Accident Analysis and Prevention 49 (2012) 366–377

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

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

The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation

Haojie Li, Daniel J. Graham*, Arnab Majumdar

Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, London SW7 2AZ, UK

ARTICLE INFO

Article history:
Received 25 August 2011
Received in revised form 17 January 2012
Accepted 14 February 2012

Keywords:
London congestion charge
Causal analysis
Difference-in-difference

ABSTRACT

This paper aims to identify the impacts of the London congestion charge on road casualties within the central London charging zone. It develops a full difference-in-difference (DID) model that is integrated with generalized linear models, such as Poisson and Negative Binomial regression models. Covariates are included in the model to adjust for factors that violate the parallel trend assumption, which is critical in the DID model. The lower Bayesian Information Criterion value suggests that the full difference-in-difference model performs well in evaluating the relationship between road accidents and the London congestion charge as well as other socio-economic factors. After adjusting for a time trend and regional effects, the results show that the introduction of the London congestion charge has a significant influence on the incidence of road casualties. The congestion charge reduces the total number of car accidents, but is associated with an increase in two wheeled vehicle accidents.

1. Introduction

Since the first congestion charging scheme was introduced by Singapore in 1975, several studies have been conducted to better understand the effects induced by congestion charging. It has been shown that congestion charging can decrease congestion effectively (Olszewski and Xie, 2005) and consequently affect the traffic flow conditions such as the traffic volume, volume-to-capacity ratio (V/C) and traffic flow speed, all of which have direct impacts on the occurrence likelihood and severity of road traffic casualties (Lord et al., 2005).

The objective of this paper is to test the causal effect of the London congestion charge (LCC) on road accidents. The difference-in-difference (DID) method is introduced as an evaluation tool to make causal inferences and this method is frequently applied to evaluate the impact of policies (Ashenfelter and Card, 1985; Card and Krueger, 1994; Finkelstein, 2002; Donald and Lang, 2007; Athey and Imbens, 2006; Abadie et al., 2010) but, to the best of our knowledge, has not been used for road accident related transport research. The DID approach is applied using generalized linear models (GLMs), such as Poisson and Negative Binomial models, to estimate the effect of congestion charging on the counts of accidents, which are categorized by casualty type and severity. Covariates are introduced to the DID model to adjust for factors that might lead to the violation of the parallel trend assumption in DID estimation.

The paper is organized as follows. The literature review is presented in Section 2. Section 3 describes the method and the data sources used in this analysis and this is followed by the results and discussion in Section 4. The conclusions are given in the final section.

2. Literature review

Causal relationships can be distinguished from pure statistical relationships given that a plausible mechanism underpinning the relationship between target variables and the treatment (Elvik, 2011) is outlined. To estimate the causal relationship between the LCC and road accidents, it is necessary to understand and outline the mechanisms by which road pricing may affect road traffic accidents.

Most previous research on road traffic accidents has explored various factors contributing to risk, including the traffic characteristics, road characteristics, demographics and the environment (Wier et al., 2009; Quddus, 2008a,b; Dissanayake et al., 2009). In recent years, researchers have paid close attention to economic factors, such as fuel prices and road taxation. As fuel prices and road users' taxes rise, cars will tend to be driven less and consequently there will be less traffic congestion. It may be hypothesized that since a higher tax leads to fewer miles travelled, roads will be emptier of traffic and probably safer. However, this line of argument is ambiguous, because when road taxation is more expensive, travellers will switch to other travel modes like bicycle and motorcycle, which may be more vulnerable to severe accidents (White, 2004;

* Corresponding author. Tel.: +44 20 7594 6088; fax: +44 20 7594 6107.
E-mail address: d.j.graham@imperial.ac.uk (D.J. Graham).

0001-4575/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2012.02.013

Leigh and Wilkinson, 1991; Crandall and Graham, 1989). Nevertheless, many researchers have shown that there does seem to be a relationship between fuel prices and road casualties, although the direction of the relationship depends on the type of accidents. Hyatt et al. (2009) investigated the relationship between motor vehicle injury and mortality rates and fuel prices. By using monthly fuel price and fatality panel data, they found higher gasoline prices were related to increased motorcycle casualties and the authors explained this increase was more a factor of the increasing number of motorcycles on the road. Grabowski et al. (2006) generated panel data on the total number of traffic fatalities for the 48 continental U.S. states during the period 1982–2000. Their results suggested that exogenous increases in state gasoline taxes were plausibly associated with fewer traffic fatalities. Many other studies have also presented similar results (Leigh and Wilkinson, 1991; Haughton and Sarkar, 1996; Grabowski and Morrisey, 2004). All the above studies suggest that fuel tax, as one mode of road taxation, has an influence on road casualties. Similarly, we could hypothesize that congestion charging, another such mode of road taxation which aims to alleviate congestion, may also affect traffic accidents.

London provides a unique opportunity to study this hypothesis. The London congestion charging (LCC) scheme was introduced in central London on 17 February 2003 at a flat rate of £5 per day between the hours of 7:00 am and 6:30 pm, Monday to Friday. The charge was then raised from £5 to £8 on 4 July 2005. A western extension of the charging zone was implemented on 19 February 2007 and the charging hours were reduced to 7:00 am to 6:00 pm. The western charging zone was removed and the charge was increased to £10 on 4 January 2011. Fig. 1 shows the area of the initial central London charging zone and the western extension. The LCC scheme aims to reduce congestion and travel delay and thereby improve journey quality. Congestion in central London reduced by up to 30% and average traffic speeds increased from 13 km/h to 17 km/h during the initial charging period (TfL, 2004). It has also been showed that the number of traffic accidents reduced significantly in both the original and extended charging zone (TfL, 2007). Periods and treatments are shown below:

(1) 2003–2004: Initial congestion charge in central London £5.
(2) 2005–2006: Congestion fee increase from £5 to £8.
(3) 2007–2010: Western extension of charging zone.

Tuerk and Graham (2010) conducted research on the impacts of the LCC scheme on traffic volumes. Traffic volumes crossing central London were measured by automatic traffic counters and aggregated at the hourly level. DID estimation was used to analyse the traffic data and the results indicate a reduction in traffic due to the increase in the congestion fee for the LCC (7.8% for inbound vehicles).

Other papers have focused on relationship between the congestion charge and the travel modes, environment and business activity matters (e.g. Eliasson and Mattsson, 2006; Wichiensin et al., 2007). All these studies show that the congestion charge does have effects on travel costs, travel time and the transit market. However, little research has directly investigated the relationship between the LCC and road accidents in London.

Quddus (2008a,b) conducted a time series analysis of traffic accidents in Great Britain and his results for the London congestion charge suggested an average 33% reduction of casualties in each month after the LCC. Noland et al. (2008) examined the effects of

