# The City-Wide Effects of Tolling Downtown Drivers: Evidence from London's Congestion Charge.

Ian Herzog*

January 2022

Click here for latest version

**Abstract**

I study the effects of London England's Congestion Charge on regional traffic, commuting, and economic activity's spatial distribution. In 2003, London began tolling drivers to enter its central business district on workdays and this policy immediately reduced downtown traffic. I measure exposure to tolled traffic throughout London's regional road network and show that the policy also reduced rush-hour traffic on untolled roads leading downtown. I then use London's Congestion Charge as a natural experiment to identify ensuing effects of traffic on commuting by car and public transit and find that increasing a route's traffic decreases both the number of commuters and their driving rates at the margin. In a quantitative spatial model, traffic's effect on the number of commuters suggests that changing traffic patterns make certain neighbourhoods more desirable places to live and work, influencing the congestion charge's equilibrium effects, so I use a model with endogenous traffic externalities and mode choice for policy analysis. Simulations suggest that London's Congestion Charge causes untolled commuters to substitute from public transit to driving and gives the region's commuters positive net benefits that disproportionately accrue to low-skill workers.
JEL codes: H23, L91, R13, R40, R48

*Department of Economics, Huron University College, 1349 Western Rd, London, Ontario, Canada, N6G 1H3. E-mail iherzog2@uwo.ca. Thanks to Nathaniel Baum-Snow, Jonathan Hall, and Stephan Heblich for their guidance and support. I also thank Rosa Sanchis-Guarner for invaluable assistance and seminar participants at the 2020 Virtual Meeting of the Urban Economics Association, the 2020 UEA PhD workshop, and the University of Toronto for comments and suggestions. This research is supported in part by funding from the Social Sciences and Humanities Research Council of Canada.

# 1 Introduction

Urban road congestion costs billions in deadweight loss each year (Couture et al., 2018) but perfectly internalizing these externalities requires unwieldy road pricing schemes that vary smoothly across time and space (Zhang and Kockelman, 2016; Hall, 2020). London England was among the first of a growing list of cities turning to cordon fees, tolling entry to physically small but economically important areas, but these policies can inadvertently tax productive employment clusters (Zhang and Kockelman, 2016; Brinkman, 2016) and their distributional effects are unclear (Ecola and Light, 2009). This paper examines London's Congestion Charge using a new data set, empirical strategy, and spatial equilibrium model to provide the first empirical analysis of a cordon fee's effect on regional patterns of traffic, commuting, and labour market outcomes.

My analysis builds on the argument that Central London's Congestion Charge not only reduces traffic in the tolled downtown area but also affects radial roads used to get downtown, influencing commute costs and employment patterns throughout the region. I track these effects with a new measure of Congestion Charge Exposure that uses routing software to identify roads frequently used en-route to central neighbourhoods. I combine this measure with panel data on regional traffic and commute flows to identify the policy's effect on traffic and ensuing effects of traffic on commuters' residences, workplaces, and public transit use. These patterns then calibrate a structural model that accounts for endogenous wages, rents, and traffic externalities between overlapping commutes for welfare analysis.

I find that by discouraging driving on tolled routes, London's Congestion Charge reallocates roadspace to cross-town commuters, impacts decisions between driving and taking public transit, and affects employment throughout the city. Taking a regional view, the policy gives commuters revenue-neutral gains of £140 million per year, an individually small

but diffuse benefit that aggregates to exceed the program's annual costs. I also find that these gains were progressive because tolled workplaces are particularly important for high-skill professionals while the low-skill often depend on suburban commutes that are less well served by London's hub-and-spoke public transit system.

I build to these conclusions with new data linking decennial census commute flows to London's dense traffic monitor network. I combine traffic and commuting data with the OpenStreetMap project's Open Source Routing Machine to compute optimal routes through the city and attribute traffic counts to particular commutes. Routing software also reveal where commutes intersect, allowing me to identify how tolled traffic spills over onto untolled commutes. The result is the first data set of long-run changes in commuting and traffic patterns within a city that is suitable for estimating traffic's effect on commuting and employment.

I use these data to develop an Index of Congestion Charge Exposure measuring how Central London's toll affects traffic throughout the city. This index breaks Greater London's road network into links and measures pre-toll traffic that would be tolled after the policy begins. Congestion Charge Exposure's spatial distribution reveals that radial roads are most affected because they are frequently used by initially high-traffic routes to Central London. I then show that Congestion Charge Exposure predicts relative traffic declines after tolling began and argue that this relationship is causal, establishing that Central London's Congestion Charge affected traffic on untolled commutes throughout the region.

I then develop a model of commuters' travel demand and estimate central parameters. The model features heterogeneous commuters solving a nested discrete choice problem; first choosing where to live and work based on endogenous wages, rents, and transportation opportunities, then choosing whether to commute by car or public transit between given locations. Travel demand aggregates across utility-maximizing commuters who pay tolls

to drive on certain commutes and incur a utility cost from driving in traffic. Since traffic is costly, increasing it makes commuters reduce driving rates or choose different residences and workplaces to the extent that they can substitute to public transit or sort towards locations linked by low-traffic routes.

Commuting's traffic elasticities are critical inputs to policy analysis and speak to the necessity of studying London's Congestion Charge in a general equilibrium framework. However, correlations between traffic growth and commute flows are uninformative since they conflate supply and demand. With this in mind, I identify reduced-form effects and calibrate the structural model using Congestion Charge Exposure to isolate exogenous variation in the way Central London's toll affects traffic for untolled commuters. Reduced-form estimates show that exogenously reducing traffic between a residence-workplace pair causes commuters to sort towards those locations and increases their driving rates. The effect on driving rates is large and the significant commuting elasticity suggests that traffic affects residence and workplace locations, motivating structural analysis to account for general equilibrium effects.

To quantify the welfare effects of London's Congestion Charge, I specify the structure of labour demand, housing supply, and traffic externalities and run counterfactual policy simulations. Commuting decisions determine neighbourhoods' labour supply while labour demand reflects agglomeration externalities as in Ahlfeldt et al. (2015) and differences between high- and low-skill labour as in Tsivanidis (2019) and Lee (2019). The model also assumes commuters impose traffic on the roads comprising their routes to work, generating endogenous traffic externalities along a realistic road network.

I estimate the model's central parameters following reduced-form analysis and calibrate location and road characteristics to match Greater London's commuting and traffic patterns. The model also provides a framework for integrating census counts of high- and low-skill

population and employment with commute flows that do not separate by worker skill. Given commute costs, the model implies unique workplace and residential fixed effects rationalizing each skill group's spatial distribution of population and employment. I use this relationship to impute skill-specific commute flows and run counterfactual simulations that identify unequal welfare effects.

Simulations show that London's Congestion Charge increases mean utility by 0.11 percent and disproportionately benefits low-skill commuters. Aggregate gains come alongside a 0.091 percent reduction in welfare inequality as high- and low-skill workers both benefit from traffic reductions but tolled workplaces are more important for high-skill commuters. At the same time, the policy shifts employment immediately outside the Congestion Charge Zone, suburbanizes population, and increases commuters' driving rates by 0.531 percent. All told, I find that London's Congestion Charge reduced the region's road traffic by 1.007 percent and that this effect would be larger without endogenous sorting and substitution away from public transit.

This paper contributes to a literature on the costs of urban road congestion and benefits of road pricing. Previous work considers tolls to alleviate congestion when drivers respond by rescheduling their trips (Vickrey, 1967; Arnott et al., 1993; Parry, 2009; Hall, 2020), analyses travel surveys (Couture et al., 2018; Akbar and Duranton, 2017), exploits temporary shocks (Gibson and Carnovale, 2015; Hanna et al., 2017), or experimentally simulates congestion pricing policies (Martin and Thornton, 2017; Kreindler, 2018). Despite theoretical interest in a cordon fee's effects on labour and housing market equilibria (Verhoef, 2005; Brinkman, 2016; Zhang and Kockelman, 2016) empirical analysis of this issue is scant (Proost and Thisse, 2019). This paper offers the first estimates of traffic on commuters' location and mode choices and embeds them in a spatial equilibrium model to show that cordon fees can create progressive welfare gains on a regional scale.

I also contribute to a literature using quantitative general equilibrium models and commuting data to study transportation policies (Tsivanidis, 2019; Tyndall, 2018; Severen, 2019; Heblich et al., 2020). This work builds on the literature estimating the effects of urban transportation on economic activity (Baum-Snow et al., 2005; Baum-Snow, 2007; Duranton and Turner, 2011; Gonzalez-Navarro and Turner, 2018; Jerch et al., 2020) using discrete choice models adapted from international trade (Eaton and Kortum, 2002; Ahlfeldt et al., 2015; Monte et al., 2018). My work also relates to Fajgelbaum and Schaal (2017) and Allen and Arkolakis (2020) who consider endogenous road congestion in spatial equilibrium models. I adapt this framework to estimate road pricing's regional effects and provide a new way to incorporate endogenous traffic externalities using a realistic road network.

Finally, I measure an important and previously unexamined source of welfare gains from London's canonical policy experiment. Leape (2006) summarises early analysis of London's Congestion Charge and more recent analysis finds effects on downtown traffic accidents (Green et al., 2016) and air pollution (Green et al., 2018). Tang (2018) summarises downtown amenity improvements by showing that the policy increased home prices in the tolled area. Despite these benefits, Zhang and Kockelman (2016) and Brinkman (2016) argue that cordon fees can have negative welfare effects if they break productive employment clusters. My analysis of London's Congestion Charge finds that sorting attenuates the policy's value but agglomeration externalities are not the culprit. Cordon fees are limited because they create incentives to drive on untolled commutes that share roadspace with tolled traffic. I also find that this limitation comes with the advantage of disproportionately benefiting low-skill commuters who rely on driving to untolled suburban workplaces.

This paper proceeds by discussing the data in section 2. Section 3 then describes London's economic geography and congestion charge policy. Section 4 introduces an Index of Congestion Charge Exposure and establishes that tolling downtown affected traffic through-

out Greater London. Section 5 fixes ideas in a model of commuters' travel demand and empirically validates its mechanisms. Section 6 embeds this demand system in a structural model that I calibrate in section 7 and use to run counterfactual simulations in section 8. Finally, section 9 offers concluding remarks.

# 2 Data

I use data from three main sources. Commute flows, employment, and population come from the United Kingdom's 2001 and 2011 decennial censuses. Traffic data come from the UK Department for Transport's (DfT) network of traffic count points. I spatially join commute and traffic data using routes computed by the OpenStreetMap project's Open Source Routing Machine (OSRM). Appendix A.1.1 describes additional data which include boundary shapefiles and public transit information from Transportation for London and the Greater London Authority's open data repositories.

I restrict my study area to the Greater London Region and develop a panel of outcomes in 2001 and 2011 for middle super output areas (MSOAs) which are census geographies designed to contain 2,000 to 6,000 resident households. Normalizing to constant geographic boundaries delivers a balanced panel of 967 MSOAs.

## 2.1 Commute flows

Commute flows between MSOA pairs come from the 2001 and 2011 UK census interaction data sets. I use these data to measure the number of workers commuting between MSOA pairs by car and other means including taxis, public transit, or walking. My sample is a balanced panel of 934,122 directional MSOA pairs, excluding commuters living and working

in the same MSOA.[1] I observe each pair's car and non-car commuting in both 2001 and 2011 for a total of 3,736,488 commute flow counts.[2]

## 2.2 Employment and population

The structural model uses data describing the spatial distribution of skills to account for unequal outcomes across workers. To this end, I define each area's residential population as the number of employed individuals living there who do not work from home and employment is the same universe measured at usual workplaces reported to the census. Unlike commute flows, population and employment counts are skill-specific. The final data describe employment and population counts by two skill groups (high and low) and two modes of transportation (car and non-car) for all MSOAs in Greater London.

I define high- and low-skill workers by aggregating the occupation-based three-class UK National Statistics Socio-economic Classification (NS-SeC) into two groups that approximate college and non-college workers. High-skill occupations include managers, professionals, small business owners, intermediate occupations (e.g. nurse, salesperson), and technical occupations (e.g. plumber, chef), which all require some form of post-secondary education or training. Low-skill occupations include routine and semi-routine (e.g. receptionists, care workers) occupations.[3]

[1] Excluding own-MSOA commuting flows avoids complications of working from home and measuring commute costs for travel within a neighbourhood. I also exclude counts of individuals who report no fixed workplace address.

[2] The UK census changed their method for anonymizing small cells between 2001 and 2011. The data appendix describes my procedure for adjusting these data so that they are comparable over time.

[3] High skill is NS-SeC 1 through 5 and low skill is NS-SeC 6 and 7. Unfortunately, I do not separately observe intermediate and technical occupations.

## 2.3 Traffic

Traffic data come from the UK department for transport (DfT) who count traffic flows at roughly 10,000 count points (CPs) across England's major road network. Trained enumerators count hourly traffic from 07:00 to 19:00 on a non-winter weekday annually or on two, four, or eight year cycles.[4] My raw data cover all years from 2000 to 2015 and are an unbalanced panel of 1,200 to 2,200 counts in Greater London depending on the year.[5] To control for composition effects caused by high-traffic CPs being counted most frequently, I restrict my sample to those observed at least once from 2000 to 2002 (pre-toll) and again from 2003 to 2015 (post-toll). This gives a balanced panel of 1,399 count points whose spatial distribution is shown in figure 1 which confirms their comprehensive coverage of Greater London's road network.

I define a road link as a travel direction-count point pair and observe each link's hourly traffic volume each time it is counted. I also observe link length, end point locations, and enumerators' coordinates. I aggregate across hours to calculate morning and evening rush hour traffic each time a link is counted. To this end, I define morning rush-hour traffic as total traffic volume counted between 07:00 and 10:00 and evening rush-hour traffic as volume counted from 16:00 to 19:00. Finally, I average across years from 2000 to 2002 and again from 2003 to 2015 to compute pre- and post-toll traffic volumes during each rush hour.

## 2.4 Routes

I use optimal routes between MSOAs to associate traffic on road links with particular commute through the city. I compute time-minimizing paths between all MSOA pairs in

[4] Counts are restricted to weekdays between March and October excluding public and school holidays.

[5] I observe a total of unique 2,991 count points located in the London Region and restrict to manually counted observations.

Greater London using the OpenStreetMap project's Open Source Routing Machine version 5.22.0 (Luxen and Vetter, 2011). I run the router 934,122 times through a fall 2019 snapshot of the London Region's road network to compute fastest paths between between centroids of all directed MSOA pairs.[6] I observe a detailed polyline representing the route's path through London's roads.

I define a commute's route in each rush hour as the set of count points within 50 metres of its path. For each directed MSOA pair, the algorithm builds a 100 metre buffer centred along its path and intersects that buffer with traffic enumerator coordinates. Road links with enumerators located within the buffer are matched with the MSOA pair. The algorithm assigns a travel direction to each link-path intersection by finding the path's closest vertex to each matched link and noting the dominant zonal or meridional direction travelled one vertex towards the destination.[7] A commute's morning route $R_{ij}^{morn}$ is the set of road links matched to the path from $i$ to $j$ and the evening route $R_{ij}^{eve}$ is the set used for return travel.[8]

Appendix figure A.1 presents an example of the spatial join. The bottom right corner is a residence MSOA centroid and the top left corner is a workplace MSOA centroid. Black dots indicate count points and highlight those matched to one direction of travel between this MSOA pair. Matching the 934,122 routes to the balanced panel of road links preserves 922,550 MSOA pairs with traffic measured both pre- and post-toll.

I aggregate across road links to compute rush hour traffic volume for each residence-workplace pair in the census data. An MSOA pair's morning traffic is the sum of morning rush-hour traffic volume on road links comprising the workplace-bound route and evening

[6] The road data come from the Geofabrik online repository.

[7] If a CP's closest vertex is the route's destination, I take travel direction from the second last vertex to the destination.

[8] Morning and evening a commute's workplace- and residence-bound paths can differ due to street elements such as one-way roads and turn restrictions but it is always the case that $R_{ij}^{morn} = R_{ji}^{eve}$.

traffic is the sum of evening traffic volume on the residence-bound route. Formally, those living in $i$ and working in $j$ in period $t$ face morning traffic $\sum_\ell 1\{\ell \in R_{ij}^{morn}\} \sum_{h=7}^{9} traffic_{h\ell t}$ and evening traffic $\sum_\ell 1\{\ell \in R_{ij}^{eve}\} \sum_{h=16}^{18} traffic_{h\ell t}$ where $traffic_{\ell t}^{h}$ is the traffic volume counted on link $\ell$ in hour $h$, each road link $\ell$ is a directed count point, and period $t$ is either pre- or post-toll. I then measure traffic faced by commuters living in $i$ and working in $j$ in period $t$ as the average of their morning and evening rush-hour traffic.

Section 4 describes a related procedure for computing commuters' Congestion Charge Exposure and I control outliers by dropping MSOA pairs with traffic growth or charge exposure in the top or bottom 2.5 percent of the distribution. Appendix table A.1 summarises pairwise traffic and commuting data in the main analysis sample.

# 3 Background

## 3.1 London's economic geography

The Greater London Region extends over 25 kilometres from the city centre to cover 33 local authorities and is surrounded by a greenbelt that has restricted urban expansion since the 1960s. Central London has had an active congestion charge since 2003 which section 3.2 describes in detail. Figure 1 maps Greater London and its major arterial roads; line widths correspond to speed limits and dots are count points used to measure traffic. The purple area is Outer London, which contains most of the region's high-capacity motorways, and Inner London is subdivided into the yellow Congestion Charge Zone (CCZ) and green adjacent area.

Table 1 shows that one third of London commuters drove to work in 2001 and this share fell to one quarter in 2011. Driving is more popular among those who live and work in Outer London, but even in the suburbs driving's share of commutes has fallen from about

two thirds in 2001 to just under half in 2011. Table 2 shows that high-skill employment is evenly split between Outer London, Inner London, and the CCZ while more than half of low-skill employment is in Outer London, where commuters are never tolled and are most likely to drive in suburban traffic.[9]

### 3.2 London's Congestion Charge

Road pricing proposals in the UK date back to the Smeed commission's 1964 report to the Department for Transport who were studying economic and technical possibilities for future policies (Smeed, 1964). The Smeed report was ahead of its time but by the late 1990s political will and technological advances made congestion charging a serious consideration in Central London. Enabling legislation came in 1999 when the Greater London Authority Act created the region's metropolitan government and gave future mayors power to enact road user charging (Greater London Authority, 1999). Ken Livingstone was elected London's first mayor in 2000 and his campaign promises included downtown road pricing. Transportation for London (TfL) subsequently began studying policy options and began charging vehicles to enter Central London in February of 2003.

Central London's Congestion Charge started at £5, grew over time, and was levied on motor vehicles entering the CCZ between 07:00 and 18:30 on weekdays using automatic license plate recognition. The CCZ comprises approximately 20 $km^2$ within London's Inner Ring Road and contains most of the city's major attractions and central business district. The Inner Ring Road itself remains untolled and used by through-traffic to avoid the CCZ. Congestion charging was temporarily expanded westward in 2007 to include Chelsea and Kensington, but the Western Extension Zone (WEZ) charge was unpopular with residents and abandoned in 2010.[10]

[9] The spatial distributions of high- and low-skill employment are similar in 2001.

[10] Appendix figure A.2 focuses on a subset of Inner London to detail the CCZ and WEZ boundaries and

Transportation for London (2003) documented a 27 percent drop in motor vehicles entering Central London after the toll's introduction and Leape (2006) asserts that most of this is substitution to public transit and detouring around the tolled zone. All told, Tang (2018) estimates a causal effect on Central London traffic closer to a nine percent drop in traffic that capitalized into a three percent rise in home values. In a follow-up survey, Outer London residents reported that they reduced driving to the CCZ for a range of activities after its introduction (Transportation for London, 2005).

Figure 1 shows that the CCZ is a small but central piece of the London Region and table 2 shows that it contains thirty-three percent of the region's high-skill employment but just fifteen percent of low-skill employment. Table 2 also shows that CCZ workplaces are destinations for just 6 percent of the region's car commuters. Combined, the CCZ's skill bias and good public transit access suggest that commuters taxed by London's toll are overwhelmingly working high-skill jobs.

# 4 Congestion Charge Exposure

This section develops a continuous measure of Congestion Charge Exposure to track the ripple effects of Central London's toll on suburban traffic and establish that this cordon fee affected regional traffic patterns. I then argue that Congestion Charge Exposure captures exogenous variation in traffic caused by the policy and provides a basis for estimating ensuing effects of traffic on commuting to ground counterfactual simulations.

contents. The 2007 WEZ experiment also shortened the charging day to end at 18:00 where it remained throughout my study period. The policy was eventually extended to end at 22:00 each day, run seven days per week, and charge private vehicles for hire including minicabs and Uber after my study period. Appendix figure A.3 presents a timeline of relevant policy developments during my study period. Throughout my analysis, I call a trip tolled if it ends in the CCZ and occurs after 2003.

### 4.1 Measurement

Congestion Charge Exposure measures the pre-2003 traffic crossing each road link that would be tolled in the post-policy period. Congestion Charge Exposure on road link $\ell$ in rush hour $h \in \{morn, eve\}$ is

$$CCE_{\ell}^{h} = \sum_{r,s} CCZ_{rs} 1\{\ell \in R_{rs}^{h}\} \frac{traffic_{rs0}^{h}}{length_{rs}^{h}} \tag{1}$$

where $1\{\cdot\}$ is the indicator function, $R_{ij}^{h}$ is the set of road links connecting residence $i$ and workplace $j$ at rush hour $h$, $CCZ_{ij}$ indicates tolled commutes, $traffic_{ij0}^{h}$ is pre-toll traffic associated with commuting between residence $i$ and workplace $j$, and $length_{ij}^{h}$ is length of monitored roads on that commute.

Figure 2 demonstrates the intuition behind CCE with an illustrative example. The left panel shows an Inner London road link crossed by tolled routes shown as blue paths and origin MSOAs are shaded in proportion to the amount of morning traffic contributed to this link's morning charge exposure.[11] Origins are all east of the link and fan outward as the road network leads drivers to this link en-route to the CCZ.

Figure 3 maps Congestion Charge Exposure's spatial distribution. Lines are road segments monitored by each count point with thickness and darkness corresponding to link-level CCE values averaged across directions and rush hours. The map shows that CCE is largest on radial roads running east-west across the city. In addition, several southern road links headed towards the central business district are very exposed to Central London's toll. Meanwhile, London's North Circular Road does not stand out as it is not exposed to substantial amounts of tolled traffic.

From a commuter's perspective, Congestion Charge Exposure aggregates across road

[11] Pre-toll morning traffic contributed to link $\ell$ by origin $i$ is $\sum_j CCZ_{ij} 1\{\ell \in R_{ij}^{morn}\} \frac{traffic_{ij0}^{morn}}{length_{ij}}$.

links in the same fashion as traffic. Specifically, Congestion Charge Exposure of those living in $i$ and working in $j$ is

$$CCE_{ij} = \frac{1}{2} \sum_{h} \sum_{\ell} 1\{\ell \in R_{ij}^{h}\} CCE_{\ell}^{h} \tag{2}$$

which sums exposure across roads en-route and averages across the morning trip to work and evening trip home. An MSOA pair's CCE summarises its exposure to traffic generated by drivers travelling to the CCZ. Looking forward, differences in Congestion Charge Exposure across otherwise similar commutes establish that Central London's toll affects regional traffic patterns, causes sorting and general equilibrium effects, and calibrate commuters' roadspace demand for counterfactual policy simulations.

The right hand panel of figure 2 sheds light on the source of cross-commute variation by contrasting commutes with high and low Congestion Charge Exposure. The red commute is more exposed as it shares the road with CCZ-bound traffic on the westbound A13. Meanwhile, the yellow commute has relatively low Congestion Charge Exposure as it mostly travels parallel to the CCZ and spends less time on major radial roads.

### 4.2 Effect on regional traffic

I establish that Central London's Congestion Charge affected regional traffic patterns using the quasi-difference-in-differences regression

$$ln(traffic_{ijt}) = \varphi CCE_{ij} \times Post_{t} + X_{ij}' \Gamma_{t} + \gamma_{it} + \gamma_{jt} + \gamma_{ij} + \nu_{ijt} \tag{3}$$

where $traffic_{ijt}$ is traffic faced when choosing to commute between residence $i$ and workplace $j$ in year $t$, pre- or post-toll, and $Post_{t}$ is a post-toll dummy. The vector $X_{ij}$ includes fixed effects for one kilometre bins of driving distance, measures of public transit availabil-

ity, road characteristics, and indicators of commutes into the CCZ or WEZ, all interacted with post-2003 dummies.[12] Fixed effects $\alpha_{it}$ and $\alpha_{jt}$ capture changes at residences and workplaces and $\alpha_{ij}$ forces identification based on changes over time. The semi-elasticity $\varphi$ captures the effect of Central London's toll on the amount of traffic a commuter faces independent of their own tolled status. This parameter excludes the toll's direct effect since I explicitly control for commutes crossing either the WEZ or CCZ boundary.

