Accident Analysis and Prevention 220 (2025) 108168

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Safe streets for cyclists? Quantifying the causal impact of cycling infrastructure interventions on safety

Anupriya [a,*], Xiaowei Zhu [b], Emma McCoy [c], Daniel J. Graham [a,b]

[a] Transport Strategy Centre, Imperial College London, London SW7 2AZ, UK
[b] Department of Mathematics, Imperial College London, London SW7 2AZ, UK
[c] Department of Statistics, London School of Economics, London WC2A 2AE, UK

ARTICLE INFO

Keywords:
Cycleway investments
Demand
Safety
Causal analysis
Heterogeneous impact

ABSTRACT

London's Cycle Superhighways (CS) form a network of cycle routes connecting central London to outer boroughs, introduced in 2010 to promote cycling and improve safety. This paper examines their causal impact on cycling volume and safety using detailed road traffic and road safety data from the UK's Department for Transport. To estimate these effects, we employ propensity score-matched difference-in-differences and panel outcome regression models, comparing two distinct infrastructure types: segregated and non-segregated CS. A key contribution of this study is the development of a novel safety indicator — the normalised collision rate — that accounts for changes in cyclist volume (exposure) while incorporating expected non-linearities in the relationship between collisions and exposure. Our findings indicate that non-segregated CS did not increase cycling volume but led to a substantially higher collision rate. This increase appears to be driven by a post-intervention surge in the proportion of new, inexperienced cyclists along these routes. In contrast, segregated CS effectively increased cycling volume without increasing collision rates. Further, an evaluation of a major segregation upgrade along an existing non-segregated CS route revealed a notable reduction in collision rates. These results highlight the crucial role of segregated infrastructure in not only encouraging cycling but also ensuring it remains a safe and viable urban transport option.

1. Introduction

In recent years, governments worldwide have increasingly prioritised sustainable transport solutions, such as cycling and walking, as key strategies to address urban mobility challenges, reduce greenhouse gas emissions, and promote public health. Among the various initiatives, the development of Cycle Superhighways (CS) has garnered significant attention as a means to enhance cycling infrastructure, encourage greater participation, and improve cyclist safety. Substantial investments are being made in these projects, reflecting a strong commitment to transforming urban landscapes and fostering a culture of cycling. In line with these efforts, CS were introduced across London in 2008. Fig. 1 illustrates an initial route map of the CS, which are dedicated cycle pathways extending from the outskirts of London to its centre.[1] These pathways were designed to facilitate safer, quicker, and more direct travel within the city. Similar initiatives have also been launched across North America, Australia, and Europe to support long-distance cycle commutes in metropolitan areas (Pucher and Buehler, 2017).

As funding for these projects continues to rise, it is crucial to rigorously assess the effectiveness of CS in achieving their intended goals. Understanding the causal impact of these interventions on cycling volume and cyclist safety is essential for policymakers to make informed decisions about resource allocation and infrastructure development. This paper aims to explore the relationship between CS and cycling behaviour, while also examining how these interventions contribute to creating safer environments for cyclists. By analysing empirical evidence and evaluating the outcomes of different CS designs, we seek to provide insights that will guide future investments in sustainable transportation.

CS introduced in London incorporated a variety of measures to improve cyclist safety including (1) realigned traffic and bus lanes to create more space for cyclists on busy stretches of the routes, (2) redesigned junctions to make them safer for cyclists (say, by removing left-turn slip roads), (3) blind-spot visibility mirrors at signalised junctions in order to improve the visibility of cyclists to heavy goods vehicle drivers, (4) new advanced stop lines and extensions to existing ones

* Corresponding author.
E-mail address: anupriya15@imperial.ac.uk (Anupriya).

[1] https://tfl.gov.uk/modes/cycling/routes-and-maps/cycleways.

https://doi.org/10.1016/j.aap.2025.108168
Received 5 March 2025; Received in revised form 19 June 2025; Accepted 11 July 2025
Available online 30 July 2025

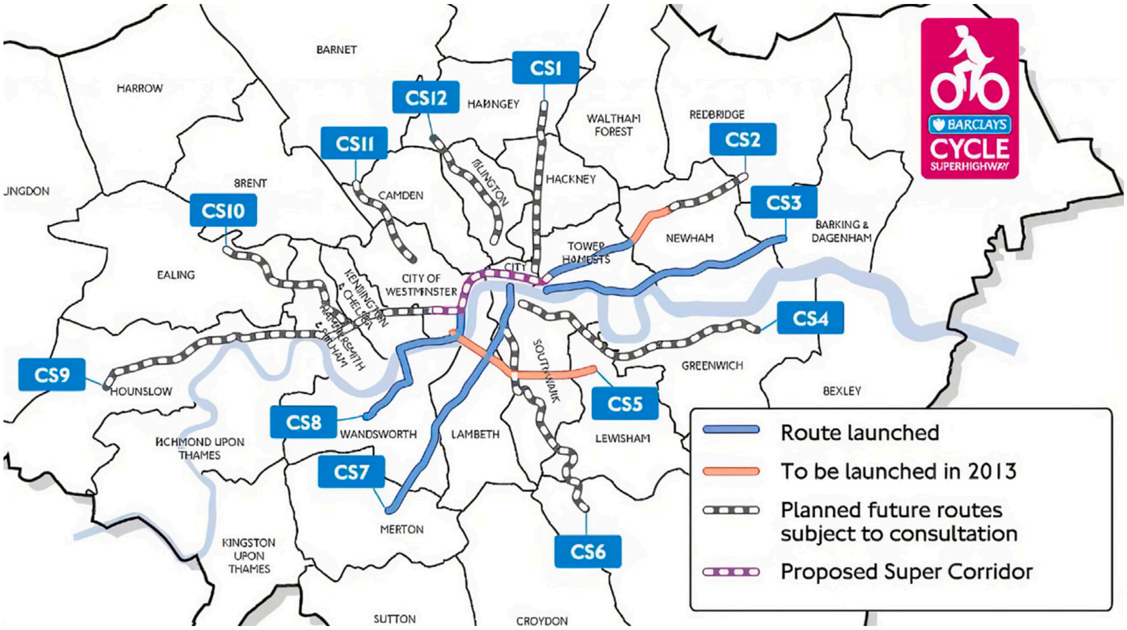

Fig. 1. Route plan map of the Cycle Superhighways in London.

(to a minimum of 5 m) in order to help cyclists move away from traffic signals before other traffic, and, (5) segregated cycle lanes at particularly busy sections of the routes, including Stockwell Gyratory and Wandsworth Bridge roundabout (Transport for London, 2011b). However, the initial implementation of CS drew widespread criticism as the promised improvements were found to be only sporadically implemented along the CS routes. Critics thus claimed the safety impacts of CS were overstated and referred to CS as "nothing but blue paint".[2] Given the infrastructure costs associated with the introduction of CS, it is imperative to understand the traffic impacts of CS, particularly those related to cyclist safety. In this paper, we investigate the causal impact of CS on cycle flow volume, cycle collision counts, and a novel cyclist safety indicator — the normalised cycle collision rate. The proposed indicator normalises observed collision counts by expected collision counts, thereby accounting for cyclist volume and mitigating bias from changes in exposure.

Several ex-post evaluations have been carried out in the past to understand the traffic impacts of cycle lanes, especially in regards to collisions (see DiGioia et al., 2017, for a detailed review). These studies mostly compare crashes before and after the deployment of cycle lanes to quantify the effects of the intervention. Studies based on such an approach can produce unreliable inference due to the presence of confounding bias and a failure to account for extraneous temporal trends (Graham, 2025; Mannering et al., 2020). Confounding can occur primarily from the non-random nature of infrastructure investments. In other words, there may exist confounding factors that determine both the likelihood of the intervention and the resulting demand and safety impacts. For instance, CS are more likely to be chosen for roads with large cycle flow volumes, however, there is an inherent scale effect with respect to the evaluation of safety impacts: more cycling usually implies higher cycle-related collisions. Thus, the estimates derived from a simple before-after comparison of demand and safety indicators may not reflect the true effect of the intervention.

In this study, we adopt two causal inference approaches, namely, (i) propensity score matched difference-in-differences, and (ii) panel outcome regression with fixed effects. Under fairly general conditions, these approaches allow unbiased estimation of the causal effect by effectively adjusting for confounding and temporal trends. Our analysis uses road traffic and road safety data from the UK Department for Transport. The closest precedent to our analysis is the study by Li et al. (2017) that quantified the causal impact of CS in London on cycling volume and collision counts at the network (that is, aggregate) level. However, we exploit the granularity of the data at hand to estimate the impact of CS on different infrastructure types, in particular, segregated versus non-segregated CS. We thus contribute novel insights on how CS segments with varying underlying features perform with respect to each other.

The structure of this paper is as follows. Section 2 presents a review of the previous studies on the demand and safety effects of cycle lanes. Section 3 describes the methodology and the data. The penultimate section presents the results of the empirical study. Conclusions and policy implications are drawn in the final section.

2. Literature review

A large volume of research exploring the impact of different cycling infrastructural interventions on cycling demand and cyclist safety has been conducted over the past decade. Extensive reviews can be found in Buehler and Dill (2016), Mölenberg et al. (2019) and DiGioia et al. (2017). In this section, we provide a brief overview of existing studies, followed by highlighting the research gaps in the literature and summarising the contributions of this study.

2.1. Studies on cycling demand

A weight of evidence in the literature suggests that cycle networks have a positive effect on increasing cycling levels (see, for instance, de Dios Ortuzar et al., 2000; Hopkinson and Wardman, 1996; Dill and Voros, 2007; Lv et al., 2022; Mattson et al., 2022; Naseri et al., 2023). Mölenberg et al. (2019) reviewed thirty-one recent studies that evaluated infrastructural interventions to promote cycling in urban areas of high-income countries and found a median increase of 62 percent in the number of cyclists due to the intervention. However, the estimates were found to vary substantially across studies (range:

[2] https://ecf.com/news-and-events/news/evolution-cycle-superhighways-london.

4 to 438 percent) primarily due to the underlying study design and the adopted methodology. Mölenberg et al. (2019) found that the estimates based on subjective measurement methods, such as surveys and direct observations of cyclists found larger changes than those based on objective measurement methods, such as GPS and accelerometers and automatic counting stations. Moreover, Mölenberg et al. (2019) noted that a majority of existing studies lack control for temporal trends and measured or unmeasured confounders (for instance, pre-treatment cycling demand at the location of the intervention), thereby delivering estimates that may only be *correlational* in nature.

Another review of various stated and revealed preference studies by Buehler and Dill (2016) highlighted the role of the features of the cycleway infrastructure on associated cycling levels. In particular, the review suggested that cyclists prefer physically separated cycle lanes over lanes or wide shoulders on roadways, particularly with high volumes of fast-moving motorised traffic. Additionally, intersections were found to have a negative effect on cycling, while cycle-specific traffic control devices at intersections, such as cycle traffic signals and cycle signal activation were found to offset this negative effect. However, Buehler and Dill (2016) highlighted two key limitations of the existing studies. First, most of the existing findings were found to be based on small samples of volunteers, members of university communities, or avid cyclists, who may be non-representative of the population. Second, a majority of the studies reviewed used cross-sectional data, thus limiting the evidence to correlations between cycleway features and cycling levels at one point in time. Moreover, such datasets may be uninformative of the long-term effects of the infrastructure. Buehler and Dill (2016) further pointed out that a handful of studies that used longitudinal data lacked control or comparison sites. Overall, the review suggested the use of systematically collected data, such as those from count points that can be tracked over time, and the adoption of appropriate research methods to establish *causality*.

