

Machine Learning from Big GPS Data about the Heterogeneous Costs of Congestion

Nicolas Koch^{1,2,3}, Nolan Ritter^{1,2}, Alexander Rohlf^{1,2}, and Ben Thies^{1,4}

¹Mercator Research Institute on Global Commons and Climate Change (MCC), Berlin, Germany.

²Potsdam Institute for Climate Impact Research (PIK), Potsdam, Germany.

³IZA Institute of Labor Economics, Bonn, Germany.

⁴Humboldt University, Berlin, Germany.

July 31, 2023

Abstract

We exploit GPS-coded vehicle movement data that records millions of road users every second of their way over a full year in Berlin to suggest a novel approach to estimate the external costs of traffic congestion from revealed choices in a big data setting. To measure travel choices, we use unsupervised machine learning to assign anonymous commuting and non-commuting trips to individual drivers and track their repeated trip and route choices. We move beyond considering particular roads in isolation and quantify externalities created by actual trips on routes that combine the diverse road technologies of main roads and side streets. Causal identification for the route-level congestion elasticity relies on exogenous increases in traffic density from rerouting induced by traffic incidences that occur on adjacent roads off a given trip's route. Our findings suggest significant temporal heterogeneity in the marginal external costs of congestion between 3.0 and 87.4 € cents per vehicle kilometer during daytime. Hour-specific congestion taxes maximize large welfare gains, while uniform taxes are only 78% as effective.

1 Introduction

The congestion of road infrastructure is a major challenge in any urban agglomeration. As populations around the globe grow and urbanize, roads become increasingly congested and households both bear and impose external costs.¹ Understanding the extent to which resources used for travel could be better allocated by addressing the missing market problem of unpriced road use is therefore an important economic endeavour. In this paper, we exploit GPS-coded vehicle movement data that describes individual mobility in hitherto unavailable detail to address key empirical challenges in quantifying the optimal policy response to road congestion: (1) the externality is highly concentrated at certain times of the day while the few existing road use prices are uniform (Fosgerau and Van Dender, 2013) and (2) data conducive to the causal quantification of real-time costs of unpriced congestion at the city level has been severely limited in the past. Despite efforts to use Pigouvian taxes to mitigate congestion that date back more than 100 years (Pigou, 1920; Verhoef, 2010), our empirical understanding of the welfare that “second-best” congestion charge regimes forgo when insufficiently differentiating tolls remains limited.

We propose a novel approach to estimate *heterogeneous* social costs of congestion and the welfare gains from *differentiated* tolls using a machine-learned panel data set of GPS-coded trips across a representative sample of drivers in Berlin. Using time-stamped and geo-coded information on 34 million trips consisting of more than 970 million high-frequency way points gathered from navigation systems, connected cars equipped with GPS, and cell phones, we know the exact locations of millions of road users every second of their way and the speeds they are driving for each trip over the full year of 2017. We identify trips taken by the same driver using advances in machine learning that exploit the uniqueness of individual mobility patterns. At the same time, we can infer the level of traffic density along any route taken from the millions of trips we observe, be they on main roads or quiet side streets, for every recurring trip of any observed driver. Altogether, we simultaneously observe how the time costs of travel of many individuals’ repeated travel choices respond to changes in aggregate real-time traffic conditions based

¹ The cost of road congestion is not limited to the loss of time but also comprises air pollution and health damages (Currie and Walker, 2011; Simeonova et al., 2019), accidents (Green et al., 2016), and subjective well-being (Kahneman and Krueger, 2006; Anderson et al., 2016).

on a single, unified source. We therefore believe that this so far unused data provided by INRIX is an important step towards the ideal of credibly measuring the costs of congestion for an entire city and our approach is readily applicable to many other contexts.

Compare our novel approach to the traditional method of choice for analyzing road use and congestion: surveys (Keeler and Small, 1977; Small et al., 2005; Couture et al., 2018). Surveys collect comprehensive socio-economic information of randomly selected households. In addition, each household member puts down information on their trips. However, such self-reported travel diaries (Karlström and Franklin, 2009) or telephone surveys (Small et al., 2005) potentially suffer from self-reporting bias and, more importantly, they are limited to a few days of travel and a small set of travelers. Zhao et al. (2015) show that individuals have large day-to-day variability in their travel behavior that single day travel surveys cannot capture. This is an important limitation because much of the travel within cities is not related to workplace commuting but rather to consumption and leisure, such as trips to grocery stores, sports venues, and other professional service providers (Miyauchi et al., 2021). While these consumption related trips are driven regularly, they most likely do not occur on a daily basis. GPS movement data can reliably measure them. Even if surveys were able to collect information on many trips, the number of other road users for any given individual trip would remain unknown. Other more recent studies such as Bento et al. (2020), Yang et al. (2017) or Adler et al. (2018) or use data from road sensors. While this method is highly accurate for a limited set of main roads, we are interested in estimates of the cost of congestion at the city-level. Therefore, we measure traffic density for the complete road network, including minor roads which may potentially congest faster and whose technology are rarely studied.

One general caveat with automatically generated but anonymous floating car data is the absence of an identification variable that assigns each trip to an individual driver. This seemingly undermines researchers' ability to control for unobserved heterogeneity in individual travel behavior, which is crucial for unbiasedness. To reconcile the advantages of big, automatically generated data with the ability to control for unobserved heterogeneity, we propose a machine learning approach that assigns groups of recurring trips characterized by similar geographical points of origin and destination to a driver. To this end, we rely on density-based spatial clustering of applications with noise (DB-

SCAN). A key feature of DBSCAN is that it accommodates any spatial shape a cluster might take at its points of origin or destination. This is important because many urban households do not own a fixed parking space and instead have to instead rely on variable road side parking conditions. Therefore, a cluster’s shape is arbitrarily determined by the road network’s urban form and the geographical distribution of available parking spaces. Other clustering algorithms, such as the popular k -means clustering, critically hinge on spherically shaped clusters and also require a priori knowledge of the total number of clusters in the data.

We refer to the identified clusters of recurring trips with highly similar points of origin and destination (OD) as *driver-OD clusters*. Trips within a given cluster serve a common purpose, e.g. commute to work, to the grocery store, or to the gym. We then accommodate that drivers may choose among different routes for any given trip purpose and assign each observed trip a specific route defined by its individual, geo-temporal GPS waypoints. We can also assign home locations at the block-level for about a quarter of our *driver-OD clusters*, which allows us to add socio-economic information and to show that our data is a representative sample of drivers in the Berlin metropolitan area. Finally, we use the universe of the 970 million waypoints in our raw data to measure traffic density by counting the number of vehicles along the route taken at the time of each observed trip. We observe private but also commercial vehicles such as taxis, delivery vehicles, and trucks.

To calculate the social cost of congestion and the optimal congestion charge, we need to estimate the wedge between individuals’ actual costs and the marginal costs of road use. Following the literature (Akbar and Duranton, 2017; Yang et al., 2017; Couture et al., 2018), we recover this wedge by estimating what we call the congestion elasticity, which describes how the time cost of travel (“price”) varies with traffic density (“quantity”). This equilibrium relationship gives rise to endogeneity because traffic density and the time cost of travel are determined simultaneously. For identification, we propose an instrumental variable (IV) strategy that exploits that drivers optimize and react to unpredictable traffic incidences such as accidents, fallen trees, or bursts of water mains. Such incidences directly cause local reductions in road capacity. However, there are important second-order effects we can exploit for identification. Drivers upstream of the affected road seek to avoid the increased time costs by switching to alternative routes

that circumvent the incident. The increased demand on alternative routes induces more traffic there. This increase in traffic density is unrelated to the local road capacity, and thus provides exogenous variation in the “quantity” of travel.

We use rich crowdsourced data on more than 16,000 time and geo-coded traffic relevant incidents to construct an IV which captures incidences that occur off a given trip’s route but exogenously increase traffic density on that route because other drivers affected by the incidences reroute. We show that our instrument is a strong predictor of a trip’s traffic density, and as we document, valid after controlling for weather conditions and a rich set of fixed effects. These include day-of-week and hour-of-day fixed effects but also a high-dimensional driver-route fixed effect based on our DBSCAN clusters. They imply that the congestion elasticity is identified from variation in traffic density caused by the repeated occurrence of exogenous traffic incidents on a given route of repeatedly taken trips from the same driver undertaken for the same given purpose.

Our results indicate that the most commonly used measure of congestion elasticity, the average elasticity of travel time costs with respect to traffic density over an entire weekday, obscures significant intra-day heterogeneity. For instance, the elasticity during the time with the highest traffic density between 16:00 and 17:59 is 0.67 compared to the average elasticity of about 0.49. A binary distinction between peak and off-peak times neither approximates the temporal heterogeneity in the elasticity in the less nor the more congested times of the day. In contrast, the variation over the spatial dimension is negligible, possibly because urban centers are evenly distributed across Berlin.

We continue to show large temporal heterogeneity in the marginal external costs of congestion, ranging between 3.0 to 87.4 € cents per vehicle kilometer during daytime. The welfare-maximizing, socially-optimal policy response are time-specific congestion charges between 3.0 and 42.3 € cents for daytime trips across Berlin. For an average trip of 12 kilometers driven between 08:00 and 09:59, a driver would have to pay 4.48 €, which is about the price of a standard public transportation ticket in Berlin. We finally show that welfare gains crucially depend on the temporal differentiation. We estimate that time-specific road charges generate annual welfare gain of about 985 M€ of which the morning and evening rush hours account for 819 M€ or about 83.1%. A “second-best,” uniform congestion charge of 27.9 € cents per kilometer achieves only about 80% of the welfare gains of the temporally differentiated charges.

Our paper speaks to three strands of the literature. First, our investigation of the congestion elasticity is fundamental to transportation and urban economics. It is the key parameter in economic theory that is required to quantify externality costs and the design of optimal congestion pricing schemes (Vickrey, 1969). We provide the first estimates of the congestion elasticity at the “route-level,” which we define as the sequence of all connected road segments that fully describes a trip’s trajectory from its origin to its destination. In comparison, the prior literature either estimates segment-level “micro-elasticities” for specific roads or “macro-elasticities” at the level of the entire city. Micro-elasticities (Small, 2007; Geroliminis and Daganzo, 2008; Yang et al., 2017) are usually high and apply exclusively to individual roads, often big ring roads and highways. Because technology on such roads very likely differs from that of minor and medium roads in a city’s diversified road network, it is questionable whether micro-elasticities are well suited for estimating city-wide costs of congestion. In addition, micro elasticities may be biased because they do not accommodate for network spillovers. Especially when congestion is high, the level of traffic on adjacent roads may be an important confounder. Macro-elasticities (Akbar and Duranton, 2017; Couture et al., 2018), in contrast, are usually low. They are identified by variation in the number of vehicles across the entire city road network instead of those on the road segment actually taken. As such, they implicitly assume very large spillovers so that any driver affects any other across an entire city although their trip’s trajectories may never even remotely cross. Our route-level elasticity provides a middle ground between micro and macro elasticities that reflects real-world driving. It applies to actual routes that combine a mix of different road types. It accounts for equilibrium responses to traffic without imposing city-wide network spillovers. At the same time, we can control for unobserved heterogeneity in individual routes using our high-dimensional driver-route fixed effects created with DBSCAN and *QuickBundles*. We believe our route-level elasticity, which turns out to be much closer in magnitude to the high level of the micro than the low level of the macro elasticities, is the welfare relevant measure to estimate externalities created by actual trips taken across the city.

Second, our work contributes to a growing empirical literature on policies that seek to alleviate traffic congestion. One selection of studies focuses on real-world congestion pricing policies from London (Santos and Fraser, 2006; Green et al., 2016; Herzog, 2021;

Tang, 2021), Milan (Gibson and Carnovale, 2015; Cornago et al., 2019) and Stockholm (Karlström and Franklin, 2009; Simeonova et al., 2019), while another analyzes policies that restrict the number of vehicles on the road in Latin American and Asian cities (Davis, 2008; Gallego et al., 2013; Viard and Fu, 2015; Gu et al., 2017; Barahona et al., 2020). These papers quantify policy impacts on aggregate outcomes but do not address their welfare implications. We instead focus on estimating congestion elasticities of hitherto unavailable temporal resolution to run equilibrium policy simulations which allow us to assess the relative welfare impacts of alternative policy designs. In particular, we quantify the welfare that different policy designs forgo when insufficiently differentiating road charges by the hour of the day, which is the case with most real-world congestion pricing schemes. In this respect, our paper connects with the theoretical and empirical literature on the efficiency gains from *spatially* differentiated regulation regarding e.g. air pollution (Muller and Mendelsohn, 2009; Fowlie and Muller, 2019) or electric vehicles (Holland et al., 2016) or appliances (Jacobsen et al., 2020). We complement this literature with evidence for the efficiency gains from *temporal* differentiation in the context of road congestion.

Third, our paper contributes to a new and growing strand of literature that takes advantage of big data in urban economics. Several recent studies collect data on the “usual” traffic speed in developing country cities from Google Maps to study congestion impacts (Akbar and Duranton, 2017; Akbar et al., 2018; Hanna et al., 2017). In an experimental setting, Martin and Thornton (2017) and Kreindler (2018) provide drivers with smartphone GPS devices to monitor their travel activity and evaluate congestion policies. Another group of studies uses data from ride hailing platforms to study important traffic shocks (Mosquera, 2021) and the value of time (Buchholz et al., 2020). While Mangrum and Molnar (2020) use taxi trip records to measure traffic speed, Herzog (2023) exploits the ubiquity of ride hailing service Uber for the same purpose. Research using cellphone data highlights the importance of non-commuting trips in urban mobility (Miyauchi et al., 2021). We extend the literature by demonstrating how a combination of unsupervised machine learning and newly available, scalable floating car data allows us to *simultaneously* observe many individuals’ repeated commuting and non-commuting travel choices and aggregate real-time traffic conditions on every part of the road network over a long period of time based on a single, unified source that is representative for the

car-owning population. We believe this so far unused data² is an important step towards better understanding the consequences of individual mobility behavior in shaping city-level outcomes. Moreover, our IV strategy is near universally applicable. In contrast to prior papers that credibly address endogeneity issues by leveraging city-specific policy restrictions as natural experiments (Adler et al., 2018; Yang et al., 2017; Mangrum and Molnar, 2020), traffic incidences are a pervasive phenomenon across cities. For instance, Bento et al. (2021) rely on accidents as an exogenous shifter of traffic density. Our INRIX data allows for a broader set of incidences, including *inter alia* fallen trees or broken water mains, and this data is available for hundreds of cities in dozens of countries worldwide from INRIX.

This is the structure of our paper. The upcoming section describes the data underlying our analysis. Section 3 describes our empirical strategy, Section 4 tests our instrument while we present our estimation results in Section 5. Section 6 investigates the welfare effects including marginal external costs of congestion, optimal congestion charges, and policy-induced welfare gains. The final section concludes.

2 Data

2.1 Floating Car Data

Our analysis uses a rich data set provided by commercial mobility data provider INRIX which distinguishes itself by combining a diverse set of sources: 1) connected cars, 2) mobile devices, 3) satnavs, and 4) connected commercial fleets.³ Overall, a selection of data sources prevents sample selection. For example, the subsample of connected cars may comprise newer and more expensive vehicles, while satnavs most likely remain popular with older drivers.

The automatically collected data comprises the time and the location of each trip in terms of longitude and latitude of its end points and regularly updated and time-stamped waypoints collected along the way. We also know whether vehicles are private or

² To the best of our knowledge, Cunningham et al. (2021) is the only other study that uses aggregate data by INRIX to study the economic impacts from decaying water infrastructure.

³ Connected cars are fitted with wireless connections that permanently transmit high-frequency GPS coordinates. Mobile devices transmit high-frequency location information via apps or cellular networks. Historically, satnavs were the most widely used devices for GPS navigation but are now considerably less popular. Commercial fleets comprise taxis, delivery vehicles or trucks with on-board GPS receivers.

commercial, each trip distance, and each trip duration and speed. Our sample includes all trips recorded by INRIX that either originate, terminate, or pass through Berlin, Germany, for the period between January 1, 2017 and December 31, 2017. Altogether, our raw data comprises 34,208,105 unique trips with about 970 million corresponding waypoints. In comparison, the raw 2017 U.S. NHTS data holds 923,572 trips.

There are three reasons why automatically generated data is better suited for analyzing congestion than data collected by traditional survey methods. First, the automation leads to precise, geo-coded, and time-stamped observations and prevents well-known response biases. Second, the data is scalable over long periods of time and, hence, does not suffer from biases inherent to single day travel diaries. Third, we can directly infer traffic conditions at the level of the individual road segment by aggregating micro-level information across all individual trips from a single, unified data source. For a comprehensive comparison of our approach to traditional methods see Appendix A.1.

Figure 1 shows that a 2% random sample of 8,149 trips with 411,353 waypoints observed in the last week of June 2017 recovers the complete Berlin road network and suffices to differentiate major arteries from minor roads. Panel A of Table 1 provides descriptive statistics for the universe of all trips from commercial and non-commercial vehicles with at least one endpoint in Berlin. We use this raw data to infer traffic density.

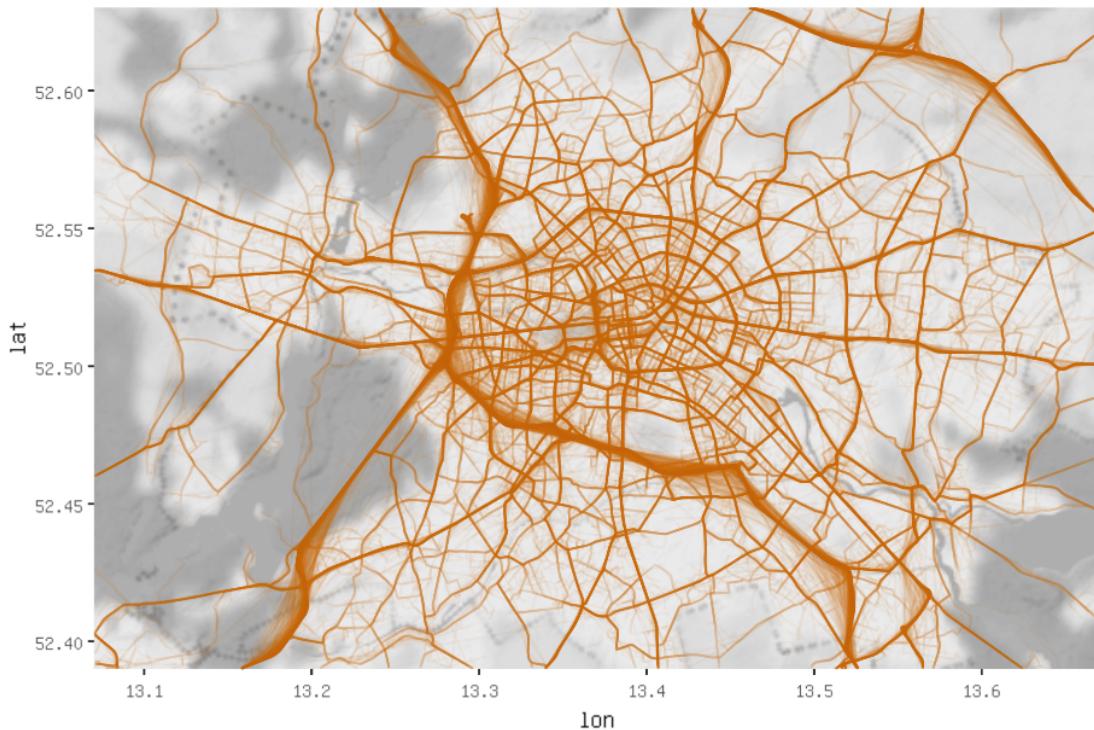
We refer the reader to Appendix A.2 for a description of all additional data used (including weather data provided by German Weather Service DWD, household travel information provided by the German National Household Travel Survey MiD, socio-economic status of zip-codes by Acxiom, and land-use data by EU Copernicus Program's Urban Atlas).