Revisiting figure 2 elucidates how least squares estimates of equation 3 identify CCE's effect on traffic. The right panel shows similar length commutes to the same workplace where the red commute has higher CCE because of its reliance on the westbound A13, which happens to be an important route downtown. A regression with only pair and workplace-by-year fixed effects would identify CCE's effect by comparing traffic trends on the red and yellow commutes. For additional robustness, my preferred specification includes residence-by-year fixed effects that force comparisons relative to the average change in traffic from each residence. So, the comparison illustrated here identifies a negative effect of CCE on traffic only if the red commute's traffic falls or the yellow commute's traffic grows relative to other trips from their residences.

Table 3 presents least squares estimates of equation 3. CCE is scaled to have a standard deviation of one and standard errors are two-way clustered at both residence and workplace levels. All regressions include MSOA-pair fixed effects and are weighted by one plus initial commute flows.

[12] Public transit controls include dummies for residence and workplace centroids both 1,500 metres from a rail transit station, 1,500 metres from a Tramlink light rail station, and additional dummies for pairs on the same rail transit or Tramlink line. Tramlink is a suburban light rail system using a mixture of dedicated rights of way and at-grade track. Rail transit includes the London Underground, Overground, Docklands Light Rail, TfL rail, and Thameslink. I also include dummies for residence and workplace within 750 metres of the same express bus line, 500 metres of the same 24 hour bus line, and 500 metres from the same daytime-only bus line and control for the share of monitors observed each year, crossing the WEZ boundary, and crossing the CCZ boundary. Finally, I control for the natural logarithm of the total number of count points behind commute in each period, this is not interacted with a post dummy and captures additional measurement error.

Negative effects throughout table 3 indicate that Central London's Congestion Charge affected regional traffic patterns and that CCE is a reasonable measure of effects on untolled commutes. Comparing columns 1 and 2 shows that controlling for distance, workplace, and residence-by-year fixed effects attenuates CCE's effect on traffic growth. Column 3, which includes a complete vector of fixed effects and controls, associates a one standard deviation increase in Congestion Charge Exposure with a 2.6 percent decrease in traffic. Change in log-traffic has a standard deviation of 0.205 in this sample, so a one standard deviation increase in CCE moves a commute 13 percent of a standard deviation down in the distribution of traffic growth.

Appendix table A.2 shows that CCE's effect on traffic is smaller for far apart MSOA pairs, diminishes in CCE, is independent of whether the trip involves Outer London, and is larger for pairs with poor transit access. Appendix table A.3 shows that a similar relationship holds at the road link level and that CCE's effect on traffic is larger during the evening rush hour. Appendix table A.4 goes on to show that estimates are robust to dropping MSOAs in the CCZ, WEZ, or both; controlling for trends on pairs with new rail transit connections after 2001; and controlling for trends on the interaction of residence and workplace distance to London's central business district. Appendix table A.5 shows that results are robust to using proxies for Congestion Charge Exposure that are based an unweighted sum of tolled routes crossing each link or that treat the WEZ as tolled. Finally, appendix A.1.5 presents complementary event-study evidence that Central London's toll shifted traffic away from CCZ-bound radial roads.

## 5 Commuters' travel demand

I now build and empirically evaluate a discrete choice model of commuting between city neighbourhoods as a basis for estimating effects of London's Congestion Charge on economic

geography, public transit use, and welfare. In the model, commuters care about wages, rents, amenities, and commute costs and face a two-step discrete choice problem in which they choose locations to live and work before choosing a mode of transportation. These assumptions lead to a nested discrete choice demand system over locations and commute modes with a log-linear commuter-gravity equation (Ahlfeldt et al., 2015) and a logit mode choice inner nest (McFadden, 1974). This section concludes by discussing the model's empirical implications, providing reduced-form evidence of their relevance, and justifying assumptions used to calibrate central parameters for policy analysis.

### 5.1 Preferences

The city is partitioned into $N$ neighbourhoods that act as both workplaces and residences for a fixed population $\Omega$ with $\pi_g\Omega$ workers in each skill group $g$. Residence $i$ provides group $g$ an amenity $B_{ig}$, costs $Q_i$ per unit floorspace in rent, and workplace $j$ offers income $w_{jg}$. All workers have Cobb-Douglas preferences over residential floorspace and a consumption good comprising $\beta$ of their spending and incur a disutility of $d_{ijm} \geq 1$ from commuting between $i$ and $j$ on mode $m$.

Worker $\omega$ has idiosyncratic preference $\epsilon_{ij\omega}$ over locations and $\nu_{m\omega}$ over modes of transportation distributed Fréchet with cumulative distribution functions $G(\nu) = \mathrm{e}^{-\nu^{-\eta}}$ and $F(\epsilon) = \mathrm{e}^{-\epsilon^{-\theta}}$ independent of skill. Workers form these preferences sequentially, first observing $\epsilon_{ij\omega}$ and choosing a residence-workplace pair $ij$ then observing $\nu_{m\omega}$ and choosing a commute mode.[13]

Ex-post, worker $\omega$ of skill $g$ living in $i$, working in $j$, and commuting on mode $m$ has

[13] Gonzalez-Navarro and Turner (2018) provide motivating evidence of commuters' nested decision making by finding that public transit expansion increases ridership by causing residents to substitute across modes rather than cities.

indirect utility

$$u_{ijmg\omega} = \frac{B_{ig}}{Q_i^{1-\beta}} \frac{w_{jg}}{d_{ijm}} \nu_{m\omega} \epsilon_{ij\omega}$$

which is the result of Cobb-Douglas preferences fixing the optimal consumption bundle for each commute $ij$. Workers then choose $i$, $j$, and $m$ to maximize utility given wages, rents, amenities, and commute costs.

### 5.2 Commute costs

Each commute $ij$ is served by two modes of transportation: driving a car $c$ and other modes $o$. The cost of commuting from $i$ to $j$ on mode $m$ is

$$d_{ijm} = toll_{ij}^{1\{m=c\}} traffic_{ij}^{\kappa 1\{m=c\}} e^{\alpha_{im}+\alpha_{jm}+d_m(ij)} \tag{4}$$

where $\kappa$ is driving's commute cost elasticity of road traffic and $d_m(ij)$ is an arbitrary function of distance.[14]

Related literature typically assumes commute costs are an exponential function of travel time such as $d_{ij} = e^{\kappa_d Time_{ij}}$ where $\kappa_d$ is cost's travel time semi-elasticity (Ahlfeldt et al., 2015; Tsivanidis, 2019; Severen, 2019). Assuming travel time is a linear in log-traffic so that $time_{ij} = \kappa_0 + \kappa_T ln(traffic_{ij})$ implies that equation 4 is isomorphic to the standard model where traffic's marginal disutility is $\kappa = \kappa_d \times \kappa_T$. Therefore, abstracting from other issues such as travel time uncertainty implies that $\kappa$ combines traffic's marginal time costs and travel time's marginal disutility.[15]

With many modes of transportation, appendix A.2.1 shows that commuters' nested

[14] The quantitative model simulates recycling toll revenues and floorspace rents into infrastructure that delivers spatially uniform commute cost reductions as described in section 6.5.

[15] Since the data aggregate traffic over complete rush hour periods, accommodating more rush hour traffic on a given road link requires decreasing speeds as feasible headways fall. In the case of a bottleneck, increasing total rush hour traffic increases the amount of time with an active queue and even commuters who respond by altering departure times prefer less road traffic (Arnott et al., 1993).

discrete choice problem summarises commute costs between locations $ij$ with the cost index

$$\bar{d}_{ij} = \gamma \left( \sum_m d_{ijm}^{-\eta} \right)^{-\frac{1}{\eta}} \tag{5}$$

which is the mean cost of commuting between $i$ and $j$ including idiosyncratic preferences $\nu_{m\omega}$ given individually rational mode choices and $\gamma = \Gamma\left(1 - \frac{1}{\eta}\right)^{-1}$.

### 5.3 Commuting

Since workers choose locations before mode of transportation, skill does not directly affect mode choice and appendix A.2.1 shows that the probability a worker chooses mode $m$ to commute between residence $i$ and workplace $j$ is

$$\pi_{m|ij} = \sum_g \pi_{m|ijg} \pi_g = \frac{d_{ijm}^{-\eta}}{\sum_m d_{ijm}^{-\eta}} \sum_g \pi_g = \frac{d_{ijm}^{-\eta}}{\sum_m d_{ijm}^{-\eta}} \tag{6}$$

and the share of type $g$ commuters living in $i$ and working in $j$ is

$$\pi_{ij|g} = \frac{1}{\Phi_g} (w_{jg} B_{ig} Q_i^{\beta-1})^{\theta} \bar{d}_{ij}^{-\theta} \tag{7}$$

where $\Phi_g = \sum_{r'} \sum_s (w_{sg} B_{rg} Q_r^{\beta-1})^{\theta} \bar{d}_{rs}^{-\theta}$.

### 5.4 Empirical implications

The following theorem shows that commuting elasticities reflect traffic's effects on sorting, mode choice, and welfare:

**Theorem 1** (Traffic's effect on commuting) *Traffic's first-order effect on the number com-*

*muting between residence $i$ and workplace $j$ is*

$$\frac{\partial ln(\pi_{ij})}{\partial ln(traffic_{ij})} = -\theta\kappa\pi_{c|ij}(1-\pi_{ij}) \tag{8}$$

*and its first-order effect on mode choice is*

$$\frac{\partial ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right)}{\partial ln(traffic_{ij})} = -\eta\kappa \tag{9}$$

*where $\pi_{ij}$ is the probability of living in $i$ and working in $j$ and $\pi_{c|ij}$ is the conditional probability of driving on commute $ij$.*

**Proof:** See appendix A.2.1. □

Equation 8 shows that commute flow elasticities reflect location preferences through $\theta$ and increase in driving's conditional choice probability $\pi_{c|ij}$ to reflect heterogeneity in road quality and transit infrastructure. The location choice probability $(1-\pi_{ij})$ enters to reflect the weight of outside options and $\kappa$ reflects traffic's marginal disutility. For a given value of $\kappa$, commuting's traffic elasticity increases in $\theta$ to reflect the variance of commuters' idiosyncratic valuations of particular locations. Empirically, if increasing traffic reduces the number commuting, it must be that they choose different residences, find jobs in different workplaces, or both—evidence of sorting.

Equation 9 shows that traffic's effect on the log-odds of driving reflects its marginal disutility through $\kappa$ and the density of workers near the margin of switching modes through $\eta$. If increasing traffic decreases the log-odds of driving between given locations, it must be because public transit becomes relatively more attractive—endogenous mode choice.

As a practical matter, appendix A.2.2 argues that a non-zero commuting elasticity is necessary and sufficient for commute costs to have general equilibrium effects as they determine where people live and work, which then determines wages and rents.

### 5.5 Reduced form evidence

To establish the model's empirical relevance and justify the exclusion restrictions that calibrate commuters' travel demand system for policy analysis, I estimate traffic-elasticities of census counts of the total number commuting by all modes and odds of driving using regressions of the form

$$ln(\pi_{ijt}) = \delta^n ln(traffic_{ijt}) + X'_{ij}\beta_t + \alpha_{it} + \alpha_{jt} + \alpha_{ij} + u_{ijt} \tag{10}$$

$$ln(\pi_{ijtc}) - ln(\pi_{ijto}) = \delta^c ln(traffic_{ijt}) + X'_{ij}\beta^c_t + \alpha^c_{it} + \alpha^c_{jt} + \alpha^c_{ij} + u^c_{ijt} \tag{11}$$

where $\pi_{ijt}$ is the number of commuters living in $i$, working in $j$, in year $t$ (by all modes of transportation) and $\pi_{ijtm}$ is the subset choosing mode $m$ which takes values $c$ for drivers and $o$ for all other modes. The vector $X_{ij}$ includes covariates described in section 4.2, $\alpha_{it}$ and $\alpha_{jt}$ are MSOA-by-year fixed effects capturing changes at residences and workplaces, and pair fixed effects $\alpha_{ij}$ force identification based on changes over time.

The goal is to identify traffic elasticities $\delta^n$ and $\delta^c$ which are proportional to traffic's marginal cost and the number of commuters near the margin of each decision. Intuitively, increasing a commute's rush-hour traffic increases drivers' travel times and creates uncertainty that burdens commuters targeting early arrival times to avoid being late. Equation 11 is fully consistent with the structural model's mode choice problem in which $\delta^c$ is a parameter cluster reflecting substitution between modes of transportation. The commuting elasticity $\delta^n$ is proportional to commute costs' influence on sorting across locations, which in turn determines housing demand and labour supply.

### Identification and estimation

Before discussing the instrumental variables that identify traffic elasticities, it is useful to note that fixed effects in equation 10 account for most salient omitted variables. Pair fixed effects force identification based on changes over time and control for constant route characteristics such as long-lived infrastructure. Residence-by-year fixed effects control for differential trends in rent or amenities such as pollution, shop openings, and transit stations. Workplace-by-year fixed effects control for trends in workplace productivity or attractiveness including changing parking availability, input access, and floorspace availability. Baseline observables account for the toll itself, traffic measurement issues, and changing effects of distance and public transit. However, least squares elasticity estimates are biased by traffic's correlation with unmeasured determinants of commuting.

Even in panel data, unmeasured changes in commute characteristics will increase both commuting and traffic. In terms of mode choice, new public transit connections can decrease both road traffic and the odds of driving on a route. There could also be workplaces that are complements with certain residences due to parking or transit availability, both attracting a certain type of commuter, or workers and their clients driving in from proximate origins. These complementarities would increase both commuting and traffic between certain origin-destination pairs, positively biasing estimates of $\delta^n$ and $\delta^c$.

To identify traffic's effect on commuter demand, I instrument traffic with each route's Congestion Charge Exposure Index. CCE is defined to capture latent demand for tolled trips sharing the road with each commuter and section 4 shows that variation is largely driven by geography and idiosyncrasies in London's road network. Since I control for the toll's direct effect, commutes with high CCE see traffic decline after 2003 because of reduced driving on intersecting routes to tolled destinations. From a commuter's perspective, Central London's toll exogenously frees up roads serving highly exposed commutes, making

them faster, more reliable, and more attractive. In this sense, Congestion Charge Exposure predicts changes in traffic that are exogenous to commuters' location and mode choices conditional on fixed effects and controls.

I estimate parameters of interest using instrumental variables regressions with the linear first stage described by equation 3 in section 4, which confirms that CCE is a relevant instrument for traffic.[16] My preferred econometric estimator is two stage least squares (TSLS) adding one to each mode's commute flows before taking logarithms. I prefer this to a non-linear maximum likliehood approach because the data's small count adjustment makes observations with one and zero commuters indistinguishable.[17] Further, poisson and negative binomial estimators drop clusters of observations in which the dependent variable is always zero, a particularly undesirable property in this setting.[18] I also weight all regressions by 2001 commuters plus one, base inference on covariance matrices that are two-way clustered by residence and workplace, and appendix A.1.6 presents corresponding maximum likelihood estimates.

### Main results

Table 4 presents estimates of the number commuting's elasticity with respect to traffic. The first column presents OLS estimates and the remaining columns present TSLS estimates instrumenting traffic with Congestion Charge Exposure, all regressions include MSOA pair fixed effects and are weighted by one plus initial total flows, standard errors and F-statistics are two-way clustered at residence and workplace levels.

Column 1 of table 4 shows that the expected attenuation bias in OLS estimates is

[16] Table 3 presents first stage effects of CCE on traffic that verify the relevance condition $E\left[CCE_{ij} \times traffic_{ij} | X_{ij}, \theta_{b(ij)t}, \alpha_{it}, \alpha_{jt}, \alpha_{ij}\right] \neq 0$.

[17] The data appendix describes how the UK data service treats small counts in commuting matrices.

[18] Dropping zero-flow clusters is necessary because the poisson model would perfectly fit their outcomes by setting the associated fixed effect to $-\infty$. Unfortunately, current techniques for managing sparse commuting data focus on cross-sections and do not address this issue (e.g. Dingel and Tintelnot (2020)).

severe. Column 2 presents TSLS estimates without controlling for distance or residence and workplace trends, finding a strong negative relationship between changes in commuter counts and traffic. Column 3 adds residence and workplace-by-year fixed effects, attenuating traffic's effect on commuting. Column 4 adds controls which increase the elasticity estimate so that a ten percent decrease in a commute's road traffic increases the number of commuters choosing that location pair by 9.18 percent.[19] In terms of observed variation, this amounts to a standard deviation increase in log-traffic decreasing log-commuting by one standard deviation.

Table 5 presents similar estimates of traffic's effect on log-drivers relative to non-drivers, the mode choice regression. Again, OLS estimates display positive bias and the preferred estimates are negative and economically substantial. Comparing columns 2 and 3 shows that controlling for workplace and residence trends turns the relationship between traffic and drivers' relative commute share from positive to negative.[20] This suggests that routes connecting places becoming more car oriented have low Congestion Charge Exposure. Column 4 controls for route characteristics to find that a ten percent decrease in traffic increases the share of drivers relative to other modes by 12.1 percent.

In terms of observed variation, the mode choice elasticity of table 5 column 4 implies that a standard deviation decrease in log-traffic increases the log-odds of driving by 1.6 standard deviations. Evaluating for an MSOA pair where half of commuters drive associates a 3 point increase in the share of commuters driving with a ten percent decrease in traffic. In context, this is a large negative effect of increasing traffic on the share of commuters choosing to drive.

[19] Unreported results suggest that traffic monitoring network controls can account for the difference between columns 3 and 4 of table 4, but independently adding transit controls also increases elasticity estimates.

[20] Unreported results suggest that bias from workplace and residence trends are largely due to transit infrastructure. Adding the full vector of controls without workplace and residence-by-year effects yields $\delta^c = -0.729$ with a two-way clustered standard error of 0.535.

**Robustness**

Appendix A.1.6 presents poisson and negative binomial control function instrumental variables (CFIV) estimates of traffic's effect on commuting. CFIV regressions drop a number of location pairs without commuting to give smaller and less powerful elasticity estimates than TSLS. Despite their drawbacks, CFIV estimates support the conclusion that traffic affects commuting—a weighted poisson estimate implies that a ten percent decrease in a commute's traffic increases the number choosing that location pair by 3.2 percent.

Appendix table A.4 presents additional robustness tests. The first three rows show that commuting elasticities are insensitive to dropping MSOAs in the CCZ, WEZ, or both and that dropping the CCZ accentuates effects on mode choice. The final two rows show that elasticities are also unaffected by controlling for a post-toll dummy interacted with rail transit connections opening after 2001 or the interaction of residence and workplace distance from the central business district.

Appendix figure A.5 shows that CCE is related to certain measures of transit connections but not systematically predicted by better or worse public transit in 2001.[21] To test the importance of rail transit expansion between 2001 and 2011, I identify newly connected pairs as those with both workplace and residence within 1,500 metres of a station in 2011 where at least one of these stations opened after 2001. Appendix table A.6 shows that new connections are positively correlated with CCE but this relationship evaporates when controlling for residence and workplace fixed effects. The same table defines an indicator for pairs with both workplace and residence within 750 metres of stops on the same express or 24 hour bus route to show that CCE's spatial distribution is conditionally uncorrelated with important bus routes.

[21] Appendix figure A.6 shows that the CCZ is not systematically related to the toll or years underlying pairwise traffic observations conditional on origin, destination, and distance fixed effects. However, there is a positive conditional correlation between CCE and the number of road links monitored post-toll.

Finally, appendix table A.7 breaks down mode choice effects by separately estimating traffic's effect on the number of commuters driving and using other modes. Results indicate that decreasing a route's traffic significantly increases the number of commuters driving with no effect on the number taking other modes. This is consistent with assuming that road traffic has no direct effect on the cost of commuting by public transit.

**Heterogeneous elasticities**

Appendix A.1.7 investigates skill-specific effects of traffic on commuting by classifying MSOA pairs according to the skill intensity of each residence and workplace. Regressions interacting MSOA pairs' traffic and skill intensity provide some evidence of larger mode choice elasticities between high-skill locations and larger effects of traffic on the location choices of low-skill commuters. This evidence of a larger location-choice elasticity for the low-skill is consistent with commuter gravity estimates from Tsivanidis (2019), Zárate (2019), and Lee (2019). While my data cannot pin down the mechanism, I propose that this pattern reflects differences in sorting behaviour and labour market outcomes rather than time preferences. This is consistent with the heterogeneous location fundamentals and homogeneous commute costs in the travel demand system I use for policy analysis.

Appendix tables A.8 and A.9 investigate additional sources of treatment effect heterogeneity. Table A.8 interacts log-traffic and instruments with distance tercile dummies to find that elasticities are largely independent of trip distance, consistent with the structural model's constant elasticity disutility of traffic. Table A.9 interacts traffic with a dummy indicating commutes within Inner London, finding that traffic reduces the number commuting most within Inner London and mode choice elasticities are only negative and statistically significant if the workplace or residence is in Outer London.

# 6 A quantitative general equilibrium model

Traffic's significant effect on location choices motivates using a general equilibrium model to quantitatively assess congestion charging's implications for city structure, inequality, and welfare. This section extends the travel demand model in section 5 to account for general equilibrium effects on road traffic, housing, and labour markets. The full model also leverages data describing employment and population of high- and low-skill workers to estimate congestion charging's unequal effects.

## 6.1 Labour supply, residential population, and wages

Commuters' travel demand, described in section 5, jointly determines labour supply and housing demand in each neighbourhood. Labour supply of skill $g$ to workplace $j$ is $L_{jg} = \Omega \sum_i \pi_{ijg}$ and group $g$ population in neighbourhood $i$ is $N_{ig} = \Omega \sum_j \pi_{ijg}$. As in Ahlfeldt et al. (2015) and Lee (2019), each group's wages satisfy

$$L_{jg} = w_{jg}^{\theta} \sum_i \frac{\bar{d}_{ij}^{-\theta}}{\sum_s w_{sg}^{\theta} \bar{d}_{is}^{-\theta}} N_{ig} \tag{12}$$

which defines a unique vector of wages rationalizing location decisions and transportation costs.

## 6.2 Housing and rents

Housing costs are determined by residential floorspace demand

$$H_i = \frac{1-\beta}{Q_i} \Omega \sum_g \sum_j \pi_{ijg} w_{jg}.$$

To close the housing market, I assume that an immutable zoning code fixes each MSOA's supply of residential floorspace at $H_i$. This simplification seems reasonable given London's scarcity of developable land and limited opportunities to convert between residential and commercial uses (Cheshire and Hilber, 2008; Hilber and Vermeulen, 2016). Housing rents are combined with firm's floorspace spending and invested in transportation infrastructure that uniformly decreases commute costs.[22]

### 6.3 Labour demand

Competitive firms use a Cobb-Douglas function of floorspace and a constant elasticity of substitution (CES) aggregate of both skill groups' labour to produce a traded output whose price is normalized to one. I assume commercial floorspace available in each workplace is fixed and simplify notation by normalizing floorspace inputs and assuming total factor productivity varies across workplaces.

Firms in location $j$ produce $A_j L_j^{\alpha}$ where $\alpha$ is labour's share of production and total labour input is $L_j = \left(\sum_g a_{jg} L_{jg}^{\rho}\right)^{\frac{1}{\rho}}$. Firms substitute workers' skills with elasticity $\sigma = \frac{1}{1-\rho}$ and $a_{jg}$ is skill $g$'s input intensity at workplace $j$ defined so that $\sum_g a_{jg} = 1$. Demand for group $g$ labour in workplace $j$ satisfies

$$w_{jg} = \alpha a_{jg} A_j L_{jg}^{-\frac{1}{\sigma}} L_j^{\alpha-\rho} \tag{13}$$

and free entry ensures firms spend $F_j = \frac{1-\alpha}{\alpha} \sum_g w_{jg} L_{jg}$ on floorspace. In my empirical application with two skill groups $h$ and $l$, equilibrium wages and employment reveal each workplace's skill intensity as

$$\frac{a_{jh}}{1 - a_{jh}} = \frac{w_{jh}}{w_{jl}} \left(\frac{L_{jh}}{L_{jl}}\right)^{\frac{1}{\sigma}}. \tag{14}$$

[22] Section 6.5 details how rents are returned to commuters.

### 6.4 Road traffic

The route from $i$ to $j$ is an exogenously determined set of road links with endogenous traffic levels. Road link $\ell$ corresponds to a travel direction along a stretch of road and its traffic in rush hour $h$ is the sum of commute flows and other traffic so that $traffic_\ell^h = \Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\} + \mathcal{O}_\ell^h$. The first term, $\Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\}$, is the endogenous number of commuters whose routes include link $\ell$ and $\mathcal{O}_\ell^h$ is non-commuter traffic.[23]

Each link-hour's non-commuter traffic is a reduced-form function of Congestion Charge Exposure $\mathcal{O}_\ell^h = \varphi^h CCE_\ell^h + e_\ell^h$. This gives interior solutions to road link traffic

$$traffic_\ell^h = \Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\} + \varphi^h CCE_\ell^h + e_\ell^h.$$

I impose non-negativity constraints by assuming that non-commute traffic endogenously responds to low demand from commuters such that $\mathcal{O}_\ell^h = \underline{traffic_\ell^h}$ if $\Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\} + \varphi^h CCE_\ell^h + e_\ell^h < \underline{traffic_\ell^h}$.[24] Finally, commuters between $i$ and $j$ face the sum of all traffic on the links they cross going to work in the morning and returning in the evening averaged across these two trips.

