Year: 2018

Does expanding regional train service reduce air pollution?

Lalive, Rafael ; Luechinger, Simon ; Schmutzler, Armin

DOI: https://doi.org/10.1016/j.jeem.2017.09.003

Originally published at:
Lalive, Rafael; Luechinger, Simon; Schmutzler, Armin (2018). Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management, 92:744-764.
DOI: https://doi.org/10.1016/j.jeem.2017.09.003

Journal of Environmental Economics and Management ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Journal of
Environmental Economics and Management

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

Does expanding regional train service reduce air pollution?

Rafael Lalive [a], Simon Luechinger [b], Armin Schmutzler [c,*]

[a] University of Lausanne, CEPR, CESIfo, and IZA, Switzerland
[b] University of Lucerne and KOF Swiss Economic Institute, Switzerland
[c] University of Zurich and CEPR, Blümlisalpstrasse 10, 8006 Zurich, Switzerland

A R T I C L E I N F O

Article history:
Received 3 June 2016
Received in revised form
8 September 2017
Accepted 8 September 2017

JEL classification:
Q53
R41
R48

Keywords:
Railways
Pollution
Subsidies
Transportation
Environmental policy

A B S T R A C T

We assess how regional rail service affects air pollution in Germany, where rail service is procured in auctions or negotiations. We argue that the procurement mode is plausibly exogenous, and show that auctions deliver stronger rail service growth than negotiations. Instrumenting rail service growth with procurement mode, we find that increasing rail service by 10% reduces carbon monoxide and nitrogen oxide pollution by around 1% and 2%, respectively. Sulfur dioxide and ozone, pollutants with no clear link to rail, are not affected. Expanding rail service reduces car and motorcycle use, especially on leisure and shopping trips. The effects of railway services on car and motorcycle use and on air pollution build up over time. Lives saved by reducing pollution via rail service growth are worth substantially more than the required subsidies.

1. Introduction

Road transportation causes a large fraction of local and global pollution, accidents, and congestion. Countries all over the world address these externalities with a wide range of policies, including support for public transportation. Yet, not much is known about the environmental effects of improvements in public transportation. Public transportation can only reduce air pollution if improvements in public transportation induce substitution of traffic from roads to public transportation and if public transportation generates less air pollution than road transportation.

Both conditions, substitution from road to rail and superior environmental performance of railways, are potentially questionable. First, it is neither clear how strongly rail patronage reacts to improved services nor whether an increase in patronage necessarily comes at the expense of road transportation rather than from additional trips. Second, while engineering studies show that public transportation performs better than road transportation with respect to many environmental externalities (see Section 2), such comparisons usually assume otherwise identical conditions (e.g., speed, capacity utilization, etc.). It is again not obvious that this assumption is adequate to evaluate concrete policy measures. For instance, in rural regions, the number of passengers per train may be so small that the emissions of most major pollutants could well be higher than if these passengers had traveled by cars. Hence, a public transport policy that targets rural regions might lead to increases in pollution.

* Corresponding author.
E-mail address: armin.schmutzler@uzh.ch (A. Schmutzler).

http://dx.doi.org/10.1016/j.jeem.2017.09.003

In Germany, the average frequency of services on regional passenger lines grew by 28% between 1994 and 2004. During the same period, most types of local air pollution declined. Yet clearly, this simple before-and-after comparison says little about the causal effect of expanding regional passenger services. This decline may reflect improvements in emission control of vehicles which are unrelated to improvements in public transport. Also, Germany witnessed changes in the economic and political landscape and the implementation of important environmental policies in the nineteen nineties.

Our paper investigates how expanding the frequency of regional rail service affects air pollution, a joint test of substitution effects and favorable environmental performance of railways. We collect data on the evolution of the service frequency on all regional passenger railway lines in Germany in 1994 and 2004.[1] We then combine this information on public transport capacity with county-level data on pollution: carbon monoxide (CO), nitrogen oxides (NO_x), sulfur dioxide (SO_2), and ground-level ozone (O_3). To assess effects on other health-relevant pollutants, we also consider infant mortality rates.[2]

To isolate the effects of passenger railway support on air quality, we start from the idea that areas with greater improvements in railway services should also experience a more favorable evolution of air quality. However, there is an obvious endogeneity problem: Public transport may have been supported most in regions in which authorities expected an increase in traffic volume and, thus, a deterioration of air quality.[3] The positive effects of an increase in the frequency in rail services on air quality would then have been obscured by the adverse underlying trend.

We address endogeneity of rail service using a railway reform in 1994. This reform allowed regional passenger service agencies to choose between procuring the service competitively via public tenders or negotiating with the incumbent supplier. Interestingly, lines that were procured competitively at some point between 1994 and 2004 grew more strongly than lines served after negotiations, by 52% compared to 23%.[4]

We use the procurement mode as an instrument for frequency of service along a rail line. Our analysis supports the hypothesis that improvements in the quality of local passenger transportation cause improvements in air quality. We find that a 10% increase in the frequency of service reduces CO and NO_x pollution and infant mortality by 1-2%. These effects are imprecisely estimated and only the reduced form effect for NO_x is statistically significant. However, one important reason for this is that the effects of improved railway service need time to materialize. Apparently, transportation demand takes time to adjust to the improved supply.

Our empirical strategy requires that the procurement mode is unrelated to other changes in air pollution along train lines and affects pollution only through increased service frequency. We offer five arguments in support of this key assumption. First, the agencies in charge of selecting the procurement mode have no say in other policies affecting our outcomes, suggesting that environmental considerations were not taken into account when choosing lines to be auctioned (Section 3 for details).[5] Second, the institutional framework and anecdotal evidence suggests that procurement has no effects on air pollution via changes in ticket prices or quality aspects other than service frequency (Section 3 for details). Third, levels of air pollution (though not infant mortality rates) and changes in air pollution and infant mortality rates are similar for competitive and non-competitive lines before the reform (Section 6 for details). Fourth, our main results are similar, but statistically significant for NO_x, with an additional instrument based on track sharing with long-distance trains in 1994 (Sections 4 and 6 for details). We test joint exogeneity of the two instruments and find that the null hypothesis is not rejected. Fifth, we find no effects on two pollutants, SO_2 and O_3, for which we would not expect a clear effect of the quality of railway services (Section 6 for details). Taken together, this pattern of evidence is consistent with procurement mode being a valid instrument for rail service growth.

As discussed before, the effects of auctions on CO, NO_x, and infant mortality are larger the more time has elapsed since the auction. Using both the procurement mode and the time of competitive procurement as instruments, we obtain estimates that are very similar to the baseline results, but statistically significant for CO and NO_x. Further, we generally find stronger effects for lines running through cities and lines where the density of train stations is high. Summing up, our analysis supports the idea that improvements in the quality of railway services can substantially reduce road traffic externalities.

Our results indicate that expanding rail service reduces air pollution. Did households really substitute from road to rail? We use survey data from the German Socio Economic Panel on travel mode to assess substitution. We find that car and motorcycle use for shopping and leisure activities fell more sharply along lines with a larger increase in regional passenger trains. These results indicate that expansions of local passenger railways reduce air pollution due to substitution from cars and motorcycles to trains. Again, these effects are increasing in the number of years since the auction.

The paper is organized as follows. Section 2 discusses the related literature. Section 3 provides a brief account of the relevant institutions. Section 4 introduces the data and Section 5 the framework for the empirical analysis. Section 6 contains the results. Section 7 provides rough estimates of the monetary value of the improvements. Section 8 concludes.

[1] Lalive et al. (2016) also use these data to estimate the effects of procurement mode on frequency of rail service.

[2] Luechinger (2009) and Luechinger (2014) use data on SO_2 pollution at the county level for Germany in the period 1985-2003 to estimate the general effects on well-being and infant mortality, respectively, without discussing the contribution of transportation.

[3] Indeed, Section 16 in LNVG (2010) clearly indicates that favorable demand projections are decisive for supply increases.

[4] Lalive and Schmutzler (2008) provide a similar result for the state of Baden-Württemberg, Lalive et al. (2016) refine the analysis and extend it to the entire country, arguing that the correlation between the mode of procurement and the frequency of service can be given a causal interpretation.

[5] Consistent with this, we find no correlation between the procurement mode of a line and the left vote share in the federal election.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

2. Related literature

A large literature deals with the effects of transportation policies on the environment (see Schmutzler, 2011 for a review). To relate our contribution to this literature, note that our paper tests whether the following conditions hold jointly:

(C1) Improvements in public transportation induce a substantial substitution of traffic from roads to public transportation.
(C2) Public transportation involves lower externalities than road transportation per passenger transported.

Moreover, our results with survey data suggest that (C1) holds individually. We have grouped the existing literature according to (C1) and (C2) as follows.

(C1) Many papers attempt to quantify the effect of public transport improvements on ridership and the induced reductions in car transportation. Typically, these studies either suffer from endogeneity problems, or they estimate short-term elasticities, ignoring the long-term adjustments. Evans (2004) reviews evidence on the elasticities of rail ridership with respect to service frequency and reports elasticities of between 0.5 and 0.9. He also discusses findings of experiments from the nineteen sixties in the Boston region. These suggest that most additional rail users previously used their own car rather than carpools or buses.[6] By contrast, Duranton and Turner (2011) find no effects of expanding peak bus service on vehicle kilometers traveled on interstate highways across metropolitan statistical areas in the U.S. However, they look at traffic within metropolitan areas in the U.S. – a context in which bus service is marginal.

(C2) Engineering studies provide “emissions coefficients” which give information on the emissions characteristics of cars and railways. Together with information on capacity utilization, such information allows comparing the emissions per passenger kilometer. For instance, IFEU (2010) provides such a comparison for Germany, while INFRAS (2010) provides a detailed analysis for automobiles. The analysis broadly supports the view that, for most pollutants, the specific emissions for railways are lower than those for cars. It also shows, however, that there are substantial differences across pollutants. The problem with emissions coefficients is that they depend highly on the specific context, such as the type of road, topography, traffic situation and driving behavior. Also, they are typically average measures and not the marginal measures that are relevant when policy effects are studied.

