparing the change in outcomes of both groups, we can properly identify the effect of the rail systems. Hence, the identification of a causal impact of the new rail systems requires a negative statistically significant impact only after treatment.

$$Y_{it} = \alpha + \delta_1 D_{it}^{Rail} + \delta_2 D_{it}^{Rail\_high\_capacity} + \beta' X_{it} + \mu'_t + \lambda'_i + \varepsilon_{it} \quad (2a)$$

$$Y_{it} = \alpha + \delta_1 D_{it}^{Rail} + \delta_2 D_{it}^{Rail\_medium\_capacity} + \beta' X_{it} + \mu'_t + \lambda'_i + \varepsilon_{it} \quad (2b)$$

$$Y_{it} = \alpha + \sum_{k=-3}^{3} \delta' D_{it}^{Rail} + \beta' X_{it} + \mu'_t + \lambda'_i + \varepsilon_{it} \quad (2c)$$

I expect a negative significant impact of rail systems on congestion and travel time if they imply a shift in mobility from cars to rails. Such a shift could be explained mainly by faster journey times, greater convenience and more reliability than buses. However, this substitution effect could be at least partially compensated by the fact that rail systems may spur economic activity and hence car traffic. The statistical impact of rail systems on PM 2.5 pollutants is less clear a priori. The positive relationship between congestion and pollution is well-documented, with prolonged car circulation at reduced speeds having a notable effect on the emission of polluting substances (Barth and Boriboonsomsin, 2008; Parry et al., 2007). However, according to the European Environmental Agency, road transport represents around 10% of total PM 2.5 emissions, which is relatively low in comparison with other pollutants like nitrogen oxides (28%) and carbon monoxide (18%). Thus, we could expect rail systems to significantly improve air quality in terms of PM 2.5 only if they imply a substantial reduction in car traffic.

Estimates may present problems of heteroscedasticity and temporal autocorrelation in the error term. The Breusch-Pagan/Cook-Weisberg test shows heteroscedasticity problems so that standard errors are robust to heteroscedasticity. Likewise, the Wooldridge test for autocorrelation in panel data reveals that a problem of serial autocorrelation may exist so that I allow for an arbitrary variance–covariance structure by computing the standard errors in clusters by city to correct for autocorrelation in the error term both at the cross-sectional and temporal levels (Bertrand et al., 2004).

Table 1 shows the descriptive statistics and the sources of all variables used in the empirical analysis.

Table 2
Cities with new rail systems in the sample period.

City	System type	Starting date of operation	Kms at starting year	Stations	Annual ridership (million passengers)	Cost of the project (million euros)
Alicante	Tram	September 2013	7.1	14	4.8	100
Avignon	Tram	October 2019	5.2	10	N.A	135
Brescia	Light metro	March 2013	13.1	17	18.7	830
Edinburgh	Tram	May 2014	13.5	16	7.5	858
Florence	Tram	February 2010	7.4*	20	25.1	721
Malaga	Light metro	July 2014	11.3	17	6.3	870
Murcia	Light rail	May 2011	17.5	28	5.0	211
Seville	Light metro	April 2009	18.5	22	16.9	658
Tours	Tram	August 2013	15	29	23.7	369
Zaragoza	Light rail	April 2011	12.8	25	27.8	355
Palermo	Tram	December 2015	17	44	9.1	322

Sources: Urban rail (http://www.urbanrail.net/), World Metro database (http://mic-ro.com/metro/table.html) and operators' websites. *Line extensions in 2018 (11.5 kms and 26 stations) and 2019 (16.9 kms and 38 stations). Annual ridership for Alicante, Brescia, Edinburgh, Florence and Murcia refers to 2019. Annual ridership for Malaga, Seville and Zaragoza refers to 2018, and annual ridership for Palermo and Tours refers to 2017 and 2016, respectively.

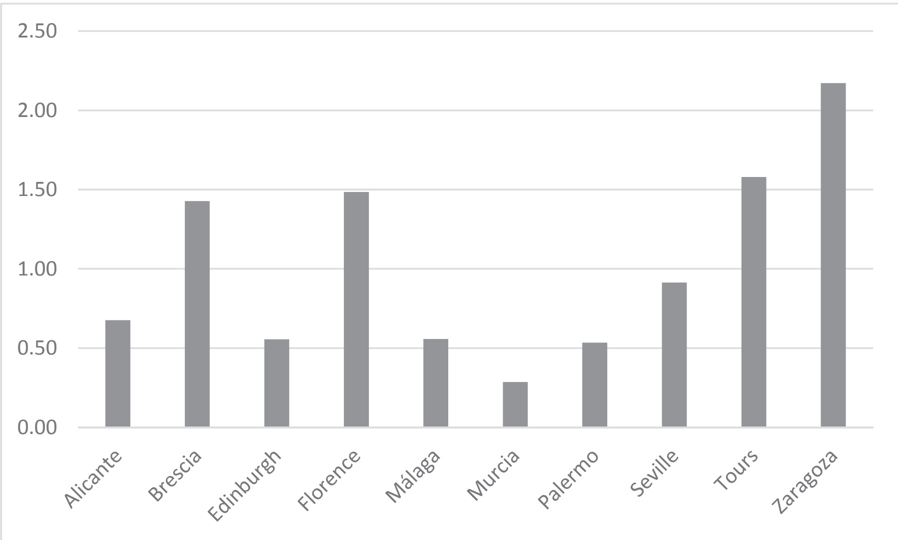

Fig. 1. Million passengers per km. Note: Annual ridership for Alicante, Brescia, Edinburgh, Florence and Murcia refer to 2019. Annual ridership data for Málaga, Seville and Zaragoza are for 2018, and annual ridership data for Palermo and Tours are for 2017 and 2016, respectively.

4. Data

Table 2 provides information about the 11 cities with new rail systems in the sample period. For most cities, the length of the rail network is from 11 to 18 km of lines and from 16 to 29 stations. The exceptions are Alicante and Avignon, which have fewer than 10 km of network, and Palermo with 44 stations. In all cases, the costs of the investment are greater than 100 million euros although differences across cities are remarkable. Costs have been particularly high in Brescia, Edinburgh, Florence, Malaga and Seville with values that range from 658 to 870 million euros. Costs in Murcia, Tours, Palermo, Tours and Zaragoza range from 200 to 370 million euros, while the lower bound is for the smaller networks of Alicante and Avignon with costs of 100 to 135 million euros.

