Bikeability and the induced demand for cycling

Fosgerau, Mogens; ukawska, Mirosawa; Paulsen, Mads; Rasmussen, Thomas Kjær

Published in:
Proceedings of the National Academy of Sciences of the United States of America

DOI:
10.1073/pnas.2220515120

Publication date:
2023

Document version
Early version, also known as pre-print

Document license:
Unspecified

Citation for published version (APA):
Fosgerau, M., ukawska, M., Paulsen, M., & Rasmussen, T. K. (2023). Bikeability and the induced demand for cycling. Proceedings of the National Academy of Sciences of the United States of America, 120(16), e2220515120. https://doi.org/10.1073/pnas.2220515120

Download date: 23. jul., 2024
Bikeability and the induced demand for cycling

Fosgerau, Mogens; Lukawska, Miroslawa; Paulsen, Mads; Rasmussen, Thomas Kjær

Publication date:
2022

Citation for published version (APA):
Fosgerau, M., Lukawska, M., Paulsen, M., & Rasmussen, T. K. (2022). Bikeability and the induced demand for cycling.
Bikeability and the induced demand for cycling

Mogens Fosgerau\textsuperscript{a}, Mirosława Łukawska\textsuperscript{b}, Mads Paulsen\textsuperscript{b}, and Thomas Kjær Rasmussen\textsuperscript{b}

\textsuperscript{a}University of Copenhagen

\textsuperscript{b}Technical University of Denmark

October 7, 2022

Significance statement

Across the world, there is considerable interest in promoting bicycle use. In addition to reducing the climate impact of urban transportation, it will help to improve public health, and reduce traffic congestion, noise and air pollution. Provision of bicycle-friendly infrastructure is a primary means to achieving this. Using a large dataset of observed bicycle trip trajectories and fine-grained network data covering the city of Copenhagen, Denmark, this paper finds a large effect of infrastructure provision on the volume of bicycle traffic.

Abstract

How much is the volume of urban bicycle traffic affected by the provision of bicycle infrastructure? We exploit a large dataset of observed bicycle trajectories in combination with a fine-grained representation of the Copenhagen bicycle-relevant network. We apply a novel model for the bicyclist choice of route from origin to destination that takes the complete network into account. This enables us to back out the bicyclist preferences for a range of infrastructure and land-use types. We use the estimated preferences to compute a subjective cost of bicycle travel, which we correlate with the number of bicycle trips across a large number of origin-destination pairs. Simulations suggest that the extensive Copenhagen bicycle lane network has increased the number of bicycle trips by 40 percent and the volume of bicycle km by 60 percent, compared to a counterfactual without the bicycle lane network. This translates into an annual benefit worth 0.4 M EUR per km of bicycle lane due to changes in subjective travel cost, health, and accidents. Our results thus strongly support the provision of bicycle infrastructure.

Introduction

Provision of bicycle infrastructure is a means to encourage cycling, which may in turn improve the urban environment and reduce the climate impact of transport [1]. On the other hand, infrastructure is costly and takes up valuable urban space. It is therefore of interest to be able to assess the impact of providing bicycle infrastructure on cycling, taking into account that the location and type of infrastructure matter a lot [2].

Copenhagen has extensive bicycle infrastructure and a high level of bicycle usage for everyday urban travel [3]. We have a dataset of unprecedented size available, comprising 218,489 bicycle trajectories in Copenhagen obtained from users of Hövding airbag helmets.\textsuperscript{1} Matching these trajectories with very detailed network information (see Figure 1a) allows us to track observed bicycle route choices across a range of infrastructure and land-use types.

To make inference regarding the factors influencing bicyclists' choice of routes, we must compare the observed chosen routes to the alternative possibilities. However, the number of possible routes between two points in a large network is extremely large and literally impossible to enumerate. To overcome this, we exploit the recent perturbed utility route choice model [4], which allows the whole network to be taken into account while being computationally feasible.