Existing studies therefore support the notion that each of the statements (C1) and (C2) holds individually. Chen and Whalley (2012) test (C1) and (C2) jointly. They find that the introduction of a mass transportation system in Taipei in 1996 significantly lowered CO pollution, reduced NO_x pollution, but, as expected, had only small effects on O_3. Their regression discontinuity approach is ideally suited to capture the short-term effects of an abrupt change in passenger services. Our approach is complementary in that it allows us to capture longer-term effects. Moreover, our paper deals with a nation-wide service improvement in an existing system rather than the introduction of one completely new metro line. The fact that we do not only focus exclusively on agglomerations may be relevant because of differences in road-rail substitution as well as in the environmental effects of substitution.[7] Environmental effects between agglomerations and less populated areas could differ because of different baseline pollution levels. Indeed, we find some evidence for heterogeneous effects (see Section 6). Bauernschuster et al. (2017) test (C1) and (C2) jointly by analyzing public transport strikes. They show that strikes increase road traffic and particulate and NO_2 pollution. Strikes allow credibly assessing the relationship between public transport and road traffic externalities in the short run, but they are not informative about long-run effects either.[8]

Another strand of literature investigates whether pollution reductions have positive effects on infant health (for reviews, see Graff Zivin and Neidell, 2013; Currie et al., 2014). Studies on traffic-related pollution are particularly relevant for our context. Using short-term local pollution fluctuations in California and New Jersey, Currie and Neidell (2005) and Currie et al. (2009) find that infant mortality increases with CO, but not with particulate pollution and O_3. To estimate the effects on infant mortality, Knittel et al. (2016) exploit local air pollution fluctuations in California caused by changes in weather and traffic situations, whereas Arceo et al. (2016) rely on pollution peaks induced by weather changes in Mexico. Both studies find effects for CO and particulate pollution, though the effect for CO is not statistically significant in Knittel et al. (2016). Bauernschuster et al. (2017) find adverse effects of public transport strikes on the health of young children. Finally, focusing on long-run pollution changes, Currie and Walker (2011) show how the launch of electronic toll collection improved the health of infants in the immediate vicinity by reducing road congestion and emissions.

Our paper contributes to the existing literature in four ways. First, we use the type of procurement as an instrument to identify the policy effects. Our results can be interpreted as the local average treatment effect from using competitive procurement rather than bilateral negotiations between the agency and the incumbent. This effect is relevant, as rail services are procured using those two modes in many countries. Second, this paper analyzes the effects of a nation-wide policy in a Western industrialized country with comparably low levels of air pollution. This setting, which is particularly relevant

[6] Balcombe et al. (2004, Table 9.25) report values of 30–68% for several new urban rail schemes. Parry and Small (2009) use diversion ratios of 0.7 (peak) and 0.6 (off-peak) for Washington D.C.

[7] There is mixed evidence on the relative size of the substitution effects in agglomerations and elsewhere (Evans, 2004; Wardman, 2012).

[8] Complementary work by Anderson (2014) and Adler and van Ommeren (2016) analyzes the channels of the adverse effects of strikes or, conversely, the beneficial effects of public transport in more detail. They identify a substantial effect of strikes on peak-time congestion that goes beyond the mere traffic increase.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

for similar countries, makes it harder to find effects. Moreover, it allows us to show how the environmental effects of improving railway services differ between rural regions and agglomerations and how the density of railway stations enhances the positive effects of increased railway services. Third, the paper directly identifies substitution from road to rail as the mechanism leading to reductions of externalities. Finally, we provide suggestive evidence on how the improvements develop over time.

3. Institutional background

Until 1994, railways in Germany were run by two vertically integrated state monopolists (*Deutsche Bundesbahn* and *Deutsche Reichs-bahn*), one in West Germany and one in East Germany. In 1994, a major railway reform became effective, which not only created *Deutsche Bahn AG* as a successor of *Deutsche Bundesbahn* and *Deutsche Reichsbahn,* but also affected regional passenger transportation in several important ways.

The overall budget for regional passenger transportation increased. Starting in 1996, the annual transfers from the central government to the 16 states amounted to around 6–7 billion EUR. The states can use the money to procure otherwise non-profitable railway services.[9] They assign responsibility for planning and procurement to agencies which act on the basis of state law.[10] On every local network a long-term contract defines the service level and the transfer payments from agencies to train operators. This contract is not necessarily the result of negotiations between the agency and the supplier, but it can also be procured competitively, i.e., in a procurement auction. In the simplest case, the agency pre-specifies the desired quantity of railway services and quality aspects such as properties of the vehicles; the bid submitted by each firm is the level of transfer payments they demand to procure the desired service.[11]

Procurement auctions have been used quite often between 1994 and 2004. Fig. 1 displays a map of all regional passenger railway lines shaded according to procurement mode. Bold lines were procured competitively (134 out of 559), whereas thin lines (425 out of 559) were procured in direct negotiations with the incumbent. Both lines that serve big cities and lines that do not were auctioned. Moreover, the map also shows that competitive and non-competitive lines co-exist within the same state. The empirical analysis will focus on this within-state variation. Further, on competitively procured lines, auctions took place at different points in time. This is another fact that we will exploit in our empirical analysis.

Competitive procurement has had a substantial impact on service frequency. Between 1994 and 2004, the average frequency of service on the lines with competitive procurement grew from 11,512 trains in 1994 to 17,526 in 2004, i.e., by 52%. The corresponding figures for lines without competitive procurement were 15,691 in 1994 and 19,261 in 2004 (23%). Thus, not only was there a substantial overall growth, but this was more pronounced on the lines that were exposed to competition than on the remaining part of the network (see also Lalive and Schmutzler, 2008; Lalive et al., 2016). This correlation between competitive procurement and frequency of service growth will play an important role in our identification strategy.

One might be concerned that procurement mode decisions are related to other policies affecting our outcome variables. Therefore, it is important to collect some facts about the decision-making authority regarding these policies. First, decisions on the quantity and quality of roads are taken at all levels of the jurisdictional hierarchy, depending on the type of road under considerations. The federal transportation ministry is responsible for motorways (*Bundesautobahnen*) and other major roads (*Bundesstrassen*).[12] The state governments set priorities for roads of intermediate importance.[13] They also define the duties of the local governments which typically decide on issues relating to smaller roads, parking policies, etc.[14] Second, the taxes that are most likely to affect road transportation (vehicle taxes and gasoline taxes) are set at the federal level. Finally, while decisions regarding air quality are made both at the federal and at the state level, federal decisions dominate in conflicting cases.

Thus, while the decisions on the procurement mode are made at the state level, the role of the states for other policies affecting transportation and the environment is much more limited. Also, even if the states are responsible, the agency in charge is not identical with the one that deals with procurement. All told, it seems unlikely that decisions on local railway procurement and other determinants of local environmental quality are made by the same parties and are therefore highly correlated.[15]

[9] The value for 2004 was approximately 6.8 billion EUR (SCI Verkehr GmbH, 2006). The funds cannot be used to procure long-distance passenger services (*Intercity* and *Intercity-Express*); the dominant firm *DB Fernverkehr* is expected to supply these services profitably.

[10] Almost all states have laws (*Nahverkehrsgesetze*) governing local public passenger transportation. This assignment of decision-making authority and financial responsibilities is unlikely to conform to public finance principles (Olson, 1969). However, in the following, we focus on the effects of increases in railway service frequency within the current system.

[11] In other cases, the firms can submit bids that contain not only transfer payments, but also some aspects of quality, which are then weighted in a suitable scoring rule.

[12] For instance, the ministry provides five-year plans for major investments (an example is the *Investitionsrahmenplan bis 2010 für die Verkehrsinfrastruktur des Bundes*) and decides on major projects.

[13] To this end, they use long-term plans (*Landesstrassenbedarfspläne*).

[14] Local governments typically formulate a long-term plan (*Verkehrsentwicklungsplan*) that identifies priorities for road infrastructure. They also lay down fee structures for parking, which are potentially of great importance for the development of passenger transportation. For instance paragraph 1(2) in the *Verordnung über Zuständigkeiten im Bereich Verkehr* of the State of Lower Saxony stipulates that local jurisdictions have the right to define parking fees.

[15] It is also important to note that the procurement decision is typically the result of negotiations between the agency and incumbent rather than a decision of the agency alone, and that the lines on which an agency wants an auction most are typically exactly those for which the incumbent wants it least: For instance, an agency benefits most from auctioning lines where competition is expected to be intense, whereas the incumbent loses most in such cases. These conflicting interests make it hard to predict how the outcome of the negotiations depends on line characteristics.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

One might also argue that the mode of procurement could affect the extent of substitution from road to rail through other channels than service frequency.[16] For instance, it is conceivable that competition might affect ticket prices. This could happen either because an incumbent firm applies different tariffs in cases where it has to win the market in competitive tendering or because the competitors who win a market set lower ticket prices. Specific features of the German market prevent this, however.[17] The German railway companies are required by law to work towards a unified tariff structure. In practice, this leads to very similar prices across companies. First, in many regions, local public transport organizations rather than railway companies set ticket prices which are identical regardless of the mode of procurement.[18] Second, outside those regions, competitors usually apply the tariffs of *DB Regio*.[19]

Similarly, the procurement mode could affect other dimensions of quality rather than just the frequency of service. However, Lalive et al. (2016) provide descriptive evidence for some of the larger agencies suggesting that this concern is not warranted. In particular, there seems to be no difference between the punctuality and reliability of trains on competitive and non-competitive lines.

4. Data

We combine railway data with county-level data on air pollution, infant mortality, and travel modes. As our procurement and service frequency data refer to railway lines, a data point will correspond to a particular railway line. For outcome variables, we first construct county level data. We then average over all counties through which a railway line runs to obtain a measure of the outcome at the level of the railway line.

4.1. Outcome variables

We have data on annual mean concentrations of CO, NO_x, SO_2, and O_3 from the *Umweltbundesamt* (UBA), the federal environmental agency. CO monitoring started only in 1997; we calculate the NO_x concentration as a weighted sum of NO and NO_2 concentrations with the relative molecular mass as weight.[20] We interpolate the monitor data with inverse distance weighting to a one square kilometer grid and then compute county averages.[21]

Our data on air pollution refer to an important subset of pollutants associated with road traffic. We therefore also analyze whether infant mortality is affected by changes in the frequency of rail service. Infant mortality is an indicator that will reflect all local pollutants that matter for infant health, not only those included in our analysis. Therefore, we view the infant mortality estimates as complementary to the pollution estimates. Infant mortality is reported as the number of deaths within one year after birth per 1000 live births. The infant mortality data are from the German Youth Institute based on information from the state statistical agencies.[22] We use the simple three year average 1993–1995 to estimate infant mortality in 1994 and the three year average 2003-2005 to assess infant mortality in 2004.

4.2. Train service

Our main explanatory variable is (the log of one plus) the frequency of service, defined as the number of passenger trains per year on a line. In tedious work, we collected this information on the basis of published railway timetables for the years 1994 and 2004, i.e., the last year before and the year one decade after the start of the reform process (Lalive et al., 2016).[23]

[16] On a related note, Mouwen and van Ommeren (2016) have analyzed the relation between procurement mode and passenger kilometers in the Dutch bus industry, suggesting that competition itself does not have a positive effect, whereas the threat of competition does.

[17] See Monopolkommission (2009), Section 2.4.3) for a detailed discussion.

[18] "[…] many agencies prescribe unified tariffs in their contracts […]"(Monopolkommission, 2009, p. 49, own translation).