Fig. 1 shows remarkable differences across cities in terms of million passengers per kilometer of line (data are for the last year with available information). These numbers are above one million for Brescia, Florence, Tours and Zaragoza, and close to one for Seville. For the remaining cities, the number of passengers per km is well below one million. Such heterogeneity in the ridership performance of cities could have several explanations. For example, Carpintero and Siemiatyck (2016) analyzed the influence of political factors in the decision-making process of light rail projects in Zaragoza and Murcia. They showed that the final decision about route choice was based on political factors. In both cases, there was public opposition to the rail system running through the city center, which was the option with the highest estimated ridership. In Zaragoza, the rail system ended up crossing the city center, but this was not the case in

Table 3
Cities with relevant rail expansions in in the sample period.

City	System type	Extension year	Kilometers added	% Increase in kilometers	Total length (kms)	New stations	Total stations
Bordeaux	Light rail/Tram	2015/2016	14	27%	66.1	25	90
Cagliari	Light rail	2015	6	100%	12	5	12
Charleroi	Light rail	2012/2013	10	43%	33	18	48
Catania	Metro	2015	5	132%	8.8	5	11
Grenoble	Tram	2014/2015	13	58%	42	10	81
Montpellier	Tram	2012	32	116%	60.5	47	84
Nottingham	Light rail/Tram	2015	17.5	125%	31.5	15	32
Padova	Tram	2010	3.6	54%	10.3	5	13
P. Mallorca	Light metro	2013	8.3	114%	15.6	7	16
Saint Etienne	Tram	2019	4	33%	16	6	43
Toulouse	Metro, tram	2010/2014	16.6	59%	43.3	27	65

Sources: Urban rail (http://www.urbanrail.net/), World Metro database (http://mic-ro.com/metro/table.html) and websites of operators.

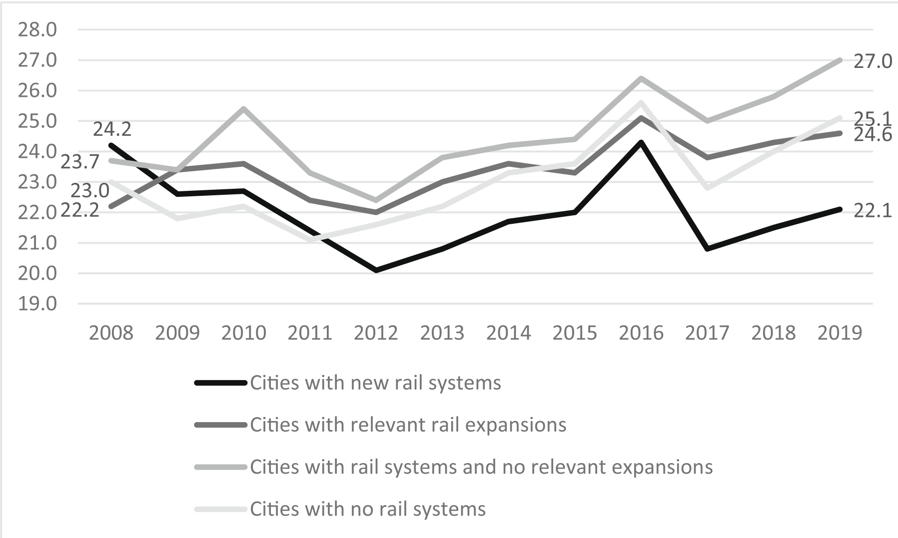

Fig. 2. Evolution of congestion in different groups of cities.

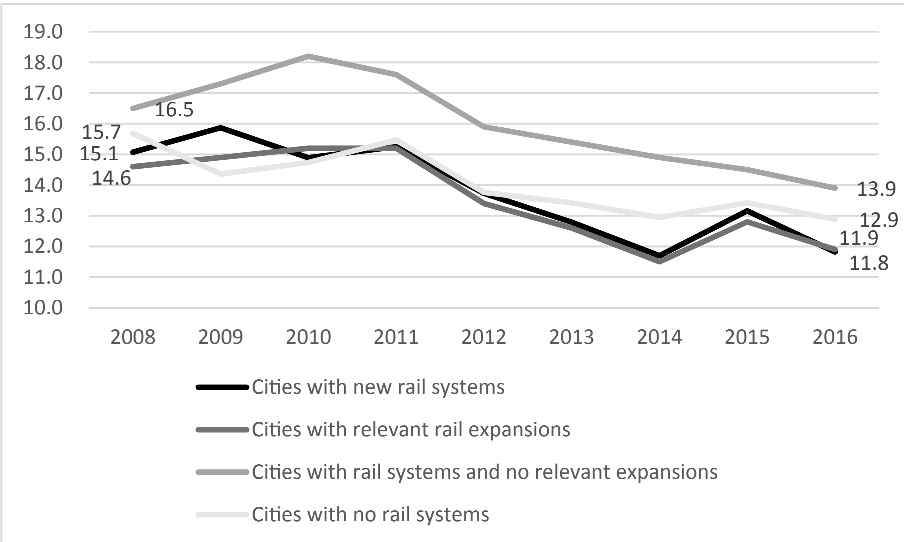

Fig. 3. Evolution of pollution in different groups of cities.

Table 4
Descriptive statistics by groups of cities.

	Cities with new rail systems	Cities with relevant rail expansions	Cities with rail systems and no relevant expansions	Cities with no rail systems
Congestion	22.0227 (8.8787)	23.4848 (5.5855)	24.5625 (7.2981)	23.07592 (7.1019)
Travel time	24.4045 (1.7757)	24.6969 (1.1171)	24.9125 (1.4596)	24.6151 (1.4203)
Pollution	13.3159 (4.5213)	13.1818 (4.6880)	15.5340 (4.8321)	13.7876 (5.0923)
Length of rail	7.7340 (7.3102)	25.3639 (18.6494)	69.5952 (46.6597)	0
Population	552.1020 (158.2890)	582.8370 (192.1238)	576.2847 (205.5581)	475.8242 (177.8755)
Density	2983.7270 (2009.2150)	3480.3560 (1986.8930)	2788.1330 (1838.9730)	2594.6690 (1347.5940)
Income	26.4469 (10.8868)	26.8636 (5.0956)	36.3901 (18.5443)	30.2213 (9.4208)
HDD	1860.3050 (786.2093)	2037.2940 (737.7210)	2804.5440 (538.6806)	2447.8360 (824.3264)
CDD	231.0374 (157.1573)	141.2849 (136.3480)	32.1258 (51.3444)	103.7481 (143.0930)
Rain	59168.48 (26495.1500)	70079.4900 (19761.5800)	74861.5300 (19273.6500)	73647.8360 (29835.96)

Note: I report mean values and standard errors (standard errors in parenthesis)