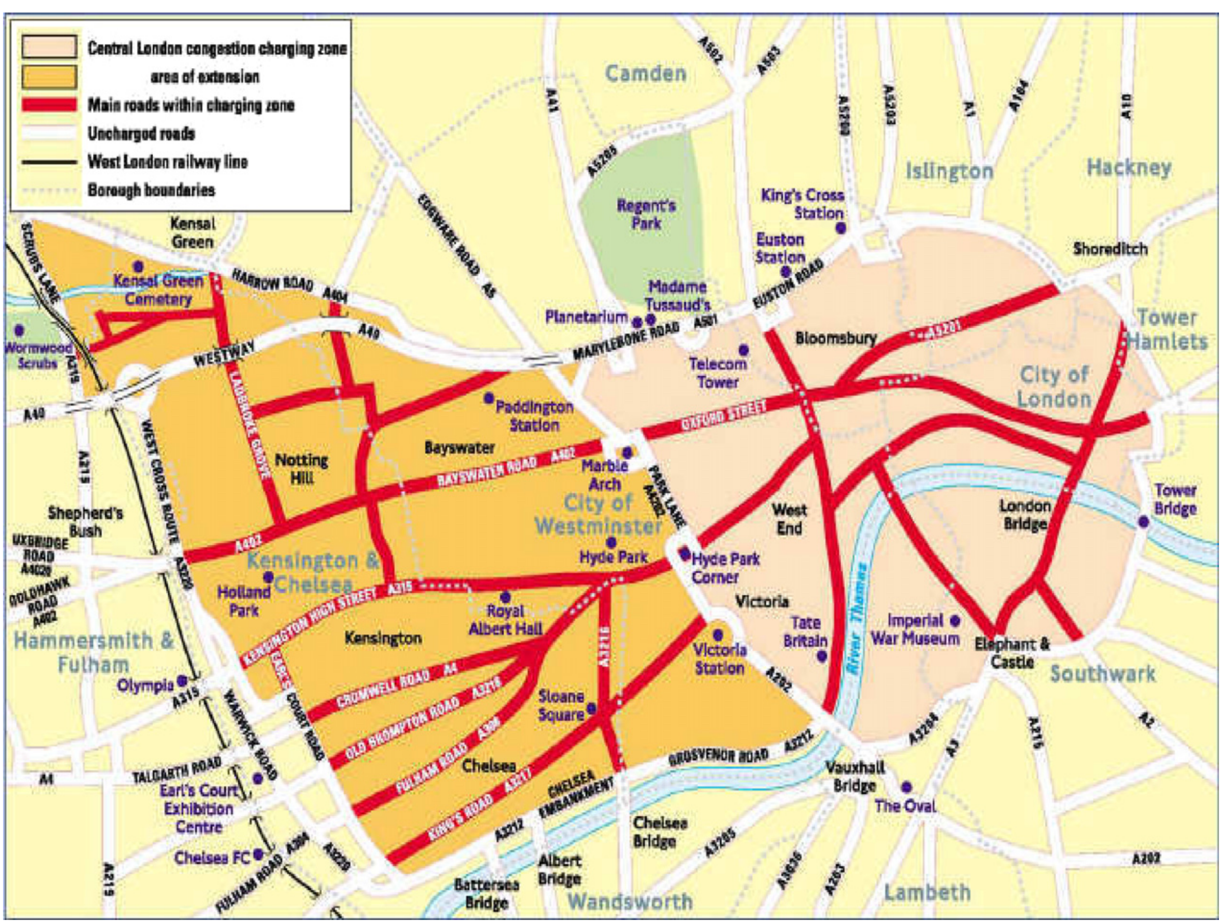

Fig. 1. Map of the London congestion charging zone.

the LCC on traffic casualties by employing an intervention analysis. In this study, data on traffic casualties from 1991 to 2004, covering the 33 London boroughs was analysed. To account for serial correlation and seasonality effects, the intervention model (Box and Tiao, 1975) was used to analyze the effect of the congestion charge on traffic casualties. Although no significant effect was found for total casualties in Greater London area, their results suggested a significant drop in vehicle casualties and an increase in cycle casualties, which can be due to the switch in commuting modes. It is worth noting that the intervention model applied in their study cannot correct for effects due to nationwide trends which could have some broadly universal influence on accident counts. Alternative evaluation methods therefore need to be applied to verify the existing findings.

The DID method has been employed in many economic evaluation studies. Card and Krueger (1994, 2000) estimate the impact of increase in New Jersey's minimum wage on employment in New Jersey's fast food restaurant. Finkelstein (2002) uses a difference-in-difference methodology to investigate how the evolution of a tax subsidy to employer-provided health insurance affects coverage by such insurance. Abadie and Gardeazabal (2003) evaluate the effect of the terrorist conflict on the economic evolution in the Basque Country using Spanish regions unaffected by the conflict as a synthetic control group. Other research is conducted to estimate the impact of a tobacco control program in California (Abadie et al., 2010). All these studies show that the DID method is able to control for effects due to the general time trend, regional differences and other confounders. However, this method has yet to be used in any road accidents analysis. An enhanced DID model appropriate for road accidents data analysis is introduced in the following section.

3. Method and data

3.1. Method

3.1.1. The difference-in-difference model

The use of DID methods has become widespread in estimating causal relationships, since the work by Ashenfelter and Card (1985). The basic idea is that observations are collected for two groups for two periods. One of the groups is the treatment group which is exposed to the treatment in one period. The other group is the control group which receives no treatment during both periods. In the case where the same units within a group are observed in each time period, the average gain over time in the non-exposed (control) group is extracted from the gain over time in the exposed (treatment) group. This double differencing, the so called "difference-in-difference" methods, removes biases in the second period comparison between the treatment and control groups that could be the result from permanent differences between those groups, as well as biases from comparison over time in the treatment group that could be the result of time trends unrelated to the treatment (see Abadie, 2005; Finkelstein, 2002; Card and Krueger, 1994 for more detailed discussion).

Assume that we observe n individuals in two time periods, $t = 0, 1$ where 0 indicates a time period before the treatment group receives treatment, i.e. pre-treatment, and 1 indicates a time period after the treatment group receives treatment, i.e. post-treatment. Every group is indexed by the letter $i = T, C$ where T indicates the treatment group, and C indicates the control group; let Y_{0T} and Y_{1T} be the outcome for the treatment group before and after treatment, respectively, and let Y_{0C} and Y_{1C} be the corresponding outcome for the control group.

3.1.2. DID model

Under the basic DID approach, the outcome Y_{it} is modelled by the following equation

$$Y_{it} = \alpha + \beta T_{it} + \gamma G_{it} + d(T_{it} \cdot G_{it}) + e_{it} \tag{1}$$

where α is the constant term, β is the time trend, γ is the specific group effect, δ is the treatment effect we are interested in and ε_{it} is a random, unobserved term which contains the error caused by omitted covariates. T_{it} is the time-specific component, which takes the value 1 if Y_{it} is observed in the post-treatment period and 0 otherwise. G_{it} is a group-specific component, which is 1 if Y_{it} is an observation from the treatment group and 0 otherwise. $T_{it} \cdot G_{it}$ is an interaction term which indicates a treated individual after the intervention.

3.1.3. Assumptions

Three assumptions are necessary for DID to provide an unbiased consistent estimate of the treatment effect.

(1) The model is correctly constructed in the sense that the function and covariates added into the equation are correct.
(2) The error term has expectation zero and is distributed independently of the covariates.
(3) The third assumption, which is also critical in DID estimation, is that the treatment group and control group will follow the same trend over time in the absence of the treatment. This is also known the as parallel trend assumption,

$$E[Y'_{1T} - Y'_{0T}] = E[Y_{1C} - Y_{0C}] \tag{2}$$

where Y'_{it} represents outcomes of the treatment group in absence of treatment.

3.1.4. DID estimator

According to these assumptions, we could attain expected values of outcomes Y_{it} given by following equations.

$$E\,[Y_{0T}] = \alpha + \gamma$$
$$E\,[Y_{1T}] = \alpha + \beta + \gamma + \delta$$
$$E\,[Y_{0C}] = \alpha$$
$$E\,[Y_{1C}] = \alpha + \beta$$

First we consider a single difference estimator, which compares the difference only in the treatment group before and after treatment.

$$\widehat{\delta}_1 = \bar{Y}_{iT} - \bar{Y}_{0T}$$

$$E[\widehat{\delta}_1] = E[\bar{Y}_{1T}] - E[\bar{Y}_{0T}] = \alpha + \beta + \gamma + \delta - (\alpha + \gamma) = \beta + \delta$$

So we can conclude that the estimator $\widehat{\delta}_1$ will be biased if a time trend exists, because we may treat the time trend as part of the treatment effect.

Next consider another estimator based on comparing the average difference between the treatment and control groups.

$$\widehat{\delta}_2 = \bar{Y}_{1T} - \bar{Y}_{1C}$$

$$E[\widehat{\delta}_2] = E[Y_{1T}] - E[Y_{1C}] = \alpha + \beta + \gamma + \delta - (\alpha + \beta) = \gamma + \delta$$

This estimator is also biased due to the specific group effect.