2.2 Sample Selection

2.2.1 Driver-OD Identification

One seemingly insurmountable problem with INRIX's automatically generated data is the absence of an identification variable that assign each trip to an individual driver without which unobserved heterogeneity cannot be accounted for. The latter is important given evidence for considerable differences in preferences and willingness-to-pay across individuals making travel decisions (Kreindler, 2018; Amador et al., 2005).

Figure 1: Two Weeks of Trip Data Reveal Berlin Road Network



Notes: This map shows a 2% random sample of 8,149 trips with 411,353 waypoints observed in the last week of June 2017. We connected the waypoints in each given trip with orange colored lines. This small sample suffices to highlight the Berlin network and differentiate major arteries from minor roads.

However, the inherent uniqueness of individual trips in terms of similar origins and destinations allows us to group together a driver's repeatedly driven trips. A large body of the data science literature shows the relative ease with which this can be achieved with respect to anonymous mobility patterns. This is because individual travel behavior is both highly *regular* in the sense that the same trip is taken many times and *unique* in the sense that individuals drive highly specific routes that are unlike those of others (De Montjoye et al., 2013; Rossi et al., 2015).

Following this literature, we use unsupervised machine learning to cluster recurring trips with highly similar points of origin and destination (OD). Based on these clusters of unique, repeatedly driven trips, we assign driver-OD identifiers to obtain a panel with an identifying variable. Trips within a given driver-OD cluster likely serve a common purpose.

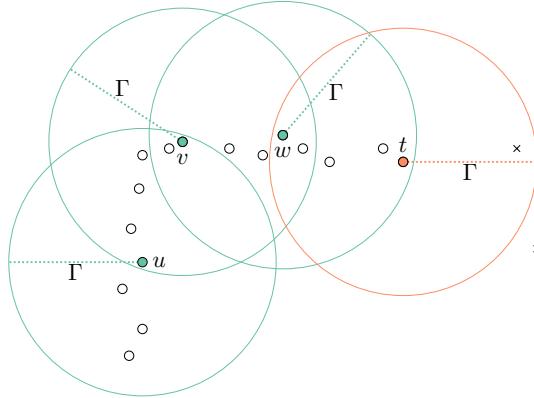
Table 1: Summary Statistics for the Floating Car Data

	Mean	Percentiles		
	Median	25th	75th	
Panel A: All trips in Raw Data				
$N = 34,208,105$				
Trip duration (min)	29.30	17.52	8.13	33.94
Trip distance (km)	26.18	7.25	2.56	19.07
Trip speed (kph)	34.34	26.45	18.14	41.35
Panel B: Non-commercial Trips in Berlin-Brandenburg				
$N = 26,023,170$				
Trip duration (min)	23.57	18.07	9.13	31.53
Trip distance (km)	15.42	7.57	2.96	17.69
Trip speed (kph)	32.69	26.69	18.63	39.90
Panel C: DBSCAN Clustered Trips in Berlin-Brandenburg				
$N = 1,712,138$, clusters = 36,449, routes = 81,692				
trips per cluster = 46.97, routes per cluster = 2.24				
Trip duration (min)	15.53	11.82	6.07	21.64
Trip distance (km)	6.91	3.35	1.49	8.73
Trip speed (kph)	23.88	21.24	15.40	29.02
Panel D: Traffic Density				
Number of concurrent route specific trips	601.28	254.69	85.44	823.00

Notes: Summary statistics for the floating car data provided by INRIX for Berlin for 2017. Panel A: Raw data based on a sample of all commercial and non-commercial trips with at least one waypoint in Berlin recorded by INRIX. Panel B: Sub-sample of non-commercial trips over half a kilometer length within Berlin or Brandenburg. Panel C: Sub-sample of repeatedly driven non-commercial trips characterized by highly similar geographical points of departure and destination clustered by DBSCAN on weekdays. For details see Section 2.2.2. Panel D: Traffic density for all trips in Panel C calculated from all trips from Panel A. For details consult Section 2.2.3

Unsupervised Machine Learning Algorithm DBSCAN The partitioning of an unlabeled sample into a number of similar subsamples, such as repeatedly driven trips, is a clustering problem for which a selection of unsupervised learning algorithms exist (Athey and Imbens, 2019). We use *Density-Based Spatial Clustering of Applications with Noise* (DBSCAN), an algorithm that is specifically designed for spatial data (Ester et al., 1996; Wang et al., 2018). We motivate our choice in Section Appendix B.2.

Figure 2: The Concept of Density-Reachability



Notes: This diagram illustrates the concept of density-reachability which is the cornerstone of DBSCAN. The hollow points exemplify the parking spots of an hypothetical driver who either parks her car on a road that runs north to south or an adjacent road that runs east to west. All points within a radius Γ of a given point are considered directly density-reachable. For $\kappa \leq 9$, u , v and w are core points because there are at least 9 points in each Γ environment. This applies to other unnamed hollow points as well. Points u and w are mutually *density reachable* because there is a chain of directly density reachable core points between them. The diagram also shows a border point t that is characterized by an insufficient number of points within its Γ environment to be considered a core point. Instead, it is only *density connected*. The crosses indicate outliers not directly density-reachable from any core-point. Both core and border points form an identified cluster.

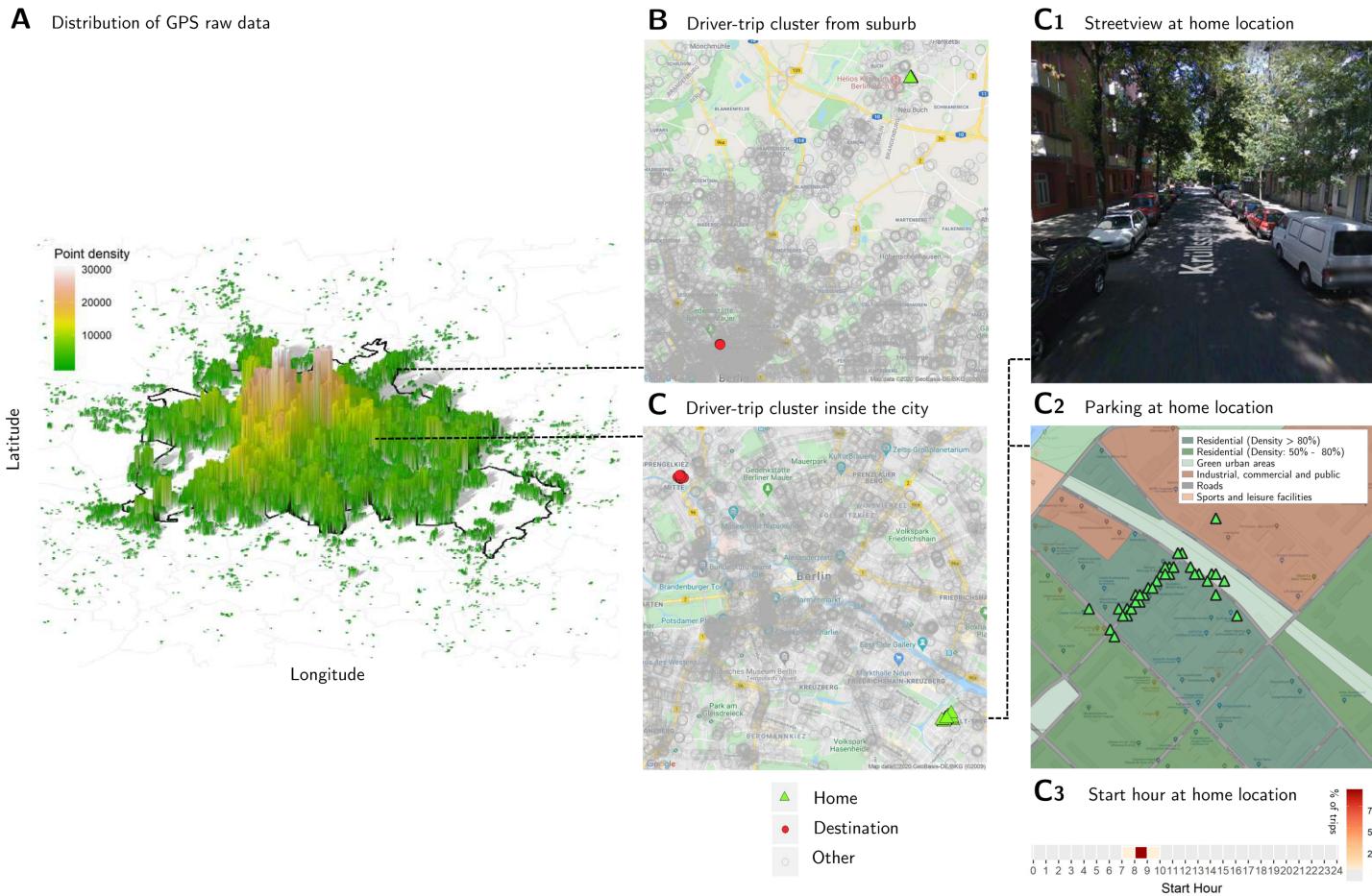
Given a set of potentially multidimensional points, DBSCAN groups together those points that are close to each other based on two user-specified parameters: A distance measure Γ and a minimum number of points κ within this Γ environment necessary for identifying a core point. A core point therefore has at least $|N_\Gamma| \geq \kappa$ points within a distance of Γ , including the point itself. The hollow points in Figure 2 exemplify the parking spots of an hypothetical driver who either parks her car on a road that runs north to south or an adjoining road that runs east to west. For $\kappa \leq 9$, u , v and w are core points because there are at least 9 points in each Γ environment. v is *directly density reachable* from u because the distance is less than Γ . u and w are *density reachable* because there is a chain of directly density reachable core points between them. t is a *density connected* border point because it has fewer than κ points in its Γ environment. Points other than core or border points are considered outliers that are not associated with any cluster and indicated with a cross. We provide exact definitions for the concept of density reachability in Appendix B.1.

Practical Application of DBSCAN Before running DBSCAN, we impose three sample restrictions on our raw data from INRIX (Panel A of Table 1). First, we remove trips shorter than half a kilometer to reduce measurement error. Second, we limit our sample to non-commercial trips which is in line with Couture et al. (2018).⁴ Third, given our focus on urban congestion, we limit trips to those within the metropolitan Berlin and Brandenburg area. Panel B of Table 1 provides summary statistics for these trips. We use the R package *dbSCAN* (Hahsler et al., 2018) to run DBSCAN on the coordinates of the endpoints of all trips in this sample of 26,023,170 non-commercial trips. Panel A of Figure 3 highlights these endpoints in the sample.

For our main analysis, we set $\Gamma = 100$ meters for the combined Euclidian distance between both trip start and trip end points as well as $\kappa = 20$, which guarantees that any cluster of repeatedly driven trips holds at least 20 trips. DBSCAN identifies 36,449 clusters comprising a total of 1,712,138 trips with an average of 47 trips per cluster. Panel C of Table 1 presents descriptive statistics for the clustered data. We refer to this data set as *clustered data*. With more than 1.7 million individual trips, our data set remains many orders of magnitude larger than those found in traditional travel surveys. It includes both commuting trips but also any regularly driven non-commuting trip, such as trips to grocery stores or sport venues. Panel B and C of Figure 3 visualize two clusters of similar trips embedded in an ocean of endpoints indicated in gray. We assign trips deemed similar in terms of their origin and destination (OD) an identification variable which we refer to as *driver-OD cluster*. We assume that trips within a given cluster serve a common purpose, e.g. commute to work, to the grocery store, to the gym, etc.). In subsequent robustness checks, we show that our results are robust to our choices of Γ and κ .

⁴ The marginal external costs of congestion for taxis and commercial trucks are arguably higher than those for private cars. Accommodating this heterogeneity in our welfare analysis is beyond the scope of this paper.

Figure 3: Visualization of Driver-OD Identification with DBSCAN



Notes: This figure visualizes the identification of driver-OD clusters from the universe of raw trip data via DBSCAN. Panel A shows the number of trip departures from a random 10% sample of all non-commercial trips (Panel B of Table 1). The x -axis indicates the longitude, the y -axis the latitude, and the z axis indicates the number of trip departures. We anonymized all the endpoints we show by randomly shifting them along their x and y dimensions. The overall interpretation of the data, however, remains unchanged. The thick black line indicates the city border of Berlin. Black dashed lines link two select areas displayed in Panel B and C to Panel A. From an ocean of trips (20% random sample) indicated in gray, DBSCAN identifies trips from the same driver. With additional information from the EU Copernicus Program's Urban Atlas, we are able to assign home locations marked with green triangles. All endpoints have again been randomly shifted along the x and y dimensions for anonymization. Panels C1 through C3 provide additional information on the home location from Panel C. Panel C1 shows a picture from the same neighborhood. Panel C2 shows the departure points of the home location which indicate that the driver relies on road side parking. In the highlighted area, there are two main roads running from north west to south east that form square blocks with parallel roads running in between at regular intervals. For the purpose of anonymization, we shifted the points of departure from their original block to a random one in the vicinity. This has neither impact on the general interpretation nor on the tightness of fit between observed parking positions and actual road side parking spaces. Finally, Panel C3 clarifies that trips associated with this home most often start between 08:00 and 09:00.

Some of the identified clusters of unique, repeatedly driven trips start from individuals' homes. Therefore, in a second step, we assign "home" locations to endpoints if these fulfill two conditions. First, in driver-OD clusters with only one-way trips, an endpoint with at least one trip occurring between 05:00 and 10:00 is considered a potential home location. In driver-OD clusters that include trips in both directions, we consider the endpoint with the higher share of starts between 05:00 and 10:00 a potential home location. Second, we consider a potential home location in a residential area as a true home location. To this end, we rely on block-level land-use data from the Urban Atlas provided by the EU Copernicus Land Monitoring Service.⁵ Altogether, we identify home locations for 27,966 of the 36,449 driver-OD clusters. Note that home locations are at the block-level and do not allow us to pinpoint any individual home addresses.

Panel B and C of Figure 3 illustrate identified home locations and destinations for two clusters. The cluster highlighted in Panel B consists of 202 repeated trips with points of departure identifying a home location in the suburbs of Berlin. Narrowly clustered GPS coordinates may indicate the availability of private parking at either endpoint. In contrast, the 105 repeated trips clustered in Panel C start in a dense urban home location area, where on-street parking is scarce as indicated by the picture in Panel C_1 . Therefore, the clustered endpoints are more widely distributed across neighboring streets. Panel C_2 demonstrates the ability of DBSCAN to account for the urban form and the geographical distribution of parking spaces. It also illustrates the Urban Atlas land-use layers that we use together with departure times (Panel C_3) to assign true home locations in residential areas.

2.2.2 Identification of Routes

So far, we used DBSCAN to identify repeatedly driven trips for a given purpose defined by similar points of origin and destination. Next, we must accommodate that drivers may choose among different routes for any given trip purpose. A route is defined by its individual, geo-temporal GPS waypoints. However, when GPS coordinates are either

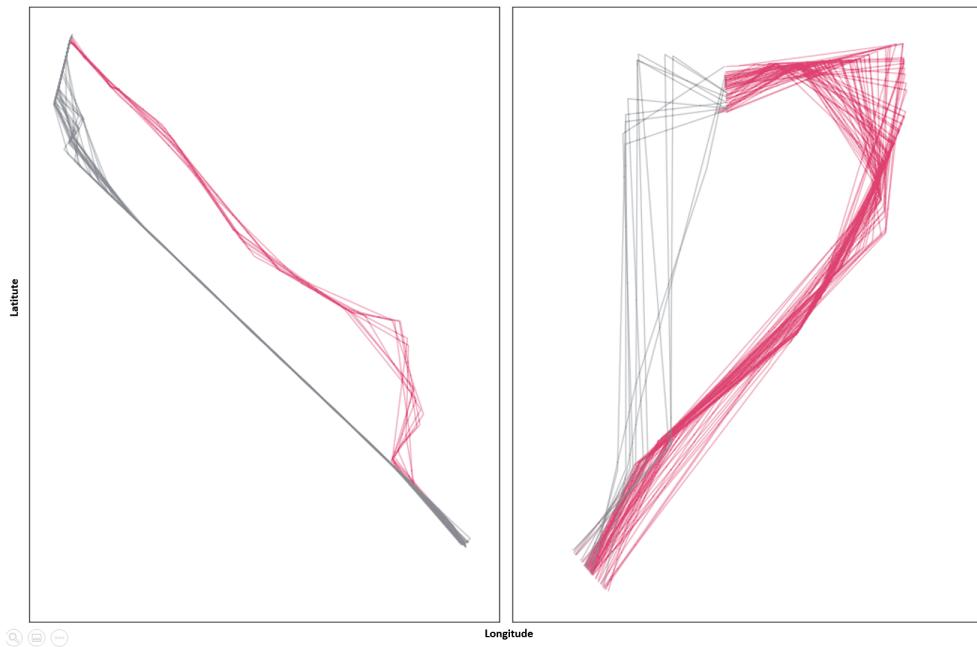
⁵ The Urban Atlas provides high-resolution land use maps for urban areas for the reference year 2012. The maps are derived from earth observation satellite images backed by other reference data, such as commercial off-the-shelf or open source navigation data, such as Open Street Map, and topographic maps. It can be downloaded at <http://land.copernicus.eu/local/urban-atlas/urban-atlas-2012/view>

imprecise or only recorded at low frequency, then identical routes may not be easily identified.

We use the QuickBundles algorithm (Garyfallidis et al., 2012) to address this issue when identifying distinct routes within driver-OD clusters. The algorithm first linearly connects all waypoints for each trip in a given driver-OD cluster in chronological order to form trajectories. In a second step, it divides these linearly interpolated trajectories for tractability into a specified number of equidistant points. Third, the algorithm compares these equidistant points to identify similar trajectories and assign them to bundles (see Figure 15 and the discussion in Appendix Appendix E). We refer to these bundles as routes.

Figure 4 presents results of the QuickBundle algorithm for two select driver-OD clusters from our dataset. Here, QuickBundles allocates the trips in each cluster to two routes, one indicated in grey, the other in red. On average, we identify 2.24 distinct routes across our 36,449 driver-OD clusters. We assign each observed trip in our *clustered data* a cluster-specific route fixed effect.

Figure 4: Allocating Trips to Routes with QuickBundles



Notes: This figure illustrates how QuickBundle splits the trips observed in two select driver-OD clusters into two distinct routes, one indicated in grey, the other in red. Each trip is visualized using all its waypoints. The left panel shows the trips of a driver-OD cluster characterized by a high degree of precision in the GPS coordinates. Therefore, the trips in each identified route share a highly similar shape. The trips in the right panel are recorded at a lower degree of precision which manifests in a higher level of heterogeneity in the shape of the trips assigned to a route.

2.2.3 Quantification of Traffic Density

We use the universe of the 970 million waypoints in our raw data to approximate traffic density along the route taken for each observed trip in our *clustered data*. To this end, we first place a convex hull around each route. This hull is shown as a red shaded area in Panel A of Figure 5 for the example of one route of a select driver-OD cluster. Second, we approximate traffic density for every trip in our *clustered data* by counting the number of unique commercial and non-commercial trips in our raw data that intersect the convex hull at the time of an individual trip (see Panel B of Figure 5).⁶ We focus on the hour which covers most of the trip. Our vehicle count is equivalent to a density because our route fixed effects control for the length of a route. Summary statistics for our density measure are provided in Panel D of Table 1.

