Air pollution and CO2 from daily mobility: Who emits and Why? Evidence from Paris

Marion Leroutier, Philippe Quirion

Energy Economics 109 (2022) 105941

Contents lists available at ScienceDirect

Energy Economics

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

Air pollution and CO_2 from daily mobility: Who emits and Why? Evidence from Paris

Marion Leroutier [a,*], Philippe Quirion [b,c]

[a] Misum, Stockholm School of Economics, Sweden
[b] CIRED, France
[c] CNRS, France

ARTICLE INFO

Dataset link: https://osf.io/pnyzk/

JEL classification:
D62
R40
Q52
Q53

Keywords:
Transport externalities
Environmental inequalities
LMDI

ABSTRACT

Urban road transport is an important source of local pollution and carbon emissions. Designing effective and fair policies tackling these externalities requires understanding who contributes to emissions today. We estimate individual transport-induced pollution footprints combining a travel demand survey from the Paris area with NOx, $PM_{2.5}$ and CO_2 emission factors. We find that the top 20% emitters contribute 75%–85% of emissions on a representative weekday. They combine longer distances travelled, a high car modal share and, especially for local pollutants, a higher emission intensity of car trips. Living in the suburbs, being a man and being employed are the most important characteristics associated with top emissions. Among the employed, those commuting from suburbs to suburbs, working at a factory, with atypical working hours or with a manual, shopkeeping or top executive occupation are more likely to be top emitters. Finally, policies targeting local pollution may be more regressive than those targeting CO_2 emissions, due to the different correlation between income and the local pollutant vs. CO_2 emission intensity of car trips.

1. Introduction

Road transport is responsible for several well-documented environmental externalities (Parry et al., 2007). First, it contributes to outdoor air pollution, which has been identified by the WHO as the world's "largest single environmental health risk", accounting for an estimated 4.2 million deaths per year (WHO, 2014). Beside its impact on physical health, air pollution negatively impacts mental health (Bishop et al., 2018; Braithwaite et al., 2019), the formation of human capital (Currie et al., 2014) and productivity (Chang et al., 2019). Road transport also contributes to greenhouse gas emissions, mostly carbon dioxide (CO_2), with an increasing contribution relative to other economic sectors in most developed countries (IEA 2019). This trend needs to be reverted to achieve emission reductions consistent with the Paris agreement.

Yet, policy proposals aiming at increasing the cost of driving polluting cars, whether motivated by air quality or climate mitigation concerns, are controversial. The recent Yellow Vest movement in France revealed the low acceptability of a specific measure, the carbon tax; but other policy instruments such as low emission zones or congestion charges have also met opposition across Europe (Viegas, 2001; Le Parisien, 2019; Delhaes and Kersting, 2019; Isaksen and Johansen, 2020). It is then crucial to understand who the high emitters are, since they are the most likely to oppose these measures.

In this paper, we estimate how much individuals contribute to transport-related emissions of local pollutants and CO_2 in their daily travels. We do so in the context of a large urban area, where emissions are both more detrimental to health and possibly easier to tackle than in rural areas. On the first point, many urban areas suffer from high levels of pollution, including in developed countries subject to relatively strict environmental regulations: in Europe, France, Germany and the UK were condemned in 2018 for failing to meet air quality standards in several cities (European Commission, 2018). On the second point, urban areas present more alternatives to cars: the higher density makes active modes more attractive, and public transport is more widespread (Creutzig et al., 2020). We combine individual travel information from a large representative survey conducted in the Paris area with mode-specific and vehicle-specific emission factors to uncover the magnitude of inequalities in local pollutant and carbon footprints, the mediating mobility patterns underlying top emissions and the individual characteristics associated with top emitters. We focus on two local pollutants having detrimental effects on health, nitrogen oxide (NOx) and fine particulate matter ($PM_{2.5}$), and the main greenhouse gas, carbon dioxide (CO_2). Our analysis only includes trips made within the Paris area. Long-distance trips made by car, rail or

* Correspondence to: Misum, Stockholm School of Economics, Sveavagen 65, 113 83 Stockholm, Sweden.
E-mail address: marion.leroutier@hhs.se (M. Leroutier).

https://doi.org/10.1016/j.eneco.2022.105941
Received 20 October 2021; Received in revised form 16 February 2022; Accepted 2 March 2022
Available online 25 March 2022
0140-9883/

aviation are excluded, such that our analysis sheds light on urban and suburban mobility patterns only.

We find strong inequalities in emissions among individuals, with the top 20% of emitters contributing 75%–85% of emissions on a representative weekday, depending on the pollutant. Applying an exact factor decomposition analysis on emissions quintiles, we show that top emissions result from the combination of longer distances travelled, a higher reliance on car, and a higher emission intensity within modes. While all three channels contribute equally to the difference in local pollutant emissions between a top and an average emitter, for CO_2 the difference is mostly explained by longer distances and a high reliance on cars, and less by differences in emission intensities. In a second step, we highlight the individual socio-economic and locational characteristics associated with being in the top 20% of emitters, and with each of the distance, modal share and emission intensity channels. Beside the characteristics already well-identified in the literature, such as being employed or living far away from the city centre, we highlight the role of gender and, for those in employment, the role of job characteristics: having atypical working hours, working in a factory, being a shopkeeper, manual worker or company head are associated with a higher likelihood to be a top emitter. Finally, we show the ambivalent role of income, which is associated with higher distances, a higher probability to use a car and a higher CO_2 emission intensity of cars, but not with a higher NOx and $PM_{2.5}$ emission intensity.

Our paper contributes to several strands of the literature: first, we contribute to the literature on environmental inequalities by investigating individual contribution to transport-related local pollutants and CO_2. On local air pollutant emissions, there is a vast literature examining cross-country inequalities in emissions – in relation to the Environmental Kuznets Curve hypothesis (Dinda, 2004) – and a more limited literature examining inequalities at the individual or household level (Levinson and O'Brien, 2018; Barnes et al., 2019). On CO_2 emissions, there is also flourishing literature looking at inequalities in individual carbon footprint at the country or regional scale (Sager, 2019; Ivanova and Wood, 2020; Büchs and Schnepf, 2013), or, closely linked, examining carbon tax incidence by socio-economic group (Douenne, 2020; Cronin et al., 2018). Most of these studies estimating individual emissions rely on input–output methodologies combined with micro-level consumer expenditure surveys, which provide limited information on travel behaviour (mostly the purchase of fuel and public transport tickets and subscription) and spatial location.

Our paper is closer in spirit to studies relying on detailed travel diaries from a sample of individuals to estimate individual emissions from transport (see for example Brand and Preston (2010), Barla et al. (2011), Ko et al. (2011), Bel and Rosell (2017), Yang et al. (2018), Brand et al. (2021)). An important limitation of most of these studies, however, is to rely on low sample sizes, and, often, on non-representative surveys where highly educated individuals are over-represented. In contrast, we use a large representative survey (N=23,690), similar to Bel and Rosell (2017) in the case of Barcelona or Ko et al. (2011) in the case of Seoul. We add to these papers in at least two ways. First, our large sample size enables us to examine the association between different job characteristics and top emissions for the subsample of employed individuals, a group who emits more than the rest and has more constrained trips. Second, having rich information on car characteristics allows us to apply different emission factors *within* modes for personal vehicles, while the aforementioned papers only use different emission factors *between* modes. This allows us to investigate how much differences in emission intensity contribute to differences in total emissions across individuals, and to examine correlations between individual characteristics and the emission intensity of car trips. Given that several popular policies directly or indirectly target the local pollution or CO_2 emission intensity of cars, it seems particularly important to understand these correlations and examine how they vary by pollutant.

Second, to the best of our knowledge, our paper is the first to jointly examine inequalities in the contribution to local pollutants and CO_2 emissions in the context of urban mobility. Given that local transport policies may be primarily motivated by either one or the other concern, it seems crucial to understand to what extent global and local pollution are caused by the same groups of individuals and the same travel behaviours. In that sense, we contribute to the literature examining the trade-offs and complementarities in tackling both CO_2 and local pollution (see Ambec and Coria (2013) for theoretical insights, Durrmeyer (2021) and Linn (2019) for empirical assessments in the transport sector). Durrmeyer (2021) and Linn (2019) show that while effective in decreasing CO_2 emissions, CO_2-based vehicle taxes are likely to increase the emission of damaging air pollutants (NOx and $PM_{2.5}$), because they increase the share of diesel cars, less CO_2-intensive but more intensive in NOx and $PM_{2.5}$. The reverse trade-off may exist in the case of local transport policies driven by air pollution concerns, and low-emission zones indeed tend to be more restrictive for diesel cars than for gasoline cars. Our results suggest that a policy targeting cars' local pollutant emission intensity may also have different distributional impacts from a policy targeting the CO_2 emission intensity, since we find different associations between household income and the $PM_{2.5}$ vs. CO_2 emission intensity of car trips.

Finally, our paper contributes to the literature on exact decomposition analysis. Most exact decomposition analyses using the Log-Mean-Divisia-index developed by Ang (2004, 2005) have aimed at understanding the components underlying the evolution of CO_2 emissions over time (for example, Wang et al. (2005), Mahony (2013)). They have been applied to aggregate time series data at the national or regional level. We instead aim at understanding the factors underlying differences in emissions across groups of individuals at a given point in time. While LMDI decompositions have been applied to cross-sectional data at the regional level (Ang et al., 2015; Liu et al., 2017), we adapt the method to the analysis of individual-level micro data.

The paper is organized as follows: Section 2 presents the local context; Section 3 presents the data and methods used; Section 4 presents the results and sectionSection 5 discusses their policy implications and concludes.

