Understanding the cycle traffic impacts of Cycle Superhighways in London

Xiaowei Zhu[1], Daniel J. Graham[1,2,*], and Anupriya[2]

[1]Department of Mathematics, Imperial College London, London, SW7 2AZ, UK.
[2]Transport Strategy Centre, Imperial College London, London, SW7 2AZ, UK.

Short summary

Cycle Superhighways (CS) are the cycle routes that run between central London and outer London. They were introduced in 2008 as a way to encourage cycling and improve safety. This paper investigates the causal cycling demand and safety impacts arising from the introduction of CS. The analysis uses road traffic and accident data from the Department for Transport in the UK. Propensity score matching and panel outcome regression models are employed and compared to estimate the effects of CS for two different infrastructure types - segregated and non-segregated. Our results suggest that, on average, the intervention had a positive effect on cycle flow volume and cycle accidents, but no statistically significant effect on the cycle accident rate. Nevertheless, we find that segregated CS show a statistically significant decrease in cycle accident rate.
Keywords: cycleway investments, demand, safety, causal analysis, heterogeneous impact.

1 Introduction

Cycling has long been regarded as a healthy, economic, and environmental-friendly way of fulfilling one's day-to-day travel needs. With the aim to increase cycling in London, Cycle Superhighways (CS) were introduced across London in 2008. Figure 1 presents an initial route map of the CS. CS are cycle pathways extending from the outer parts of London to its centre[1], that were developed to enable safer, quicker, and more direct travel within the city. Several variants of CS have been also introduced across North America, Australia and Europe to serve the longer distance cycle commutes in metropolitan centres (Pucher & Buehler, 2017). The overarching aim of this paper is to contribute to the growing empirical evidence on the impact of such cycling infrastructure investments on cycle traffic and cyclist safety.

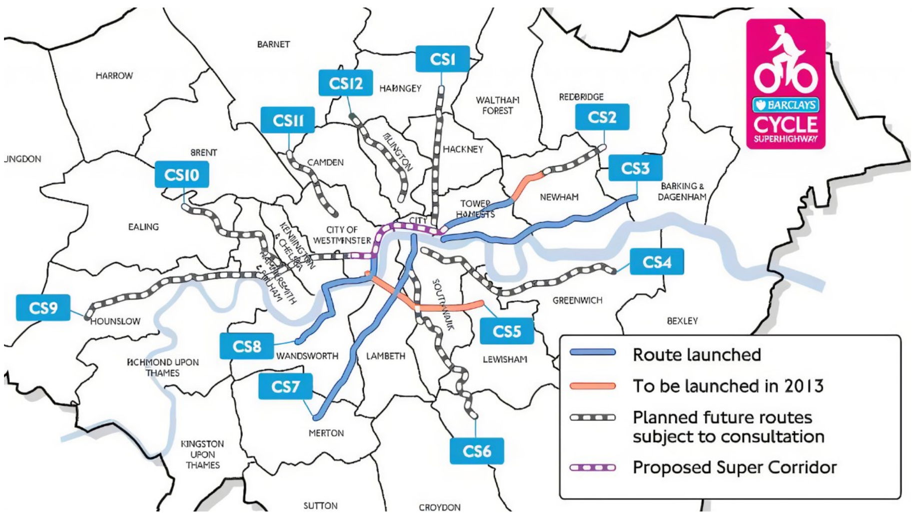

Figure 1: Route plan map of the Cycle Superhighways in London.

CS incorporated a variety of measures to improve cyclist safety including (Transport for London, 2011): (1) realigned traffic and bus lanes to create more space for cyclists on busy stretches of the

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

routes, (2) re-designed 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 (to a minimum of 5 meters) 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. However, given the associated infrastructure costs, the initial implementation of CS drew widespread criticism. Opponents claimed that the safety impacts of CS were overstated and referred to CS as nothing but blue paint[2]. It is, therefore, imperative to understand the traffic impacts of CS, particularly those related to cyclist safety. In this paper, we investigate the causal effect of CS on cycle flow volume, number of cycle accidents, and cycle accident rate.

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. We argue the impact estimated in these studies may suffer from confounding biases which occur primarily from the non-random nature of such 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: more cycling usually implies higher cycle-related accidents. Additional biases may emerge from temporal trends in the data. Thus, the estimates derived from a simple before-after comparison of demand and safety indicators may not reflect the true intervention effect.