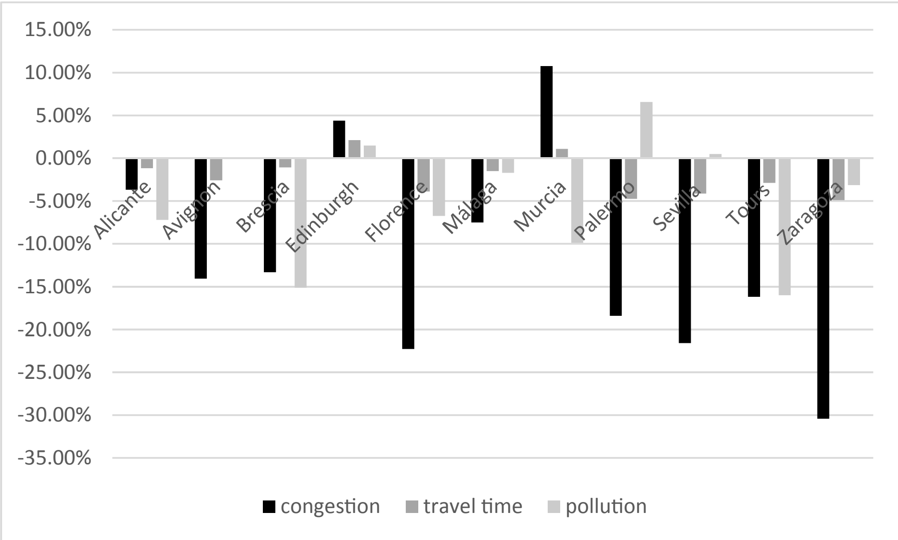

Fig. 4. Evolution of congestion, travel time and pollution in treated vs. control cities from the year before the rail start up.

Murcia. A European Commission study suggests that the poor performance of the tram in Palermo could be explained by the lack of tariff integration between public transport operators.

Table 3 provides some information about cities in the sample that have undertaken relevant expansions of their rail systems in the study period. In most cases, the expansion of the rail system represents more than a 50% increase in relation to the network length before the expansion. However, the number of kilometers added ranges from 3 to 17 km and the number of new stations from 5 to 47.

Fig. 2 shows the evolution of congestion in groups of cities according to the changes in their rail systems. In the initial year of the study period, cities that had launched new rail services were slightly more congested than the rest of cities. At the end of the period, they were less congested than the rest of cities. In fact, this is the only group of cities in which congestion fell in the study period. The poorest performance was for cities with rail systems and no relevant expansions in the study period, while cities with relevant rail expansions and cities with no rail systems experienced a slight rise in congestion levels in the study period. Thus, the numbers in this figure show that congestion falls for cities with new rail systems but does not fall for cities with additions to an existing rail system. This may suggest diminishing returns to transit size as opposed to network effects from increased connection possibilities.

Fig. 3 shows the evolution of pollution for the same groups of cities. Pollution levels dropped in all cities in the study period. However, this drop was greater in cities with new rail systems. The decreasing trend in pollution in all cities could be explained by the fact that new cars have more efficient fuel consumption and pollute less in kilometers driven than older ones. Cities with rail investments in the study period (either cities with new projects or expanded existing projects) had slightly lower levels of pollution than the other group of cities in the initial year of the study period. Cities with rail systems and no relevant expansions were those with the highest levels of pollution.

Up to this point, a potential endogeneity bias could underestimate the impact of rail systems, as cities that invest in rails may be those with the highest levels of pollution, travel time and congestion. However, differences in congestion between groups of cities are

Table 5
Results of estimates: impacts of rail length network.

	Congestion	Travel time	Pollution
Length of rail (coefficient)	−0.0040 (0.0012)***	−0.0007 (0.0002)***	−0.0015 (0.0007)**
Length of rail (marginal impact)	−0.1428 (0.0438)***.	−0.0270 (0.0075)***	−0.0527 (0.0254)**.
Log (Population)	1.0003 (0.25)***	0.1902 (0.0491)***	0.0301 (0.2633)
Log (Density)	0.3437 (0.1486)**	0.0560 (0.0254)**	−0.0569 (0.1462)
Log (Income per capita)	0.1789 (0.1439)	0.0291 (0.0273)	0.3485 (0.1120)***
CDD	–	–	0.0005 (0.0002)***
HDD	–	–	0.00006 (0.00003)*
Log (*Rain*)	–	–	0.0027 (0.0314)
Intercept	−10.8284 (0.9591)***	1.9247 (0.1981)***	1.6289 (1.2486)
R^2	0.23	0.22	0.57
# Observations	1174	1174	873
# Cities	98	98	97

Notes: Standard errors in parentheses (robust to heterocedasticity and clustered at the city level). All regressions include city fixed effects and year fixed effects. Statistical significance at 1% (***), 5% (**), 10% (*).

low in the baseline year and cities that have invested in rails have even lower levels of pollution than the rest of cities in the baseline year. In particular, differences between treated and control cities in the subsample of cities that did not have urban rails in the baseline year are small enough to assume that the endogeneity bias (if any) should be modest.

In this regard, table 4 shows the descriptive statistics for the 4 groups of cities considering all variables of the empirical analysis (both dependent and explanatory variables). Differences between treated and control cities are small, which is particularly relevant for those cities included in the DiD model. To confirm this, table A1 in the appendix shows the estimates of the determinants of being treated (i.e.; to get a new rail system during the study period) considering jointly and separately dependent and explanatory variables as regressors. In these regressions, I use data for the first year of the considered period and the cities of the DiD sample. Results of these regressions show that there are no systematic differences between treated and control cities before treatment in terms of the covariates (population, income, density, weather variables) and more importantly in terms of congestion, travel time and pollution. This analysis does not exclude the existence of unobserved time-varying differences between treated and control cities but these potential differences are not generating relevant differences in terms of congestion, travel time and pollution.

Data in Figs. 2 and 3 represent an average for the groups of cities, but some heterogeneity may arise between cities in the same group. In this regard, the comparison between treated and control cities in the subsample of cities that do not have rail systems in the baseline year is particularly illustrative.

Fig. 4 shows the evolution of congestion, travel time and pollution in each of the 11 treated cities (cities with new rail systems during the study period) in comparison with control cities (cities that do not have rail systems in any year of the study period). For each city, I compare the evolution of congestion and pollution from the year before the rail start up to the last year with data available. For example, the year previous to the launch of rail services in Zaragoza was 2010, so I compare the evolution of congestion in Zaragoza with that in control cities for 2010–2019 and the evolution of pollution in Zaragoza in relation to control cities for 2010–2016. Note that we cannot infer from this figure any differences between the evolution of congestion and pollution, given that they are not comparable. Indeed, the last year with data available is 2019 for congestion and 2016 for pollution.

Most of the treated cities show a good performance in terms of congestion in comparison to control cities. Differences are smaller in terms of travel time, which it may be explained by the lower variability of the travel time variable in relation to the congestion measure.