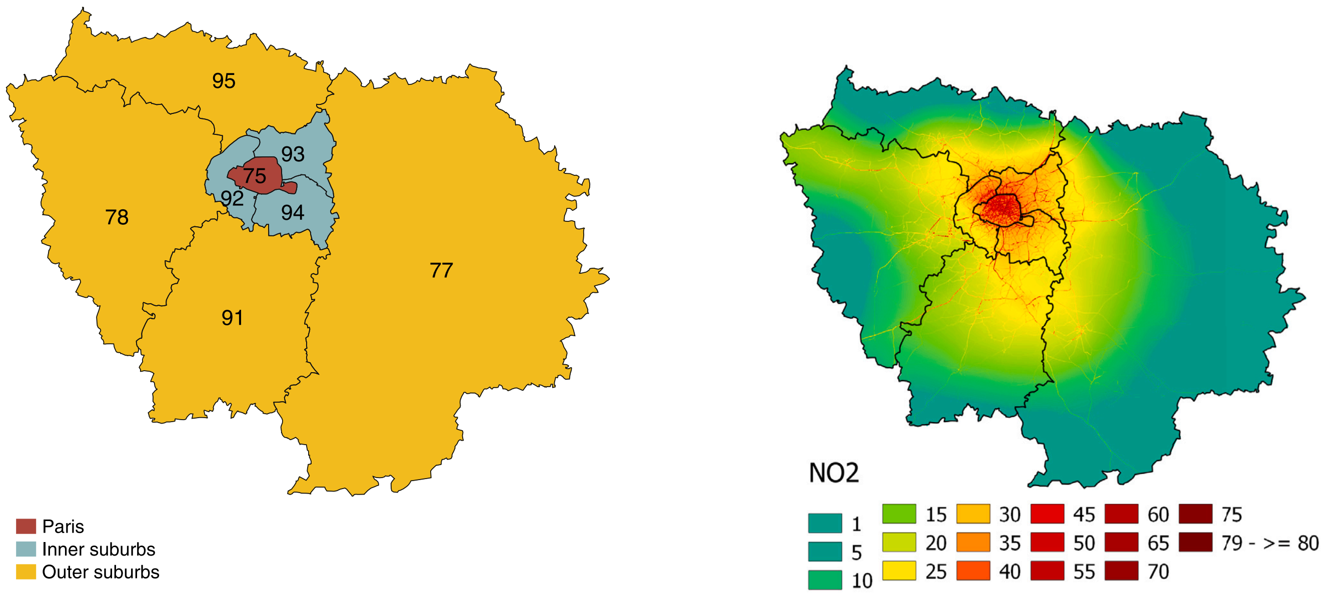
2. Air pollution and transport emissions in the Paris area

We consider the Paris area, which we define here as the administrative *region* of Ile de France (IdF), represented on Fig. 1(a) — the *region* is the first level of administrative subdivision in France.[1] The IdF region has a population of 12.2 million inhabitants and is made of three layers: the city of Paris in the centre (red), a first layer around Paris called the "inner suburbs", made of three small *départements* (blue) — the second level of administrative subdivision in France, and a second layer called the "outer suburbs", made of four larger but less dense *départements* (yellow).

We consider two types of transport emissions in this paper: local air pollutants contributing to ambient air pollution, and greenhouse gases contributing to climate change. Ambient air pollution levels regularly exceed recommended and legal thresholds in the Paris area. While concentrations of the main regulated pollutants,[2] have been decreasing throughout the area over the past ten years, they remain high, especially in the city centre. Fig. 1(b) shows NO_2 concentrations in 2015, a pollutant to which long-term exposure is associated with increases of bronchitis in asthmatic children and reduced lung function growth (WHO, 2018). The legal threshold of 40 $\mu g/m^3$ is exceeded in

[1] The Paris metropolitan area as defined by the French statistical institute does not include all the IdF region; it excludes a small part of the outer suburbs. We consider the whole region because our transport data are representative of the population from the entire region.

[2] Nitrogen dioxide NO_2 ozone O_3, and particulate matter PM_{10}.



(a) The Paris Metropolitan Area

(b) NO_2 in 2015, $\mu g/m^3$

Fig. 1. The Paris area: administrative subdivisions and pollution levels.
Note: The black line shows the division of each area in *département*. The numbers are the administrative identifiers for each *département*: 75-Paris; 92-Hauts-de-Seine; 93-Seine-Saint-Denis; 94-Val-de-Marne; 77-Seine-et-Marne; 78-Yvelines; 91-Essonne; 95-Val d'Oise. Source for NO_2 concentrations: Airparif.

Paris and the majority of the inner suburbs. Furthermore, despite the improvement in air quality, air pollution is the number one environmental concern in the Paris area according to a 2018 survey (Airparif - IFOP, 2018), and 61% of the respondents think that air pollution has increased in the past ten years. Regarding transport-related contribution to emissions, we consider two different local pollutants here: NOx, a generic category of pollutants including NO_2, and $PM_{2.5}$. Exposure to $PM_{2.5}$ has detrimental effects on health and increases mortality risk in the short- (Deryugina et al., 2019) and long-term (Lepeule et al., 2012), without evidence of a threshold below which exposure would be harmless (WHO, 2018). Road traffic is responsible for a sizeable share of local pollutant emissions in the Paris area in 2018: 53% of nitrogen oxides (NOx) and 19% of the $PM_{2.5}$. Besides, road traffic is responsible for 29% of the region's CO_2 emissions (Airparif, 2021).

Several regional and local policies have been implemented to tackle local pollution and CO2 emissions from cars. To dampen local pollution specifically, short-term driving restrictions based on license plate numbers have been systematically imposed since 2014 during pollution peaks. Long-term measures advertised by the regional authority include developing the public transport network – the Paris area is a typical monocentric city where most public transport lines converge to the centre – building more cycling lanes, reserving lanes for buses, clean vehicles and car-pooling, as well as speed reduction on the ring road (Région Ile de France, 2016). By far, the most ambitious policy specifically targeting air pollution is the Low Emission Zone (LEZ) projected to be rolled-out in Paris and the surrounding municipalities between 2017 and 2024, which should progressively ban all polluting vehicles – defined by their age and fuel type – from the city centre. However, this policy has met political opposition from some municipal authorities (Le Parisien, 2019). To reduce both local air pollution and CO_2 emissions from cars, the Paris metropolitan area also announced the complete ban of diesel cars by 2024 and of gasoline cars by 2030 (Le Monde, 2018).

3. Data and methodology

3.1. The data

Our main source of data is the 2010 wave of the EGT (*Enquête générale des transports* — EGT 2010-STIF-OMNIL-DRIEA), a survey conducted every 8 to 10 years in the Paris area. The 2010 wave is the last available,[3] and was conducted between October 2009 and May 2010, and between October 2010 and May 2011. The survey contains detailed information on the transport choices of 35,175 individuals from 14,885 households.[4] on a given weekday[5] and many socio-economic characteristics. The sample is representative of the Paris area (=IdF region) population as characterized in the 2008 census in terms of household size, type of housing and individual socio-economic and demographic profiles.[6] The EGT is also broadly representative of the 2011 Paris area population.[7]

For the present analysis, we use the subsample of mobile individuals, that is, adults having done at least one trip during the weekday (N=23,690). This represents 93.07% of the surveyed adults. Table 1 shows descriptive statistics for this subsample.[8] The average daily

[3] A new wave was planned to be carried out between 2018 and 2022. The data collection was interrupted in late 2019 because of a large public transport strike, and was subsequently stalled due to the Covid-19 crisis.

[4] The sampling rate at the household level is 1/330. In 2010, the Paris area had a population of 11.79 millions inhabitants.

[5] The respondents are asked about all their trips from the day preceding the interview, which can correspond to a day between Monday and Friday. We include survey day-of-week fixed effects in all our regression analyses because more people stay at home some days of the week, especially on Mondays (because many shops are closed) and Wednesdays (to take care of the children who have no school that day). If some types of households are surveyed more often during some days of the week, results could be biased without these fixed effects. Our results are almost identical without these fixed effects.

[6] Based on 30 categories combining gender, age, socio-professional category and main occupation.

[7] See Table A.4, comparing average household characteristics from the EGT and from administrative data for the year 2011.

[8] See Table 1 for household-level descriptive statistics on the whole sample.

distance is 29 kilometres. The average self-reported travel time is 107 min. The average number of trips per day is 4.3, with an average trip distance of 8.3 kilometres and a trip duration of 29 min. The average distances are comparable to the national patterns reported in the 2019 wave of the French national transport survey (SDES, 2021): in terms of daily mobility (excluding long-distance trips), the average French person travels 26 kilometres per day and the average trip is 8.7 kilometres long.[9] Travel time is higher in the Paris area than on average for France, presumably due to a higher congestion: The average trip duration is 20 min in France but 29 min in the Paris area. Finally, the modal shares of public transport and active modes are higher in the Paris area than the national average, and that of car lower: the average Paris area resident makes 27% of her trips by public transport and 33% by active mode, versus only 9% for public transport and 26% for active mode for the average French person. She uses the car for 39% of her trips versus 63% for the average French person. In In table A.5, we compare mean observed characteristics for the full sample and the sample of adults with at least one trip recorded. Mobile individuals – which is our population of interest – are representative of the whole sample in terms of locational characteristics, but they are on average more educated and richer, and are more likely to be full-time employed.

The survey records and geolocates all the places visited by each individual during the day with a grid size of 100 metres*100 m. For each trip defined by an origin and destination, the data describes each journey stage, a journey stage being defined as a single travelling mode.[10] Only the trips starting or finishing within the Paris area boundaries are geolocated. For the 0.8% of trips starting (finishing) in the Paris area but finishing (starting) in another region, we do not know the departure (arrival) point's location, nor the trip distance.

We add three variables not readily available in the EGT data:

- **Actual distances travelled:** The EGT data only contains as-the-crow-flies distances for each trip and journey stage. We obtained estimated actual distances based on a shortest-path algorithm from the regional transport authority Ile de France Mobilités. Both as-the-crow-flies and actual distances are not available for the 0.8% of trips made outside the Paris area.
- **Income quintiles:** in the EGT data, household income is self-declared and interval-coded in nine income brackets, with a non-response rate of 6%. In order to estimate the relationship between income quintiles and contribution to emissions, we estimate the full distribution of income using an interval regression imputation method (Royston, 2007). Since the method assumes an underlying normal model for the partially observed imputed variable – given other predictors – and the distribution of income is usually log-normal, we apply a log transformation to the income brackets declared in the EGT. We then estimate the continuous income variable by including several socio-economic factors known to be correlated with income in the interval-coded regression.[11] For households with a missing income bracket, we use a predictive mean matching imputation method (Little, 1988), using the same predictors and similarly predict their continuous income. Finally, we transform the obtained continuous variable of household monthly income into a variable of annual income per consumption unit, using the OECD equivalence scale. Table A.4 shows that the average income per consumption unit obtained with this imputation is close to the average income per consumption unit in the Paris area in 2011 obtained from administrative data. In the regression analyses, we build quintiles of income based on this continuous income variable.
- **Rail public transport stops within a one kilometre radius:** We create a binary variable indicating whether a household lives less than one kilometre away from a rail public transport stop. To do so, we combine geocoded information on the location of each rail public transport stop in 2010, including subway, regional train and streetcar, with information on households' place of residence.

Beside the transport survey, we use emission factor data by transport mode (and by type of vehicle for cars and two-wheelers) coming from a variety of sources, detailed in the next section and in Appendix A.1.