2.2. Studies on cyclist safety

Several existing studies, such as Abdel-Aty et al. (2014), Agerholm et al. (2008), Buckley and Wilke (2000), Chen et al. (2012), DiGioia et al. (2017), Fowler and Koorey (2006), Jensen et al. (2008), Ling et al. (2020), Lusk et al. (2011), Nolan et al. (2021), Park et al. (2015), Pulugurtha and Thakur (2015), Reynolds et al. (2009), Teschke et al. (2012), Thomas and DeRobertis (2013) and Zangenehpour et al. (2016) have explored the impact of cycle lanes on cycle collisions. Three previously conducted review studies highlight different aspects of the safety of cycle lanes. Reynolds et al. (2009) reviewed twenty-three papers and discussed two categories of infrastructure: one related to intersections and the other related to straight-ways. The results suggest that a separate cycle track can reduce the risk to cyclists. Thomas and DeRobertis (2013) examined studies of cycle tracks from different countries and found that one-way cycle tracks are generally safer at intersections than two-way ones, and the construction of cycle tracks can reduce collisions if effective intersection treatments are employed. DiGioia et al. (2017) examined twenty-two bicycle treatments and concluded that most interventions are still in need of rigorous research. In particular, DiGioia et al. (2017) highlighted that there are fundamental questions with respect to appropriate exposure measures, crash measures, and crash data sources, that still need to be addressed.

Relatedly, DiGioia et al. (2017) found that most existing studies can be classed into two groups based on the type of method adopted in these studies: (i) before-after studies (see, for instance, Abdel-Aty et al., 2014; Agerholm et al., 2008; Buckley and Wilke, 2000; Chen et al., 2012; Fowler and Koorey, 2006; Jensen et al., 2008; Ling et al., 2020; Park et al., 2015; Parsons and Koorey, 2013) and (ii) cross-sectional studies (see, for instance, Lusk et al., 2011; Nolan et al., 2021; Park et al., 2015; Pulugurtha and Thakur, 2015; Teschke et al., 2012; Zangenehpour et al., 2016). Before–after studies further comprise of simple before-after studies and studies based on full Bayes and empirical Bayes. For instance, Jensen et al. (2008) implemented a simple before-after method to study the impact of bicycle tracks on crashes in Copenhagen, Denmark. The results suggest a slight increase in both crashes and injuries. Case-control and cohort methods are two common types of cross-sectional studies (Gross et al., 2010; Zoghi, 2013). For instance, Lusk et al. (2011) compared six cycle tracks and comparable reference streets in Montreal, and concluded that the injury risk of cycling on cycle tracks is less than that on streets.

Table 1 summarises the key findings from the literature.

2.3. Research gaps

We note several research gaps in the current literature. Firstly, most existing studies do not control for unobserved confounding factors. For example, cycle lanes are often introduced such that areas with parking spaces and bus stops are avoided as they increase the risk of collision (Pai, 2011; Pei et al., 2011). Failure to control for such confounding factors may result in estimates that are biased and unrepresentative of the intervention. Secondly, both before-after and cross-sectional methods need a comparison group to compare with the treated group. Ideally, the comparison group is supposed to behave similarly to the treated group prior to the intervention. However, previous studies fail to give a plausible explanation for how to quantify this similarity and match the comparison group with the treated group. Essentially, the counterfactual, which is instrumental in determining effect size and significance, has not been rigorously justified. Thirdly, most existing studies fail to control for exposure. For example, Agerholm et al. (2008) concluded that cycle tracks in western Denmark increased the risk of injury collisions without controlling for cyclist volume. Some studies assume a linear relationship between cyclist volume and collision risk, while others omit exposure altogether, potentially biasing the estimated effect of infrastructure. Moreover, rarely are infrastructure characteristics modelled explicitly as interacting covariates, and effect heterogeneity is often discussed descriptively rather than quantified through formal subgroup analysis or parameter interaction.

2.4. Contributions

Our contribution to this line of research is three-fold.

1. To address the methodological shortcomings of the current literature, we adopt two statistically robust causal inference methods to evaluate the impact of London CS on cycling demand and cyclist safety. The methods account for potential sources of confounding biases and extraneous temporal trends.
2. We introduce a novel cyclist safety indicator — the normalised collision rate — which normalises observed collision counts by expected collision counts based on cyclist volume. This approach accounts for changes in exposure while incorporating potential non-linearities in the relationship between collisions and exposure, ensuring a more accurate assessment of safety impacts.
3. We provide novel empirical insights into how the impact of London CS varies by infrastructure type. We achieve this by estimating the impact of each CS (with different underlying features) separately and also by assessing the effect of the segregation of CS on the outcomes of interest, cycling demand, and cyclist safety.

3. Methodology

This section has three subsections. The first subsection introduces the causal inference framework. The next two subsections provide a description of the methods used in this paper: propensity score weighted difference-in-differences and panel outcome regression with fixed effects.

Table 1
Review of the key studies quantifying the safety impact of cycle lanes.

Study	Area of study	Time horizon	Response studied	Method	Key findings
Chen et al. (2012)	Roadway segments and intersections in New York City	1996 to 2006	Number of cycle collisions	Before-after analysis	Despite the likelihood of an increase in the number of cyclists, the installation of cycle lanes did not result in a higher incidence of crashes.
Fowler and Koorey (2006)	Pages Road, in the east-Christchurch suburb of Wainoni, New Zealand	1986 to 2004	Number of cycle collisions	Before-after analysis	The impact of cycle lanes on cycle safety was found to be limited.
Ling et al. (2020)	The city of Toronto	2000 to 2016	Number of cycle-motor vehicle collisions	Before-after analysis	After controlling for cycling volume, the implementation of cycle tracks was linked with improved safety for cyclists who were using the cycle tracks.
Lusk et al. (2011)	Six cycle tracks in Montreal	2002 to 2006 but only considering the cycling season, that is, 1 April to 15 November	Cyclist injury rates	Cross-sectional methods	The probability of getting injured while cycling on cycle tracks was found to be lower compared to cycling on streets.
Pulugurtha and Thakur (2015)	The city of Charlotte, North Carolina	2008 to 2010	Number of bicycle crashes per annual million vehicle miles travelled	Cross-sectional methods	After analysing all types of crashes, it was concluded that on-street bicycle lanes do not have a significant negative impact on overall safety.
Teschke et al. (2012)	Toronto and Vancouver, Canada	2008 to 2009	Cyclist injury rates	Case-crossover method	Amongst the 14 different types of cycling routes, it was found that cycle tracks had the lowest level of risk.

3.1. Causal inference framework

Let $Z_{it} = (Y_{it}, D_{it}, X_{it})$ represent the observed data comprising N road segments or units indexed with $i = 1, 2, \dots, N$ over years $t = 1, 2, \dots, T$. Here, Y_{it} is an outcome of interest for unit i in year t. D_{it} denotes a binary treatment indicator. If $D_{it} = 1$, the unit i receives the treatment, that is, the implementation of CS; otherwise, $D_{it} = 0$. X_{it} is a vector of covariates describing the characteristics of unit i in year t. Let $Y_{it}(1)$ and $Y_{it}(0)$ represent the potential outcomes for unit i in year t, in other words, the outcomes that would have occurred for unit i in year t under treatment and control status, respectively. Note that for each unit i, we can only observe one of the potential outcomes, $Y_{it}(1)$ or $Y_{it}(0)$. As a result, we cannot directly estimate the unit-specific treatment effect $\tau_i = Y_{it}(1) - Y_{it}(0)$. The estimand of interest is, therefore, the average treatment effect (ATE), that is, the treatment effect averaged over the population (that is, all road segments), given by,

$$\tau_{ATE} = E[Y_{it}(1) - Y_{it}(0)] \tag{1}$$

There are three key identifying assumptions required for the estimation of the ATE (for details, refer to Imbens and Wooldridge, 2009; Graham, 2025).

- *Conditional Independence Assumption (CIA)*, which requires the potential outcomes to be independent of the treatment, given the observed covariates X, that is, $Y(0), Y(1) \perp\!\!\!\perp D|X$. In other words, it assumes that if all observed differences in characteristics between the treated and untreated units are controlled for, the outcomes that would result in the absence of treatment are the same for both groups.
- *Common Support* requires that each unit i has a positive probability of both receiving the treatment or not, that is, $0 < P(D = 1|X) < 1$. The assumption ensures that there is sufficient overlap in the characteristics of the treated and control units to generate counterfactual outcomes for the treated units.
- *Stable Unit Treatment Value Assumption (SUTVA)* requires the potential outcomes of unit i to not vary with the treatments assigned to any other unit $j \neq i$. Moreover, the observed outcomes under a given treatment allocation must be equivalent to potential outcomes under that allocation, that is, $Y_{it} = I_1(D_{it})Y_{it}(1) + (1 - I_1(D_{it}))Y_{it}(0), \forall i = 1, 2, \dots, N$, where $I_1(D_{it})$ in the indicator function for receiving the treatment.

3.2. Propensity score matched difference-in-differences

We aim to estimate the ATE of the introduction of CS on cycling demand and safety. A straightforward approach might compare outcomes between treated and control groups. However, as outlined in Section 2, the assignment of CS to individual road segments is unlikely to be random. Confounding variables X, representing pre-treatment characteristics of road segments, may influence both the likelihood of receiving the treatment and the potential outcomes, introducing selection bias. To address this, we adopt a propensity score-matched difference-in-differences (PSM-DID) approach, which adjusts for X and enables consistent estimation of the ATE. The approach was first proposed in Heckman et al. (1998) and then extensively discussed and applied in studies such as (Smith and Todd, 2005), Caliendo and Kopeinig (2008) and Gebel and Voßemer (2014).

The PSM-DID approach integrates propensity score matching with a difference-in-differences model to address selection bias and confounding from both observed and unobserved factors. We begin by estimating propensity scores (PS), which represent the probability of treatment assignment based on observed covariates X^{obs}. Using these scores, treated and control units are matched, creating a subsample where the distribution of observed covariates is balanced across the two groups. This balancing reduces selection bias and allows us to approximate the conditions of a randomised experiment while ensuring the validity of the Common Support assumption, as units with no overlap in propensity scores are excluded.

After matching, we employ the difference-in-differences (DID) model as the framework for outcome analysis. Unlike naïve regression-based models, which depend on the Conditional Independence Assumption (CIA) for causal inference, an assumption that may be violated when unobserved confounding factors are present, DID relies on the "parallel trends" assumption. This assumption posits that, in the absence of treatment, the treated and control groups would exhibit similar changes in outcomes over time. By leveraging this assumption, the DiD approach effectively accounts for time-invariant unobserved heterogeneity and any unobserved time-varying factors that influence both groups in the same way.

In sum, the PSM-DID framework provides a robust method for estimating the causal impacts of interest by combining the strengths of PSM for observed confounders and DID for unobserved heterogeneity. Together, these methods address selection bias comprehensively, ensuring reliable estimates of the ATE.

Below, we delve deeper into the specifics of the two key components of this approach as applied to the context of this study.