As defined in the previous section, the difference-in-difference estimator is the difference in average outcome in the treatment group before and after the treatment minus the difference in average outcome in the control group before and after the treatment.

$$\widehat{\delta}_{DID} = \bar{Y}_{1T} - \bar{Y}_{0T} - (\bar{Y}'_{1T} - \bar{Y}'_{0T}) = \bar{Y}_{1T} - \bar{Y}_{0T} - (\bar{Y}_{1C} - \bar{Y}_{0C})$$

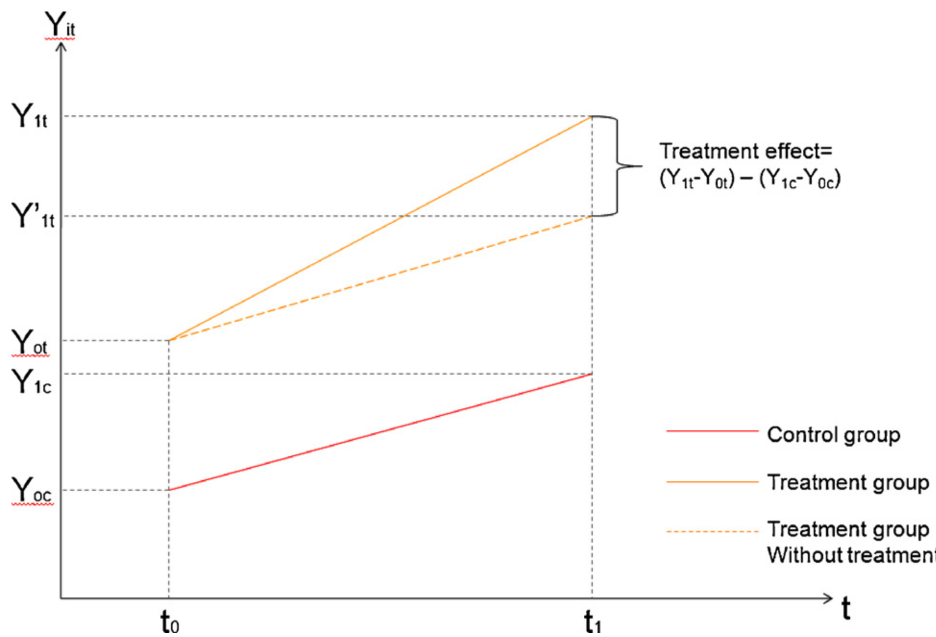

Fig. 2. Graphic illustration for DID estimator.

$$E[\widehat{\delta}_{DID}] = \alpha + \beta + \gamma + \delta - (\alpha + \gamma) - (\alpha + \beta) - \alpha = \delta$$

We can see that this is an unbiased estimator. The DID estimator is shown graphically in Fig. 2. According to the parallel trend assumption, the dotted line would have to hold in the absence of the treatment.

3.2. Parallel trend assumption

The conventional DID approach strongly relies on assumption (3), i.e. that the average outcomes for the treatment and control groups would have the parallel time trend. However, compositional differences between the treatment and control groups can cause non-parallel trend in the outcomes. If assumption (3) cannot be assumed to hold then:

$$E[Y'_{1T} - Y'_{0T}] = E[Y_{1C} - Y_{0C}] + \Delta$$

and thus the DID estimator will be biased as:

$$\widehat{\delta}_{DID} = \bar{Y}_{1T} - \bar{Y}_{0T} - (Y'_{1T} - \bar{Y}_{0T}) = \bar{Y}_{1T} - \bar{Y}_{0T} - (\bar{Y}_{1C} - \bar{Y}_{0C}) + \Delta$$

$$E[\widehat{\delta}_{DID}] = \alpha + \beta + \gamma + \delta - (\alpha + \gamma) - (\alpha + \beta) - \alpha + \Delta = \delta + \Delta$$

To account for this issue, a vector of covariates $\boldsymbol{X}$ can be introduced to the DID model to adjust for factors that might lead to a violation of the parallel trend assumption (Abadie, 2005).

$$Y_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \delta (T_{it} \cdot G_{it}) + \boldsymbol{\pi X_{it}} + \varepsilon_{it} \tag{3}$$

where $\boldsymbol{X_{it}}$ is the vector of covariates uncorrelated with ε_{it} and $\boldsymbol{\pi}$ is the vector of coefficients.

The key point is that assumption (3) is now still assumed to hold, but is now conditional on covariates X:

$$E[Y'_{1T} - Y'_{0T}|X] = E[Y'_{1C} - Y'_{0C}|X] \tag{4}$$

An additional advantage of including covariates is that the effect of the treatment is allowed to differ among individuals. Finkelstein (2002) and Heckman et al. (1998a,b) have proved the plausibility of this assumption using experimental data.

In addition to using covariates, information on outcomes not just in the periods immediately before and after the intervention, but also in earlier periods can be included to shed light on whether the parallel assumption is reasonable. With observations from only two years it is difficult to know whether it is plausible. However, it is realizable with observations on multiple years (Hastings, 2004). Similar trends of the treatment and control groups in periods prior to the treatment indicate the assumption is valid.

3.3. Dependent and independent variables

3.3.1. Dependent variables

The data used for our analysis includes road accidents recorded in the UK from 2001 to 2004. The casualty data are based on police records and collected by the UK Department for Transport (DfT) and are known as "Road accident data – GB", or the STATS 19 database. The location of an accident is recorded using coordinates which are in accordance with the British National Grid coordinate system. Geographical Information System (GIS) software, MapInfo, was used to locate every individual accident on a map and these accidents were further aggregated at the ward level. A ward is the primary unit of British administrative and electoral geography with an average area of 14 km^2.

3.3.2. Independent variables

To strengthen the parallel trend assumption, we add the following covariates to the DID model:

(1) The data for road network is available from EDINA Digimap. We calculate the length of each type of road and the number of junctions and roundabouts in every ward. Road characteristics have been proved to be related to road casualties by previous research (Noland and Quddus, 2004).
(2) The data for traffic exposure is not available at the ward level. We have employed proxy variables for the potential ward-level exposure. It is assumed that the internal traffic generation of ward i is proportionate to the population (P_i) and the number of employees (E_i) in ward i. We also presume that the external traffic generation of ward i is affected by the population (PP_i) and employee (PE_i) of its proximate ward j ($i \neq j$). The variable,

proximate population and employees (PPE_i), which reflects the external generation of ward i can be constructed as follows:

$$PPE_i = PP_i + PE_i = \sum_{j}^{i\neq j}\frac{P_j}{d_{ij}} + \sum_{j}^{i\neq j}\frac{E_j}{d_{ij}} = \sum_{j}^{i\neq j}\frac{P_j+E_j}{d_{ij}}$$

where d_{ij} is the centroid distance from ward i to ward j. The data for population and employment at the ward level is obtained from Office for National Statistics (ONS), with the former further disaggregated by age cohorts and age percentage.

(3) Recent research suggests child injuries are influenced by factors related to area deprivation (Graham and Stephens, 2008). Therefore the Index of Multiple Deprivation, published by the Office of the Deputy Prime Minister (ODPM, 2004), is used as a control variable. The Index of Multiple Deprivation integrates data on the following seven deprivation domain indices into one overall deprivation score: income, employment, housing and services, health, education, crime and the environment.

3.4. Statistical model

Due to the non negative integer nature of road traffic casualty count data, generalized linear models, including the Poisson and the Negative Binomial models, have been widely used to establish the relationship between traffic casualties and various risk factors.

A generalized linear model with the Poisson distribution is

$$\ln \mu_{it} = \boldsymbol{\pi X_{it}} \tag{5}$$

The assumption that the variance is equal to the mean will be violated when the variance is significantly greater than the mean, also known as over-dispersion. To deal with this problem, the Negative Binomial model has been developed and a Gamma distributed error term is introduced to the Poisson regression model. The structure of NB regression model is:

$$\ln \mu_{it} = \boldsymbol{\pi X_{it}} + \varepsilon_{it} \tag{6}$$

McDonald et al. (2000) considered the DID estimator (referred to as the Before-After Control-impact in their study) when dependent variables are counts. Untransformed, log-transformed data and generalized linear model (GLM) were applied in this study. The authors recommended the GLM for the analysis of count data because assumptions are more likely to be satisfied and interpretation of the estimated parameters is straightforward.

Hence, we adopt the same specification of the DID model as McDonald et al. (2000) Based on Eqs. (1) and (6) we obtain the basic DID model:

$$\ln \mu_{it} = \alpha + \beta T_{it} + G_{it} + \delta(T_{it}\cdot G_{it}) + \varepsilon_{it} \tag{7}$$

Then we further introduce covariates to Eq. (7) and call this model the full DID model, compared with the basic one. In our study, the full DID model is described as:

$$\ln \mu_{it} = \alpha + \beta T_{it} + \gamma G_{it} + \delta(T_{it}\cdot G_{it}) + \pi X_{it} + \varepsilon_{it} \tag{8}$$

The percentage change in the number of accidents due to the effect of the LCC can be obtained as

$$\frac{\mu_{1T}-\mu_{0T}}{\mu_{0T}} = e^{\delta} - 1 \tag{9}$$

3.5. Groups and periods selection for DID

To date, there have been three stages of the LCC: (1) an initial congestion charge in central London from 2003 to 2005; (2) an increase in the congestion charging fee from £5 to £8 from 2005 to 2007; (3) a western extension of the charging zone from 2007 to 2010. In our study, only stage 1 (the initial congestion charge in central London) is investigated because of the problem of analysing correlated multiple treatments. For instance, if we want to estimate the effect of increasing the congestion fee, since the initial LCC may have impacts on the congestion charging zone which persists over time, it is difficult to eliminate this source of confoundness and the consequent estimation may be biased. Thus, we focus on the initial LCC regarding data for 2002 as pre-treatment and for 2003 as post-treatment.