3.2. Methodology

Building individual measures of contribution to pollution. We estimate individual- and trip-level contributions to local and global pollution based on the detailed information contained in the EGT. For local pollutants, we use NOx and $PM_{2.5}$. For global pollution, we use CO_2 emissions. The total emissions of pollutant P for individual i during the day are the sum of her emissions at the trip level, with T the total number of trips made during the day:

$$E_{P,i} = \sum_{t \in T} E_{P,i,t} \tag{1}$$

Emissions at the trip level $E_{P,i,t}$ are themselves the sum of emissions for each journey stage j that t is made of. Note that we cannot calculate emissions for the trips starting or finishing outside the Paris area, for which we do not have trip distances. For each individual i and each journey stage j, we know the estimated journey distance in kilometres $d_{j,i}$, the travel mode used m, the mode-specific, or, for personal vehicles, the vehicle-specific emission factor $e_{P,j,i}$ in grams per kilometre, and the number of passengers $n_{j,i}$ if the mode used is a private vehicle (car or two-wheeler). For all the journey stages done with a collective transport mode, the number of passengers is set to one, as an average occupancy rate is included in the estimation of their emission factor. Emissions at the journey stage are simply the product of distance and the emission factor, divided by the number of passengers:

$$E_{P,i,t} = \sum_{j \in J}^{J} d_{j,i} e_{P,j,i} \frac{1}{n_{j,i}} \tag{2}$$

Appendix A.1 details the sources used and data processing steps to obtain emission factors that are comparable across modes for each of the three pollutants considered. To summarize, active modes (walking, cycling, skate-boarding, etc.) have a zero emission factor for all three pollutants. The train and subway have a zero emission factor for NOx and CO_2,[12] but not for $PM_{2.5}$, due to the emissions from train brakes. For transportation modes with positive emission factors — buses, two-wheelers and cars for NOx and CO_2, plus electric public transport for $PM_{2.5}$, we use data from different sources, described in Appendix A.1.

Emission factors can exist in two versions: the "true", on-road emission factor, which varies with the vehicle speed, quality of the road and driving conditions; and the type-approval values reported by

[9] The sample is slightly different because individuals not travelling during the day and individuals aged 6 to 17 are included in the national sample.

[10] For example, a work commuting trip by subway including one change will include four journey stages: the first stage is the journey by foot from home to the subway station; the second stage is the subway journey with the first metro line, finishing at the subway station where the commuter changes lines; the third stage is the subway journey with the second metro line, finishing at the subway station near the workplace; the fourth stage is the journey by foot from the subway station to the workplace.

[11] List of predictors: age, age squared, gender, education level and socio-economic class of the household head; socio-economic category of her partner; number of household members working full-time and number working part-time; housing status of the household; dummy for whether the household is eligible to family allowances based on the number and age of children, to proxy for social transfers.

[12] These modes embody some NOx and CO_2 emissions, but given our focus on air pollution mitigation *in the Paris area* we think it is satisfying to focus on exhaust emissions only.

Table 1
Summary statistics — Individuals ≥18 years old with at least one trip recorded.

	Mean	Sd	N
Residence: Paris	21%		23,690
Inner suburbs	37%		
Outer suburbs	42%		
Education: Primary school	6%		23,636
Secondary education	39%		
Higher education < 3 years	14%		
Higher education ≥ 3 years	35%		
Still in education	7%		
SES: Farmers	0%		22,495
Manual workers	11%		
Office workers	19%		
Intermediate professions	19%		
Traders and craftspeople	3%		
Managers and executives	20%		
Pensioner	20%		
Other	7%		
Age	45.72	16.62	23,690
Net household income (€ 2010)	40,910.90	26,462.14	23,683
Net household income per consumption unit (€ 2010)	24,298.50	14,725.03	23,683
Actual distance to workplace (km)[a]	14.77	14.35	8,374[a]
Nb of trips prev. day	4.32	2.40	
Modal share for trips: Car	39%		23,690
Collective transportation	27%		
Bicycle	2%		
Two-wheeler	2%		
Walking	31%		
Other mode	<1%		
Daily distance travelled (km)	28.88	31.60	23,690
Daily travel time (min)	107.19	76.06	23,690
Average trip distance (km)	8.26	10.53	23,444
Average trip duration (min)	29.30	24.26	23,458

Note: Source: EGT data. Observations weighted with EGT individual-level sampling weights. SES stands for Socio-Economic Status. The eight categories follow the aggregate classification of the French Statistical Institute. Household income is estimated with a predictive mean matching imputation method.

[a]Actual distance to workplace is only observed for workers making one commuting trip starting exactly at home and finishing exactly at work during the day, hence the lower sample size.

car manufacturers, subject to a maximum value under the EU emission standards regulation. For NOx and $PM_{2.5}$, we use on-road emission factors because the discrepancy between type-approval and real-world emissions is large,[13] For $PM_{2.5}$ specifically, using on-road emission factors also allows us to take into account emissions from tyres and brakes – rather than only those from exhaust – which represent a substantial share of emissions (OECD, 2020). For CO_2, type-approval values seem more relevant for two reasons. First, there exist car model-specific CO_2 emission factor data, which we can link to the information on the vehicles owned by EGT households to estimate precise emission factors varying by fuel type, age and horsepower. Second, while for local pollutants, type-approval values are drastically underestimating real-world emissions, for CO_2 the difference between type-approval and real-world emissions is relatively small.[14]

Table 2 shows the emission factors obtained for each pollutant and transport mode. The car emission factor reported in the table is the one imputed when an individual travels with a car that she does not own. For journey stages done with a car owned by the household, we find a large variation in emission intensity values, as illustrated in Figures A.1, A.2 and A.3, showing emission intensity values by transport mode and pollutant.[15] The heterogeneity in emission intensities is the highest for NOx and for private cars, with few extremely high values corresponding to old light-commercial vehicles (included in the car category).

Exact factor decomposition analysis:. Starting from Eq. (2), we re-write individual emissions in the form of an extended Kaya identity (see Wang et al. (2005), Mahony (2013), Bigo (2019) for other examples), as the product of distance, modal share and emission intensity by mode. Note D_i the total distance travelled by individual i, $S_{m,i}$ the modal share of mode m, and $I_{P,m,i}$ the average emission intensity of mode m used by individual i for pollutant P (using the notations from Eq. (2), $I_{P,m,i} = e_{P,m,i}\frac{1}{n_{m,i}}$). If we call $d_{m,i}$ the total distance travelled by individual i with mode m and $E_{P,m,i}$ the total emissions of pollutant P from using mode m, we have:

$$E_{P,i} = \sum_{m\in M} D_i \frac{d_{m,i}}{D_i}\frac{E_{P,m,i}}{d_{m,i}} = \sum_{m\in M} D_i S_{mi} I_{P,m,i} \quad (3)$$

Given this multiplicative structure, we can use the Log Mean Divisia Index (LMDI) developed by Ang (2004) and Ang (2005) to decompose differences in emissions into differences in distance, modal choice, and emission intensity. We group individuals by quintile of emissions, and calculate how much each of these three components explains the observed difference in emissions between a reference individual from the middle quintile, and reference individuals from quintiles 1, 2, 4 and 5 of emissions. The LMDI decomposition has been originally developed to explain changes in emissions over time and this is how it has been applied mostly in the literature. Ang et al. (2015) suggest that the LMDI is also appropriate to compare emissions between countries or regions at a given point in time, and this cross-country version has been used in some applications (Liu et al., 2017). Although the method has, to our knowledge, not been applied to individual-level data as we do here, our decomposition across quintiles of individuals is mathematically equivalent to the cross-country case.

[13] Baldino et al. (2017) compare on-road and type-approval emission factors for a sample of diesel cars registered after 2011, brought under the spotlight by the 2015 Volkswagen scandal. They report an average factor of 4 between the type-approval and real-world NOx values.

[14] For the same sample of diesel cars, Baldino et al. (2017) find that on-road CO_2 emissions are on average only 30% higher than type-approval values.

[15] For collective transportation, these emission intensity values are equal to the emission factors reported on Table 2. For private transportation, they are defined as the emission factor divided by the number of passengers.

Table 2
Emission factors by mode.

Type of emission value	Unit	NOx (mg) Real-world	$PM_{2.5}$ (mg) Real-world	CO_2 (g) Type-approval
Walking	per passenger-km	0	0	0
Cycling	per passenger-km	0	0	0
Street-car	per passenger-km	0	7	0
Metro	per passenger-km	0	7	0
Train	per passenger-km	0	7	0
Bus	per passenger-km	242	5	117
Taxi	per passenger-km	1,178	127	332
Car not owned by the household	per vehicle-km	589	63	166
Two-wheeler not owned by the household	per vehicle-km	86	21	65

Notes: All the assumptions are explained in Appendix A.1.

For each pollutant P, we generate a reference individual by quintile of emissions Qk, which we defines as an individual having the average distance D_{Qk}, modal share $S_{m,Qk}$, and emission intensity $I_{m,Qk}$ of her quintile Qk, $k = 1..5$.[16] For the reference individual of quintile Qk, the extended Kaya equation reads:

$$E_{P,Qk} = \sum_{m\in M} D_{Qk} S_{m,Qk} I_{P,m,Qk} \tag{4}$$

As recommended in Ang et al. (2015), we define a benchmark individual, here the reference individual from quintile 3, to which we compare the reference individuals from each quintile. We then apply the LMDI decomposition. The total (tot) difference in emissions between Qk, $k = 1, 2, 4, 5$ and Q3 can be decomposed into the difference in the distance (D), modal share (S) and intensity (I) components:

$$E_{P,Qk} - E_{P,Q3} = \Delta E_{P,Qk-Q3,tot} = \Delta E_{P,Qk-Q3,D} + \Delta E_{P,Qk-Q3,S} + \Delta E_{P,Qk-Q3,I} \tag{5}$$

Following Ang (2005), this can be rewritten:

$$E_{P,Qk} - E_{P,Q3} = \sum_{m\in M} w_m ln(\frac{D_{Qk}}{D_3}) + \sum_{m\in M} w_m ln(\frac{S_{m,Qk}}{S_{m,3}}) + \sum_{m\in M} w_m ln(\frac{I_{m,Qk}}{I_{m,3}}) \tag{6}$$

Where w_m is defined as:

$$w_m = \frac{E_{P,Qk,m} - E_{P,Q3,m}}{ln(E_{P,Qk,m}) - ln(E_{P,Q3,m})} \tag{7}$$

And $E_{P,Qk,m}$ are the emissions of pollutant P associated with mode m for quintile Qk.[17]

Regression analysis:. We investigate the individual socio-economic and demographic characteristics associated with emissions in two steps: first, we examine for each pollutant the characteristics associated with being a top emitter, which we define as being in the top quintile (top 20%) of the emission distribution. The reason to look at this discrete outcome – being a top emitter – rather than at the continuous emission variable is twofold: first, emissions are fat-tailed and the normality assumption of the residuals is likely to be violated under a standard linear model. On the other hand, the high number of zeros makes a log-transformation of the emission variable challenging (Bellégo et al., 2021). Second, it seems relevant to focus on the high emitters from a policy perspective, since this group is more likely to bear the cost of policies making emissions more costly and oppose them.