### 6.5 Equilibrium with exogenous productivity

In equilibrium the mass of group $g$ workers living in $i$ and working in $j$ is

$$\Omega \pi_{ijg} = \Omega \frac{\pi_g}{\Phi_g} (w_{jg} B_{ig} Q_i^{\beta-1})^\theta \bar{d}_{ij}^{-\theta} \tag{15}$$

[23] Appendix A.1.4 discusses an empirical counterpart to the accounting identity describing links' traffic, provides evidence of non-commuter traffic's first-order importance, shows that measurement error precludes identifying each road link's non-commuter traffic, and argues that these errors are random across road links.

[24] Link-specific lower bounds on traffic can be modelled as the result of a version of induced roadspace demand documented by Duranton and Turner (2011).

where $\bar{d}_{ij} = \gamma^{-1}\left(\sum_m d_{ijm}^{-\eta}\right)^{-\frac{1}{\eta}}$. Labour market clearing implies that wages satisfy

$$w_{jg} = \alpha a_{jg} A_j \left(\Omega \sum_i \pi_{ijg}\right)^{-\frac{1}{\sigma}} \left(\sum_{g'} a_{jg'} \left(\Omega \sum_i \pi_{ijg'}\right)^{\rho}\right)^{\frac{\alpha-\rho}{\rho}} \tag{16}$$

floorspace market clearing implies

$$Q_i = \frac{1-\beta}{H_i} \Omega \sum_g \sum_j \pi_{ijg} w_{jg} \tag{17}$$

and

$$F_j = \frac{1-\alpha}{\alpha} \Omega \sum_g \sum_i \pi_{ijg} w_{jg}. \tag{18}$$

Given parameters $\{\theta,\ \sigma,\ \alpha,\ \beta\}$, a closed city labour market equilibrium is endogenous allocations $\{\pi_{ijg}\}$, prices $\{w_{jg},\ F_j,\ Q_i\}$, and expected utility (equation 31) determined by fundamentals $\{B_{ig},\ A_j,\ a_{jg}\}$, commute costs indices $\{\bar{d}_{ij}\}$, aggregate skill shares $\{\pi_g\}$, and population $\Omega$ so that equations 15, 16, 17, and 18 hold.

The labour market determines locations $\pi_{ij} = \sum_g \pi_{ijg}$ and commute costs with toll and rent recycling are

$$\frac{\bar{d}_{ij}}{G \cdot Q} = \frac{\gamma}{G \cdot Q}\left(d_{ijc}^{-\eta} + d_{ijo}^{-\eta}\right)^{-\frac{1}{\eta}} \tag{19}$$

where $d_{ijc} = toll_{ij} \cdot traffic_{ij}^{\kappa} \cdot e^{\alpha_{ic}+\alpha_{jc}+d_c(ij)}$ are driving-specific costs, $G = \left(\sum_{ij} \pi_{ijc} d_{ijc} + \pi_{ijo} d_{ijo}\right)$ $\times \left(\sum_{ij} \pi_{ijc} \frac{d_{ijc}}{toll_{ij}} + \pi_{ijo} d_{ijo}\right)^{-1}$ recycles toll revenues, and $Q = \ddot{c}\left(\sum_i H_i Q_i + \sum_j F_j\right)$ recycles floorspace rents. I assume that adjustment factors $G$ and $Q$ recycle revenues and rents by improving transportation infrastructure serving each commute without affecting relative prices. This specification accommodates a fixed efficiency factor $\ddot{c}$ that converts rent

receipts into commute cost reductions.[25]

Commute costs depend on endogenous allocations through road traffic $traffic_{ij} = \frac{1}{2}\sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} traffic_\ell^h$ where

$$traffic_\ell^h = \max\left(\varphi^h CCE_\ell^h + e_\ell^h + \Omega \sum_{r,s} 1\{\ell \in R_{rs}^h\}\pi_{c|rs}\pi_{rs}, \underline{traffic}_\ell^h\right) \tag{20}$$

and $h \in \{morn, eve\}$. Combining with equation 6 gives

$$\begin{aligned} ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right) =& \eta\alpha_{ic} + \eta\alpha_{jc} + \eta d_c(ij) - \eta ln(toll_{ij}) + \eta ln(d_{ijo}) \\ & - \eta\kappa ln\left[\frac{1}{2}\sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} \max\left(\varphi^h CCE_\ell^h + e_\ell^h + \Omega \sum_{r,s} 1\{\ell \in R_{rs}^h\}\pi_{c|rs}\pi_{rs}, \underline{traffic}_\ell^h\right)\right] \end{aligned} \tag{21}$$

which implicitly defines mode shares $\{\pi_{c|ij}\}$ as a function of location choices. Mean worker welfare is expected utility

$$E(u) = \gamma'\gamma \cdot G \cdot Q \sum_g \pi_g \left[\sum_i \sum_j \left(\frac{B_{ig} w_{jg}}{\bar{d}_{ij} Q_i^{1-\beta}}\right)^\theta\right]^{\frac{1}{\theta}} \tag{22}$$

where $\gamma'\gamma = \Gamma\left(1-\frac{1}{\theta}\right)\Gamma\left(1-\frac{1}{\eta}\right)^{-1}$.

Given parameters $\{\theta, \eta, \kappa, \varphi^{morn}, \varphi^{eve}, \sigma, \alpha, \beta\}$, a closed city general equilibrium is a labour market equilibrium plus traffic and mode shares $\{traffic_\ell^h, \pi_{c|ij}\}$ so that equations 15, 16, 17, 18, 19, 20, and 21 hold given routes, labour market fundamentals, transit costs, mode share fixed effects, cost residuals, charge exposure, and tolls $\{R_{rs}^h, B_{ig}, A_j, a_{jg}, \alpha_{ic}, \alpha_{jc}, d_{ijo}, e_\ell^h, CCE_\ell^h, toll_{ij}\}$. Appendix A.2.3 discusses the existence and uniqueness of a

[25] This form of rent recycling implies a unit elasticity of each commuter's equilibrium utility with respect to total rent and toll revenue recycling adjusts commute costs so that their mean is equal to its value with equilibrium traffic but no tolls.

closed city general equilibrium with exogenous productivity.

### 6.6 Endogenous productivity

I introduce symmetric agglomeration externalities that decay across neighbourhoods as in Ahlfeldt et al. (2015), allowing total factor productivity to increase in nearby labour inputs such that

$$A_j = a_j \Upsilon_j^{\lambda} \text{ s.t. } \Upsilon_j = \sum_s e^{-\delta d(s,j)} \frac{L_s}{Area_s}$$

where $Area_s$ is the land area of neighbourhood $s$ and $d(s,j)$ is the euclidean distance between centroids of $s$ and $j$ set to $\sqrt{\frac{Area_j}{3.14}}$ when $s = j$. The parameter $\delta > 0$ governs spillovers' spatial decay and $\lambda$ is productivity's elasticity of nearby CES employment indices. This externality adapts Ahlfeldt et al.'s (2015) single-skill model to allow for cross-neighbourhood agglomeration externalities from firms' labour inputs.

## 7 Model estimation and calibration

For counterfactual policy analysis, I require estimates of parameters $\{\theta, \eta, \kappa, \varphi^{morn}, \varphi^{eve}, \sigma, \alpha, \beta\}$, commute costs $\{d_{ijm}\}$, amenities $\{B_{ig}\}$, wages $\{w_{jg}\}$, skill intensities $\{a_{jg}\}$, regional skill shares $\{\pi_g\}$, and traffic residuals $\{e_\ell^h\}$. Incorporating endogenous productivity also requires taking a stand on agglomeration's decay $\delta$ and elasticity $\lambda$. This section describes how I calibrate the model using link-level traffic volumes and Congestion Charge Exposure, commute flows between MSOA pairs, and MSOA population and employment counts by skill.

### 7.1 Sequential parameter estimation

I use London's Congestion Charge to estimate location and mode choice elasticities ($\theta$ and $\eta$) from commute flows following the reduced form analysis in section 5.5. The location

choice elasticity ($\theta$) reflects unobserved heterogeneity in preferences across residences and workplaces and governs commuters' propensity to sort across locations. The mode choice elasticity ($\eta$) reflects heterogeneity in preferences for driving over taking public transit and governs commuters' propensity to switch modes as commute costs change. After setting the commute cost elasticity of traffic ($\kappa$) to reflect travel time costs, I estimate location and mode choice elasticities sequentially using Congestion Charge Exposure as instruments for the model's commuter gravity equations.

Aggregating across groups' location choices, taking logarithms, and substituting for commute costs using the identity $\bar{d}_{ij} = \frac{\gamma}{G \cdot Q}(1 - \pi_{c|ij})^{\frac{1}{\eta}} d_{ijo}$ gives the structural gravity equation

$$ln(\pi_{ij}) = ln\left(\sum_g \frac{\pi_g}{\Phi_g}(w_{jg}B_{ig}Q_i^{\beta-1})^\theta\right) - \frac{\theta}{\eta}ln(1 - \pi_{c|ij}) + \theta ln\left(\frac{G \cdot Q}{\gamma d_{ijo}}\right).$$

The first term summarises amenities, rents, and wages that I capture with time-varying residence and workplace fixed effects. Adding time subscripts and decomposing $\theta ln\left(\frac{G \cdot Q}{\gamma d_{ijo}}\right)$ into controls, fixed unobservables, and a residual gives the estimating equation

$$ln(\pi_{ijt}) = \omega_{jt} + b_{it} - \frac{\theta}{\eta}ln(1 - \pi_{c|ijt}) + X'_{ij}\delta_t + \delta_{ij} + u_{ijt}. \tag{23}$$

Assuming $E[CCE_{ij} \times Post_t \times u_{ijt}] = 0$, I use a weighted TSLS regression to identify the parameter ratio $\frac{\theta}{\eta}$.[26] Appendix table A.10 presents detailed TSLS estimates of equation 23 weighted by 2001 commuting and adding $\frac{1}{N \times (N-1)}$ to choice probabilities before taking logarithms. Results are insensitive to controls and I set $\frac{\theta}{\eta} = 2.229$ to match the full

[26] Linearly decomposing the structural unobservable gives $ln\left(\sum_g \frac{\pi_g}{\Phi_g}(w_{jg}B_{ig}Q_i^{\beta-1})^\theta\right) = \omega_{jt} + b_{it} + ln\left(\sum_g \frac{\pi_g}{\Phi_g}\frac{B_{ig}^\theta}{\sum_{g'}\pi_{g'}B_{ig'}^\theta}\frac{w_{jg}^\theta}{\sum_{g'}\pi_{g'}w_{jg'}^\theta}\right)$ where the final term is included in the residual of equation 23. This omitted variable will only bias parameter estimates if Congestion Charge Exposure is systematically correlated with residence-workplace pairs that offer high amenities and wages to one group relative to others.

specification in table A.10's second column.

I then use equation 6 to identify $\eta\kappa$ from the TSLS regression

$$ln(\pi_{c|ijt}) - ln(1 - \pi_{c|ijt}) = \alpha^c_{it} + \alpha^c_{jt} - \eta\kappa ln(traffic_{ijt}) + X'_{ij}\beta^c_t + \theta^c_{b(ij)t} + \alpha^c_{ij} + u^c_{ijt} \quad (24)$$

under the assumption $E[CCE_{ij} \times Post_t \times u^c_{ijt}] = 0$. Section 5.5 presents results and I set $\eta\kappa = 1.208$ to match column 4 of table 5.

Assuming commute costs are an exponential function of travel time, that travel time's marginal disutility per minute is 0.01 (Ahlfeldt et al., 2015; Tsivanidis, 2019; Severen, 2019), and that log-traffic has a constant effect on travel time given by $\kappa_T$ implies that the commute cost elasticity of traffic is $\kappa = 0.01 \times \kappa_T$. Herzog (2021a) shows that a log-point increase in traffic increases travel times by 0.16 seconds per metre and converting seconds to minutes and evaluating for a 14 kilometre trip delivers $\kappa_T = 37.33$.[27] Multiplying by 0.01 gives traffic's commute cost elasticity $\kappa = 0.3733$—absorbing ten percent more traffic reduces utility by 3.7 percent—which implies $\eta = 3.24$ and $\theta = 7.21$. The location choice elasticity ($\theta$) is similar to Ahlfeldt et al. (2015) and larger than in Tsivanidis (2019) and Severen (2019), suggesting that locations are more homogeneous on unobservables in my setting. The mode choice elasticity ($\eta$) lacks counterparts in recent literature but my estimate suggests that commuters view driving and transit as fairly close substitutes.[28]

## 7.2 Commute costs

I identify commute cost indices up to scale by adjusting 2011 commute flows for traffic using elasticity estimates from section 5.5 and regressing them on controls, distance, residence,

[27] Herzog (2021a) also demonstrates that a linear-in-logs functional form is a reasonable fit for London's link-level travel supply function.

[28] The mode choice elasticity ($\eta$) is equal to commuters' elasticity of substitution between driving and other modes with respect to commute cost changes.

and workplace fixed effects $ln(\pi_{ij2011}) + 0.918ln(traffic_{ij2011}) = \theta_{b(ij)} + \alpha_i + \alpha_j + e_{ij}$. I recover commute cost indices as the structural residual $\bar{d}_{ij} = \exp\left(\frac{1}{\theta}(\alpha_i + \alpha_j - ln(\pi_{ij2011}))\right)$ which includes the non-linear distance function captured by distance band fixed effects. I then substitute mode shares into equation 5 to get transit costs $d_{ijo} = \bar{d}_{ij}\pi_{o|ij}^{-\frac{1}{\eta}}$ and driving costs $d_{ijc} = \bar{d}_{ij}\pi_{c|ij}^{-\frac{1}{\eta}}$ up to scale.[29]

### 7.3 Wages and adjusted amenities

I calibrate each group's wages and amenities to rationalize their spatial distributions of employment and population as in Ahlfeldt et al. (2015) and Lee (2019). Given $\theta$, $\bar{d}_{ij}$, and data on high- and low-skill employment and population in 2011, I solve the system of equations described by equation 12 for high- and low-skill wages that rationalize locations and commute costs. Given wages, I use equilibrium populations to identify adjusted amenities $\widetilde{B}_{ig} = \Phi_g^{-1}(B_{ig}Q_i^{\beta-1})^{\theta}$ by aggregating across each group's gravity equation and re-arranging to compute $\widetilde{B}_{ig} = \frac{\pi_{i|g}}{\sum_j w_{jg}^{\theta}\bar{d}_{ij}^{-\theta}}$ up to scale.[30]

### 7.4 Commuting by skill

I use the model to separately estimate commute flows of high- and low-skill workers. Taking the ratio of each group's location choices (equation 7) gives

$$\frac{\pi_{ij|h}}{\pi_{ij|l}} = \frac{w_{jh}^{\theta}\widetilde{B}_{ih}}{w_{jl}^{\theta}\widetilde{B}_{il}}. \tag{25}$$

Given wages and adjusted amenities, I solve for skill specific commute flows that satisfy equation 25 and add up to observed 2011 commuting $\pi_{ij} = \pi_h\pi_{ij|h} + \pi_l\pi_{ij|l}$ for each MSOA

[29] Commute costs are identified up to a common scale by assuming $\gamma = \frac{1}{G \cdot Q}$.

[30] I scale wages to have geometric means of £120 and £160 for low- and high-skill respectively. After identifying skill-specific commute flows, an expected wage of £170 rationalizes 2011 location choices as a labour market equilibrium.

pair.[31]

### 7.5 Labour and housing demand parameters

I set housing and labour demand parameters according to existing estimates. I set housing's expenditure share $(1-\beta) = 0.4$ (as in Severen (2019)) a relatively large value that is reasonable in a city as expensive as London. On the labour demand side, I set firms' floorspace share to $(1-\alpha) = 0.2$ and the elasticity of substitution between high- and low-skill workers to $\sigma = 1.3$ following Card (2009), Tsivanidis (2019), and Lee (2019). Given wages and the elasticity of substitution, I evaluate equation 14 to identify skill intensities that rationalize high- and low-skill workers' relative employment and wages in each MSOA.

I set the agglomeration elasticity $\lambda = 0.07$ and decay $\delta = 4.2$ to match estimates from Ahlfeldt et al. (2015) given a 5 km-per-hour walking speed. These parameters are consistent with the rapidly decaying agglomeration forces found in similar contexts (Arzaghi and Henderson, 2008; Ahlfeldt et al., 2015; Lee, 2019).

### 7.6 Traffic residuals

I set traffic residuals $e^h_\ell$ that rationalize the distributions of traffic and CCE across road links in each rush hour. First, I run panel regressions of link-level morning and evening traffic on Congestion Charge Exposure to identify $\varphi^h$ in each rush hour from

$$traffic^h_{\ell t} = \varphi^h CCE^h_\ell \times Post_t + X_i' \Gamma^h_t + \gamma^h_t Post_t + \gamma^h_\ell + e^h_{\ell t} \tag{26}$$

[31] Groups' population and employment are imbalanced by those commuting into and out of the region. In 2011 the ratio of high-skill commuter population to employment in the region is 1.03 and this ratio is 2.12 for low-skill workers. Therefore, regional skill shares $\pi_g$ are the average of each group's total commuter population and employment in 2011 and normalized to sum to one. This procedure sets $\pi_h = 0.755$, London's local labour market is approximately 76 percent high-skill workers.

where $h \in \{morn, eve\}$ denotes morning and evening rush hour traffic, $\gamma_\ell^h$ are road-link fixed effects, and $X_i$ are baseline controls indicating links within the CCZ, WEZ, year of measurement, and log-distance from the boundary interacted with a CCZ dummy. I extract each hour's marginal effect $\varphi^h$ and set $e_\ell^h = traffic_{\ell 2011}^h - \varphi^h CCE_\ell^h - \Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}\}$ which is non-commuter traffic without the toll. Finally, I impose non-negativity constraints by calibrating induced non-commuter roadspace demand so that counterfactual traffic cannot fall below its observed level. Appendix table A.3 presents least squares estimates of equation 26 which suggest that a standard deviation increase in link-level charge exposure decreases daily rush hour traffic volume by 66 vehicles on average and that this effect is larger in the evening rush hour.[32]

### 7.7 Model fit

The model predicts high- and low-skill commuters using cars and public transit at the same rate holding all else equal. Interestingly, the data suggest that this assumption is a reasonable approximation in Greater London. In particular, section 5.5 finds little evidence of heterogeneous mode choice elasticities of traffic and workplace data suggest roughly equal car shares of low- and high-skill commuters. In 2011, 37 percent of low-skill workers drive to work and 35 percent of high-skill workers drive.

Figure 4 presents several assesments of the model's ability to fit commuter's decisions. In each panel, colours delineate skill groups and residential or workplace counts, large points are binned means, and the red dashed 45 degree line indicates a perfect fit. The top row presents overidentification tests on 2011 data and the bottom row recalibrates to 2001 and tests the model's ability to match changes over time. The top-left panel shows that the model generally fits 2011 driver shares well except for moderately over-predicting low-skill

[32] CCE has a standard deviation of one and a minimum value of zero in each rush hour.

driving and the lower-left panel shows that this issue does not occur in changes. The top-right panel shows that the model makes accurate predictions for both groups' employment but has some trouble capturing residential patterns. Fortunately, the lower-right panel shows that the model fits residential changes better than levels with the exception of a few outliers.

# 8 Counterfactuals

I now simulate removing Central London's toll to compute counterfactual changes in commuters' locations, driving rates, road traffic, and welfare. Simulations shock commute costs by removing the toll's monetary cost from commutes into the CCZ and setting each road link's Congestion Charge Exposure to its normalized minimum of zero. Removing the congestion charge adds non-commuter traffic to roads in proportion to CCE's spatial distribution (shown in figure 3) creating an initial traffic shock that reallocates roadspace away from commuters.[33] I then solve for relative prices and allocations using an iterative hat algebra procedure following Dekle et al. (2007) and report the policy's effects as percent differences between actual and counterfactual outcomes.[34] Since I fix Greater London's population, transportation policy can improve commuter welfare to the extent that it increases real wages and decreases commute costs.

## 8.1 Welfare, mode choice, and road traffic

Table 6 presents percent change in equilibrium utility, inequality, land rent, traffic, and driving's share of commutes caused by London's Congestion Charge. Iceberg tolls reflect

[33] The initial traffic shock sets each link-hour's traffic equal to its residual plus commuter traffic $e_\ell^h + \Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\}$ and I account for cross-commute spillovers as the computational algorithm converges to a new equilibrium. Appendix A.2.5 details quantitative assumptions.

[34] Appendix A.2.4 defines a hat algebra equilibrium and describes a computational procedure for arriving at one.

an £8 daily toll, revenue is expressed as the percent change expected utility from spending to reduce commute costs, and inequality is the ratio of high- and low-skill expected utility. Columns sequentially add model mechanisms that section 8.3 describes in detail and the final column presents results using the full model with agglomeration externalities.

The final column of table 6 shows that London's Congestion Charge increased expected utility by 0.11 percent. This effect is a similar order of magnitude to early time savings calculations (Leape, 2006), potential short-run gains from an optimal congestion charge in Bangalore (Kreindler, 2018), and recent estimates of gross welfare gains from public transit expansion (Severen, 2019).[35] Multiplying the mean welfare gain by 2.9 million commuters in the 2011 flow data, a £170 mean daily wage, and 260 work days per year gives an annual benefit of approximately £140 million, exceeding Leape's (2006) estimate of £130 in operating and amortized capital costs.[36]

In terms of inequality, the second and third rows of table 6's final column show that the congestion charge favours low-skill commuters and decreases utility inequality by 0.091 percent. The policy is progressive because high-skill commuters are more likely to work in tolled workplaces and the low-skill rely more on suburban commutes that are less well served by London's hub-and-spoke public transit system. This follows from census data discussed in section 3 which shows that high-skill jobs are disproportionately concentrated in the CCZ while suburban workplaces are particularly important for the low-skill. Calibrated location

[35] Leape (2006) associates London's Congestion Charge with £200 million in annual time savings to all drivers, commuters and otherwise. Kreindler (2018) combines a field experiment and model to find that optimal congestion charging in Bangalore would increase commuter welfare by 0.5 percent in the short run where origin, destination, and mode are fixed but departure time and driving route adjust. Severen (2019) estimates that Los Angeles Metro Rail increases welfare between 0.051 to 0.10 percent (depending on the time horizon and time savings for drivers) and that this benefit fails to exceed annual operating subsidies. Estimated benefits of public transit improvement in developing world cities typically exceed those from richer cities (Tsivanidis, 2019; Zárate, 2019).

[36] Leape (2006) estimates that in 2005, the CCZ's cost included £85 million for operation, £5 million for administration, and £23 million for amortized set-up costs. Inflating to 2011 GBP using the UK CPIH all items consumer price index gives approximately £130 million.

fundamentals reflect this pattern and comparing across table 6's columns shows that the policy is progressive regardless of sorting and externalities.

Table 6's fourth row shows that spending toll revenues on uniform transportation improvements increases utility by 0.11 percent. Table 6's fifth row shows that gains from recycling toll revenues are partially offset by a 0.061 percent decline in total land rent. This implies that commuters would enjoy a direct utility improvement of 0.171 percent from the policy in a model without rent recycling.

The sixth row of table 6 shows that London's Congestion Charge increases the driving rate of the region's commuters by 0.531 percent. Comparing columns 2 and 3 shows that endogenous traffic attenuates this mode shift as commuters endogenously sort and crowd each-other onto public transit. The final row of table 6 shows that London's Congestion Charge reduces total traffic by 1.007 percent. Comparing to the initial traffic shock (columns 1 and 2) shows that traffic would have fallen by an additional 0.183 percentage points if commuters did not endogenously sort and substitute from public transit to driving.

## 8.2 Economic geography

Figure 5 maps percent change in each skill group's MSOA employment and population caused by London's Congestion Charge, red hues indicate declines and green indicates growth. The left panel shows that the policy decreases both groups' employment in the CCZ and increases employment in Inner London's CCZ-adjacent MSOAs. Interestingly, low-skill employment also falls throughout the suburbs and concentrates in a ring around Central London. Effects on high-skill suburban employment are less prominent, suggesting their workplaces are less geographically flexible than low-skill workplaces.

Figure 5's right panel shows that London's Congestion Charge causes population growth throughout the suburbs. In fact, it appears that the policy reallocates high- and low-

skill population to different suburban neighbourhoods—low-skill population grows more in northern and eastern suburbs while high-skill population growth favours the west and eastern neighbourhoods along the Thames.