3.2.1. Propensity score matching

The concept of PSM was introduced by Rosenbaum and Rubin (1983) and further developed by Heckman et al. (1997) and is one of the most popularly used matching methods in the literature. The main advantage of PSM is that it reduces the multiple dimensions of matching to a single dimension, namely, the PS.

For the purpose of this study, we estimate the PS, $e(X)$, using a logistic regression model, defined as:

$$\begin{aligned} \widehat{e_i(X_i, \boldsymbol{\gamma})} =& P(D_i = 1 | X_{i1}, \dots, X_{iK}) \\ =& \frac{1}{1 + \exp(-(\gamma_0 + X_{i1}\gamma_1 + \cdots + X_{iK}\gamma_K))}, \end{aligned} \tag{2}$$

where $\boldsymbol{\gamma}$ represents the estimated parameters, and the covariates X are pre-treatment characteristics discussed in Section 4.4. The PS model ensures that treated and control units with similar PS values are comparable in terms of their covariates.

We use the estimated PS to perform matching using a flexible and robust matching technique named *full matching* (Hansen, 2004). Unlike simpler matching methods, such as nearest-neighbour, full matching aims to create a more comprehensive balance by ensuring that each treated unit is matched with one or more control units in a way that balances the covariates across the entire sample. The method involves creating groups where treated and control units are paired based on their estimated PS and covariates, allowing each treated unit to be matched with multiple control units and vice versa. The goal is to minimise the imbalance in covariates and PS across these groups, utilising all available data and enhancing the comparability of treatment and control units. Full matching is implemented through an optimisation process that seeks to minimise a weighted sum of covariate imbalances and PS distances.

3.2.2. Difference-in-differences estimation

After matching, we fit a DID model to estimate the ATE, specified as:

$$g(Y_{it}) = X_{it}^T \boldsymbol{\beta} + I_{D_i}\alpha + I_{D_i} \times I_{T_t}\tau_{DID} + \boldsymbol{\delta}^T \boldsymbol{\lambda} + \epsilon_{it}, \tag{3}$$

where, Y_{it} is the outcome of interest for unit i in year t (refer to Section 4.3 for details on the outcomes considered in this study); $g(.)$ is the appropriate link function based on the distribution of the outcome; α, $\boldsymbol{\beta}$, λ and τ_{DID} are parameters to be estimated; I_{D_i} is an indicator variable representing assignment of the treatment, $I_{D_i} = 1$ if $D_i = 1$, zero otherwise; I_{T_t} is an indicator variable representing the post-treatment period, $I_{T_t} = 1$ if year t falls in the post-treatment period, zero otherwise; $\boldsymbol{\delta}$ is a $T \times 1$ vector that captures year fixed effects; and ϵ_{it} is an idiosyncratic error term, $\epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2)$. Depending on the nature of the outcome variable (say, count or continuous), appropriate link functions (for instance, log for Negative Binomial, identity for continuous outcomes) are specified as detailed in Table 3 attached in Section 4.3. The inclusion of covariates X_{it} (discussed in Section 4.4) ensures that any remaining imbalances in observed characteristics are controlled. The parameter α measures the difference between the expected pre-treatment responses of the treated and control groups. The effect of the treatment is captured by the parameter τ_{DID}, which provides the sample counterpart to

$$g(\tau_{DID}) = E[Y_{i,1} - Y_{i,0} \mid I_{D_i}(1), e(X)] - E[Y_{i,1} - Y_{i,0} \mid I_{D_i}(0), e(X)] \tag{4}$$

where the first term on the right-hand side represents the difference in the average response of the treated group between the post-treatment and the pre-treatment periods, conditional on and the propensity score. The second term represents the corresponding difference for the control group. For a detailed discussion on the DID approach and example applications in the transportation literature, refer to Ashenfelter and Card (1984), Finkelstein (2002), Li et al. (2012) and Anupriya et al. (2020).

A key assumption for valid identification in DID estimation is the 'parallel trend' assumption, which stipulates that, in the absence of the treatment, the treated and control groups would exhibit the same trend over time. Mathematically, this can be expressed as:

$$\begin{aligned} &g(E[Y_{i,1}(0) \mid I_{D_i}(1), e(X)]) - g(E[Y_{i,0}(0) \mid I_{D_i}(1), e(X)]) \\ &= g(E[Y_{i,1}(0) \mid I_{D_i}(0), e(X)]) - g(E[Y_{i,0}(0) \mid I_{D_i}(0), e(X)]). \end{aligned} \tag{5}$$

3.3. Panel outcome regression with fixed effects

The performance of the PSM-DID approach hinges on the correct specification of the propensity score (PS) model, which affects the sample size available for estimating the average treatment effect (ATE), as well as the satisfaction of the parallel trends assumption, which is critical for addressing unobserved confounding. Errors in these steps or violations of the underlying assumptions can undermine the reliability of PSM-DID estimates, as biases from both observed and unobserved confounding cannot be adequately addressed.

To evaluate the robustness of the ATE estimates derived from PSM-DiD, we compare them with estimates obtained from a panel outcome regression with fixed effects (FE). This comparison provides a benchmark for assessing the reliability of PSM-DiD. Previous studies, such as Burbidge and Goulias (2009), have also utilised the FE model to evaluate the impact of cycleway investments on cycling demand. The FE model offers consistent ATE estimates under the assumption that any unobserved confounders are time-invariant. By controlling for both measured covariates and unobserved time-invariant factors, the FE model enhances the robustness of causal inference, particularly when the CIA assumption may be questioned.

Suppose that the data-generating process is:

$$g(Y_{it}) = D_{it}\tau_{ATE} + X_{it}^T\beta + W_i^T\Delta + \delta_t + \psi_{it} \tag{6}$$

where X_{it}^T is a $K \times 1$ vector of observed time-variant covariates and W_i^T is an $J \times 1$ vector of observed and unobserved time-invariant covariates. δ_t is the year-specific effect and ψ_{it} is the error term. $E[\psi_{it} | X_{it}, W_i] = 0$, $i = 1, 2, \dots, N$, $t = 1, 2, \dots, T$. The FE model assumes that each unit i has a unique attribute α_i that is constant through time. The resulting model is

$$g(Y_{it}) = \alpha_i + D_{it}\tau_{FE} + X_{it}^T\beta + \delta_t + \varepsilon_{it} \tag{7}$$

with τ_{FE} being the estimand of interest. Note that while the FE model can effectively deal with unobserved time-invariant confounding factors, it fails to control for time-varying confounding factors.

4. Data

In this section, we describe the datasets, the variables used, and the detailed data preparation procedures implemented to ensure robustness in our empirical analysis.

4.1. Data sources and preparation

We primarily use four datasets: road safety data, road traffic data, cycleway data, and socioeconomic data.

- **Road safety data**: The road safety data originates from STATS19, published by the Department for Transport.[3] This dataset provides details of road collisions in Great Britain, including date, location, vehicle types, casualty details, and severity. We specifically focus on cycle-related collisions occurring between 2000 and 2019 in Greater London. The collision records were cleaned for duplicates and also classified by collision type (single-cycle incidents, cycle-to-cycle collisions, cycle-to-vehicle collisions) and collision location (junction versus non-junction) for further analysis.

[3] https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data.

Table 2
Cycle superhighways under study.

CS No.	Year	Length	Route	Segregation
CS1	2016 (opened)	10.8 km approx.	Runs from Tottenham to the City, primarily along quieter streets parallel to the A10	Utilises existing streets with limited segregation; focuses on traffic calming and signage
CS2	2011 (opened); 2016 (upgraded)	6.8 km approx.	Connects Aldgate to Stratford along the A11	Before Upgrade: predominantly unsegregated cycle lanes; After Upgrade: significant portions feature segregated cycle tracks
CS3	2010 (opened); 2016–2018 (extended)	23 km	Extends from Barking to Lancaster Gate, passing through central London	Mostly segregated two-way cycle tracks
CS5	2015 (opened)	1.4 km	Connects Oval to Pimlico via the Vauxhall Bridge	Completely segregated two-way cycle tracks
CS6	2016 (opened)	5 km	Connects Elephant and Castle to King's Cross through central London	Completely segregated two-way cycle tracks
CS7	2010 (opened); 2016 (upgraded)	13.7 km	Extends from Colliers Wood to Cannon Street, primarily along the A24 and A3	Before Upgrade: predominantly unsegregated cycle lanes, mostly shared with buses; After Upgrade: Heavy traffic portions feature segregated cycle tracks
CS8	2011 (opened)	8.2 km	Runs from Wandworth to Westminster	Unsegregated cycle tracks

- **Road traffic data**: The road traffic data, sourced from the Department for Transport,[4] includes annual average daily traffic flow of different vehicle types, recorded at fixed count points across Greater London from 2000 to 2019. Annualised daily traffic volumes for all motorised vehicles (AADT) and bicycles (AADB) were calculated and aggregated by count point and year to align temporally with the collision data.
- **Cycleway data**: Cycleway data was obtained from Transport for London's publicly available repository.[5] This dataset provides geographic coordinates and information on cycling infrastructure, including the presence of Cycle Superhighways (CS) and associated route numbers across London. We calculated the geographic distance from each traffic count point to the nearest CS route, subsequently classifying segments into treated and control groups based on proximity (see Section 4.2). Treated segments were further categorised according to CS route numbers, facilitating detailed subgroup (route-level) analyses of cycle volume and collision outcomes. Additionally, the segregation status of each CS segment (segregated versus unsegregated) was manually determined using historic Google Maps street view imagery (refer to Section 5.5 for details).
- **Socioeconomic data**: Socioeconomic indicators were sourced from the Office for National Statistics (ONS)[6] at the Lower Layer Super Output Area (LSOA) level, including population density, employee numbers, and the Index of Multiple Deprivation (IMD). These socioeconomic variables were spatially matched to traffic count points by linking each point to its nearest LSOA centroid. These variables are included as key covariates (X^{obs}) in our models to adjust explicitly for potential confounding influences (see Section 4.4 for details).

Our data integration procedure systematically combined these datasets. First, traffic count points served as spatial reference locations. We matched cycle collision records from STATS19 to the nearest traffic count points within a spatial threshold of 0.4 km. This threshold was chosen to ensure that matched collisions occurred reasonably close to the reference count point, balancing spatial precision against retaining sufficient sample size. Collisions falling outside this threshold were excluded to maintain robust spatial alignment. Similarly, socioeconomic data were spatially matched to count points based on proximity to LSOA centroids. Finally, we created a comprehensive panel dataset at the count point-year level that integrates motor traffic and cycle volumes, collision data, socioeconomic indicators and CS proximity indicators.

4.2. Control and treated groups

Although twelve CS routes were initially planned, as illustrated in Fig. 1, only seven routes were operational by 2019. Our analysis focuses explicitly on these seven CS routes. Relevant details, including implementation years, route lengths, and segregation levels, are summarised in Table 2.

As described in Section 4.1, treatment and control groups were defined based on proximity to the CS routes. Road segments located within 0.5 km of implemented CS routes formed the treatment group (comprising 111 segments), whereas segments located beyond 1.5 km from any CS route formed the potential control group (comprising 584 randomly selected segments). Fig. 2 illustrates the spatial distribution of treated (red) and control segments (blue). Control segments were explicitly chosen to avoid adjacency to corresponding treated CS routes, thus mitigating potential spillover effects and ensuring compliance with the Stable Unit Treatment Value Assumption (SUTVA). Observations from road segments located between 0.5 km and 1.5 km of any CS route were excluded from the dataset as unmatched observations, thereby maintaining consistency and robustness in the spatial matching process.