We estimate a logit model for the three pollutants NOx, $PM_{2.5}$ and CO_2. For pollutant P, writing x the vector of covariates and β the vector of parameters to estimate, the model writes:

$$Pr(E_{P,i} \in Q_5 | x) = \Lambda(x\beta) = \frac{exp(x\beta)}{1 + exp(x\beta)} \tag{8}$$

In a second step, we seek to understand the role of the distance, modal choice and emission intensity in mediating the association between individual characteristics and emissions. We run separate regressions examining the relationship between individual characteristics and distance, modal choice (as captured by the likelihood to use a car at least once in the day) and emission intensity (as captured by the average emission intensity of car trips made during the day).

For the distance regression, we estimate a log-linear model. Writing $ln(y)$ the natural logarithm of total distance travelled during the day, β_1 the vector of parameters to estimate and ϵ an error term, the model writes:

$$ln(y) = x\beta_1 + \epsilon \tag{9}$$

For the modal choice regression, we take as an outcome a binary variable equal to one when the individual has a strictly positive car modal share, and estimate a logit model. Writing S_{car} the modal share of car, and β_2 the vector of parameters to estimate, the model writes:

$$Pr(S_{car} > 0 | x) = \Lambda(x\beta_2) = \frac{exp(x\beta_2)}{1 + exp(x\beta_2)} \tag{10}$$

For the emission intensity regression, the outcome variable is the average emission intensity of car trips, and the sample is restricted to individuals with a positive car modal share. We estimate a linear model, and our results should be interpreted conditionally on driving a car on that day. Writing $I_{P,car}$ the average emission intensity of the car trips for pollutant P, β_3 the vector of parameters to estimate, and μ an error term, we estimate the following model for the three pollutants NOx, $PM_{2.5}$ and CO_2:

$$I_{P,car} = x\beta_3 + \mu \tag{11}$$

We run these four regressions on two samples: the full sample of individuals, and the sample of individuals in employment. Beside emitting more on average than non-working individuals, individuals in employment have more constrained trips, so it seems particularly important to understand the job characteristics associated with emissions. The characteristics of interest for the full sample of individuals are location, public transport availability (as proxied by proximity to a rail public transport stop), car availability,[18] gender, household size, household income (we define a "low-income" category for the bottom income quintile and a "top-income" category for the top income quintile), and employment status. For the regression on the sample of individuals in employment, the characteristics of interest are public transport and car

[16] This reference individual has emissions $E_{P,i}$ that differ from the average emissions of her quintile, given the multiplicative form of the decomposition formula: the product of averages is not the average of the product.

[17] The modal share of bus, two-wheeler and car is 0 for the bottom quintile of NOx emissions. To be able to apply the log formula, we apply the "Small Value" strategy suggested in Ang and Liu (2007), that is, we replace the zero values by $\delta = 10^{-100}$.

[18] The vehicle availability variable is defined at the individual level and concerns the reference day, it is different from the variables of car ownership defined at the household level.

availability, age,[19] gender, household size, household income, and the following job characteristics: type of commute flow, distance to work, type of workplace, type of occupation, and a dummy for having atypical working hours.[20] In all the regressions, we also control for survey day-specific effects with three variables: day-of-the-week fixed effects (we do not have information on the exact survey date); a dummy variable indicating whether the individual encountered a problem with taking transport that day (such as a car breakdown, a public transport strike, or bad weather conditions); and a dummy variable indicating whether the individual was on holidays or on sickness leave that day.

4. Results

4.1. How unequal are contributions to emissions?

Fig. 2 illustrates the large inequalities in daily emissions at the individual level using Lorenz curves: on a representative weekday, the top 20% of NOx emitters contribute 85% of NOx emissions, the middle 48% contribute 15%, and the bottom 32% have a zero contribution[21] (Fig. 2(a)). The top 20% of PM2.5 emitters contribute 78% of PM2.5 emissions, the middle 62% contribute 22%, and the bottom 18% have a zero contribution (Fig. 2(b)). The top 20% of CO_2 emitters contribute 75% of emissions, the middle 48% contribute 25%, while 32% have a zero contribution (Fig. 2(c)). Top emitters are not exactly the same across pollutants but the correlation is high, with a correlation coefficient between individual-level NOx and CO_2 emissions of 0.82. Inequalities of contribution to emissions at the trip level (as defined by Eq. (2)) are higher than at the individual level, reflecting the high dispersion of trip distances (see Figure A.4).

The concentration of daily-mobility-induced CO_2 emissions that we find is close to Bel and Rosell (2017)'s results on Barcelona, where the top 20% of emitters contribute 74% of CO_2 emissions. We further document that the distribution of local pollutant emissions is even more unequal than that of carbon emissions. In the next section, we investigate whether top emitters emit more because of longer distances travel, because of a higher reliance on high-emitting modes, because of more polluting vehicles within modes, or a combination of these three factors.

4.2. What explains high emissions?

Fig. 3 show the results of the LMDI decomposition for NOx, $PM_{2.5}$ and CO_2 emissions (see tables A.6, A.7, A.10, A.11, A.9 and A.8 in Appendix for the components' values for each quintile and the LMDI Deltas, $\Delta E_{P,Qk-Q3,D}$, $\Delta E_{P,Qk-Q3,S}$, and $\Delta E_{P,Qk-Q3,I}$ from Eq. (5)). For the local pollutants NOx and $PM_{2.5}$, emission intensity, distance and modal share contribute about the same way in explaining the difference between Q5 and Q3. For example, for NOx, differences in emission intensity contribute 36%, differences in distance 34%, and differences in modal share 30%. To give an idea of the differences, individuals from the top NOx quintile Q5 travel on average 62 km a day against 26 km for those in the middle quintile Q3; they travel by car for 92% of this distance against 37% for Q3, and have car trips emitting 794 mg/km against 300 mg/km for Q3 (see Table A.6). In contrast, for CO_2 emissions distance and modal share are more important than emission intensity: differences in distances explain 58% of the difference in emissions between Q5 and Q3, differences in modal share explain 36%, while differences in emission intensity explain only 6%. To summarize, the top 20% of NOx and $PM_{2.5}$ emitters are individuals combining long distances, a high car modal share, and more polluting cars, while the top 20% of CO_2 emitters combine long distances and a high car modal share, but have cars that are only slightly more CO_2 intensive than the average car.

Regarding the factors contributing to the difference between the bottom quintiles and Q3, the main difference is the more important role of the distance component for $PM_{2.5}$, compared to NOx and CO_2. This is because subway and train, which are the only transport modes taken by 32% of the individuals, do not have a zero $PM_{2.5}$ emission factor while they have a zero NOx and CO_2 emission factor. The bottom quintile for $PM_{2.5}$ emissions then includes more individuals travelling very short distances and not relying on subway and train but only on walking (the average distance of the bottom $PM_{2.5}$ quintile is only 3 km, versus 16 km for the bottom CO_2 and NOx quintile).

The distance, modal choice and emission intensity components are of course not independent from each other: the correlation coefficient between distance and the car modal share is 0.2, and the emission intensity component is by definition only calculated for modes with a strictly positive modal share. However, conditional on having a positive car modal share, the emission intensity of those trips is barely correlated with distance[22] and with the modal share[23] Given this absence of correlation, different groups of people may be affected by mitigation policies aiming at a reduction in cars' emission intensity (such as vintage-based low-emission-zones or subsidies for electric cars) compared to mitigation policies tackling distance (such as policies to increase urban density) or aiming a reduction in the car modal share (such as public transport subsidies). This justifies investigating separately the correlation between socio-economic characteristics and distance, modal choice and emission intensity, as we do in the next section.

4.3. Who emits pollution?

4.3.1. Analysis on the sample of all individuals travelling

The first subgraph of Fig. 4 shows, for each pollutant outcome, the average marginal effects of the different individual characteristics on the propensity to be in the top 20% of emitters[24] (see regression outputs in Table A.12). The three next subgraphs show the association between these characteristics and each of the three channels distance, modal share and emission intensity (see regression outputs in tables A.14 and A.15). All the coefficients should be interpreted as correlations rather than causal effects: it is easy to think of omitted variables that could influence at the same time some covariates and the outcome variable. For example, a preference for driving is likely to decrease the propensity to live close to a rail public transport stop and increase the propensity to use a car.

We make four observations. First, most marginal effects are close in magnitude across the three pollutant outcomes and only few of them have opposite directions. This is logical given the high correlation between being a top NOx, top $PM_{2.5}$, and top CO_2 emitter. One exception is the income variable, which plays a different role for local pollutants and for CO_2 emissions: being in the high-income category is associated with a 2.8 percentage point increase in the likelihood to be a top CO_2

[19] We do not control for age in the regression on all individuals because the employment status already captures some age effects, with the distinct categories for students, employed individuals and pensioners.

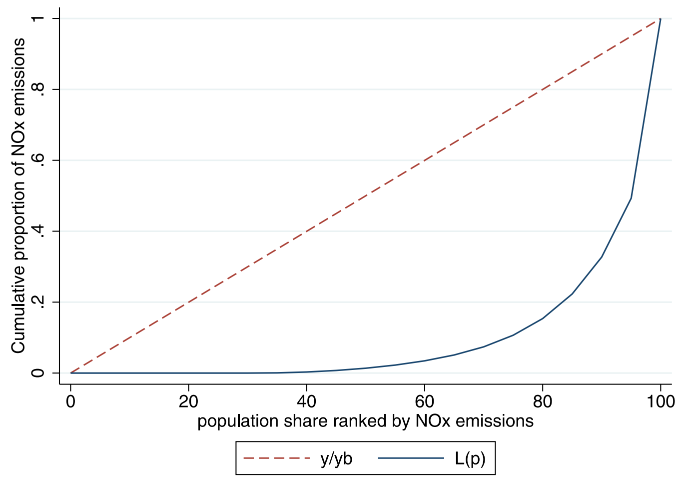
[20] Atypical working hours are defined as going to work or coming back from work before 5 am, or going to work after 4 pm.

[21] Only individuals with at least one trip are in the sample, so those with zero emissions are the ones travelling only with active modes, electric collective transportation or electric car.

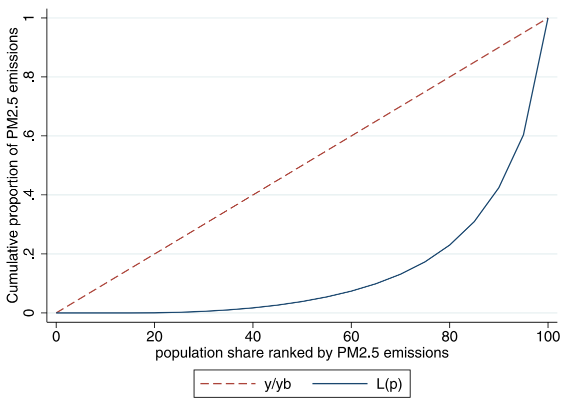
[22] $\rho_{\text{distance,NOx emission intensity}} = 0.05$, $\rho_{\text{distance,PM2.5 emission intensity}} = 0.08$ and $\rho_{\text{distance,CO2 emission intensity}} = -0.002$.