### 8.3 Mechanisms

In the model, three mechanisms mediate transportation policy's welfare effects: commuters' location and mode choices, endogenous wages and rents, and endogenous road traffic caused by commuters who drive. Table 6 unpacks these mechanisms by presenting the congestion charge's effect on aggregate outcomes in four simplified models.

Column 1 of table 6 eliminates all three margins of adjustment and aggregates across commuters to compute the initial policy shock's benefit. This partial equilibrium effect is equivalent to a traditional value of time savings calculation holding commuters' location and mode choices as they are with the toll in place.[37] Column 2 allows locations, wages, and rents to adjust, shutting down all externalities and exogenously fixing traffic after the initial shock. Column 3 endogenizes road traffic to commuters' decisions but omits cross-neighbourhood productivity externalities.

Column 1 of table 6 shows that a partial equilibrium calculation overstates the congestion charge's aggregate welfare effect. Repeating the calculation from section 8.1 using this partial equilibrium measure would associate a £235 million annual welfare gain with the policy, nearly double the £140 million equilibrium benefit. Column 2 shows that sorting and endogenous price responses attenuate commuter gains relative to the partial equilibrium measure. Moving to column 3 shows that adding endogenous road traffic further reduces gains and attenuates the policy's effect on equilibrium driving rates.

[37] Specifically, column 1 of table 6 presents $dlnE(u|g) = dln(G) - \sum_{i,j} \pi_{ij|g} \frac{d_{ijc}}{d_{ij}} (dln(toll_{ij}) + \kappa dln(traffic_{ij}))$. See Hall (2020), Yang et al. (2020), and Herzog (2021b) for recent applications of related time savings calculations.

Comparing columns 3 and 4 shows that agglomeration externalities cause a small decrease in the CCZ's benefits to commuters. This is because agglomeration lost to endogenous employment decentralization is partially offset by decreased suburban employment and jobs coalescing in Inner London's CCZ adjacent area. The model highlights this mechanism by capitalizing lost agglomeration into floorspace rents; excluding rent reductions in row 5 from welfare gains in row 1 shows that agglomeration would accentuate the policy's benefits in a model where absentee landlords do not return rents to commuters.

For a lower bound on welfare gains, it is useful to compare results with simulations that do not assume London's Congestion Charge reallocates roadspace to commuters. To this end, appendix A.2.6 presents an alternative simulation where removing tolls decreases the cost of tolled commutes but leaves non-commuter traffic unaffected before determining equilibrium locations, modes of transportation, and traffic.[38] Unsurprisingly, these assumptions yield a similar spatial distribution of traffic reductions to baseline results but the aggregate traffic decline is an order of magnitude smaller. As a result, appendix A.2.6 reports a small decrease in commuters' driving rates and a £70 million welfare gain, about half of the preferred estimate.[39] Encouragingly, this alternative simulation resembles baseline results in terms of the spatial distributions of employment and population and agrees that the policy is progressive.

# 9 Conclusion

This paper studies the effects of Central London's Congestion Charge on regional traffic, commuting, and labour market outcomes. To this end, I develop a method to measure

[38] Specifically, appendix A.2.6 preforms simulations where $\varphi^h = 0$ and all other parameters are the same. CCE is still used for estimation but now plays no role in determining counterfactual commute costs.

[39] Appendix table A.17 presents 0.054 percent increase in expected utility with toll and rent recycling. Applying the calculation from section 8.1 to this smaller welafre gain yields approximately £70 million.

effects of a centralized cordon fee on traffic patterns throughout the city and use a general equilibrium model for policy analysis. In the process, I present reduced-form evidence that Central London's Congestion Charge reduced traffic on roads leading downtown and use this variation to estimate elasticities of commuters' travel demand. These elasticities then calibrate a general equilibrium model with heterogeneous worker skills, productivity externalities, mode choice, and endogenous traffic externalities to estimate the aggregate effects of London's Congestion Charge.

My results show that tolling drivers to enter a regional employment centre generates widespread traffic reductions that benefit commuters in the long-run. I also find that exogenously reducing traffic between neighbourhoods attracts additional commuters, increases their odds of driving over taking transit, and that sorting across residences and workplaces has quantitatively meaningful implications. In fact, I find that accounting for traffic spillovers across routes and central workplace's skill-bias makes cordon fees progressive policies.

All things considered, I find that London's Congestion Charge gives commuters £140 million in annual welfare gains and disproportionately benefits the region's low-skill workers. These gains add to improved downtown amenities valued by Tang (2018) and considering the policy's reasonable costs makes similar, or even more ambitious, cordon fees an attractive option for other centralized cities looking to alleviate congestion's costs.

## References

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., and Wolf, N. (2015). The economics of density: Evidence from the Berlin wall. *Econometrica*, 83(6):2127–2189.

Akbar, P. and Duranton, G. (2017). Measuring the cost of congestion in highly congested

city: Bogotá.

Allen, T. and Arkolakis, C. (2020). The welfare effects of transportation infrastructure improvements. Technical report, National Bureau of Economic Research.

Allen, T., Arkolakis, C., and Li, X. (2015). On the existence and uniqueness of trade equilibria. *Manuscript, Yale Univ.*

Arnott, R., De Palma, A., and Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, pages 161–179.

Arzaghi, M. and Henderson, J. V. (2008). Networking off madison avenue. *The Review of Economic Studies*, 75(4):1011–1038.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Baum-Snow, N., Kahn, M. E., and Voith, R. (2005). Effects of urban rail transit expansions: Evidence from sixteen cities, 1970-2000 [with comment]. *Brookings-Wharton papers on urban affairs*, pages 147–206.

Berge, L. et al. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. Technical report, Center for Research in Economic Analysis, University of Luxembourg.

Brinkman, J. C. (2016). Congestion, agglomeration, and the structure of cities. *Journal of Urban Economics*, 94:13–31.

Card, D. (2009). Immigration and inequality. *American Economic Review*, 99(2):1–21.

Cheshire, P. C. and Hilber, C. A. (2008). Office space supply restrictions in Britain: the political economy of market revenge. *The Economic Journal*, 118(529):F185–F221.

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

Dekle, R., Eaton, J., and Kortum, S. (2007). Unbalanced trade. *American Economic Review*, 97(2):351–355.

Dingel, J. I. and Tintelnot, F. (2020). Spatial economics for granular settings.

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from us cities. *American Economic Review*, 101(6):2616–52.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Ecola, L. and Light, T. (2009). Equity and congestion pricing. *Rand Corporation*, pages 1–45.

Fajgelbaum, P. D. and Schaal, E. (2017). Optimal transport networks in spatial equilibrium. Technical report, National Bureau of Economic Research.

Gibson, M. and Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, 89:62–73.

Gonzalez-Navarro, M. and Turner, M. A. (2018). Subways and urban growth: Evidence from Earth. *Journal of Urban Economics*, 108:85–106.

Google (2020). Google API Terms of Service | Google Developers. `https://developers.google.com/terms`.

Greater London Authority (1999). Greater London authority act 1999.

Green, C., Heywood, J. S., and Navarro Paniagua, M. (2018). Did the London congestion charge reduce pollution?

Green, C. P., Heywood, J. S., and Navarro, M. (2016). Traffic accidents and the London congestion charge. *Journal of Public Economics*, 133:11–22.

Hall, J. D. (2020). Can tolling help everyone? estimating the aggregate and distributional consequences of congestion pricing. *Journal of the European Economic Association*.

Hanna, R., Kreindler, G., and Olken, B. A. (2017). Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta. *Science*, 357(6346):89–93.

Heblich, S., Redding, S. J., and Sturm, D. M. (2020). The making of the modern metropolis: evidence from London.

Herzog, I. (2021a). *The Economic Effects of Transportation Policy*. PhD thesis.

Herzog, I. (2021b). The marginal external cost of traffic in Greater London.

Hilber, C. A. and Vermeulen, W. (2016). The impact of supply constraints on house prices in England. *The Economic Journal*, 126(591):358–405.

Jerch, R., Barwick, P., Li, S., Wu, J., et al. (2020). Road rationing policies and housing markets. Technical report, Department of Economics, Temple University.

Kreindler, G. E. (2018). The welfare effect of road congestion pricing: Experimental evidence and equilibrium implications. *Unpublished paper*.

Leape, J. (2006). The London congestion charge. *Journal of Economic Perspectives*, 20(4):157–176.

Lee, K. (2019). Benefits of working near college workers: Human capital spillovers across neighborhoods.

Luxen, D. and Vetter, C. (2011). Real-time routing with OpenStreetMap data. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '11, pages 513–516, New York, NY, USA. ACM.

Martin, L. A. and Thornton, S. (2017). Can road charges alleviate congestion? *Available at SSRN 3055428*.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4):303–328.

Monte, F., Redding, S. J., and Rossi-Hansberg, E. (2018). Commuting, migration, and local employment elasticities. *American Economic Review*, 108(12):3855–90.

Office for National Statistics (2012). Annual survey of hours and earnings: 2011. Technical report.

Parry, I. W. (2009). Pricing urban congestion. *Annu. Rev. Resour. Econ.*, 1(1):461–484.

Proost, S. and Thisse, J.-F. (2019). What can be learned from spatial economics? *Journal of Economic Literature*, 57(3):575–643.

Rivers, D. and Vuong, Q. H. (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of econometrics*, 39(3):347–366.

Sanderson, E. and Windmeijer, F. (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics*, 190(2):212–221.

Severen, C. (2019). Commuting, labor, and housing market effects of mass transportation: Welfare and identification.

Small, K. A., Verhoef, E. T., and Lindsey, R. (2007). *The economics of urban transportation*. Routledge.

Smeed, R. (1964). Road pricing: the economic and technical possibilities. *The Smeed Report, Ministry of Transport. London*.

Stillwell, J. and Duke-Williams, O. (2007). Understanding the 2001 UK census migration and commuting data: the effect of small cell adjustment and problems of comparison with 1991. *Journal of the Royal Statistical Society: Series A (statistics in society)*, 170(2):425–445.

Tang, C. K. T. (2018). The cost of traffic: Evidence from the London congestion charge. Technical report.

Transportation for London (2003). Congestion charging: Six months on. Technical report.

Transportation for London (2005). Central London congestion charging scheme impacts monitoring. Technical report.

Tsivanidis, N. (2019). The aggregate and distributional effects of urban transit infrastructure: Evidence from Bogotá's transmilenio. *Unpublished manuscript*.

Tyndall, J. (2018). The local labour market effects of light rail transit.

Verhoef, E. T. (2005). Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics*, 58(3):367–388.

Vickrey, W. (1967). Optimization of traffic and facilities. *Journal of Transport Economics and Policy*, pages 123–136.

Wardrop, J. G. (1952). Road paper. some theoretical aspects of road traffic research. *Proceedings of the institution of civil engineers*, 1(3):325–362.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Yang, J., Purevjav, A.-O., and Li, S. (2020). The marginal cost of traffic congestion and road pricing: Evidence from a natural experiment in Beijing. *American Economic Journal: Economic Policy*, 12(1):418–53.

Zárate, R. D. (2019). Factor allocation, informality and transit improvements: Evidence from Mexico City.

Zhang, W. and Kockelman, K. M. (2016). Optimal policies in cities with congestion and agglomeration externalities: Congestion tolls, labor subsidies, and place-based strategies. *Journal of Urban Economics*, 95:64–86.

# 10 Tables

Table 1: Commute flows by mode of transportation

| | 2001 | 2011 | log-change |
|---|---|---|---|
| Entire London Region | | | |
| Driver share | 0.362 | 0.254 | -0.354 |
| Other share | 0.638 | 0.746 | 0.156 |
| Total flows | 2,447,652 | 2,936,727 | 0.182 |
| Within Outer London | | | |
| Driver share | 0.627 | 0.445 | -0.343 |
| Other share | 0.373 | 0.555 | 0.397 |
| Total flows | 792,876 | 1,205,940 | 0.419 |
| Outer London flow share | 0.324 | 0.411 | 0.237 |

Top panel summarises commute flows between all MSOA pairs in the London Region and the bottom panel summarises flows for pairs with both MSOAs in Outer London boroughs listed in appendix A.1.1.

Table 2: Workplaces employment shares

| | High-skill | Low-skill | Driver | Area |
|---|---|---|---|---|
| Congestion Charge Zone | 0.33 | 0.15 | 0.06 | 0.01 |
| Adjacent | 0.31 | 0.31 | 0.24 | 0.19 |
| Outer London | 0.36 | 0.53 | 0.70 | 0.79 |
| Total | 0.826 | 0.174 | – | – |

Columns present shares of Greater London's high-skill employment, low-skill employment, employees driving to work, and land area in each city section as of the 2011 census. The final row shows the entire region's skill and mode shares in workplace counts.

Table 3: Congestion Charge Exposure and traffic

| | *Dependent variable:* | | |
|---|---|---|---|
| | ln(Traffic) | | |
| | (1) | (2) | (3) |
| CCE × Post | −0.042*** | −0.040*** | −0.026*** |
| | (0.004) | (0.003) | (0.002) |
| Dist × Post FE | No | Yes | Yes |
| Pow/Por × Post FE | None | Both | Both |
| Controls | No | No | Yes |
| Observations | 1,709,220 | 1,709,220 | 1,709,220 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Each column presents results of an OLS regression using 2001 total flows plus one as weights and all regressions include pair fixed effects. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. CCE is scaled to have a standard deviation of one.

Table 4: Traffic's effect on the number commuting

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | ln(Number of commuters) | | | |
| | (1) | (2) | (3) | (4) |
| ln(Traffic) | $-0.096^{***}$ | $-0.966^{***}$ | $-0.619^{***}$ | $-0.918^{***}$ |
| | (0.026) | (0.260) | (0.121) | (0.188) |
| Dist × Post FE | Yes | No | Yes | Yes |
| Pow/Por × Post FE | Both | None | Both | Both |
| Controls | Yes | No | No | Yes |
| Excl. IV | None | CCE × Post | CCE × Post | CCE × Post |
| First Stage F-stat | None | 106.09 | 224.54 | 274.51 |
| Observations | 1,709,220 | 1,709,220 | 1,709,220 | 1,709,220 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

Each column presents results of an OLS or TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. The dependent variable is the natural logarithm of total flows plus one. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. First stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

Table 5: Traffic's effect on mode shares

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | ln(Drivers) - ln(Non-drivers) | | | |
| | (1) | (2) | (3) | (4) |
| ln(Traffic) | 0.031 | 0.894** | −1.019*** | −1.208*** |
| | (0.054) | (0.449) | (0.219) | (0.352) |
| Dist × Post FE | Yes | No | Yes | Yes |
| Pow/Por × Post FE | Both | None | Both | Both |
| Controls | Yes | No | No | Yes |
| Excl. IV | None | CCE × Post | CCE × Post | CCE × Post |
| First Stage F-stat | None | 106.09 | 224.54 | 274.51 |
| Observations | 1,709,220 | 1,709,220 | 1,709,220 | 1,709,220 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Each column presents results of an OLS or TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. The dependent variable is the natural logarithm of drivers plus one minus the natural logarithm of commuters by other modes plus one. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. First stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

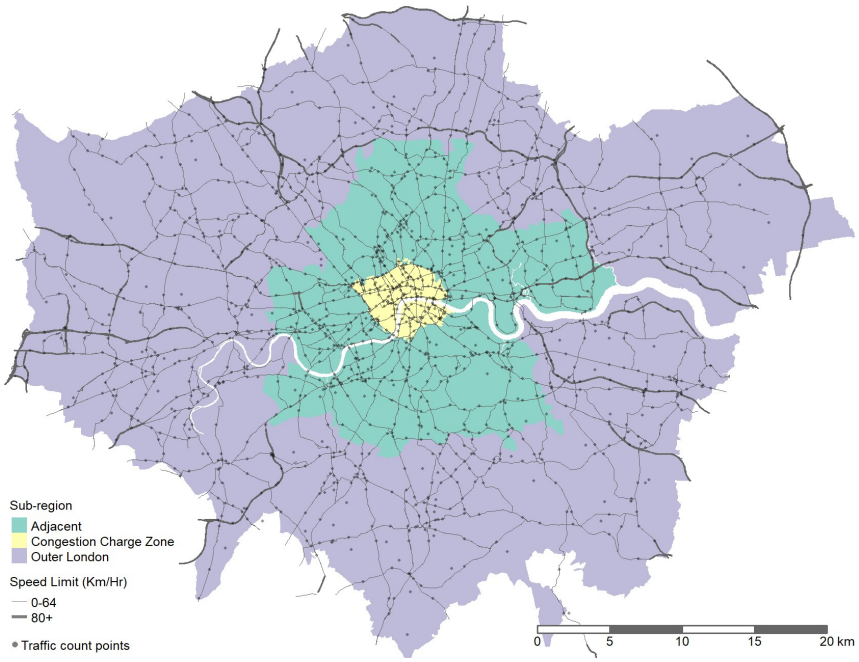
Table 6: Aggregate effects of the congestion charge.

| | | | Endogenous adjustment mechanisms | | | |
|---|---|---|---|---|---|---|
| | | | (1) None | (2) Sorting, wages, and rents | (3) Full model exog. TFP | (4) Full model endog. TFP |
| (1) | Utility | Mean | 0.183 | 0.147 | 0.120 | 0.110 |
| (2) | | Low ; High | 0.223 ; 0.174 | 0.22 ; 0.129 | 0.181 ; 0.106 | 0.184 ; 0.092 |
| (3) | | Inequality | -0.049 | -0.090 | -0.075 | -0.091 |
| (4) | | Toll revenue | 0.110 | 0.110 | 0.110 | 0.110 |
| (5) | | Rent | – | -0.053 | -0.045 | -0.061 |
| (6) | | Car share | – | 0.860 | 0.517 | 0.531 |
| (7) | | Traffic | -1.190 | -1.190 | -1.011 | -1.007 |

Entries are percent difference between actual and counterfactual outcomes, toll revenue and rents are measured as percent change in utility from recycling, and each column is a unique counterfactual simulation. All simulations use the same calibrated parameters and fundamentals with toll revenues and rents recycled into uniform commute cost reductions.

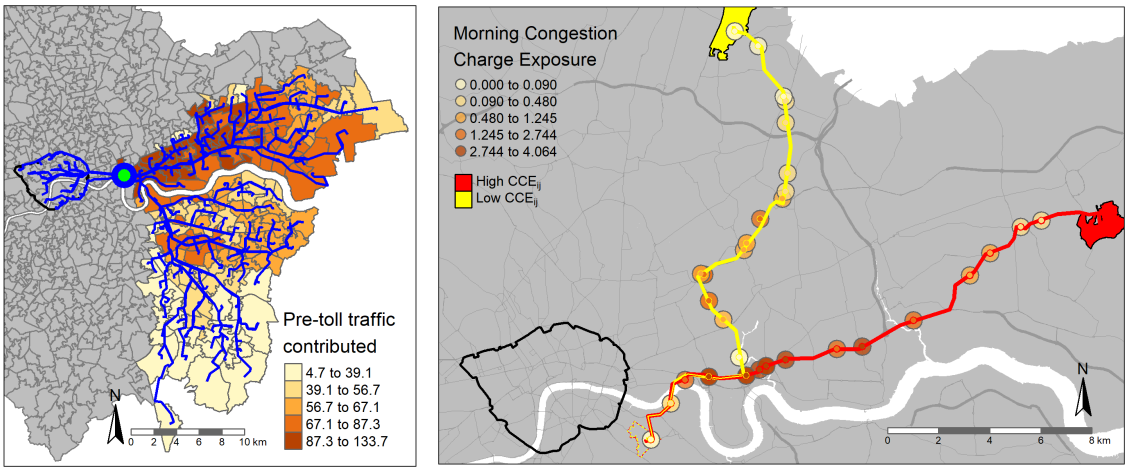
# 11 Figures

Figure 1: The London Region




Points are DfT traffic count points in the balanced panel and lines are a simplified representation of region's major roads taken from the Geofabrik OpenStreetMaps repository (side roads not shown).

Figure 2: Computing Congestion Charge Exposure

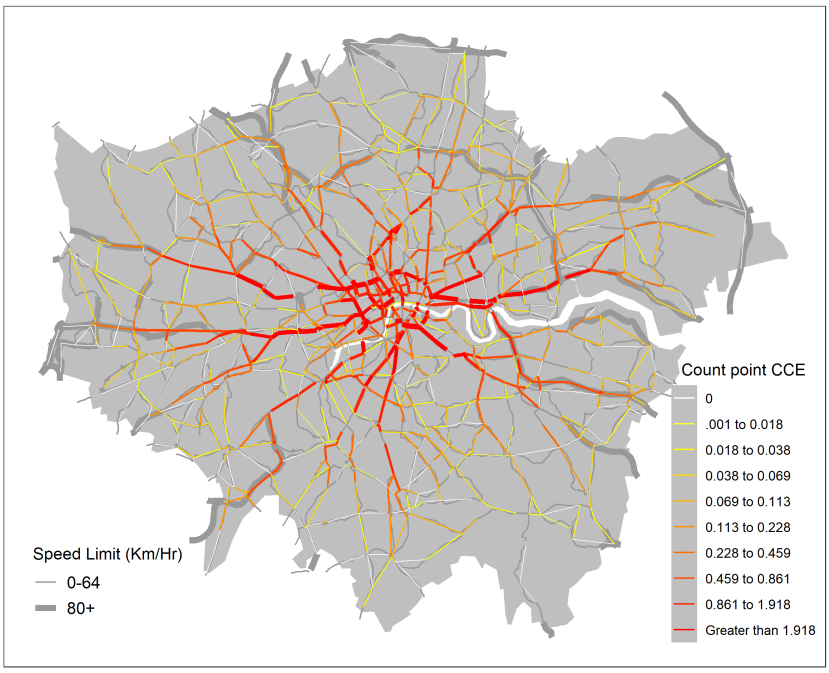



(a) Road link exposure

(b) Variation across commutes

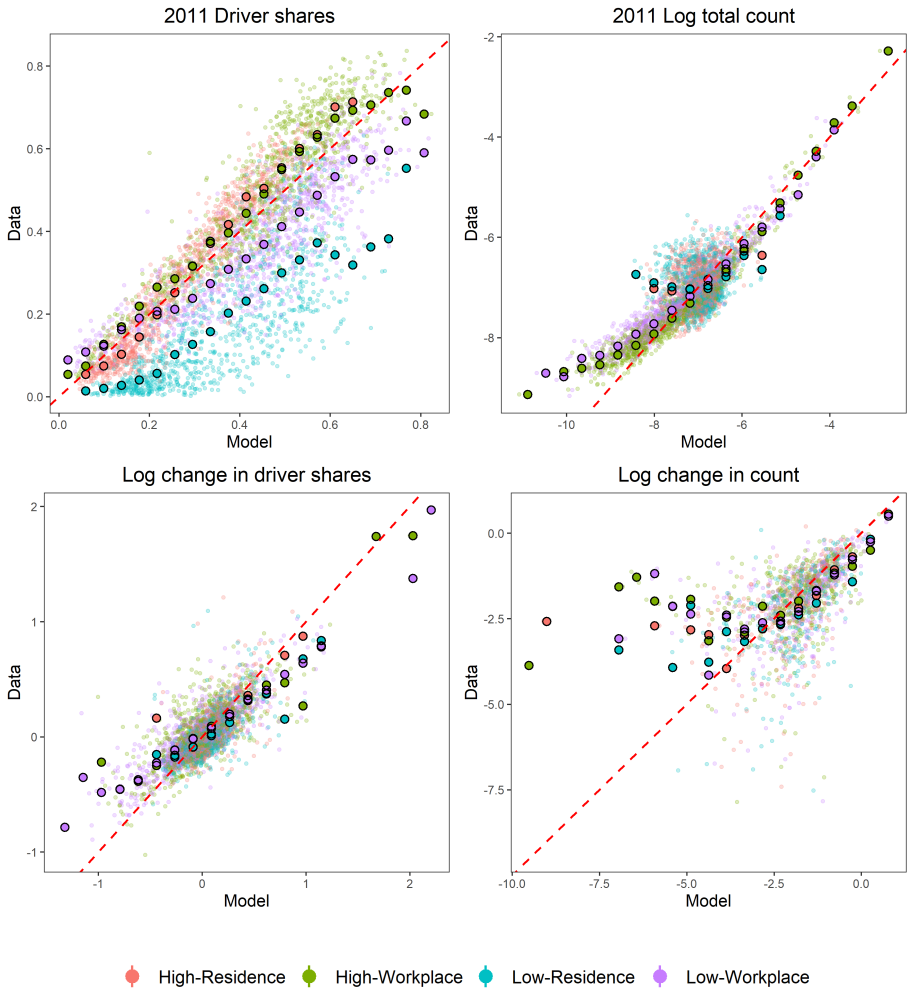
The left panel demonstrates calculating Congestion Charge Exposure for the green highlighted road link where blue paths are OSRM routes contributing to Congestion Charge Exposure and MSOA colours map to quintiles of the amount of morning traffic crossing this link attributed to each origin. The right panel demonstrates the source of cross-commute variation in Congestion Charge Exposure with links' representative points shaded in proportion to their CCE and major roads shown as grey lines with thickness proportional to their speed limits.