In this study, we adopt causal inference approaches; (i) propensity score matching and (ii) panel outcome regression with fixed effects; that allow an unbiased estimation of the causal effect by effectively adjusting for such confounding factors. Our analysis uses road traffic and accident 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 rate at the network (that is, aggregate) level. However, we exploit the granularity of the data in hand to estimate the impact of CS for different infrastructure types - segregated versus non-segregated CS. We thus contribute with novel insights on how segregated CS segments perform with respect to the non-segregated ones.

2 Methodology

This section has two main subsections. The first subsection introduces the causal inference framework, which is followed by a description of the methods used in this paper: propensity score matching and panel outcome regression. Both of the methods are applied to compare the reliability of estimation. In the second subsection, we summarise the relevant details of the data that we use to estimate the impact of CS.

Causal Inference Framework

We use Rubin's Causal Model Rubin (1974) to develop a causal inference framework as follows: Let $Z_i = (Y_i, T_i, X_i)$ represents the observed data, where $i = 1, 2, \cdots, N$. N is the total population. Here, Y_i is the outcome of interest for an individual unit i. T_i denotes the binary treatment indicator. If $T_i = 1$, the unit i receives the treatment, otherwise, $T_i = 0$. X_i is the covariates, describing the characteristics of unit i. The potential outcome is defined as $Y_i(T)$ for each unit i. $Y_i(1)$ and $Y_i(0)$ represent the potential outcomes for unit i under treatment and control respectively. Then, the treatment effect for each unit i can be defined as:

$$\tau_i = Y_i(1) - Y_i(0) \tag{1}$$

However, we can only observe one of potential outcome of $Y_i(1)$ or $Y_i(0)$. As a result, we can not directly estimate each unit treatment effect τ_i. Instead, we can estimate the average treatment effect (ATE).

$$\tau_{ATE} = E[Y_i(1) - Y_i(0)] \tag{2}$$

We can estimate the ATE if the following three assumptions hold.

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

- Conditional Independence Assumption: This assumption means that given the observed covariates X, the potential outcomes are independent of the treatment, i.e. $Y(0), Y(1) \perp\!\!\!\perp T|X$.
- Common Support - This condition requires that each unit i has a positive probability of both receiving the treatment or not. There is no probability that one unit is always treated or untreated: $0 < P(T = 1|X) < 1$.
- Stable Unit Treatment Value Assumption(SUTVA) - The SUTVA requires that the observed outcomes under a given treatment allocation must be equivalent to potential outcomes under that allocation. i.e. $Y_i = I_1(T_i)Y_i(1) + (1 - I_1(T_i))Y_i(0), \forall i = 1, 2, \cdots, n$, where $I_1(T_i)$ in the indicator function for receiving the treatment.

All the three assumptions above are known as strong ignorability. Under this condition, the ATE (2) can be estimated from the observational data. This can be demonstrated as follows:

$$\begin{aligned}
\tau_{ATE} &= E[Y_i(1) - Y_i(0)] \\
&= E_X[E(Y_i(1)|X_i = x) - E(Y_i(0)|X_i = x)] \\
&= E_X[E(Y_i(1)|X_i = x, T_i = 1) - E(Y_i(0)|X_i = x, T_i = 0)] \\
&= E_X[E(Y_i|X_i = x, T_i = 1) - E(Y_i|X_i = x, T_i = 0)]
\end{aligned} \tag{3}$$

Propensity Score Matching

Propensity Score Matching (PSM) is a statistical method to estimate an intervention or treatment. The concept of PSM was first introduced by Rosenbaum and Rubin (1983) Rosenbaum & Rubin (1983) and further developed by Heckman et al. (1997) Heckman et al. (1997). Suppose that we want to compare the treatment effect between a control group and a treatment group. For example, in this study, the treatment is the construction of Cycle Superhighways(CS) or not. The treatment group contains those road segments with the installation of CS, while the control group includes other road segments. A naive method to estimate the treatment effect is to directly compare the difference between the two original groups. However, usually, the treatment is not assigned randomly on each individual. There exist confounding variables which affect both the treatment and the outcome.