The decrease in congestion and travel time is particularly pronounced in cities with high levels of rail ridership like Brescia, Florence, Palermo, Seville, Tours and Zaragoza and to lower extent Avignon for which data on annual ridership is not available. Alicante and Malaga show modest decreases in congestion and travel time that correspond to moderate levels of annual ridership. The performance of Edinburgh and Murcia is poor in terms of congestion and travel time. Both cities had low levels of traffic, particularly Murcia.

Furthermore, most treated cities performed better in terms of pollution than control cities. The main exception is Palermo, although the tram in this city started to operate in December 2015, so only a short period is available to assess the impact on pollution. Finally, Edinburgh and Seville followed a similar evolution in pollution to the control cities.

5. Results

Table 5 shows the results of estimates of equation (1), where the main variable of interest is the length of the rail system and the entire sample of cities is considered. Note that I report the results of the estimated coefficient and the elasticity, given that the variable of rail length is not in logs. When the dependent variable is the measure of congestion, the coefficient of the rail variable is negative and statistically significant at the 1% level. In terms of elasticities, a 1% increase in rail supply led to a −0.14% decrease in congestion. Similarly, when the dependent variable is travel time, the coefficient of the rail variable is negative and statistically significant at the 1% level but the estimated elasticity is smaller. This may be explained by the lower variability of the travel time variable in relation to the congestion measure. While the variable of travel time is based on the assumption that the average trip is 20 min for all cities, it is

Table 6
Results of estimates for congestion and travel time: DiD.

	Dependent variable: Congestion				Dependent variable: Travel time			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
D^{rail}	−0.0772 (0.0165)***	–		–	−0.0142 (0.0035)***	–		–
$D^{rail} + D^{rail\_high\_capacity}$	–	−0.1006 (0.0369)***	–	–	–	−0.0217 (0.0070)***	–	–
$D^{rail} + D^{rail\_medium\_capacity}$	–	–	−0.0662 (0.0131)***	–	–	–	−0.0115 (0.0031)***	–
$D^{rail(t\text{-}3)}$ - $D^{rail(t\text{-}1)}$	–	–	–	0.0016 (0.0581)	–	–	–	0.0006 (0.0094)
$D^{rail(t\text{-}2)}$ - $D^{rail(t\text{-}1)}$	–	–	–	0.0036 (0.0227)	–	–	–	0.0018 (0.0041)
$D^{rail(t0)}$ - $D^{rail(t\text{-}1)}$	–	–	–	−0.0132 (0.0287)	–	–	–	−0.0004 (0.0053)
$D^{rail(t+1)}$ - $D^{rail(t\text{-}1)}$	–	–	–	−0.0360 (0.0130)***	–	–	–	−0.0071 (0.0030)**
$D^{rail(t+2)}$ - $D^{rail(t\text{-}1)}$	–	–	–	−0.0609 (0.0148)***	–	–	–	−0.0122 (0.0037)***
$D^{rail(t+3)}$ - $D^{rail(t\text{-}1)}$	–	–	–	−0.0821 (0.0265)***	–	–	–	−0.0153 (0.0050)***
R^2	0.34	0.34	0.34	0.34	0.33	0.33	0.33	0.33
# Observations	514	514	514	514	514	514	514	514
# Cities	43	43	43	43	43	43	43	43

Notes: Standard errors in parentheses (robust to heterocedasticity and clustered at the city level). All regressions include city fixed effects and year fixed effects. Statistical significance at 1% (***), 5% (**), 10% (*).

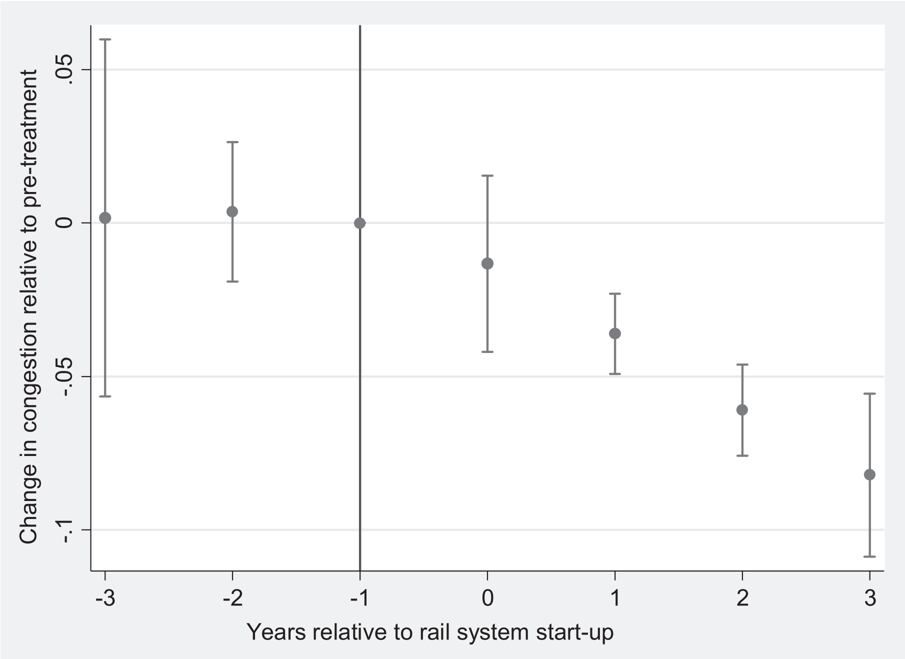

Fig. 5. Impact of urban rail systems on congestion. Note: This figure shows the main results of the event study analysis in which the treatment variable is based on dummies for three years before and three years after the launch of urban rail services. The figure shows the differences in the estimated coefficients of the impact of rails on congestion relative to the year previous to entry in the city. I also show the 95% confidence interval based on standard errors.