[19] "This model has been replaced by a system of ticket price cooperation in which railway operators cooperate with Deutsche Bahn AG to jointly offer nationwide prices. […] Ticket price cooperation typically requires the application of the prices of Deutsche Bahn AG." (Monopolkommission, 2009, p. 49, own translation).

[20] Data are available for 310 (1997) and 207 (2004) monitors for CO, 323 (1994) and 410 (2004) monitors for NO, 404 (1994) and 410 (2004) monitors for NO_2, 513 (1994) and 265 (2004) monitors for SO_2, and 319 (1994) and 305 (2004) monitors for O_3. We aggregate the NO and NO_2 concentrations at the line level at the end of the data setup process.

[21] The method of inverse distance weighting requires the choice of two parameters: the number of monitors used in the interpolation and the power parameter. We have chosen these parameters so as to get a large correlation between actual readings at a monitor and the pollution concentration that would be predicted without this monitor (Currie and Neidell, 2005 suggest this approach to judge the accuracy of the interpolation procedure).

[22] Three corrections are necessary. First, zeroes are reported as missings in some years. Second, we eliminate observations with obvious coding errors where infant mortality rates decline with time since birth, e.g., where the rate for deaths in the first week is higher than for those in the first year. Third, for the early years, we use simple averages for those counties in the former East Germany that later merged.

[23] We decompose the rail network following the classification of lines in the official timetables, with two exceptions. First, to avoid double-counting of trains, we deleted parts of lines that are also contained in other lines. Second, to have a clear assignment of a line to a state, we divided lines that do not lie completely within one state.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

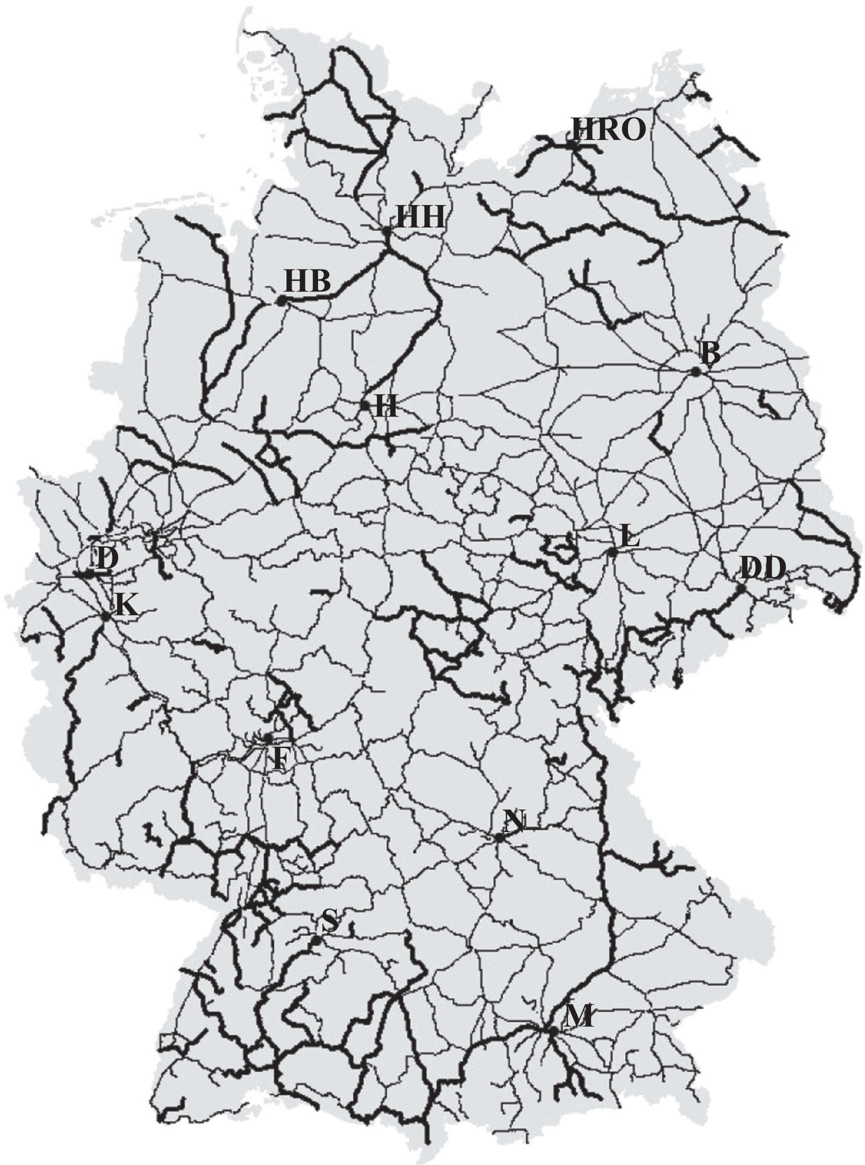

Fig. 1. Map of regional railway lines. *Notes:* Thin lines denote non-competitively procured train lines, bold lines competitively procured train lines, and points important cities: B: Berlin, D: Düsseldorf, DD: Dresden, F: Frankfurt (Main), H: Hanover, HB: Bremen, HH: Hamburg, HRO: Rostock, K: Cologne, L: Leipzig, M: Munich, N: Nuremberg, and S: Stuttgart.

4.3. *Instruments for train service*

We instrument train service with the procurement mode and capacity utilization by long-distance trains in 1994. Following Lalive et al. (2016), competitive lines are those with a share of competitively procured services of at least 20%.[24] The overwhelming majority of these cases correspond to competitive tenders which were either won by *DB Regio* itself or some competitor, but we also included cases where an ownership change took place for other reasons.[25] In extensions of our baseline results, we show how changes in outcomes differ with the time since an auction took place and use this duration in addition to the procurement mode as another instrument.

Further, we construct another instrument on capacity utilization by long-distance trains (ICE trains) in 1994, prior to the

[24] The precise value of this cut-off is unlikely to be important for the classification, because most lines were either served almost exclusively by competitively procured trains, or there was just a small number of such trains, usually extended from adjacent lines.

[25] We are grateful to Felix Berschin for information on competitive tenders; in addition, we obtained information on ownership changes from the *Deutsche Bahn* timetables.

rail service expansion we analyze. Such capacity utilization is, arguably, exogenous to regional rail service expansion, but it dampens regional rail service growth as lines have a fixed capacity. We thus expect this instrument to have a negative effect on regional rail service expansion. Specifically, we set a dummy for the capacity utilization by long-distance trains as one if at least six long-distance passenger trains per weekday (Monday to Friday) run on at least 25% of the line. We used the 1994 timetable to obtain the relevant number of long-distance passenger trains.

4.4. Travel mode

We use data from the German Socio Economic Panel (SOEP), an annual survey on about 11,000 households (see Wagner et al., 2007), containing information on the type of travel mode normally used for commuting, leisure activities (cinema, soccer game, etc.), shopping, weekend trips, and taking children to school. Individuals choose between the following travel modes: public transport, car, motorcycle, bicycle, or by foot. We aggregate their responses to the county level and calculate the probability of traveling by car or motorcycle conditional on using at least one of the travel modes. The closest year to the beginning of the reform for which information on travel mode is available is 1998, the closest year to the end of the period under investigation is 2003.[26] The SOEP does not include data from all counties in the respective years.

The variables on outcomes and travel mode are at the county level (*Landkreise*). As our data points correspond to lines which might run through several counties, we take simple averages of these variables across all relevant counties for a particular line.

4.5. Control variables

Control variables: The data also include controls for important line-level characteristics, i.e., distance to nearest large city, length of line, and information on electrification. All regressions include fixed effects of federal states.[27]

5. Empirical strategy

We now introduce a simple model from which we derive hypotheses on the determinants of pollution and the resulting effects on infant mortality.

This model will capture the relation between environmental outcomes at the line level (i.e., in all adjacent counties) and the frequency of service on the line, which is mediated through an effect of the frequency of service on car traffic. We exclude any effect on near-by counties through which the line does not run. Such effects could arise because improved services in one county reduce pollution concentration in neighboring counties. Focusing on counties along railway lines is justified for two reasons. First, counties are typically large and contributions of tailpipe emissions to pollution levels decline rapidly with distance from the source. The average county has a size of 818 km^2, implying an average radius of 16 km. At this distance, contributions of tailpipe emissions are reduced to 18% of their initial value, according to the distance decay function used by Knittel et al. (2016). Thus, it is unlikely that reduced road traffic has measurable effects in neighboring counties on average. Second, while we know the location of railway lines, we do not know where exactly road traffic decreases.

Let Y_{it}^s capture the concentration of pollutant s along a line i at time t. It is affected by total passenger transportation, T_{it}, the share of car traffic, CAR_{it}, and potentially by other variables. We assume that an increased frequency of service of trains, FOS_{it}, reduces car traffic. Specifically, we suppose the share of car traffic is

$$CAR_{it} = \frac{1}{(1 + FOS_{it})^\gamma}$$

where γ is a parameter to be estimated. Note that $CAR_{it} = 1$ if there are no train services, and CAR_{it} decreases in FOS_{it} if $\gamma > 0$.

The modal split is important because cars and trains differ in their pollution intensities. θ_s^C measures the emissions of pollutant s per unit of car traffic, θ_s^T is the corresponding parameter for train traffic. Thus, we can write the outcome variable as follows[28]:

$$Y_{it}^s = (\theta_s^C CAR_{it} + \theta_s^T (1 - CAR_{it})) T_{it} \exp(\alpha_i + \tau_{jt} + \varepsilon_{it}) \tag{1}$$

where α_i is a location specific unobserved effect, ε_{it} reflects other sources of emissions, τ_{jt} are time effects which we allow to differ across states j to capture changes in background pollution. The (natural) logarithmic version of this equation is

$$log Y_{it}^s = \log\left(\theta_s^C CAR_{it} + \theta_s^T (1 - CAR_{it})\right) + log T_{it} + \alpha_i + \tau_{jt} + \varepsilon_{it} \tag{2}$$

[26] In analyses relying on SOEP data, we therefore only classify a railway line as competitively procured if the auction took place in the years 1998–2002.

[27] We use only one fixed effect for Bremen and Lower Saxony and for Berlin and Brandenburg, respectively, because these states share railway agencies.

[28] Eq. (1) refers to emissions, not concentration levels. To get concentrations, we need to divide both sides of (1) by a measure of the volume of air that is affected by emissions of type s. However, as this volume does not change in time, it will cancel out after log-differencing.