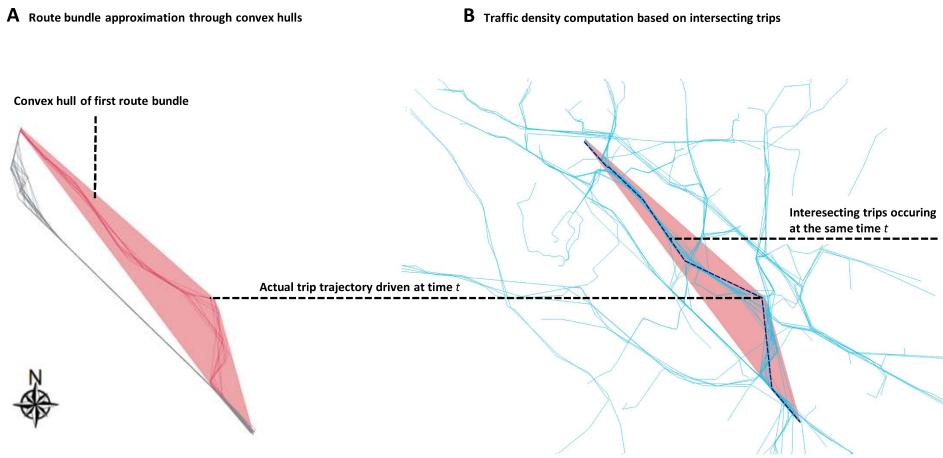
Our approach has the advantage that we can infer traffic for the complete road network, including minor roads which may potentially congest faster and whose technology are rarely studied. Most prior studies rely on road sensor data that provide accurate measurements for a limited set of main roads (Adler et al., 2018; Yang et al., 2017; Bento et al., 2020). We instead rely on measurements for routes, i.e. the connected sequence of all road segments that fully describes a trip from its origin to its destination. By measuring traffic density within the convex hull of a specific route taken, we account for potential network spillovers from nearby traffic on adjacent roads along the route.

2.3 Representativeness

Although it is a distinguishing feature of INRIX that it relies on a diverse set of sources, we test whether our *clustered data* is a representative sample of drivers in the Berlin metropolitan area. Because our data is devoid of socio-economic information at the level of the individual, we must restrict the analysis to our clusters with assigned home locations and compare the distribution of home clusters across geographical areas, such

⁶ The sheer number of waypoints in our data makes counting all intersections between the actual waypoint trajectories of a given trip with all waypoint trajectories of all others an intractable computational endeavor.

Figure 5: Computation of Traffic density



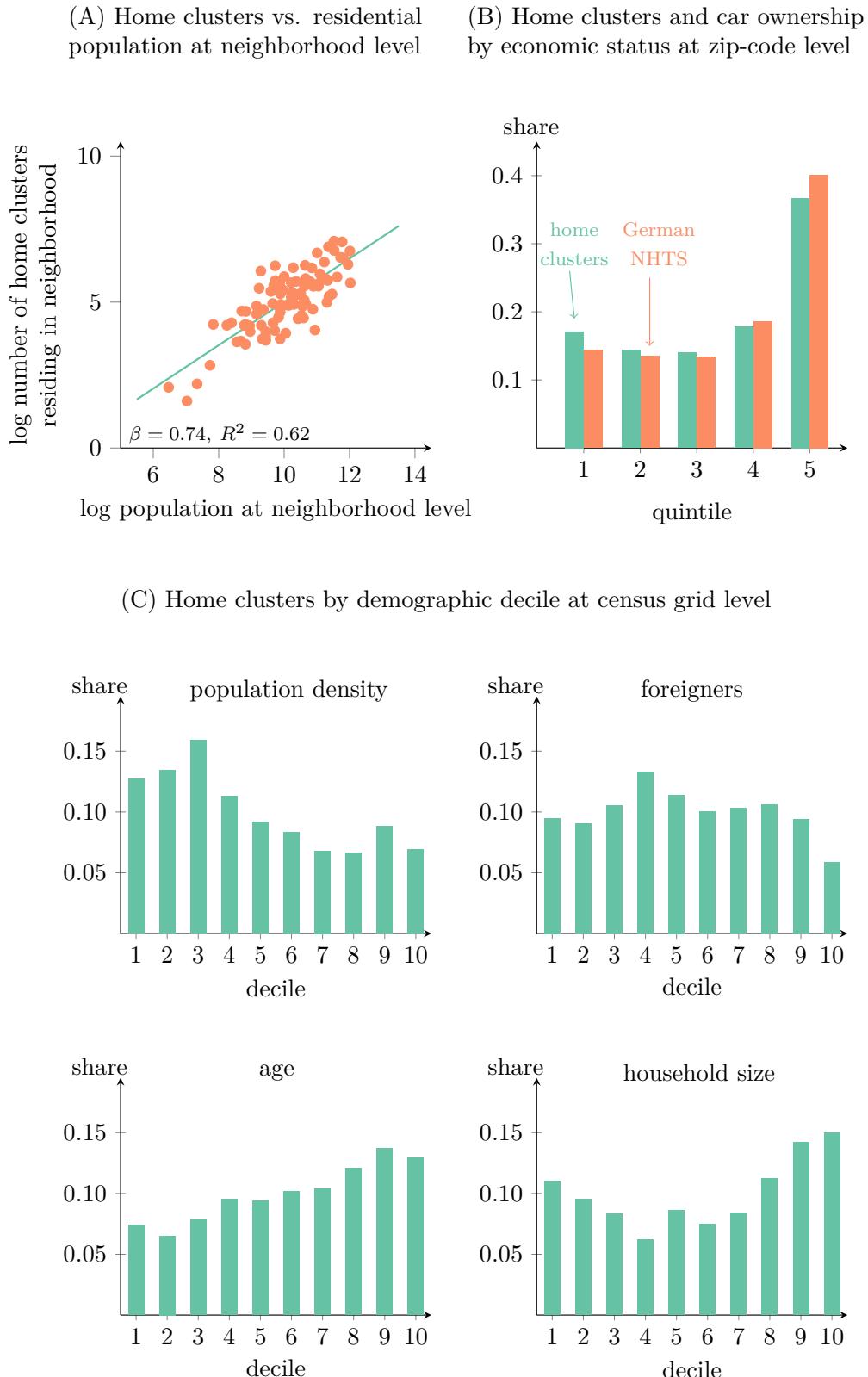
Notes: This figure illustrates how we calculate traffic density. Panel A shows the convex hull around one route as a red triangle of a select driver-OD cluster (shown already in Figure 4). Panel B demonstrates the procedure of computing traffic density for a trip driven at time t (black dashed line) by intersecting its convex hull with the universe of commercial and non-commercial trips at time t (light blue lines).

as residential census grid cells, neighborhoods, or ZIP codes for which representative official data on population characteristics exists.

Figure 6(A) shows a high correlation between the number of home clusters in the 96 administrative neighborhoods of Berlin and the local population.⁷ However, there is no one-to-one relationship because our home clusters consist of households who own at least one vehicle, while the car ownership rate in the Berlin metropolitan area is moderate. In Figure 6(B), we demonstrate that our sample of home clusters closely follows the distribution of car-owning households in Berlin across socio-economic status groups. The green bars show the shares of home clusters by socio-economic status group, while the orange bars demonstrate the distribution of car ownership across status groups according to the German NHTS for Berlin (MiD, 2018). The status quintiles are calculated at the zip-code level using data by Acxiom. Differences in distribution between our clustered sample and car ownership patterns in Berlin are negligible.

⁷ <https://daten.berlin.de/datasets/einwohnerinnen-und-einwohner-den-ortsteilen-berlins-am-31122018>.

Figure 6: Representativeness



Notes: This figure demonstrates the representativeness of our clustered dataset. Panel (A) shows the correlation between the log number of home clusters and of the residential population in the neighborhood (“Ortsteil”) of residence. β and R^2 are from a bivariate linear regression. Panel (B) compares the share of home clusters distributed across socio-economic status quintiles at the ZIP code level based on data by Acxiom to the share of car-owning households distributed across economic status groups as indicated by the German NHTS for Berlin (MiD, 2018). Panel (C) consists of four diagrams that each show the distribution of home clusters across population deciles of four socio-economic variables: Population density, share of foreigners, age, and household size. All of these variables are provided by the German 2011 census and are available at the level of the $1km^2$ census grid. A uniform distribution of home cluster shares across the deciles of population weighted grid cells indicates a high level of representativeness with respect to the given characteristic.

Finally, we scrutinize socio-demographic information at the 1km^2 census grid level⁸. The Figures in Panel (C) show the shares of home clusters for four variables across census grid deciles from 1,498 grid cells in the Berlin metropolitan area. A uniform distribution across deciles indicates equal representation. Overall, we find only slight over-representation within our *custered data* regarding individuals from low density areas and areas with large households as well as under-representation in areas with a high share of foreigners. Given that we compare our sample of vehicle owning households to the overall population, discrepancies occur where expected.

3 Empirical Strategy

The speed-density relationship is the centerpiece in the analysis of congestion that describes how the time cost of travel given by the inverse of speed (“price”) changes with additional vehicles per road kilometer (“quantity”) (Akbar and Duranton, 2017; Couture et al., 2018; Yang et al., 2017). It allows us to derive the negative congestion externality, which results from drivers optimizing their travel behavior according to prevailing traffic conditions while disregarding the time costs they impose on others (see Section 6.1 for a formal economic framework).

The empirical challenge is that the observed speed-density relationship is an equilibrium outcome of travel demand and supply. For causal identification, we need to carefully address the endogeneity from the simultaneous determination of the number of vehicles on the road and the time cost of travel.

3.1 OLS setup

We specify the inverse speed-density relationship as follows:

$$\log C_{ijrt} = \sigma \cdot \log D_{ijrt} + \beta X_t' + \gamma_{jr} + \psi_w + \omega_h + \epsilon_{ijrt}, \quad (1)$$

where C_{ijrt} is the time cost of travel for trip i by driver j on route r at time t . Drivers j are identified by the driver-OD clusters DBSCAN identified. Quickbundles identifies routes r within driver-OD clusters. Time t is given at the hour level. D_{ijrt} measures traffic density given by the number of commercial and non-commercial vehicles along the

⁸ <https://www.zensus2011.de/DE/Home/Aktuelles/DemografischeGrunddaten.html>.

route taken at the time of an individual trip.⁹ Matrix X_t holds indicators for holidays and a host of weather characteristics such air temperature, precipitation, sunshine duration, wind speed, and visibility. γ_{jr} is a driver-route fixed effect. ψ_w and ω_h are day-of-week fixed effects and hour-of-day fixed effects, respectively. ϵ_{ijrt} is an error term. Standard errors are clustered at the driver-route level. σ is our coefficient of interest that measures what we call the congestion elasticity, i.e. the elasticity of travel time with respect to traffic density.

Our main concern in estimating the congestion elasticity is endogeneity from shocks ϵ_{ijrt} that simultaneously affect the time cost of travel and the number of travelers on the road. Our specification in equation (1) accounts for various sources of endogeneity. First, the inclusion of weather controls at the time a trip is taken avoids bias from weather shocks that affect driving in a city, e.g. fog or heavy rain are likely to lead to high time costs of travel and fewer number of travelers, resulting in downward bias. Second, our time fixed effects account for unobserved hour-of-day and day-of-week specific phenomena. For instance, hour-of-the-day fixed effects control for variation in traffic light settings, which are likely correlated with time costs and traffic density. Day-of-the-week fixed effects control for the possibility that roadworks and other events such as waste removal or street-cleaning tend to occur only on certain days, which may affect both time costs and traffic density. Finally, our driver-route fixed effects imply that σ is identified from variation in D_{ijrt} for a given route of a repeatedly taken trip by the same driver for a given purpose. This accounts for unobserved characteristics of drivers (i.e. age, gender, or driving style), trips (i.e. purpose), as well as routes (i.e. distance and infrastructure quality). These characteristics may be correlated with time costs and traffic density. For instance, prevailing infrastructure quality may render certain routes inherently slower or drivers with a preference for high travel speed may select into residential locations where travel costs are low.

3.2 Instrumental Variable Approach

While our OLS specification in (1) accounts for important sources of endogeneity, the concern remains that yet uncaptured shocks to the time costs may affect traffic density

⁹ The vehicle count is equivalent to a density because we control for road kilometers by including driver-route fixed effect. Using vehicles per km² yields identical results.

because drivers avoid traveling when they are aware of them. In addition, shocks to time costs *mechanically* affect traffic density precisely because drivers reach their destinations either earlier or later. To break these correlations, we rely on a demand-shifting instrument.

Our instrumental variables (IV) approach exploits that drivers optimize and react to unpredictable traffic incidences, such as accidents or fallen trees. All else equal, the first-order effect of such incidences is a local reduction in road capacity, which increases the time cost of travel on the affected road. Second-order effects occur because drivers upstream of the affected road seek to avoid these costs by switching to alternative routes that circumvent the incidence area. The upward shift in the demand on alternative routes induces more traffic there. This increase in traffic density is unrelated to the local road capacity, and thus provides exogenous variation in the “quantity” of travel. Therefore, our instrument are traffic relevant incidences that occur off a given trip’s route, but exogenously increase traffic density there because other drivers affected by the incidences reroute.¹⁰

When instrumenting, the first stage explains the log of traffic density $\log D_{ijrt}$ by

$$\log D_{ijrt} = \alpha \cdot \log Incidents_{ij-rt} + \beta X'_t + \gamma_{jr} + \psi_w + \omega_h + \nu_{ijrt} \quad (2)$$

where $Incidents_{ij-rt}$ is the number of traffic relevant incidents at time t of driver j ’s trip i in the vicinity of but not on route r , i.e. $-r$ denotes the set of roads in the vicinity. Otherwise, the set of control variables is identical to equation (1). We replace $\log D_{ijrt}$ in equation (1) with the predicted traffic density \hat{D}_{ijrt} from equation (2) to estimate our second stage.

IV Construction The construction of our IV is based on 16,514 time and geo-coded traffic relevant incidents in Berlin in 2017 collected by INRIX. We restrict ourselves to 6,680 unpredictable incidents (see Figure 14), defined as accidents, water main breaks, fallen trees, or stalled vehicles. It is unlikely that drivers could have predicted the spatial and temporal occurrences of these events. Using INRIX data has two advantages. First,

¹⁰Note that our IV approach also allows us to eliminate simultaneity issues related to hypercongestion (Verhoef, 2003; Fosgerau and Small, 2013; Anderson and Davis, 2020) because we only use variation of vehicle density that is caused by conditions on alternative routes, which is unrelated to bottleneck conditions on the route of trips studied.

it is crowdsourced data that draws on various twitter feeds and road user reports in addition to conventional police and road agency records. This reduces the risk of under-reporting, i.e missing incidents that actually happen, which might vary by traffic density. Second, INRIX uses their real-time traffic flow data to filter out all incidences that do no alter traffic conditions. A fallen tree, for instance, is only considered when it blocks a lane. Altogether, this increases the relevance of our IV.

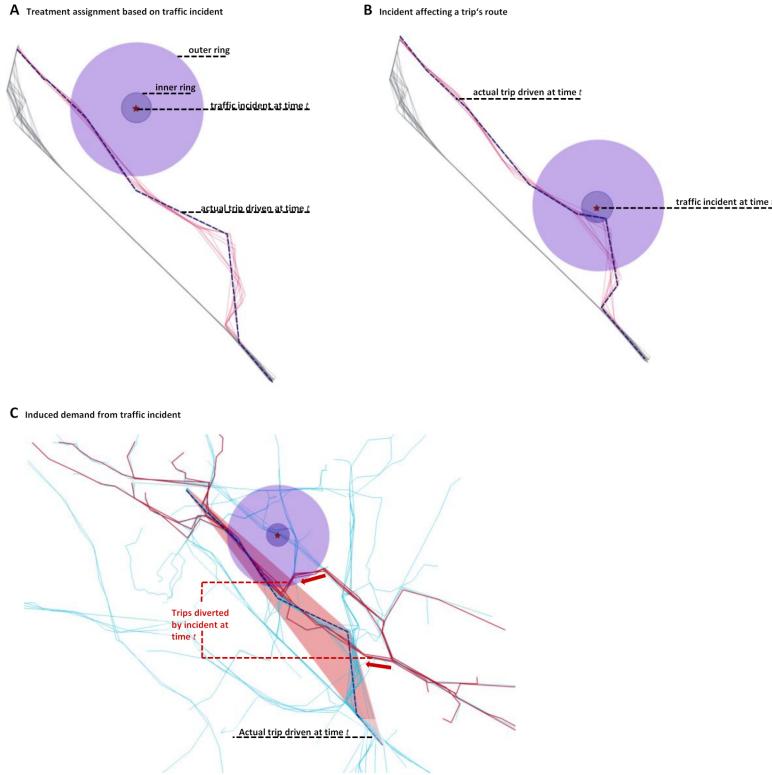
To assign an incident to a given trip, we draw two concentric circles around the coordinate of an incident (see Figure 7). The outer circle has a radius of 1,000 meters whereas the inner circle has a radius of 50 meters. We consider all trips that intersect the area between the inner and the outer circle as subject to potential rerouting caused by an incident on roads in the vicinity (Panel A). If a route, however, intersects with the inner circle (Panel B), we consider that the incident may directly reduce the road capacity on the route taken. In this case, the increase in traffic density is endogenous. Therefore, we discard trips intersecting with the inner circle. Our robustness checks reveal that the choice of the two circles' radii has no effect on our findings.

In Panel C, we add to trip i driven at time t from Panel A a selection of concurrent trips from other drivers in our data set in light blue color. The trips indicated in red are from other drivers that change to another route to circumnavigate the incidence. The red arrows indicate that other drivers choose to switch to i 's route and thereby increases trip i 's traffic density.¹¹ We exploit this exogenous increase in traffic density caused by rerouted traffic from nearby incidences to estimate the congestion elasticity.

Trips in our clustered data are frequently treated. 21,995 of all 36,449 driver-OD clusters (60.3%) are subject to at least one treatment in terms of traffic incidents on other roads in the vicinity. For the 81,692 distinct routes within these driver-OD clusters, we observe on average about two treatments per route. If we focus only on those routes with at least one treatment, this number increases to five. The repeated occurrence of exogenous incidents on the same route of a given driver-OD cluster enables our IV strategy.

¹¹Recall that, for computational tractability, we measure traffic density by counting the number of concurrent trips that intersect the red shaded convex hull of driver i 's route.

Figure 7: Constructing the Instrumental Variable



Notes: This figure illustrates how we build our instrument that identifies exogenous variation in traffic density. In Panel A, we show how we assign an incident to a given trip. We draw two concentric circles around the coordinate of an incident. The outer circle has a radius of 1,000 meters whereas the inner circle has a radius of 50 meters. We consider all trips that intersect the area between the inner and the outer circle as subject to potential rerouting caused by an incident in the vicinity. In Panel B, we show a trip whose route intersects with the inner circle. We discard this trip because the incident in this case may directly reduce the road capacity on the route taken, and the increase in traffic density would be endogenous. In Panel C, we add to the given trip from Panel A a selection of concurrent trips from other drivers in our data set in light blue color. The trips indicated in red are from other drivers that change to another route to circumnavigate the incidence. The red arrows indicate that other drivers choose to switch to i 's route and thereby increases trip i 's traffic density. Our IV strategy exploits this exogenous increase in traffic density measured by counting the number of concurrent trips that intersect the red shaded convex hull.