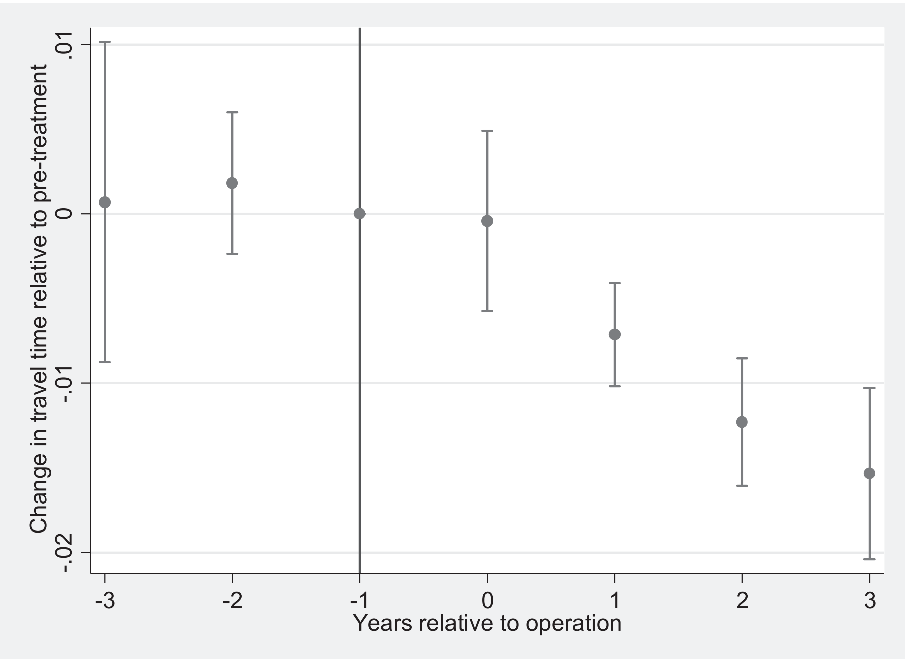

Fig. 6. Impact of urban rail systems on travel time. Note: This figure shows the main results of the event study analysis in which the treatment variable is based on dummies for three years before and three years after the launch of urban rail services. The figure shows the differences in the estimated coefficients of the impact of rails on travel time relative to the year previous to entry in the city. I also show the 95% confidence interval based on standard errors.

Table 7
Results of estimates for pollution: DiD.

	(1)	(2)	(3)
D^{rail}	−0.0345 (0.0121)***	–	–
$D^{rail} + D^{rail\_high\_capacity}$	–	−0.0410 (0.0208)**	–
$D^{rail} + D^{rail\_medium\_capacity}$	–	−0.0319 (0.0139)**	–
$D^{rail(t-3)}$ - $D^{rail(t-1)}$	–	–	−0.0106 (0.0253)
$D^{rail(t-2)}$ - $D^{rail(t-1)}$	–	–	0.0178 (0.0181)
$D^{rail(t0)}$ - $D^{rail(t-1)}$	–	–	0.0177 (0.0238)
$D^{rail(t+1)}$ - $D^{rail(t-1)}$	–	–	0.0044 (0.0330)
$D^{rail(t+2)}$ - $D^{rail(t-1)}$	–	–	−0.0302 (0.0145)**
$D^{rail(t+3)}$ - $D^{rail(t-1)}$	–	–	−0.0624 (0.0178)***
R^2	0.54	0.54	0.54
# Observations	378	378	378
# Cities	42	42	42

Notes: Standard errors in parentheses (robust to heterocedasticity and clustered at the city level). All regressions include city fixed effects and year fixed effects. Statistical significance at 1% (***), 5% (**), 10% (*).

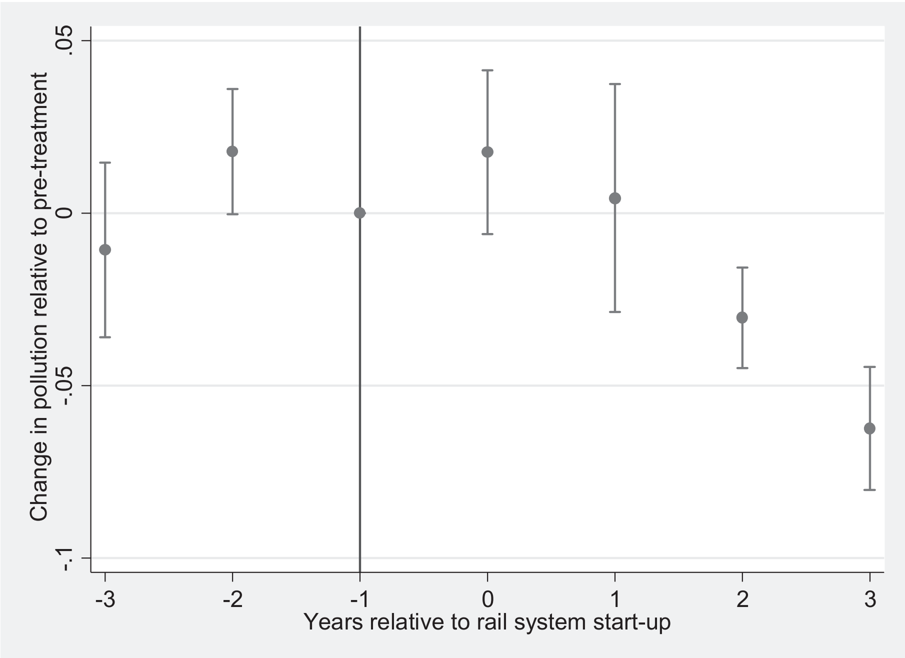

Fig. 7. Impact of urban rails on pollution. Note: This figure shows the main results of the event study analysis in which the treatment variable is based on dummies for three years before and three years after the launch of urban rail services. The figure shows the differences in the estimated coefficients of the impact of rails on pollution relative to the year previous to entry in the city. I also show the 95% confidence interval based on standard errors.

useful to show that the magnitude of the impact of rails is smaller when considering travel time instead of the congestion measure.

When the dependent variable is the measure of pollution, the coefficient of the rail variable is negative and statistically significant at the 5% level. In terms of elasticities, a 1% increase in rail supply led to a −0.05% decrease in pollution in terms of PM 2.5. Thus, I find evidence that an increase in rail supply leads on average to less congestion, less travel time and less pollution. In terms of controls, congestion and travel time is higher in larger and denser cities. Pollution is higher in richer cities and in cities with more adverse weather in terms of temperature.

Table 6 shows the results of the DiD estimates when the dependent variable is congestion and travel time. In both cases, the DiD coefficient is negative and statistically significant at the 1% level. Cities with a new rail system benefit on average from a 7.7% decrease in congestion and a 1.4% decrease in travel time in comparison to cities with no rail systems. The impact is greater for high-capacity than medium-capacity rail systems, although in both cases the effect is statistically significant at the 1% level. The impact is about −10%/-2% for high-capacity systems and −6%/- 1% for medium-capacity systems.

The difference in results between the congestion and travel time measures is remarkable but it may be explained by the lower variability of the travel time measure. This lower variability can be explained by means of an example.

The congestion measure is based on an indicator of extra travel time as a percentage. For example, the mean extra travel time in

Table 8
Results of estimates excluding control variables.