The pre- and post-intervention periods were defined based on the specific opening years of the corresponding CS routes.

4.3. Outcomes of interest

Our outcome variables of interest include cycle flow volume, the number of cycle collisions and the normalised collision rate. The former two measures correspond to the AADB and total collision count over a given year. When analysing motorised-traffic collisions, it is a standard practice in the literature to assess collision counts relative to traffic volume. Accordingly, previous studies such as Strauss et al. (2013) and Li et al. (2017) have considered a simple cycle collision rate, defined as

$$\text{cycle collision rate} = \frac{\text{number of collisions}}{\text{AADB} \times 365},$$

to control for the changes in cycle volume when assessing safety, thereby allowing for a more appropriate assessment of the relative risk of an collision before and after the intervention. It is worth noting that this normalisation implicitly assumes a linear relationship between cycle volumes and the number of collisions. We argue that if the true relationship is non-linear, that is, if the number of collisions does not increase proportionally with the number of cyclists, then dividing by cyclist counts distorts the interpretation of risk. For instance, if the underlying relationship is sub-linear, that is, collisions increase at a slower rate than the number of cyclists due to a *safety in numbers* effect (more cyclists make roads safer by increasing driver awareness or slower speeds due to congestion), a direct normalisation may underestimate

[4] https://roadtraffic.dft.gov.uk/regions/6.
[5] https://cycling.data.tfl.gov.uk/.
[6] https://www.ons.gov.uk/.

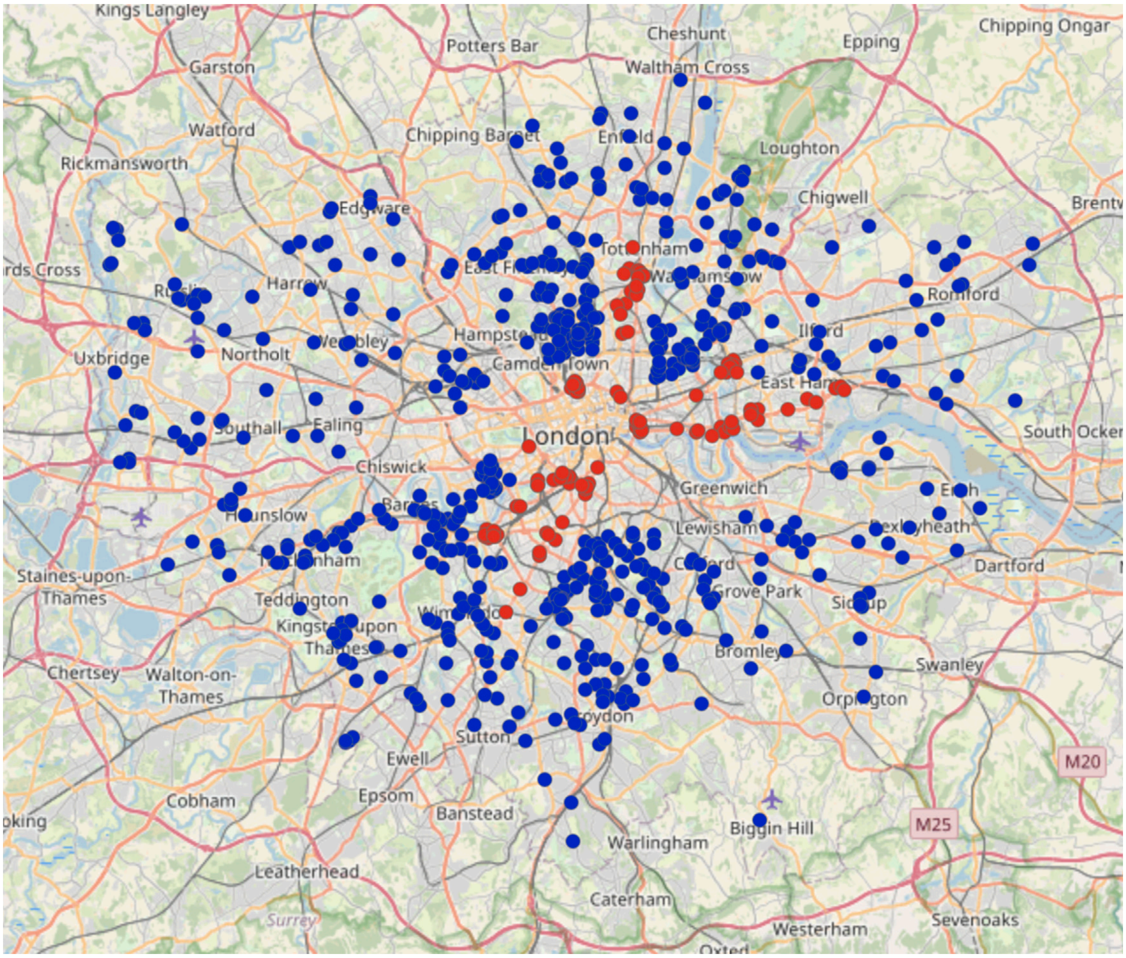

Fig. 2. The distribution of treatment and control segments.

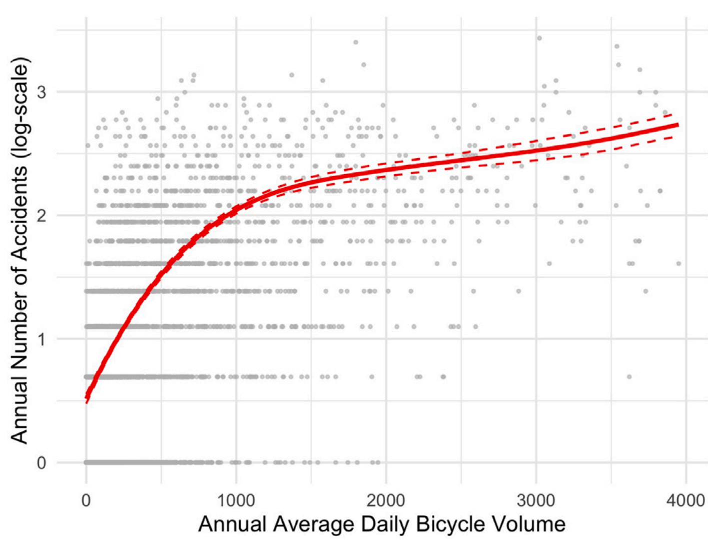

Fig. 3. A penalised spline fit of the pre-treatment collision counts versus AADB data.

risk if there is a substantial increase in cycle volumes post-intervention. Fig. 3 presents a scatterplot of annualised collision counts against AADB for the pre-treatment period, overlaid with a penalised spline fit. The estimated fit provides clear evidence of non-linearities in the relationship between collision counts and AADB.

Instead of the traditional linear collision-rate normalisation, we propose a novel normalised collision rate defined as:

$$\text{normalised collision rate} = \frac{\text{observed number of collisions}}{\text{expected number of collisions}}, \tag{8}$$

Table 3
Summary of outcomes and link functions.

Outcome	Distribution	Link function $g(.)$
Cycling volume (AADB)	$Y_{it} \sim$ Negative Binomial(μ_{it}, θ)	$g(Y_{it}) = \log(\mu_{it})$
Number of collisions	$Y_{it} \sim$ Negative Binomial(μ_{it}, θ)	$g(Y_{it}) = \log(\mu_{it})$
Normalised collision rate	$Y_{it} \sim$ Log-Normal(μ_{it}, σ^2)	$g(Y_{it}) = \log(Y_{it})$

Table 4
Descriptive statistics of the covariates.

Covariate	Description	Mean	Std. Dev.	Min	Max
AADB_pre	Annual average daily bicycle volume in the pre-intervention period	487.17	714.27	0.00	5098.30
AADT	Annual average daily traffic volume in the pre-intervention period	25 015	19 703.92	101	150 203
Collision_pre	Total number of collisions in the pre-intervention period	3.63	3.56	1.00	20.12
MED	A measure of accessibility of LSOA	6915	2553.32	2356	15 630
IMD	The index of multiple deprivation	28.79	13.82	3.76	63.43

Table 5
Summary of balance for the matched data.

	Std. Mean difference	Maximum threshold	Variance ratio
Distance	−0.00	Balanced, <0.1	0.99
AADB_pre	−0.05	Balanced, <0.1	0.95
AADT	0.05	Balanced, <0.1	0.74
Collision_pre	−0.03	Balanced, <0.1	0.68
IMD	0.02	Balanced, <0.1	1.45
MED	−0.03	Balanced, <0.1	1.77

where the expected number of collisions is modelled using a flexible function of cyclist volume (AADB) and other relevant covariates. This approach explicitly accounts for potential non-linearities in the relationship between cyclist volumes and collision counts. The interpretation of the index is straightforward: if the normalised collision rate equals 1, the location experiences collisions at the expected rate; values greater than 1 indicate higher-than-expected collision risk, while values below 1 indicate lower-than-expected risk.

Recognising potential non-linear relationships such as the 'safety in numbers' effect, we empirically model the expected collisions using a Generalised Additive Model (GAM). Specifically, we employ penalised spline regression fitted to the pre-treatment data to flexibly capture the relationship between collision counts and cyclist volumes (Fig. 3). Formally, the expected collisions are modelled as:

$$\text{expected number of collisions} = l^{-1}[f(\text{AADB})] + \xi, \tag{9}$$

where $l(\cdot)$ is the log-link function consistent with Negative Binomial models, ensuring predicted counts remain positive; $f(\text{AADB})$ represents the penalised spline fit capturing non-linearities in collision-cyclist volume relationships; and ξ captures remaining unexplained variability.

Given our aggregated data structure, that is, annual collision events recorded at fixed locations, cyclist volume (AADB) remains the most directly relevant measure of exposure. Alternative metrics, such as trip distance or duration, while potentially useful in other contexts, are less appropriate due to the spatially aggregated nature of the collision data analysed in this study. Additionally, as we focus on annual collision counts, variables such as weather and seasonal factors become less pertinent as covariates in Eq. (9). Furthermore, we have deliberately excluded infrastructure characteristics (for instance, segregated versus shared cycling lanes, speed limits, cycle lane width) from the expected collision model. Such infrastructure attributes typically changed following the CS intervention, posing significant risks of extrapolation beyond observed pre-treatment conditions.

Consequently, we propose the normalised collision rate as a robust, model-based exposure adjustment measure. Specifically, by using GAM-based predictions of expected collision counts as the denominator in our normalised collision rate, we explicitly account for the non-linear relationship between cyclist volume and collision risk. This integration ensures that our causal inference framework directly incorporates these complex non-linear dynamics into the outcome measure itself.

Table 3 summarises the distribution of each outcome of interest, along with the corresponding link function used in the modelling framework.

To account for heterogeneity in infrastructure design across different CS, we implement a multi-level analytical approach. Specifically, we conduct (i) an overall analysis pooling data across all CS routes, (ii) a route-level analysis to explore variation across individual CS routes, and (iii) a subgroup analysis comparing outcomes between segments with segregated and unsegregated CS. While the primary outcome definitions remain consistent, this layered structure allows us to examine how differences in design, particularly the degree of segregation, may influence both cycling volumes and collision risk.