4 Assessing the Instrument

4.1 Relevance

We use traffic relevant, quasi-random incidences such as accidents as an instrumental variable for traffic density. Our instrument is relevant if and only if these incidences cause increases in traffic density for treated trips. Table 2 shows estimates for our first

stage. In Models 1 and 2, we regress the log of traffic density for the individual trip on the log of the number of traffic incidences in the trip’s vicinity. In Models 3 and 4, we instead use a binary incidence indicator that takes on a value of 1 in case that there is an incidence in the vicinity of a given trip. We either include the same controls as in our main specification (Models 1 and 3) or include additional control in the form of quadratic and cubic weather controls or controls for the daylight share or the season (Models 2 and 4). The first-stage coefficients are statistically significant at any conventional level of significance across all models. Models 3 and 4 tell us that if there is at least one incident along a trip, traffic density increases by 2.5% or by about 16 vehicles on average. While the interpretation of the binary incident measure is more straightforward, we prefer using the log of traffic incidences because it better captures the intensity of incidences. In this case, Model 1 suggests that a 10% increase in incidences raises traffic density by about 0.35% or by about 16 vehicles as in case of the binary model. Kleibergen-Paap F -statistics exceed at least 1,819 and provide evidence for a strong instrument. Notably, the fist stage F passes the threshold of 104.7 recently suggested by Lee et al. (2022) for a conventional t -test to have the correct size.

4.2 Validity

Conditional independence and exclusion restriction First, our instrument is valid if there are no unmeasured factors that affect a given trip’s time costs and simultaneously increase the number of traffic incidences in that trip’s vicinity. This conditional independence could be violated if there are unobserved weather shocks such as rainstorms that affect both traffic speed and traffic incidences or if we do not control for the fact that some routes and times are particularly prone to the occurrence of incidents and high time costs. Therefore, we condition on weather controls, route fixed effects but also holiday and weekday- and hour-specific phenomena to block potential violations of the conditional independence.

Second, for a causal interpretation to our estimates, our instrument must exclusively affect the outcome via the treatment. In our case, incidences in a given trip’s vicinity but not on its route can only affect its time cost of travel by increasing traffic density via rerouting. We condition on route and time fixed effects as well as weather controls to control for potential violations of this exclusion restriction.

Table 2: First stage

model	Dep. variable: log(traffic density)			
	1	2	3	4
log(incidents)	0.035*** (0.001)	0.034*** (0.001)		
binary incidents			0.025*** (0.001)	0.025*** (0.001)
baseline controls	yes	yes	yes	yes
quadratic & cubic weather	no	yes	no	yes
daylight	no	yes	no	yes
season	no	yes	no	yes
Kleibergen-Paap <i>F</i> -statistic	2,003	1,997	1,915	1,839

Notes: This table presents results for the first stage on workdays between 04:00 and 21:59 CET following equation (1). The dependent variable is the logarithm of traffic density. Traffic density is the number of all commercial and non-commercial trips that intersect with a given trip's route. All specifications include fixed effects for the hour-of-day, the day-of-week, the route, and for public holidays. The baseline controls include air temperature, total precipitation, wind speed, sunshine duration, and visibility. In Models 2 and 4, we also include quadratic and cubic weather weather controls, the share of daylight, and the season. The ultimate row shows the Kleibergen-Paap *F*-statistic for weak instruments. The sample size for all regressions is 1,712,138. Standard errors are clustered at the level of the route. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% level, respectively.

A testable implication of these assumptions is that a larger set of control variables should not significantly change the first stage estimates, as they should be uncorrelated with the instrument. Columns 2 and 4 of Table 2 confirm stable first stage coefficients if we add more controls. In a similar spirit, Cinelli and Hazlett (2022) suggest to evaluate the validity of the conditional independence and exclusion restriction by investigating how strongly an unmeasured confounder must predict residual variation in both the instrument and the outcome in order to reduce (i) the estimated effect of the instrument on the treatment (first stage) and (ii) the estimated effect of the instrument on the outcome (reduced form). The test results reveal that omitted confounders between the instrument and the outcome or omitted side-effects of the instrument via paths other than the treatment need to be significantly stronger than any of our included control variables in order to explain away the first stage and reduced form estimates (see Tables 7 and 8 in the Appendix). This is suggestive evidence in favor of instrument validity.

4.3 Monotonicity

If the causal effect of traffic density is heterogeneous, we must also assume that our instrument has a monotonic effect. This implies that traffic incidents may either leave unchanged or increase a given trip’s traffic density but never decrease it. One testable implication of the monotonicity assumption is that the first stage estimates must be non-negative for any subsample (Bhuller et al., 2020). Therefore, we estimate the first stage on a total of 72 strata defined along the categories of our control variables and, additionally, along the categories of the socio-economic controls that we used to demonstrate representativeness (see Section 2.3). For all these subsamples, the first stage estimates are positive and consistent with the monotonicity assumption (Table 9 in the Appendix).

5 Estimation Results

5.1 Average Elasticity of Congestion

Table 3 presents our estimates of the congestion elasticity, i.e. the elasticity of travel time costs with respect to traffic density, following equation (1). The dependent variable is the log of the time costs of travel. Models 1 and 2 in Table 3 presents coefficients from OLS while Models 3 and 4 are for the IV regression. Altogether, we use two different sets of control variables. Our results demonstrate that an increase in traffic density increases the time cost of travel. All coefficients of interest are clustered at the route level and are statistically significant at any conventional level level.

While our OLS model with baseline controls yields an estimate of 0.067, our preferred IV estimate with baseline controls returns a higher elasticity of 0.492. The latter indicates that a 1% increase in traffic density increases the time cost of travel on average by 0.492%. The difference between the IV and the OLS estimate is statistically significant at any conventional level. The comparison suggests that the OLS results are biased toward zero. Our assessment of our instrument in Section 4 shows that it is conditionally independent and that the exclusion restriction is likely to hold. Therefore, the attenuation of OLS likely occurs because of potential reductions in traffic density in response to uncaptured shocks to time costs that drivers are aware of, such as construction works, road closures, or sporting events.

Table 3: Elasticity of Congestion

model	Dep. variable: log(time costs)			
	OLS	OLS	IV	IV
	1	2	3	4
log(traffic density)	0.067*** (0.002)	0.069*** (0.002)	0.492*** (0.031)	0.485*** (0.032)
baseline controls	yes	yes	yes	yes
quadratic & cubic weather	no	yes	no	yes
daylight	no	yes	no	yes
season	no	yes	no	yes
Kleibergen-Paap <i>F</i> -statistic		2,003	1,997	

Notes: This table presents results for the first stage on workdays between 04:00 and 21:59 CET following equation (1). The dependent variable is the logarithm of traffic density. Traffic density is the number of all commercial and non-commercial trips that intersect with a given trip's route. All specifications include fixed effects for the hour-of-day, the day-of-week, the route, and public holidays. The baseline controls include air temperature, total precipitation, wind speed, sunshine duration, and visibility. In Models 2 and 4, we also include quadratic and cubic weather controls, the share of daylight, and the season. The ultimate row shows the Kleibergen-Paap *F*-statistic for weak instruments. The sample size for all regressions is 1,712,138. Standard errors are clustered at the level of the route. *, **, and *** indicate statistical significance at the 5%, 1%, and 0.1% level, respectively.

Overall, the magnitude of our elasticity estimate is comparable to recent contributions to the congestion literature. For instance, using Uber Movement Speeds for London, Herzog (2023) recovers an average elasticity of 0.41 at the trip level. However, the average estimate of 0.73 in Mangrum and Molnar (2020) who use the roll-out of “boro” taxis in New York as an instrument to estimate the effect of vehicle density on travel times considerably exceeds our estimate. Additionally, Geroliminis and Daganzo (2008) find a higher elasticity of 0.65 based on GPS data from 140 taxis in Yokohama, Japan. There are also studies that recover estimates that are lower. First, analyses focusing on highways tend to find lower congestion elasticities. For instance, Yang et al. (2017) exploit driving restrictions as an instrument to estimate an elasticity of 0.14 for major ring roads in Beijing Second, studies based on data from travel surveys often recover lower elasticities. For instance, Couture et al. (2018) estimate an average elasticity of about 0.15 for US cities while Akbar and Duranton (2017) find an average elasticity of 0.06 for Bogotá, Colombia. One potential reason for these smaller elasticities is attenuation from recall bias in surveys. In particular, survey respondents are likely to indicate

typical departure times that overestimate the concentration of actual departure times in the morning and evening,¹² which may lead to the overestimation of peak-hour traffic density. Finally, studies that instrument for traffic density with hour-of-day indicators to capture shifts in demand generally estimate lower elasticities. This applies to Akbar and Duranton (2017) but also Kreindler (2018) who estimate elasticities of about 0.14 for Bangalore, India (see also Russo et al. (2021) for evidence from Rome).¹³

Robustness Checks

Our baseline estimates are robust to various data and specification changes. To illustrate this, we build a specification chart by holding all baseline regression features of our preferred IV estimate from Model 3 of Table 3 constant and then varying only one feature at a time. Altogether, we vary 7 design elements: (1) how we use DBSCAN to assign individual trips to driver-OD clusters, (2) how we use QuickBundles to generate route fixed effects, (3) how we attribute whether a trip is treated by accidents along but not on its route, (4) what kind of instrument we use for estimation, (5) the stringency of our fixed effects, (6) how we accommodate for weather conditions and seasons, and (7) how we deal with potential outliers.

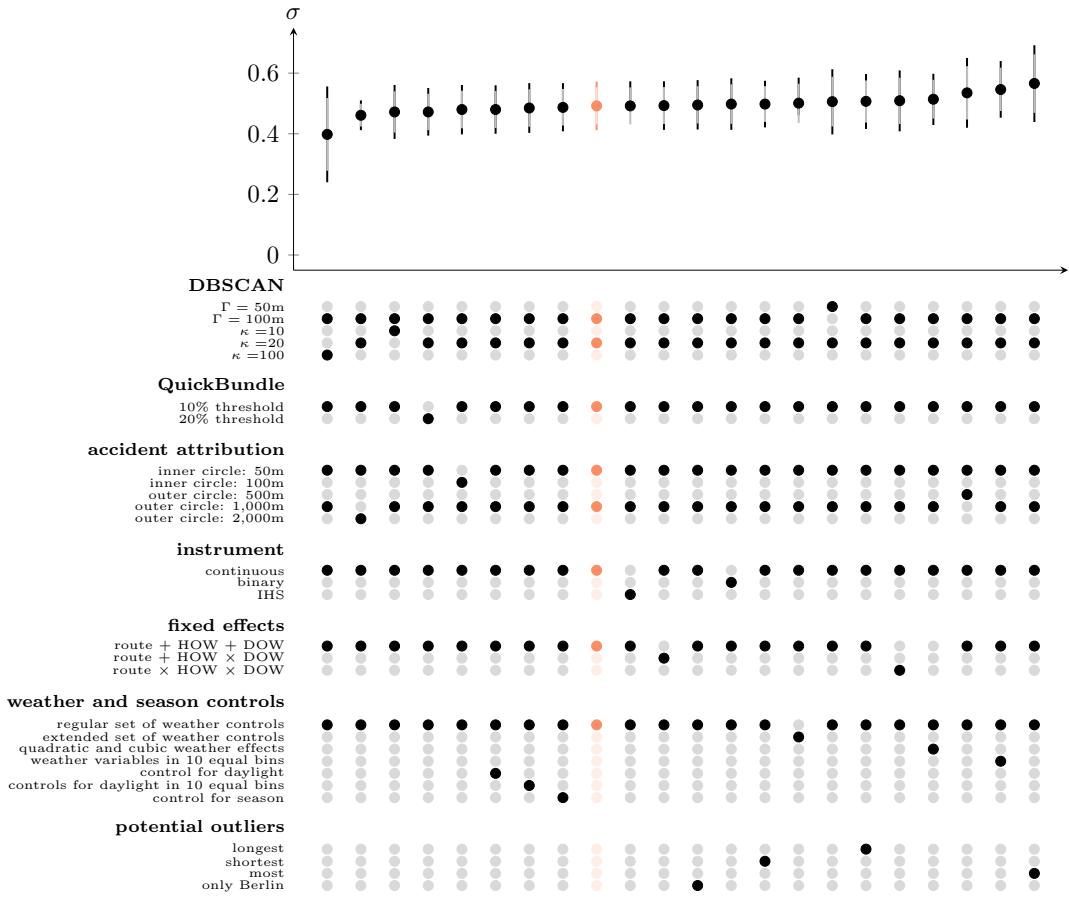
Figure 8 shows that all instrumental variable point estimates are between 0.398 and 0.566. The confidence intervals indicate that there are no significant differences across model specifications. We elaborate on selected robustness checks in the following paragraphs.

Choice of DBSCAN Parameters We vary the parameters Γ and κ that define how DBSCAN identifies driver-OD clusters. The smaller Γ , the closer the starting points have to be and the closer the end points have to be for trips to be considered directly density-reachable and to be assigned to the same driver-OD cluster. κ defines the minimum number of trips a driver-OD cluster needs to comprise. Consequently, the number of identified driver-OD clusters decreases when κ increases or when Γ decreases. Although the variation of our DBSCAN parameters has a significant impact on the number of driver-

¹² Zhao et al. (2015) document this bias by collecting and comparing data on the same sample and the same travels using two different methods: 1) a survey and 2) recording precise GPS coordinates via smartphone app in Singapore.

¹³ The elasticities discussed here for Geroliminis and Daganzo (2008) and Kreindler (2018) are not explicitly reported in the studies but inferred as an arc elasticity using their plots.

Figure 8: Specification Chart for the Elasticity of Congestion



Notes: The specification chart shows point estimates for congestion elasticities (full circles) and their corresponding 95% confidence intervals (light colored) and 99% confidence intervals (dark colored) for a host of different model specifications. The confidence interval of our preferred specification is indicated in an orange color. The range of our elasticities from instrumental variable regressions is between 0.398 and 0.566. Model specifications vary along the lines indicated in the table below the chart.

OD clusters in our regression sample¹⁴, the differences in our congestion elasticities based on these different samples are nevertheless small and insignificant. This alleviates concerns about potential sample selection when allocating trips to driver-OD clusters.

¹⁴E.g. when we focus on clusters with highly regular trips approximated by a five times higher minimum number of observed trips than in our main sample by selecting $\kappa = 100$ while holding the size of the radius of start- and endpoints constant at $\Gamma = 100$. This limits the sample to 2,629 driver OD-clusters with 277,044 observations. We also evaluate the effect of congestion on a subsample of clusters with regularly available parking spots approximated by tighter start- and endpoints by halving the radius to $\Gamma = 50$ while holding the minimum number of trips at the level of our main analysis at $\kappa = 20$. This limits our sample to 16,875 driver-OD clusters with 541,347 observations.

Choice of QuickBundles Parameters When applying QuickBundles to further subdivide driver-OD clusters into individual routes, one must specify a threshold. In our main specification, we consider that two trip share a common route if they share the same origin and destination and for which the deviation at the waypoint level does not deviate by more than 10% of the length of the trip. To evaluate the choice of this parameter on our results, we also ran our regressions with a threshold of 20%. This reduces the number of different routes we identify because some highly similar routes are now combined together. However, the choice of the threshold has virtually no impact on our estimate. This alleviates our concern that our baseline choice might lead to the detection of spurious and infrequently treated routes that may potentially bias our estimate of the congestion elasticity.

Incident Attribution When attributing incidents, we made choices regarding the radii of the inner and the outer circle when attributing treatments (see Figure 7). Here, we explore the effect of varying these radii. First, we explore the effect of increasing the size of the inner circle from our preference of 50 to 100 meters to address concerns about bias from mistakenly including incidences that actually occur on a trip’s route, in which case the variation in traffic density would not be exogenous anymore. Second, we explore the effect of decreasing the outer ring’s radius from our preference of 1,000 meters to 500 to focus only on incidents on roads in the very close vicinity or increasing it to 2,000 meters to allow for effects from incidents further away. Overall, we find that these variations have little effect on our estimate.

Instrument While our main specification uses the log of the number of traffic incidences as instrument, we already showed the first stage results when using a binary instrument (see Table 2). We now also show IV estimates for the congestion elasticity with a binary instruments and when using an inverse-hyperbolic sine (IHS) transformation. Altogether, these variations have little bearing on our estimate’s magnitude.

More stringent fixed effects We interact our fixed effects to increase our specification’s stringency. Our main specification comprises three separate, additive fixed effects for the route, the hour of the day, and the day of the week. To better accommodate for unobserved shocks that only occur at certain times on certain days, we first interact the

hour-of-day and the day-of-week indicators. For instance, traffic density may be higher in the early hours on Mondays or in the late hours on Fridays because of long distance commuters. However, this has hardly any bearing on our estimate.

Second, because individual routes may be subject to unobserved shocks at certain times on certain days, we interact the fixed effects for the route, the hour-of-day, and the day-of-week indicators for a highly flexible model. For instance, trips on a given route may exhibit higher time costs of travel because garbage collection regularly impedes the flow of traffic in the early hours on Tuesdays. Although this significantly increases the number of fixed effects, our estimate of the elasticity of congestion remains unchanged.

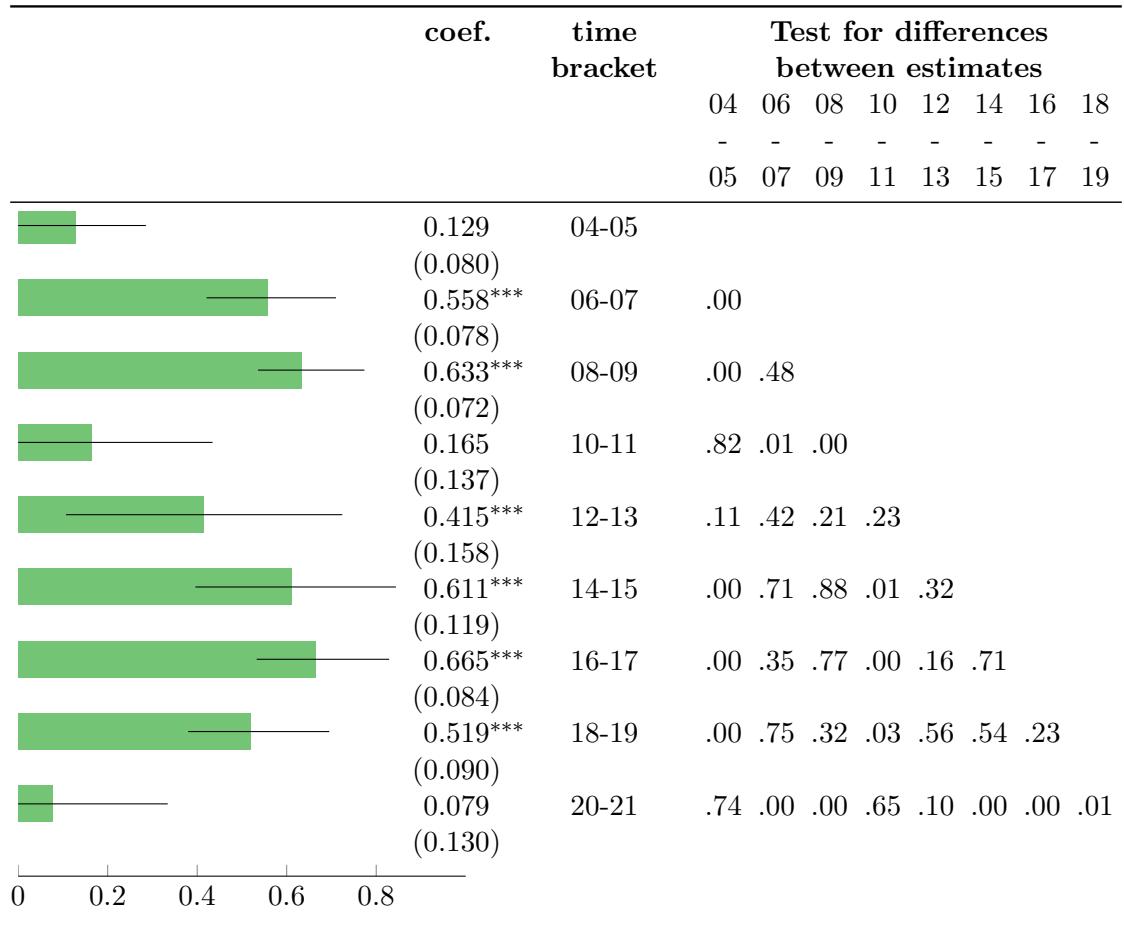
Weather and seasonal controls We also assess whether additional controls for weather or daylight conditions, or seasonal controls in linear, cubic and quadratic, or binned form inform our estimate of the elasticity of congestion. Controlling for daylight or seasons reduces the magnitude of our congestion elasticity minutely. Adding additional weather controls or controlling more flexibly increases the elasticity of congestion. However, neither variation in weather nor seasonal controls significantly changes our estimates.