In the empirical analysis, we rely on the following log-linear approximation of Eq. (2):[29]

$$\begin{aligned} logY_{it}^{s} &= \delta logCAR_{it} + logT_{it} + \alpha_i + \tau_{jt} + \varepsilon_{it} \\ &= -\delta\gamma log(1 + FOS_{it}) + logT_{it} + \alpha_i + \tau_{jt} + \varepsilon_{it} \end{aligned} \tag{3}$$

This approximation provides information on the relative contribution of cars and trains to environmental quality s. In particular, one can show that

$$\delta = \frac{\partial logY_{it}^{s}}{\partial logCAR_{it}} = \frac{(\theta_s^C - \theta_s^T)CAR_{it}}{\theta_s^T + (\theta_s^C - \theta_s^T)CAR_{it}},$$

which is zero if trains and cars contribute equally, and one if trains do not contribute at all. Thus, if the train-specific contribution to s were zero ($\theta_s^T = 0$), the coefficient associated with log frequency of service would directly measure the substitution parameter γ. When the train-specific contribution to the outcome variable is not zero, the coefficient relating outcome to log frequency of service should be smaller, so that γ provides an upper bound.

To remove the fixed region effect, we look at changes between $t-1 = 1994$ and $t = 2004$. Using $\Delta logY_i^s \equiv logY_{it}^s - logY_{it-1}^s$, $\Delta \log(1 + FOS_i) \equiv \log(1 + FOS_{it}) - log(1 + FOS_{it-1})$, $\Delta logT_i \equiv \log T_{it} - logT_{it-1}$, $\Delta\tau_j \equiv \tau_{jt} - \tau_{jt-1}$, and $\Delta\varepsilon_i \equiv \varepsilon_{it} - \varepsilon_{it-1}$, the difference specification corresponding to (3) is

$$\Delta logY_i^s = -\,\delta\gamma\Delta\log(1 + FOS_i) + \Delta logT_i + \Delta\tau_j + \Delta\varepsilon_i \tag{4}$$

We estimate specification (4) without changes in transport demand, $\Delta logT_i$, which we cannot observe.[30] The frequency of service is endogenous, as it is likely to be correlated with the change in transport demand. To address this issue, we add time invariant line characteristics to reflect differences in the evolution of transport demand along lines with different characteristics. More importantly, we instrument frequency of service with the procurement mode. We chose this instrument because competitive procurement predicts service frequency growth.[31]

For the procurement mode to be a valid instrument, it has to be independent of unrelated changes in environmental quality, working only through increases in service frequency. As explained in Section 3, the institutional structure makes it unlikely that the decisions on procurement mode are directly related to environmental policy decisions. Consistent with the latter point, the procurement mode of a line is unrelated to voter preferences as measured by the proportion of voters in favor of the left-wing parties in the 1994 federal election (see Table 3 below).[32] Further, Section 3 shows that competition for the market does not seem to affect ridership via reduced ticket prices or improved quality. In addition, in Section 6 we find that the procurement mode of a line is neither related to levels (except for infant mortality) nor trends in environmental quality before the reform. Moreover, it is not related to changes in pollutants for which no effects are to be expected. Finally, the results are robust to adding our other instrument related to long-distance train service. We also explore testing joint exogeneity of our procurement mode instrument and the other instrument.

We report two sets of standard errors to conduct inference. The first set of standard errors allows for arbitrary heteroskedasticity but assumes independence of error terms in rail service across train lines (White, 1980). In rail networks, independence is possibly violated as shocks to rail service on one line may spill over to adjacent lines, i.e., lines connected by at least one railway station. To address possible correlation across the network, we use standard techniques from spatial econometrics and calculate standard errors that explicitly allow for errors of adjacent lines to be correlated. We report tests based on both sets of standard errors and assess to what extent our statistical inference is robust to assuming independence.

6. Results

We first present descriptive statistics. After that, we discuss the results that rely on procurement mode as an instrument, assess the validity of this instrument and the robustness of our results to using our alternative instruments, and present placebo estimates. We then investigate the mechanisms by which public transport affects outcome variables. Finally, we examine how the effects change with the year of the auction and other line characteristics.

[29] Note that this log-linear Eq. (3) can be studied using standard linear least squares or instrumental variables methods. Estimation of the non-linear Eq. (2) is challenging since the non-linear part of (2) contains three parameters, but only one regressor, i.e., the equation is not identified. Moreover, the log-linear approximation (3) provides an estimate of how expanding regional passenger services translates into improvements of air quality. This is the relevant policy parameter.

[30] We do not explicitly model the sources of the level of transport demand. These could be plausibly related to GDP per capita and population. However, GDP and population are potentially endogenous to the expansion of regional passenger service.

[31] Lalive et al. (2016) show that the positive effect of competition on the frequency of services is the mirror image of a negative effect of competition on the procurement prices paid by the agencies: Where competitive procurement takes place, agencies expect to pay less than when there are direct negotiations with the incumbent. Accordingly, they are prepared to procure greater quantities. There are two reasons why agencies nevertheless do not procure all lines competitively: First, the large administrative costs of auctions and, second, the influence activities of the incumbent.

[32] Table 11 in the Appendix analyzes the determinants of the procurement mode, emphasizing in particular the negative role of electrification.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 1
Descriptive statistics. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, state statistical agencies, and the SOEP.

	N	Mean	Median	Std. Dev.
A. Outcomes (log-differences, 2004–1994)				
CO	559	−0.424	−0.444	0.233
NO_x	559	−0.095	−0.128	0.257
Infant mortality	559	−0.335	−0.297	0.349
B. Placebo outcomes (log-differences, 2004–1994)				
SO_2	559	−1.291	−1.079	0.608
O_3	559	0.087	0.089	0.073
C. Railway service (log-differences, 2004–1994)				
Frequency of service	559	0.282	0.236	0.367
D. Instruments				
Auction	559	0.240	0.000	0.427
Years since auction	134	3.918	4.000	2.317
Long distance trains 1994	559	0.054	0.000	0.226
E. Line characteristics				
Electric traction	559	0.456	0.000	0.499
Length (log km)	559	3.800	3.829	0.788
Distance to city (100 km)	559	0.192	0.000	0.301
F. Car and motorcycle use (differences, 2003–1998)				
Work	553	−0.022	−0.029	0.131
Leisure	554	−0.103	−0.110	0.119
Shop	555	−0.003	−0.010	0.101
Weekend	556	−0.130	−0.136	0.111
Children	546	0.092	0.071	0.219

Notes: This table reports descriptive statistics. Panel A shows our main outcomes, Panel B placebo outcomes, and Panel C our main explanatory variable, the frequency of service. All these variables are log-differences between 1994 and 2004. Panel D depicts our instruments. Auction takes the value 1 if the line was procured using an auction rather than in direct negotiations with the incumbent between 1994 and 2004. Years since auction is 2004 minus the year in which the auction took place. Long-distance trains in 1994 takes the value 1 if, in 1994, at least 6 long-distance high-speed trains (ICE trains) per day used at least 25% of the line. Panel E shows other line characteristics. Electric traction is 1 if the line has electric traction, 0 if it has diesel traction. Length is the log of length of the line in kilometers. Distance to city is 0 if the line runs through a city of at least 100,000 inhabitants in 1994 and equal to the distance (along the rail network) to the nearest city with at least 100,000 inhabitants otherwise. Panel F depicts changes in the share of individuals using the car or the motorcycle for an activity between 1998 and 2003.

6.1. Descriptive statistics

Panel A and B of Table 1 provide descriptive statistics on our dependent variables. Air quality has improved substantially for all measures except O_3. For instance, the median reduction of CO was 36% $(100(\exp(-0.444)-1)=-36)$. For SO_2, which is essentially unrelated to road transportation, the median reduction was 66%. Finally, infant mortality also decreased by 26%, a considerable improvement. For all variables, the standard deviation is substantial, indicating large regional variation.

Panel C of Table 1 contains our key explanatory variable, namely the evolution of regional passenger railway services. It identifies a substantial overall increase, with a median growth of 27%. The large standard deviation suggests the possibility that differences in the evolution of service quality may have contributed to the different developments in the outcome variables. Below, we will investigate whether this is indeed true.

Panel D of Table 1, provides information on the instruments. It shows that 24% of the railway lines (134 lines) are classified as competitive. On average, in 2004, four years have elapsed since the auction. In 1994, on about 5% of the lines, at least six long-distance trains per day used at least part of the line.

Panel E of Table 1 also shows that train lines differ with respect to traction, length, and distance to large cities. 46% of all railway lines are electrified. The mean length of a line is 3.8 log km, i.e., 45 km. Distance to city measures the remoteness of a line. Distance to city is 0 if the line runs through a city of at least 100,000 inhabitants in 1994. For railway lines that do not run through a city the distance is the length of the shortest passenger railway connection between a city and one of the stations on the line. The average distance is 19 km, the median is 0.

Finally, Panel F of Table 1 depicts reductions in the use of the car for most purposes, except for the transport of children.

6.2. Baseline results

Fig. 2 visualizes our identification strategy and anticipates our baseline results. It depicts the mean residual growth rates

ARTICLE IN PRESS

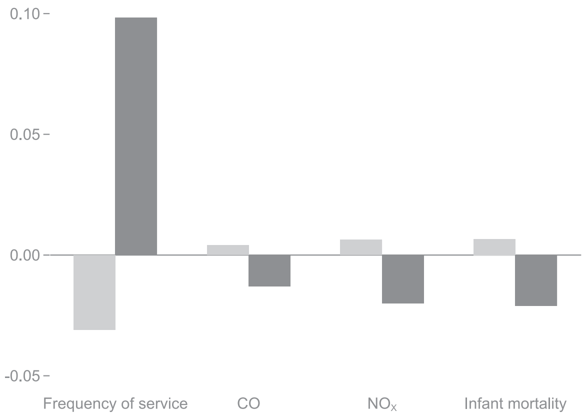

Fig. 2. Higher growth rates of frequency of service on auctioned lines, but lower growth rates of CO and NO_x pollution and infant mortality rates. *Notes:* Bars depict the mean residual change in the log of frequency of service, CO and NO_x pollution, and infant mortality rates on auctioned (dark grey) and non-auctioned (light grey) lines, respectively. The residual growth rates are corrected for the influence of electric traction, track length, distance to city, and state effects. See notes to Table 1 for the definition of the control variables. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the depicted variables across these line characteristics and states or characteristic- and state-specific time effects.

of frequency of service, CO and NO_x pollution, and infant mortality for auctioned (dark grey) and non-auctioned (light grey) lines, respectively. The growth rates are corrected for the influence of the important line characteristics electric traction, track length, and distance to city and of state effects. Thus, the residual growth rates indicate if a variable grew faster or slower between 1994 and 2004 on a particular line than on comparable lines in the same state. Fig. 2 demonstrates that auctioned lines experienced a larger increase in frequency of service than non-auctioned lines, which translated into a larger decrease in CO and NO_x pollution and infant mortality along these lines.

Table 2 presents our baseline results with the IV approach that uses the procurement mode as an instrument.[33] It displays both types of standard errors, with and without correlation between adjacent railway lines. The former are larger than the latter.