4.4. Covariates

We require an appropriate vector of covariates X^{obs} that determines both the propensity of the treatment as well as the outcomes. Guided by the existing literature and the availability of segment-level data, we consider the following covariates, with values from the pre-treatment period: average traffic flow volume (AADT), average cycle flow volume (AADB_pre), the average number of collisions (collision_pre), the index of multiple deprivation (IMD), and access to economic mass or accessibility measured via mean effective density (MED)[7] A detailed

[7] The MED ρ_j for zone, $j, j = (1, \ldots, n)$, is calculated as follows:

$$\rho_j = \frac{1}{n}\sum_{j=1}^{n} m_j h(d_{ij})$$

where m_j represents a measure of economic activity in each zone j and $h(.)$ denotes the deterrence function, which is a decreasing function of the cost of travelling from origin j to destination k. The measure is designed to capture the effects of the geographic centrality of the zones, their size distribution, and the spatial distribution of economic mass. We consider the zonal employment

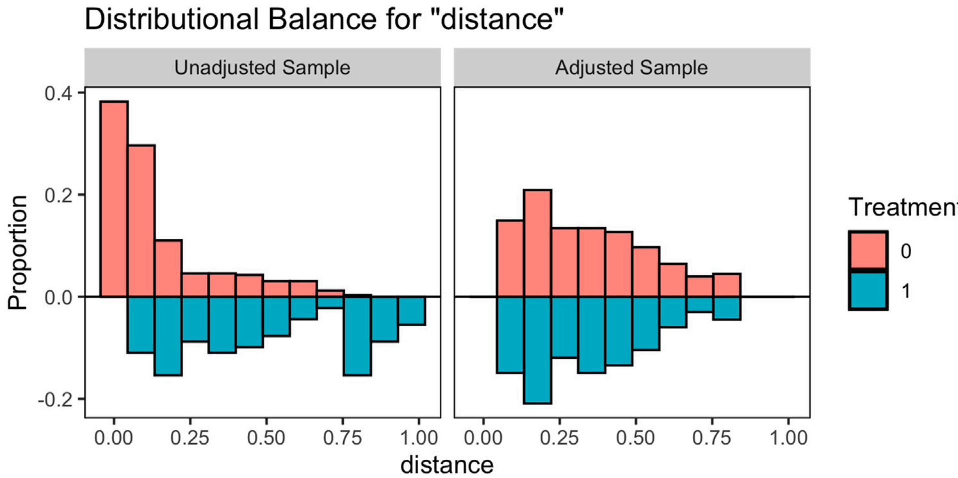

Fig. 4. Overlap test.

description of these covariates alongside their summary statistics is presented in Table 4.

5. Results and discussion

This section is divided into five subsections. In the first subsection, we present the results from a set of tests that assess the success of the propensity scores in removing selection bias. In the second subsection, we present results from testing the parallel trend assumption. In the third subsection, we summarise the estimated causal impact of Cycle Superhighways (CS) on an aggregate level. The penultimate section details the results for each CS, which is followed by an evaluation of the impact of the segregation of CS on the outcome variables of interest.

5.1. Covariate balancing and overlap tests

As described in Section 3.2, estimating the effect of treatment on outcomes in non-randomised studies may be subject to a selection bias in which treated subjects differ systematically from untreated subjects. Nevertheless, once conditioned on the true propensity score, treatment status becomes independent of measured baseline covariates. In other words, we expect the treated and control units with the same propensity score to have similar distributions of observed baseline covariates (Austin, 2009). Nonetheless, certain imbalances may still exist if the statistical model used to calculate the propensity score is misspecified. Thus, it is of vital importance to carry out balance diagnostics after PSM.

To gauge the performance of the propensity scores in removing selection bias, we perform a covariate balancing test. The `MatchIt` package (Ho et al., 2018) in R is applied to perform the propensity score matching between treated and control units. The package consists of several matching methods. As discussed in Section 3.2, we apply *full matching* because it matches every treated unit to at least one control, and every control to at least one treated unit performs quite well, thus assuring a better overlap (Hansen, 2004; Stuart and Green, 2008). We first report the results of covariate balancing (refer to Table 5).

Table 5 shows that all the variables are well-balanced.

Next, we check the overlap by comparing the distribution plot of the estimated propensity scores. The plot is presented in Fig. 4. The figure illustrates that before matching, the distributions of propensity scores between treatment and control groups are quite different. However, upon matching, these distributions become very similar to each other. The figure, therefore, provides sufficient evidence to support that the *common support* assumption holds.

5.2. Parallel trend tests

To evaluate the validity of the parallel trend assumption, we adopt the approach outlined by Hastings (2004), conducting a pre-test that examines data from pre-intervention years. This test assesses whether the temporal patterns in outcomes of interest from the control group align closely with those of the treatment group. Figs. 5, 6, and 7 illustrate the temporal trends for log-transformed AADB, the log-transformed number of collisions, and the log-transformed normalised collision rate, respectively, across the two groups. The trends are broadly similar, indicating that the parallel trend assumption is reasonably satisfied.

5.3. Aggregate-level impacts of CS

Tables 6–8 summarise the aggregate-level impact of Cycle Superhighways (CS) on cycle volume, number of collisions, and the normalised collision rate, respectively, estimated via propensity score matched difference-in-differences model (PSM-DiD) and the panel outcome regression (OR) model with fixed effects (FE). Overall, the results suggest that, on the aggregate, the introduction of CS had a statistically significant and positive (in magnitude) impact (at 95 percent level) on all three outcomes.

Our estimate from PSM-DiD suggests that the estimated impact of the treatment on cycle volume is 0.20 log-points with an associated standard error of 0.09 log-points (refer to Table 6). The estimate indicates a 22.68% (standard error: 10.61%) increase in cycling volume on the treated segments relative to the average cycling volume during the pre-treatment period. The corresponding estimate from the panel OR regression with FE is 0.13 log-points with an associated standard error of 0.06 log-points. This result implies that relative to the control units, the treated road segments experienced an average increase of 14.25% (standard error: 6.79%) in cycle volume post-intervention, given that other covariates remained fixed. Thus, the introduction of CS had a statistically significant and positive effect on cycling demand.

level E_{jt} as the measure of the economic activity of zone j and year t and the inverse Euclidean distance between the centroids of each zone d_{jk}^{α} for the construction of the deterrence function, where α is the distance decay parameter, generally assumed to take a value of 1.0.

Table 6
The estimated impact of CS on cycle volume.

Estimated coefficient	PSM-DiD	Panel OR with FE
CS indicator	0.08 (0.08)	
Post-treatment indicator	0.01 (0.13)	
CS indicator × Post-treatment indicator	0.20 (0.09)**	0.13 (0.06)**
AADT	−0.00 (0.00)**	0.00 (0.00)
AADB_pre	0.00 (0.00)***	
Collision_pre	0.03 (0.02)	
IMD	−0.00 (0.00)	
MED	0.00 (0.00)***	
Year-effects included	Yes	Yes

Figures in brackets represent clustered robust standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent.

Our results are consistent with the Transport for London (TfL) 2011 report (Transport for London, 2011a) and a previous study by Li et al. (2017), both of which suggest an increase in cycle demand following the intervention. The estimated increase could either be a consequence of more cycling or route switching by existing cyclists or any induced cycling demand as a result of the cycling intervention. For instance, the Transport for London (2011a) study found that the introduction of CS3 and CS7 increased the frequency of cycling amongst those cycling five or more times per week by over two percentage points.

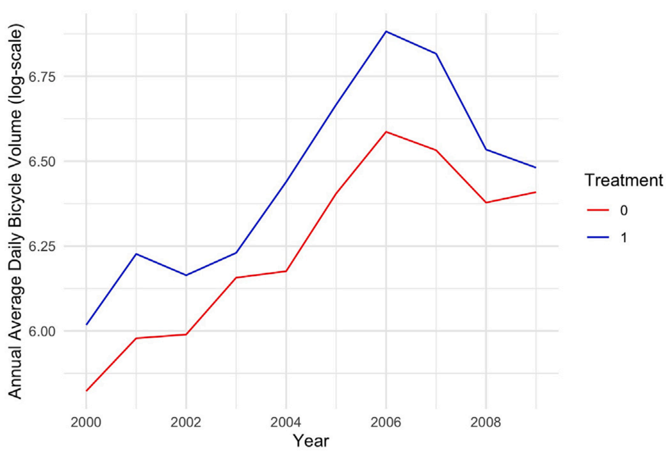

Fig. 5. Temporal trend for log-transformed AADB in the pre-intervention years.

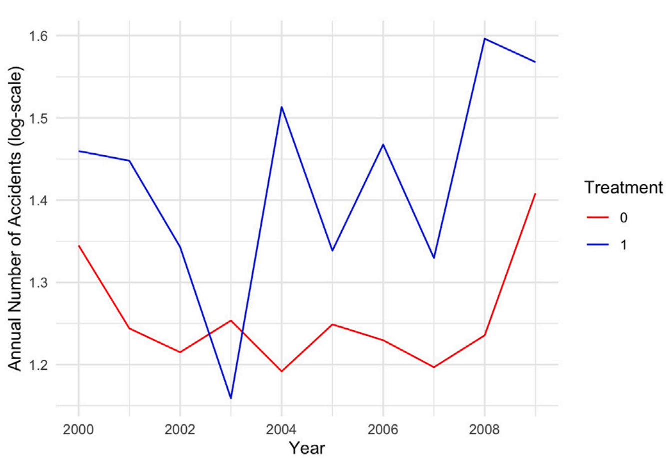

Fig. 6. Temporal trend for log-transformed number of collisions in the pre-intervention years.

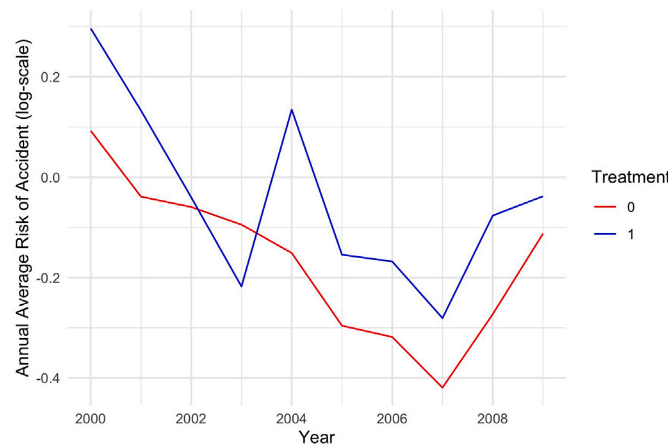

Fig. 7. Temporal trend for log-transformed normalised collision rate in the pre-intervention years.

Moreover, another case study[8] by the TfL suggests that at certain locations on CS6, the cycling volume increased by 124% between 2014 and 2017, whereas the corresponding number for CS3 is 200% over the same period. Further, two waves of interviews conducted among people who have the potential to cycle suggested that the proportions of respondents who started cycling as a result of the cycling intervention were 20% for CS3 and 32% for CS7.