[23] $\rho_{\text{car modal share,NOx emission intensity}} = 0.03$, $\rho_{\text{car modal share,PM2.5 emission intensity}} = 0.006$, $\rho_{\text{car modal share,CO2 emission intensity}} = -0.05$.

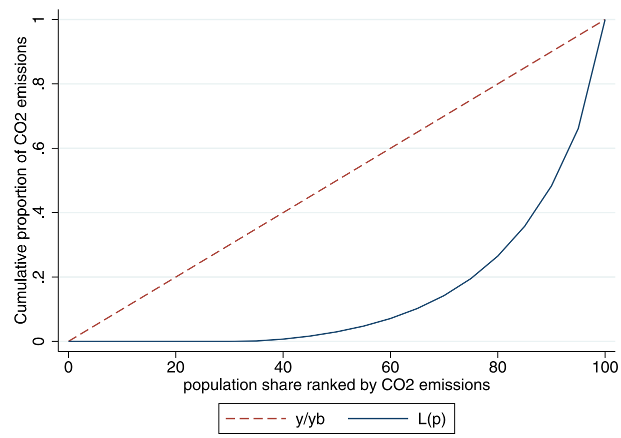
[24] The omitted categories for the categorical variables present in the model are: for the place of residence, we omit living in the inner suburbs; for gender, we omit male; for income, we take as reference the middle 60% and group individuals from the bottom quintile in the "Low-Income" category and individuals from the top quintile in the "High-Income" category. For the activity status, we omit unemployed individuals.



(a) NOx emissions



(b) $PM_{2.5}$ emissions



(c) CO_2 emissions

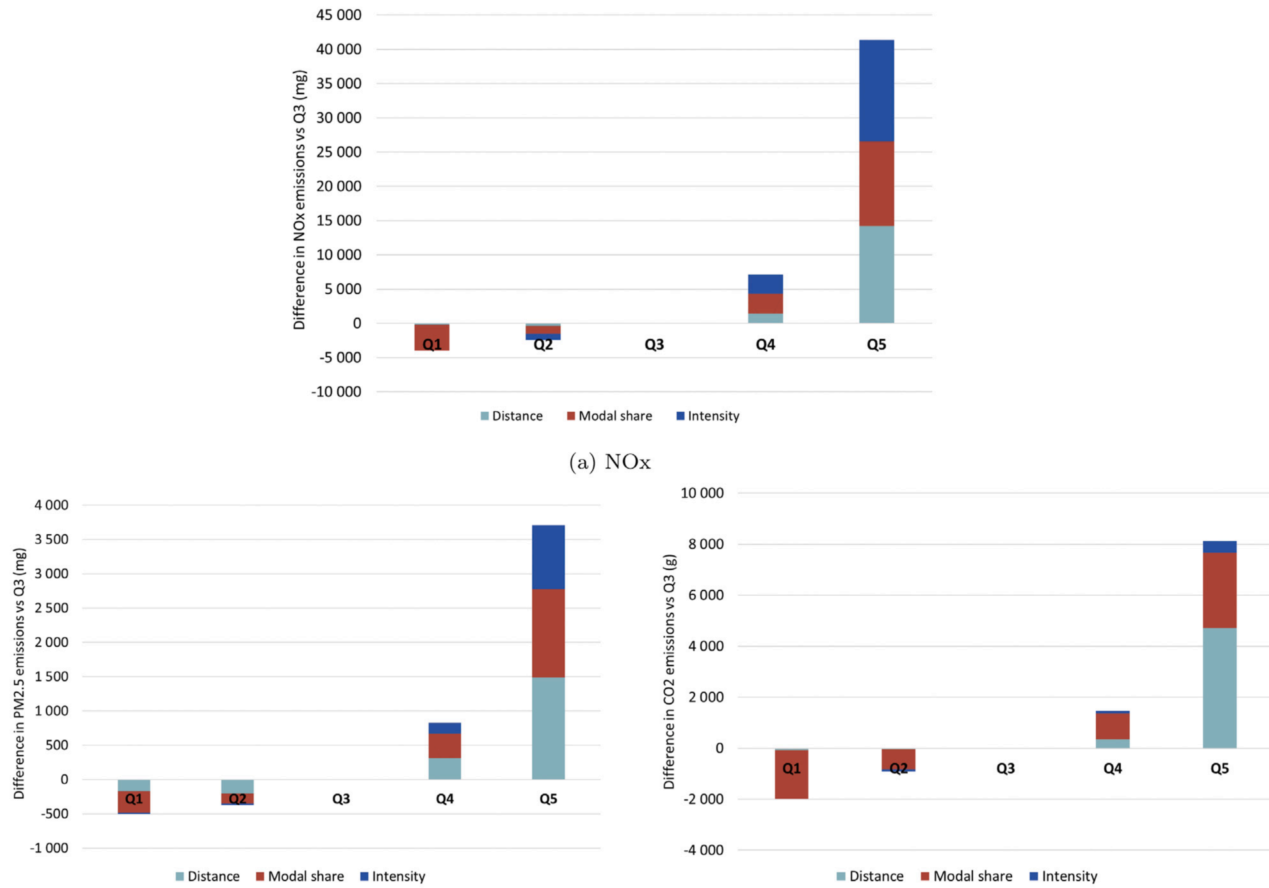
Fig. 2. Lorenz curves for contributions to emissions at the individual level.
Note: the *x*-axis shows the percentiles of individual-level emissions and the *y*-axis shows the share of total emissions generated by all the individuals below that percentile. Observations are weighted with EGT individual-level sampling weights. The red curve shows how the distribution would look like if everyone contributed equally to emissions. Source: EGT data. Sample: all adults with at least one trip on the day.

emitter, while it has no significant effect on the likelihood to be a top NOx or $PM_{2.5}$ emitter. Since both local pollutant and CO_2 emissions are the product of distance, modal choice and emission intensity, and only emission intensity differ across the two types of pollutants, the correlation between income and vehicles' emission intensity must differ across the two types of pollutant. This is indeed what we see in the last subgraph: being in the top income quintile is associated with a lower local pollution emission intensity of car trips but with a higher CO_2 emission intensity. This is true both before and after controlling for the type of car owned.[25] The positive correlation between high income and CO_2 emission intensity can be explained by the fact that richer households generally own heavier, larger and more powerful cars, attributes that correlate positively with the CO_2 emission factor.

On the other hand, being in the low-income category is associated with a higher emission intensity of car trips across the three pollutants, which may be due to the fact that the cars owned by poorer households are older and more often light-commercial vehicles, two attributes positively correlated with emission intensity. This positive association somehow contrasts previous findings from Barnes et al. (2019), where in the UK, the areas with the highest poverty rate are the ones where the cars owned have the lowest NOx, PM and CO_2 emission factor. The difference between our results may be due to differences in the context considered – the Paris area vs. the entire UK – in the data scope – car trips from daily mobility vs. all car trips – in the type of relationship examined – partial correlation holding other characteristics constant in our case vs. bivariate analysis in the case of Barnes et al. (2019) – or in the methodology used to estimate emission intensity — where we take into account cars' occupancy rate and have a specific emission factor for light-commercial vehicles, while Barnes et al. (2019) do not. Note that in our case, the association fades out for NOx and $PM_{2.5}$ when the type of car owned by a household is controlled for (see columns (2) and (4) of table A.15).

A second observation is that the associations between individual characteristics and top emissions hide mediating channels sometimes having conflicting effects. For example, the null association between low-income and top emissions hides a negative association between low-income and distance, and car use, combined with a positive association with the emission intensity of car trips. Similarly, living in central Paris is associated with shorter distances travelled and a lower propensity to use a car, but for those who do, a much higher CO_2 emission intensity, while living in the outer suburbs is associated with

[25] The last subgraph of Fig. 4 shows the coefficient estimates without controlling for the type of car owned. Columns (2), (4) and (6) of table A.15 show coefficient estimates after controlling for car type and fuel type.



(a) NOx

(b) $PM_{2.5}$

(c) CO_2

Fig. 3. Contribution of distance, modal choice and emission intensity to the differences in emissions, by pollutant.
Note: These graphs show, for each pollutant, the difference in emissions between the reference individuals from quintiles 1, 2, 4 and 5 and the benchmark individual from quintile 3 (total length of the bars), decomposed into differences in total distance travelled, modal shares, and the emission intensity of a given mode. The LMDI formula used is the additive decomposition (Ang, 2004), shown in Eqs. (5) and (6).

longer distances and a higher propensity to use a car but a lower CO_2 emission intensity. Given these associations going in opposite directions depending on the channel considered, a policy aiming at reducing distances travelled will not affect the same group of individuals as one aiming at reducing the emission intensity of cars. The only characteristics for which the association is significant and goes in the same direction is being employed, which is positively correlated with distance, car modal share and emission intensity.

Our third observation concerns the type of characteristics associated with top emissions. Living in the far suburbs, being employed, being a man and having a motorized vehicle available are associated with a higher likelihood to be a top emitter. Living in central Paris, being unemployed or inactive, being a woman and living within one kilometre of a rail public transport stop are associated with a lower likelihood to be a top emitter. The role of employment status, income or household location is consistent with previous findings focussing on carbon emissions at the household or individual level (Nicolas and David, 2009; Barla et al., 2011; Brand et al., 2013; Bel and Rosell, 2017; Blaudin de Thé et al., 2021). We also document an important role for gender, with women having a lower likelihood to be top emitters than men. This dimension has been less frequently investigated, because analysing gender differences in emissions requires having individual-level rather than household-level data. The negative association between being a woman and emissions had been reported by Brand et al. (2013) in the case of motorized passenger travel specifically, and by Bel and Rosell (2017) and Barla et al. (2011) in the case of daily mobility in Barcelona and Quebec city respectively. In contrast, Brand and Preston (2010) found that being a woman was not significantly associated with total CO_2 emissions from transport.

Our analysis of the distance, modal share and emission intensity channels enables us to better understand the underlying mechanisms of this gender difference in the context of Paris: first, conditional on the other covariates, being a woman is associated with 25% shorter distances than being a man. This result can be linked to the urban planning literature emphasizing gender differences in distances travelled (MacDonald, 2016) and, in the case of employed women, to the economic literature finding that women have a shorter maximum acceptable commute than men (Le Barbanchon et al., 2021). A second reason is that being a woman is associated with a lower emission intensity of car trips for all pollutants, due to the combination of a higher occupancy rate and a lower emission factor of the cars used. In contrast, the result does not seem driven by gender differences in car use, conditional on the characteristics controlled for. Indeed, once we control for car availability – which we do in all the regressions presented on Fig. 4 – women are not less likely to use a car than men. They are even more likely to do so in the subsample of employed individuals (see column (2) of table A.17). Had we not controlled for car availability though, being a woman would have been even more negatively correlated with emissions, since only 60% of women have a car available on the survey day, against 75% of men.