Column 1 in Panel A presents the first stage results. As already shown in Fig. 2, the frequency of service grew more strongly on lines that were competitively procured than on lines that were not. This result is both quantitatively important and statistically significant. The F-statistic on the instrumental variable is 12.143, suggesting that procurement mode is not a weak instrument. Service frequency increases 16% more strongly on competitively procured lines than under direct negotiations.

Columns 2–4 in Panel A present the reduced form estimates of procurement mode on the change in air pollution and infant mortality between 1994 and 2004. Results indicate negative effects on CO and NO_x concentrations and infant mortality, though the effect is only significant for NO_x and robust standard errors.

Panel B depicts the IV results. We find substantial, though statistically insignificant (p-value for NO_x with robust standard errors is 0.1004), effects for all outcomes. As will be shown later, adding additional instruments improves the precision of the estimates for CO and NO_x without having large effects on the point estimates. The point estimates for CO and NO_x indicate that a 10% increase in the frequency of service reduces concentration levels along the line by 1.2% and 1.9%, respectively. This finding is likely to reflect reductions in road traffic, which is a substantial source of CO and NO_x emissions.

Generally, our results are slightly weaker for CO than for NO_x. This could simply reflect the fact that data quality is inferior for CO than for NO_x in our case.[34] Alternatively, the effect for CO may indeed be weaker. Interestingly, previous evidence regarding CO is conflicting: While Chen and Whalley (2012) found a significantly negative effect on CO pollution of the introduction of a mass transportation system, Knittel et al. (2016) and Bauernschuster et al. (2017) found no significant effects of road traffic and public transport strikes, respectively.

6.3. Validation, robustness, and placebos

While the requirement that the procurement mode is unrelated to the evolution of environmental quality cannot be tested, we can test whether the outcome variables were different on competitive lines in 1994 and, more importantly, whether they evolved more favorably before the railway reform, in the 1990 to 1994 period. The analysis of pre-trends focuses on the evolution of NO_x and infant mortality, because CO data are not available. The results in Table 3 show that infant mortality was higher in 1994 on auctioned lines, but there was no difference between the 1994 pollution levels on competitive and non-competitive lines. Moreover, the change in NO_x concentration and infant mortality between 1990 and

[33] Table 10 in the Appendix shows that, as expected, there is no systematic relationship between frequency of service and environmental quality in OLS regressions.

[34] CO pollution is only available from 310 monitors in 1997 and 207 monitors in 2004, whereas NO_x is measured by 400 monitors or more (except for NO in 1994), providing a more disaggregate picture of local air pollution.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 2
Baseline results. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, and state statistical agencies.

	Frequency of service	CO	NO_x	Infant mortality
A. First stage and reduced form				
Auction	0.149	−0.019	−0.030	−0.032
	(0.043)***	(0.013)	(0.017)*	(0.033)
	[0.049]***	[0.015]	[0.020]	[0.035]
Electric traction	−0.020	−0.011	0.018	−0.066
	(0.034)	(0.014)	(0.013)	(0.030)**
	[0.038]	[0.015]	[0.015]	[0.037]*
Length (log km)	−0.067	−0.001	−0.018	−0.016
	(0.024)***	(0.010)	(0.008)**	(0.019)
	[0.027]**	[0.011]	[0.011]*	[0.021]
Distance to city (100 km)	−0.027	0.041	−0.089	−0.037
	(0.064)	(0.028)	(0.028)***	(0.070)
	[0.074]	[0.043]	[0.042]**	[0.098]
F-stat instruments	12.143			
Adj. R-squared	0.062	0.635	0.661	0.121
B. IV estimates				
Frequency of service		−0.131	−0.203	−0.214
		(0.094)	(0.124)	(0.227)
		[0.107]	[0.159]	[0.252]
Overall significance (p-value)		0.000	0.000	0.000
Train lines		559	559	559

Notes: This table reports the results of OLS regressions of the change in the log of frequency of service, CO and NO_x pollution, and infant mortality between 1994 and 2004 on the procurement mode of the line in Panel A. The change in the log of frequency of service is instrumented using the procurement mode of the line in Panel B. See notes to Table 1 for the definition of the instrument and control variables. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

1994 was the same along both types of lines. This is consistent with the main identifying assumption that air quality would have improved in the same fashion on competitively procured lines as on the remaining lines.

Table 4 presents estimates with capacity utilization by long-distance trains in 1994 as an additional instrument. Column 1 in Panel A presents the first stage, Columns 2 to 4 the reduced form. Interestingly, lines that were frequently served by long-distance trains in 1994 grew 7% less strongly over the 1994 to 2004 period, consistent with the intuition that long-distance trains act as a capacity constraint. The Kleibergen-Paap (2006) weak identification statistic, a generalization of the standard F statistic to a setting with heteroskedasticity, is 7.64. This value is low, indicating that our set of instruments may be weak. The overidentification test does not reject the null that the instruments are valid. The effects on pollution are similar to the baseline results with just one instrument (Table 2), but slightly more precisely estimated.

Table 5 contains the results of placebo regressions for pollutants for which it is implausible that road traffic should have a clear effect on overall concentration. First, we consider SO_2 pollution. As diesel engines in cars and locomotives produce SO_2 emissions, one might expect a positive or negative effect of road-rail substitution on SO_2 pollution. However, direct emissions from passenger transportation are a negligible source of SO_2 emission. For instance, in 2000, *total* road transportation was responsible for 3.1% of all SO_2 emissions in Germany according to the UBA; the corresponding figure for *total* other transportation was 2.4%. It is thus safe to assume that the contribution of each mode of regional passenger transportation was below 2%. This would make it highly unlikely that the supply of passenger railway services could have had a non-negligible (positive or negative) effect on SO_2 pollution.

Second, we consider O_3. While NO_x is an important precursor of O_3, this does not imply a positive monotone relation between NO_x emissions and O_3 concentration. The point is made clearly by Geddes et al. (2009) who argue that, among many

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 3
Validation of auction instrument. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, state statistical agencies, and the Federal Returning Officer.

	Levels, 1994				Log-differences, 1994–1990	
	CO	NO_x	Infant mortality	Left vote share	NO_x	Infant mortality
Auction	0.005	−0.897	0.275	−0.006	0.004	0.019
	(0.010)	(1.312)	(0.115)**	(0.004)	(0.015)	(0.024)
	[0.012]	[1.692]	[0.128]**	[0.005]	[0.019]	[0.025]
Electric traction	0.022	2.940	0.123	0.005	0.001	−0.016
	(0.009)**	(1.110)***	(0.094)	(0.004)	(0.013)	(0.020)
	[0.010]**	[1.217]**	[0.100]	[0.004]	[0.016]	[0.021]
Length (log km)	−0.010	−1.676	0.180	−0.007	−0.014	0.030
	(0.006)*	(0.673)**	(0.062)***	(0.003)**	(0.008)*	(0.014)**
	[0.007]	[0.836]**	[0.062]***	[0.003]**	[0.011]	[0.014]**
Distance to city (100 km)	−0.098	−16.597	0.348	−0.037	−0.068	0.005
	(0.016)***	(2.037)***	(0.257)	(0.008)***	(0.032)**	(0.042)
	[0.027]***	[3.570]***	[0.311]	[0.011]***	[0.046]	[0.049]
Adj. R-squared	0.736	0.752	0.178	0.784	0.516	0.030
Train lines	559	559	559	559	559	559

Notes: This table reports the results of OLS regressions of the 1994 level of CO, NO_x, infant mortality, left vote share, and the change between 1990 and 1994 in the log of NO_x and infant mortality on the procurement mode of the line and line characteristics. CO data is not available before 1997. The left vote share is the vote share (second vote) of the Social Democratic Party (SPD), the Greens (GRÜNE), and the Party of Democratic Socialism (PDS) in the 1994 federal elections. See notes to Table 1 for the definition of the instrument and control variables. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects in regressions in columns 5 and 6.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

other factors, the effect of an increase in NO_x depends on the initial concentration: O_3 production can be inversely proportional to NO_x levels (for similar arguments, see Han et al., 2011). This observation is closely related to the fact that NO_x is not only important for O_3 formation, but also for its decay: Once O_3 has been formed, it can react with NO to form NO_2 and O_2.

Neither SO_2 nor O_3 is significantly affected by the expansion of regional train service, reinforcing the plausibility of the main results.[35]

6.4. Effects on modal choice

The analysis so far suggests that increasing the frequency of service on regional passenger train lines is beneficial for the environment. We have interpreted this as evidence of substitution from cars to trains. In this section, we provide direct support for this interpretation. Table 6 presents IV estimates of the effect of increasing frequency of regional passenger train service on the change in the probability of using cars or motorcycles for commuting, leisure activities, shopping, weekend trips or taking children to school. Results indicate that an increase in service frequency of 10% reduces the probability of using cars or motorcycles for shopping or leisure activities (e.g., watching a movie or soccer game) by 1.6 and 2.3 percentage points, respectively.[36]

We now use a simple thought experiment to argue that the results on modal choice are high, but in the right order of magnitude. To see this, our point estimates suggest that a 28% increase in the frequency of rail services would reduce the mode share of road transportation for commuting, leisure activities, shopping, and weekend trips by 3.6, 5.9, 4.2, and 0.05 percentage points, respectively, and increase this share for taking children to school by 6.7 percentage points. Assuming that these activities are responsible for 31% (commuting), 25% (leisure activities), 11% (shopping), 25% (weekend trips), and 8% (taking children to school) of all transportation demand,[37] the overall mode share of road transportation should fall by 2.5 percentage points. Assuming initial mode shares of railway and road transportation of 10% and 60%, respectively, and assuming a constant

[35] The result for O_3 is consistent with Chen and Whalley (2012). Other papers using SO_2 as a placebo outcome in the context of road transportation are Currie and Walker (2011), Chen and Whalley (2012), and Bauernschuster et al. (2017).

[36] For shopping trips, estimates reject the null of instrument exogeneity. Recall that the overidentification statistic tests a joint null of exogeneity of instruments and homogenous treatment effects. Rejection of the null for shopping trips could be evidence that auctions and long distance trains affect shopping trips differently.

[37] According to ADAC (2010, p. 12) leisure is responsible for 40% of all passenger kilometers, commuting to work or education for 25%, business trips for 12%, shopping for 9%, errands for 8%, and accompaniment for 6%. In contrast to this breakdown of activities, we have no estimates for business trips and errands, but separate estimates for weekend trips and other leisure activities. Assuming that the latter two activities are equally important and rescaling the shares of all activities except business trips and errands so that these shares sum to 100%, yields the figures mentioned in the text.

Table 4
Additional instrument on capacity. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, and state statistical agencies.