To avoid this selection bias, we can do a matching between the control and treatment groups. PSM is one of the matching methods. It first uses models like logit or probit models to estimate the probability that each individual receives the treatment, which is called the propensity score (PS). Here we use logit model, which is defined as:

$$P(T = 1|X) = \frac{1}{1 + \exp(-(\alpha + X\beta))} \tag{4}$$

where α is the intercept and β is the coefficient vector. Then, each individual in the treatment group is matched to the individual in the control group with similar propensity score. There are four main matching algorithms: nearest neighbour matching, caliper and radius matching, stratification and Interval matching, kernel and local linear matching. Finally, the treatment effect is estimated by the difference between the two matched groups. If the strong ignorability assumption is satisfied and the matching algorithm is the nearest neighbour matching, then the ATE can be estimated by the following equation:

$$\hat{\tau}_{ATE} = \frac{1}{N_T} \sum_{i=1}^{N_T} \left(Y_i(1) - \hat{Y}_i(0) \right) \tag{5}$$

where N_T represents the total units of treatment group, $Y_i(1)$ is the outcome of ith unit in the treatment group, $\hat{Y}_i(0)$ is the closest unit in the control group in terms of propensity score distance that is matched to the ith treatment unit. It should be noted that different matching algorithm has different estimation equation. The main advantage of PSM is that it reduces the multiple dimension of matching to a single dimension, i.e. propensity score. More details of PSM can be found in Caliendo and Kopeinig (2008) Caliendo & Kopeinig (2008)

Panel outcome regression with fixed effects

One drawback of the propensity score matching method is that its performance highly relies on the choice of confounding factors. If there exist significant confounding factors but unobserved, the result may be unreliable. Thus, in order to control for these unobserved confounding factors,

we also implement panel outcome regression to compare with PSM. There are four main kinds of panel outcome regression models: pooled model, first differences, random effects and fixed effects. Here we introduce panel outcome regression with fixed effects. For more details of other models, see Wooldridge(2010) Wooldridge (2010).
Suppose that the data generating process is

$$y_{it} = X_{it}^T \beta + W_i^T \gamma + \epsilon_{it} \tag{6}$$

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 unobserved time-invariant covariates. ϵ_{it} is the error term. $E[\epsilon_{it}|X_{it}, W_i] = 0$, $i = 1, 2, \cdots, N$, $t = 1, 2, \cdots, T$. The fixed effects model assumes that each individual has a unique attribute that is constant through time. The panel model is

$$y_{it} = \alpha_i + X_{it}^T \beta + \epsilon_{it} \tag{7}$$

One possible way to estimate is to use the within estimator. The formula is as follows:

$$y_{it} - \bar{y_i} = (X_{it} - \bar{X_i})^T \beta + (\epsilon_{it} - \bar{\epsilon}_i)$$

where $\bar{y_i}$, $\bar{X_i}$, $\bar{\epsilon}_i$ are the respective mean over time. The advantage of the fixed effect model is that it can effectively deal with the unobserved time-invariant confounding factors. However, it fails to control for the time-varying confounding factors.

Data

In this section, we describe the relevant datasets and variables we used in this study. There are four main datasets which are highly related to this study: accident data, road data, cycleway data, socioeconomic data.

- Accident data

  The accident data is from STATS19, published by Department for Transport [3]. This dataset gives a detailed description of road accidents in Great Britain, including the date and location of accident, vehicle type, casualty details and severity. In this study, we only focus on the cycle-related accident data from 2000 to 2020 in Greater London.

- Road data

  The road data is from road traffic statistics, published by Department for Transport[4]. This dataset gives number of vehicles that travel past the count point on an average day of the year. Here, we use the annual average daily traffic volume (AADT) and annual average daily bicycle volume (AADB) from 2000 to 2020 in Greater London.

- Cycleway data

  The cycleway data is from public TfL data[5]. The dataset records the position and type of cycleway in London. In this paper, we mainly focus on the Cycle Superhighway.

- Socioeconomic data

  The socioeconomic data is from Office for National Statistics (ONS)[6] provides the data related to economy, population and society at national, regional and local levels in United Kingdom. Here, we use the population density, employee numbers, index of multiple deprivation (IMD) at the level of Lower Layer Super Output Areas (LSOA).

In this study, the observation unit is LSOA. We use the count points in road data as basis and link other dataset to the road data. To be more specific, for each record of the accident data, we calculate its nearest count point in road data. If the nearest distance is less than a pre-defined threshold (here we use 0.4 kilometer), then we can allocate the record to its nearest count point. Similarly, for each record of the cycleway data and socioeconomic data, we can allocate it to its nearest count point. And if the nearest distance of one record of cycleway data is greater than

[3] https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data.
[4] https://roadtraffic.dft.gov.uk/regions/6.
[5] https://cycling.data.tfl.gov.uk/.
[6] https://www.ons.gov.uk/.

a pre-defined threshold (e.g 1.5 kilometer), we can assume that there is no construction of Cycle Superhighway.
Although 12 CS routes had been planned, only part of them were put into use. In this paper, we studied 6 CS routes. The detailed description of these 6 CS routes is in Table 1.