	Dependent variable: Congestion		Dependent variable: Travel time		Dependent variable: Pollution	
	All sample	DiD	All sample	DiD	All sample	DiD
Length of rail (coefficient)	−0.0037 (0.0012) ***	–	−0.0007 (0.0002) ***	–	−0.0018 (0.0007) ***	–
Length of rail (marginal impact)	−0.1314 (0.0451) ***	–	−0.0244 (0.0077) ***	–	−0.0636 (0.0260) ***	–
D^{rail}	–	−0.0923 (0.0248) ***	–	−0.0169 (0.0046) ***	–	−0.0354 (0.0159) ***
R^2	0.18	0.21	0.17	0.19	0.52	0.43
# Observations	1174	514	1174	514	873	378
# Cities	98	43	98	43	97	42

Notes: Standard errors in parentheses (robust to heterocedasticity and clustered at the city level). All regressions include city fixed effects and year fixed effects. Statistical significance at 1% (***), 5% (**), 10% (*).

Table 9
Results of estimates for congestion, travel time and pollution: City specific coefficients (DiD).

	Dependent variable: congestion	Dependent variable: travel time	Dependent variable: pollution
D^{rail} (Alicante)	−0.1425 (0.0451)***	−0.0242 (0.0079)***	−0.0642 (0.0406)
D^{rail} (Avignon)	0.0218 (0.0382)	0.0011 (0.0055)	–
D^{rail} (Brescia)	0.0107 (0.0433)	−0.0045 (0.0050)	−0.0181 (0.0261)
D^{rail} (Edinburgh)	−0.0306 (0.0204)	0.0099 (0.0068)	0.0469 (0.0342)
D^{rail} (Florence)	−0.1136 (0.0510)**	−0.0216 (0.0099)**	0.0048 (0.0105)
D^{rail} (Malaga)	−0.2625 (0.0944)***	−0.0486 (0.0184)***	−0.0827 (0.0212)***
D^{rail} (Murcia)	−0.2663 (0.0345)***	−0.0418 (0.0067)***	−0.0881 (0.0321)***
D^{rail} (Seville)	0.0127 (0.0180)	0.0012 (0.0032)	−0.0293 (0.0400)
D^{rail} (Tours)	0.0601 (0.0711)	0.0020 (0.0094)	−0.1739 (0.0676)***
D^{rail} (Zaragoza)	−0.1563 (0.0964)	−0.0227 (0.0151)	−0.0372 (0.0541
D^{rail} (Palermo)	−0.0936 (0.0184)***	−0.0291 (0.0077)***	0.064 (0.0356)*
R^2	0.38	0.34	0.34
# Observations	514	514	514
# Cities	43	43	43

Notes: Standard errors in parentheses (robust to heterocedasticity and clustered at the city level). All regressions include all controls, city fixed effects and year fixed effects. Statistical significance at 1% (***), 5% (**), 10% (*).

Alicante was 21% in 2008 and 18% in 2019. This implies that the reduction in congestion between the two periods was −14.3%. In terms of average travel time, the numbers are 24.2 min (20 + 0.21*20) in 2008 and 23.6 (20 + 0.18*20) minutes in 2019 that means that the reduction in average travel time between the two periods was −2.5%. Therefore, the coefficient of the light rail variable is lower when the dependent variable is travel time instead of extra travel time even though both estimates are capturing essentially the same impact. The difference comes from the different measurement scale of the two dependent variables.

Columns (4) and (8) of Table 6 and Figs. 5 and 6 show the timing effects of the new rail systems. Recall that in this event study analysis the treatment variable is based on dummies for three years before and three years after the launch of the rail service. All estimates are expressed as changes from event date k = -1, or the year just prior to treatment, the estimates for which are normalized to 0.

The difference is not significant between the coefficients for the pre-treatment year and previous years. Thus, different pre-trends between treated and control cities do not seem to have biased the results. Furthermore, the difference is not significant between coefficients for the year previous to treatment and the year in which rail services are launched. The data are annual but the launch of a rail service may take place at the middle or end of the year. Furthermore, it is reasonable to assume that it could take some months for the shift from cars to trains to reach significant numbers. This may explain why the impact of rails on congestion and travel time is not statistically significant in the year in which the service is launched. The difference is statistically significant (at 1% level) between coefficients for the year previous to treatment and coefficients for the years following the launch of the rail service in the city. The magnitude of the negative impact of rails on congestion and travel time is increasing over time. Thus, the evidence shows that the impact of rails on congestion and travel time is gradual as it increases over time, and it is only relevant after a full year of offering services in the city.

Table 7 shows the results of the DiD estimates when the dependent variable is pollution. The DiD coefficient is negative and statistically significant at the 1% level. I find that cities with a new rail system benefit on average from a −3% decrease in pollution in terms of PM 2.5 in comparison to cities with no rail systems. The impact is similar for high-capacity and medium-capacity rail systems.

Column (3) of Table 6 and Fig. 7 shows the timing effects of the new rail systems for pollution. I do not find evidence of a negative pre-trend biasing my results. However, a negative, significant impact of the rail variable is found from the second full year in which the rail service is in operation.

Thus, the negative impact of light rails on pollution take longer to be significant than on congestion and travel time. However, results for pollution are remarkable since cars represent a modest amount of total emissions in terms of PM 2.5 and all sample cities had a decreasing trend in terms of pollution in the considered period. Although no data are available for testing, it is possible that rails have had a greater impact on other types of pollutants such as nitrogen oxides or carbon monoxide.

As a robustness check, table 8 show the results of the estimates excluding the control variables. Results are very similar (essentially identical) regardless of adding or not the control variables. Furthermore, table 9 report the results of the DiD coefficients for each treated city. While the interpretation of the coefficients must be interpreted here with caution because they refer to a low number of observations, they provide a robustness check since we should not expect positive and statistically significant impacts of the rail variable. In this regard, the city-specific coefficients are negative and statistically significant or non-statistically significant. The only exception is the coefficient for Palermo that is positive and statistically significant at the 10% level in the equation for pollution. However, this latter result may be explained by the fact that the rail-system in Palermo was launched in December of 2015 and data for pollution ends in 2016.

Finally, a potential limitation of the data must be taken into account in the interpretation of the elasticity estimates. The sample used in this analysis considers comparable treated and control cities in terms of the covariates and the pre-levels of the outcome variables. Furthermore, I provide evidence that the parallel trends assumption holds. However, there may be time-varying unobservable differences across cities in the treatment and control groups that could be contemporaneous with rail investments. Rail investments are not randomly determined temporally and spatially and are often part of other policies that may occur in the same year. Thus, the rail variable may be picking up some of these other effects if there are systematic differences for cities that get new rail systems and for cities that do not get new rail systems.

6. Discussion and concluding remarks

In this paper, I have examined the impact of light rails systems on congestion, travel time and pollution using panel data for a large sample of cities in Europe. I have found that rail supply increases lead to less congestion, travel time and pollution. This result held in the analysis that consider the entire sample of cities and in the analysis that focuses on the sample of cities that did not have rail systems in the baseline year. Furthermore, the impact on congestion and travel time is stronger for high-capacity than medium-capacity systems, while no clear differences were found regarding the type of rail system for pollution. Finally, I do not find evidence of pre-trends of treated and control cities and the negative impact of rails on congestion, travel time and pollution is gradual, given that its impact increased over time.