	Frequency of service	CO	NO_x	Infant mortality
A. First stage and reduced form				
Auction	0.149	−0.020	−0.030	−0.032
	(0.043)***	(0.013)	(0.017)*	(0.033)
	[0.049]***	[0.015]	[0.020]	[0.035]
Long distance trains 1994	−0.077	0.018	0.038	0.049
	(0.045)*	(0.024)	(0.023)*	(0.057)
	[0.056]	[0.025]	[0.028]	[0.080]
Electric traction	−0.013	−0.013	0.015	−0.070
	(0.034)	(0.014)	(0.014)	(0.031)**
	[0.039]	[0.016]	[0.015]	[0.039]*
Length (log km)	−0.065	−0.002	−0.019	−0.017
	(0.024)***	(0.010)	(0.008)**	(0.019)
	[0.026]**	[0.011]	[0.011]*	[0.021]
Distance to city (100 km)	−0.029	0.041	−0.088	−0.035
	(0.064)	(0.028)	(0.028)***	(0.070)
	[0.074]	[0.043]	[0.042]**	[0.098]
F-stat instruments	7.637			
Adj. R-squared	0.063	0.635	0.662	0.121
B. IV estimates				
Frequency of service		−0.138	−0.224	−0.244
		(0.089)	(0.119)*	(0.219)
		[0.108]	[0.154]	[0.258]
Hansen J (p-value)		0.755	0.416	0.602
Overall significance (p-value)		0.000	0.000	0.000
Train lines		559	559	559

Notes: This table reports the results of OLS regressions of the change in the log of frequency of service, CO and NO_x pollution, and infant mortality between 1994 and 2004 on two instruments, namely the procurement mode of the line and track sharing with long-distance trains in 1994 in Panel A. The change in the log of frequency of service is instrumented using these two instruments in Panel B. See notes to Table 1 for the definition of the instruments and control variables. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

transport demand, this amounts to a reduction in road transportation of around 4%. To put this rough estimate in perspective, suppose further that the 28% increase in railway services led to a 25% increase in passenger kilometers (elasticity of about 0.9). Under these assumptions, which are clearly optimistic in view of previous literature (see Section 2), we would obtain mode shares of about 13% for railways and 57% for cars after the increase in the frequency of service, which would roughly correspond to the previously mentioned reduction of car transportation by 2.5 percentage points.

To sum up, the response of the modal split implied by our estimates is high, but not implausibly so. In fact, the high service elasticity of passenger rail demand fits quite well with various observations on the evolution on passenger services. For instance, consistent with the analysis presented here, Böttger and Pörner (2007) report an increase in the usage of local passenger railways in Germany from 30.3 billion passenger kilometers in 1994 to 40.2 billion passenger kilometer in 2004, that is, by more than 30%. LNVG (2010) even reports growth rates of passenger transportation that are often substantially above the growth rates of supply, measured in train kilometers.[38] Similarly, Allianz Pro Schiene, a lobbying organisation supporting "safe and environmentally friendly rail transportation" published a list of 15 (mostly competitively procured) railway lines that were particularly successful in attracting passengers in the first 10–15 years after the railway reform. In several cases, patronage increased by a factor of two to three or even higher.[39] Such strong demand effects are necessary conditions for substantial substitution effects of supporting public transport.

[38] For instance, on the *Weser-Ems Netz* run by *Nord-West-Bahn*, the number of passengers grew by 248% between 1998 and 2006, with a concomitant growth in the number of passengers per train by more than 50% (LNVG, 2010).

[39] See https://www.allianz-pro-schiene.de/presse/pressemitteilungen/2009-2009-47/, visited February 18, 2017.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 5
Placebo results. Sources: Own calculations based on data from published railway timetables and the federal environmental agency.

	SO_2	O_3	SO_2	O_3
A. Reduced form				
Auction	0.021	0.006	0.021	0.006
	(0.020)	(0.005)	(0.020)	(0.005)
	[0.026]	[0.006]	[0.026]	[0.006]
Long distance trains 1994			0.013	0.013
			(0.029)	(0.009)
			[0.035]	[0.011]
Electric traction	0.008	−0.000	0.007	−0.001
	(0.020)	(0.004)	(0.021)	(0.004)
	[0.024]	[0.005]	[0.025]	[0.005]
Length (log km)	0.028	−0.004	0.028	−0.004
	(0.013)**	(0.003)	(0.013)**	(0.003)
	[0.015]*	[0.004]	[0.015]*	[0.004]
Distance to city (100 km)	0.062	−0.012	0.062	−0.011
	(0.046)	(0.011)	(0.046)	(0.011)
	[0.065]	[0.017]	[0.065]	[0.017]
Adj. R-squared	0.880	0.581	0.880	0.581
B. IV estimates				
Frequency of service	0.139	0.043	0.117	0.028
	(0.139)	(0.034)	(0.131)	(0.031)
	[0.181]	[0.043]	[0.166]	[0.040]
Hansen J (p-value)			0.437	0.107
F-stat instruments	12.143	12.143	7.637	7.637
Overall significance (p-value)	0.000	0.000	0.000	0.000
Train lines	559	559	559	559

Notes: This table reports the results of OLS regressions of the change in the log of SO_2 and O_3 pollution between 1994 and 2004 on the procurement mode of the line in Columns (1) and (2) and on the procurement mode and track sharing with long-distance trains in 1994 in Columns (3) and (4) of Panel A. The change in the log of frequency of service is instrumented using these instruments in Panel B. See notes to Table 1 for the definition of the instruments and control variables. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols * and ** refer to statistical significance at the 10% and 5%-level, respectively.

6.5. Duration and heterogeneity

In Table 7 we examine whether the effects on air pollution and infant mortality increase with the time elapsed since the auction took place. In Panel A, the first stage and reduced form regressions include not only the procurement mode, but also the number of years since the auction took place. The duration variable is demeaned for all competitive lines and 0 for all non-competitive lines. Thus, the coefficient for auctions captures the effect of procurement mode for the average competitive line. According to Panel A, the number of years since the auction has a negative effect on pollution. The most plausible explanation is that the environmental effects need time to materialize.[40] Panel B of Table 7 presents IV estimates with both procurement mode and the duration variable as instruments. The estimated elasticities are similar to the ones reported in Tables 2 and 4.

Arguably, the main reason for the delayed effects of competitive procurement on environmental quality is that individuals need time to adapt their travelling habits to an increases in the frequency of service. The results in Table 8 support this notion. Not only does competitive procurement have a negative effect on car use for leisure activities and shopping on average, the effect is larger the earlier the auction took place. IV estimates with the procurement mode and the duration variable as instruments confirm the negative effect of the frequency of service on car use for leisure activities and shopping.

Table 9 suggests that increases in frequency of service reduce infant mortality more on lines serving big cities than on rural lines. As discussed in Section 1, railway transportation may have less favorable environmental properties in rural areas due to lower capacity utilization. Note, however, that we find larger effects on CO pollution along rural lines. Table 9 further suggests that the effects on air pollution and health are considerably larger on lines with a high density of stations. Thus, easy access to railway stations and high frequency of service complement each other.

[40] However, in our setup, it is not possible to disentangle cohort and duration effects.

Table 6
Effect on modal choice. Sources: Own calculations based on data from published railway timetables and the SOEP.

	Work	Leisure	Shop	Weekend	Children
A. Reduced form					
Auction	−0.020	−0.035	−0.029	0.001	0.040
	(0.014)	(0.014)**	(0.012)**	(0.012)	(0.026)
	[0.016]	[0.015]**	[0.012]**	[0.013]	[0.029]
Long distance trains 1994	0.007	−0.001	−0.019	0.004	−0.004
	(0.023)	(0.020)	(0.013)	(0.020)	(0.037)
	[0.030]	[0.030]	[0.014]	[0.027]	[0.039]
Electric traction	0.006	−0.005	0.007	0.003	0.021
	(0.012)	(0.010)	(0.009)	(0.010)	(0.018)
	[0.012]	[0.011]	[0.010]	[0.011]	[0.020]
Length (log km)	−0.006	−0.016	0.005	−0.001	−0.019
	(0.008)	(0.007)**	(0.005)	(0.006)	(0.015)
	[0.008]	[0.008]*	[0.005]	[0.007]	[0.014]
Distance to city (100 km)	0.030	−0.053	−0.012	0.021	0.023
	(0.026)	(0.022)**	(0.018)	(0.020)	(0.043)
	[0.032]	[0.029]*	[0.021]	[0.024]	[0.053]
Adj. R-squared	0.062	0.090	0.053	0.085	0.050
B. IV estimates					
Frequency of service	−0.146	−0.237	−0.172	−0.002	0.270
	(0.113)	(0.131)*	(0.097)*	(0.084)	(0.199)
	[0.135]	[0.153]	[0.106]	[0.092]	[0.224]
Hansen J (p-value)	0.873	0.382	0.038	0.827	0.668
F-stat instruments	5.084	5.079	5.064	5.000	5.254
Overall significance (p-value)	0.001	0.000	0.000	0.000	0.000
Train lines	553	554	555	556	546

Notes: This table reports the results of OLS regressions of the changes in the share of individuals using car or motorcycle for an activity between 1998 and 2003 on the procurement mode of the line and track sharing with long-distance trains in 1994 in Panel A. The change in the log of frequency of service is instrumented using these two instruments in Panel B. See notes to Table 1 for the definition of the instruments and control variables. Only lines auctioned in the years 1998–2002 are classified as competitively procured lines. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols * and ** refer to statistical significance at the 10% and 5%-level, respectively.

7. Size and value of benefits

We now use back-of-the-envelope calculations to illustrate the size of the CO and NO_x pollution reductions. Because reduced NO_x pollution is important from a health perspective, we also provide a highly tentative assessment of the value of this reduction.

According to the UBA, German CO emissions in 1994 were 6.81 megatons from all sources and 3.70 megatons (54%) from road transportation. The emissions from all sources fell by 2.89 megatons between 1994 and 2004. Road traffic emissions declined by 2.17 megatons, so that the sector contributed to 75% of the total emission reduction. NO_x emissions in 1994 from all sources and road transportation were 2.20 megatons and 1.14 megatons (52%), respectively. NO_x emissions from all sources fell by 0.55 megatons. The reduction for road transportation was 0.34 megatons, corresponding to 61% of the total reduction. Pollution reduction in Germany was thus mainly the result of improvements in road transportation. Our analysis suggests that expanding regional rail service contributed to these reductions in CO and NO_x emissions. Our point estimates indicate that the 28% increase in the frequency of service between 1994 and 2004 reduced car and motorcycle by about 4% (Table 6 and Section 6.4) and CO and NO_x pollution by 3.2% and 4.9% (Table 2), respectively. Thus, the estimates for pollution are somewhat larger than what the estimate for road traffic would suggest, but the estimates are nevertheless quite similar.[41] They also suggest that the contribution of improved railway services to the reduced emissions from road transportation is plausible. Moreover, if we apply the percentage pollution reductions to emissions rather than to concentrations,

[41] To put the size of the pollution reduction in perspective, one should note that the traffic reductions resulting from better public transport are likely to be particularly pronounced in peak hours (Anderson, 2014; Adler and van Ommeren, 2016). As stop-and-go traffic is very pollution intensive, strong pollution reductions are plausible.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 7
Does time since auction matter? Sources: Own calculations based on data from published railway timetables, the federal environmental agency, and state statistical agencies.