CS No.	Open time	Length	Route	Type
CS1	2016 April	9.5km	The city to Tottenham	Partly segregated two-way cycle tracks. Most shared with bus
CS2	2011 July	6.8km	Stratford to Aldgate	Segregated cycle tracks are rarely seen. (Note: An upgradtion was added in 2016)
CS3	2010 July	12.3km	Barking to Tower Hill	Mostly segregated two-way cycle tracks.
CS5	2015 Autumn	1.4km	Oval to Pimlico	Completely segregated two-way cycle tracks.
CS7	2010 July	13.7km	Merton to the City	Segregated cycle tracks are rarely seen. Mostly shared with buses.
CS8	2011 July	8.2km	Wandworth to Wesminster	Segregated cycle tracks are rarely seen.

Table 1: The characteristics of each Cycle Superhighways

The distribution of these 6 CS routes can be seen in Figure 2.

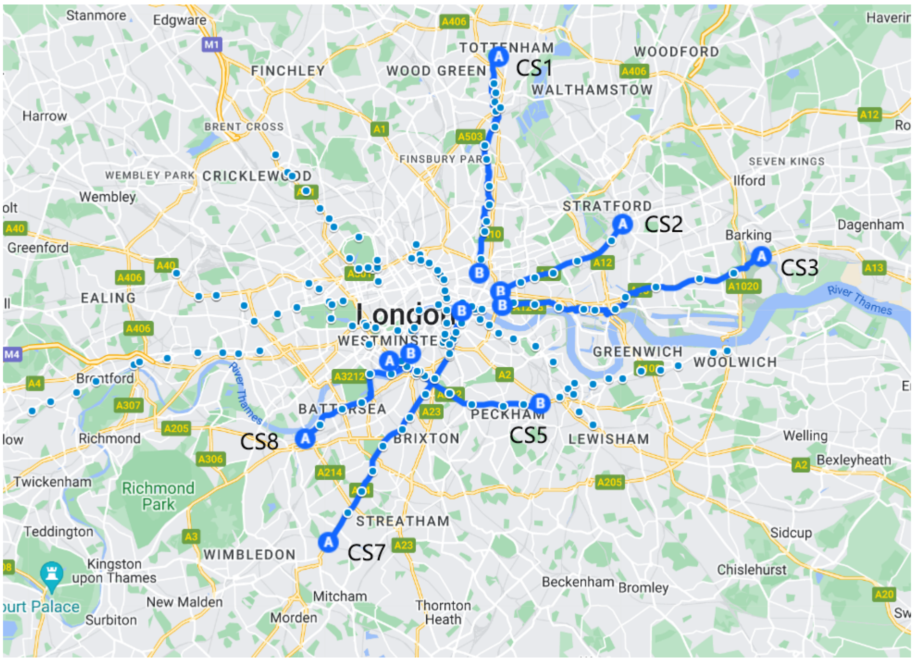

Figure 2: The distribution of 6 CS routes

Here, we choose 80 road segments with construction of CS as the treatment group, and select 434 road segments as the control group. Their distribution can be found in Figure 3. The red points in Figure 3 represent the CS segments (treatment group) while the blue points represent the control segments.
It should be noted that the installation of Cycle Superhighways is not randomly assigned. There exist some confounding variables which could not only affect the construction of Cycle Superhighways, but also influence the cycle accident rate. Here, we consider the following covariates: the traffic flow volume, the bicycle flow volume, the number of previous accidents, population density, employees, IMD, bus density. The choice of covariates is based on empirical findings. The description of the covariates is in Table 2.
In this study, the treatment variable is binary, representing presence of a Cycle Superhighway or not. If there is no CS around a road segment in a pre-defined distance (e.g 1.5 kilometers), we assume that there is no construction of CS in this road segment and the treatment variable is 0, otherwise, it is 1. We define the pre-intervention period as the three years before the CS is open, the post-intervention period as the three years after the open time. The outcome variables we are interested, include average cycle accident rate, cycle flow volume, the number of cycle accidents over the post-intervention period. The average cycle accident rate is defined as the mean of yearly cycle accidents divided by AADB during the post-intervention period. The cycle flow volume is reflected by the average AADB during the post-intervention period. The number of cycle accidents is the total number of cycle related accidents that happened during the post-intervention period.