Table 7 illustrates that as per PSM-DiD, the intervention increased the number of cycle collisions on treated road segments by 0.30 (standard error: 0.06) log-points relative to the non-treated road segments, conditional on the other covariates in the model. The estimate indicates a 34.88% (standard error: 7.61%) increase in the number of collisions on the treated segments relative to the average number of collisions during the pre-treatment period. The corresponding estimate from the panel OR regression model with FE is 0.19 (standard error: 0.03) log points, or equivalently, 18.68% (standard error: 3.38%). This result is again consistent with Li et al. (2017), which also estimated a positive impact of CS on the number of cycle collisions. This increase may be a result of the increase in cycling traffic brought about by CS, resulting in more cycling collisions than before. Moreover, our study period corresponds to the time when other complementary schemes were introduced in London to promote cycling (for instance, Barclays Cycle Hire and Biking Boroughs), bringing in more inexperienced cyclists on the road. As such inexperienced cyclists were more prone to use CS routes due to their increased perception of safety (see Liu et al., 2025,

8 https://www.gov.uk/government/case-studies/developing-londons-cycle-infrastructure.

Table 7
The estimated impact of CS on the number of cycle collisions.

Estimated coefficient	PSM-DiD	Panel OR with FE
CS indicator	−0.00 (0.04)	
Post-treatment indicator	−0.02 (0.08)	
CS indicator × Post-treatment indicator	0.30 (0.06)***	0.19 (0.03)***
AADT	−0.00 (0.00)**	−0.00 (0.00)
AADB_pre	0.00 (0.00)***	
Collision_pre	0.17 (0.01)***	
IMD	0.00 (0.00)	
MED	0.00 (0.00)***	
Year-effects included	Yes	Yes

Figures in brackets represent clustered robust standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent.

Table 8
The estimated impact of CS on the normalised collision rate.

Estimated coefficient	PSM-DiD	Panel OR with FE
CS indicator	−0.02 (0.01)	
Post-treatment indicator	−0.09 (0.01)***	
CS indicator × Post-treatment indicator	0.20 (0.06)***	0.26 (0.08)***
AADT	0.00 (0.00)	−0.00 (0.00)***
AADB_pre	−0.00 (0.00)***	
Collision_pre	0.18 (0.00)***	
IMD	0.00 (0.00)***	
MED	−0.00 (0.00)	
Year-effects included	Yes	Yes

Figures in the table represent bootstrapped estimates and standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent.

for a discussion on the role of perception in traffic safety) with respect to CS routes,[9] this may have led to more collisions on CS routes.

Further, our estimates from both the PSM-DiD approach and the panel OR regression model with FE suggest that the intervention increased the normalised collision rate in the treated segments relative to the control road segments (see Table 8).

The results indicate that the observed number of collisions post-intervention exceeds the expected level, suggesting an increase in collision risk following the intervention. Specifically, the PSM-DiD approach estimates an increase of 0.20 log-points (standard error: 0.06), while the panel OR regression model with fixed effects yields an increase of 0.26 log-points (standard error: 0.08). These correspond to percentage increases of 22.11% (0.76%) and 30.33% (10.04%), respectively. It is worth noting that new cycling infrastructure, particularly those incorporating enhanced safety features, is likely to attract inexperienced cyclists, even for regular commuting purposes. This shift in user demographics may have contributed to the increased collision risk post-intervention. However, it remains crucial to explore whether different routes and infrastructure types have differentiated impacts. We investigate this heterogeneity in the next subsections.

5.4. The impacts of CS along different CS routes

In the previous subsection, we quantified the impacts of CS aggregated over all CS routes. It is worth noting that each CS route is characteristically distinct, for instance, by design (route length, level of segregation and separation from bus lanes, see Table 2) and by regions and populations of cyclists (experienced versus inexperienced, commuters versus leisure travellers or tourists) served, which may lead to heterogeneity in impacts across different CS routes. Similar heterogeneity in road safety outcomes has been explored in studies such as Azimian et al. (2021) and Han et al. (2018). Therefore, in this sub-section, we evaluate the route-level impacts of the intervention. Tables 9–11 summarise the results. Note that these results have been obtained by estimating Eqs. (2), (3) and (7) separately for each route; however, for a concise presentation of the corresponding results, we have excluded the other covariates from these tables. Further, due to the limited number of treated segments along CS5 and CS6, we estimate their combined effect. This adjustment does not impact our conclusions, as the two routes share similar characteristics, with CS5 covering a subset of the regions served by CS6.

From Table 9, we note that the impact of the intervention on cycle volumes remains statistically significant at the 95% confidence level only for four CS routes: CS1, CS2, CS5 & CS6 and CS7. The estimated changes in cycling volume for the four routes are +52.25%, +68.60%, +55.38% and +27.25%, respectively. These increases may be a result of several factors like strategic position, early implementation, segregation, perceived safety of the route, or significant upgrades.

CS1, although unsegregated, runs via quieter streets, which may have been perceived by cyclists as a safer alternative to busy main roads, thus encouraging more cycling along the route. CS2 connects East London, including Stratford and Bow, to the City of London, serving rapidly growing areas post-2012 Olympics. Major upgrades in 2016 introduced fully segregated lanes and cycle-priority junctions, enhancing safety and usability. CS5 features a fully segregated two-way cycle track through Vauxhall, catering to commuters in this high-traffic area. CS6, also fully segregated, provides a continuous north-south route from Elephant and Castle to King's Cross, linking key transport hubs and central London destinations. Their strategic location and high-quality infrastructure may have made them particularly attractive to commuters and local residents. CS7 links high-density residential neighbourhoods in South London, like Tooting, Balham, and Clapham, to central London, following the busy A24 commuter corridor, which may have provided a robust alternative to public transport in these areas. Upgrades in 2017 improved segregation and widened lanes, which may have further boosted its appeal. Amongst the four routes, the estimated increase in cycle volume post-intervention is highest for CS2 (+68.60%), which is followed by CS5&CS6 (+55.38%).

Our estimates for CS3 and CS8 lack statistical significance. For CS3, it is important to note that the treated segments considered in this

[9] TfL surveys found that 84% of CS3 users felt safe during their journey (source: refer to the previous footnote).

Table 9
The estimated route-level impact of CS on cycle volume.

Route	PSM-DiD		Panel OR model with FE
	Estimate	Percent change	Estimate
CS1	0.42 (0.11)***	52.25% (16.60%)	0.24 (0.13)*
CS2	0.52 (0.15)***	68.60% (25.77%)	0.60 (0.14)***
CS3	0.16 (0.19)	n.s.	0.13 (0.20)
CS5&CS6	0.44 (0.16)***	55.38% (24.54%)	0.29 (0.12)**
CS7	0.24 (0.11)**	27.25% (13.73%)	0.29 (0.10)***
CS8	0.09 (0.10)	n.s.	−0.16 (0.08)**

Figures in brackets represent clustered robust standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

Table 10
The estimated route-level impact of CS on the number of cycle collisions.

Route	PSM-DiD		Panel OR model with FE
	Estimate	Percent change	Estimate
CS1	0.25 (0.08)***	28.91% (10.04%)	0.21 (0.04)***
CS2	0.43 (0.09)***	54.28% (14.08%)	0.39 (0.08)***
CS3	0.42 (0.07)***	52.76% (10.58%)	0.38 (0.05)***
CS5&CS6	0.14 (0.05)***	15.15% (6.11%)	0.11 (0.04)***
CS7	0.42 (0.12)***	52.76% (17.58%)	0.21 (0.11)*
CS8	−0.12 (0.05)**	−11.68% (4.85%)	−0.05 (0.07)

Figures in brackets represent clustered robust standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

study are primarily located along its initial stretch, between Barking and Tower Hill (see Fig. 2). This corridor may not have experienced a notable increase in cycling volume, as it is well-served by frequent and efficient public transport options, including the District Line, C2C rail services, and buses, which may have reduced the relative appeal of cycling as a commuting option. Meanwhile, although CS8 connects residential and commuter-heavy areas like Wandsworth and Battersea to central London, it primarily consisted of painted cycle lanes with minimal protection during the study period. This lack of segregation likely made it less attractive, especially for less experienced cyclists. Additionally, the route did not pass through rapidly growing or regenerating areas until 2019, further limiting its potential for significant growth in cyclist volume.

Next, Table 10 shows that all routes exhibit a statistically significant impact of the intervention on the number of cycling collisions at the 95% confidence level. While the estimated impacts for all routes except CS8 are positive, indicating an increase in cycling collisions post-intervention, CS8 experienced a reduction of 11.68% in the number of collisions. Notably, there was no observed change in cycling volume on CS8 following the intervention. This reduction in collisions may reflect cyclists on the route becoming more experienced over time. This interpretation aligns with the nature of CS8, which primarily serves commuter-heavy areas connecting southwest London to the city centre. Further, amongst all routes, CS2, CS3 and CS7 experienced the highest increase in the number of collisions post-intervention, the magnitude of this increase being +54.28%, +52.76%, and +52.76%, respectively. Meanwhile, the increase in CS5&CS6 is the lowest (+15.15%).

Finally, Table 11 indicates that CS8 exhibited a reduction in normalised collision rate post-intervention at the 95% confidence level. For CS8, the reduction in collision risk could result from its predominant use by experienced cyclists as noted above. Moreover, according to our estimates, while normalised collision rate post-intervention increased for CS1, CS2, CS3 and CS7, the estimates for CS5&CS6 remained statistically insignificant, indicating that these routes follow the expected relationship between collision counts and cycle volume. It is worth noting that unlike other routes, CS5&CS6 have been fully segregated since their introduction in 2015–16.

Taken together, these patterns suggest that infrastructure quality, particularly the presence of physical segregation, is likely a key driver of both the scale and safety of post-intervention cycling activity. While other contextual factors may also contribute, segregation appears to offer consistent protective effects, both by encouraging uptake and by moderating increases in collision risk. We explore this relationship further through a focused subgroup analysis in the next section.

5.5. The effect of segregation of CS

In this subsection, we examine the impact of segregation of Cycle Superhighway (CS) on cycling volume and cyclist safety. Segregated cycle lanes dedicate a portion of the road exclusively for cyclists, offering several notable benefits. For instance, the shift to segregated lanes has been shown to increase the carrying capacity of congested streets (Aldred et al., 2017). Additionally, studies from Denmark indicate that segregated cycle lanes can reduce cyclist fatalities by 35%.[10] Segregated lanes are also significantly more effective than non-segregated ones in encouraging cycling, particularly among women.

The data on whether a CS segment is segregated is collected using Google Maps. The platform provides not only recent street views of the locations but also historical images. By utilising the coordinates of the 111 CS segments included in the study and manually reviewing historical Google Street View images, we encode the segregation status of each segment over time. An assessment of the collected data suggests that CS2 received substantial segregation upgrades along its entire length a few years after its introduction in 2010. Using the available data, we examine the cycling volume and safety effects of segregated CS segments relative to non-segregated ones through two approaches. First, we analyse subsets of the data comprising segregated and non-segregated CS segments separately, along with their respective control segments, and compare the estimated impacts. Second, for CS2, we treat the upgrade as the new intervention and estimate its demand and safety effects.

5.5.1. The impacts for segregated and non-segregated CS segments

We estimate Eqs. (2), (3) and (7) separately for the above-described subsamples of the data. The results are summarised in Table 12, where once again for brevity, we have excluded the other covariates from our presentation.

From Table 12, we observe that the intervention resulted in a statistically significant increase (at the 95% confidence level) in cycling

[10] https://www.trafficchoices.co.uk/traffic-schemes/segregated-cycle-lanes.shtml.