Finally, even after including a rich set of socio-economic, spatial and demographic factors as well as controls relative to the survey day, the McFadden's pseudo R-squared of the top emitter regression and the R-squared of the distance and emission intensity regressions are quite low, never exceeding 0.2 (see Tables A.12, A.14 and A.15). This observation suggests an important role for other, potentially unobserved factors in explaining the variation in emissions across individuals, and can be linked to previous findings from similar analyses in other contexts (Brand and Boardman, 2008; Ko et al., 2011; Bel and Rosell, 2017), as well as findings in the tax incidence literature reporting a vast heterogeneity in carbon tax incidence, poorly explained by observable household characteristics (Cronin et al., 2018; Douenne, 2020).

To understand the difference between the partial correlations captured in our multivariate regressions and the unconditional correlations between each characteristic and the outcome, figure A.5 shows the coefficient estimates for the same characteristic and outcomes as in Fig. 4, but based on regressions with only one individual characteristic of interest and no other covariate, except for the survey-day specific variables. We call this second set of coefficient estimates the "unconditional correlations", although we still control for the survey-day specific variables. The results indicate three main differences between the partial and unconditional correlations: (i) being a low-income individual is strongly associated with a lower propensity to be a top emitter and a lower propensity to use a car, when no other characteristics are controlled for; (ii) being a woman is associated with a lower propensity to use a car, while there was no significant difference in car use between men and women when other characteristics were controlled for. This probably reflects the negative correlation between being a woman and having a car available; (iii) being employed is associated with a higher propensity to use a car while being a student is associated with a lower probability while there were no significant effect when other variables were controlled for. This probably also reflects the positive (negative) correlation between being employed (student) and having a car available (and is consistent with the sign of the coefficient when having a car available is not controlled for, as in the second column of table A.14).

4.3.2. Analysis on the subsample of individuals in employment

We next examine the association between employment characteristics and top emissions. The first subgraph of Fig. 5 shows, for each pollutant, the average marginal effects of selected characteristics on the propensity to be in the top 20% of the emitters for the subsample of individuals in employment[26] (see regression outputs in Table A.13). The three next subgraphs show the association between these characteristics and the three channels of distance, modal share and emission intensity (see regression outputs in tables A.16, A.17 and A.18).

Some observations made based on the analysis of the entire sample still hold: the correlations between employment characteristics and being a top emitter are close in direction and magnitude across the three pollutants. Furthermore, many characteristics correlate positively with one channel and negatively with another. For example, compared to living in the suburbs and working in Paris, living and working in the suburbs is associated with shorter distances travelled but a higher propensity to use the car, which results in a higher propensity to be a top emitter. On the other hand, compared to having an intermediate profession, being a craftsworker is associated with shorter distances travelled but a much higher emission intensity of car trips (probably due to the more widespread use of light-commercial vehicles), which results in a null association with being a top emitter. Finally, although the explanatory power of the different characteristics is slightly higher than for the analysis of the full sample, it remains limited.

We highlight some associations between job characteristics and the likelihood to be a top emitter which, to the best of our knowledge, had not been documented before. Compared to commutes between the suburbs and Paris centre, having to commute from suburbs to suburbs is associated with a 13 to 17 percentage point increase in the likelihood to be a top emitter, depending on the pollutant. This was expected and probably reflects the low density of the radial Parisian public transport network in the suburbs, which constrains car use for this commute type.

Having atypical working hours – which is the case for 2.6% of employed individuals in our sample – is associated with an increase by around 5 percentage points in the propensity to be a top emitter across the different pollutant outcomes. The result is driven by a higher propensity to use a car, likely reflecting the lower availability of public transport at night or very early in the morning. Compared to working in an office, working in a factory is also associated with an increase by around 5 percentage points in the likelihood to be a top emitter, also driven by a higher likelihood to use a car. This may reflect the relative poor public transport accessibility of industrial zones, compared to areas with a high density of office space.

Concerning the type of occupation, being a technician, qualified manual worker, shopkeeper or company head is associated with a higher likelihood to be a top emitter of local pollutants, compared to having an intermediate profession. Being a technician, shopkeeper or company head is associated with a higher likelihood to be a top CO_2 emitter. These associations are driven by the higher propensity to use a car for all these occupations. Being a technician or company head is also associated with longer distances travelled. Being a manual worker or shopkeeper is also associated with a higher emission intensity of car trips, maybe partly due to a more widespread use of light-commercial vehicles for these professions.

That some professional categories seem highly reliant on car could play a role in the political economy of opposition to policies regulating car use. We find a partial overlap between the occupations associated with top emissions and the occupations overrepresented in the Yellow Vest movement, a highly publicized wave of protests that took place in France in 2018 and 2019, initially as a reaction against the increase in fuel costs induced by the planned increase in the carbon tax. According to a face-to-face survey carried out on the entire French territory (N=863)[27] (André et al., 2019), seven occupations were overrepresented during Yellow Vest protests on roundabouts: craftsworkers, qualified manual workers, farmers and public sector office clerks were overrepresented among men compared to their share in the labour force, and shopkeepers, nurses and personal domestic service workers were overrepresented among women. Of these, four are associated with a higher likelihood to be a top emitter or/and to use a car in our analysis: craftsworkers, qualified manual workers, shopkeepers and farmers.[28] Their participation to the Yellow Vest Movement may be influenced by an objectively higher carbon tax incidence, or at least the perception of a high reliance on car. The three remaining categories are not associated with higher emissions in our analysis focused on the Paris area: nurses and public sector office clerks are part of the reference category or not significantly different from it regarding emissions and car use; workers from the Personal Domestic Service sector are associated with a significantly lower likelihood to be top emitters, due to a combination of much shorter distances travelled and a lower propensity to use a car. Several reasons may underpin the over-representation of these three categories in the nationwide Yellow Vest movement: they may be associated with higher emissions

[26] The omitted reference categories for employment characteristics are: for the place of residence combined with the place of work: individuals living in the suburbs and working in Paris; for the workplace type: working in an office; for socio-professional category: intermediate professions, which include public sector jobs such as school teacher or nurse and private sector jobs such as customer service managers.

[27] Out of 1,333 survey answers collected on the roundabouts where Yellow vests gathered and during demonstrations, the occupation could be retrieved for 883 individuals. Although the sample was not randomly chosen, several techniques were deployed to try and reach a representative sample of Yellow Vest participants, such as varying survey times or randomly selecting participants at different locations during demonstrations. The response rate was high at 87%.

[28] Included in the shopkeeper category in our analysis, given the low sample size of this category in the Paris area.

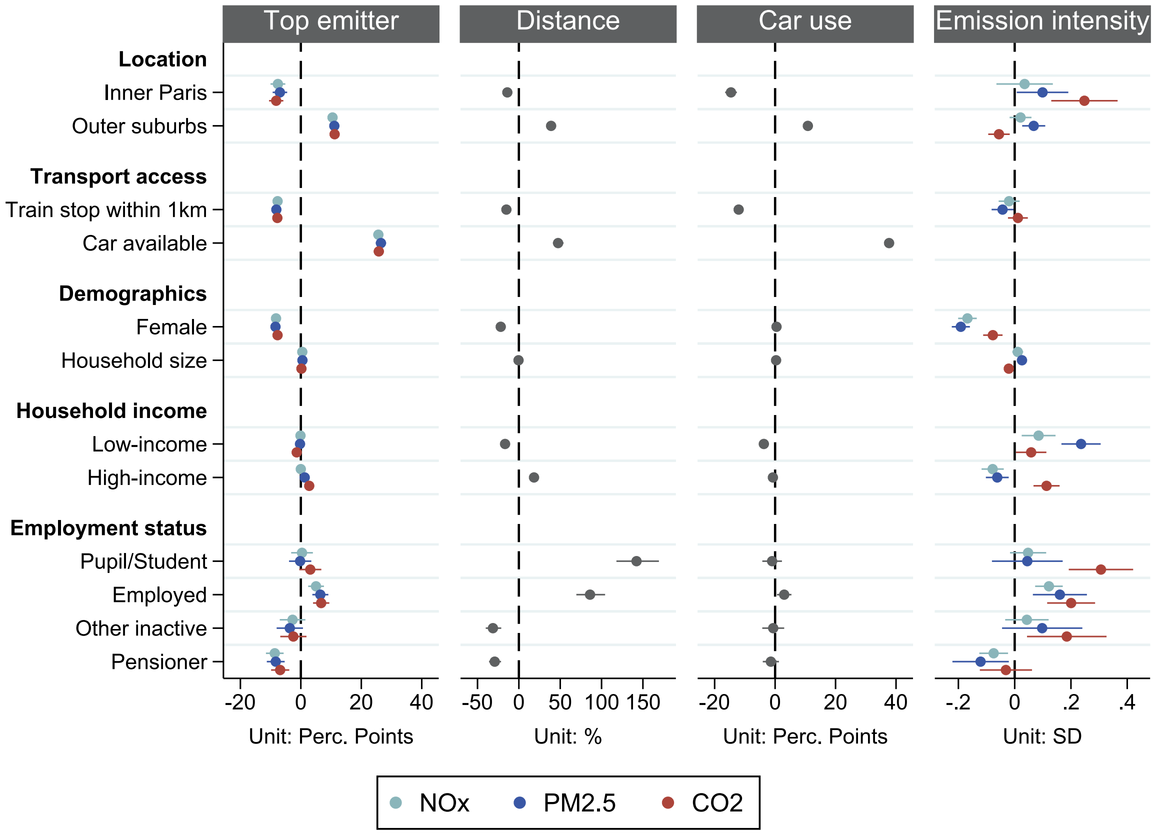


Fig. 4. Regression coefficients of the top emitter, distance, car use and cars' emission intensity regressions, all individuals sample.
Notes: from left to right: selected x covariates are listed on the left, by category. Omitted categories for categorical variables: Location: inner suburbs; Gender: male; Employment status: unemployed. All the regression models also include survey-day fixed effects and control variables for problems with taking transport, being on leave or on sickness leave on the survey day. Standard errors are clustered at the household level. The first panel shows the average marginal effect of each characteristic on the likelihood to be among the top 20% of NOx (in light blue), $PM_{2.5}$ (dark blue) and CO_2 (red) emitters, expressed in percentage points. The second panel shows the percent change in the total daily distance travelled associated with each characteristic, in %. We have transformed the β coefficients from the log-linear model to be able to interpret them as percent changes, knowing that a 1-unit change in x corresponds to an increase in distance by $(e^{\hat{\beta}} - 1) * 100$. The third panel shows the average marginal effect of each characteristic on the likelihood to use the car at least once during the day, expressed in percentage points. The fourth panel shows the change in the NOx (in light blue), $PM_{2.5}$ (dark blue) and CO_2 (red) emission intensity of the car trips made by the individual, expressed in standard deviation units, associated with each characteristic. Regressions are unweighted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in other French regions, characterized by a lower density of public transport compared to the Paris area; they may be more affected by fuel cost increases despite having lower emissions than other categories, due to their relatively low income levels; or their participation may be motivated by the other reasons put forward in the sociological literature on the movement (such as claims on social and fiscal justice or occupation-specific claims, see André et al. (2019) for a summary of these reasons).