For the geographic units, we choose to define treated and control groups. Since our aim is to evaluate the effect of the London congestion charge, it is logical we choose the central London area as the treatment group.

An important property the control group should have is the independence of treatment, i.e. the control groups should not have received treatment either directly or indirectly through proximity to or interaction with treated groups. A typical example of when this condition is violated is the area outside of the charging zone. It is reasonable to speculate that the travelling and living habits of residents from outside London can be influenced by the LCC. Because of the data limitation, Wales and Scotland are excluded in this research. Although the DID method does not require the treatment and control groups to have the same demographic or traffic characteristics, cities with large population and urban area are preferred as the control groups. Thus, data for conurbations geographically distinct from London is extracted as candidates for the control group, including the central area of the cities of Manchester, Birmingham and Leeds.

As discussed before, the parallel trend assumption is valid conditional on the covariates X. To insure the validity of this assumption, we follow Hastings (2004) to conduct a pre-test which uses more observations from pre-intervention years to assess whether the control group is able to mimic the temporal path of the treatment group. To give an example, Fig. 3 shows the time trend of car casualties in London (the treatment group), Leeds, Manchester and Birmingham (potential control groups). It can be seen that Leeds best reflects the time trend of car casualties in London prior to the introduction of London congestion charge in 2003. It is reasonable that we choose Leeds as the control group in the analysis of car killed and seriously injured (KSI) casualties. However, when we conduct the pre-test for bicycle and motorcycle casualties, Manchester and Birmingham are found to be the most suitable control groups. The result for the control groups is presented below:

(1) Control group for car casualties analysis: Leeds.
(2) Control group for bicycle casualties analysis: Manchester.
(3) Control group for motorcycle casualties analysis: Birmingham.

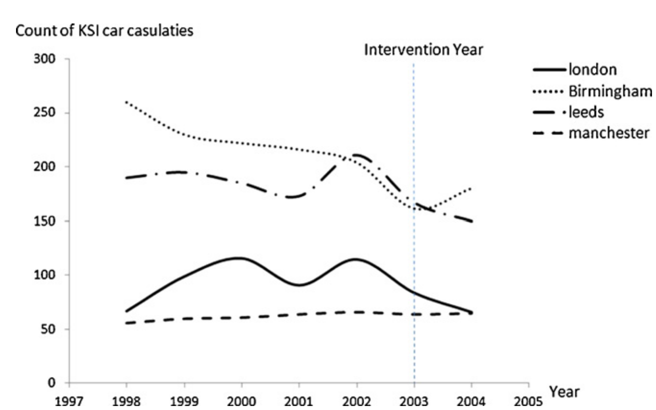

Fig. 3. Time trend for count of car KSI casualties.

3.6. Issues in the DID method

There are two issues that need to be considered when using repeated data sets. The first is regression to the mean (RTM), also known as selection bias. This is a type of bias due to a flaw in the sample selection process, where a systematic difference exists between the characteristics of those samples selected for the study and those which are not. In the context of road safety, selection bias happens when evaluating the effect of treatments that aim to make dangerous sites safer. Black spot sites with high recent accident records are often chosen and their accidents will tend to be lower in subsequent years. The selection bias effect can make random variation appear as a real change caused by interventions and therefore overestimate the effect of a safety treatment.

In a simple before-after control study, the observational effect will consist of:

(1) Changes in accidents attributed to the impact of the intervention on traffic flow and speed.
(2) The national accidents trend.
(3) The RTM effect.

Only the first one is the real intervention effect, while the national accidents trend and the RTM effect are the non-intervention effects or the confounders.

In a naïve before and after study, the national accidents trend can be controlled for by employing a control group which is constructed by a sample of untreated individuals. However, it is unable to control for the RTM effect due to the nature of non-random experiments.

The traditional way to deal with RTM is to apply Empirical Bayes (EB). As a treated-control based approach, the full DID model we introduce in this paper shares the same idea that "accident counts are not the only clue to the safety of an entity; another clue is in what is known about the safety of similar entities" (Hauer et al., 2002). The DID method can control for the RTM effect conditional on the parallel assumption. Let N_i denote the national accidents trend and R_i denote the RTM effect, where $i = T, C$ is as defined before.

(1) Instead of treating the national accidents trend and the RTM effect separately, we consider these two effects as a whole, the time trend effect β, which is as defined before.

$$\beta = N + R$$

(2) Different from the naïve before and after method, in the DID approach, the control group is chosen with a limitation that the treatment and control groups have the parallel path for the accident number in the pre-intervention years. Conditional on the parallel assumption, the time trend effect for the treatment and control groups is the same over the pre- and post-intervention periods. This can be described as:

$$\beta_T = \beta_C$$

which is the same as Eq. (2).

(3) The national accidents trend N is the same for both groups. Then we obtain following equations:

$$\beta_T = N_T + R_T; \quad \beta_C = N_C + R_C$$
$$\beta_T = \beta_C$$
$$N_T = N_C$$
$$R_T = R_C$$

Therefore, conditional on the parallel assumption, the treatment and control groups will have the same national accidents trend and the RTM effect.

(4) Finally, by taking the double differences, the bias due to the time trend effect β is removed thereby controlling for the RTM effect.

Although EB has been used in many control studies, this method is not employed in this paper for two reasons. First, although EB procedures can account for RTM and other effects over time, it relies on a large sample of reference groups, which ideally should have the same or similar characteristics as the treatment group but are unaffected by the treatment. In reality, such types of reference groups are not always available in practice. London is so unique in the UK that no comparable city can be found with the same or similar characteristics. In this case, the EB method is infeasible. Moreover, the word "similar" in the EB method is very ambiguous and clearer justification for selecting reference groups is needed. In contrast, a clear definition of the control group is given in the DID method. In previous sections, we discussed two properties of the control group: (1) it should be independent of the treatment; (2) the treatment and control groups have the similar time trend of accident count. Therefore, the DID method is more flexible and tractable than the EB.

The second issue, as noted by Bertrand et al. (2004), suggests that the conventional DID estimation relying on repeated data sets may suffer from the problem of serial correlation. The standard errors will be underestimated and the *t*-statistics will bias upwards in the presence a positive correlation. Over-rejection of the null hypothesis can cause false inference regarding the effect of treatment. The Durbin–Watson (DW) test is applied to test the presence of serial correlation in the residuals. The value of DW for all models is more than 1.5, which suggests no significant serial correlation is found in this study.

4. Results and discussion

This section presents the results of our DID models. The STATS 19 data classifies the casualty by severity: KSI and slightly injured, which enables the estimation of different types of casualties separately. In this study, one casualty is defined as one accident with one or more persons injured or killed.

4.1. Model selection

Results are obtained from two different specifications of the DID model. First, the basic DID model without any covariate adjustment is applied. The basic model includes only the time fixed effect (CCYear), group fixed effect (CCZone), and the variable of interest (CCYear × CCZone). Next, Eq. (8), or the full DID model, is regressed to compare with the basic DID model by using Bayesian Information Criterion (BIC).

BIC values are obtained from all basic and full DID models for each class of injury. The lower BIC values indicate that the full DID model is superior in interpreting the causal relationship between casualties and the LCC scheme. One possible reason is that only using the dummy variable CCZone cannot adequately explain the internal heterogeneity within the charging zone.

Both Poisson and Negative Binomial models are estimated for car, bicycle and motorcycle accidents. If the dispersion parameter is significantly greater than zero, then the Negative Binomial model provides a better fit than the Poisson model. A likelihood ratio test examines if the dispersion parameter equals zero. The associated chi-squared value and *p*-value are included in the result and these together with the BIC value strongly indicates that the Negative

Table 1
Pearson correlation coefficients of covariates.