Figure 3: The Congestion Charge Exposure Index



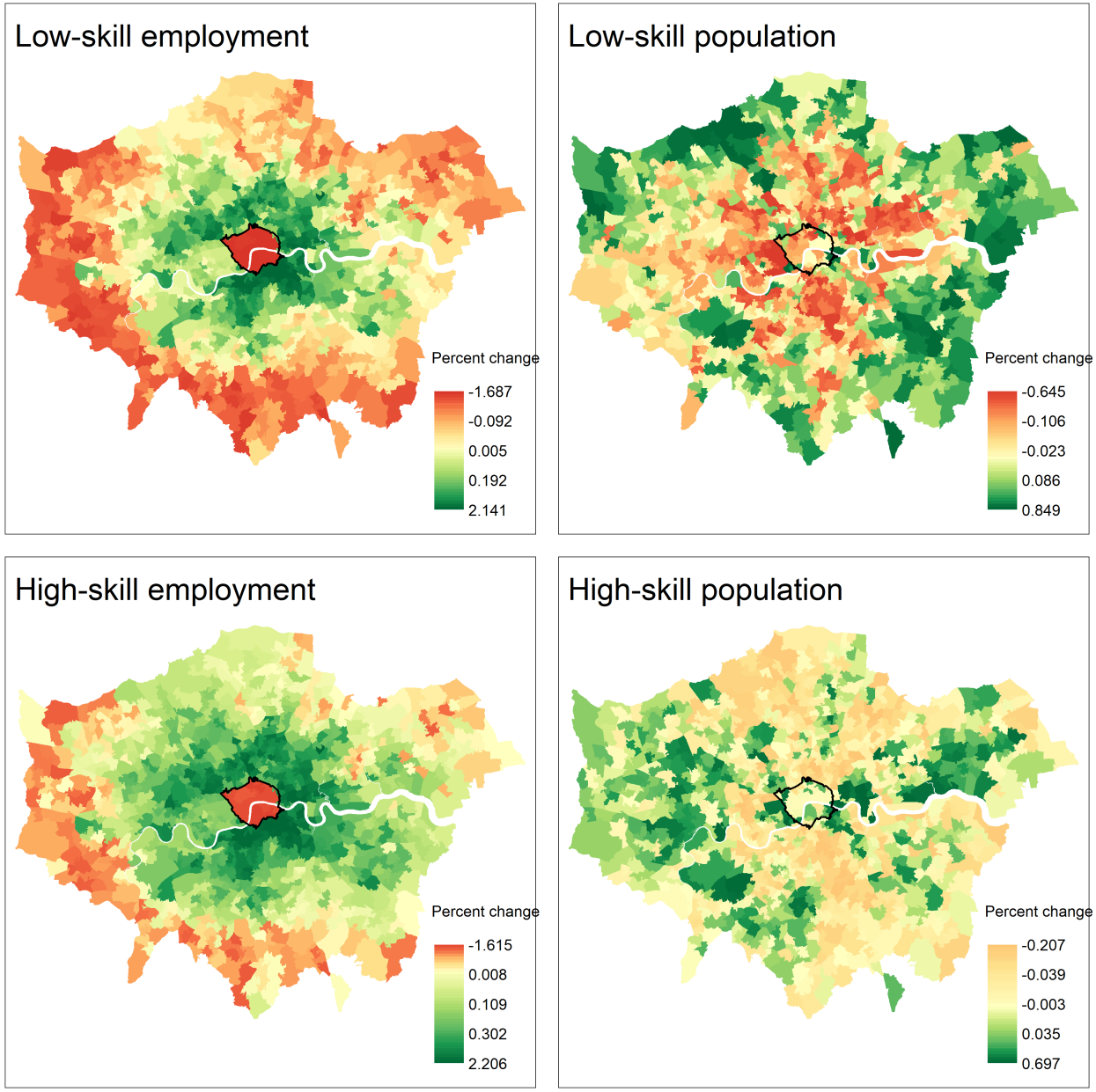

Line color and thickness are both mapped to CCE averaged across travel directions and rush hours at each count point. Count points appear as linear representations of the road links they monitor and grey lines are a subset of the major road network.

Figure 4: Model predicted and actual MSOA outcomes.




Small points plot observed and simulated driver shares and natural logarithm of total counts and large points are binned means. Panel A presents predictions and data for 2011 and changes in panel B are normalized so that there is no total population growth over time.

Figure 5: Percent change employment and population caused by the congestion charge



# A Appendices

## A.1 Data and econometric appendix

This section describes details of the data construction and additional econometric results. First, appendix table A.1 presents summary statistics of MSOA pair commute and traffic data for the main analysis sample. The table summarises raw counts of drivers and non-drivers and transforms their by first adding one to all counts as in linear regressions of sections 5.5 and 7. An MSOA pair's ln(traffic) is the natural logarithm of the mean of its morning and evening traffic.

### A.1.1 Additional data sources

The study area is the London census region which is comprised of 33 boroughs. Following the Office for National Statistics, Inner London is MSOAs in the City of London, Camden, Hackney, Hammersmith and Fulham, Haringey, Kensington and Chelsea, Islington, Lambeth, Lewisham, Newham, Southwark, Tower Hamlets, Wandsworth, and Westminster; Outer London is the remaining 19 boroughs. MSOA shapefiles and cross-walks to boroughs, to OAs, and across time come from the Office for National Statistics Open Geography Portal. I define constant boundary MSOAs by aggregating splitters and mergers to the union of their 2001 and 2011 geographies.

Shapefiles describing Congestion Charge and Western Extension Zone boundaries come from Transportation for London. I define an MSOA as inside the CCZ or WEZ if it has a non-empty intersection with the relevant charge zone with the exception of ONS code E02000967 (contains Regent's park), E02000890, E02000878, and E02000371 which I exclude from the CCZ. Rare cases of constant boundary MSOAs intersecting both the CCZ and the WEZ default to inclusion in the CCZ and not the WEZ; the only exception is ONS

code E02000978 (contains Hyde Park and defined in the WEZ not the CEZ).

Bus routes come from Transportation for London's Open Data Portal as of September 2019. The data contain the locations of all stops in the London Buses network and defines routes as sequences of stops. I identify express bus routes as those with names containing a capital "X" and identify night bus routes as those with names containing a capital "N."

My rail transit database builds on stop locations released by Transportation for London in response to public data request FOI-1451-1819. I augment these data by manually characterizing opening dates, stop locations, and routes of the London Underground, Overground, Thameslink, TfL Rail, Docklands Light Rail (DLR), and Tramlink systems. I define heavy rail transit as all rail systems operating entirely in exclusive rights of way to separate differentiate and Tramlink, a suburban light rail system using a mixture of dedicated rights of way and at-grade track.

Baseline controls exclude rail stations built after 2001, but several stations and lines opened during my study period. This includes Wood Lane tube station and substantial expansions to Overground and Dockland's Light Railway systems. Fortunately, reduced-form results are robust to controlling for post-2001 Overground, DLR, Underground extensions.

### A.1.2 Commute flows

For census each year (2001 and 2011) I create two 967 by 967 commute flow matrices for constant boundary MSOAs in Greater London.

The raw data are measured for census blocks known as output areas (OAs) and workplace zones (WZs). In each year, I obtain three commute flow matrices that each describe one of three mode choices: drives car or van, other (includes transit, car passenger, walk, etc), and works from home. The universe of interest is Greater London residents aged 16 and over in employment the week before the census in 2011. The 2001 census set an upper

age bound of 74, but this discrepancy over time should be unimportant in practice. In 2001, the data are square matrices of output area to output area flows and in 2011 the workplaces are given by workplace zones instead of output areas.[40] I aggregate flows to constant boundary MSOAs before merging the data over time.[41]

The 2001 census randomly adjusts observations of 1 and 2 commuters to 0 and 3 at a highly granular level before aggregating to three modes of transportation and releasing the data. The 2011 data are not adjusted and are instead available for less granular mode categories. Therefore, 2001 data available for research contain synthetic counts of zero and three so that naively aggregating to MSOAs and merging over time creates synthetic multiples of three that are inconsistent with 2011 data.

To make the data comparable over time, I adjust 2011 counts to match the 2001 census small cell adjustment method and bin both years to multiples of three before aggregating to MSOA pairs. The adjustment procedure is as follows:

1. Randomly adjust observations of 1 and 2 to 0 and 3 in 2011 data using the adjustment probability matrix proposed by Stillwell and Duke-Williams (2007):

| | 0 | 3 |
|---|---|---|
| 1 | 2/3 | 1/3 |
| 2 | 1/3 | 2/3 |

   where cell $ij$ is the probability the value in row $i$ is adjusted to the value in column $j$. For example, a cell with one commuter has a $2/3$ probability of being adjusted to zero.

2. Bin both years' data around multiples of three using the function $bin(x) = 3 \times$

[40] Output areas are defined to contain 40 to 125 households and the newer workplace zone geography is designed to contain a roughly constant number of workers.

[41] 2011 data are from census table WU03BUK and 2001 are from the level 3 special workplace statistics.

$round(x/3)$ where *round* rounds to the nearest integer.

3. Aggregate both years to constant boundary MSOAs.

4. Create internally consistent commute flows by aggregating across modes of transportation for each MSOA pair.

Throughout, I omit commuters who report working from home or within their residence MSOA. The result is two comparable flow matrices per year with zeros along the diagonals.

Table A.11 presents the share of London Region employment captured by my adjusted MSOA pair commuting data in each year to assess two sources of measurement error. First, commute flows mechanically undercount employment by excluding workers who live outside of the London Region. Second, small cell adjustments produce random measurement error which should cancel out in aggregate. Commute flows account for 72 percent of regional employment in 2001 and 75 percent in 2011. Undercounting is more prominent for drivers and roughly constant across years, suggesting that it is driven by unobserved inflows to the region rather than unexpected small cell adjustment errors.

### A.1.3 Validating routing assumptions

I compute paths between MSOA pairs using OSRM version 5.22.0, which minimizes driving times along an uncongested road network accounting for turn restrictions, one-way roads, speed limits etc. However, these paths may deviate from drivers' preferred routes in rush-hour traffic. Routes extracted from Google Maps assuming typical rush-hour traffic would be ideal but their terms of service forbid scraping Google's routing Application Programming Interface (API) (Google, 2020). This section uses a limited sample of fastest-in-traffic paths to validate the assumptions implicit in OSRM routing.

I benchmark OSRM paths against a non-random sample of fastest-in-traffic paths be-

tween MSOA pair centroids obtained from the API of an authoritative mapping and routing service during January 2020. These are travel-time minimizing paths given traffic predicted on January 2020 for a 09:00 arrival time on an upcoming October Wednesday. I use these paths to define fastest-in-traffic routes, pairwise traffic, monitored road length, and I trim outliers as in the main text for morning rush hour in the pre- and post-toll periods.

Figure A.7 plots pairwise traffic computed with OSRM routes against fastest-in-traffic counterparts in levels and log-changes for the 36,565 MSOA pairs covered by the latter data. Small points are MSOA pairs, large points are binned means, and the red 45 degree line indicates a perfect fit. Prominent diagonals in both plots suggest a good fit but OSRM slightly under-estimates traffic at low volumes.

Table A.12 shows that corresponding ordinary and weighted least squares regressions of fastest-in-traffic routed against OSRM traffic have slopes near one and an $R^2$ of 0.690 in levels. The fit is slightly worse in changes which give a slope of 0.76 and $R^2$ of 0.566. Adding origin, destination, and distance band fixed effects attenuates this slope to 0.57 but increases the residual $R^2$ and the final column weights by one plus baseline total commuting to find a similar slope and a residual $R^2$ of 0.697. Finally, table A.13 shows that similar patterns emerge in regressions using monitored road length in place of traffic.

All told, these patterns suggest that OSRM routes give a realistic view of the traffic commuters face. In hindsight, this seems reasonable given that *Wardrop's First Principle* predicts that identical drivers will arbitrage away differences in travel costs across routes serving an origin-destination pair (Wardrop, 1952; Small et al., 2007).

### A.1.4 Accounting for traffic

It is tempting to decompose road link traffic between commuters and other traffic, but my data are unsuited for this purpose. In particular, the accounting identity that theoretically

separates endogenous and exogenous traffic (see section 6.4) does not hold for all road links in my data. This appendix compares the distributions of traffic and commuting and argues that incompatibilities stem from random measurement errors that do not bias aggregate statistics or this paper's central results.

Without measurement error, the number of car commuters crossing road link $\ell$ in hour $h$ is $n_\ell^h = \sum_{i,j} 1\{\ell \in R_{ij}^h\} n_{ijc}$ and daily mean traffic is $traffic_\ell = \frac{1}{H} \sum_h (n_\ell^h + \mathcal{O}_\ell^h)$ with $H$ hours each day and $\mathcal{O}_\ell^h$ non-commuters each hour. This accounting identity is broken by measurement error in $traffic_\ell$, routes $R_{ij}^h$, and commute flows $n_{ijc}$. First, I only observe traffic from 7 am to 7 pm and cannot guarantee that commuters drive during my specified rush hour periods. Second, I cannot guarantee that all commuters use the OSRM's optimal routes. Third, enumeration error and daily idiosyncrasies in commuters' decisions affect traffic counts. Fourth, there is rounding in commute flow data.[42] Fortunately, all of these errors are plausibly random across road links.

Figure A.8 plots road link $traffic_\ell$ against estimated car commuters $\frac{1}{H} \sum_h \sum_{i,j} 1\{\ell \in R_{ij}^h\} n_{ijc}$, large points are binned means by year and red dashes denote the 45 degree line. While there is considerable variation around the conditional expectation function, conditional mean traffic tends to exceed car commuting with the exception of a small number of high outliers in estimated commute flows. Ignoring outliers also reveals that traffic and commuting are positively correlated throughout the distribution, providing support for the proposition that measurement error is random across road links.

Figure A.9 presents pre- and post-toll empirical cumulative density functions of road link traffic and commuter estimates. In total, car commuters account for 79 percent of 2001 traffic and 66 percent of 2011 traffic. Encouragingly, the distributions across road links show that total traffic stochastically dominates estimated commuters, leaving a role

[42] Section A.1.2 discusses rounding in commute flows.

for non-commuter traffic at all points in the distribution.

### A.1.5 Regional traffic declines

Figure A.10 maps Greater London's widespread traffic decline over the 2000s. Each line segment is the roadway monitored by a single count point, links intersecting London's Outer Ring Road are highlighted in blue, line colours denote deciles of log-traffic change, and thickness is proportional to traffic change's distance from zero. London's median count point saw traffic fall by approximately ten percent in the years following the toll's introduction. Traffic declined most on central roads and suburban links on radial roads. Meanwhile, traffic grew in the eastern suburbs and on major circumferentials including London's North Circular Road.

To provide suggestive evidence that Central London's toll shifted traffic away from CCZ-bound roads and to motivate the main paper's systematic evaluation, I define radial road links as those not intersecting the Outer Ring Road and run the event study regression:

$$ln(traffic_{\ell t}) = \sum_{\substack{\tau=2000,\\ \tau\neq 2002}}^{2015} D_\tau^t \times \beta_\tau radial_\ell + \sum_{\substack{\tau=2000,\\ \tau\neq 2002}}^{2015} D_\tau^t \times f_\tau(lat_\ell, lon_\ell) + \alpha_t + \alpha_\ell + e_{\ell t} \quad (27)$$

where $traffic_{\ell t}$ is rush-hour traffic volume observed on link $\ell$ in year $t$ from 2000 to 2015, $radial_\ell$ is indicates links that are not on the Outer Ring Road, $D_\tau^t = 1\{\tau = t\}$ are annual dummies excluding 2002, $f_\tau(lat_\ell, lon_\ell)$ is a year-specific second-order polynomial in longitude and latitude, and $\alpha_t$ and $\alpha_\ell$ are year and link fixed effects. The semi-elasticities $\beta_\tau$ measure traffic trends on radial roads relative to nearby links on the Outer Ring Road. Section 4 of the main paper shows that many radial roads are on important routes to the CCZ, suggesting that negative values of $\beta_\tau$ after 2002 indicate Central London's Congestion Charge reduced radial road traffic.

Figure A.11 presents ordinary least squares (OLS) estimates of $\beta_\tau$ normalized to zero in 2002—the final year without Central London's toll—error bars denote link-clustered 95 percent confidence intervals. The plot shows that the Outer Ring Road was on a similar traffic trend to nearby radial roads before 2003 and that relative declines in radial-road traffic emerge after Central London's toll began. This pattern persists for several years and a corresponding difference-in-differences estimate shows that radial-road traffic fell by 6.8 percent relative to ring-road traffic in the post-toll period. Increasingly negative $\beta_\tau$ estimates after 2006 may reflect the 2005 CCZ toll increase and the Western Extension Zone charge that was active from 2007 to 2011.[43] The main paper captures this trend with a continuous measure of Congestion Charge Exposure and an identification strategy that provides direct evidence that Central London's toll affected regional traffic patterns.

### A.1.6 Maximum likelihood commuting elasticities

This appendix describes estimates of traffic's effect on commuting using the high dimensional poisson maximum likelihood estimator (MLE) implemented by Berge et al. (2018) in concert with the two-step control function instrumental variables (CFIV) estimator proposed by Rivers and Vuong (1988).

The estimating equation for location choice is

$$\pi_{ijt} = \exp\left(\delta^n ln(traffic_{ijt}) + X'_{ij}\beta_t + \theta_{b(ij)t} + \alpha_{it} + \alpha_{jt} + \alpha_{ij} + \psi\hat{\nu}_{ijt} + u_{ijt}\right)$$

where $\hat{\nu}_{ijt}$ is the first stage residual fit by applying 2001 flow weighted OLS to equation 3. Mode choice estimates recast equation 11 into the equivalent poisson regression to avoid running a fractional logit regression. Specifically, I lengthen the data so that observations

[43] Appendix figure A.4 shows a similar, albeit less precisely estimated, pattern using a variant of equation 27 saturated with year-specific controls for being in the CCZ, WEZ, or Inner London.

are workplace-residence-year-mode quartets and estimate:

$$\pi_{ijtm} = \exp\Big(\delta^c ln(traffic_{ijt})D_m + X'_{ij}\beta_{tm} + \theta_{b(ij)tm} + \alpha_{itm} + \alpha_{jtm} + \alpha_{ijt} + \alpha_{ijm} + \psi\hat{\nu}_{ijtm} + u_{ijtm}\Big) \tag{28}$$

Table A.14 presents non-linear CFIV estimates of commuting's elasticity of traffic. All regressions are fit by maximum likelihood, the first row presents coefficients on log-traffic, and the second row presents coefficients on the first stage residuals. The first three columns are weighted poisson estimates using 2001 flows as weights and the final column replicates column 3 using a negative binomial model. Standard errors in round parenthesis are clustered by both residence and workplace and square brakets are clustered by pair. Two-way clustered standard errors are robust to arbitrary residual correlation across routes that share an origin or destination to give a high estimate on confidence interval width while pair clustered standard errors give a low estimate.[44]

Poisson estimates support the hypothesis that traffic effects commuting the pattern across specifications mirrors TSLS estimates. However, MLE estimates are smaller and less powerful because the estimator drops observations in clusters that always exhibit zero flows. The weighted poisson estimate in column 3 associates a 3.2 percent decrease in commuting with a ten percent increase in traffic and this effect is statistically significant using pair clustered inference. The negative binomial estimator produces a smaller but more precisely estimated negative effect.

Table A.15 presents similar non-linear CFIV estimates of driver share's elasticity of traffic. Each column is a separate regression based on equation 28 where an observation

[44]CFIV standard errors presented here are not asymptotically valid since computational burden precludes computing bootstrap standard errors recommended in Wooldridge (2010). I present QMLE-style standard errors to give a sense of uncertainty in point estimates. See Berge et al. (2018) for computational details of the high-dimensional fixed effects maximum likelihood estimators.

is a unique combination of an MSOA pair, year, and mode of transportation for a total of 3,435,604 data points. Standard errors in round parenthesis are two-way clustered by residence-mode and workplace-mode and square brackets present pair-mode clustered standard errors.

Poisson estimates of $\delta^c$ omit a substantial number of observations whose cluster never has any drivers. This is particularly problematic because many fixed effects are needed to force identification based on relative mode shares. As a result, MLE mode choice regressions with the full set of fixed effects fail to pick up the variation in the data that delivers a negative elasticity in the linear case.[45]

### A.1.7 Heterogeneous commuting elasticities

To find heterogeneous effects of traffic on commuting, I begin by classifying each MSOA as a low-skill workplace if its low skill employment share is above the sample median. Similarly, low-skill residences are those with above median low-skill population share. I then define low-skill MSOA pairs as those where both the workplace and residence are classified as low-skill.

Table A.16 presents variants of equations 10 and 11 with log-traffic and instruments interacted with dummies indicating low-skill pairs. All regressions are fit by weighted TSLS, include full sets of controls and a low-skill-post toll dummy, columns 3 and 4 include separate workplace and residence-by-year fixed effects for low-skill and non-low-skill pairs. The two-way clustered Sanderson and Windmeijer (2016) weak IV F-statistics in columns 1 and 2 are smaller than in the homogeneous case but the first stage remains strong enough to provide reliable estimates.

[45]The fourth column of table A.15 shows that the negative binomial overdispersion parameter reached its predefined maximum value at optimal parameter values. This suggests no overdispersion in the mode choice regression.

The first two columns of table A.16 suggest that traffic has a larger effect on the location choices of low-skill commuters. Mode choice elasticities are slightly larger between high-skill locations but this difference is not statistically significant. Columns 3 and 4 allow different workplace and residence trends at high- and low-skill MSOA pairs to control for differential wage and amenity growth and give a similar pattern of location effects to columns 1 and 2, but the first stage is weak and results should be interpreted cautiously.

## A.2 Theory and quantitative appendix

### A.2.1 Deriving choice probabilities and proving theorem 1

This appendix provides detailed derivations of commuters choice probabilities under the assumptions described in section 5 and concludes by proving theorem 1.

**Mode choice:** First, define the ex-ante expected mean utility of commuting from $i$ to $j$ for a worker with skill $g$ as $\mu_{ijg} = \frac{w_{jg}B_{ig}}{Q_i^{1-\beta}\bar{d}_{ij}}$ where $\bar{d}_{ij}$ is a commute cost index (often called the inclusive value in nested logit applications) equal to each commute's expected value of $\max_m \frac{\nu_{m\omega}}{d_{ijm}}$. With many workers, mode $m$'s share of skill $g$ choices is the conditional probability of a skill $g$ worker taking mode $m$ after choosing to commute from $i$ to $j$ and

realizing the preference $\nu_{m\omega}$:

$$\begin{aligned}
\pi_{m|ijg} &= Pr\big(u_{ijgm} \geq u_{ijgm'} \ \forall \ m' \neq m\big) \\
&= Pr\big(\frac{\mu_{ijg}}{d_{ijm}}\nu_{m\omega}\epsilon_{ij\omega} \geq \frac{\mu_{ijg}}{d_{ijm'}}\nu_{m'\omega}\epsilon_{ij\omega} \ \forall \ m' \neq m\big) \\
&= Pr\big(\frac{\nu_{m\omega}}{d_{ijm}} \geq \frac{\nu_{m'\omega}}{d_{ijm'}} \ \forall \ m' \neq m\big) \\
&= \int_0^\infty \prod_{m' \neq m} Pr\big(u \geq \frac{\nu_{m'\omega}}{d_{ijm'}}\big) Pr(u = \frac{\nu_{m\omega}}{d_{ijm}})du \\
&= \int_0^\infty \exp\left(-u^{-\eta}\sum_{m' \neq m} d_{ijm'}^{-\eta}\right) \exp\left(-\left(ud_{ijm}\right)^{-\eta}\right) d_{ijm}^{-\eta}\eta u^{-\eta-1}du \\
&= \int_0^\infty \exp\left(-u^{-\eta}\sum_{m'} d_{ijm'}^{-\eta}\right) d_{ijm}^{-\eta}\eta u^{-\eta-1}du.
\end{aligned}$$

To evaluate this integral, define $y = -u^{-\eta}\sum_{m'} d_{ijm'}^{-\eta}$ which implies $dy = \eta u^{-\eta-1}\sum_{m'} d_{ijm'}^{-\eta}du$ or $\eta u^{-\eta-1}du = \frac{dy}{\sum_{m'} d_{ijm'}^{-\eta}}$. This, paired with the monotonic negative relationship of $y$ and $u$, allows the change of variables:

$$\begin{aligned}
\pi_{m|ijg} &= \int_0^\infty \exp\left(-u^{-\eta}\sum_{m'} d_{ijm'}^{-\eta}\right) d_{ijm}^{-\eta}\eta u^{-\eta-1}du. \\
&= \int_{-\infty}^0 e^y \frac{d_{ijm}^{-\eta}}{\sum_{m'} d_{ijm'}^{-\eta}}dy \\
&= \frac{d_{ijm}^{-\eta}}{\sum_{m'} d_{ijm'}^{-\eta}}.
\end{aligned} \tag{29}$$

Since mode choice is independent of skill, the conditional probability of any worker choosing mode $m$ given commute $ij$ is $\pi_{m|ij} = \frac{d_{ijm}^{-\eta}}{\sum_{m'} d_{ijm'}^{-\eta}}$.