Like for the analysis of the full sample, the coefficient estimates reported in Fig. 5 reflect partial correlations. They can be compared to those obtained in figure A.6 when only the individual characteristic of interest and the survey-day specific variables are included, and no other covariate. The main difference is that working in a factory as opposed to an office, and being a qualified manual worker, technician, crafts worker or company head is more clearly associated with a higher propensity to be a top emitter when other characteristics are not controlled for.

5. Policy implications and conclusion

Inequalities in contribution to transport-related emissions are large in the Paris area, both for carbon and local pollutant emissions. What are the implications of such a high concentration in emissions? First, policies tackling daily mobility emissions may represent a large cost for a small number of individuals, while leaving the majority of individuals unaffected. How much the top emitters are car-dependent and whether they have low-emission alternatives is crucial to estimate the distributional impacts of these policies, and more research is needed to characterize the potential to substitute away from high-emission trips. Second, policy-makers may want to target top emitters specifically. Yet, our regression analysis suggests that such targeting may be challenging, because top emitters are quite a heterogeneous group. Interestingly, although the role of income is at the heart of many policy debates and central to assess the distributional impacts of mitigation policies, we find that in the context of the Paris area, it is only poorly correlated with emissions once other characteristics are accounted for.

Since top emitters combine large distances, a high car modal share and a high emission intensity of car trips, relevant measures to tackle emissions may include the three types of policies included in the Avoid-Shift-Improve framework (Creutzig et al., 2018): that is, policies aiming at reducing distances, at a modal shift, or at a decrease in the emission intensity of car trips. But these policies are expected to affect different groups of individuals, given the different characteristics associated with distance, car use, and the emission intensity of car trips. The mechanisms put in place to compensate the affected groups and avoid fairness issues should therefore be tailored to each policy type. For example, given the positive association between low-income and pollution intensity and the negative association between low-income and distance and low-income and car use, a policy banning the most pollution-intensive cars with no regard for the number of kilometres driven could be socially unfair: it would affect low-income individuals with a pollution-intensive car but driving only few kilometres, but

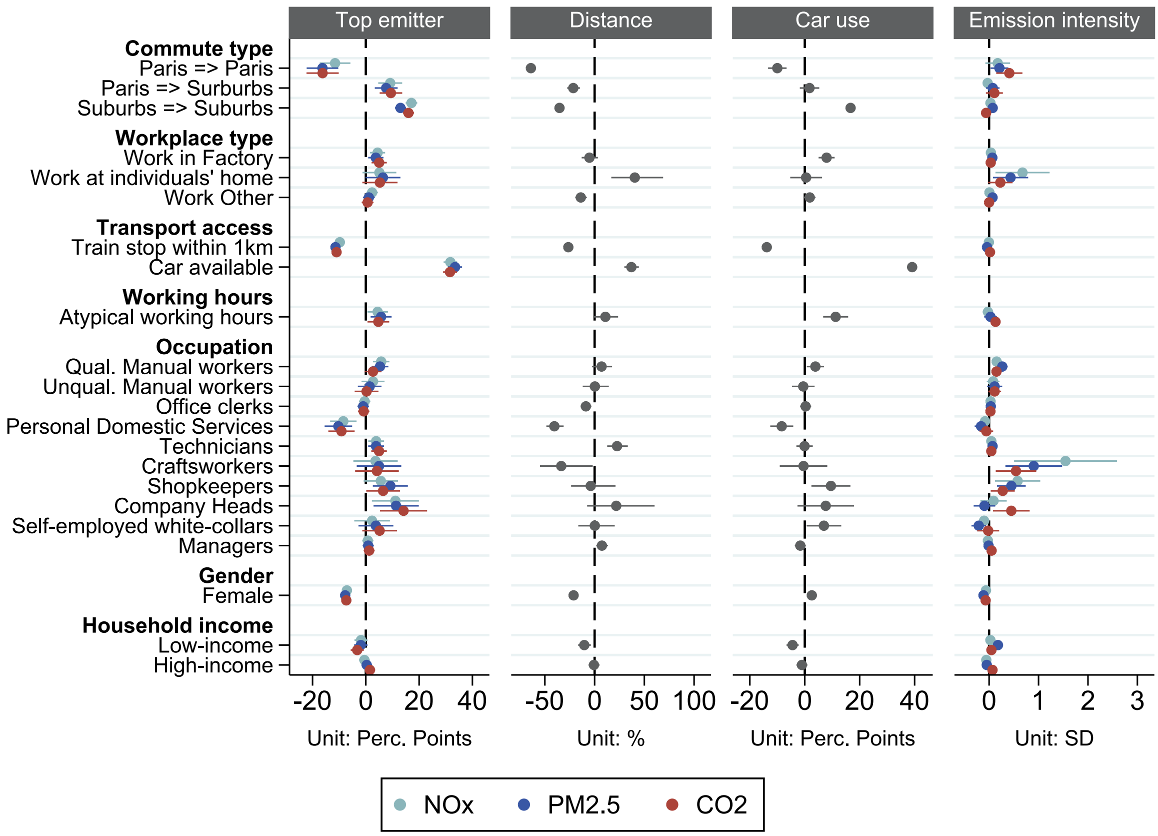


Fig. 5. Regression coefficients of the top emitter, distance, car use and cars' emission intensity regressions, individuals in employment sample.
Notes: from left to right: selected *X* covariates are listed on the left, by category. Omitted categories for the categorical variables: Commute type: Suburbs => Paris; Workplace type: Work in office; Occupation: Intermediate professions; Gender: male. All the regression models also include survey-day fixed effects and control variables for age, age squared, household size, problems with taking transport, being on leave or on sickness leave on the survey day, which coefficients are not included. Standard errors are clustered at the household level. The first panel shows the average marginal effect of each characteristic on the likelihood to be among the top 20% of NOx (in light blue), $PM_{2.5}$ (dark blue) and CO_2 (red) emitters, expressed in percentage points. The second panel shows the percent change in the total daily distance travelled associated with each characteristic, in %. The estimated coefficients from a log-linear model are that a 1-unit change in X corresponds to an increase in Y by $(e^{\hat{\beta}} - 1) * 100$, so we have transformed the obtained coefficients to be able to interpret them as percent changes. The third panel shows the average marginal effect of each characteristic on the likelihood to use the car at least once during the day, expressed in percentage points. The fourth panel shows the change in the NOx (in light blue), $PM_{2.5}$ (dark blue) and CO_2 (red) emission intensity of the car trips made by the individual, expressed in standard deviation units, associated with each characteristic. Regressions are unweighted. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

not high-income individuals having a less pollution-intensive car but driving more kilometres with it, and having higher emissions overall.

Furthermore, policies motivated by air quality goals may have different distributional effects from those motivated by climate goals, because low-income people have more pollution intensive cars across all pollutant types while high-income people have cars with a lower local pollutant intensity but a significantly higher CO_2 emission intensity. As a consequence, policies based on the local pollution intensity of vehicles, such as Low-emission zones, could be more regressive than policies regulating the CO_2 emission intensity of vehicles, such as CO_2 emission standards. More research is needed to compare the actual distributional impacts of the two types of policies.

A caveat to our inequality calculations is that we only take into account emissions on weekdays. Weekday inequalities seem relevant to analyse air pollution mitigation in the Paris area, because ambient pollution tends to be higher on weekdays, where car traffic and economic activity are higher. In contrast, estimating total transport-induced carbon footprints requires examining long-distance trips and weekends as well: residents from the city centre tend to take the plane more often and emit more during their long-distance trips, such that their lower carbon footprint on weekdays may be offset by a higher carbon footprint the rest of the time (Pottier et al., 2020).

Although we use data from 2010, we think that our results are still relevant to explain today's distribution of emissions in Paris. Preliminary results from the new wave of the EGT survey (planned to be carried out between 2018 and 2022, but currently stalled due to the Covid-19 crisis) suggest that the average number of trips, time and distances spent travelling have not changed since 2010 (Omnil-Ile de France Mobilites, 2019). The average modal share changed only slightly, with a small decrease in car use (from 37.8% of the trips in 2010 to 34.4% in 2018), compensated by an increase in active transportation modes and collective transportation.

CRediT authorship contribution statement

Marion Leroutier: Conceptualization, Methodology, Software, Formal analysis, Writing – original draft. **Philippe Quirion:** Conceptualization, Methodology, Resources, Writing – review & editing.

Data and code availability

The codes used for the analysis are publicly available on the following Open Science Frame-work page: https://osf.io/pnyzk/. Some of the data we use have a restricted access, such that we are not able to make the data public. The Readme document of the OSF page explains how to access the raw data.

Acknowledgements

We thank Louis-Gaëtan Giraudet, Francois Bareille, Francois Libois, Ulrich Wagner, Laure de Preux, Olivier Chanel, Mouez Fodha, Nicolas Taconet, an anonymous referee from FAERE, two anonymous referees, seminar participants at CIRED and AgroParisTech, and participants to the AFSE, Mannheim Energy, EAERE and LAGV conferences, for helpful comments and discussion; we thank Olivier Perrussel from Airparif for his advice on emissions estimation, and Olivier Mahieu from Ile de France Mobilités for providing estimated distance data. Marion Leroutier thanks ANR, France for the support of the EUR grant ANR-17-EURE-0001, and the Mistra Foundation for financial support.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.eneco.2022.105941.

References

Airparif, 2021. Inventaire 2018 des émissions atmosphériques en Île-de-France | airparif.

Airparif - IFOP, 2018. Etude De Perception Des Franciliens à L'égard De La Qualité De L'air (The Perceptions of Francilians Towards Air Quality). Technical Report.

Ambec, S., Coria, J., 2013. Prices vs quantities with multiple pollutants. J. Environ. Econ. Manag. 66 (1), 123–140.

André, A., Bedock, C., Bendali, Z., Bernard De Raymond, A., Beurier, A.-G., Blavier, P., Bonin, L., Courty, A., Della Sudda, M., Delozière, G., Dondeyne, C., Elalaoui, C., Flipo, F., Frémont, C., Gaborit, M., Girardin, A., Grémion, T., Guillemin, P., Havard, M., Ivanovic, M., Larrouqué, D., Levain, A., Liochon, P., Pauchard, L., Ormière-Noblet, L., Neyrat, F., Pigot, F., Soulié, F., Reungoat, E., Schnatterer, T., Spéter-Lejeune, L., Alfieri, V., Walker, E., 2019. Enquêter in situ par questionnaire sur une mobilisation en cours : une étude sur les gilets jaunes. Rev. Fr. Sci. Politique 69 (5–6), 869–892, Publisher: Fondation Nationale des Sciences Politiques.