Potential Outliers Finally, we test for the influence of potential outliers. To this end, we remove either the 10% of driver-OD clusters with the longest average trip distance, or the 10% with the shortest average trip distance, as well as the 10% of driver-OD clusters with the most trips taken. We also explore the effect of including only those trip that originate or ends outside in Berlin. Neither of these variations has any bearing on our results.

5.2 Temporal Heterogeneity

So far, we estimated σ assuming a constant elasticity over the entire day. Yet, traffic is highly concentrated at certain times of the day, and the congestion elasticity may therefore be subject to significant temporal heterogeneity. While the literature abounds with the use of peak and off-peak times to account for time-varying elasticities, our rich data allows us to provide estimates for the elasticity of congestion for individual hour

Table 4: *Temporal heterogeneity*



Notes: The Table shows the coefficients as bars and the 95% confidence interval as black lines. Standard errors in parentheses. ***(**, *) indicate statistical significance at the 0.1% (1%, 5%) level. *p*-values for test for differences between estimates.

brackets using our preferred specification (Model 3 in Table 3). For sufficient power, we divide the daytime hours 04:00 through 21:59 into two-hour brackets.

Table 4 shows significant temporal heterogeneity in the elasticity of congestion. The power of our analysis is sufficiently strong to detect significant differences in 15 out of 36 comparisons across the hour brackets. Four relevant patterns show up in the variation throughout the day. First, the morning is divided into two distinct regimes. The first half indicates a clear morning peak with a high congestion elasticity of up to 0.63 between 08:00 and 09:59. However, in the second half between 10:00 and 11:59, we find a very strong reduction in the elasticity to a level that is no longer statistically

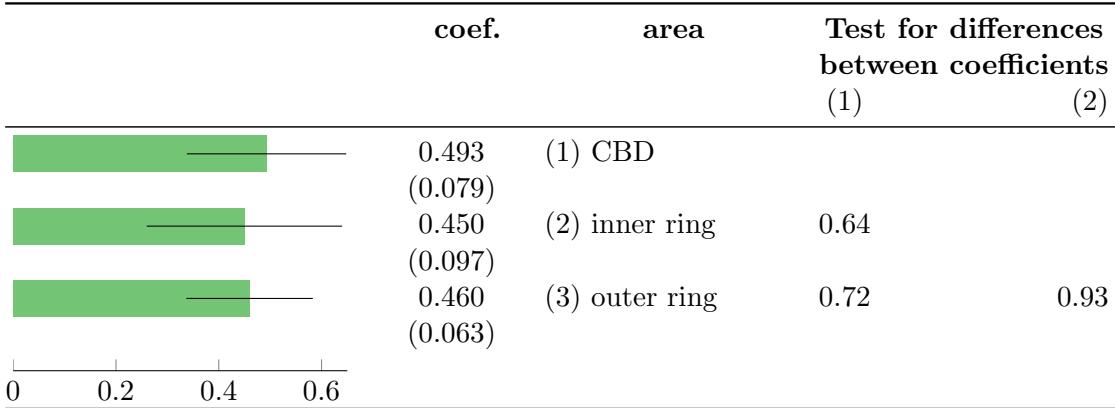
significant. This suggests that road capacity quickly recovers after the first morning rush and any congestion pricing policy would need to take into account this pronounced temporal heterogeneity. Second, our findings suggest that the afternoon is much more homogeneous in terms of the estimated elasticities. The highest estimates occur between 14:00 and 16:59, when σ peaks at 0.67. After noon between 12:00 and 13:59 and in the early evening between 18:00 and 19:59 the elasticity remains relatively high and our t -tests cannot detect significant differences to the peak hours between 14:00 and 16:59. Third, the difference in the magnitude of the elasticities between the afternoon and the morning peak are statistically insignificant. Finally, as expected, we find the lowest and statistically insignificant estimates for the congestion elasticity in the early hours of the day before 06:00 and in the late afternoon after 20:00. We do not provide estimates for the uncongested night hours for the sake of brevity.

We also estimate congestion elasticities for the popular binary distinction of peak and off-peak hours. Based on the official definition of Berlin's transport authorities, peak hours are between 6:00 and 9:59 and between 14:00 and 19:59. We estimate a corresponding elasticity of 0.551 with a standard error of 0.035 for this time period. For off-peak hours for the intervals between 04:00 and 05:59, 10:00 and 11:59, and 20:00 and 21:59, we recover a congestion elasticity of 0.261 with a standard error of 0.066. While we do not have the power to detect significant differences to our hour-bracket specific elasticities discussed above, a comparison still provides suggestive evidence that the binary distinction between peak and off-peak fails to capture the temporal highs and lows in the congestion elasticity.

5.3 Spatial Heterogeneity

In addition to temporal heterogeneity there might also exist heterogeneity across space. To measure this spatial heterogeneity, we estimate separate congestion elasticities for three areas. First, we estimate congestion elasticities for trips that occur in the area enclosed by Berlin's circular railway, the "inner ring." With a length of about 37 km, the circular railway encompasses Berlin's inner city which accounts for roughly 10% of the entire city area but about 20% of its population. Because of its density, one could hypothesize that congestion might be worse there. Second, we estimate elasticities for trips that relate exclusively in the "outer ring," the suburban area of Berlin not enclosed

Table 5: *Spatially differentiated Congestion Elasticities*



Notes: The figure shows the magnitudes of spatially differentiated elasticities as bars and the 95% confidence interval as black lines. CBD refers to Berlin’s central business district define by the boroughs of Mitte, Charlottenburg-Wilmersdorf, and Friedrichshain-Kreuzberg. The remaining coefficients are for trips driven exclusively within Berlin’s “inner ring” defined by the circular railway, and trips driven exclusively in the “outer ring,” the suburban area of Berlin. Standard errors in parentheses. ***(**, *, †) indicate statistical significance at the 0.1% (1%, 5%, 10%) level. *p*-values for test for differences between coefficients.

by Berlin’s circular railway. Third, we estimate an elasticity for Berlin’s central business district (CBD).

Table 5 shows that level of spatial heterogeneity is negligible. The elasticities for the different parts of the city are statistically indistinguishable. This finding may be explained by the fact that Berlin has evenly distributed sub-centers.

6 Welfare Effects

6.1 Theoretical Framework for Optimal Policy Design

From our empirical model in equation (1), we derive a supply function of travel. We define $\log \Omega = \beta X'_t + \gamma_i + \delta_s + \psi_w + \varphi_d + \omega_h + \epsilon_{it}$. Substitution into (1), while suppressing the subscript it , returns $C(D) = \Omega \cdot D^\sigma$, a function that describes how the average time cost of a kilometer of travel is affected by the aggregate number of vehicles per kilometer. Traffic conditions can also be measured in terms of the flow of vehicles defined as $V = D/C$. Substituting this expression into $C(D)$ and introducing a time indicator for different periods of the day T , we derive an aggregate inverse supply curve for automobile travel

$$C(V_T) = \Omega_T^{\frac{1}{1-\sigma_T}} \cdot V_T^{\frac{\sigma_T}{1-\sigma_T}}. \quad (3)$$

Equation (3) gives the average time cost of a kilometer of travel as a function of the aggregate flow of vehicles measured in terms of vehicles per kilometer per hour.¹⁵ The magnitude of σ_T determines the cost curves as illustrated in Figure 9 for the simplified case with $T \in \{0, 1\}$, where we treat $T = 0$ as off-peak and $T = 1$ as peak traffic hours. In our welfare calculation below, we consider the nine hour brackets T for which we recovered σ_T .

Drivers optimize their travel behavior according to their private costs described by equation (3) while disregarding the costs they impose on others. The marginal cost function MC reflects the private costs of an additional vehicle per kilometer per hour *and* the extent to which this additional vehicle slows other drivers. Multiplying the average cost of travel $C(V_T)$ in (3) by V_T and differentiating returns

$$MC(V_T) = \frac{C(V_T)}{1 - \sigma_T}. \quad (4)$$

This is the marginal costs curve in Figure 9 illustrated for off-peak times in the lower panel and for peak times in the upper panel.

For a welfare analysis, we need to make assumptions about the demand for travel. We define demand as $V_T = \Phi_T \cdot C_T^{-\eta}$. Φ_T is a time-varying demand shifter and η is the price elasticity of travel demand which we assume is constant. Rearranging this expression yields the inverse demand curve

$$C(V_T) = \Phi_T^{\frac{1}{\eta}} \cdot V_T^{-\frac{1}{\eta}}. \quad (5)$$

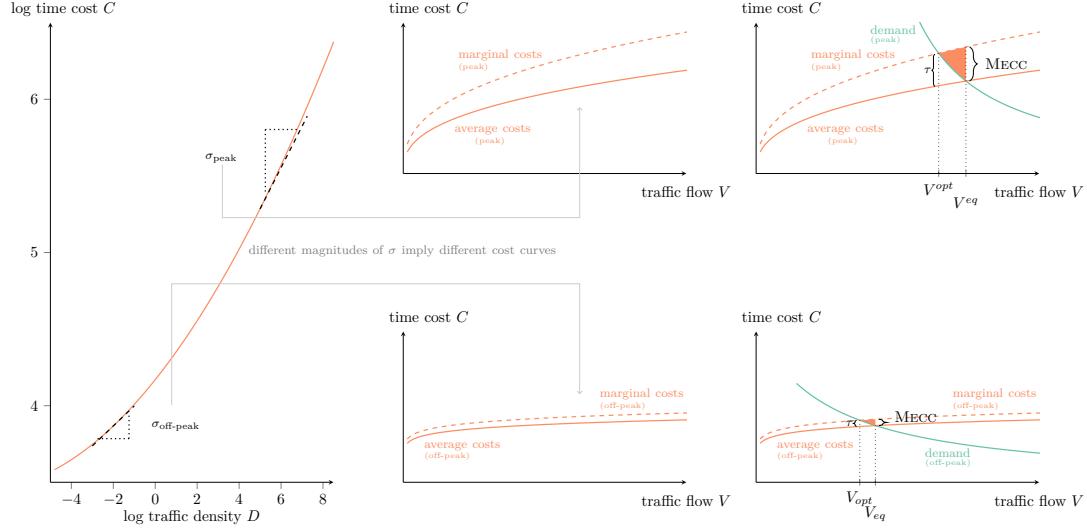
In the lower rightmost panel, Figure 9 shows the demand curve for off-peak hours when Φ_T is small. We present peak hours in the upper rightmost panel when Φ_T is high.

Following Vickrey (1969), the rightmost panels of Figure 9 illustrate the resulting partial equilibrium. The intersection of the average cost and the demand curves indicate the equilibrium travel flow $V_{T,eq}$. The social optimal level of travel flow $V_{T,opt}$ occurs

¹⁵In line with the prior literature, we recover σ from the effect of traffic density on the time cost of travel because speed and travel time are monotonic functions of density regardless of the existence of hypercongestion, which would lead to a non-monotonic relationship between flow and travel time.

Figure 9: Supply Elasticity and Congestion

(A) Inverse speed-density curve and the supply of travel



Notes: This figure links traffic density D and the flow of traffic V to the time cost of travel C , the marginal external costs of congestion (MECC), the deadweight loss caused by congestion, and its remedy in form of a congestion charge τ . The leftmost panel shows a stylized relationship between the log of traffic density D and time costs C . The elasticity of travel time costs with respect to traffic density σ measures the slope of the curve. It becomes steeper as D increases. The two diagrams in the center illustrate that the shape of the cost curves depends on the magnitude of σ . They show that the wedge between the average and the marginal cost curve is larger the higher the magnitude of σ becomes. Because $\sigma_{\text{peak}} > \sigma_{\text{off-peak}}$, the wedge is higher during peak times. Finally, the rightmost panel links the shape of the cost curves to the demand of travel to infer the size of the deadweight loss, the MECC, and the magnitude of the optimal congestion charge τ . The lower panel shows how a low demand combined with a shallow-sloped cost curve during off-peak times causes a small MECC and a small deadweight loss that requires a small τ to alleviate. The upper panel shows how a high demand during peak times combined with a steeper-sloped cost curve causes a high MECC and a high deadweight loss that requires a high τ .

at the intersection of marginal cost curve and demand curve. The difference between $MC(V_{T,eq})$ and $C(V_{T,eq})$ is the marginal external cost of traffic congestion $MECC(V_{T,eq}) = \frac{\sigma_T}{1-\sigma_T} \cdot C(V_{T,eq})$ in the unregulated equilibrium in period T .

The lower rightmost panel of Figure 9 highlights that during off-peak hours the combination of a low level of demand and a low supply elasticity $\sigma_{\text{off-peak}}$ results in a small welfare loss from unpriced congestion indicated by the orange triangle. However, when a high demand during peak hours coincides with a high supply elasticity σ_{peak} , the

resulting level of welfare loss is also high. Therefore, different hours of the day require different levels of congestion charges τ_T to restore social equilibria.

In Appendix D, we derive specific expressions for the equilibria and the optimal level of travel illustrated in Figure 9. We also derive the following two expressions for the optimal congestion charge and the welfare gain it provides. First, the optimal congestion charge τ_T measured in hours per kilometer is a function of the supply elasticity σ_T , the demand elasticity η and the equilibrium cost of travel $C(V_{T,eq})$:

$$\tau_T = \sigma \cdot C(V_{T,eq}) \cdot (1 - \sigma_T)^{\left(\frac{\eta \cdot \sigma_T}{1 - \sigma_T \cdot (1 - \eta)}\right) - 1}. \quad (6)$$

Second, the welfare gain from congestion pricing equals the deadweight loss resulting from the excess equilibrium congestion in the absence of the policy. As a share of total travel time it only depends on σ_T and η :

$$\Delta_T = 1 - (1 - \sigma_T)^{\frac{\eta}{1 - \sigma_T \cdot (1 - \eta)}} - \frac{\eta}{\eta - 1} \cdot \left(1 - (1 - \sigma_T)^{\frac{(1 - \sigma_T)(\eta - 1)}{1 - \sigma_T \cdot (1 - \eta)}}\right). \quad (7)$$

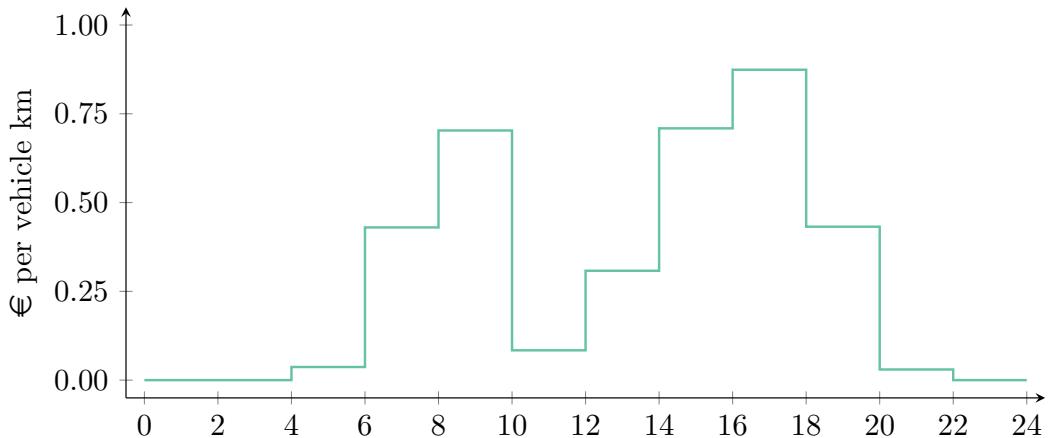
We rely on the literature for a long-term estimate of η . Akbar and Duranton (2017) is one of the few papers that directly estimate the elasticity of travel demand with respect to the time costs of travel, but only for the short-term. Based on the willingness of drivers to take a trip at a given time and at given time costs, they estimate a short-term elasticity with a magnitude of 0.40. The seminal textbook of Small (2007) instead draws on case studies for a few cities to suggest a higher magnitude of about 1 for the elasticity of time costs. There is a much more extensive literature on the elasticity of travel demand with respect to fuel costs. The central tendency in this literature for the long-term fuel price elasticity is a magnitude of 0.3 (Parry, 2009; Anderson et al., 2016). Following Anderson (2014) and assuming that penalized time costs are 240% higher than fuel costs, this fuel price elasticity implies a magnitude for the long-term elasticity of travel demand with respect to total travel costs (i.e. fuel and time costs) of about 1.32 (see also Wang et al., 2018; Yang et al., 2017). Altogether, we subsequently use $\eta = 1$ as our baseline but conduct robustness checks for 0.5 and 1.5.

6.2 Marginal External Cost of Congestion

To calculate the MECC, we use our time-specific estimates for σ_T from Table 4. We monetize $\frac{\sigma_T}{1-\sigma_T} \cdot C(V_{T,eq})$ combining official statistics on the 2017 median hourly wage¹⁶ with a value of time of 33% of the median hourly wage underlying Germany's federal transport planning model (Axhausen et al., 2015). We calculate $C(V_{T,eq})$ by solving the equation system given by (3) and (5) after deriving Ω_T and Φ_T . Assuming one person per vehicle for a conservative cost estimate, Figure 10 shows the resulting MECC in € cent per vehicle kilometer.

The heterogeneity in the time-specific estimates of σ_T is reflected in the significant level of heterogeneity in the MECC (Figure 10). Accordingly, external costs accrue only between the times of 06:00 to 09:59 and 14:00 to 19:59. The peak hours in the morning (08:00 to 09:59) and afternoon (14:00 to 17:59) are the key cost drivers with MECC of more than 70€ cents per km. There are hardly any costs of congestion before noon (10:00 to 11:59) and in the early morning (before 6:00) or evening (after 20:00).