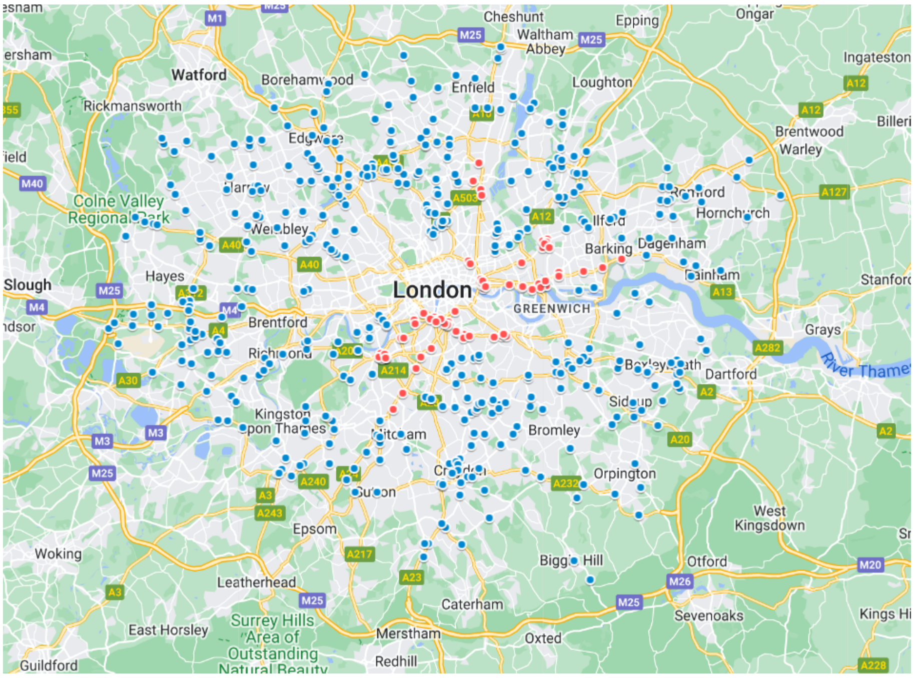

Figure 3: The distribution of treatment and control samples

variable	description	mean	std	min	max
AADB_pre	Annual average daily bicycle volume in the pre-intervention period	387.66	728.87	0.00	7167.33
AADT	Annual average daily traffic volume in the pre-intervention period	29666.31	24101.18	1825.66	202912.33
accident_pre	Total accident nums in the pre-intervention period	2.59	4.27	0.00	28.00
Population Density	population density in LSOA	75.77	51.06	1.16	437.04
Employees	Number of employees in LSOA	1468.48	5565.53	0.00	52000.00
IMD	The index of multiple deprivation	25.37	13.44	2.68	62.57
bus density	Number of bus stop within 0.5 km	13.38	6.47	0.00	35.00

Table 2: Descriptive statistics of the covariates

3 Results and discussion

In this section, we will estimate the effect of Cycle Superhighway (CS) on cycle accident rate, cycle volume and the number of cycle accidents. We will implement both PSM and panel outcome regression with fixed effects and compare the relative results. "MatchIt" package (?) in R is applied to perform the PSM. There are various kinds of matching methods in this package. Here, we use full matching because it performs quite well for the balance and overlap test. Next we will check the covariate balancing. The summary is in Table 3. Table 3 shows that all the variables are

	Std. Mean difference	M.Threshold	Variance Ratio
distance	0.0123	Balanced, <0.1	0.9936
AADB_pre	-0.0389	Balanced, <0.1	1.1077
AADT	0.0481	Balanced, <0.1	0.4973
accident_pre	0.0454	Balanced, <0.1	1.2428
Population.Density	-0.2404	Not Balanced, >0.1	1.3790
Employees	0.0606	Balanced, <0.1	0.3864
IMD	0.0386	Balanced, <0.1	1.1947
bus.density	-0.0576	Balanced, <0.1	0.7224

Table 3: Summary of balance for matched data

well-balanced except population density. The standardized mean difference of population density

is -0.2404. The absolute value is still not too big and population density may be considered as less important compared to other road characteristic variables. As a result, we can assume that the covariates achieve balance after matching.
Then, we check the overlap test by comparing the distribution plot of propensity score. The plot is in Figure 4. Figure 4 shows that before matching, the distributions of propensity score between