	Frequency of service	CO	NO_x	Infant mortality
A. First stage and reduced form				
Auction	0.150	−0.020	−0.031	−0.033
	(0.042)***	(0.013)	(0.016)*	(0.033)
	[0.048]***	[0.014]	[0.020]	[0.035]
Years since auction	0.021	−0.009	−0.009	−0.017
	(0.015)	(0.005)*	(0.005)*	(0.012)
	[0.019]	[0.006]	[0.008]	[0.016]
Electric traction	−0.018	−0.012	0.018	−0.067
	(0.034)	(0.014)	(0.013)	(0.030)**
	[0.038]	[0.015]	[0.015]	[0.037]*
Length (log km)	−0.065	−0.002	−0.019	−0.018
	(0.024)***	(0.010)	(0.008)**	(0.019)
	[0.026]**	[0.011]	[0.011]*	[0.021]
Distance to city (100 km)	−0.027	0.041	−0.089	−0.037
	(0.064)	(0.027)	(0.027)***	(0.069)
	[0.075]	[0.042]	[0.041]**	[0.097]
F-stat instruments	6.564			
Adj. R-squared	0.065	0.637	0.662	0.123
B. IV estimates				
Frequency of service		−0.173	−0.235	−0.295
		(0.085)**	(0.118)**	(0.228)
		[0.100]*	[0.150]	[0.280]
Hansen J (p-value)		0.260	0.492	0.373
Overall significance (p-value)		0.000	0.000	0.000
Train lines		559	559	559

Notes: This table reports the results of OLS regressions of the change in the log of frequency of service, CO and NO_x pollution, and infant mortality between 1994 and 2004 on two instruments, namely the procurement mode of the line and the number of years since the auction, in Panel A. The change in the log of frequency of service is instrumented using these two instruments in Panel B. See notes to Table 1 for the definition of the instruments and control variables. Years since auctions is demeaned for competitively procured lines and 0 for non-competitively procured lines. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

they would correspond to 0.129 megatons of CO emissions and 0.085 megatons of NO_x emissions.[42] This would suggest that 6% of reduced CO emissions from road transportation (and 4% of total reductions) and 25% of reduced NO_x emissions from road transportation (or 15% of total reductions) could be attributed to the substitution from road to rail. Again, this seems plausible.

For our valuation of the benefit from the reduction of NO_x pollution by 4.9%, we rely on results from studies on the mortality effects of NO_x pollution and on the value of a statistical life. We proceed in three steps. First, we calculate the absolute reduction in concentration levels due to the 28% increase in the frequency of service as 2.980 $\mu g/m^3$.[43]

Second, we calculate the number of prevented deaths. For this purpose, we use a recent recommendation for the U.K. government which refers to European studies (COMEAP, 2015). The analysis suggests using hazard ratios of 1.02–1.027 for the group of persons aged 30 or more (for deaths from all causes). To address the substantial uncertainty in such estimates, we use the lower bound of this interval and thus assume that an increase in the exposure to NO_2 by 10 $\mu g/m^3$ increases the mortality risk by the (still substantial) amount of 2%. According to the Federal Statistical Office, there were 807,562 deaths of persons aged 30 or more from all causes in Germany in 2004. Hence, the number of prevented deaths in 2004 is 4813

[42] Denote the percentage reductions as z. The absolute emission reduction then is $(z/(100+z)) \times$ total 2004 emissions.

[43] Again, denote the resulting percentage reduction in pollution of −4.9% as z. As the average level of NO_x pollution in 2004 is 57.993 $\mu g/m^3$, the absolute reduction in concentration levels is $(z/(100+z)) \times 57.993\ \mu g/m^3 = -2.980\ \mu g/m^3$.

Table 8
Modal choice and time since auction. Sources: Own calculations based on data from published railway timetables and the SOEP.

	Work	Leisure	Shop	Weekend	Children
A. Reduced form					
Auction	−0.020	−0.035	−0.029	0.000	0.040
	(0.014)	(0.014)***	(0.011)**	(0.012)	(0.026)
	[0.016]	[0.015]**	[0.012]**	[0.013]	[0.029]
Years since auction	−0.011	−0.010	−0.017	−0.002	0.005
	(0.008)	(0.008)	(0.007)**	(0.007)	(0.014)
	[0.009]	[0.009]	[0.008]**	[0.007]	[0.013]
Electric traction	0.006	−0.005	0.005	0.003	0.020
	(0.012)	(0.010)	(0.009)	(0.010)	(0.018)
	[0.011]	[0.010]	[0.010]	[0.011]	[0.020]
Length (log km)	−0.006	−0.016	0.005	−0.000	−0.019
	(0.008)	(0.007)**	(0.005)	(0.006)	(0.015)
	[0.008]	[0.008]*	[0.005]	[0.007]	[0.014]
Distance to city (100 km)	0.028	−0.054	−0.013	0.021	0.023
	(0.026)	(0.023)**	(0.018)	(0.020)	(0.043)
	[0.033]	[0.029]*	[0.021]	[0.025]	[0.053]
Adj. R-squared	0.065	0.093	0.062	0.085	0.050
B. IV estimates					
Frequency of service	−0.167	−0.276	−0.240	0.000	0.296
	(0.126)	(0.156)*	(0.123)*	(0.091)	(0.212)
	[0.148]	[0.181]	[0.139]*	[0.097]	[0.238]
Hansen J (p-value)	0.284	0.520	0.131	0.758	0.891
F-stat instruments	3.582	3.566	3.564	3.552	3.702
Overall significance (p-value)	0.001	0.000	0.000	0.000	0.000
Train lines	553	554	555	556	546

Notes: This table reports the results of OLS regressions of the changes in the share of individuals using car or motorcycle for an activity between 1998 and 2003 on two instruments, namely the procurement mode of the line and the number of years since the auction, in Panel A. The change in the log of frequency of service is instrumented using these two instruments in Panel B. See notes to Table 1 for the definition of the instruments and control variables. Only lines auctioned in the years 1998–2002 are classified as competitively procured lines. Years since auctions is demeaned for competitively procured lines and 0 for non-competitively procured lines. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state- specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

$(= (1.02 - 1) \times (2.980/10) \times 807, 562)$. Third, with the figure for the value of a statistical life of EUR 5,364,661 (2008 EUR; 7.9 million USD in 2008 USD) from the 2014 update of the 2010 guidelines of the Environmental Protection Agency (EPA, 2014), we translate the prevented deaths into annual health benefits of approximately 26 billion EUR.

Although this figure may well be biased upwards, the health benefits seem to compare favorably to the costs of increasing the frequency of service. Taking the subsidies of 6–7 billion EUR spent for the entire provision of regional passenger services as a reference and assuming that the 28% increase in railway services corresponds to a cost increase of similar size, the costs for the additional services are likely to be in the order of magnitude of 1.3–1.5 billion EUR. Moreover, other health and non-health benefits from air quality improvements and the benefits from reducing other road traffic externalities are not accounted for. The substitution from cars to railways most likely has considerable effects on other externalities such as reductions in emissions of carbon dioxide, less noise, and lower congestion. Increased emissions from electricity generation counterbalance some of these effects.

Our results appear quite high compared to Chen and Whalley (2012). There are many differences between the two studies, including the size of the measure (one new metro project vs. an overall increase in transportation services), the age group considered (infants vs. adults) and the basis for the monetary valuation of saved lives. The main reason for the difference is the approach to translating pollution reduction into saved lives. While Chen and Whalley (2012) extrapolate the effects from studies on infant mortality (for CO), we use the results of epidemiological studies that account directly for deaths of older people.

Please cite this article as: Lalive, R., et al., Does expanding regional train service reduce air pollution? Journal of Environmental Economics and Management (2017), http://dx.doi.org/10.1016/j.jeem.2017.09.003

Table 9
Effects for sub-groups. Auction and time since auction. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, and state statistical agencies.

	Rural			City		
	CO	NO_x	Infant mortality	CO	NO_x	Infant mortality
Frequency of service	−0.199	−0.143	0.074	−0.103	−0.163	−0.377
	(0.114)*	(0.141)	(0.368)	(0.083)	(0.100)	(0.159)**
	[0.132]	[0.178]	[0.420]	[0.090]	[0.112]	[0.161]**
Electric traction	0.004	0.056	−0.123	−0.025	−0.011	−0.073
	(0.024)	(0.026)**	(0.078)	(0.019)	(0.018)	(0.034)**
	[0.027]	[0.029]*	[0.101]	[0.020]	[0.020]	[0.035]**
Length (log km)	−0.005	−0.032	0.057	−0.021	−0.026	−0.084
	(0.016)	(0.017)*	(0.047)	(0.015)	(0.014)*	(0.023)***
	[0.016]	[0.020]	[0.053]	[0.017]	[0.019]	[0.025]***
Distance to city (100 km)	0.111	−0.039	−0.004			
	(0.034)***	(0.036)	(0.108)			
	[0.040]***	[0.044]	[0.123]			
Hansen J (p-value)	0.351	0.420	0.489	0.903	0.858	0.269
F-stat instruments	3.040	3.040	3.040	6.493	6.493	6.493
Overall significance (p-value)	0.000	0.000	0.025	0.000	0.000	0.000
Train lines	221	221	221	338	338	338
	Many stations			Few stations		
	CO	NO_x	Infant mortality	CO	NO_x	Infant mortality
Frequency of service	−0.307	−0.778	−1.533	−0.133	−0.115	0.007
	(0.330)	(0.574)	(1.059)	(0.072)*	(0.088)	(0.203)
	[0.337]	[0.573]	[1.097]	[0.084]	[0.093]	[0.202]
Electric traction	−0.042	−0.051	−0.184	−0.014	0.023	−0.061
	(0.036)	(0.054)	(0.099)*	(0.016)	(0.018)	(0.039)
	[0.041]	[0.060]	[0.103]*	[0.020]	[0.023]	[0.046]
Length (log km)	0.003	−0.002	0.027	−0.025	−0.034	−0.038
	(0.020)	(0.028)	(0.052)	(0.015)*	(0.015)**	(0.034)
	[0.027]	[0.029]	[0.052]	[0.016]	[0.017]**	[0.038]
Distance to city (100 km)	−0.018	−0.114	0.063	0.092	−0.050	−0.093
	(0.057)	(0.074)	(0.115)	(0.030)***	(0.032)	(0.087)
	[0.081]	[0.074]	[0.111]	[0.040]**	[0.042]	[0.107]
Hansen J (p-value)	0.694	0.301	0.719	0.119	0.372	0.473
F-stat instruments	1.127	1.127	1.127	6.250	6.250	6.250
Overall significance (p-value)	0.000	0.000	0.049	0.000	0.000	0.000
Train lines	279	279	279	280	280	280