Figure 10: Marginal External Cost of Congestion



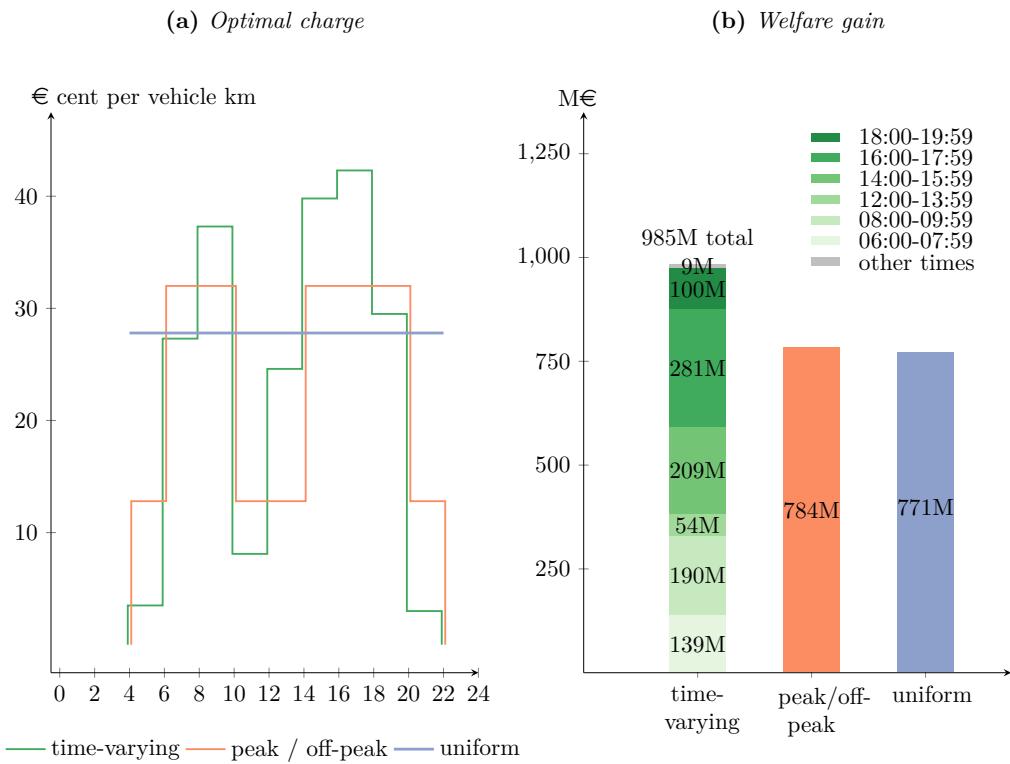
Notes: This figure shows the marginal external costs of congestion (MECC) by hour bracket according our preferred IV estimates.

¹⁶According to official statistics by “Statistische Ämter des Bundes und der Länder,” the 2017 median hourly wage in Berlin was 25.58€.

6.3 Optimal Congestion Charges

Next, we evaluate the optimal congestion charge that restores the social optimum based on Equation (6). As before, we monetize τ measured in minutes per kilometer from Equation (6) with 33% of the median hourly wage of 25.58 €. Just as in the case of the MECC calculation, we calculate $C(V_{T,eq})$ by solving the equation system given by (3) and (5).

Figure 11: Optimal Congestion Charges and Welfare Gains



Notes: The left panel of this figure shows the optimal congestion charge τ for $\eta = 1$ for three different charging regimes. The uniform charge of 27.9 € cents per km results from our mean estimate $\sigma = 0.493$ for an entire weekday (Table 3). The time-specific charges for peak and off-peak times and the hour brackets are based on the time-specific coefficients (see Section 5.2). Specifically, the charge for the peak times is 32.0 € cents per km

while it is 12.8 € cents per km for off-peak times. The charges for the hour brackets may be as low as 3.0 € cent per km between 20:00 and 21:59 or as high as 42.3 € cent per km between 16:00 and 17:59. While peak/off peak road pricing schemes are easy to implement, they do not address congestion optimally. In our case, they either overprice road use when congestion is low, as is here the case in the hours between 04:00 to 05:59, 10:00 to 11:59, and 20:00 to 21:59. Or it underprices road use in the hours between 08:00 to 09:59 and 14:00 to 17:59 when congestion is high.

To evaluate the economic significance of the suggested congestion charge it is also of interest to consider the per trip costs imposed by the per kilometer charge. Data from the German National Household Travel Survey MiD shows that the average trip length in Berlin is 12 kilometers. This means that the average driver would have to pay about 3.35 € for such a trip under uniform road charges. This is about the price of a standard public transportation ticket in Berlin. However, the same trip taken between 08:00 and 09:59 under a time-specific charging regime would cost about 4.48 €.

6.4 Welfare Gains

We calculate the welfare gains from implementing road congestion charges following (7). Our estimate for σ and our assumptions for η allow us to calculate Δ , which measures the welfare gain as a percentage of total travel time. Combining this share with Berlin specific information on travel times and the number of affected travelers returns the absolute welfare gain in hours per year, which we monetize.

For simplicity, we walk the reader through the calculation of the potential gains using a uniform congestion charge and $\eta = 1$. The calculations for the time-specific cases are identical with the exception that we use either peak/off-peak or time-bracket specific information on travel behavior. SrV (2018) indicates that individuals in metropolitan areas travel about 84 minutes per week day. Cyclists (18%), individuals in private vehicles (26%), and users of buses and trams (7%) are subject to congestion.¹⁷ The welfare gain from congestion charging is about $\delta = 14.9\%$ of total travel time. With about 49% of travel subject to congestion, the welfare gain is $14.9\% \cdot 84 \text{ minutes} \cdot 49\% \cdot 60 \text{ min/hour} = 0.1$ hours per individual and week day. With approximately 3.5M individuals in Berlin

¹⁷The share of individuals relying on public transport, including trains and subways that operate independently from road traffic conditions, is 26%. All mode shares are based on SrV (2018).

(AfSBB, 2018), the gain from a uniform congestion charge is about 90.4M hours per year. If we value time lost in traffic at 33% of the median hourly wage of 25.58 €, the annual welfare gain is about 771M€.

The right panel of Figure 11 indicates the magnitude of annual welfare gains from road congestion charges for the three charging regimes. While a uniform charge increases welfare by 771M€, peak charging increases welfare by about 784M€. Time-specific charging further increases welfare gains to about 985M€. Only about 78% of this welfare gain is achieved by the second-best, uniform congestion charge. Notably, the aggregate welfare gains for the rush-hours between 17:00 and 18:59 when traffic is worst alone account for 281M€ of the total 985M€ or about 28.5%. Overall, our findings suggest that a uniform charge is the least efficient of the three charging regimes at improving social welfare.

6.5 Robustness Checks and Discussion

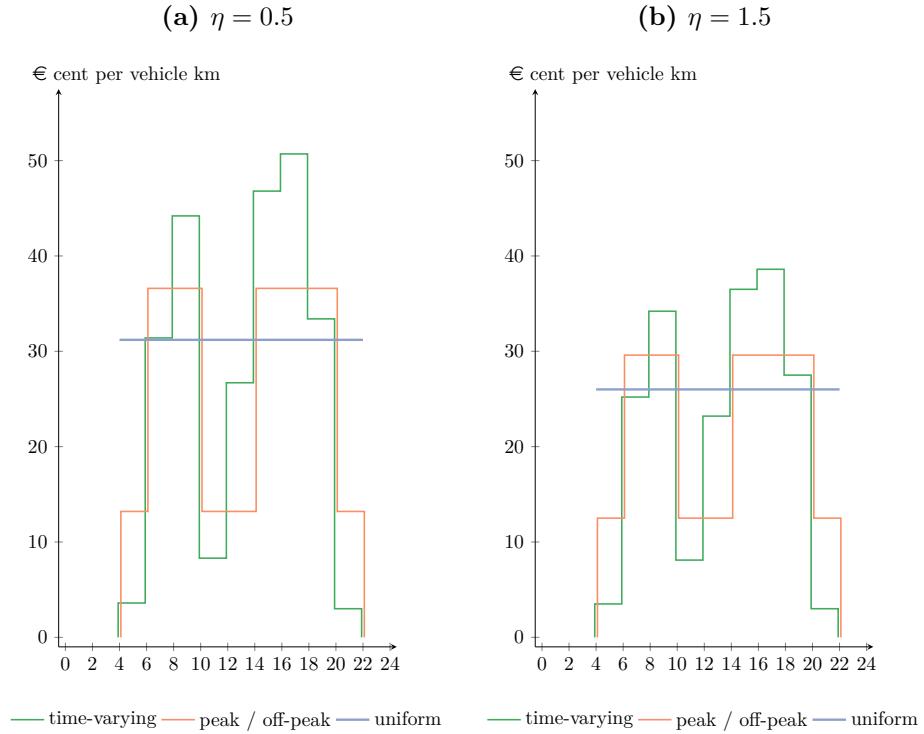
To check the robustness of our welfare estimates, we examine the effect of choosing two alternative demand elasticities from the literature. On the one hand, we examine the effect of a demand elasticity half as elastic as our baseline with $\eta = 0.5$. On the other hand, we examine the effect of a more elastic demand with $\eta = 1.5$.

Figure 12 indicates optimal congestion charges for $\eta = 0.5$ and $\eta = 1.5$. The comparison with the optimal charges in case of $\eta = 1$ (Figure 11) reveals that the optimal congestion charge decreases with increases in the magnitude of the demand elasticity.¹⁸ This is intuitive because the same travel reductions can be achieved with lower penalties in the form of congestion charges when the elasticity of demand increases. Overall, the choice of η has implications for the magnitude of the optimal charge for most the day. The difference is slightly more pronounced for the rush-hours when charges are highest.

The choice of the demand elasticity also has implications regarding the magnitude of the welfare gains achieved by imposing road charges. Figure 13 indicates that welfare gains increase with increases in the magnitude of η . This is intuitive because a more elastic demand implies that a lower congestion charge is needed in order to induce the same travel reduction, which produces a larger improvement in terms of speed and welfare gains. For $\eta = 1.5$, annual welfare gains are between 877M€ for peak-time

¹⁸We provide the specific values for all congestion charges for $\eta \in \{0.5, 1, 1.5\}$ in Table 11 in Appendix C.

Figure 12: Optimal Congestion Charges for Alternative Demand Elasticities

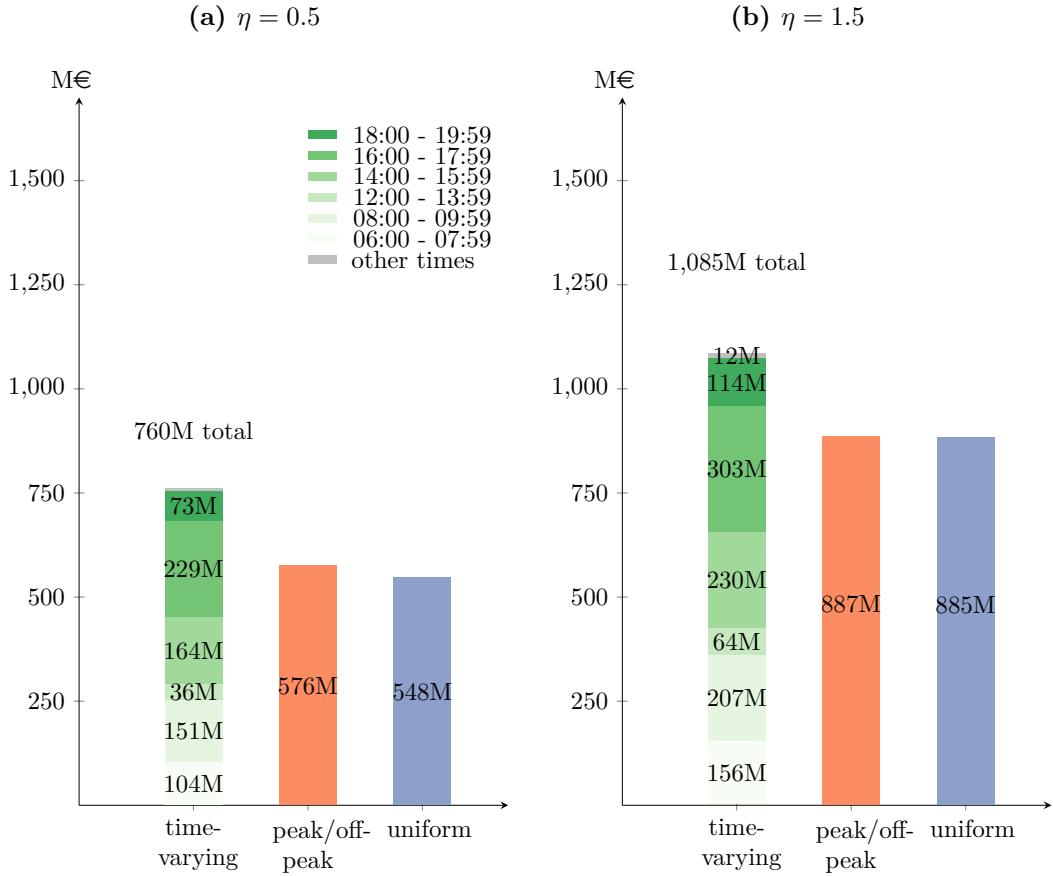


Notes: The left panel of this figure shows the congestion charges for $\eta = 0.5$ while the right panel does the same for $\eta = 1.5$. Three different congestion charging regimes are considered: uniform charges (blue), charges for peak/off-peak times (orange), and time-varying charges by hour-brackets (green).

charges and 1,085M€ for time-varying charges. More importantly, even in the case of a fairly inelastic demand, the magnitude of the welfare gains achieved by levying road charges is economically significant. For $\eta = 0.5$, annual welfare gains are between 885M€ for the uniform charges and 887M€ for time-varying charges.

We calculate that optimal, hour-bracket specific road charges induce welfare gains between between 0% and 29.9% of total travel time. This range includes the values indicated in Couture et al. (2018) who indicate a range between 0.9% and 6.5% in terms of total travel time for 100 U.S. metropolitan areas. However, our results are larger than those found by Akbar and Duranton (2017), who report a deadweight loss of 0.06% for their preferred estimates for Bogotá, Colombia. The difference is mainly driven by the less elastic supply in their study. For Beijing, Yang et al. (2017) estimate an annual welfare loss of 0.7 to 1.5 billion Yuan or 90 to 192 million Euro. Estimates for Europe

Figure 13: Welfare Gains for Alternative Demand Elasticities



Notes: The left panel of this figure shows the annual welfare gains for $\eta = 0.5$ while the right panel does the same for $\eta = 1.5$. Three different congestion charging regimes are considered: uniform charges (blue), charges for peak/off-peak times (orange), and time-varying charges by hour-brackets (green). Contributions to welfare gains outside of the times indicated are negligible.

indicate deadweight losses for Greater London of about 207 million Pounds Sterling or about 243 million Euro Herzog (2022) and for Prague of about 123 Million Euro annually Buchholz et al. (2020).

7 Conclusion

Using a representative machine-learned panel data set of GPS-coded trips, we propose a novel approach to estimate temporally heterogeneous costs of road congestion and to assess the benefits of optimal, time-varying, city-wide congestion charges for Berlin. We

pursue a multi-pronged approach. First, we use unsupervised machine learning to assign anonymized trips to individual drivers and track their repeated travel behavior under different traffic conditions throughout a full year. Second, we infer the level of traffic density for each observed trip at its respective point in time using the universe of the 970 million time-stamped and geo-coded trip waypoints in our sample from moving vehicles in Berlin. Third, to calculate the social cost of congestion and the optimal congestion charge, we estimate the causal relationship between the time cost of travel and traffic density. The identification of the causal effect of traffic density on the time cost of travel relies on exogenous increases from rerouting caused by traffic incidences on adjacent roads. Combining our estimated congestion elasticity with assumptions about travel demand allows us to evaluate the welfare implications of different congestion charging schemes.

Our results indicate that the average elasticity of travel time cost with respect to traffic density over an entire weekday is 0.492. This average elasticity, however, obscures significant intra-day variation. In particular, the congestion elasticity can be as low as 0.165 before noon and as high as 0.665 for the afternoon rush hours. Accordingly, we find large temporal heterogeneity in the marginal external costs of congestion between 3.0 and 87.4 € cents per vehicle kilometer during daytime. We calculate that optimal, time-specific congestion charges between 3.0 to 42.3 € cents per km increase annual welfare by 985 M€. Only about 78% of this welfare gain is achievable by a “second-best,” uniform congestion charge regime that levies 27.9 € cents per kilometer that overcharges drivers outside of the rush hours while it undercharges them during rush hours.

Overall, our findings suggest that the choice between a uniform or a time-varying congestion charge scheme is consequential. Uniform charges are less efficient in improving social welfare. These welfare gains are economically significant even in the case of a relatively inelastic demand. However, remaining uncertainty about the demand elasticity precludes more precision on the magnitude achievable in a real-world setting. Another limitation is that our congestion elasticity only considers the value of lost time. Accommodating the full range of the social costs of congestion such as air pollution, fuel consumption, or noise is beyond the scope of our paper. Therefore, we most likely underestimate the full benefits of time-varying congestion charges.

Our finding that time-varying charges improve welfare more than uniform ones raises important policy questions, especially because the implementation of time-specific road charges requires additional technical effort and political capital. Policy makers are likely to pay particular attention to these direct costs if voters underrate the beneficial equilibrium effect of the charge in terms of a lower traffic density (Dal Bó et al., 2018). Leveraging advances in mobile communication and computer technology provide a major lever to reduce implementation costs (Cramton et al., 2018). Revenue recycling is crucial for securing public support, for instance earmarking revenues for improving public transport (Anas and Lindsey, 2011).

We conclude by indicating directions for future research. First, the research on optimal congestion charges would benefit greatly from precise *long-term* demand estimates. However, the requirements for a full modeling of demand in terms of the choice of destinations, departure times, travel modes, and routes are high. Even the comparatively easier task of estimating the *short-term* effect of the time cost of travel on the willingness of individual drivers to travel at a given time faces serious obstacles. For example, modeling the decision to drive requires a measure of counterfactual travel time costs for times at which the trip was not taken. We would also need information regarding a trip's urgency and any time constraints that determine when it starts. We hope that future research can overcome these data and identification obstacles, maybe along the suggestions by Akbar and Duranton (2017).

Second, we use DBSCAN to identify repeated trips with highly similar origin and destination by the same individual. We further subdivide trips of similar original and destination into distinct routes with the help of QuickBundles. Other approaches or algorithms to derive similar identifiers may also prove worthwhile, especially if these allow the detection of trip chaining, e.g. using probabilistic chaining algorithms (citation).

Third, while our data is simultaneously very rich in trip observations and very precise with respect to time and spatial information, we lack socio-economic information at the level of the individual driver. Socio-economic data could help identify who is most affected by congestion, who bears the burden of congestion charging, and who benefits most from reduced congestion levels. Finally, floating car data is becoming more readily available which convinces us that our approach is applicable to other research questions and contexts.