**Commute cost indices are the expected costs of each commute:** A Fréchet random variable $x$ with CDF $F(x) = e^{-\left(\frac{x}{s}\right)^{-\eta}}$ has expectation $E(x) = s\Gamma\left(1 - \frac{1}{\eta}\right)$. Given route choice, workers know $\mu_{ijg}\epsilon_{ij\omega}$ and $d_{ijm}$, but must form expectations about transport costs

since they do not know yet know what mode they will choose after learning their $\nu_{m\omega}$. At this stage, inverse shock-inclusive costs of commute $ij$ have the CDF:

$$\begin{aligned} G_{ij}(d^{-1}) &= Pr\left(\frac{1}{d} \geq \frac{\nu_{mw}}{d_{ijm}} \ \forall \ m\right) \\ &= \prod_m Pr\left(\frac{d_{ijm}}{d} \geq \nu_{m\omega}\right) = \exp\left(-d^{\eta} \sum_m d_{ijm}^{-\eta}\right) \\ &= \exp\left(-\left[\frac{d^{-1}}{\left(\sum_m d_{ijm}^{-\eta}\right)^{\frac{1}{\eta}}}\right]^{-\eta}\right) \\ &\implies E(\nu_{m\omega}/d_{ijm}|ij) = \left[\sum_m d_{ijm}^{-\eta}\right]^{\frac{1}{\eta}} \Gamma\left(1-\frac{1}{\eta}\right) \end{aligned}$$

and expected commute costs are

$$\begin{aligned} \bar{d}_{ij} &= E(\nu_{m\omega}/d_{ijm}|ij)^{-1} \\ &= \left(\Gamma\left(1-\frac{1}{\eta}\right)\right)^{-1} \left[\sum_m d_{ijm}^{-\eta}\right]^{-\frac{1}{\eta}} \end{aligned}$$

which are independent of skill.

**Location choice:** Before choosing a mode of transportation, workers jointly choose locations based on location pair shocks $\epsilon_{ijw}$ and ex-ante expected mean utility $\mu_{ijg}$. The share

of skill $g$ workers commuting from $i$ to $j$ is

$$\begin{aligned}
\pi_{ij|g} &= Pr\big(\mu_{ijg}\epsilon_{ijw} \geq \mu_{rsg}\epsilon_{rsw} \ \forall \ rs \neq ij\big) \\
&= \int_0^\infty \prod_{rs \neq ij} Pr\big(u \geq \mu_{rsg}\epsilon_{rsw}\big) Pr(u = \mu_{ijg}\epsilon_{ijw}) du \\
&= \int_0^\infty \prod_{rs \neq ij} F\left(\frac{u}{\mu_{rsg}}\right) \frac{d}{du} F\left(\frac{u}{\mu_{ijg}}\right) du \\
&= \int_0^\infty \exp\left(-u^{-\theta} \sum_{rs \neq ij} \mu_{rsg}^\theta\right) \exp\left(-u^{-\theta}\mu_{ijg}^\theta\right) \mu_{ijg}^\theta \theta u^{-\theta-1} du \\
&= \frac{\mu_{ijg}}{\sum_{rs} \mu_{rsg}} = \frac{(w_{jg}B_{ig}Q_i^{\beta-1})^\theta \bar{d}_{ij}^{-\theta}}{\sum_r \sum_s (w_{sg}B_{rg}Q_r^{\beta-1})^\theta \bar{d}_{rs}^{-\theta}}
\end{aligned} \tag{30}$$

where the final line follows from evaluating the integral as in the mode choice nest.

**Expected utility:** The properties of the Fréchet distribution imply that mean utility of skill $g$ workers is

$$E(u|g) = \gamma'\gamma \left[\sum_i \sum_j \left(\frac{B_{ig}w_{jg}}{\bar{d}_{ij}Q_i^{1-\beta}}\right)^\theta\right]^{\frac{1}{\theta}} \tag{31}$$

and the law of iterated expectations shows that equation 22 is equilibrium expected utility. This lays the necessary foundations for deriving empirical implications.

**Proof of theorem 1:** First note that commute flows from $i$ to $j$ are $\pi_{ij} = \sum_g \pi_g \pi_{ij|g}$. Combining with equations 4, 5, 29, 30, and partially differentiating gives elasticities in equations 8 and 9. □

### A.2.2 Negative commuting elasticity implies general equilibrium effects

In this section, I show that a negative causal effect of commute costs on commuting is necessary and sufficient for an exogenous shock to commute costs to have spatially dependent

general equilibrium effects through wages and rents. It naturally follows that traffic, a determinant of commute costs, has general equilibrium effects if and only if it affects commute flows. This is true in a wide class of models generating a commuter gravity equation.

**Proposition 1** *If commute flows between each location pair $i, j$ in a finite set of locations $\mathcal{I} = \{1, 2, .., N\}$ s.t. $N > 3$ are determined by the following system of equations*

$$\begin{aligned}
\pi_{ij} &= w_j q_i d_{ij}^{\delta} \text{ (gravity)} \\
\pi_i &= \sum_j \pi_{ij} \text{ (housing demand)} \\
\pi_j &= \sum_i \pi_{ij} \text{ (labour supply)} \\
w_j &= A_j \pi_j^{\alpha} \text{ (labour demand)} \\
q_i &= B_i \pi_i^{\beta} \text{ (housing supply)}
\end{aligned}$$

*where $\{d_{ij}\}$ are potentially endogenous commute costs, $\{\pi_{ij}, \pi_i, \pi_j\}$ are endogenous allocations, $\{w_j, q_i\}$ are endogenous prices, $\{\delta, \alpha, \beta\}$ are exogenous parameters satisfying $\delta < 0$, $\alpha \notin \{0, 1\}$, $\beta \notin \{0, 1\}$, and fundamentals $\{A_j, B_i\}$ satisfy $A_j > 0$, and $B_i > 0$ for at least four locations $i, j, r, s \in \mathcal{I}$, then*

$$\delta \neq 0 \iff \exists\ i, j, r, s \in \mathcal{I} \text{ s.t. } \frac{dq_i}{dd_{is}} \neq 0 \text{ and } \frac{dw_j}{dd_{rj}} \neq 0.$$

**Proof:** I prove both necessity and sufficiency by renaming location indices and reducing

the system of equations to

$$w_j = A_j^{\frac{1}{1-\alpha}} \left( \sum_r q_r d_{rj}^{\delta} \right)^{\frac{\alpha}{1-\alpha}}$$
$$q_i = B_i^{\frac{1}{1-\beta}} \left( \sum_s w_s d_{is}^{\delta} \right)^{\frac{\beta}{1-\beta}}.$$

For necessity, setting $\delta = 0$ and combining shows that all $w_j$ and $q_i$ are uniquely determined by the set of fundamentals $\{A_j, B_i\}$ if commute costs have no causal effect on total flows. For sufficiency, note that with $\delta \neq 0$ perturbing any commute cost to $j$, such as $d_{rj}$, will have first-order effects on $w_j$ since $\frac{\alpha}{1-\alpha} \neq 0$.[46] A similar argument applies to the relationship between $q_i$ and any commute cost $d_{is}$. □

**Remark:** Necessity of $\delta \neq 0$ only applies to the case of a shock to pairwise commute costs such as a congestion charge reducing non-commuter traffic. If there were an initial shock to some fundamental $A_j$ or $B_i$, such as reduced noise and air pollution in some location, there could be general equilibrium effects even if $\delta = 0$. However, these effects would be spatially independent in that they affect each neighbourhood identically.

### A.2.3 Equilibrium properties

With exogenous productivity, arguments developed by Tsivanidis (2019) guarantee the existence of at least one labour market equilibrium.[47] The following lemma appropriates this fact for my context.

**Lemma 1** (Existence of a labour market equilibrium with exogenous productivity) *Given commute costs* $\{\bar{d}_{ij}\}$*, parameters* $\{\theta, \sigma, \alpha, \beta\}$*, and fundamentals* $\{B_{ig}, A_j, a_{jg}, \pi_g\}$ *there*

[46] This perturbation has higher order effects via $q_r$, but this is not necessary to prove the current proposition.

[47] Labour market equilibrium's existence follows the proof in an early appendix to Tsivanidis (2019) accessible from the University of Chicago's dissertation archive as DOI: 10.6082/dk7g-zf31.

*exist allocations* $\{\pi_{ijg}\}$ *and prices* $\{w_{jg},\ F_j,\ Q_i\}$ *satisfying the conditions of a closed city labour market equilibrium.*

**Proof:** Equations 15, 16, 17, and 18 can be combined into a continuous and homogeneous of degree one map from choice probabilities onto themselves. Applying Brouwer's fixed point theorem shows that given parameters $\{\theta,\ \eta,\ \sigma,\ \alpha,\ \beta\}$, fundamentals $\{B_{ig},\ A_j,\ a_{jg}\}$, population $\{\Omega,\ \pi_g\}$, and commute costs $\{\bar{d}_{ij}\}$, there exist equilibrium allocations $\{\pi_{ijg}\}$ and prices $\{w_{jg},\ F_j,\ Q_i\}$ satisfying equations 15, 16, 17, and 18. □

Lemma 1 establishes that a labour market equilibrium is determined by location choices, which are a continuous function of commute costs $G(\{\bar{d}_{ij}\})$. I now establish the existence of equilibrium mode shares that pin down traffic and commute costs following from any given labour market equilibrium.

**Lemma 2** (Existence of mode shares and commute costs) *Given location choices* $\{\pi_{ijg}\}$, *parameters* $\{\eta,\ \kappa,\ \varphi^{morn},\ \varphi^{eve}\}$, *fundamentals* $\{\pi_g,\ \Omega,\ \alpha_{ic},\ \alpha_{jc},\ d_{ijo},\ e^h_\ell,\ CCE^h_\ell,\ toll_{ij},\ R^h_{ij}\}$ *and a normalization ensuring* $\varphi^h CCE^h_\ell + e^h_\ell \geq \check{e}_{\ell,h}\ \forall\ \ell, h$, *there exist traffic, mode shares, and commute costs* $\{traffic^h_\ell,\ \pi_{c|ij},\ \bar{d}_{ij}\}$ *satisfying the conditions of a closed city general equilibrium.*

**Proof:** The normalization $\varphi^h CCE^h_\ell + e^h_\ell \geq \check{e}_{\ell,h}$ allows re-writing equation equation 21 as

$$\begin{aligned} ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right) =& \eta\alpha_{ic} + \eta\alpha_{jc} + \eta d_c(ij) - \eta ln(toll_{ij}) + \eta ln(d_{ijo}) \\ & - \eta\kappa ln\left[\frac{1}{2}\sum_h \sum_\ell 1\{\ell \in R^h_{ij}\}\left(\varphi^h CCE^h_\ell + e^h_\ell + \Omega \sum_{r,s} 1\{\ell \in R^h_{rs}\}\pi_{c|rs}\pi_{rs}\right)\right]. \end{aligned} \tag{32}$$

Conditional on $\{\pi_{ijg}\}$ and fundamentals, equation 32 is a continuous function mapping $N^2$ mode shares onto their log-odds. Exponentiating and rearranging to isolate $\pi_{c|ij}$ on the

left hand side produces a continuous mapping from mode shares onto themselves. Since $\pi_{c|ij} \in [0,1]\ \forall\ i,j$, Brouwer's fixed point theorem implies that there exists a vector of mode shares $\{\pi_{c|ij}\}$ satisfying equation 21. Combining with location choices $\pi_{ijg}$ and substituting into equation 20 defines $traffic_\ell^h$ and equation 19 determines $\bar{d}_{ij}$. □

Lemma 2 guarantees the existence of mode shares and traffic externalities consistent with any labour equilibrium and shows that equilibrium commute costs can be written as function of location choices $D(\{\pi_{ijg}\})$ because $\pi_{ij} = \sum_g \pi_{ijg}$. Further inspection of equation 32 shows that $\mathrm{D}(\{\pi_{ijg}\})$ is continuous and combining with lemma 1 gives the equilibrium operator $\ddot{D}(\{\bar{d}_{ij}\}) = \{\bar{d}_{ij}\} - D\big(G(\{\bar{d}_{ij}\})\big)$ which is an $N \times N$ matrix of zeros at equilibrium commute costs $\{\bar{d}_{ij}\}$. The following theorem guarantees that these commute costs exist and characterizes a closed city general equilibrium.

**Theorem 2** (Existence of a general equilibrium with exogenous productivity) *Given {$\theta$, $\sigma$, $\alpha$, $\beta$, $\varphi^{morn}$, $\varphi^{eve}$, $B_{ig}$, $A_j$, $a_{jg}$, $\pi_g$, $\alpha_{ic}$, $\alpha_{jc}$, $d_{ijo}$, $e_\ell$, $CCE_\ell$, $toll_{ij}$, $R_{ij}^h$} satisfying $\varphi^h CCE_\ell^h + e_\ell^h \geq \check{e}_{\ell,h}\ \forall\ \ell$ and $h$, there exist allocations {$\pi_{ijg}$, $\pi_{c|ij}$, $traffic_\ell^h$} and prices {$w_{jg}$, $F_j$, $Q_i$} satisfying the conditions of a closed city general equilibrium.*

**Proof:** Lemmas 1 and 2 show that equilibrium commute costs are a fixed point on the continuous operator $D\big(G(\{\bar{d}_{ij}\})\big)$ and that they pin down endogenous variables $\{\pi_{ijg}, \pi_{c|ij}, traffic_{ij}, w_{jg}, F_j, Q_i\}$. It remains to show that $D\big(G(\{\bar{d}_{ij}\})\big)$ satisfies the conditions of Brouwer's fixed point theorem, which is true if the $\{\bar{d}_{ij}\}$=$D(\cdot)$ given by lemma 2 are confined to a compact and convex set.

Potential commute costs' compactness and convexity turns out to hold without normalization since a fixed population $\Omega$ naturally bounds traffic. To see this, first note that $\bar{d}_{ij} = \gamma^{-1}\left(d_{ijc}^{-\eta} + d_{ijo}^{-\eta}\right)^{-\frac{1}{\eta}}$ is a continuous and increasing function of $traffic_{ij} = \frac{1}{2}\sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} traffic_\ell^h$. Since $\sum_h \sum_\ell traffic_\ell^h \geq \sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} traffic_\ell^h\ \forall\ i,j$, the existence of a finite upper bound on total traffic guarantees a finite upper bound on

each element of the commute cost vector $\{\bar{d}_{ij}\}$.

Total traffic is bounded because it is the sum over a finite number of road links $\ell$, each with a finite maximum possible $traffic_\ell^h$ in each rush hour. Formally, $traffic_\ell^h = \Omega \sum_{i,j} \pi_{ijc} 1\{\ell \in R_{ij}^h\} + \varphi^h CCE_\ell^h + e_\ell^h$ has the finite upper bound $\Omega \sum_{i,j} 1\{\ell \in R_{ij}^h\} + \varphi^h CCE_\ell^h + e_\ell^h$. Since zero trivially bounds $traffic_{ij}$ from below, it follows that the vector $\{\bar{d}_{ij}\}$ is confined to a compact subset of $\mathbb{R}^{N^2}$ and Brouwer's fixed point theorem guarantees the existence of a fixed point on the equilibrium commute cost operator. □

**Multiplicity:** As is common in settings with externalities, I cannot guarantee uniqueness of an equilibrium with exogenous productivity. As a heuristic, it is useful to focus on the mode choice problem that determines equilibrium driver shares and road traffic for a given labour market equilibrium. Abstracting from hours of the day costs no generality and gives

$$ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right) = \eta\alpha_{ic} + \eta\alpha_{jc} + \eta d_c(ij) - \eta ln(toll_{ij}) + ln(\frac{\gamma^\theta}{\Phi}) + \eta ln(d_{ijo}) \tag{33}$$
$$- \eta\kappa ln\left(\sum_\ell 1\{\ell \in R_{ij}\}\left[\varphi CCE_\ell + e_\ell + \sum_{r,s,g} 1\{\ell \in R_{rs}\}\pi_g \Omega (w_{sg} B_{rg} Q_r^{\beta-1})^\theta \pi_{c|rs}(1-\pi_{c|rs})^{\frac{-\theta}{\eta}} d_{rso}^{-\theta}\right]\right)$$

assuming interior solutions on traffic. Equation 33 defines location pair $ij$ car share demand where commute costs are prices and wages and rents are demand shifters. I now show that standard homogeneity and gross substitution conditions cannot guarantee a unique vector of mode shares satisfying equation 32, leaving the possibility that multiple vectors of mode shares and traffic externalities can follow from a labour market equilibrium.

To see the problem at hand, it is useful to re-write equation 32 as an equilibrium system $f : \mathbb{R}_{++}^{N^2} \to \mathbb{R}^{N^2}$ with elements $f_{ij}(\{\pi_{c|rs}\}) = ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right) - \ddot{\Lambda}_{ij}(\{\pi_{c|rs}\})$ where $\ddot{\Lambda}_{ij}(\{\pi_{c|rs}\})$ is the right hand side of equation 32 and an equilibrium requires $f(\{\pi_{c|rs}\}) = 0$. Theorem 2 of Allen et al. (2015) establishes uniqueness of $\{\pi_{c|rs}\}$ under a homogeneity condition and the gross substitutes condition $\frac{\partial f_{ij}}{\partial \pi_{c|rs}} > 0 \ \forall \ r, s \neq i, j$.

I cannot rule out multiple equilibria since $ln\left(\frac{\pi_{c|ij}}{1-\pi_{c|ij}}\right)$ is not a homogeneous function of $\pi_{c|ij}$. However, it is interesting to note that the equilibrium system exhibits (weak) gross substitution. Differentiating equation 32 with respect to drivers' share on an arbitrary commute gives

$$\frac{\partial \ddot{\Lambda}_{ij}}{\partial \pi_{c|rs}} = -\frac{\kappa}{\ddot{\Phi}_{ij}} \frac{1\{\ell \in R_{ij}\}1\{\ell \in R_{rs}\}(w_{sg}B_{rg}Q_r^{\beta-1})^{\theta}}{d_{rso}^{\theta}(1-\pi_{c|rs})^{\frac{\theta+\eta}{\eta}}} \left[\theta\pi_{c|rs} + \eta(1-\pi_{c|rs})\right]$$

$$\text{s.t. } \ddot{\Phi}_{ij} = \sum_{\ell} 1\{\ell \in R_{ij}\}\left[\varphi CCE_{\ell} + e_{\ell} + \sum_{r,s,g} 1\{\ell \in R_{rs}\}\Omega\pi_g(w_{sg}B_{rg}Q_r^{\beta-1})^{\theta}\pi_{c|rs}(1-\pi_{c|rs})^{\frac{-\theta}{\eta}} d_{rso}^{-\theta}\right]$$

which is always negative when routes $R_{ij}$ and $R_{rs}$ intersect.[48] It follows that $\frac{\partial f_{ij}}{\partial \pi_{c|rs}} = -\frac{\partial \ddot{\Lambda}_{ij}}{\partial \pi_{c|rs}} \geq 0 \; \forall \; r,s \neq i,j$, which holds strictly when $R_{ij} \cap R_{rs} \neq \emptyset$.

In conclusion, there can be more than one vector of mode shares consistent with a given set of wages and rents. It follows that these wages and rents can generate more than one vector of endogenous commute costs.

### A.2.4 Counterfactual Procedure

I use hat algebra to compute responses to a counterfactual road tolling policy defined by tolls $toll'_{ij}$, charge exposure $CCE'^{h}_{\ell}$, and transit costs $d'_{ijo}$(Dekle et al., 2007). The model admits multiple equilibria and this procedure gives counterfactual outcomes that are natural progressions from the observed equilibrium.

For any endogenous variable $x$ with counterfactual value $x'$, I solve for relative changes $\hat{x} = x'/x$. In a closed city, counterfactual wages, rents, mode shares, and transportation costs are determined by following system of equations.

[48] $\frac{\partial \ddot{\Lambda}_{ij}}{\partial \pi_{c|rs}}$ must be weakly negative because $\ddot{\Phi}_{ij} > 0$, $\theta > 0$, $\eta > 0$, and $\pi_{c|rs} \in (0,1)$ by definition.

$$\hat{A}_j = \begin{cases} 1 & \text{with exogenous productivity} \\ \left( \frac{\sum_s e^{-\delta d(s,j)} \frac{1}{Area_s} \left( \sum_g a_{jg} (\sum_i \pi_{ijg} \hat{\pi}_{ijg})^{\rho} \right)^{\frac{1}{\rho}}}{\sum_s e^{-\delta d(s,j)} \frac{1}{Area_s} \left( \sum_g a_{jg} (\sum_i \pi_{ijg})^{\rho} \right)^{\frac{1}{\rho}}} \right)^{\lambda} & \text{with agglomeration} \end{cases} \tag{34}$$

$$\hat{w}_{jg} = \hat{A}_j \left( \frac{\sum_i \pi_{ijg} \hat{\pi}_{ijg}}{\sum_i \pi_{ijg}} \right)^{-\frac{1}{\sigma}} \left( \frac{\sum_{g'} a_{jg'} \left( \sum_i \pi_{ijg'} \hat{\pi}_{ijg'} \right)^{\rho}}{\sum_{g'} a_{jg'} \left( \sum_i \pi_{ijg'} \right)^{\rho}} \right)^{\frac{\alpha - \rho}{\rho}} \tag{35}$$

$$\hat{Q}_i = \frac{\sum_j \sum_g w_{jg} \hat{w}_{jg} \pi_{ijg} \hat{\pi}_{ijg}}{\sum_j \sum_g w_{jg} \pi_{ijg}} \tag{36}$$

$$\hat{\pi}_{ijg} = \pi_g \frac{\hat{w}_{jg}^{\theta} \hat{Q}_i^{\theta(\beta-1)} \hat{\bar{d}}_{ij}^{-\theta}}{\sum_{r,s} \pi_{rs|g} \hat{w}_{jg}^{\theta} \hat{Q}_r^{\theta(\beta-1)} \hat{\bar{d}}_{rs}^{-\theta}} \tag{37}$$

$$\hat{\pi}_{ij} = \frac{\sum_g \pi_{ijg} \hat{\pi}_{ijg}}{\sum_g \pi_{ijg}} \tag{38}$$

$$\widehat{traffic}_{ij} = \frac{\sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} \max\left( \varphi^h CCE_\ell'^h + e_\ell^h + \Omega \sum_{r,s} 1\{\ell \in R_{rs}^h\} \pi_{c|rs} \pi_{rs} \hat{\pi}_{c|rs} \hat{\pi}_{rs}, \underline{traffic}_\ell^h \right)}{\sum_h \sum_\ell 1\{\ell \in R_{ij}^h\} traffic_\ell^h} \tag{39}$$

$$\hat{d}_{ijc} = \widehat{toll}_{ij} \left( \widehat{traffic}_{ij} \right)^{\kappa} \tag{40}$$

$$\hat{\pi}_{c|ij} = \frac{\hat{d}_{ijc}^{-\eta}}{(1-\pi_{c|ij})\hat{d}_{ijo}^{-\eta} + \pi_{c|ij}\hat{d}_{ijc}^{-\eta}} \tag{41}$$

$$\hat{\bar{d}}_{ij} = \left[(1-\pi_{c|ij})\hat{d}_{ijo}^{-\eta} + \pi_{c|ij}\hat{d}_{ijc}^{-\eta}\right]^{-\frac{1}{\eta}} \tag{42}$$

Relative commuter welfare, aggregate floorspace rent, toll revenue, and traffic are

$$\widehat{E(u)} = \widehat{G}\widehat{Q}\frac{\sum_g \pi_g \Phi_g^{\frac{1}{\theta}} \left[\sum_i \sum_j \pi_{ij|g} \hat{w}_{jg}^{\theta} \hat{Q}_i^{\theta(\beta-1)} \hat{\bar{d}}_{ij}^{-\theta}\right]^{\frac{1}{\theta}}}{\sum_{g'} \pi_{g'} \Phi_{g'}^{\frac{1}{\theta}}} \tag{43}$$

$$\widehat{G} = \left(\frac{\sum_{ij} \pi'_{ijc} d'_{ijc} + \pi'_{ijo} d'_{ijo}}{\sum_{ij} \pi'_{ijc} \frac{d'_{ijc}}{toll'_{ij}} + \pi'_{ijo} d'_{ijo}}\right)\left(\frac{\sum_{ij} \pi_{ijc} \frac{d_{ijc}}{toll_{ij}} + \pi_{ijo} d_{ijo}}{\sum_{ij} \pi_{ijc} d_{ijc} + \pi_{ijo} d_{ijo}}\right) \tag{44}$$

$$\widehat{Q} = \frac{\sum_j F_j \hat{F}_j + \sum_i H_i Q_i \hat{Q}_i}{\sum_j F_j + \sum_i H_i Q_i} = \frac{\sum_i \sum_j \sum_g w_{jg} \hat{w}_{jg} \pi_{ijg} \hat{\pi}_{ijg}}{\sum_i \sum_j \sum_g w_{jg} \pi_{ijg}} \tag{45}$$

$$\widehat{traffic} = \frac{\sum_h \sum_\ell \left[\varphi^h CCE'^h_\ell + e^h_\ell + \Omega \sum_{r,s} 1\{\ell \in R^h_{rs}\} \pi_{c|rs} \pi_{rs} \hat{\pi}_{c|rs} \hat{\pi}_{rs}\right]}{\sum_h \sum_\ell traffic^h_\ell}. \tag{46}$$

Given the initial shock to commute costs, $\hat{\bar{d}}^0_{ij}$, the following iterative procedure identifies the equilibrium prices and allocations that are closest to the initial equilibrium:

Guess starting values of $\hat{w}^0_{jg}$, $\hat{Q}^0_i$, $\hat{\pi}^0_{c|ij}$, and initiate commute flows $\hat{\pi}^0_{ijg}$ by evaluating equation 37 given observed flows $\pi_{ijg}$ and initial shocks $\hat{\bar{d}}^0_{ij}$.