Ang, B.W., 2004. Decomposition analysis for policymaking in energy:: which is the preferred method? Energy Policy 32 (9), 1131–1139.

Ang, B.W., 2005. The LMDI approach to decomposition analysis: a practical guide. Energy Policy 33 (7), 867–871.

Ang, B.W., Liu, N., 2007. Handling zero values in the logarithmic mean divisia index decomposition approach. Energy Policy 35 (1), 238–246.

Ang, B.W., Xu, X.Y., Su, B., 2015. Multi-country comparisons of energy performance: The index decomposition analysis approach. Energy Econ. 47, 68–76.

Baldino, C., Titge, U., Muncrief, R., Bernard, Y., Mock, P., 2017. Road Tested: Comparative Overview of Real-World Versus Type-Approval NOX and CO2 Emissions from Diesel Cars in Europe, ICCT White Paper, International Council on Clean Transportation.

Barla, P., Miranda-Moreno, L.F., Lee-Gosselin, M., 2011. Urban travel CO2 emissions and land use: A case study for quebec city. Transp. Res. D 6 (16), 423–428.

Barnes, J.H., Chatterton, T.J., Longhurst, J.W.S., 2019. Emissions vs exposure: Increasing injustice from road traffic-related air pollution in the united kingdom. Transp. Res. D Trans. Environ. 73, 56–66.

Bel, G., Rosell, J., 2017. The impact of socioeconomic characteristics on CO2 emissions associated with urban mobility: Inequality across individuals. Energy Econ. 64 (C), 251–261, Publisher: Elsevier.

Bellégo, C., Benatia, D., Pape, L.-D., 2021. Dealing with Logs and Zeros in Regression Models. SSRN Scholarly Paper ID 3444996, Social Science Research Network, Rochester, NY.

Bigo, A., 2019. How to explain the past trends in transport CO2 emissions in France?. Chair Energy and Prosperity Working Paper.

Bishop, K.C., Ketcham, J.D., Kuminoff, N.V., 2018. Hazed and Confused: The Effect of Air Pollution on Dementia. Technical Report 24970, National Bureau of Economic Research, Inc.

Braithwaite, I., Zhang, S., Kirkbride, J.B., Osborn, D.P.J., Hayes, J.F., 2019. Air pollution (particulate matter) exposure and associations with depression, anxiety, bipolar, psychosis and suicide risk: A systematic review and meta-analysis. Environ. Health Perspect. 127 (12), 126002.

Brand, C., Boardman, B., 2008. Taming of the few—The unequal distribution of greenhouse gas emissions from personal travel in the UK. Energy Policy 36 (1), 224–238.

Brand, C., Goodman, A., Rutter, H., Song, Y., Ogilvie, D., 2013. Associations of individual, household and environmental characteristics with carbon dioxide emissions from motorised passenger travel. Appl. Energy 104, 158–169.

Brand, C., Götschi, T., Dons, E., Gerike, R., Anaya-Boig, E., Avila-Palencia, I., de Nazelle, A., Gascon, M., Gaupp-Berghausen, M., Iacorossi, F., Kahlmeier, S., Int Panis, L., Racioppi, F., Rojas-Rueda, D., Standaert, A., Stigell, E., Sulikova, S., Wegener, S., Nieuwenhuijsen, M.J., 2021. The climate change mitigation impacts of active travel: Evidence from a longitudinal panel study in seven European cities. Global Environ. Change 67, 102224.

Brand, C., Preston, J.M., 2010. '60-20 Emission'—The unequal distribution of greenhouse gas emissions from personal, non-business travel in the UK. Transp. Policy 17 (1), 9–19.

Büchs, M., Schnepf, S.V., 2013. Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO2 emissions. Ecol. Econom. 90, 114–123.

Chang, T.Y., Zivin, J.G., Gross, T., Neidell, M., 2019. The effect of pollution on worker productivity: Evidence from call center workers in China. Am. Econ. J. Appl. Econ. 11 (1), 151–172.

Creutzig, F., Javaid, A., Koch, N., Knopf, B., Mattioli, G., Edenhofer, O., 2020. Adjust urban and rural road pricing for fair mobility. Nature Clim. Change 10 (7), 591–594.

Creutzig, F., Roy, J., Lamb, W.F., Azevedo, I.M.L., Bruine de Bruin, W., Dalkmann, H., Edelenbosch, O.Y., Geels, F.W., Grubler, A., Hepburn, C., Hertwich, E.G., Khosla, R., Mattauch, L., Minx, J.C., Ramakrishnan, A., Rao, N.D., Steinberger, J.K., Tavoni, M., Ürge-Vorsatz, D., Weber, E.U., 2018. Towards demand-side solutions for mitigating climate change. Nature Clim. Change 8 (4), 260–263.

Cronin, J.A., Fullerton, D., Sexton, S., 2018. Vertical and horizontal redistributions from a carbon tax and rebate. J. Assoc. Environ. Resour. Econ. 6 (S1), S169–S208.

Currie, J., Zivin, J.G., Mullins, J., Neidell, M., 2014. What do we know about short- and long-term effects of early-life exposure to pollution? Ann. Rev. Resour. Econ. 6 (1), 217–247.

Delhaes, D., Kersting, S., 2019. Yellow vests: Ban on diesel vehicles drives germans to the street – to protest. Handelsblatt.

Deryugina, T., Heutel, G., Miller, N.H., Molitor, D., Reif, J., 2019. The mortality and medical costs of air pollution: Evidence from changes in wind direction. Amer. Econ. Rev. 109 (12), 4178–4219.

Dinda, S., 2004. Environmental kuznets curve hypothesis: A survey. Ecol. Econom. 49 (4), 431–455.

Douenne, T., 2020. The vertical and horizontal distributive effects of energy taxes: A case study of a french policy. Energy J. 41 (3).

Durrmeyer, I., 2021. Winners and losers: the distributional effects of the french feebate on the automobile market. Econ. J. ueab084.

European Commission, 2018. Air quality: Commission takes action to protect citizens from air pollution. European Commission - European Commission.

Isaksen, E., Johansen, B., 2020. Congestion pricing, air pollution, and individual-level behavioral responses.

Ivanova, D., Wood, R., 2020. The unequal distribution of household carbon footprints in Europe and its link to sustainability. Global Sustain. 3.

Ko, J., Park, D., Lim, H., Hwang, I.C., 2011. Who produces the most CO2 emissions for trips in the seoul metropolis area? Transp. Res. D Trans. Environ. 16 (5), 358–364.

Le Barbanchon, T., Rathelot, R., Roulet, A., 2021. Gender differences in job search: Trading off commute against wage. Q. J. Econ..

Le Monde, 2018. L'interdiction des véhicules les plus polluants étendue au grand Paris à partir de juillet. Le Monde.Fr.

Le Parisien, 2019. Grand Paris : moins d'une ville sur deux s'engage contre les véhicules polluants. Leparisien.Fr.

Lepeule, J., Laden, F., Dockery, D., Schwartz, J., 2012. Chronic exposure to fine particles and mortality: an extended follow-up of the harvard six cities study from 1974 to 2009. Environ. Health Perspect. 120 (7), 965–970.

Levinson, A., O'Brien, J., 2018. Environmental engel curves: Indirect emissions of common air pollutants. Rev. Econ. Stat. 101 (1), 121–133.

Linn, J., 2019. Interactions between climate and local air pollution policies: The case of European passenger cars. J. Assoc. Environ. Resour. Econ. 6 (4), 709–740.

Little, R.J.A., 1988. Missing-data adjustments in large surveys. J. Bus. Econom. Statist. 6 (3), 287–296, Publisher: [American Statistical Association, Taylor & Francis, Ltd.].

Liu, N., Ma, Z., Kang, J., 2017. A regional analysis of carbon intensities of electricity generation in China. Energy Econ. 67, 268–277.

MacDonald, H.I., 2016. Women's employment and commuting: Explaining the links:. J. Plan. Lit..

Mahony, T.O., 2013. Decomposition of Ireland's carbon emissions from 1990 to 2010: An extended kaya identity. Energy Policy 59, 573–581.

Nicolas, J.-P., David, D., 2009. Passenger transport and CO2 emissions: What does the french transport survey tell us? Air Pollut. Relat. to Transp., Atmos. Environ. Air Pollut. Relat. to Transp., 43 (5).1015–1020,

OECD, 2020. Non-exhaust Particulate Emissions from Road Transport: An Ignored Environmental Policy Challenge. Technical report.

Omnil-Ile de France Mobilites, 2019. La Nouvelle Enquête Globale Transport Présentation Des Premiers Resultats 2018. Technical Report.

Parry, I.W.H., Walls, M., Harrington, W., 2007. Automobile externalities and policies. J. Econ. Lit. 45 (2), 373–399.

Pottier, A., Combet, E., Cayla, J.-M., Lauretis, S.d., Nadaud, F., 2020. Qui Emet Du CO2? panorama Critique Des Inegalites Ecologiques En France. Technical Report 2020.02, FAERE - French Association of Environmental and Resource Economists.

Région Ile de France, 2016. Changeons D'air En Ile De France: plan Régional Pour La Qualité De L'air (2016–2021). Technical Report.

Royston, P., 2007. Multiple imputation of missing values: Further update of ice, with an emphasis on interval censoring. Stata J. 7 (4), 445–464, Publisher: SAGE Publications.

Sager, L., 2019. Income inequality and carbon consumption: Evidence from environmental engel curves. Eighth Atlantic Workshop on Energy and Environmental Economics, Energy Econ. Eighth Atlantic Workshop on Energy and Environmental Economics, 84.104507,

SDES, 2021. Résultats détaillés de l'enquête mobilité des personnes 2019. Données Et études Statistiques Pour Le Changement Climatique, L'énergie, L'Environnement, Le Logement, Et Les Transports.

Blaudin de Thé, C., Carantino, B., Lafourcade, M., 2021. The carbon 'carprint' of urbanization: New evidence from french cities. Reg. Sci. Urban Econ. 89, 103693.

Viegas, J.M., 2001. Making urban road pricing acceptable and effective: searching for quality and equity in urban mobility. Transp. Policy 8 (4), 289–294.

Wang, C., Chen, J., Zou, J., 2005. Decomposition of energy-related CO2 emission in China: 1957–2000. Energy 30 (1), 73–83.

WHO, 2014. 7 million premature deaths annually linked to air pollution. WHO.

WHO, 2018. Ambient (outdoor) air pollution.

Yang, Y., Wang, C., Liu, W., 2018. Urban daily travel carbon emissions accounting and mitigation potential analysis using surveyed individual data. J. Cleaner Prod. 192, 821–834.