Notes: This table reports the results of IV regressions of the change in the log of CO and NO_x pollution and infant mortality on the change in the log of frequency of service between 1994 and 2004. The change in the log of frequency of service is instrumented using the procurement mode of the line and the number of years since the auction. Separate results are shown for rural lines (with distance to city >0), lines running through cities (with distance to city =0), lines with many stations (with station density above median), and lines with few stations (with station density below median). Station density is the number of stations divided by line length. See notes to Table 1 for the definition of the instruments and control variables. Years since auctions is demeaned for competitively procured lines and 0 for non-competitively procured lines. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state- specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets. Tests of model specification are based on robust standard errors, not allowing for possible clustering of adjacent lines. Hansen J (p-value) reports the p-value of the Hansen J statistic, a joint test of exogeneity of instruments. F-stat instruments reports the robust first stage F-statistic, the Kleibergen-Paap rk Wald F statistic for a setting with one endogenous regressor. Overall significance p-value is the p-value of the robust joint test of significance of all regressors in the second stage equation.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

8. Concluding remarks

This paper exploits regional variation in the supply of railway services to identify the effects of support for passenger railways on road traffic externalities. We find substantial benefits of improving public transport. For instance, NO_x pollution decreases by around 2% for every 10% increase in regional passenger railway service.

Regional passenger railway service is heavily subsidized in Germany. Are these subsidies worthwhile? Using previous estimates of the health costs from NO_x pollution, we provide rough estimates of the monetary benefits of the 28% expansion in regional railway services between 1994 and 2004. The results suggest that these monetary benefits might well exceed the costs by a considerable amount.

Is our empirical approach applicable to other countries with local variation in the development of railway services? We addressed endogeneity concerns by using the procurement mode as an instrument. In some countries this approach cannot be applied, because procurement either relies exclusively on competition (as in Great Britain) or not at all (as in Switzerland). Yet other countries, such as Denmark and the Netherlands increasingly use auctions (Nash, 2008). Thus, the empirical approach could be used in other countries as well.

Acknowledgements

We are grateful to Jérôme Adda, Gerard van den Berg, Paolo Masella, Moritz Meyer, Alex Whalley, and to seminar participants in Berlin (Free University), EUI Florence, Geneva, Lucerne, Mannheim, and Zurich and at meetings of the German Economic Association (Göttingen), the environmental economics group of the German Economic Association (Wageningen), the Swiss Society of Economics and Statistics (Lucerne), and the Workshop “Beyond Basic Questions” (Schiermonnikoog) for helpful comments. Jérémy Zuchuat provided excellent research assistance.

Appendix. Additional results

Table 10
OLS results. Sources: Own calculations based on data from published railway timetables, the federal environmental agency, and state statistical agencies.

	CO	NO_x	Infant mortality
Frequency of service	−0.022	0.004	0.005
	(0.015)	(0.016)	(0.048)
	[0.019]	[0.018]	[0.047]
Electric traction	−0.009	0.024	−0.060
	(0.013)	(0.013)*	(0.030)**
	[0.015]	[0.015]	[0.036]*
Length (log km)	−0.003	−0.018	−0.016
	(0.010)	(0.008)**	(0.019)
	[0.012]	[0.011]*	[0.021]
Distance to city (100 km)	0.039	−0.091	−0.039
	(0.028)	(0.028)***	(0.069)
	[0.043]	[0.042]**	[0.098]
Adj. R-squared	0.636	0.659	0.120
Train lines	559	559	559

Notes: This table reports the results of OLS regressions of the change in the log of CO and NO_x pollution and infant mortality on the change in the log of frequency of service between 1994 and 2004. See notes to Table 1 for the definition of control variables. All regressions control for fixed effects of federal states. The variables electric traction, track length, distance to city, and state effects are time invariant and, thus, capture differential trends in the dependent variable across these line characteristics and states or characteristic- and state-specific time effects. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets.

The symbols *, **, *** refer to statistical significance at the 10%, 5%, and 1%-level, respectively.

Table 11
Determinants of auctions. Sources: Own calculations based on data from published railway timetables and the Federal Returning Officer.

	Auction
Left vote share	−0.589 (0.399) [0.469]
Population in 1994 (log)	−0.042 (0.033) [0.041]
Electric traction	−0.158 (0.038)*** [0.048]***
Length (log km)	0.014 (0.029) [0.030]
Distance to city (100 km)	−0.002 (0.079) [0.091]
Adj. R-squared	0.106
Train lines	559

Notes: This table reports the results of a linear probability model explaining the procurement mode of the line. The left vote share is the vote share (second vote) of the Social Democratic Party (SPD), the Greens (GRÜNE), and the Party of Democratic Socialism (PDS) in the 1994 federal elections. See notes to Table 1 for the definition of control variables. All regressions control for fixed effects of federal states. Robust standard errors in parentheses and standard errors that account for clustering along network links in brackets.

The symbol *** refers to statistical significance at the 1%-level.

References

ADAC, 2010. Mobilität in Deutschland. Ausgewählte Ergebnisse, München.
Adler, M.W., van Ommeren, J.N., 2016. Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes. J. Urban Econ. 92, 106–119.
Anderson, M.L., 2014. Subways, strikes, and slowdowns: the impacts of public transit on traffic congestion. Am. Econ. Rev. 104, 2763–2796.
Arceo, E., Hanna, R., Oliva, P., 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. Econ. J. 126, 257–280.
Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. The Demand for Public Transport: A Practical Guide, TRL Report TRL 593, Crowthorne.
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. Am. Econ. J.: Econ. Policy 9, 1–37.
Böttger, C., Pörner, R., Der SPNV in Deutschland – Eine Erfolgsstory mit Potenzialen? Jahrbuch des Bahnwesens 2006/2007, 2007, pp. 132–140.
Chen, Y., Whalley, A., 2012. Green infrastructure: the effects of urban rail transit on air quality. Am. Econ. J.: Econ. Policy 4, 58–97.
COMEAP, 2015. Interim Statement on Quantifying the Association of Long-term Average Concentrations of Nitrogen Dioxide and Mortality, Guidance Statement from the Committee on the Medical Effects of Air Pollutants.
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., Walker, R., 2011. Traffic congestion and infant health: evidence from E-ZPass. Am. Econ. J.: Appl. Econ. 3, 65–90.
Currie, J., Graff Zivin, J., Mullins, J., Neidell, M., 2014. What do we know about short- and long-term effects of early-life exposure to pollution? Annu. Rev. Resour. Econ. 6, 217–247.
Currie, J., Neidell, M., Schmieder, J.F., 2009. Air pollution and infant health: lessons from New Jersey. J. Health Econ. 28, 688–703.
Duranton, G., Turner, M.A., 2011. The fundamental law of road congestion: evidence from US cities. Am. Econ. Rev. 101, 2616–2652.
EPA, 2014. Guidelines for Preparing Economic Analyses, Washington.
Evans, J.H., 2004. Transit Scheduling and Frequency. Traveler Response to Transportation System Changes, TCRP Report 95, Chapter 9.
Geddes, J.A., Murphy, J.G., Wang, D.K., 2009. Long term changes in nitrogen oxides and volatile organic compounds in Toronto and the challenges facing local ozone control. Atmos. Environ. 43, 3407–3415.
Graff Zivin, J., Neidell, N., 2013. Environment, health, and human capital. J. Econ. Lit. 51, 689–730.
Han, S., Bian, H., Feng, Y., Liu, A., Li, A., Zeng, F., Zhang, X., 2011. Analysis of the relationship between O_3, NO, and NO_x in Tianjin, China. Aerosol Air Qual. Res. 11, 128–139.
IFEU, 2010. Fortschreibung und Erweiterung 'Daten- und Rechenmodell: Energieverbrauch und Schadstoffemissionen des motorisierten Verkehrs in Deutschland 1960–2030 (TREMOD, Version 5)', Heidelberg.
INFRAS, 2010. Handbook of Road Emission Factors, Zürich.
Kleibergen, F., Paap, R., 2006. Generalized reduced rank tests using the singular value decomposition. J. Econ. 133, 97–126.
Knittel, C.R., Miller, D.L., Sanders, N.J., 2016. Caution, drivers! Children present: traffic, pollution, and infant health. Rev. Econ. Stat. 98, 350–366.
Lalive, R., Schmutzler, A., 2008. Exploring the effects of competition for railway markets. Int. J. Ind. Organ. 26, 443–458.
Lalive, R., Schmutzler, A., Zulehner, C., 2016. Auctions vs Negotiations in Public Procurement. Which Works Better?, Working Paper. University of Zurich.
LNVG, 2010. 2013+ −, Konzept zum Schienenpersonennahverkehr, Landesnahverkehrsgesellschaft Niedersachsen.
Luechinger, S., 2009. Valuing air quality using the life satisfaction approach. Econ. J. 119, 482–515.
Luechinger, S., 2014. Air pollution and infant mortality: a natural experiment from power plant desulfurization. J. Health Econ. 37, 219–231.
Monopolkommission, 2009. Bahn 2009: Wettbewerb braucht Weichenstellung, Sondergutachten 55.
Mouwen, A., van Ommeren, J.N., 2016. The effect of contract renewal and competitive tendering on public transport costs, subsidies and ridership. Transp. Res. Part A 87, 78–89.

Nash, C., 2008. Passenger railway reform in the last 20 years – European experience reconsidered. Res. Transp. Econ. 22, 61–70.
Olson, M., 1969. The principle of 'fiscal equivalence': the division of responsibilities among different levels of government. In: American Economic Review Papers and Proceedings 59, pp. 479–487.
Parry, I.W.H., Small, K.A., 2009. Should urban transit subsidies be reduced? Am. Econ. Rev. 99, 700–724.
Schmutzler, A., 2011. Local transportation policy and the environment. Environ. Resour. Econ. 48, 511–535.
SCI Verkehr GmbH, 2006. Analyse der Mittelverwendung bei der SPNV-Finanzierung in den deutschen Bundesländern auf Grundlage des Haushaltsvollzuges für das Jahr 2004, Berlin.
Wagner, G.G., Frick, J.R., Schupp, J., 2007. The German socio-economic panel study (SOEP) – scope, evolution and enhancements. Schmollers Jahrb. 127, 139–169.
Wardman, M., 2012. Review and meta-analysis of U.K. time elasticities of travel demand. Transportation 39, 465–490.
White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica 48, 817–838.