References

- Adler, M. W., F. Liberini, A. Russo, and J. N. van Ommeren (2018). Road congestion and public transit. Working Paper.
- AfSBB (2018). Mikrozensus basisdaten. *Amt für Statistik Berlin-Brandenburg*.
- Akbar, P. A., V. Couture, G. Duranton, and A. Storeygard (2018). Mobility and congestion in urban India. *NBER Working Paper 25218*.
- Akbar, P. A. and G. Duranton (2017, April). Measuring the Cost of Congestion in Highly Congested City: Bogotá. *Working Paper*.
- Amador, F. J., R. M. González, and J. d. D. Ortúzar (2005, April). Preference Heterogeneity and Willingness to Pay for Travel Time Savings. *Transportation* 32(6), 627–647.
- Anas, A. and R. Lindsey (2011). Reducing urban road transportation externalities: Road pricing in theory and in practice. *Review of Environmental Economics and Policy* 5(1), 66–88.
- Anderson, M. L. (2014, September). Subways, Strikes, and Slowdowns: The Impacts of Public Transit on Traffic Congestion. *American Economic Review* 104(9), 2763–2796.
- Anderson, M. L. and L. W. Davis (2020). An empirical test of hypercongestion in highway bottlenecks. *Journal of Public Economics* 187, 104197.
- Anderson, M. L., F. Lu, Y. Zhang, J. Yang, and P. Qin (2016). Superstitions, street traffic, and subjective well-being. *Journal of Public Economics* 142, 1–10.
- Anderson, S. T., J. M. Sallee, et al. (2016). Designing policies to make cars greener. *Annual Review of Resource Economics* 8(1), 157–180.
- Athey, S. and G. W. Imbens (2019). Machine learning methods that economists should know about. *Annual Review of Economics* 11(1), 685–725.
- Axhausen, K. W., I. Ehreke, A. Glemser, S. Hess, C. Jödden, K. Nagel, A. Sauer, and C. Weis (2015). Schlussbericht: FE-Projekt-Nr. 96.996/2011: Ermittlung von Bewertungsansätzen für Reisezeiten und Zuverlässigkeit auf Basis der Schätzung eines

Modells für modale Verlagerungen im nicht-gewerblichen und gewerblichen Personenverkehr für die Bundesverkehrswegeplanung. Technical report, ETH Zurich.

Barahona, N., F. A. Gallego, and J.-P. Montero (2020). Vintage-specific driving restrictions. *The Review of Economic Studies* 87(4), 1646–1682.

Bento, A., J. D. Hall, and K. Heilmann (2021). Estimating Congestion Externalities using Big Data. *Working Paper*.

Bento, A. M., K. Roth, and A. Waxman (2020, December). Avoiding Traffic Congestion Externalities? The Value of Urgency. *NBER Working Paper No. 26956*.

Bhuller, M., G. B. Dahl, K. V. Løken, and M. Mogstad (2020). Incarceration, recidivism, and employment. *Journal of Political Economy* 128(4), 1269–1324.

Buchholz, N., L. Doval, J. Kastl, F. Matějka, and T. Salz (2020). The value of time: Evidence from auctioned cab rides. *NBER Working Paper 27087*.

Chandra, E. and V. Anuradha (2011, June). A Survey on Clustering Algorithms for Data in Spatial Database Management Systems. *International Journal of Computer Applications* 24(9), 19–26.

Cinelli, C. and C. Hazlett (2020). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B - Statistical Methodology* 82(1), 39–67.

Cinelli, C. and C. Hazlett (2022). An omitted variable bias framework for sensitivity analysis of instrumental variables. *SSRN Working Paper 4217915*.

Cornago, E., A. Dimitropoulos, and W. Oueslati (2019). Evaluating the impact of urban road pricing on the use of green transport modes: The case of Milan.

Couture, V., G. Duranton, and M. A. Turner (2018, October). Speed. *The Review of Economics and Statistics* 100(4), 725–739.

Cramton, P., R. R. Geddes, and A. Ockenfels (2018). Set road charges in real time to ease traffic. *Nature* 560(7716), 23–26.

- Cunningham, B., J. LaRiviere, and C. J. Wichman (2021). Clustered into control: Heterogeneous causal impacts of water infrastructure failure. *Economic Inquiry*.
- Currie, J. and R. Walker (2011). Traffic congestion and infant health: Evidence from e-zpass. *American Economic Journal: Applied Economics* 3(1), 65–90.
- Dal Bó, E., P. Dal Bó, and E. Eyster (2018). The demand for bad policy when voters underappreciate equilibrium effects. *The Review of Economic Studies* 85(2), 964–998.
- Davis, L. W. (2008). The effect of driving restrictions on air quality in Mexico City. *Journal of Political Economy* 116(1), 38–81.
- De Montjoye, Y.-A., C. A. Hidalgo, M. Verleysen, and V. D. Blondel (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific Reports* 3, 1–5.
- Ester, M., H.-P. Kriegel, J. Sander, and X. Xu (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96)*, Portland, Oregon, pp. 226–231. AAAI Press.
- Firdaus, S. and M. A. Uddin (2015, March). A Survey on Clustering Algorithms and Complexity Analysis. *IJCSI International Journal of Computer Science Issues* 12(2), 62–85.
- Fosgerau, M. and K. A. Small (2013). Hypercongestion in downtown metropolis. *Journal of Urban Economics* 76, 122–134.
- Fosgerau, M. and K. Van Dender (2013). Road pricing with complications. *Transportation* 40, 479–503.
- Fowlie, M. and N. Muller (2019). Market-based emissions regulation when damages vary across sources: What are the gains from differentiation? *Journal of the Association of Environmental and Resource Economists* 6(3), 593–632.
- Gallego, F., J.-P. Montero, and C. Salas (2013). The effect of transport policies on car use: Evidence from Latin American cities. *Journal of Public Economics* 107, 47–62.

- Garyfallidis, E., M. Brett, M. M. Correia, G. B. Williams, and I. Nimmo-Smith (2012). Quickbundles, a method for tractography simplification. *Frontiers in neuroscience* 6, 175.
- Geroliminis, N. and C. F. Daganzo (2008, November). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological* 42(9), 759–770.
- Gibson, M. and M. Carnovale (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics* 89, 62–73.
- Green, C. P., J. S. Heywood, and M. Navarro (2016). Traffic accidents and the London congestion charge. *Journal of Public Economics* 133, 11–22.
- Grengs, J., X. Wang, and L. Kostyniuk (2008, August). Using GPS Data to Understand Driving Behavior. *Journal of Urban Technology* 15(2), 33–53.
- Gu, Y., E. Deakin, and Y. Long (2017). The effects of driving restrictions on travel behavior evidence from Beijing. *Journal of Urban Economics* 102, 106–122.
- Hahsler, M., M. Piekenbrock, S. Arya, and D. Mount (2018). dbscan: Density Based Clustering of Applications with Noise (DBSCAN) and Related Algorithms.
- Hanna, R., G. Kreindler, and B. A. Olken (2017). Citywide effects of high-occupancy vehicle restrictions: Evidence from “three-in-one” in Jakarta. *Science* 357(6346), 89–93.
- Herzog, I. (2021). The City-Wide Effects of Tolling Downtown Drivers: Evidence from London’s Congestion Charge.
- Herzog, I. (2022). The marginal external cost of traffic in Greater London. *Working Paper*.
- Herzog, I. (2023). The Marginal External Cost of Traffic in Greater London. *Working Paper*.
- Holland, S. P., E. T. Mansur, N. Z. Muller, and A. J. Yates (2016). Are there environmental benefits from driving electric vehicles? The importance of local factors. *American Economic Review* 106(12), 3700–3729.

- Jacobsen, M. R., C. R. Knittel, J. M. Sallee, and A. A. Van Benthem (2020). The use of regression statistics to analyze imperfect pricing policies. *Journal of Political Economy* 128(5), 1826–1876.
- Jain, A. K. (2010, June). Data Clustering: 50 Years beyond K-means. *Pattern Recognition Letters* 31(8), 651–666.
- Kahneman, D. and A. B. Krueger (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives* 20(1), 3–24.
- Karlström, A. and J. P. Franklin (2009). Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial. *Transportation Research Part A: Policy and Practice* 43(3), 283–296.
- Karlström, A. and J. P. Franklin (2009, March). Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial. *Transportation Research Part A: Policy and Practice* 43(3), 283–296.
- Keeler, T. E. and K. A. Small (1977). Optimal peak-load pricing, investment, and service levels on urban expressways. *Journal of Political Economy* 85(1), 1–25.
- Kreindler, G. E. (2018). The Welfare Effect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications*. *Working Paper*.
- Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022, October). Valid t-ratio inference for iv. *American Economic Review* 112(10), 3260–90.
- Mangrum, D. and A. Molnar (2020). The marginal congestion of a taxi in New York City. *Working Paper*.
- Martin, L. A. and S. Thornton (2017). Can Road Charges Alleviate Congestion? *SSRN Electronic Journal*.
- MiD (2018). Mobilität in deutschland- mid: Ergebnisbericht.
- Miyauchi, Y., K. Nakajima, and S. J. Redding (2021). Consumption access and agglomeration: Evidence from smartphone data. *NBER working paper 28497*.

- Mosquera, R. (2021). Stuck in traffic: Measuring congestion externalities with negative supply shocks. *SSRN Working Paper 3400807*.
- Muller, N. Z. and R. Mendelsohn (2009). Efficient pollution regulation: Getting the prices right. *American Economic Review* 99(5), 1714–39.
- Parry, I. W. (2009, October). Pricing Urban Congestion. *Annual Review of Resource Economics* 1(1), 461–484.
- Pigou, A. (1920). *The Economics of Welfare*. London: MacMillan.
- Rossi, L., J. Walker, and M. Musolesi (2015, December). Spatio-Temporal Techniques for User Identification by Means of GPS Mobility Data. *EPJ Data Science* 4(1).
- Russo, A., M. W. Adler, F. Liberini, and J. N. van Ommeren (2021). Welfare losses of road congestion: evidence from rome. *Regional Science and Urban Economics* 89, 103692.
- Santos, G. and G. Fraser (2006). Road pricing: Lessons from London. *Economic Policy* 21(46), 264–310.
- Simeonova, E., J. Currie, P. Nilsson, and R. Walker (2019). Congestion pricing, air pollution, and children’s health. *Journal of Human Resources*, 0218–9363R2.
- Small, K. A. (2007, October). *The Economics of Urban Transportation* (2 ed.). Routledge.
- Small, K. A., C. Winston, and J. Yan (2005, July). Uncovering the Distribution of Motorists’ Preferences for Travel Time and Reliability. *Econometrica* 73(4), 1367–1382.
- SrV (2018). Tabellenbericht zum forschungsprojekt “mobilität in städten – srv 2018” in berlin.
- Tang, C. K. (2021). The cost of traffic: Evidence from the London congestion charge. *Journal of Urban Economics* 121, 103302.

- Verhoef, E. (Ed.) (2010). *The Economics of Traffic Congestion (Volumes I and II)*. Number 244 in The International Library of Critical Writings in Economics. Edward Elgar.
- Verhoef, E. T. (2003). Inside the queue:: hypercongestion and road pricing in a continuous time-continuous place model of traffic congestion. *Journal of Urban Economics* 54(3), 531–565.
- Viard, V. B. and S. Fu (2015). The effect of Beijing’s driving restrictions on pollution and economic activity. *Journal of Public Economics* 125, 98–115.
- Vickrey, W. S. (1969). Congestion Theory and Transport Investment. *The American Economic Review* 59(2), 251–260.
- Wang, Q., N. E. Phillips, M. L. Small, and R. J. Sampson (2018, July). Urban Mobility and Neighborhood Isolation in America’s 50 Largest Cities. *Proceedings of the National Academy of Sciences* 115(30), 7735–7740.
- Yang, J., A.-O. Purevjav, and S. Li (2017). The Marginal Cost of Traffic Congestion and Road Pricing: Evidence from a Natural Experiment in Beijing. *SSRN Electronic Journal*.
- Zhao, F., F. C. Pereira, R. Ball, Y. Kim, Y. Han, C. Zegras, and M. Ben-Akiva (2015, December). Exploratory Analysis of a Smartphone-Based Travel Survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board* 2, 45–56.

Appendix A Data

Appendix A.1 Traditional Methods of Data Collection

Traditional methods of data collection use questionnaires to gather self-reported information. While individuals inform precisely on socio-economic variables, their ability to reveal precise information on trips is limited. We break with traditional approaches and use automatically generated, time-stamped, and geo-coded information. This means that we have very precise information on trips, while the availability of socio-economic information is limited. There are three reasons why our approach is better suited for analyzing congestion than the traditional approach.

First, the precision afforded by automatic measurements of time and location for individual trips is superior to travel information that individuals recall from memory in surveys. Moreover, automation means that every trip is recorded and every measurement is of similar quality. This prevents well-known biases from estimated travel times¹⁹ or the preference for recording simple over complicated trips (Grengs et al., 2008).

Second, automation allows data collection over indefinite periods of time. We are able to follow drivers over the course of an entire year. Most travel diaries, however, are limited to one or two days given the reporting burden to survey respondents. This is problematic because the number and the purpose of trips varies considerably between days (Zhao et al., 2015).

Third, with more than 30 million individual trips at hand, we can infer traffic density for short time intervals across the day. This allows us to observe how an increasing number of drivers on the road affects traffic flow. Even if travel surveys were able to collect information on many trips, the number of road users at the time of the individual trips remained unknown, and the significant share of commercial vehicles (20% of all trips in our data), that we observe, would be neglected.

Appendix A.2 Additional Data Sources

Germany’s national meteorological service Deutscher Wetterdienst (DWD) provides comprehensive information on a range of weather conditions at the hourly level which we

¹⁹For instance, travel start time is often reported as the time when leaving the house, rather than the time when beginning driving.

use as additional controls. These conditions include air temperature, humidity, precipitation, air pressure, wind speed, cloud cover, and sunshine duration. They may affect travel conditions and thereby either alter the time cost of travel or the number of travelers at a given time, in particular. We create variables on hourly, Berlin-wide weather conditions by averaging across weather stations. Table 6 provides descriptive statistics.

Table 6: Weather Conditions for the main regression sample

Variable	Units	Mean	Percentiles		
			Median	25th	75th
Air Temperature	°C	11.088	10.660	5.275	16.980
Humidity	%	0.735	0.772	0.620	0.876
Total Precipitation	mm	0.096	0.000	0.000	0.000
Air Pressure	Hectopascal	1,015.909	1,016.533	1,010.267	1,021.700
Sunshine Duration	hours	0.217	0.000	0.000	0.406
Wind Speed	m/sec	5.597	5.317	3.533	7.500
Cloud Cover	%	0.712	0.870	0.500	1.000
Visibility	km	29.009	30.000	15.000	40.000

Notes: This table summarizes the weather conditions across the observations in our main regression. We obtained all hourly observations from DWD.

Finally, household travel information, for example on annual mileage, average trip lengths, and vehicle ownership are from the German National Household Travel Survey (MiD) and SrV (2018). Spatially disaggregated socio-economic status is provided by Acxiom. The EU Copernicus Programm's Urban Atlas provides detailed digital maps that reveal land-use information based on thousands of high-resolution, satellite images. We refer to sources for individual figures within the body of the text.

Appendix B Additional Information on DBSCAN

Appendix B.1 Density-Reachability Definitions

Definition 1. Let $u, v \in V$ be two points. u is *directly density-reachable* from v within V with respect to Γ and κ if and only if v is a core point and u is in its Γ -neighborhood, i.e. $u \in N_\Gamma(v)$.

Definition 2. Let $u, v \in V$ be two points. u is *density-reachable* from v within V with respect to Γ and κ if there is a chain of points p_1, \dots, p_n such that $p_1 = v, p_n = u$ and for each $i = 2, \dots, n$ it holds that p_i is directly density-reachable from p_{i-1} within V with respect to Γ and κ .

Definition 3. A point $u \in V$ is a *core point* if its Γ -neighborhood $N_\Gamma(u) = \{v \in V \mid \text{dist}(u, v) \leq \Gamma\}$, where $\text{dist}(u, v)$ is the distance between u and v , contains at least κ neighbor points ($|N_\Gamma(u)| \geq \kappa$). Points which are in the Γ -neighborhood of a core point but have fewer than κ directly density reachable points in their own Γ -neighborhood are considered border points. Border points are density-connected.

Definition 4. Let $u, v \in V$ be two points. u is *density-connected* to v within V with respect to Γ and κ if and only if there is a point $m \in V$ such that u is density-reachable from m and v is density-reachable from m .

Appendix B.2 Arguments in Favor of DBSCAN

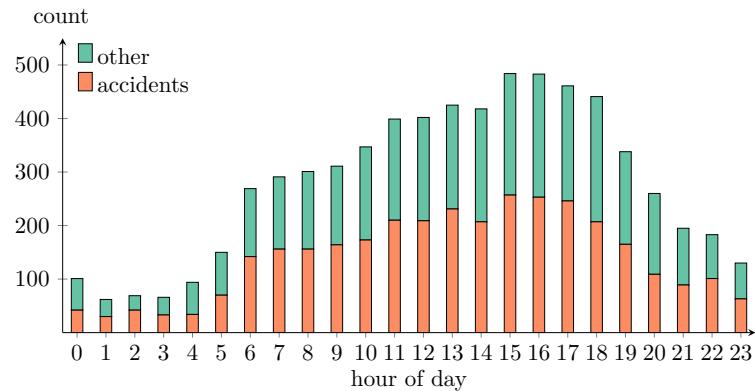
Our decision in favor of DBSCAN rests on three pillars. First, we have no a priori information on the number of clusters in our data set. DBSCAN does not require this kind of information. Alternative hierarchical or partitioning clustering algorithms such as the popular k -means clustering (Jain, 2010) critically hinge on knowledge of the correct number of clusters.

Second, the shape of a start or end cluster depends on whether a *driver* has a fixed parking space or has to rely on road-side parking. In the first case, the cluster is of approximately spherical shape. In the latter case, the shape is arbitrary and depends on the urban form and the geographical distribution of parking spaces. DBSCAN accommodates any shape, while k -means clustering only allows for spherical clusters (Firdaus and Uddin, 2015).

Third, DBSCAN is insensitive to outliers and will remove noise points, which are a common feature in geographic data. It discards all observations that are neither core nor border points. Other clustering algorithms, including k -means clustering, however, assign all points to clusters (Chandra and Anuradha, 2011), which very likely induces bias in the driver-OD identification and the fixed effects.

Appendix C Additional Results

Figure 14: *Traffic Relevant Incidences by Hour*



Notes: This Figure shows the number of traffic relevant incidences by category and hour. The category labeled “other” comprises incidences that block lanes or entire roads such as stalled vehicles, fallen trees, or breaks of water mains. INRIX provides the exact time and location of each incident. To avoid systematic reporting biases, INRIX supplements official police and road agency records with a rich selection of twitter feeds and road user reports. In addition, INRIX compares all incidences with their traffic flow data to filter out all incidences that do no alter traffic conditions. A fallen tree, for instance, is only considered when it blocks a lane. Altogether, this increases our instrument’s relevancy.

Table 7: Sensitivity tests for the first stage

outcome	coef	$R^2_{W \sim Z X}$	t-value	RV
log(incidents)	0.0351 (0.0008)	0.0012	44.6780	0.0329
if confounder was as strong as		$R^2_{Y \sim W Z,X}$	$R^2_{W \sim Z X}$	adjusted t
temperature	0.0008	0.0012	3.2166	
precipitation	0.0001	0.0000	1.9724	
sunshine duration	0.0001	0.0023	2.4566	
wind speed	0.0001	0.0008	2.3321	
visibility	0.0000	0.0002	2.0538	
holiday	0.0001	0.0073	2.7680	
school holidays (winter)	0.0003	0.0099	4.2333	
school holidays (spring)	0.0000	0.0083	2.1542	
school holidays (summer)	0.0009	0.0293	8.4603	
school holidays (autumn)	0.0003	0.0007	2.4948	

Notes: The upper section of this Table holds the estimate of our coefficient of interest, its corresponding standard error and t -value. Following Cinelli and Hazlett (2020) and Cinelli and Hazlett (2022), we also present $R^2_{W \sim Z|X}$, the partial R^2 of omitted variable W with our outcome Y controlling for variables in X . RV is the robustness value that indicates the share of the residual variation that a potential confounder needs to explain to render our coefficient estimate statistically insignificant.