	CCZoon	CCYear	CCYear × CCZone	Resident population	Resident population aged 0–15	Resident population aged 16–59
CCZoon	1					
CCYear	0	1				
CCYear × CCZone	0.5925	0.5458	1			
Resident population	−0.5605	0.0251	−0.3218	1		
Resident population aged 0–15	−0.714	−0.021	−0.4229	0.6464	1	
Resident population aged 16–59	−0.4173	0.0464	−0.2317	0.772	0.7126	1
Percentage of resident population aged 0–15	−0.5906	−0.0573	−0.3657	0.4931	0.7871	0.3275
Percentage of resident population aged 16–59	0.6589	0.0717	0.4133	−0.4502	−0.7276	−0.2543
Employee population	0.4923	0.0325	0.29	−0.1987	−0.4163	−0.098
Land area	−0.6979	0	−0.4135	0.4848	0.5468	0.3724
Employee population density	0.7153	0.001	0.4166	−0.5035	−0.6325	−0.7054
Resident population density	0.4651	0.0152	0.2874	0.2117	0.0084	0.3116
Length of minor road	−0.6636	0	−0.3932	0.6074	0.5868	0.5341
Length of motor road	−0.2944	0	−0.1744	0.1321	0.1867	0.0716
Length of A-road	0.0269	0	0.0159	0.3007	0.0811	0.3671
Length of B-road	−0.1473	0	−0.0873	0.3847	0.1798	0.4149
Density of minor road	0.5185	0	0.3072	−0.2322	−0.3477	−0.1356
Density of motor road	−0.2629	−0.0214	−0.1558	0.1137	0.1703	0.0582
Density of A-road	0.0763	−0.0064	0.0436	−0.3457	−0.3252	−0.3224
Density of B-road	0.2123	0	0.1258	0.1794	−0.0196	0.2435
Count of junctions	−0.292	0	−0.173	0.4471	0.3057	0.4536
Count of roundabout	−0.0425	0	−0.0252	0.0914	0.0573	0.0984
IMDscore	−0.6636	0	−0.3932	0.5163	0.7415	0.3825
PPE	0.5426	0	0.3585	−0.7066	−0.8116	−0.5687

	Percentage of resident population aged 0–15	Percentage of resident population aged 16–59	Employee population	Land area	Employee population density	Resident population density
Percentage of resident population aged 0–15	1					
Percentage of resident population aged 16–59	−0.6971	1				
Employee population	−0.4788	0.4491	1			
Land area	0.4409	−0.5487	−0.1248	1		
Employee population density	−0.5235	0.5109	0.5831	−0.5466	1	
Resident population density	−0.0793	0.1528	−0.0457	−0.472	−0.0609	1
Length of minor road	0.3963	−0.4471	0.0764	0.7649	−0.5016	−0.3622
Length of motor road	0.1932	−0.2588	−0.0705	0.6128	−0.2022	−0.2484
Length of A-road	−0.0747	0.1108	0.5069	0.2826	−0.0319	−0.0779
Length of B-road	−0.0073	−0.0338	0.2299	0.2754	−0.1873	0.0419
Density of minor road	−0.3664	0.4595	0.3874	−0.5853	0.4709	0.3626
Density of motor road	0.1791	−0.2304	−0.1256	0.4587	−0.1918	−0.2047
Density of A-road	−0.3078	0.3306	0.4232	0.1797	0.4555	−0.5273
Density of B-road	−0.1284	0.1137	0.1452	−0.1163	−0.0325	0.4515
Count of junctions	0.126	−0.132	0.3606	0.4593	−0.219	−0.1875
Count of roundabout	0.0466	−0.0428	0.2447	0.4668	−0.0575	−0.0849
IMDscore	0.7658	−0.7081	−0.328	0.5237	−0.6569	−0.1709
PPE	−0.6516	0.7116	0.4434	−0.7268	0.8064	0.3397

	Length of minor road	Length of motor road	Length of A-road	Length of B-road	Density of minor road	Density of motor road
Length of minor road	1					
Length of motor road	0.3785	1				
Length of A-road	0.5326	0.0342	1			
Length of B-road	0.3411	0.3744	0.3133	1		
Density of minor road	−0.1497	−0.3027	0.0086	−0.0624	1	
Density of motor road	0.2649	0.5106	−0.0464	0.4387	−0.2626	1
Density of A-road	0.2125	0.1129	0.4881	0.0134	−0.0117	0.0274
Density of B-road	−0.1433	0.1257	0.0261	0.7601	0.0811	0.2094
Count of junctions	0.864	0.1925	0.7753	0.4282	0.1169	0.1198
Count of roundabout	0.3035	0.4855	0.3085	0.2215	−0.2039	0.3391
IMDscore	0.6074	0.1909	0.1691	0.0473	−0.257	0.1596
PPE	−0.704	−0.2999	−0.0392	−0.1932	0.4937	−0.2633

	Density of A-road	Density of B-road	Count of junctions	Count of roundabout	IMDscore	PPE
Density of A-road	1					
Density of B-road	−0.2167	1				
Count of junctions	0.3579	0.0016	1			
Count of roundabout	0.2714	0.0843	0.2614	1		
IMDscore	−0.1001	−0.1525	0.4085	0.1038	1	
PPE	0.1437	0.1489	−0.3421	−0.0683	−0.7267	1

Table 2
Full DID models for car casualties.

Full DID models	Model 1 car all casualties		Model 2 car KSI (Poisson)		Model 3 car slight injured	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	2.73E−01	(4.17E−02)	1.05E+00	(1.06E−03)	2.24E−01	(5.83E−02)
CCYear	−2.54E−01	(8.22E−03)*	−1.48E−01	(3.12E−02)*	−2.52E−01	(7.91E−03)**
CCYear × CCZone	−5.36E−02	(1.25E−04)*	−1.53E−01	(4.95E−02)*	−4.67E−02	(1.73E−04)*
Resident population	4.05E−04	(8.45E−05)*	8.45E−04	(3.73E−04)**	3.37E−04	(6.02E−05)
Resident population aged 0–15	−5.50E−04	(1.43E−04)	−9.65E−04	(2.03E−04)*	−4.78E−04	(1.50E−04)
Resident population aged 16–59	−1.61E−04	(2.42E−05)	−7.80E−04	(2.57E−05)	−1.16E−04	(1.48E−05)
Percentage of resident population aged 0–15	4.06E+00	(8.93E−01)	7.78E+00	(2.08E+00)	3.30E+00	(9.70E−01)**
Percentage of resident population aged 16–59	2.19E+00	(9.84E−01)**	7.32E+00	(2.52E+00)	1.87E+00	(1.06E+00)
Employee population	5.30E−06	(1.21E−06)*	7.71E−06	(3.61E−07)*	5.52E−06	(1.31E−06)*
Land area	1.31E−01	(4.05E−02)	2.60E−01	(1.49E−02)**	1.23E−01	(4.82E−02)
Employee population density	−2.31E−06	(1.33E−06)***	−2.25E−06	(2.13E−06)	−3.69E−06	(1.61E−06)*
Resident population density	−8.01E−05	(1.50E−05)*	6.70E−05	(2.06E−05)*	−8.52E−05	(1.59E−05)*
Length of minor road	2.39E−05	(1.97E−05)	9.88E−05	(1.42E−05)*	2.32E−05	(2.25E−05)
Length of motor road	2.17E−04	(7.60E−05)*	−7.08E−04	(1.97E−04)*	−2.00E−04	(7.20E−05)
Length of A-road	1.08E−04	(3.74E−06)*	1.99E−04	(5.10E−05)*	1.02E−04	(8.51E−06)*
Length of B-road	−1.99E−04	(4.58E−05)	−1.94E−04	(1.22E−04)	−1.98E−04	(4.37E−05)
Density of minor road	1.17E−04	(1.26E−05)*	4.87E−05	(1.66E−06)*	1.19E−04	(1.47E−05)*
Density of motor road	2.40E−01	(9.79E−02)**	6.05E−01	(1.35E−01)*	2.24E−01	(9.32E−02)**
Density of A-road	8.48E−02	(7.80E−03)	−1.54E−02	(3.77E−02)	9.74E−02	(7.53E−03)
Density of B-road	6.37E+00	(4.40E−01)*	3.82E+00	(1.79E−01)*	6.45E+00	(4.16E−01)*
Count of junctions	−4.26E−03	(9.28E−04)	−9.25E−03	(1.48E−03)***	−4.14E−03	(9.80E−04)*
Count of roundabout	9.57E−03	(1.96E−02)	−2.76E−02	(4.51E−02)	8.07E−03	(1.82E−02)
IMDscore	1.43E−02	(4.44E−03)*	1.16E−02	(3.09E−03)*	1.40E−02	(4.90E−03)*
PPE	−4.41E−04	(4.28E−04)	1.37E−04	(3.12E−04)**	−4.44E−04	(4.44E−04)
Constant	−1.80E+00	(3.49E−01)**	−9.08E+00	(2.44E+00)*	−1.48E+00	(3.65E−01)*
Obs	244		244		244	
BIC	446.55		349.22		444.55	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 76.73 Prob ≥ chibar2 = 0.000		chibar2(01) = 0.0E+00 Prob ≥ chibar2 = 0.500		chibar2(01) = 67.96 Prob ≥ chibar2 = 0.000	

* Figures are significant at: 99%.
** Figures are significant at: 95%.
*** Figures are significant at: 90%.