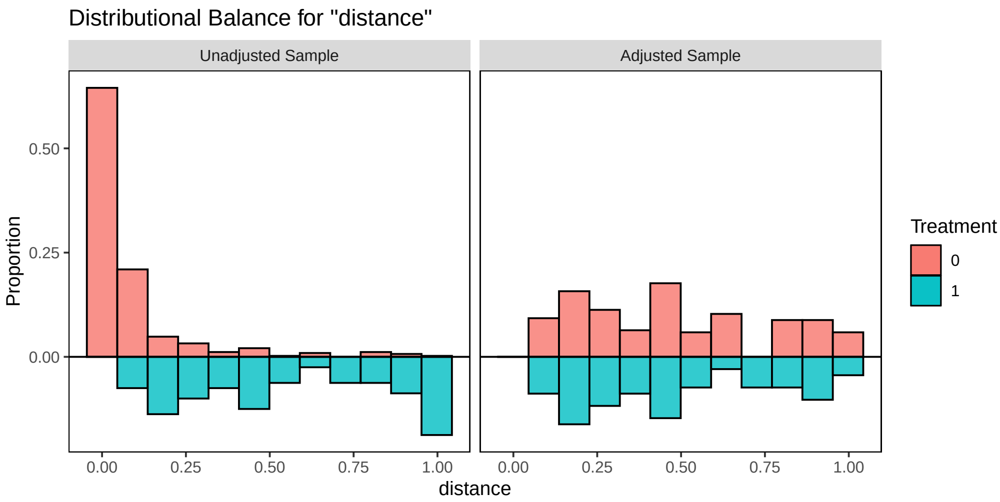

Figure 4: Overlap test

treatment and control groups are quite different. However, after matching they are similar to each other. Thus, there is support for the overlap assumption being verified.

Effect of CS on cycle accident rate

We can estimate the treatment effect ATE using the standard linear regression with matching weights. The coefficients and standard errors can be estimated by the lmtest (?) and sandwich packages (?) in R. First, we consider the effects of CS on cycle accident rate. The result is in Table 4.

	Estimate	Std. Error	t value	Pr(> $\lvert t \rvert$)
(Intercept)	62.0057	32.2701	1.9215	0.0561
CS_or_not	11.0589	23.6364	0.4679	0.6404
AADB_pre	-0.0401	0.0130	-3.0847	0.0023
AADT	0.0009	0.0003	2.7053	0.0074
accident_pre	1.8630	1.0624	1.7536	0.0811
Population.Density	0.0455	0.1462	0.3113	0.7559
Employees	-0.0037	0.0015	-2.5259	0.0123
IMD	-1.5655	0.7903	-1.9808	0.0490
bus.density	0.5267	0.4907	1.0733	0.2845

Table 4: t test of coefficients (CS on cycle accident rate)

From Table 4, we can see that the $p-$value of CS_or_not is 0.6404, which indicates that the CS has no significant effect on the cycle accident rate. Then, we will use panel outcome regression with fixed effect to estimate the effect of CS on cycle accident rate. This can be achieved by using the "plm" package (?) in R. The result is in Table 5.
We can see that in Table 5, the $p-$value for CS_or_not is 0.51 and the R^2 is 0.00119. This means that the effect of CS on cycle accident rate is also not significant and the model does not fit well. The result of panel outcome regression is similar to the one of PSM. Both of them suggest that there is no significant effect of CS on cycle accident rate.

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
CS_or_not	1.50e+02	2.25e+02	0.67	0.51
AADT	9.74e-03	1.48e-02	0.66	0.51
R^2	0.00119			

Table 5: Summary of panel OR(CS on accident rate)

Effect of CS on the cycle volume

In this section, we will examine the effect of CS on the cycle volume. The result for PSM is in Table 6. From Table 6, the $p-$value for CS_or_not is 0.0029 and the estimated coefficient is 273.