**Main loop:** for iteration $t > 0$:

1. **Labour market loop:** for iteration $k > 0$:

(a) Evaluate the relevant case of equation 34 at $\hat{\pi}_{ijg}^{k-1}$ to define $\hat{A}_j(\hat{\pi}_{ijg}^{k-1})$.

(b) Evaluate equation 35 at $\hat{\pi}_{ijg}^{k-1}$ and $\hat{w}_{jg}^{k-1}$ to define $\hat{w}_{jg}(\hat{\pi}_{ijg}^{k-1})$.

(c) Evaluate equation 36 at $\hat{\pi}_{ijg}^{k-1}$ and $\hat{w}_{jg}(\hat{\pi}_{ijg}^{k-1})$ to define $\hat{Q}_i(\hat{\pi}_{ijg}^{k-1})$.

(d) Evaluate equation 37 at $\hat{\bar{d}}_{ij}^{t-1}$, $\hat{Q}_i(\hat{\pi}_{ijg}^{k-1})$, and $\hat{w}_{jg}(\hat{\pi}_{ijg}^{k-1})$ to define $\hat{\pi}_{ijg}(\hat{\bar{d}}_{ij}^{t-1})$.

(e) Update endogenous variables $\hat{x}^k = \zeta\hat{x}^{k-1} + (1-\zeta)\hat{x}(\hat{\pi}_{ijg}^{k-1})$ for each $x \in \{\hat{w}_{jg}, \hat{Q}_i, \hat{\pi}_{ijg}\}$ using a weight $\zeta \in [0, 1]$.

(f) If $|\hat{x}^k - \hat{x}^{k-1}|$ is sufficiently small, stop and define $\hat{w}_{jg}^t$, $\hat{Q}_j^t$, and temporary commute flows $\hat{\pi}_{ijg}(\hat{\bar{d}}_{ij}^{t-1})$ as each variable's $k$th values.

2. Evaluate equation 38 for commute flows $\hat{\pi}_{ij}(\hat{\bar{d}}_{ij}^{t-1})$.

3. **Road traffic loop:** for iteration $k > 0$:

(a) Evaluate equations 39 and 40 at $\hat{\pi}_{ij}(\hat{\bar{d}}_{ij}^{t-1})$ and $\hat{\pi}_{c|ij}^{k-1}$ to update driving costs $\hat{d}_{ijc}^k$.

(b) Evaluate equation 41 at $\hat{d}_{ijc}^k$ to define $\hat{\pi}_{c|ij}(\hat{d}_{ijc}^k)$ and update mode shares $\hat{\pi}_{c|ij}^k = \zeta\hat{\pi}_{c|ij}^{k-1} + (1-\zeta)\hat{\pi}_{c|ij}(\hat{d}_{ijc}^k)$.

(c) If $|\hat{\pi}_{c|ij}^k - \hat{\pi}_{c|ij}^{k-1}|$ is sufficiently small, stop.

4. Evaluate equation 42 at $\hat{d}_{ijc}^k$ to update commute cost indices $\hat{\bar{d}}_{ij}^t$.

5. Evaluate equation 37 at at $\hat{\bar{d}}_{ij}^t$, $\hat{Q}_j^t$, and $\hat{w}_{jg}^t$ to define $\hat{\pi}_{ijg}^{temp}$ and update $\hat{\pi}_{ijg}^t = \zeta\hat{\pi}_{ijg}^{temp} + (1-\zeta)\hat{\pi}_{ijg}^{t-1}$.

6. If $|\hat{\pi}_{ijg}^t - \hat{\pi}_{ijg}^{t-1}|$ is sufficiently small, stop. Otherwise, update $t = t+1$ and repeat main loop.

### A.2.5 Quantifying congestion charge removal

To simulate removing the CCZ, I increase traffic according to routes' Congestion Charge Exposure, remove all road tolls, and confiscate recycled toll revenue. The initial shock increases each road link's traffic to $traffic'^{h}_{\ell} = traffic^{h}_{\ell} - \varphi^{h} CCE^{h}_{\ell}$ in both morning and evening rush hours. Exogenous routes for each rush hour aggregate shocks to MSOA pairs so that $traffic'^{h}_{ij} = \sum_{\ell} 1\{\ell \in R_{ij}\} traffic'^{h}_{\ell}$. I then average across morning and evening rush hours to compute counterfactual traffic $traffic'_{ij} = \frac{1}{2}(traffic'^{morn}_{ij} + traffic'^{eve}_{ij})$.

I scale daily wages $w_{jg}$ to have cross-MSOA geometric means of £120 for low-skill and £160 for high-skill workers, set $\Omega = 1,943,475$ to match the sum of commute flows in the 2011 analysis sample, and parametrize $toll_{ij} = \frac{1}{1 - \frac{8CCZ_{ij}}{\bar{w}}}$ so that $\frac{w_{jg}}{toll_{ij}} \approx w_{jg} - 8CCZ_{ij}$ which holds exactly if $w_{jg}$ is the average wage. Quantitatively, paying this toll increases commute costs by 4.87 percent.[49,50]

Removing the CCZ initiates commute cost shocks $\hat{d}^{0}_{ijc} = \frac{1}{toll_{ij}} \left( \frac{traffic'_{ij}}{traffic_{ij}} \right)^{\kappa}$ and $\widehat{G} = \frac{\left( \sum_{ij} \pi_{ijc} \frac{d_{ijc}}{toll_{ij}} + \pi_{ijo} d_{ijo} \right)}{\left( \sum_{ij} \pi_{ijc} d_{ijc} + \pi_{ijo} d_{ijo} \right)}$ since $toll'_{ij} = 1 \forall i, j$. With exogenous productivity, rents and mode shares begin at their observed 2011 values and commute flows begin at values consistent with commute cost shocks, wages, and rents. With endogenous productivity, all prices and allocations begin at values constituting an equilibrium with exogenous productivity.

### A.2.6 Simulations without shocking non-commuter traffic

In benchmark simulations, traffic falls before commuters adjust because non-commuter traffic leaves the road in proportion to the road links' Congestion Charge Exposure; this section discusses an alternative simulation where only commuters are initially affected by the

[49] Wages and tolls reflect their values in early 2011 (Office for National Statistics, 2012).

[50] The analytical sample contains fewer commuters than raw data because it excludes origin destination pairs that are unmatched to the DfT traffic monitoring network or are outliers in pairwise traffic or CCE as well as dropping counts of individuals working from home or within their residence MSOA.

toll. This simulation initiates commute cost shocks at $\hat{d}^0_{ijc} = \frac{1}{toll_{ij}}$ and traffic changes to the extent that commuters alter location and mode choices in equilibrium. These assumptions give smaller welfare gains and different implications for transit use, but broad conclusions about progressivity and the spatial distributions of employment, population, and traffic resemble baseline results.

Panel A of table A.17 presents correlation coefficients of hat changes in location and traffic patterns simulated with and without shocking non-commuter traffic. Equilibrium traffic changes are highly correlated across simulations and this remains true when excluding a small number of road links where the congestion charge causes particularly large reductions in traffic. Unsurprisingly, changes in residential and workplace location choices are also similar across simulations.

Panel B of table A.17 presents percent change in equilibrium utility, inequality, land rent, traffic, and driving's share of commutes caused by London's Congestion Charge in simulations that do not shock non-commuter traffic. Welfare gains are smaller than in the main text by approximately half but remain positive and progressive. Smaller welfare gains reflect the smaller equilibrium traffic reduction and sustained progressivity reflects the same location fundamentals as in the main text.

The most notable distinction from the main text is that the congestion charge reduces commuters' driving rate in this simulation. This occurs because commuters still face tolls on central commutes but no longer have roadspace reallocated to them in the initial shock. However, it is important to note that effects on mode choice are small relative to the baseline simulation—without shocking non-commuter traffic, the congestion charge reduces the number of commuters driving by 0.03 percent.

### A.2.7 Shipped inputs

Firms may use roads for deliveries, creating an unobserved mechanism through which traffic affects labour demand. I consider this possibility by assuming productivity decreases in travel costs to an exogenously located supplier so that $A_j = a_j \overline{traffic}_j^{\iota}$ where $\overline{traffic}_j$ is the amount of traffic incurred between workplace $j$ and its supplier and $\iota < 0$ is a simplification of firms' elasticity of intermediate input costs with respect to traffic. This is isomorphic to traffic impeding firms' access to customers visiting from a fixed origin.

In this case, labour market clearing requires

$$w_{jg} = \alpha a_{jg} a_j \overline{traffic}_j^{\iota} \left( \Omega \sum_i \pi_{ijg} \right)^{-\frac{1}{\sigma}} \left( \sum_{g'} a_{jg'} \left( \Omega \sum_i \pi_{ijg'} \right)^{\rho} \right)^{\frac{\alpha-\rho}{\rho}}.$$

Recall that $\pi_{ijg}$ decreases in traffic so the standard labour market channel increases wages to compensate commuters for driving in traffic. It follows that counterfactuals without intermediate input access overstate traffic's positive effect on wages to the extent that $\overline{traffic}_j$ is correlated with firms' labour market access $\sum_i (B_{ig} Q_i^{\beta-1})^{\theta} \bar{d}_{ij}^{-\theta}$ for each group.

## Appendix tables

Table A.1: Additional summary statistics for MSOA pairs in the main analysis sample

| | 2001 | | 2011 | | Full Sample | |
|---|---|---|---|---|---|---|
| | Mean | St. Dev | Mean | St. Dev | Mean | St. Dev |
| Commute flows: | | | | | | |
| Drivers | 0.761 | 3.826 | 0.637 | 3.19 | 0.699 | 3.523 |
| Other modes | 1.227 | 8.448 | 1.587 | 10.174 | 1.407 | 9.353 |
| ln(Commuters) | 0.379 | 0.844 | 0.416 | 0.88 | 0.397 | 0.862 |
| ln(Drivers) -ln(Other modes) | -0.035 | 0.662 | -0.125 | 0.723 | -0.08 | 0.695 |
| Pairwise traffic: | | | | | | |
| Morning | 81,148.32 | 56,384.97 | 106,171.30 | 82,321.41 | 93,659.82 | 71,655.95 |
| Evening | 86,735.20 | 59,597.29 | 112,385.90 | 87,644.30 | 99,560.56 | 76,033.98 |
| ln(Traffic) | 11.045 | 0.876 | 11.245 | 0.958 | 11.145 | 0.924 |

Table A.2: Heterogeneous effects of Congestion Charge Exposure on traffic

| | *Interaction variable:* | | | |
|---|---|---|---|---|
| | Distance | CCE | Both Inner London | Both heavy rail $<$ 1,500m |
| CCE $\times$ Post $\times$ Q1 or $1\{d = 1\}$ | $-0.027^{***}$ (0.002) | $-0.186^{***}$ (0.055) | $-0.024^{***}$ (0.002) | $-0.025^{***}$ (0.002) |
| CCE $\times$ Post $\times$ Q2 or $1\{d = 0\}$ | $-0.027^{***}$ (0.003) | $-0.026^{***}$ (0.008) | $-0.028^{***}$ (0.002) | $-0.032^{***}$ (0.002) |
| CCE $\times$ Post $\times$ Q3 | $-0.019^{***}$ (0.003) | $-0.027^{***}$ (0.002) | | |
| Observations | 1,709,220 | 1,709,220 | 1,709,220 | 1,709,220 |

Each column presents results of an OLS regression using 2001 total flows plus one as weights and interacting CCE with indicators describing a fixed observable. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Column 1 divides pairs into distance terciles using DfT monitored road length. Column 2 divides pairs into CCE terciles. Column 3 interacts CCE with an indicator for pairs within Inner London and column 4 interacts with an indicator for both origin and destination centroids within 1,500m of a heavy rail station. For terciles, Q1 is the lowest third and Q3 is the highest, for indicators, $1\{d = 1\}$ indicates when the condition is true. All regressions include full sets of fixed effects and controls described in section 4 footnote 12 and column three adds a both Inner London by year interaction. CCE is scaled to have a standard deviation of one.

Table A.3: Congestion Charge Exposure's effect on road link traffic

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | $Traffic$ | | $Traffic^{morn}$ | $Traffic^{eve}$ |
| | (1) | (2) | (3) | (4) |
| CCE × Post | −109.755*** | −65.527** | | |
| | (23.255) | (30.266) | | |
| $CCE^{morn}$ × Post | | | −65.319** | |
| | | | (32.242) | |
| $CCE^{eve}$ × Post | | | | −102.659*** |
| | | | | (31.623) |
| Controls | No | Yes | Yes | Yes |
| Observations | 5,698 | 5,698 | 5,698 | 5,698 |
| Dep. var. mean | 2819 | 2819 | 2739 | 2898 |
| Dep. var. std. dev. | 2289 | 2289 | 2282 | 2414 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Standard errors clustered by road in parenthesis. Controls are dummies indicating whether a road link was observed in each year from 2000 to 2015, is at least partially within the WEZ, within the CCZ, and interactions of CCZ inclusion with log-distance to CCZ boundary, all interacted with a post period dummy.

Table A.4: Robustness of reduced-form evidence

| Dep. Var | ln(Commuters) | $\ln\left(\frac{\text{Drivers}}{\text{Non-drivers}}\right)$ | ln(Traffic) |
|---|---|---|---|
| Indep. Var | ln(Traffic) | ln(Traffic) | CCE × Post |
| Drop CCZ | −0.977*** | −1.950*** | −0.026*** |
| | (0.221) | (0.246) | (0.002) |
| Drop WEZ | −0.991*** | −1.339*** | −0.026*** |
| | (0.193) | (0.381) | (0.002) |
| Drop WEZ and CCZ | −1.007*** | −2.177*** | −0.026*** |
| | (0.228) | (0.261) | (0.002) |
| New rail control | −0.922*** | −1.202*** | −0.027*** |
| | (0.188) | (0.351) | (0.002) |
| Dist to CBD interaction | −0.726*** | −1.439*** | −0.026*** |
| | (0.190) | (0.332) | (0.002) |
| Excl. IV | CCE × Post | CCE × Post | None |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Each cell presents results of a separate TSLS regression using 2001 total flows plus one as weights and all regressions include full sets of fixed effects and controls. Dependent variables add one to commute counts before taking logarithms and the final column presents first stage estimates. The first three rows iteratively exclude MSOAs in the CCZ, WEZ, or both, the fourth row controls for rail transit built after 2001, and the the final row controls for a post-toll dummy interacted with the interaction of origin and destination distances to Charing Cross. Standard errors in parenthesis are two-way clustered at the residence and workplace levels.

Table A.5: Alternative CCE definitions

| Dep. Var | ln(Commuters) | $\ln\left(\frac{\text{Drivers}}{\text{Non-drivers}}\right)$ | ln(Traffic) |
|---|---|---|---|
| Indep. Var | ln(Traffic) | ln(Traffic) | Standardized CCE Alternative × Post |
| Unweighted tolls | −0.330 | −3.996*** | −0.018*** |
| | (0.280) | (0.572) | (0.003) |
| Include WEZ | −0.801*** | −1.664*** | −0.021*** |
| | (0.226) | (0.398) | (0.002) |
| Per road link | −0.946*** | −2.345*** | −0.015*** |
| | (0.337) | (0.535) | (0.002) |
| Excl. IV | CCE × Post | CCE × Post | None |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Each cell presents results of a separate TSLS regression using 2001 total flows plus one as weights and all regressions include full sets of fixed effects and controls. Dependent variables add one to commute counts before taking logarithms, the final column presents first stage estimates, the final row computes both traffic and CCE on a per-road link basis, and CCE is normalized to have a cross-sectional standard deviation of one throughout. Standard errors in parenthesis are two-way clustered at the residence and workplace levels.

Table A.6: Congestion Charge Exposure and transit expansion.

| | *Dependent variable:* congestion charge exposure | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| New rail transit connection | 0.021** | −0.005* | −0.003 | | | |
| | (0.010) | (0.003) | (0.003) | | | |
| Same express or 24hr bus route | | | | −0.794** | −0.224 | 0.096 |
| | | | | (0.351) | (0.152) | (0.172) |
| Dist FE | No | Yes | Yes | No | Yes | Yes |
| Pow/Por FE | None | Both | Both | None | Both | Both |
| Controls | None | None | All | None | None | No transit |
| Observations | 854,610 | 854,610 | 854,610 | 854,610 | 854,610 | 854,610 |

Each column presents results of an OLS regression of the form $CCE_{ij} = \alpha NewConnection_{ij} + \alpha_i + \alpha_j + \alpha_{d(ij)} + X'_{ij}\beta + e_{ij}$ using 2001 total flows plus one as weights. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. CCE is scaled to have a standard deviation of one.

Table A.7: Traffic's separate effects on commuting by mode

| | *Dependent variable:* | |
|---|---|---|
| | ln(Drivers) | ln(Non-drivers) |
| | (1) | (2) |
| ln(Traffic) | $-1.338^{***}$ | $-0.130$ |
| | (0.292) | (0.161) |
| Observations | 1,709,220 | 1,709,220 |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

Table A.8: Traffic's effect on commuting by trip distance

| | *Dependent variable:* | |
|---|---|---|
| | ln(Commuters) | $\ln\left(\frac{\text{Drivers}}{\text{Non-drivers}}\right)$ |
| | (1) | (2) |
| ln(Traffic) | $-0.749^{***}$ | $-1.171^{***}$ |
| $\times$ (dist $<$ 15.1 km) | (0.165) | (0.307) |
| ln(Traffic) | $-0.748^{***}$ | $-1.126^{***}$ |
| $\times$ (15.1 km $\leq$ dist $<$ 26.4 km) | (0.163) | (0.306) |
| ln(Traffic) | $-0.752^{***}$ | $-1.106^{***}$ |
| $\times$ (26.4 km $\geq$ dist ) | (0.162) | (0.307) |
| Dist $\times$ Post FE | Yes | Yes |
| Pow/Por $\times$ Post FE | Both | Both |
| First Stage F-stat (dist $<$ 15.1 km) | 190.41 | 190.41 |
| First Stage F-stat (15.1 km $\leq$ dist $<$ 26.4 km) | 312.85 | 312.85 |
| First Stage F-stat (26.4 km $<$ dist ) | 363.24 | 363.24 |
| Observations | 1,709,220 | 1,709,220 |
| *Note:* | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Each column presents results of a TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. Dependent variables add one to flow variables before taking logarithms and interaction terms are dummies separating the sample into equal sized groups by euclidean distance between workplace and residence. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. Excluded instruments are CCE-distance tercile interactions and partial first stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

Table A.9: Traffic's effect on commuting by Inner London

| | *Dependent variable:* | |
|---|---|---|
| | ln(Commuters) | $\ln(\frac{\text{Drivers}}{\text{Non-drivers}})$ |
| | (1) | (2) |
| ln(Traffic) × | −2.702*** | 0.282 |
| (Both Inner London) | (0.437) | (0.736) |
| ln(Traffic) × | −0.761*** | −1.569*** |
| (1 - Both Inner London) | (0.196) | (0.334) |
| Dist × Post FE | Yes | Yes |
| Pow/Por × Post FE | Both | Both |
| First Stage F-stat (Both Inner London) | 23.75 | 23.75 |
| First Stage F-stat (Not both Inner London) | 61.83 | 61.83 |
| Observations | 1,709,220 | 1,709,220 |
| *Note:* | *p<0.1; **p<0.05; ***p<0.01 | |

Each column presents results of a TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. Dependent variables add one to flow variables before taking logarithms and the interaction term specifies MSOA pairs where both workplace and residence MSOAs are in the Inner London boroughs defined in appendix A.1.1. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12 and add Inner London-post-toll interactions. Excluded instruments are CCE-Inner London interactions and partial first stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

Table A.10: Structural gravity

| | *Dependent variable:* | |
|---|---|---|
| | $\ln(\pi_{ij})$ | |
| | (1) | (2) |
| $\ln(1\text{-}\pi_{c\|ij})$ | $-2.074^{***}$ | $-2.229^{***}$ |
| | (0.370) | (0.392) |
| Controls | No | Yes |
| Dist × Post FE | Yes | Yes |
| Pow/Por × Post FE | Both | Both |
| Excl. IV | CCE × Post | CCE × Post |
| First Stage F-stat | 126.59 | 106.31 |
| Observations | 1,709,220 | 1,709,220 |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

Each column presents results of a TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. The dependent variable is the natural logarithm of the MSOA pair's share of annual commuters plus $\frac{1}{N\times(1-N)}$ and independent variable is the natural logarithm of the pair's non-car commute share plus one. Standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12. First stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

Table A.11: Share of London Region employment captured by commute flows

| | Drivers | Other modes | Total |
|---|---|---|---|
| 2001 | 0.65 | 0.77 | 0.72 |
| 2011 | 0.60 | 0.81 | 0.75 |

Table A.12: Validating routes using pairwise traffic

| | *Fastest-in-traffic routed:* | | | | | |
|---|---|---|---|---|---|---|
| | ln(Morning traffic) | | $\Delta$ ln(Morning traffic) | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ln(Morning traffic) | 0.920*** | | | | | |
| *OSRM routed* | (0.011) | | | | | |
| Post-toll | | 0.207*** | | | | |
| | | (0.050) | | | | |
| ln(Morning traffic) $\times$ | | 0.926*** | | | | |
| Pre-toll, *OSRM routed* | | (0.012) | | | | |
| ln(Morning traffic) $\times$ | | 0.910*** | | | | |
| Post-toll, *OSRM routed* | | (0.012) | | | | |
| $\Delta$ ln(Morning traffic) | | | 0.763*** | 0.763*** | 0.570*** | 0.529*** |
| *OSRM routed* | | | (0.014) | (0.014) | (0.017) | (0.027) |
| Constant | 0.972*** | 0.898*** | 0.070*** | 0.070*** | | |
| | (0.129) | (0.128) | (0.005) | (0.005) | | |
| Observations | 73,130 | 73,130 | 73,130 | 73,130 | 73,130 | 73,130 |
| Por/Pow/Dist FEs | No | No | No | No | Yes | Yes |
| Weights | None | None | None | None | None | $\pi_{ij2001} + 1$ |
| Projected $R^2$ | 0.689 | 0.690 | 0.566 | 0.566 | 0.690 | 0.697 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Table A.13: Validating routes using monitored road length (cont)

| | *Fastest-in-traffic routed:* | | | | | |
|---|---|---|---|---|---|---|
| | ln(Monitorred metres) | | $\Delta$ ln(Monitorred metres) | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| ln(monitorred metres) | 0.922*** | | | | | |
| *OSRM routed* | (0.012) | | | | | |
| Post-toll | | −0.038 | | | | |
| | | (0.053) | | | | |
| ln(monitorred metres)× | | 0.911*** | | | | |
| Pre-toll, *OSRM routed* | | (0.012) | | | | |
| ln(monitorred metres)× | | 0.918*** | | | | |
| Post-toll, *OSRM routed* | | (0.013) | | | | |
| $\Delta$ln(Monitorred metres) | | | 0.686*** | 0.686*** | 0.572*** | 0.595*** |
| *OSRM routed* | | | (0.018) | (0.018) | (0.017) | (0.018) |
| Constant | 0.751*** | 0.843*** | 0.094*** | 0.094*** | | |
| | (0.117) | (0.121) | (0.006) | (0.006) | | |
| Observations | 73,130 | 73,130 | 73,130 | 73,130 | 73,130 | 73,130 |
| Por/Pow/Dist FEs | No | No | No | No | Yes | Yes |
| Weights | None | None | None | None | None | $\pi_{ij2001} + 1$ |
| Projected $R^2$ | 0.690 | 0.691 | 0.441 | 0.441 | 0.618 | 0.654 |

*Note:* *p<0.1; **p<0.05; ***p<0.01

Table A.14: Maximum Likelihood Estimates of traffic's effect on the number commuting

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Number of commuters | | | |
| | (1) | (2) | (3) | (4) |
| ln(Traffic) | -0.9760 | -0.1024 | -0.3236 | -0.2258 |
| | (0.4320) | (0.1589) | (0.2375) | (0.2062) |
| | [0.1609] | [0.1178] | [0.1852] | [0.1419] |
| First stage residual | 0.8764 | 0.0922 | 0.2218 | 0.1935 |
| | (0.8764) | (0.1544) | (0.2405) | (0.2086) |
| | [0.1783] | [0.1178] | [0.1886] | [0.1439] |
| Dist × Post FE | No | Yes | Yes | Yes |
| Pow/Por × Post FE | None | Both | Both | Both |
| Controls | No | No | Yes | Yes |
| Model | Poisson | Poisson | Poisson | Negative binomial |
| Overdispersion parameter | | | | 9.227 |
| Observations | 493,976 | 493,976 | 493,976 | 493,976 |

Poisson regressions use 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. Standard errors in round parenthesis are two-way clustered at the residence and workplace levels and square parenthesis are clustered by pair and p-value are omitted as I present QMLE-style errors rather than CFIV bootstrap inference. Controls are described in section 4 footnote 12.