Binomial fits better than Poisson for most models, except for model (2), (7), (8), (9) and (14) (see Table 1).

4.2. London congestion charge

The results show a distinct reduction in car casualties, despite an increase in bicycle and motorcycle accidents.

In terms of total car casualties, significant effects (at the 95% level) exist for the LCC with a coefficient (standard error) of −0.054 (0.001), suggesting that the introduction of the LCC scheme reduced car casualties within the charging zone by 5.2% (Table 2). This is most likely due to the fact that the traffic volume has decreased since the introduction of the LCC (Tuerk and Graham, 2010), coupled with the fact that the total number of accidents decreases as the traffic volume decreases (Lord et al., 2005). The result for car KSI also shows a remarkable drop of 14.2% due to the LCC, while the reduction in slightly injured accidents is 4.6% (see Table 3).

According to a report from TfL (2004), there is a decrease in the number of two-wheelers involved in accidents. The original data also suggests a reduction in the absolute number of cycle-related accidents from 1353 accidents (year 2002) to 1254 accidents (year 2003). However, after controlling for the time trend and regional effects, we find that an increase of 5.7% in total motorcycle casualties is related to the LCC, while the increase for bicycle casualties is 13.3% (Tables 4 and 6). The numbers of KSI for bicycle and motorcycle increase by 2.7% and 17.3% respectively during 2003. It is suggested that the year-on-year decrease in slightly injured two-wheeled continues and is even greater in 2003 (TfL, 2004). We again fit separate models to bicycle and motorcycle slight injured data sets. The model shows the LCC scheme has increased both bicycle and motorcycle slight injured casualties by 13.5% and 1.8% respectively (see Tables 5 and 7).

The result is, to a large extent, consistent with the conclusion of previous research by Noland et al. (2008). They hypothesized that this effect was down to the increasing number of two-wheeled commuters and the increased average network traffic speeds. And according to the annual report from TfL (2004), inbound two-wheeled vehicles have increased by 15% and the average traffic speed has increased by 31% after introduction of the LCC.

These results therefore suggest that the LCC scheme plays an important role in influencing traffic casualties in the London congestion charging zone.

4.3. Control variables in full DID models

The DID models provide some other interesting results on the effect of covariates on accidents counts. The time effect variable CCYear has a significantly negative relationship with the number of traffic casualties (at the 99% level) for all models, indicating a nationwide downtrend of traffic accidents. This is in part possibly related to other traffic laws and policies, such as speed limits, seatbelt law and improvements in the road infrastructure. The other dummy variable CCZone, which controls for regional differences, also proved to be significant in most models.

The absolute numbers as well as the density of ward population and employment are used to control for traffic exposure within each ward. Most models show there are positive effects from the scale of population and employment. That implies that more accidents may occur in wards with more residents and job opportunities. In contrast, coefficients for ward resident density in most models are negative, implying that regions of high resident density may experience fewer accidents. Considering many previous studies, this result is unsurprising. The variable PPE, which reflects external traffic generation, shows positive effects in most models.

Table 3
Basic DID models for car casualties.

Basic DID models	Model 4 car all casualties		Model 5 car KSI		Model 6 car slight injured	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	−3.81E−01	(4.92E−10)*	7.20E−01	(2.49E−09)*	−4.65E−01	(4.87E−10)*
CCYear	−2.22E−01	(2.18E−11)*	−2.01E−01	(2.18E−14)*	−2.23E−01	(2.92E−11)*
CCYear × CCZone	−3.17E−02	(5.43E−11)*	−1.10E−01	(4.34E−10)*	−1.19E−02	(1.12E−10)*
Constant	3.94E+00	(4.92E−10)*	6.93E−01	(3.47E−14)*	3.90E+00	(4.87E−10)*
Obs	244		244		244	
BIC	489.39		463.03		487.79	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 2245.44		chibar2(01) = 95.46		chibar2(01) = 2052.78	
	Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000	

* Figures are significant at: 99%.

Table 4
Full DID models for cycle casualties.

Full DID models	Model 7 bicycle all casualties (Poisson)		Model 8 bicycle KSI (Poisson)		Model 9 bicycle slight injured (Poisson)	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	1.45E+00	(3.71E−01)*	5.06E−01	(6.38E−02)*	1.63E+00	(4.33E−01)*
CCYear	−1.14E−01	(4.77E−04)*	−1.73E−01	(1.71E−02)*	−1.01E−01	(7.11E−03)*
CCYear × CCZone	1.25E−01	(4.76E−03)*	2.67E−02	(2.41E−02)	1.27E−01	(1.20E−03)*
Resident population	1.14E−03	(7.83E−04)	1.29E−04	(5.42E−04)	1.34E−03	(8.51E−04)
Resident population aged 0–15	−1.93E−04	(1.71E−03)	−1.53E−03	(8.09E−04)***	2.91E−06	(1.92E−03)
Resident population aged 16–59	2.39E−04	(4.26E−04)	−5.02E−04	(1.39E−03)	3.77E−04	(3.46E−04)
Percentage of resident population aged 0–15	6.77E+00	(2.27E+00)*	4.27E+00	(1.36E+00)*	7.27E+00	(2.31E+00)*
Percentage of resident population aged 16–59	5.52E+00	(2.53E+00)**	4.95E+00	(1.59E+00)*	5.71E+00	(2.64E+00)**
Employee population	5.70E−06	(1.72E−06)*	2.55E−06	(1.78E−06)	6.24E−06	(1.63E−06)*
Land area	3.02E−02		9.50E−04	(2.18E−02)	3.85E−02	(6.16E−03)*
Employee population density	−3.41E−07	(8.63E−07)	−3.84E−06	(3.60E−07)*	8.11E−08	(1.08E−06)
Resident population density	−5.79E−05	(1.30E−05)*	−7.08E−05	(3.28E−05)	−5.42E−05	(1.91E−05)*
Length of minor road	−3.34E−05	(1.63E−05)**	−3.21E−06	(2.01E−05)	−4.08E−05	(1.40E−05)*
Length of motor road	−5.16E−05	(3.92E−05)	−4.63E−06	(5.31E−06)	−6.21E−05	(4.20E−05)
Length of A-road	1.15E−06	(3.78E−05)	5.36E−05	(3.76E−05)	−9.09E−06	(3.56E−05)
Length of B-road	−1.15E−04	(2.26E−05)*	−2.37E−04	(8.69E−05)*	−8.94E−05	(7.91E−06)*
Density of minor road	3.39E−05	(3.55E−06)*	−1.86E−05	(6.72E−05)	4.30E−05	(1.26E−05)*
Density of motor road	4.28E−04	(1.68E−03)	1.36E−03	(1.29E−04)	−7.81E−05	(1.85E−03)
Density of A-road	−1.21E−03	(1.50E−02)	−2.72E−02	(9.21E−03)*	3.04E−03	(1.41E−02)
Density of B-road	2.52E+00	(1.03E+00)**	7.65E+00	(1.43E+00)*	1.53E+00	(1.42E+00)
Count of junctions	3.65E−03	(1.04E−03)*	2.42E−03	(1.36E−04)*	4.06E−03	(9.77E−04)*
Count of roundabout	2.11E−02	(9.76E−02)	−1.51E−01	(1.02E−01)	5.65E−02	(9.55E−02)
IMDscore	1.78E−02	(1.58E−03)**	2.17E−02	(1.69E−02)	1.69E−02	(1.58E−03)*
PPE	1.12E−03	(6.03E−05)*	1.62E−03	(5.84E−04)*	1.02E−03	(2.50E−05)*
Constant	−6.22E+00	(2.88E+00)*	−6.26E+00	(2.43E+00)*	−6.78E+00	(2.98E+00)*
Obs	244		244		244	
BIC	307.52		339.29		305.37	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 0.0E+00		chibar2(01) = 0.0E+00		chibar2(01) = 0.0E+00	
	Prob ≥ chibar2 = 0.500		Prob ≥ chibar2 = 0.500		Prob ≥ chibar2 = 0.500	

* Figures are significant at: 99%.
** Figures are significant at: 95%.
*** Figures are significant at: 90%.

Table 5
Basic DID models for cycle casualties.