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	-1.82e+02	1.78e+02	-1.02	0.3069
CS_or_not	2.73e+02	9.02e+01	3.02	0.0029
AADB_pre	1.11e+00	8.12e-02	13.68	<2e-16
AADT	-5.89e-04	7.69e-04	-0.77	0.4446
accident_pre	-1.74e+01	2.35e+01	-0.74	0.4593
Population.Density	3.22e-01	6.53e-01	0.49	0.6229
Employees	5.02e-03	2.72e-03	1.85	0.0666 .
IMD	4.00e+00	3.53e+00	1.13	0.2584
bus.density	-2.77e+00	6.43e+00	-0.43	0.6668

Table 6: t test of coefficients (CS on the cycle volume)

This indicates that the introduction of CS significantly increased the number of cycle volume. In average, it could increase 273 in the number of AADB.
The result of panel outcome regression with fixed effects is in Table 7. From Table 7, the $p-$value

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
CS_or_not	3.32e+02	9.38e+01	3.54	0.00065
AADT	6.52e-04	1.22e-02	0.05	0.95741
accident_num	-2.56e+01	1.89e+01	-1.36	0.17835
R^2	0.137			

Table 7: Summary of panel OR(CS on the cycle volume)

of CS_or_not is 0.00065 and the estimated coefficient is 332. This also indicates that the CS has a significant effect on the cycle volume. The result is consistent with the previous PSM result and the coefficient 332 is close to 273, which is the coefficient using PSM. As a result, we can conclude that CS significantly increased the cycle flow volume.

Effect of CS on the number of cycle accidents

We will next estimate the effect of CS on the number of cycle accidents. The result of using PSM is in Table 8. In Table 8, the $p-$value of CS_or_not is 0.0113 and the estimated coefficient is 1.64. This suggests that the CS significantly increased the number of cycle accidents.
The result of panel outcome regression is in Table 9. The $p-$value of CS_or_not is nearly 0 and estimated coefficient is 2.46. The result is also similar to the previous PSM result. As a result, both PSM and panel OR method suggest that the CS significantly increased the number of cycle accidents.
So far, we have found that overall, the CS significantly increase the cycle flow volume and the number of cycle accidents with no significant effect on the cycle accident rate. However, as shown in Table 2, each CS has different segregated condition. Among them, CS5 is the only fully segregated, while CS2, CS7, CS8 rarely have any segregation installation. Considering the heterogeneity in different CS routes, we will perform the causal analysis on each CS route and study the effect of segregation.

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	-4.69e-01	1.49e+00	-0.31	0.7533
CS_or_not	1.64e+00	6.42e-01	2.56	0.0113
AADB_pre	1.42e-03	5.46e-04	2.60	0.0101
AADT	-1.91e-05	8.51e-06	-2.25	0.0256
accident_pre	7.98e-01	8.29e-02	9.62	<2e-16
Population.Density	2.20e-03	4.72e-03	0.47	0.6417
Employees	9.16e-05	3.19e-05	2.87	0.0046
IMD	1.26e-02	2.22e-02	0.57	0.5690
bus.density	8.38e-02	4.52e-02	1.85	0.0654

Table 8: t test of coefficients (CS on cycle accidents)

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
CS_or_not	2.46e+00	2.22e-01	11.06	<2e-16
AADB	-7.15e-04	2.60e-04	-2.75	0.0062
AADT	-2.82e-05	1.39e-05	-2.03	0.0427
R^2	0.181			

Table 9: Summary of panel OR(CS on the cycle accidents)

The safety effect of different CS routes

In this subsection, we evaluate the impact of each CS route. The result using PSM is in Table 10. The result using panel outcome regression with fixed effects is in Table 11.

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
CS1	1.5860	3.4992	0.4532	0.6541
CS2	3.9906	2.8908	1.3804	0.1843
CS3	3.5691	4.7206	0.7561	0.4539
CS5	-5.6549	1.9912	-2.8400	0.0078
CS7	7.7870	5.2608	1.4802	0.1492
CS8	-2.1071	2.5961	-0.8116	0.4188

Table 10: Summary of the effect of each CS using PSM

	Estimate	Std. Error	t-value	Pr(> $\|t\|$)
CS1	-53.9521	335.71	-0.1607	0.8725
CS2	257.69	252.90	1.0190	0.3098
CS3	101.13	115.12	0.8785	0.3806
CS5	-16.7454	181.34	-0.0923	0.9265
CS7	353.40	274.82	1.2859	0.2003
CS8	141.57	227.91	0.6212	0.5353

Table 11: Summary of the effect of each CS using panel OR

In Table 11, the result using panel outcome regression still does not show any significance. From Table 10, we can see that although other CS routes still do not show any significant effect, the $p-$value of CS5 is less than 0.05 and the estimated coefficient is -5.6549. This implies that CS5 significantly decreases the cycle accident rate. In Table 1, we can see that CS5 is the only CS route that is fully segregated. This indicates that the existence of segregation may be a crucial factor that influences the cycle accident rate.