Thus, the empirical analysis shows that light rail systems have been successful on average in reducing the negative externalities associated with car traffic. To be more precise, light rail systems have been successful in reducing car negative externalities in mid-size European cities. Up to this point, the main conclusion that can be derived from previous literature is that public transport impacts are specific to the context. In this study, the sample is mostly based on relatively rich cities with a high use of public transport. In such context, light rails seem to be helpful in making urban mobility more sustainable.

However, several caveats must be mentioned in relation to the positive effects of light rail systems. First, the econometric analysis focuses on the average effects but the data reveal significant differences between cities that experienced large increases in rail supply in the study period. Thus, rail systems may not be so helpful in reducing the negative externalities associated with car traffic if they are not designed to maximize annual ridership.

Second, my analysis focuses on the extensive margin as the econometric exercise exploits the variability that comes from big changes in the supply of rails. In cities with a dense urban rail network, it is not clear that marginal increases in supply would have substantial effects in terms of reducing the weight of cars on urban mobility. In such cases, other policies may be necessary such as congestion tolls, low emission zones or parking restrictions.

Third, my analysis provides evidence of the short-term effects of light rail systems but it is not clear whether the positive effects that I find would hold for a longer period, as better public transport can attract economic activity and hence car traffic.

In addition, building and maintenance costs of light rail systems are high in comparison with bus rapid transit systems. Therefore, a complete social evaluation of rail investments should consider whether similar reductions in negative externalities that come from car traffic could have been achieved with lower costs.

Finally, I cannot rule out the existence of time-varying unobservable differences across cities. In particular, rail investments may be part of a package of measures to reduce car negative externalities, so the estimated effect could also include the impact of these other measures.

Appendix A:. Data details

Congestion. The congestion variable measures additional time as the percentage that a vehicle needs for any trip in the city compared to a situation characterized by free traffic flow. TomTom obtains data from anonymous drivers' travel time from every city where it is active. Based on actual GPS-based measurements for each city, data from local roads, arterial roads and highways is registered. A *baseline* of travel times was established under uncongested and free-flow conditions across each road segment in each city. Second, the *actual* average travel times were calculated considering the entire year (24/7) and every vehicle in the city network. Speed measurements were used to compute travel times on individual road segments and over the entire city network. A weight was then applied considering the number of measurements, so that busier and more important roads in the network have a higher influence on the city's congestion level. Finally, the *baseline* and *actual* travel times were compared to compute the *extra travel time*. Hence, the congestion index represents the extra travel time experienced by drivers due to traffic conditions.

Table A1
Determinants of the probability of being treated (baseline year).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log (congestion)	0.29 (1.08)	–	–	–	−1.00 (1.17)	–	–
Log (travel time)	–	2.50 (5.82)	–	–	–	−3.81 (7.35)	–
Log (pollution)	–	–	−0.32 (1.13)	–	–	–	−1.67 (2.04)
Log (Population)	–	–	–	1.23 (1.29)	2.26 (1.34)*	2.11 (1.31)	0.43 (1.35)
Log (Density)	–	–	–	0.19 (0.62)	−0.02 (0.53)	−0.02 (0.53)	0.89 (0.77)
Log (Income per capita)	–	–	–	0.12 (1.36)	−1.47 (1.18)	−1.40 (1.17)	1.82 (2.27)
CDD	–	–	–	0.003 (0.004)	–	–	−0.001 (0.007)
HDD	–	–	–	0.0002 (0.001)	–	–	−0.001 (0.001)
Log (*Rain*)	–	–	–	−1.21 (1.14)	–	–	−3.28 (2.00)*
R^2	0.22	0.22	0.13	0.12	0.08	0.08	0.06
# Cities	43	43	43	43	43	43	43

Notes: Standard errors in parentheses (robust to heterocedasticity). Statistical significance at 1% (***), 5% (**), 10% (*).

Pollution. The pollution variable was based on annual mean estimates of PM 2.5, i.e., particulate matter with aerodynamic diameters smaller or equal to 2.5 μm. This is one of the main indicators of urban pollution, given that they imply serious effects on human health due to a composition that is rich in very toxic compounds and their great capacity of penetration in the respiratory tract. PM 2.5 are associated with the exacerbation of respiratory alterations such as bronchitis and cardiovascular diseases. This type of pollution is linked to increases in the morbidity and mortality of the exposed population and the growing development of asthma and allergies among children. In addition, as these particles are very light, they generally remain in the air for long periods. Pollution data rely on the method outlined in van Donkelaar et al. (2019). I focused on the European subset that provides estimates between 33 and 80°north and −15 and 45°east, at 0.1 × 0.1°resolution (about 10 km × 10 km). PM 2.5 is one of the main pollutants because it penetrates into sensitive regions of the respiratory system and can cause or aggravate cardiovascular and lung diseases. PM 2.5 emissions from road transportation come from two sources: *i) exhaust emissions* produced primarily from the combustion of petroleum products (such as gasoline or diesel), and *ii) abrasion emissions* produced by the mechanical abrasion and corrosion of vehicle parts (such as the vehicle's tires, brakes and clutch; road surface wear; or the corrosion of the chassis, bodywork and other vehicle components). Exhaust emissions are strongly related with the type of vehicle and, particularly, with its efficiency in petroleum consumption. Instead, abrasion emissions are independent of the vehicle type. According to publicly available data from the European Environmental Agency, both sources contribute similarly to the total emissions' volume caused by road transportation.

HDD. The HDD variable measures cold severity in a specific time period, considering *outdoor* and *average room* temperature. The calculation of *HDD* relies on the *base* temperature, defined as the lowest daily mean air temperature not leading to indoor heating. Although the value of the base temperature depends on several factors associated with the building and the surrounding environment, a general climatological approach is adopted in building this index and it is set at 15 °C. If T_m^i denotes the mean air temperature of day i (measured in °C), then the *HDD* of a certain year is given by

$$\boldsymbol{HDD} = \left\{ \begin{array}{ll} \sum_i^I 18 - \boldsymbol{T}_m^i & \mathbf{for} \boldsymbol{T}_m^i \leq 15 \\ 0 & \mathbf{for} \boldsymbol{T}_m^i > 15 \end{array} \right\}$$

where I denotes the number of days in the study year. For example, if the daily mean air temperature is 12 °C, the value of the *HDD* index for that day is 6 (i.e., 18 °C–12 °C). Instead, if the daily mean air temperature is 16 °C, the *HDD* index for that day is 0.