Basic DID models	Model 10 bicycle all casualties		Model 11 bicycle KSI		Model 12 bicycle slight injured	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	1.24E+00	(2.02E−10)*	7.42E−01	(8.99E−14)*	1.32E+00	(1.18E−10)*
CCYear	−7.67E−02	(5.11E−11)*	−1.34E−01	(1.19E−17)*	−6.45E−02	(3.63E−11)*
CCYear × CCZone	2.95E−02	(2.08E−10)*	2.18E−02	(5.07E−11)*	6.81E−02	(1.20E−10)*
Constant	1.67E+00	(1.06E−12)*	−3.08E−02	(9.18E−17)*	1.47E+00	(2.68E−15)*
Obs	244		244		244	
BIC	499.39		473.27		493.89	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 460.26		chibar2(01) = 14.89		chibar2(01) = 409.68	
	Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000	

* Figures are significant at: 99%.

Table 6
Full DID models for motorcycle casualties.

Full DID models	Model 13 motorcycle all casualties		Model 14 motorcycle KSI (Poisson)		Model 15 motorcycle slight injured	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	2.55E+00	(1.63E−02)*	2.08E+00	(2.71E−01)*	2.62E+00	(7.59E−02)*
CCYear	−2.61E−01	(1.68E−02)*	−3.40E−01	(6.52E−03)*	−2.43E−01	(1.93E−02)*
CCYear × CCZone	5.57E−02	(1.15E−02)*	1.60E−01	(7.09E−03)*	1.83E−02	(4.59E−04)*
Resident population	2.48E−03	(2.80E−05)*	1.15E−03	(8.22E−04)	2.50E−03	(7.56E−04)*
Resident population aged 0–15	−3.60E−03	(2.39E−03)	−4.88E−04	(3.33E−03)	−3.34E−03	(1.51E−03)**
Resident population aged 16–59	−2.90E−03	(4.80E−04)*	−9.49E−04	(1.18E−03)	−2.92E−03	(3.71E−04)*
Percentage of resident population aged 0–15	8.35E+00	(6.01E−01)*	1.58E+00	(1.53E+00)	1.02E+01	(1.47E+00)*
Percentage of resident population aged 16–59	1.23E+01	(1.71E−01)*	4.00E−03	(5.24E−01)	1.46E+01	(1.03E+00)*
Employee population	4.07E−06	(5.46E−07)*	−3.18E−06	(1.18E−06)*	4.62E−06	(7.24E−07)*
Land area	2.27E−01	(8.87E−02)*	1.96E−01	(8.15E−02)**	2.43E−01	(1.09E−01)**
Employee population density	−7.51E−06	(1.86E−06)*	−4.82E−06	(8.48E−07)*	−4.57E−06	(1.12E−06)*
Resident population density	−4.92E−05	(1.21E−05)*	−9.38E−05	(2.47E−05)*	−4.50E−05	(2.81E−06)*
Length of minor road	−3.20E−05	(2.14E−05)	−5.20E−06	(3.79E−05)	−4.50E−05	(2.77E−05)
Length of motor road	−3.87E−04	(2.37E−05)*	−3.31E−04	(9.67E−05)*	−3.49E−04	(2.53E−06)*
Length of A-road	−8.66E−06	(1.11E−05)	7.69E−05	(8.73E−06)*	−2.87E−05	(2.85E−06)*
Length of B-road	3.22E−06	(1.10E−04)	7.47E−05	(1.00E−04)	−7.53E−06	(5.10E−05)
Density of minor road	4.34E−05	(7.40E−09)*	1.47E−04	(2.90E−05)*	1.95E−05	(5.34E−06)*
Density of motor road	2.89E−01	(6.58E−02)*	1.87E−01	(5.74E−02)*	2.79E−01	(6.51E−02)*
Density of A-road	7.39E−02	(1.15E−02)*	6.76E−02	(2.36E−02)*	5.30E−02	(1.30E−02)*
Density of B-road	1.38E+00	(1.36E+00)	2.26E+00	(1.06E+00)**	5.56E−01	(8.54E−01)
Count of junctions	2.45E−03	(1.27E−03)***	−3.00E−03	(1.11E−03)*	3.96E−03	(1.72E−03)**
Count of roundabout	7.45E−02	(4.16E−03)*	1.60E−01	(7.05E−02)**	4.56E−02	(2.25E−02)**
IMDscore	2.08E−02	(4.76E−03)*	1.44E−02	(3.96E−03)*	2.34E−02	(5.74E−03)*
PPE	2.55E−04	(9.28E−05)*	−6.27E−05*	(4.64E−04)	4.07E−04	(2.41E−04)***
Constant	−1.19E+01	(2.93E−01)	−3.68E+00	(2.00E−01)*	−1.42E+01	(7.61E−01)*
Obs	268		268		268	
BIC	418.59		334.04		414.78	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 18.07		chibar2(01) = 0.0E+00		chibar2(01) = 5.35	
	Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 1		Prob ≥ chibar2 = 0.01	

* Figures are significant at: 99%.
** Figures are significant at: 95%.
*** Figures are significant at: 90%.

Table 7
Basic DID models for motorcycle casualties.

Basic DID models	Model 16 motorcycle all casualties		Model 17 motorcycle KSI		Model 18 motorcycle slight injured	
	Coef. (std. err.)		Coef. (std. err.)		Coef. (std. err.)	
CCZoon	1.77E+00	(3.96E−08)*	1.18E+00	(1.89E−10)*	1.90E+00	(1.79E−08)*
CCYear	−2.49E−01	(4.28E−09)*	−4.05E−01	(1.03E−16)*	−2.06E−01	(3.56E−09)*
CCYear × CCZone	3.37E−02	(4.26E−09)*	4.24E−01	(2.22E−11)*	−4.90E−02	(6.65E−08)*
Constant	1.63E+00	(2.43E−08)*	1.67E−01	(1.97E−17)*	1.36E+00	(4.48E−11)*
Obs	268		268		268	
BIC	453.37		471.63		475.89	
Likelihood-ratio test of alpha = 0:	chibar2(01) = 780.68		chibar2(01) = 38.86		chibar2(01) = 654.08	
	Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000		Prob ≥ chibar2 = 0.000	

* Figures are significant at: 99%.

This result suggests that the higher level of population and employment in proximate wards is related to more accidents. The ward population is further categorized by age cohorts and percentage, although these are not significant in all models.

Another set of independent variables are the characteristics of the road infrastructure, including the length and density of motorway, A-road, B-road and minor road. We find that increased car casualties are associated with motorway and A-road length. Since motorway and A-road are designed mainly for high speed vehicles, it is reasonable to suppose that the risk for cars is relatively higher. The B-roads and minor roads are expected to have lower speeds and more cyclists and the results suggest that both B-road and minor road density have significant positive effects on cyclist casualties. The coefficients for counts of junctions and roundabouts are found to be less significant for car casualties, except for cycle-related accidents. It is significant (at the 95% level) that more roundabouts and junctions in wards result in more cycle-related injuries. One interpretation for this outcome may be the complexity of junctions and roundabouts and fewer safeguards, making cyclists more vulnerable.

Socio-economic deprivation has been proved to be positively related to road traffic casualties, which have again been confirmed by our results which indicate that IMD scores have positive effects on all types of casualties.

5. Conclusions

Our analysis has presented new evidence on the impact of the London congestion charge scheme on traffic casualties. We have employed the Difference-In-Difference estimation method which can eliminate biases due to regional differences and nationwide trends. This paper is the first use of this method in assessing the road safety impacts of the LCC.

We have further developed the DID model by combining it with generalized linear regression models. Covariates are included in the full DID to adjust for factors that violate the parallel trend assumption. By comparing the BIC value we can see that the full DID models perform better than the basic ones.

One of the original objectives of the LCC is to reduce road accidents. Our regression results suggest that with respect to car casualties at least, there has been a significant reduction within the congestion charging area. These improvements can be interpreted by the decrease in traffic flow and a better driving environment, which can be attributed to the LCC. However, there are two major areas of concern: cycle-related accidents have increased after the LCC because more motorcycles and bicycles have replaced cars. We also found other variables, including population, employment, deprivation, road infrastructure and productivity, have affected casualties significantly, thereby confirming the conclusions from previous studies.

Although the introduction of the London congestion charge has increased traffic speed and reduced congestion effectively, the authorities need to be aware of the mode shift effect to two-wheeled vehicles, which can create a large class of vulnerable road users with a greater number of accidents. More attention therefore needs to be paid to road safety strategies when introducing congestion charging, especially for cyclists.