Table 11
The estimated route-level impact of CS on the normalised collision rate.

Route	PSM-DiD		Panel OR model with FE
	Estimate	Percent change	Estimate
CS1	0.10 (0.01)***	10.18% (0.72%)	0.15 (0.03)***
CS2	0.23 (0.02)***	26.47% (1.94%)	0.22 (0.02)***
CS3	0.25 (0.01)***	28.43% (0.84%)	0.36 (0.00)***
CS5&CS6	−0.04 (0.11)	n.s.	0.09 (0.14)
CS7	0.39 (0.04)***	47.47% (5.56%)	0.89 (0.39)***
CS8	−0.14 (0.01)***	−12.83% (0.70%)	−0.04 (0.01)***

Figures in the table represent bootstrapped estimates and standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

Table 12
The estimated impacts for segregated versus non-segregated CS segments.

Segment type	PSM-DiD		Panel OR model with FE
	Estimate	Percent change	Estimate
Outcome: cycle volume			
Segregated CS	0.39 (0.17)**	47.67% (24.52%)	0.27 (0.09)***
Non-segregated CS	0.02 (0.08)	n.s.	0.08 (0.06)
Outcome: number of cycle collisions			
Segregated CS	0.30 (0.08)***	35.66% (10.98%)	0.19 (0.04)***
Non-segregated CS	0.26 (0.06)***	29.84% (8.44%)	0.20 (0.05)***
Outcome: normalised collision rate			
Segregated CS	0.06 (0.6)	n.s.	0.10 (0.18)
Non-segregated CS	0.24 (0.03)***	26.95% (3.92%)	0.38 (0.06)***

Figures in brackets represent clustered robust standard errors for the first two outcomes
Figures in the table represent the bootstrapped standard error for the third outcome
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

volume along segregated CS segments, with an estimated magnitude of 47.67% (standard error: 24.52%). In contrast, the impact on non-segregated CS segments remains statistically insignificant. These findings align with our route-level results (see Section 5.4), which show that segregated CS routes, such as CS2, CS5, CS6, and CS7, experienced statistically significant increases in cycling volume post-intervention. This suggests that segregated CS infrastructure is likely more appealing to cyclists, as it provides a greater sense of safety.

Table 12 shows that both segregated and non-segregated CS segments experienced a statistically significant increase in cycle collisions post-intervention at the 95% confidence level. The estimated impact was 35.66% (standard error: 10.98%) for segregated segments and 29.84% (standard error: 8.44%) for non-segregated segments. However, Table 12 reveals a notable divergence in the normalised collision rates. While the normalised collision rate increased by 26.95% (standard error: 3.92%) for non-segregated CS segments, the change remains statistically insignificant for segregated segments. This suggests that while segregated CS segments maintain a predictable relationship between collision counts and cycle volume, non-segregated CS segments become significantly riskier than expected.

Next, we examine the time-varying effects of the intervention by modifying the DiD model in (3) as follows:

$$Y_{it} = X_{it}^T \boldsymbol{\beta} + I_{D_i}\alpha + \sum_{tw} I_{D_i} \times I_{T_{tw}} \tau_{DID}^{tw} + \sum_{tw} \delta_{tw}\lambda^{tw} + \epsilon_{it}, \tag{10}$$

where, Y_{it} is the outcome of interest for unit i in year t; α, $\boldsymbol{\beta}$, λ^{tw} and τ_{DID}^{tw} are parameters to be estimated; I_{D_i} is an indicator variable representing assignment of the treatment, $I_{D_i} = 1$ if $D_i = 1$, zero otherwise; $I_{T_{tw}}$ is an indicator variable for each post-treatment time window; δ_{tw} captures time-window specific fixed effects; and ϵ_{it} is an idiosyncratic error term, $\epsilon_{it} \sim \mathcal{N}(0, \sigma_\epsilon^2)$. We divide the post-treatment period into three time-windows, denoted by tw:

1. Initial implementation (2010–2014), $tw = 1$: This period marks the introduction of Cycle Superhighways (CS), primarily consisting of non-segregated routes
2. Segregation expansion (2015–2017), $tw = 2$: During this period, major segregation upgrades were implemented along existing routes (for instance, CS2), and fully segregated routes like CS5 were opened
3. Later years (2018–2019), $tw = 3$: This period captures longer-term effects of the intervention

The ATE corresponding to the above time windows are captured by $\tau_{DID}^{tw=1}$, $\tau_{DID}^{tw=2}$, and $\tau_{DID}^{tw=3}$, respectively. Table 13 summarises the results.

From Table 13, we observe that the intervention did not lead to a statistically significant change in cycle volume along non-segregated CS in any of the three time windows. In contrast, for segregated CS, while no significant change is detected in time window 1, substantial increases in cycle volume emerge in time windows 2 and 3. This pattern aligns with expectations, as time window 1 (2010–2014) corresponds to a period of minimal segregation, where the segregation implementation was sporadic. In contrast, time windows 2 and 3 (2015–2019) saw the introduction of major segregation upgrades and fully segregated routes, driving the observed increases in cycling volumes.

Next, Table 13 reveals that for segregated CS, the intervention did not affect the number of cycle collisions in time window 1. However, in time windows 2 and 3, collision counts increase significantly — a trend that appears to be driven by higher cycling volumes. Importantly, despite this increase in collision counts, normalised collision rates remained unchanged, indicating that the infrastructure effectively absorbed the additional cycling volume without increasing per-cyclist risk. In contrast, for non-segregated CS, the intervention led to a significant increase in both collision counts and normalised collision rates in time windows 1 and 2, while the effects in time window 3 are statistically insignificant. These time-varying collision trends and risk profiles seem to indicate a shift in cyclist composition and consequent risk exposure along non-segregated routes. Specifically, it may well be the case that the CS intervention led to a surge in new, inexperienced cyclists on non-segregated routes, while more experienced cyclists likely switched to alternative routes.

To substantiate this hypothesis, we leverage the granularity of the STATS-19 data and examine a proxy for cyclist experience: *single-cycle incidents at junctions*. The key findings, summarised in Table 14, reveal that along non-segregated CS, there is a statistically significant

Table 13
The estimated impacts over time for segregated versus non-segregated CS segments.

Time segment	Segregated CS		Non-segregated CS	
	Estimate	Percent change	Estimate	Percent change
Outcome: cycle volume				
Time window 1	−0.10 (0.18)	n.s.	−0.02 (0.08)	n.s.
Time window 2	0.51 (0.17)***	66.15% (28.50%)	0.06 (0.10)	n.s.
Time window 3	0.60 (0.17)***	82.07% (30.40%)	−0.05 (0.12)	n.s.
Outcome: number of cycle collisions				
Time window 1	0.02 (0.09)	n.s.	0.22 (0.06)***	24.09% (7.30%)
Time window 2	0.30 (0.08)***	35.39% (11.52%)	0.28 (0.08)***	31.92% (9.97%)
Time window 3	0.37 (0.08)***	44.46% (12.38%)	0.15 (0.08)*	n.s.
Outcome: normalised collision rate				
Time window 1	0.06 (0.02)***	6.02% (2.61%)	0.28 (0.11)**	31.74% (14.13%)
Time window 2	0.01 (0.08)	n.s.	0.48 (0.13)***	61.35 (20.84)%
Time window 3	0.07 (0.08)	n.s.	0.18 (0.14)	n.s.

Figures in brackets represent clustered robust standard errors for the first two outcomes
Figures in the table represent the bootstrapped standard error for the third outcome
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

Table 14
The estimated impact on the number of single-cycle incidents for segregated versus non-segregated CS segments.

Time segment	Segregated CS		Non-segregated CS	
	Estimate	Percent change	Estimate	Percent change
Time window 1	0.14 (0.24)	n.s.	0.40 (0.16)**	48.95% (24.42%)
Time window 2	0.18 (0.28)	n.s.	−0.34 (0.30)	n.s.
Time window 3	0.63 (0.36)*	n.s.	−0.18 (0.24)	n.s.

Figures in bracket represent clustered robust standard errors
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

Table 15
The estimated impact segregation upgrades along CS2.

Outcome	PSM-DiD		Panel OR model with FE
	Estimate	Percent change	Estimate
Cycle volume	0.22 (0.09)**	24.80% (11.53%)	0.15 (0.07)**
Number of collisions	0.19 (0.09)**	21.88% (10.85%)	0.15 (0.07)**
Normalised collision rate	−0.05 (0.01)***	−5.08% (0.99%)	−0.07 (0.01)***

Figures in brackets represent clustered robust standard errors for the first two outcomes
Figures in the table represent the bootstrapped standard error for the third outcome
Significance levels - ***:99 percent, **:95 percent, *:90 percent, n.s.:not significant at 95 percent.

increase (95% confidence level) in single-cycle incidents at junctions during time window 1. In contrast, no such effect is observed along segregated CS in any of the time windows. This finding aligns with our hypothesis that non-segregated CS attracted a higher proportion of new, less experienced cyclists, who are more prone to difficulties in complex traffic navigation at junctions. Notably, while this effect is present in time window 1, it diminishes and becomes statistically insignificant in time windows 2 and 3, suggesting that either cyclist experience improved over time or route choices adjusted accordingly.

5.5.2. The impact of upgrades along CS2

Next, we treat the segregation upgrade along CS2 as a new intervention and re-estimate its effects on volume and safety. For this analysis, the study period is restricted to 2011–2019. We re-estimate Eqs. (2), (3) and (7). The results, summarised in Table 15, focus on the key parameters for brevity.

From Table 15, we observe that the upgrade led to a statistically significant increase in cycling volume (magnitude: 24.80%, standard error: 11.53%) and the number of cycle collisions (magnitude: 21.88%, standard error: 10.85%). Nevertheless, the normalised collision rate decreased significantly (magnitude: −5.08%, standard error: 0.99%), that is, the upgrade made the route less risky by 5.08% than expected.

Conclusions

London Cycle Superhighways (CS) play a pivotal role in the city's ongoing cycling revolution. In this paper, we analysed the causal impact of these superhighways on cycling volume and cyclist safety. The study examined 111 CS segments and 584 control segments over the period 2000–2019. Covariates included annual average daily traffic (AADT), pre-treatment annual average daily bicycle volume (AADB), pre-treatment cycle collisions, mean effective density (MED), and the index of multiple deprivation (IMD). To estimate the average treatment effects of interest, we employed propensity score-matched difference-in-differences and panel outcome regression with fixed effects. Our findings indicate that CS significantly increased cycling volumes, the number of cycle collisions and the normalised collision rate.

To account for heterogeneity in CS design and implementation, we adopted a multi-tiered analysis framework. We first estimated overall average treatment effects, then conducted a route-level analysis to assess variation across the seven operational CS routes. This was followed by a subgroup analysis comparing segregated versus unsegregated segments. Across all three levels, infrastructure quality, particularly the presence of physical segregation, emerged as a key contributor to differential outcomes. CS5 and CS6, which have featured fully-segregated lanes since inception, demonstrated the most favourable safety outcomes. In contrast, non-segregated routes, while often experiencing large increases in cycle volume, showed disproportionately higher increases in collision risk. This pattern was most evident on CS2, where a 2016 upgrade introducing full segregation was associated with a marked post-intervention reduction in normalised collision risk.