Table A.15: Maximum Likelihood Estimates of traffic's effect on mode shares

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | Number of commuters (by mode) | | | |
| | (1) | (2) | (3) | (4) |
| ln(Traffic) | -6.9119 | 0.4292 | 0.6405 | 0.3222 |
| × Driver | (0.9407) | (0.9511) | (1.319) | (1.2147) |
| | [0.2517] | [0.4359] | [0.6512] | [0.4969] |
| First stage residual | 6.4022 | -0.5145 | -0.5712 | -0.2815 |
| | (0.982585) | (.99713) | (1.2892) | (1.1839) |
| | [0.2566] | [0.4380] | [0.6560] | [0.4951] |
| Dist × Post × Driver FE | No | Yes | Yes | Yes |
| Pow/Por × Post × Driver FE | None | Both | Both | Both |
| Controls | No | No | Yes | Yes |
| Model | Poisson | Poisson | Poisson | Negative binomial |
| Overdispersion parameter | – | – | – | 9,999 |
| Observations | 588,910 | 588,910 | 588,910 | 588,910 |
| *Note:* | | | | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 |

Poisson regressions use 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. Standard errors in round parenthesis are two-way clustered at the residence and workplace levels and square parenthesis are clustered by pair and p-value are omitted as I present QMLE-style errors rather than CFIV bootstrap inference. Controls are described in section 4 footnote 12 and are additionally interacted with a driver dummy variable.

Table A.16: Commuting elasticities by skill

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | ln(Commuters) | $\ln(\frac{\text{Drivers}}{\text{Non-drivers}})$ | ln(Commuters) | $\ln(\frac{\text{Drivers}}{\text{Non-drivers}})$ |
| | (1) | (2) | (3) | (4) |
| ln(Traffic) × | $-1.686^{***}$ | $-0.948^{**}$ | $-3.595^{***}$ | $-0.182$ |
| low-skill | (0.281) | (0.371) | (1.238) | (1.004) |
| ln(Traffic) × | $-0.726^{***}$ | $-1.352^{***}$ | $-0.505$ | $-0.011$ |
| (1 - low-skill) | (0.201) | (0.387) | (0.320) | (0.579) |
| Pow/Por FE | by year | by year | by skill-year | by skill-year |
| First Stage F-stat (low-skill) | 24.78 | 24.78 | 6.49 | 6.49 |
| First Stage F-stat (1 - low-skill) | 40.53 | 40.53 | 84.39 | 84.39 |
| Controls | Yes | Yes | Yes | Yes |
| Observations | 1,709,220 | 1,709,220 | 1,709,220 | 1,709,220 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

Each column presents results of a TSLS regression using 2001 total flows plus one as weights and all regressions include MSOA pair fixed effects. Dependent variables add one to flow variables before taking logarithms and standard errors in parenthesis are two-way clustered at the residence and workplace levels. Controls are described in section 4 footnote 12 and add skill-group-by-year controls. Excluded instruments are CCE-skill interactions and partial first stage F-statistics are computed as in Sanderson and Windmeijer (2016) using two-way residence and workplace clustered variance covariance matrices.

Table A.17: Effects of the congestion charge without shocking non-commuter traffic

| A: Correlation with and without non-commuter traffic shock | | | | | | |
|---|---|---|---|---|---|---|
| Employment | | Population | | Road link traffic | | |
| Low | High | Low | High | | | |
| 0.8870 | 0.8871 | 0.8829 | 0.8407 | 0.9818 (0.9406) | | |
| B: Aggregate effects without non-commuter traffic shock | | | | | | |
| Utility | | | | Rent | Car share | Traffic |
| Mean | Low | High | Ineq | | | |
| 0.0540 | 0.1372 | 0.0340 | -0.0999 | -0.0600 | -0.0300 | -0.1199 |

Panel A presents cross-simulation correlations of hat changes in MSOA population, MSOA employment, and road link traffic. The traffic correlation in parenthesis excludes 25 road links where $\widehat{traffic}_\ell > 2$ in either simulation. Panel B presents percent difference between actual and counterfactual outcomes in the full model without the initial shock to non-commuter traffic.

# Appendix figures

Figure A.1: Matching pairwise commutes to traffic counts

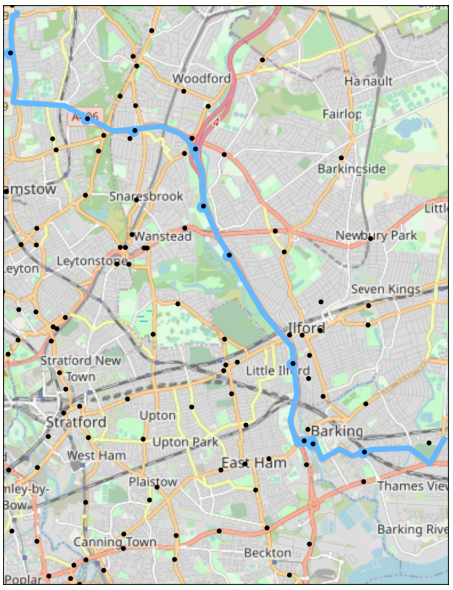


Figure A.2: Central London's Congestion Charge Zone and Western Extension

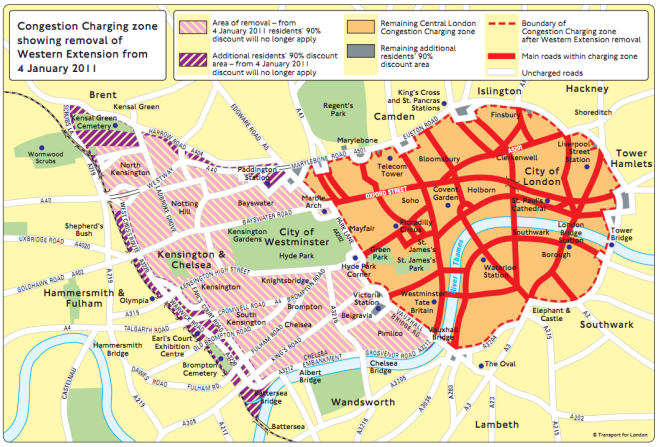



Figure produced by Transport for London and adapted from Tang (2018).

Figure A.3: Policy timeline

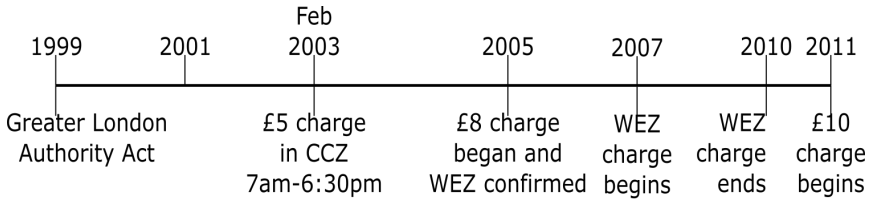
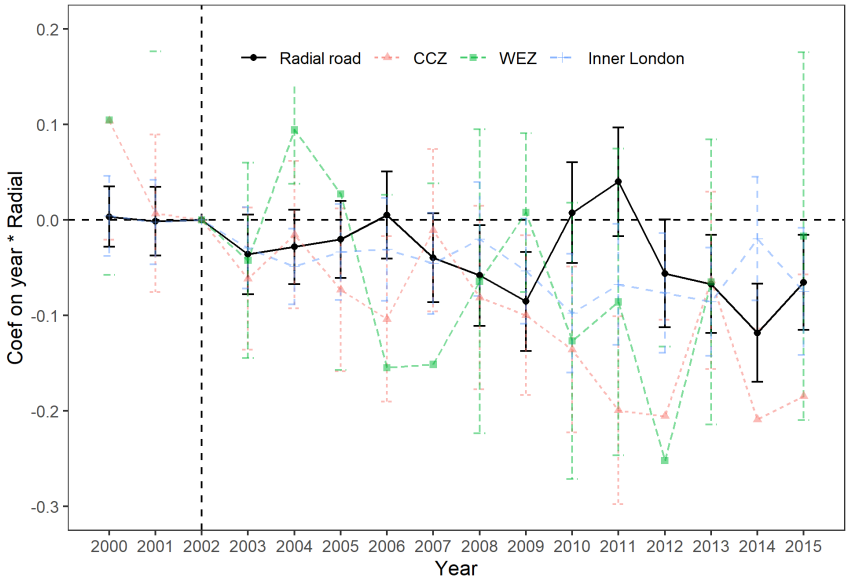
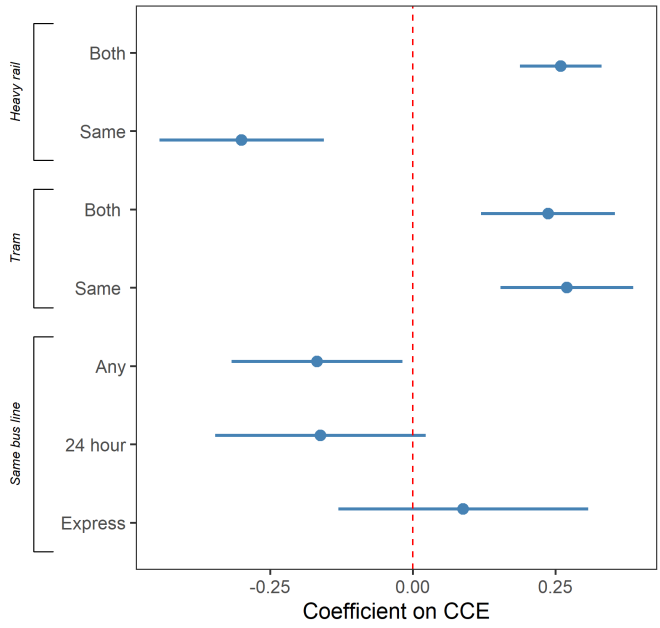


Figure A.4: London's congestion charge and radial road traffic



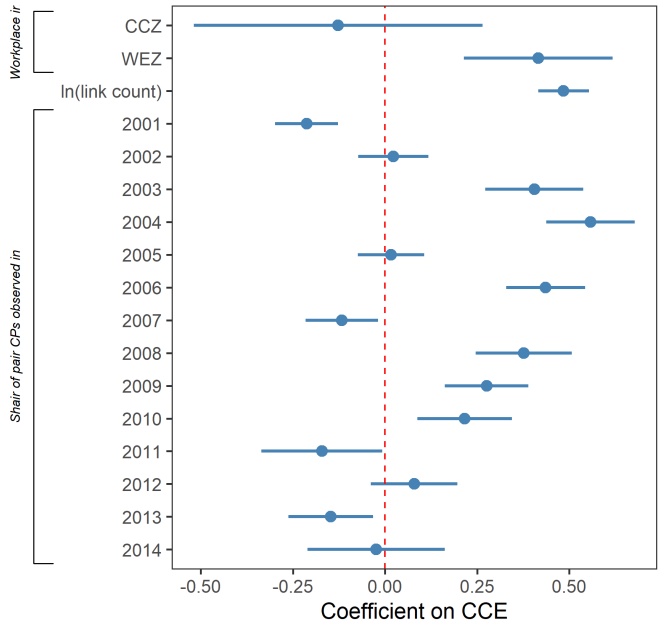

Points are OLS estimates of $\beta_\tau^{zone}$ the link-year regression $ln(Traffic_{\ell t}) = \sum_{\substack{\tau=2000,\\ \tau\neq 2002}}^{2015} \Big( D_\tau^t \beta_\tau^r Radial_\ell + D_\tau^t \beta_\tau^c CCZ_\ell + D_\tau^t \beta_\tau^w WEZ_\ell + D_\tau^t \beta_\tau^\iota InnerLondon_\ell + D_\tau^t f_\tau(lat_\ell, lon_\ell) \Big) + \alpha_t + \alpha_\ell + e_{\ell t}$, normalized to zero in 2002, $Traffic_{\ell t}$ is the mean of morning and evening rush-hour traffic volumes in that link-year, and error bars denote link-clustered 95 percent confidence intervals in zone-years where they are contained between $[-.3, .2]$.

Figure A.5: Conditional correlations between CCE and transit characteristics.




Points are estimates of $\beta$ from independent weighted OLS regressions of the form $CCE_{ij} = \beta_i + \beta_j + \beta_{d(ij)} + \beta x_{ij} + u_{ij}$ where $x_{ij}$ is a public transit characteristic for commute $ij$. "Both" indicates pairs with origin and destination both within 1500 metres of a station and "same" indicates pairs with origin and destination within 1500 metres of stations on the same transit line. Bars are 95 percent confidence intervals 2 way clustered by residence and workplace.

Figure A.6: Conditional correlations between CCE and route characteristics (cont.).




Points are estimates of $\beta$ from independent weighted OLS regressions of the form $CCE_{ij} = \beta_i + \beta_j + \beta_{d(ij)} + \beta x_{ij} + u_{ij}$. Bars are 95 percent confidence intervals 2 way clustered by residence and workplace.

Figure A.7: Validating routing assumptions

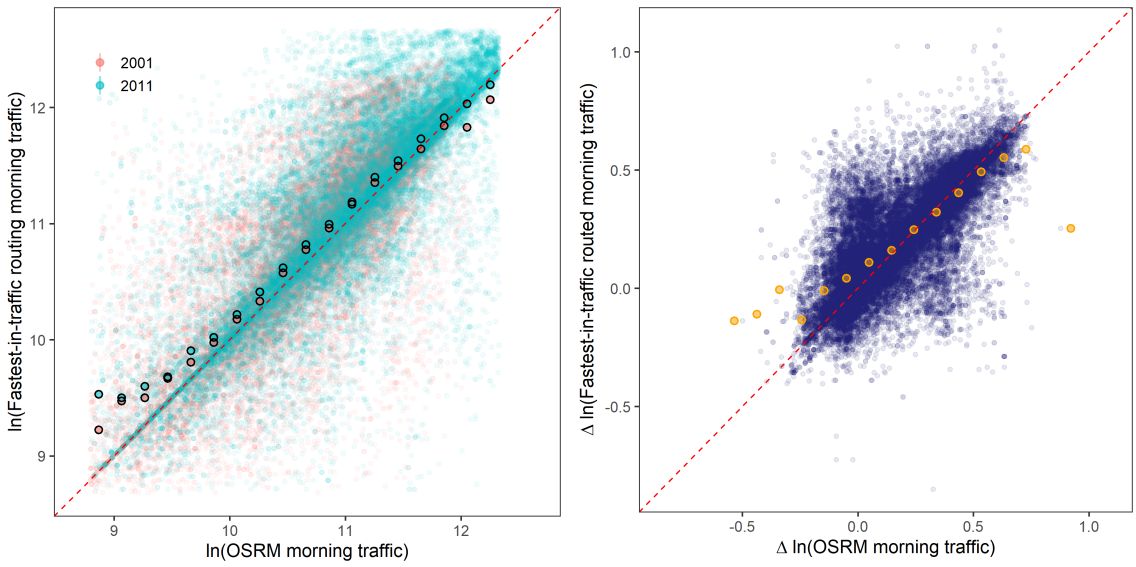
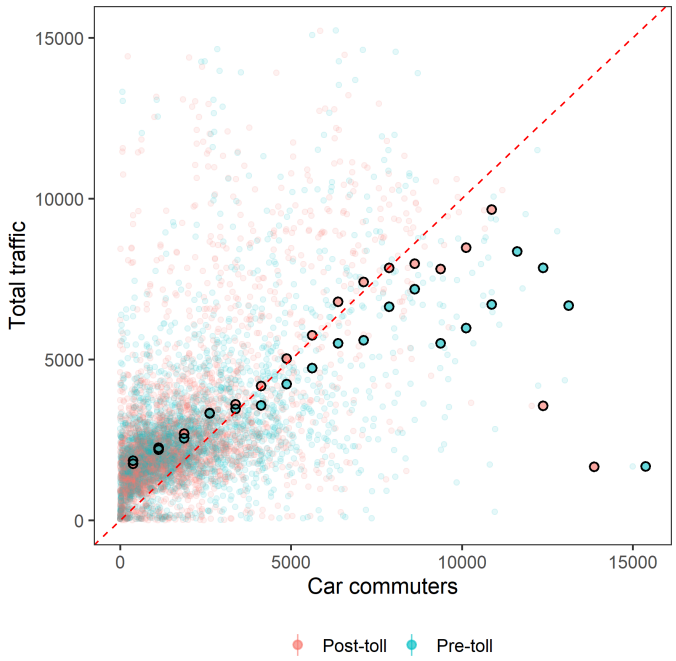
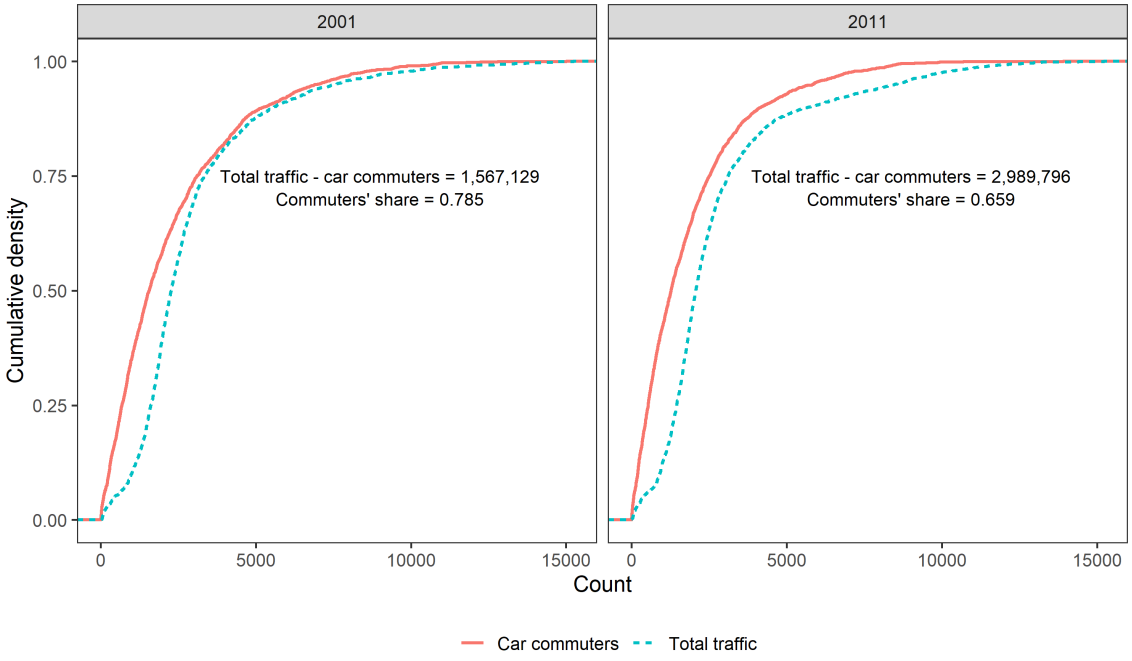


Figure A.8: Road link traffic and car commuters



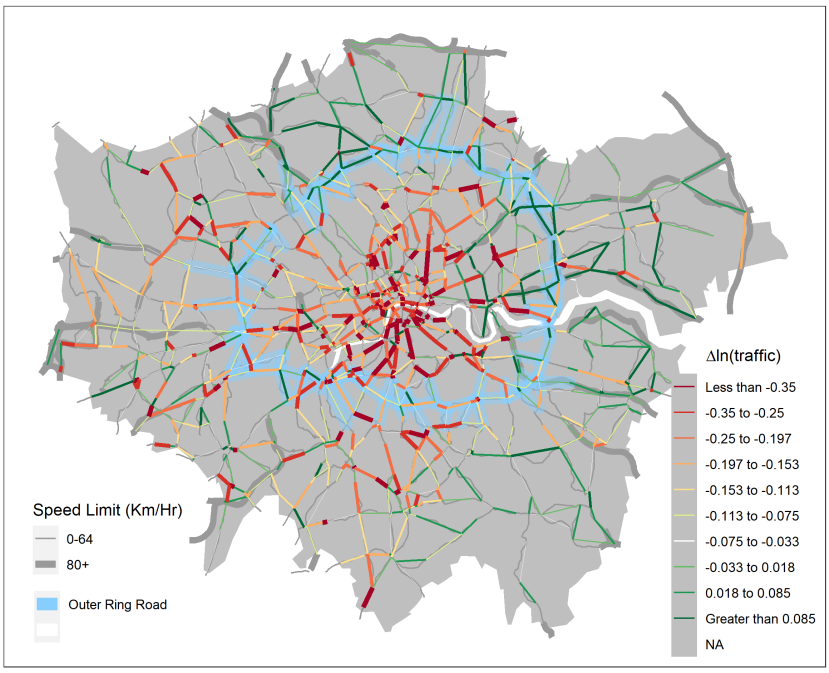

Small points are road-link daily mean rush hour traffic and estimated counts of daily car commuters, large points are binned means, and red dashes denote the 45 degree line. Small point transparency is such that ten overlapping points appear opaque.

Figure A.9: Road link traffic and car commuters (cont.)



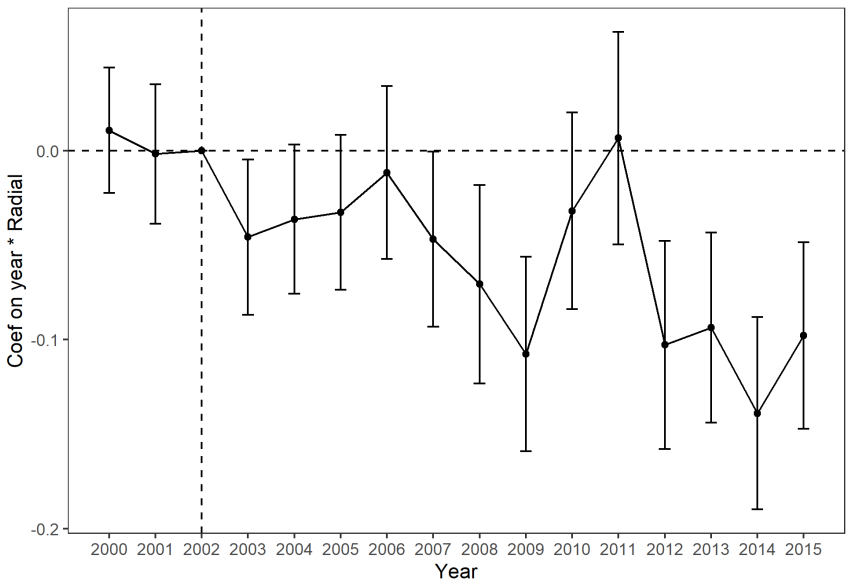

Solid lines are empirical density functions of estimated daily car commuters and dashed lines are those of road-link daily main rush hour traffic.

Figure A.10: Log-change in count point traffic




Colours and thickness are proportional difference in logarithms of post- and pre-toll traffic at each count point (relative to zero). I aggregate traffic across travel directions and rush hour to compute count point traffic in each period. Count points appear as linear representations of the road links they monitor.

Figure A.11: London's congestion charge and radial road traffic




Points are OLS estimates of $\beta_\tau$ the link-year regression $ln(traffic_{\ell t}) = \sum_{\substack{\tau=2000, \\ \tau \neq 2002}}^{2015} \left( D_\tau^t \times \beta_\tau radial_\ell + D_\tau^t \times f_\tau(lat_\ell, lon_\ell) \right) + \alpha_t + \alpha_\ell + e_{\ell t}$, normalized to zero in 2002, $Traffic_{\ell t}$ is the mean of morning and evening rush-hour traffic volumes in that link-year, and error bars denote link-clustered 95 percent confidence intervals.