The lower section of this Table presents $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$, a potential confounder's partial R^2 of outcome Y with omitted variable W controlling for X and Z . Only if both $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$ exceed robustness value RV in the upper half of this Table, can an unobserved confounder change our IV approach's conclusions. To facilitate interpretation, we also provide the reader with an adjusted, critical t -value. As long as the t -value associated with our outcome of interest from the upper part of the table exceeds this adjusted t test statistic, the potential confounder has insufficient strength to render it insignificant.

The sensitivity tests by Cinelli and Hazlett (2020) and Cinelli and Hazlett (2022) assess the effect of an omitted confounder under the assumption that it had the same properties as one of the included control variables. For instance, if an omitted confounder was as strong as the variable that controls for temperature and had the same $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$, then the adjusted, critical value would increase from $t = 1.96$ for a significance level of 0.05 to $t = 3.2166$. Overall, we find no case in which an omitted variable that mimicked any of our controls is sufficiently strong to alter our findings. Overall, these tests provide strong evidence in favor of unconfoundedness and a valid exclusion restriction.

Table 8: Sensitivity tests for the reduced form

outcome	coef	$R^2_{W \sim Z X}$	t-value	RV
log(incidents)	0.0173 (0.0011)	0.0002	16.2986	0.0112
if confounder was as strong as		$R^2_{Y \sim W Z,X}$	$R^2_{W \sim Z X}$	adjusted t
temperature	0.0008	0.0005	2.7491	
total precipitation	0.0001	0.0001	2.0710	
sunshine duration	0.0001	0.0000	1.9712	
wind speed	0.0001	0.0001	2.0785	
visibility	0.0000	0.0000	2.0079	
holiday	0.0001	0.0012	2.2933	
school holidays (winter)	0.0003	0.0013	2.7943	
school holidays (spring)	0.0000	0.0004	2.0032	
school holidays (summer)	0.0009	0.0016	3.5270	
school holidays (autumn)	0.0003	0.0001	2.1803	

Notes: The upper section of this Table holds the estimate of our coefficient of interest, its corresponding standard error and *t*-value. Following Cinelli and Hazlett (2020) and Cinelli and Hazlett (2022), we also present $R^2_{W \sim Z|X}$, the partial R^2 of omitted variable W with our outcome Y controlling for variables in X . RV is the robustness value that indicates the share of the residual variation that a potential confounder needs to explain to render our coefficient estimate statistically insignificant.

The lower section of this Table presents $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$, a potential confounder's partial R^2 of outcome Y with omitted variable W controlling for X and Z . Only if both $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$ exceed robustness value RV in the upper half of this Table, can an unobserved confounder change our IV approach's conclusions. To facilitate interpretation, we also provide the reader with an adjusted, critical *t*-value. As long as the *t*-value associated with our outcome of interest from the upper part of the table exceeds this adjusted *t* test statistic, the potential confounder has insufficient strength to render it insignificant.

The sensitivity tests by Cinelli and Hazlett (2020) and Cinelli and Hazlett (2022) assess the effect of an omitted confounder under the assumption that it had the same properties as one of the included control variables. For instance, if an omitted confounder was as strong as the variable that controls for temperature and had the same $R^2_{W \sim Z|X}$ and $R^2_{Y \sim W|X,Z}$, then the adjusted, critical value would increase from $t = 1.96$ for a significance level of 0.05 to $t = 3.2166$. Overall, we find no case in which an omitted variable that mimicked any of our controls is sufficiently strong to alter our findings. Overall, these tests provide strong evidence in favor of unconfoundedness and a valid exclusion restriction.

Table 9: First Stage Coefficients across Subsamples

No.	variable	category	coef	se	N
1	school holiday	1	0.001	0.007	70,784
2	end borough	8	0.003	0.004	67,878
3	temperature	3	0.003	0.001	427,126
4	school holiday	2	0.005	0.008	26,444
5	start borough	8	0.005	0.004	67,533
6	school holiday	4	0.014	0.006	53,297
7	sunshine duration	4	0.016	0.002	273,014
8	public holiday	1	0.016	0.020	17,068
9	visibility	4	0.017	0.002	370,758
10	start borough	3	0.019	0.003	147,878
11	sunshine duration	3	0.021	0.002	269,296
12	temperature	4	0.021	0.001	427,564
13	temperature	2	0.022	0.001	429,058
14	end borough	3	0.022	0.003	140,822
15	sunshine duration	2	0.023	0.002	269,149
16	end borough	2	0.023	0.003	80,222
17	average age	1	0.025	0.002	319,987
18	end borough	999	0.026	0.003	164,089
19	start borough	2	0.026	0.003	74,318
20	foreigners	4	0.027	0.001	316,349
21	wind speed	2	0.028	0.001	427,236
22	wind speed	3	0.028	0.002	425,984
23	trip distance	4	0.030	0.001	427,612
24	start borough	6	0.030	0.002	206,559
25	end borough	1	0.031	0.002	201,231
26	wind speed	1	0.031	0.001	432,368
27	end borough	12	0.032	0.003	118,665
28	start borough	999	0.032	0.002	203,438
29	start borough	1	0.032	0.002	166,042
30	population density	4	0.032	0.001	317,292
31	precipitation	4	0.032	0.003	91,202
32	end borough	6	0.032	0.002	191,102
33	start borough	12	0.033	0.003	117,831
34	school holiday	3	0.033	0.002	227,783
35	household size	3	0.033	0.002	303,011
36	precipitation	1	0.033	0.001	1,438,033
37	start borough	9	0.033	0.005	107,288
38	trip distance	3	0.033	0.001	428,071
39	household size	2	0.034	0.002	301,781
40	start borough	7	0.034	0.002	137,310

Notes: The first column indicates the number of the subsample for ease of reference. Together, columns two and three indicate the variable and the category with respect to which we created the subsample. The final three columns indicate the coefficient (coef), the standard error (se), and the number of a subsample's observations N . The categories of all variables are numbered consecutively from lowest to highest. The only exception is for the start and end boroughs where the thirteenth category is labeled "999" for all boroughs outside of Berlin. We split all variables measuring temperature into four categories. We categorized school holidays by season.

First Stage Coefficients across Subsamples (cont.)

No.	variable	category	coef	se	N
41	foreigners	2	0.034	0.002	309,860
42	visibility	1	0.035	0.001	476,051
43	trip distance	2	0.035	0.002	428,339
44	population density	1	0.035	0.002	317,868
45	end borough	10	0.035	0.007	105,429
46	end borough	7	0.036	0.002	140,105
47	start borough	4	0.036	0.001	208,757
48	start borough	10	0.036	0.006	107,760
49	household size	1	0.036	0.002	351,052
50	public holiday	0	0.036	0.001	1,695,070
51	population density	2	0.036	0.002	322,542
52	end borough	4	0.037	0.001	228,997
53	foreigners	3	0.037	0.002	318,695
54	school holiday	0	0.037	0.001	1,333,830
55	end borough	9	0.037	0.004	103,318
56	visibility	3	0.038	0.002	442,174
57	household size	4	0.038	0.002	315,583
58	average age	4	0.038	0.002	305,365
59	population density	3	0.038	0.002	313,725
60	average age	2	0.039	0.002	319,779
61	visibility	2	0.042	0.002	423,155
62	sunshine duration	1	0.042	0.001	900,679
63	average age	3	0.042	0.002	326,296
64	precipitation	2	0.044	0.004	93,116
65	trip distance	1	0.045	0.003	428,116
66	start borough	5	0.046	0.004	89,546
67	foreigners	1	0.047	0.003	326,523
68	precipitation	3	0.047	0.003	89,787
69	end borough	5	0.049	0.004	92,499
70	wind speed	4	0.054	0.002	426,550
71	temperature	1	0.066	0.002	428,390
72	end borough	11	0.083	0.007	77,781

Notes: The first column indicates the number of the subsample for ease of reference. Together, columns two and three indicate the variable and the category with respect to which we created the subsample. The final three columns indicate the coefficient (coef), the standard error (se), and the number of observations in the subsample *N*. The categories of all variables are numbered consecutively from lowest to highest. The only exception is for the start and end boroughs where the thirteenth category is labeled “999” for all boroughs outside of Berlin. We split all temperature or census related variables into four categories. We categorized school holidays by season.

Table 10: MECC in € cents per vehicle km by hour bracket

time bracket	MECC
04:00 - 05:59	3.6
06:00 - 07:59	43.0
08:00 - 09:59	70.3
10:00 - 11:59	8.4
12:00 - 13:59	30.8
14:00 - 15:59	70.9
16:00 - 17:59	87.4
18:00 - 19:59	43.2
20:00 - 21:59	3.0

Notes: This table shows the marginal external costs of congestion (MECC) in € cents per km based on the outcome of our main specification.

Table 11: Optimal Congestion Charges conditional on Demand Elasticities in € cents per km

time bracket	$\eta = 0.5$	$\eta = 1.0$	$\eta = 1.5$
uniform	31.2	27.9	26.0
peak	36.6	32.0	29.6
off-peak	13.3	12.8	12.5
04:00 - 05:59	3.6	3.5	3.5
06:00 - 07:59	31.4	27.3	25.2
08:00 - 09:59	44.2	37.3	34.1
10:00 - 11:59	8.3	8.1	8.1
12:00 - 13:59	26.7	24.6	23.4
14:00 - 15:59	46.8	39.8	36.5
16:00 - 17:59	50.7	42.3	38.6
18:00 - 19:59	33.4	29.5	27.5
20:00 - 21:59	3.0	3.0	3.0

Notes: This table shows the optimal congestion charges in € cents per km based on the outcome of our main specification and different values for the elasticity of demand η .

Appendix D Derivation of Welfare Gains and Optimal Road Charges

We derive how to calculate welfare gains and optimal road charges. To facilitate understanding and without loss of generality, we omit subscript T that indicates the time-period of the day. In the absence of road pricing, the equilibrium flow V_{eq} occurs at the intersection of demand and average cost curve:

$$\begin{aligned}
 \Omega \cdot V_{eq}^{\frac{\sigma}{1-\sigma}} &= \Phi^{\frac{1}{\eta}} \cdot V_{eq}^{-\frac{1}{\eta}} \\
 V_{eq}^{\frac{\sigma}{1-\sigma}} &= \Phi^{\frac{1}{\eta}} \cdot V_{eq}^{-\frac{1}{\eta}} \cdot \Omega^{-\frac{1}{1-\sigma}} \\
 V_{eq}^{\frac{\sigma}{1-\sigma} + \frac{1}{\eta}} &= \Phi^{\frac{1}{\eta}} \cdot \Omega^{-\frac{1}{1-\sigma}} \\
 V_{eq}^{\frac{1-\sigma(1-\eta)}{(1-\sigma)\eta}} &= \Phi^{\frac{1}{\eta}} \cdot \Omega^{-\frac{1}{1-\sigma}} \\
 V_{eq} &= \left(\Phi^{1-\sigma} \cdot \Omega^{-\eta} \right)^{\frac{1}{1-\sigma(1+\eta)}}
 \end{aligned} \tag{8}$$

The optimum, however, occurs at the intersection of the marginal cost and the demand curve

$$\begin{aligned}
 \frac{1}{1-\sigma} \cdot \Omega^{\frac{1}{1-\sigma}} \cdot V_{opt}^{\frac{\sigma}{1-\sigma}} &= \Phi^{1\eta} \cdot V_{opt}^{-\frac{1}{\eta}} \\
 \Omega^{\frac{1}{1-\sigma}} \cdot V_{opt}^{\frac{\sigma}{1-\sigma}} &= (1-\sigma) \cdot \Phi^{\frac{1}{\eta}} \cdot V_{opt}^{-\frac{1}{\eta}} \\
 V_{opt}^{\frac{\sigma}{1-\sigma}} \cdot V_{opt}^{\frac{1}{\eta}} &= (1-\sigma) \cdot \Phi^{\frac{1}{\eta}} \cdot \Omega^{-\frac{1}{1-\sigma}} \\
 V_{opt}^{\frac{\sigma\eta+1-\sigma}{(1-\sigma)\eta}} &= (1-\sigma) \cdot \Phi^{\frac{1}{\eta}} \cdot \Omega^{-\frac{1}{1-\sigma}} \\
 V_{opt} &= (1-\sigma)^{\frac{\eta \cdot (1-\sigma)}{1-\sigma \cdot (1-\eta)}} \cdot V^{eq}
 \end{aligned} \tag{9}$$

The deadweight loss DWL from excess congestion is the area below the marginal cost curve and above the average cost curve between the lower bound V_{opt} and the upper bound V_{eq} as shown in Figure 9:

$$\begin{aligned}
DWL &= \int_{V_{opt}}^{V_{eq}} (MC^S(V) - C^D(V))dV \\
&= \left[\Omega^{\frac{1}{1-\sigma}} \cdot V^{\frac{1}{1-\sigma}} - \frac{\eta}{1-\eta} \cdot \Phi^{\frac{1}{\eta}} \cdot V^{\frac{\eta-1}{\eta}} \right]_{V_{opt}}^{V_{eq}} \\
&= \Omega^{\frac{1}{1-\sigma}} \cdot (V_{eq} - V_{opt})^{\frac{1}{1-\sigma}} - \frac{\eta}{\eta-1} \cdot \Phi^{\frac{1}{\eta}} \cdot (V_{eq} - V_{opt})^{\frac{\eta-1}{\eta}} \\
&= \Omega^{\frac{1}{1-\sigma}} \cdot \left(V_{eq} - (1-\sigma)^{\frac{(1-\sigma)\eta}{1-\sigma(1-\eta)}} \cdot V_{eq} \right)^{\frac{1}{1-\sigma}} - \frac{\eta}{\eta-1} \cdot \Phi^{\frac{1}{\eta}} \cdot \left(V_{eq} - (1-\sigma)^{\frac{(1-\sigma)\eta}{1-\sigma(1-\eta)}} \cdot V_{eq} \right) \\
&= \Omega^{\frac{1}{1-\sigma}} \cdot \left(1 - (1-\sigma)^{\frac{\eta}{1-\sigma(1-\eta)}} \right) \cdot V_{eq}^{\frac{1}{1-\sigma}} - \frac{\eta}{\eta-1} \cdot \Phi^{\frac{1}{\eta}} \cdot \left(1 - (1-\sigma)^{\frac{(1-\sigma)(\eta-1)}{1-\sigma(1-\eta)}} \right) \cdot V_{eq}^{\frac{\eta-1}{\eta}}
\end{aligned} \tag{10}$$

Substituting $\Omega^{\frac{1}{1-\sigma}} = C(V_{eq}) \cdot V_{eq}^{-\frac{\sigma}{1-\sigma}}$, $\Phi^{\frac{1}{\eta}} = C(V_{eq}) \cdot V_{eq}^{\frac{1}{\eta}}$, and dividing by $C(V_{eq}) \cdot V_{eq}$ returns

$$\Delta = 1 - (1-\sigma)^{\frac{\eta}{1-\sigma(1-\eta)}} - \frac{\eta}{\eta-1} \cdot \left(1 - (1-\sigma)^{\frac{(1-\sigma)(\eta-1)}{1-\sigma(1-\eta)}} \right), \tag{11}$$

the deadweight loss as a share of total travel time.

To internalize the cost of congestion and to realize the social optimum, a congestion charge of τ that is equal to the difference between marginal cost and average costs of travel has to be charged:

$$\begin{aligned}
\tau &= MC(V_{opt}) - C(V_{opt}) \\
&= \frac{1}{1-\sigma} \cdot C_{opt} - C_{opt} \\
&= \frac{\sigma}{1-\sigma} \cdot C_{opt} \\
&= \sigma \cdot C_{eq} \cdot (1-\sigma)^{\left(\frac{\eta\sigma}{1-\sigma(1-\eta)}\right)-1}
\end{aligned} \tag{12}$$

Appendix E Quickbundles

Quickbundles (Garyfallidis et al., 2012) minimizes the average direct-flip distance (MDF). It divides a trip into K equidistant points s_i , so that a trip s is an ordered polyline $s = (s_1, s_2, s_3, \dots, s_K)$. $s' = (s_K, \dots, s_3, s_2, s_1)$ is the flipped version of s with a reversed order of points. The direct distance between two trips s and t is

$$\begin{aligned} d_{\text{direct}}(s, t) &= d(s, t) \\ &= \frac{1}{K} \sum_{i=1}^K |s_i - t_i| , \end{aligned} \tag{13}$$

where $|s_i - t_i|$ is the Euclidian distance between points s_i and t_i . The flipped distance is

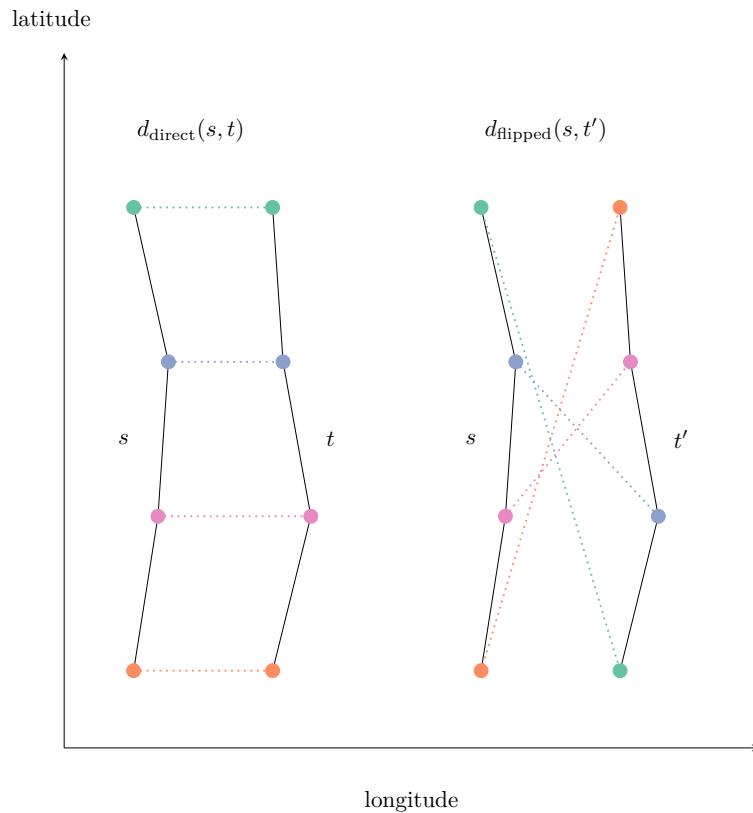
$$d_{\text{flipped}}(s, t) = d(s', t) = d(s, t') . \tag{14}$$

The minimum direct-flipped distance (MFD) is

$$MFD(s, t) = \min [d_{\text{direct}}(s, t), d_{\text{flipped}}(s, t)] . \tag{15}$$

Figures 15 illustrates how Quickbundles works. In the left panel, routes s and t share the same starting point and the color coded points indicate the same order. In the right panel, route t' is reversed so that the order of the equidistant points is reversed, too. Quickbundles solves the issue of route direction by calculating two distance measures and by choosing the minimum distance. The distances calculated are indicated by dashed lines. The first measure calculates the total Euclidean distance between the points in the order they occur. The second measure reverses the order of points of one route before recalculating the total distance (right panel). The total distance between points of similarly oriented routes is smaller than the total difference between reversed but otherwise identical routes. Altogether, we consider that trips share the same route if the difference between their trajectories does not exceed 10% of their trip length. We use 100 equidistant points.

Figure 15: Illustration of the QuickBundles algorithm



Notes: This figure illustrates how QuickBundles works. First, trips s and t are divided into K equidistant points. The colors indicate the order in which the points occur in the data. In the left panel, routes s and t share the same starting points and have the same order of points. In the right panel, trip t' is reversed. The distances calculated are indicated by dashed lines.