These findings suggest that segregation not only encourages greater uptake but also plays a protective role in managing the risks associated with higher cycling volumes. Importantly, our analytical framework integrates infrastructure heterogeneity descriptively through stratification and sub-group comparisons. Future work could further strengthen this approach by incorporating infrastructure characteristics directly into the causal models, for instance, via interaction terms or typology-based heterogeneous slope coefficients with richer and more consistently coded spatial data.

More broadly, these findings contribute to the growing discourse on urban transport policy by offering evidence-based assessments of how design choices shape cycling outcomes. As cities worldwide continue expanding active travel infrastructure, our results underscore the importance of prioritising high-quality, segregated facilities to enhance both uptake and safety.

Several limitations also point to directions for future research. While our models included key socioeconomic and traffic volume covariates, we were unable to include other potentially important factors such as traffic speeds, intersection density, or more granular measures of infrastructure quality. Our current binary segregation classification, though informative, could be extended to continuous measures of protected coverage, provided suitable data sources are available. Similarly, while our spatial matching relied on proximity-based allocation, more precise segment-level matching via reverse geocoding could improve estimation accuracy. Future studies may also benefit from adopting non-binary treatment effect estimation frameworks to better capture graded infrastructure improvements.

CRediT authorship contribution statement

Anupriya: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Xiaowei Zhu:** Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Emma McCoy:** Writing – review & editing, Project administration. **Daniel J. Graham:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was conducted as part of the Alan Turing Institute's Data-centric Engineering Challenge 1.0, headed by Professor Emma McCoy.

Data availability

Data will be made available on request.

References

Abdel-Aty, M.A., Lee, C., Park, J., Wang, J.-H., Abuzwidah, M., Al-Arifi, S., et al., 2014. Validation and Application of Highway Safety Manual (Part D) in Florida.. Technical report, Florida. Dept. of Transportation.

Agerholm, N., Caspersen, S., Lahrmann, H., 2008. Traffic safety on bicycle paths: Results from a new large scale Danish study.

Aldred, R., Elliott, B., Woodcock, J., Goodman, A., 2017. Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age. Transp. Rev. 37 (1), 29–55.

Anupriya, Graham, D.J., Hörcher, D., Anderson, R.J., Bansal, P., et al., 2020. Quantifying the ex-post causal impact of differential pricing on commuter trip scheduling in Hong Kong. Transp. Res. Part A: Policy Pr. 141, 16–34.

Ashenfelter, O.C., Card, D., 1984. Using the longitudinal structure of earnings to estimate the effect of training programs. Rev. Econ. Stat. 67 (4), 648–660.

Austin, P.C., 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. Stat. Med. 28 (25), 3083–3107.

Azimian, A., Pyrialakou, V.D., Lavrenz, S., Wen, S., 2021. Exploring the effects of area-level factors on traffic crash frequency by severity using multivariate space-time models. Anal. Methods Accid. Res. 31, 100163.

Buckley, A., Wilke, A., 2000. Cycle Lane Performance: Road Safety Effects. In: New Zealand Cycling Symposium.

Buehler, R., Dill, J., 2016. Bikeway networks: A review of effects on cycling. Transp. Rev. 36 (1), 9–27.

Burbidge, S.K., Goulias, K.G., 2009. Evaluating the impact of neighborhood trail development on active travel behavior and overall physical activity of suburban residents. Transp. Res. Rec. 2135 (1), 78–86.

Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. 22 (1), 31–72.

Chen, L., Chen, C., Srinivasan, R., McKnight, C.E., Ewing, R., Roe, M., 2012. Evaluating the safety effects of bicycle lanes in New York City. Am. J. Public Health 102 (6), 1120–1127.

de Dios Ortuzar, J., Iacobelli, A., Valeze, C., 2000. Estimating demand for a cycle-way network. Transp. Res. Part A: Policy Pr. 34 (5), 353–373.

DiGioia, J., Watkins, K.E., Xu, Y., Rodgers, M., Guensler, R., 2017. Safety impacts of bicycle infrastructure: A critical review. J. Saf. Res. 61, 105–119.

Dill, J., Voros, K., 2007. Factors affecting bicycling demand: initial survey findings from the Portland, Oregon, region. Transp. Res. Rec. 2031 (1), 9–17.

Finkelstein, A., 2002. The effect of tax subsidies to employer-provided supplementary health insurance: evidence from Canada. J. Public Econ. 84 (3), 305–339.

Fowler, M., Koorey, G., 2006. The effects of the pages road cycle lane on cyclist safety and traffic flow operations.

Gebel, M., Voßemer, J., 2014. The impact of employment transitions on health in Germany. A difference-in-differences propensity score matching approach. Soc. Sci. Med. 108, 128–136.

Graham, D.J., 2025. Causal inference for transport research. Transp. Res. Part A: Policy Pr. 192, 104324.

Gross, F., Persaud, B.N., Lyon, C., 2010. A Guide to Developing Quality Crash Modification Factors. Technical report,, United States. Federal Highway Administration. Office of Safety.

Han, C., Huang, H., Lee, J., Wang, J., 2018. Investigating varying effect of road-level factors on crash frequency across regions: a Bayesian hierarchical random parameter modeling approach. Anal. Methods Accid. Res. 20, 81–91.

Hansen, B.B., 2004. Full matching in an observational study of coaching for the SAT. J. Amer. Statist. Assoc. 99 (467), 609–618.

Hastings, J.S., 2004. Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California. Am. Econ. Rev. 94 (1), 317–328.

Heckman, J.J., Ichimura, H., Smith, J.A., Todd, P.E., 1998. Characterizing selection bias using experimental data.

Heckman, J.J., Ichimura, H., Todd, P.E., 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. Rev. Econ. Stud. 64 (4), 605–654.

Ho, D., Imai, K., King, G., Stuart, E., Whitworth, A., 2018. Package 'MatchIt'. Version.[Google Scholar].

Hopkinson, P., Wardman, M., 1996. Evaluating the demand for new cycle facilities. Transp. Policy 3 (4), 241–249.

Imbens, G.W., Wooldridge, J.M., 2009. Recent developments in the econometrics of program evaluation. J. Econ. Lit. 47 (1), 5–86.

Jensen, S.U., et al., 2008. Bicycle tracks and lanes: A before-after study. In: Transportation Research Board 87th Annual Meeting. Vol. 2095, Transportation Research Board Washington DC, United States, p. 15.

Li, H., Graham, D.J., Liu, P., 2017. Safety effects of the London cycle superhighways on cycle collisions. Accid. Anal. Prev. 99, 90–101.

Li, H., Graham, D.J., Majumdar, A., 2012. The effects of congestion charging on road traffic casualties: A causal analysis using difference-in-difference estimation. Accid. Anal. Prev. 49, 366–377.

Ling, R., Rothman, L., Cloutier, M.-S., Macarthur, C., Howard, A., 2020. Cyclist-motor vehicle collisions before and after implementation of cycle tracks in Toronto, Canada. Accid. Anal. Prev. 135, 105360.

Liu, Y., Chen, T., Chung, H., Jang, K., Xu, P., 2025. Is there an emotional dimension to road safety? A spatial analysis for traffic crashes considering streetscape perception and built environment. Anal. Methods Accid. Res. 100374.

Transport for London, 2011a. Barclays Cycle Superhighways Evaluation of Pilot Routes 3 And 7. Technical report.

Transport for London, 2011b. Barclays Cycle Superhighways FAQs.. Technical report.

Lusk, A.C., Furth, P.G., Morency, P., Miranda-Moreno, L.F., Willett, W.C., Dennerlein, J.T., 2011. Risk of injury for bicycling on cycle tracks versus in the street. Inj. Prev. 17 (2), 131–135.

Lv, H., Li, H., Sze, N., Zhang, Z., Ren, G., Zhang, Y., 2022. Analysis of factors influencing cycling count: A micro-level study using road segment count data in London. Int. J. Sustain. Transp. 1–16.

Mannering, F., Bhat, C.R., Shankar, V., Abdel-Aty, M., 2020. Big data, traditional data and the tradeoffs between prediction and causality in highway-safety analysis. Anal. Methods Accid. Res. 25, 100113.

Mattson, J., Askarzadeh, T., Ebrahimi, Z.D., 2022. Effects of Bicycle Facility Characteristics and the Built Environment on Bicycle Use: Case Study of Fargo-Moorhead. Technical report..

Mölenberg, F.J., Panter, J., Burdorf, A., van Lenthe, F.J., 2019. A systematic review of the effect of infrastructural interventions to promote cycling: strengthening causal inference from observational data. Int. J. Behav. Nutr. Phys. Act. 16, 1–31.

Naseri, M., Delbosc, A., Kamruzzaman, L., 2023. The role of neighbourhood design in cycling activity during COVID-19: An exploration of the Melbourne experience. J. Transp. Geogr. 106, 103510.

Nolan, J., Sinclair, J., Savage, J., 2021. Are bicycle lanes effective? The relationship between passing distance and road characteristics. Accid. Anal. Prev. 159, 106184.

Pai, C.-W., 2011. Overtaking, rear-end, and door crashes involving bicycles: An empirical investigation. Accid. Anal. Prev. 43 (3), 1228–1235.

Park, J., Abdel-Aty, M., Lee, J., Lee, C., 2015. Developing crash modification functions to assess safety effects of adding bike lanes for urban arterials with different roadway and socio-economic characteristics. Accid. Anal. Prev. 74, 179–191.

Parsons, J., Koorey, G., 2013. The effect of cycle lanes on cycle numbers and safety.

Pei, X., Wong, S., Sze, N.-N., 2011. A joint-probability approach to crash prediction models. Accid. Anal. Prev. 43 (3), 1160–1166.

Pucher, J., Buehler, R., 2017. Cycling towards a more sustainable transport future. Transp. Rev. 37 (6), 689–694.

Pulugurtha, S.S., Thakur, V., 2015. Evaluating the effectiveness of on-street bicycle lane and assessing risk to bicyclists in Charlotte, North Carolina. Accid. Anal. Prev. 76, 34–41.

Reynolds, C.C., Harris, M.A., Teschke, K., Cripton, P.A., Winters, M., 2009. The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. Environ. Heal. 8, 1–19.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70 (1), 41–55.

Smith, J.A., Todd, P.E., 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? J. Econometrics 125 (1–2), 305–353.

Strauss, J., Miranda-Moreno, L.F., Morency, P., 2013. Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. Accid. Anal. Prev. 59, 9–17.

Stuart, E.A., Green, K.M., 2008. Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes.. Dev. Psychol. 44 (2), 395.

Teschke, K., Harris, M.A., Reynolds, C.C., Winters, M., Babul, S., Chipman, M., Cusimano, M.D., Brubacher, J.R., Hunte, G., Friedman, S.M., et al., 2012. Route infrastructure and the risk of injuries to bicyclists: a case-crossover study. Am. J. Public Health 102 (12), 2336–2343.

Thomas, B., DeRobertis, M., 2013. The safety of urban cycle tracks: A review of the literature. Accid. Anal. Prev. 52, 219–227.

Zangenehpour, S., Strauss, J., Miranda-Moreno, L.F., Saunier, N., 2016. Are signalized intersections with cycle tracks safer? A case–control study based on automated surrogate safety analysis using video data. Accid. Anal. Prev. 86, 161–172.

Zoghi, M., 2013. The International Handbook of FRP Composites in Civil Engineering. CRC Press.