The effect of Segregation on CS road

In this subsection, we will examine the effect of segregation on the CS road. Segregated cycle lane spared a space of the road for cycle use only. It is reported that there are lots of benefits of segregation. For example, the shift to segregated cycle lane can increase the carrying capacity of congested streets Aldred et al. (2017). Also, studies from Denmark have shown that segregated cycle lane reduces cyclists deaths by 35%[7]. Segregated cycle lanes are far more likely than those non-segregated ones to encourage people to cycle, especially women.
The characteristic of each road segment can be found at Google map. It not only provides the recent street view, but also provides the past few years' photos. With the help of Google map, we can denote each road segment as segregated or not. We also perform the PSM to inspect the effect of segregation on cycle accident rate. The result is in Table 12. From Table 12, we can see that

	Estimate	Std. Error	t value	Pr(> $\|t\|$)
(Intercept)	-28.6089	131.5069	-0.2175	0.8284
CS_Seg	-58.8391	19.3721	-3.0373	0.0033
accident_rate_pre	0.6367	0.1043	6.1033	4.934e-08
AADT	0.0024	0.0021	1.1437	0.2565
Road.type	96.5759	76.8731	1.2563	0.2131
Population.Density	-0.2186	0.4600	-0.4752	0.6360
Employees	0.0116	0.0025	4.5338	2.294e-05
IMD	3.9610	0.8972	4.4145	3.543e-05
bus.density	-7.6020	1.9579	-3.8826	0.00022

Table 12: t test of coefficients(CS_seg vs CS_non_seg)

the $p-$value of "CS_Seg" is nearly 0 and the estimated coefficient is -58.839. This suggests that the segregated Cycle Superhighways are significantly safer that those without segregation.

4 Conclusions

London Cycle Superhighways are a significant part of the "cycling revolution". In this paper, we studied the safety effect of Cycle Superhighways. 80 CS segments and 434 control segments were chosen to analyze. The covariates included annual average daily traffic (AADT), annual average daily bicycle volume (AADB), previous cycle accidents, population density, employees, index of multiple deprivation (IMD), and bus density. We implemented the propensity score matching and panel outcome regression with fixed effects to estimate the safety effect. Both of the methods showed that the installation of Cycle Superhighways has no significant effect on the cycle accident rate. However, they both indicated that the Cycle Superhighways significantly increased the cycle flow volume and the number of cycle accidents.
Then, we studied the heterogeneity of different Cycle Superhighways and found that CS5 performs best among these CS routes. CS5 is reported to be the only fully segregated CS. Thus, we next examined the effect of segregation among CSs. Using propensity score matching, it turned out that the segregated CS significantly decreased the cycle accident rate compared to those non-segregated CSs. As a result, in order to improve the safety of Cycle Superhighways, more segregation implementation should be encouraged.
For further study, one can consider more choices of the covariates. Limited by the data availability, we only considered AADT, AADB, previous cycle accidents, population density, employees, IMD, and bus density. One can also include some other plausible variables, e.g. traffic speed, intersection density.
Also, in this study, we constructed the model based on each "count point". This was because the dataset we got only had the latitude and longitude information. Thus, we just simply allocated each accident to its nearest count point. This might lead to some bias. A better choice is to do the reverse geocoding. That is, to map each coordinate to the corresponding road and construct the causal model base on each road. Then, each accident takes place exactly on its respective road.
Another possible improvement is to quantify the percentile of segregation for each road segment. In this paper, we only classified each road segment as segregated or not. To better inspect the

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

effect of segregation, one can quantify the percentile of segregation. In this case, the treatment variable is not binary any more. Instead, it becomes continuous. We would need to construct the propensity score matching with continuous treatment variable.

References

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. *Transport reviews*, *37*(1), 29–55.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31–72.

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of safety research*, *61*, 105–119.

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, *64*(4), 605–654.

Li, H., Graham, D. J., & Liu, P. (2017). Safety effects of the london cycle superhighways on cycle collisions. *Accident Analysis & Prevention*, *99*, 90–101.

Pucher, J., & Buehler, R. (2017). Cycling towards a more sustainable transport future. *Transport reviews*, *37*(6), 689–694.

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.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology*, *66*(5), 688.

Transport for London. (2011). *Barclays cycle superhighways faqs.* (Tech. Rep.).

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