The example of London's effort in this area of cycling usage and safety strategies is instructive. The last year of road accidents analysed in this paper was 2004. This marked the beginning of a profound modal shift in London from car to cycling that has seen cycle trip rates increase by 46%, whilst the amount of travel by all modes remained broadly the same, between 2005/6 to 2010/11. This is equivalent to an average rate of growth for cycling of around 7% per year in this period. However, this headline figure conceals the fact that between 2009/10 and 2010/11, the number of cycle journeys made by London residents grew at the rather faster rate of 13% per year (TfL, 2011).

A number of recent initiatives have encouraged cyclists in London, e.g. 2010 was the Mayor's Year of Cycling and saw the launch of Barclays Cycle Hire, the Barclays Cycle Superhighways and 13 'Biking Boroughs', alongside a wide range of interventions to improve conditions for cyclists and to raise the profile of cycling in London. Whilst the impact of the full scale of change resulting from this activity has yet to be properly assessed, by end of 2010, there were 30,000 more cycle journeys made every day across the city (including for commuting trips), and around 90,000 new London households bought a bike for the first time in 2010, when compared to 2009.

Whilst our results would imply that the sheer increase in the numbers of people cycling is of major concern, the four Barclays Cycle Superhighways that have contributed to the increased cycle flows and encouraged existing cyclists to increase the amount they cycle, also attempt to make the cycling experience safer. TfL surveys indicate that cyclists generally agree that they benefit from an improved journey experience as a result of the introduction of the routes, e.g. improved quality of road surface, the visibility of the blue markings. In particular the majority of users agreed that the routes make them feel safer when cycling (TfL).

Our paper highlighted the impact of the increase in cycle related accidents in the early years of the LCC and the consequent need for authorities to have adequate road safety measures to account for any modal shift to cycling. Transport for London indeed appears to have considered certain measures, including better education (TfL, 2009), though recent press campaigns highlight the need for further actions.

While we developed the basic DID model, one issue outstanding is that our model cannot account for multiple treatments. For example, if we want to evaluate the effect of the western extension of London congestion charge, it is impossible to remove the effect of the initial LCC using our DID model. In addition, proxy variables are employed instead of direct exposure variables such as vehicle-miles-travelled and traffic flow and this may well affect the results.

Acknowledgement

The authors are grateful to The Lloyd's Register Educational Trust for supporting this research.

References

Abadie, A., Gardeazabal, J., 2003. The economic costs of conflict: a case study of the Basque Country. American Economic Review 93 (1), 113–132.
Abadie, A., 2005. Semiparametric difference-in-differences estimators. The Review of Economic Studies 72 (1), 1–19.
Abadie, A., Diamond, A., Hainmueller, J., 2010. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. Journal of the American Statistical Association 105, 493–505.
Ashenfelter, O., Card, D., 1985. Using the longitudinal structure of earnings to estimate the effect of training programs. Review of Economics and Statics 67, 648–660.
Athey, S., Imbens, G., 2006. Identification and inference in nonlinear difference-in-differences models. Econometrica 74, 431–497.
Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? Quarterly Journal of Economics 119, 249–275.
Box, G.E.P., Tiao, G.C., 1975. Intervention analysis with applications to economic and environmental problems. Journal of the American Statistical Association 70 (349), 70–79.
Card, D., Krueger, A.B., 1994. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. The American Economic Review 8 (44), 772–793.
Card, D., Krueger, A.B., 2000. A re-analysis of the effect of the New Jersey minimum wage with representative payroll data. American Economic Review 90, 1397–1420.
Central London Congestion Charging Impacts Monitoring Fifth Annual Report. Transport for London, 2007.
Central London Congestion Charging Impacts Monitoring Second Annual Report. Transport for London, 2004.
Cycle Safety Action Plan. Transport for London, 2009.
Crandall, R.W., Graham, J.D., 1989. The effect of fuel economy standards on automobile safety. Journal of Law and Economics 32, 97–118.
Dissanayake, D., Aryaija, J., Wedagama, D.M.P., 2009. Modelling the effects of land use and temporal factors on child pedestrian casualties. Accident Analysis and Prevention 41, 1016–1024.
Donald, S., Lang, K., 2007. Inference with difference in differences and other panel data. Review of Economics and Statics 89, 221–233.
Eliasson, J., Mattsson, L., 2006. Equity effects of congestion pricing quantitative methodology and a case study for Stockholm. Transportation Research Part A 40, 602–620.
Elvik, R., 2011. Assessing causality in multivariate accident models. Accident Analysis and Prevention 43, 253–264.
Finkelstein, A., 2002. The effect of tax subsidies to employer-provided supplementary health insurance: evidence from Canada. Journal of Public Economics 84, 305–339.
Grabowski, D.C., Morrisey, M.A., 2004. Gasoline prices and motor vehicle fatalities. Journal of Policy Analysis and Management 23 (3), 575–593.
Grabowski, D.C., Michael, T., Morrisey, A., 2006. Do higher gasoline taxes save lives? Economics Letters 90, 51–55.
Graham, D.J., Stephens, D.A., 2008. Decomposing the impact of deprivation on child pedestrian casualties in England. Accident Analysis and Prevention 40, 1351–1364.
Hastings, J.S., 2004. Vertical relationships and competition in retail gasoline markets: empirical evidence from contract changes in Southern California. The American Economic Review 94 (1), 317–328.
Hauer, E., Harwood, D.W., Council, F.M., Griffith, M.S., 2002. Estimating safety by the empirical Bayes method: a tutorial. Transportation Research Record 1784, 126–131.
Haughton, J., Sarkar, S., 1996. Gasoline tax as a corrective tax: estimates for the United States, 1970–1991. The Energy Journal 17 (2), 103–126.
Heckman, J., Ichimura, H., Todd, P., 1998a. Matching as an econometric evaluation estimator. Review of Economic Statics 65, 261–294.
Heckman, J., Ichimura, H., Smith, J., Todd, P., 1998b. Characterizing selection bias using experimental data. Econometrica 66 (5), 1017–1098.
Hyatt, E., Griffin, R., Rue III, L.W., McGwin Jr., G., 2009. The association between price of regular-grade gasoline and injury and mortality rates among occupants involved in motorcycle- and automobile-related motor vehicle collisions. Accident Analysis and Prevention 41, 1075–1079.
Leigh, J.P., Wilkinson, J.T., 1991. The effect of gasoline taxes on highway fatalities. Journal of Policy Analysis and Management 10 (3), 474–481.

Lord, D., Manar, A., Vizioli, A., 2005. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. Accident Analysis and Prevention 37, 185–199.
McDonald, T.L., Erickson, W.P., McDonald, L.L., 2000. Analysis of count data from before-after control-impact studies. Journal of Agricultural, Biological, and Environmental Statistics 5 (3), 262–279.
Noland, R.B., Quddus, M.A., 2004. A spatially disaggregate analysis of road casualties in England. Accident Analysis and Prevention 36, 973–984.
Noland, R.B., Quddus, M.A., Ochieng, W.Y., 2008. The effect of the London congestion charge on road casualties: an intervention analysis. Transportation 35, 73–91.
ODPM, 2004. The English Indices of Deprivation 2004 (revised). HMSO, London.
Olszewski, P., Xie, L., 2005. Modelling the effects of road pricing on traffic in Singapore. Transportation Research Part A 39, 755–772.
Quddus, M.A., 2008a. Modelling area-wide count outcomes with spatial correlation and heterogeneity: an analysis of London crash data. Accident Analysis and Prevention 40, 1486–1497.
Quddus, M.A., 2008b. Time series count data models: an empirical application to traffic accidents. Accident Analysis and Prevention 40, 1732–1741.
Travel in London, Report 4. Transport for London, 2011.
Tuerk, A., Graham, D.J., 2010. Traffic Impacts of the Central London Congestion Charging Scheme. Working Paper. Imperial College London, London.
White, M.J., 2004. The arms race on American roads: the effect of heavy vehicles on traffic safety and failure of liability rules. Journal of Law and Economics 47 (2), 333–356.
Wichiensin, M., Bell, M.G.H., Yang, H., 2007. Impact of congestion charging on the transit market: an inter-modal equilibrium model. Transportation Research Part A 41, 703–713.
Wier, M., Weintrauba, J., Humphreys, E.H., Seto, E., Bhatia, R., 2009. An area-level model of vehicle-pedestrian injury collisions with implications for land use and transportation planning. Accident Analysis and Prevention 41, 137–145.