CDD. The CDD variable measures heat severity in a specific period, considering *outdoor* and *average room* temperature. As before, the calculation of *CDD* relies on the *base* temperature, which is now defined as the highest daily mean air temperature not leading to indoor cooling and is set at 24 °C (by adopting a general climatological approach). Then, the *CDD* of a certain year is given by

$$\boldsymbol{CDD} = \left\{ \begin{array}{ll} 0 & \mathbf{for} \boldsymbol{T}_m^i < 24 \\ \sum_i^I \boldsymbol{T}_m^i - 21 & \mathbf{for} \boldsymbol{T}_m^i \geq 24 \end{array} \right\}$$

For example, if the daily mean air temperature is 26 °C, the value of the *CDD* index for that day is 5 (i.e., 26 °C–21 °C). Instead, if the daily mean air temperature is 22 °C, the *CDD* index for that day is 0.

Rain. The variable of rain measures the amount of precipitation. Data is for the precipitation sum with unit 0.01 (mm) and were obtained from the European Climate Assessment dataset website that uses the methodology developed by Klein Tank et al. (2002).

Appendix B:. Additional table

(See Table A1)

References

Anderson, M., 2014. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. American Economic Review 104 (9), 2763–2796.
Angrist, J.D., Pischke, J.S., 2009. Mostly Harmless Econometrics. Princeton University Press, New York.
Adler, W., van Ommeren, J.N., 2016. Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes. Journal of Urban Economics 92, 106–119.
Atmospheric Composition Analysis Group (2020). Data accessible at: http://fizz.phys.dal.ca/~atmos/martin/?page_id=140. Retrieved July 20, 2020.
Barth, M., Boriboonsomsin, K., 2008. Real-world carbon dioxide impacts of traffic congestion. Transportation Research Record: Journal of the Transportation Research Board 2058, 163–171.
Bauernschuster, S., Hener, T., Rainer, H., 2017. When labor disputes bring cities to a standstill: the impact of public transit strikes on traffic, accidents, air pollution, and health. American Economic Journal: Economic Policy 9 (1), 1–37.
Baum-Snow, N., Kahn, M., 2005. Effects of Urban Rail Transit Expansions: Evidence from Sixteen Cities Brookings-Wharton Papers on Urban Affairs, 147-206.
Baum-Snow, N., Kahn, M.E., 2000. The effects of new public projects to expand urban rail transit. Journal of Public Economics 77 (2), 241–263.
Beaudoin, J., Lawell, C.L., 2018. The effects of public transit supply on the demand for automobile travel. Journal of Environmental Economics and Management 88 (8), 447–467.
Beaudoin, J., Lin Lawell, C.-Y.C., 2017. Is public transit's "green" reputation deserved?: Evaluating the effects of transit supply on air quality. Working paper. University of California at Davis.
Beaudoin, J., Hossein Farzin, Y., Lawell, C.L., 2015. Public transit investment and sustainable transportation: A review of studies of transit's impact on traffic congestion and air quality. Research in Transportation Economics 52, 15–22.
Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics 119, 249–275.
Carpintero, S., Siemiatyck, M., 2016. The politics of delivering light rail transit projects through public-private partnerships in Spain: A case study approach. Transport Policy 49, 159–167.
Chen, Y., Whalley, A., 2012. Green infrastructure: the effects of urban rail transit on air quality. American Economic Journal: Economic Policy 4 (1), 58–97.
Commission, European, 2019. Sustainable transport infrastructure charging and internalization of transport externalities: Main findings. Directorate-General for Mobility and Transport, European Commission.
Eurostat, 2020. Data accessible at: https://ec.europa.eu/eurostat/data/database. Retrieved July 20, 2020.
Gendron-Carrier, N., Gonzalez-Navarro, M., Polloni, S., Turner, M.A., 2018. Subways and Urban Air Pollution. NBER Working paper 24183, 1–48.
Gertler, P.J., Martinez, S., Premand, P., Rawlings, L.B., Vermeersch, C.M.J., 2016. Impact Evaluation in Practice. Inter-American Development Bank and World Bank, Washington, DC.
Gu, P., He, D., Chen, Y., Zegras, P.C., Jiang, Y., 2019. Transit-oriented development and air quality in Chinese cities: A city-level examination. Transportation Research Part D: Transport and Environment 68, 10–25.
Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálsdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A.F.V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., López, J.A., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V., Petrovic, P., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. International Journal of Climatology 22, 1441–1453.
Knowles, R.D., Ferbrache, F., 2016. Evaluation of wider economic impacts of light rail investment on cities. Journal of Transport Geography 54, 430–439.
Lee, S.S., Senior, M.L., 2013. Do light rail services discourage car ownership and use? Evidence from Census data for four English cities. Journal of Transport Geography 29, 11–23.
Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015. The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525, 367–371.
Li, S., Liu, Y., Purevjav, A.O., Yang, L., 2019. Does subway expansion improve air quality? Journal of Environmental Economics and Management 96, 213–235.
Mayer, T., Trevien, C., 2017. The impact of urban public transportation evidence from the Paris region. Journal of Urban Economics 102, 1–21.
Mulley, C., Hensher, D.A., Cosgrove, D., 2017. Is rail cleaner and greener than bus? Transportation Research Part D: Transport and Environment 51, 14–28.
TomTom, 2020. TomTom traffic index. Data accessible at https://www.tomtom.com/en_gb/trafficindex. Retrieved July 2020.
Urban rail, 2020. Data accessible at: http://www.urbanrail.net/. Retrieved July 2020.
Parry, W.H., Walls, M., Harrington, W., 2007. Automobile externalities and policies. Journal of Economic Literature 45, 373–399.
van Donkelaar, A., Martin, R.V., Li, C., Burnett, R.T., 2019. Regional estimates of chemical composition of.ne particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. Environmental Science and Technology 53, 2595–2611.
Vuk, G., 2005. Transport impacts of the Copenhagen metro. Journal of Transport Geography 13 (3), 223–233.
World Metro Database, 2020. Data accessible at http://micro.com/metro/table.html. Retrieved July 2020.
Yang, J., Chen, S., Qin, P., Lu, L., Liu, A.A., 2018. The effect of subway expansions on vehicle congestion: Evidence from Beijing. Journal of Environmental Economics and Management 88, 114–133.
Zhang, Y., Zheng, S., Sun, C., Wang. R., 2017. Does subway proximity discourage automobility? Evidence from Beijing. Transportation Research Part D: Transport and Environment 52 Part B, 506-517.
Zheng, S., Zhang, X., Sun, W., Wang, J., 2019. The effect of a new subway line on local air quality: A case study in Changsha. Transportation Research Part D: Transport and Environment 68, 26–38.