The estimated route choice model indicates very substantial variation in the subjective cost of using various infrastructure types. The subjective cost per metre travelled on the most attractive infrastructure type, a cycleway, designated as a cycle superhighway, located in a green area, is 7 times lower than the subjective cost of cycling on a residential road.

There is thus very considerable scope for bicycle infrastructure to affect the subjective cost of cycling and thereby the level of bicycle use. We are able to relate the volume of trips in each origin-destination (OD) pair to the subjective cost from the route choice model. Employing a variant of the gravity model [5], we find a clear relationship whereby lower subjective cost is associated with more bicycle trips. We use this model to simulate a range of counterfactual scenarios, exploring the impact of the bicycle network on the volume of bicycle use.

\textsuperscript{1}https://www.hovding.com
Results

Bicycle route choice

We employ the perturbed utility route choice model [4]. It is a perturbed utility model [6, 7, 8], adapted to describe the choice of path through a network. For each OD pair, the predicted behaviour of bicyclists is a vector $x \in \mathbb{R}^{|E|}$ that represents the distribution of flow across the network links $e \in E$. The model assumes that the observed flow $x$ minimises a convex cost function under the constraint that the flow $x$ is physically consistent with a mass one flow through the network from origin to destination. The cost does not involve monetary elements, but is a subjective cost that represents the bicyclists preferences for various infrastructure types. The cost function has the form

$$C(x) = \sum_{e \in E} l_e (c_e x_e + F(x_e)), \quad (1)$$

where $l_e$ are link lengths, $x_e$ are link flows, and link cost rates are specified as $c_e = z_e' \beta$, where $z_e$ is a vector of link characteristics and $\beta$ is a vector of parameters to be estimated. The perturbation function $F(\cdot) : \mathbb{R}_+ \to \mathbb{R}$ is defined as $F(x_e) = (1 + x_e) \ln(1 + x_e) - x_e$, which is a convex function with $F(0) = F'(0) = 0$. The perturbation term provides incentive to distribute flow on more than one route, while allowing the cost-minimizing flow to be zero on most links.

The materials and methods section explains the data and estimation methods with further details given in the SI Appendix. The full specification of the link characteristics vector $z_e$ and the corresponding parameter estimates are given in SI Appendix Table 4.

Figure 1b shows the estimated link cost rates on a map of the network. The main bicycle network is clearly visible and seems to be quite dense and well connected. Such maps may be used by urban planners to suggest areas where the bicycle network could be improved. The specification of the cost rate comprises variables that indicate the type of infrastructure for each link in the network including information about the type of bicycle infrastructure, and the nearby land-use. We will discuss the most important parameters in turn.

Figure 1: Maps representing the network used in the case study.
The reference category is residential roads without specific bicycle infrastructure in low-rise urban areas. Compared to the reference, bicyclists associate a 11 % higher cost rate with large roads (roads with at least two lanes in one direction), while the difference to medium roads (roads with at most one lane in each direction) is small and statistically insignificant.

Provision of dedicated bicycle infrastructure reduces the subjective cost of bicycling quite substantially. Cycleways (bicycle paths in own trace) reduce the subjective cost by 20 %. On residential and medium roads, bicycle lanes, whether protected or just painted, reduce the cost rate by 14 % and 22 %, respectively. The type of bicycle lane matters a lot for the large roads category: painted bicycle lanes have only a small and statistically insignificant effect, while protected bicycle lanes reduce the cost rate by 34 %. It makes clear intuitive sense that the impact of bicycle lanes is larger, the larger the road is, and that only protected lanes have impact on the largest roads where car traffic is heavier.

A number of routes are branded as so-called bicycle superhighways. This is a label that is applied to high-quality, continuous bicycle routes, catering to commuter cyclists [9]. Another set of routes are planned to become bicycle superhighways in the future, but have not yet received the label [10]. We estimate a cost rate reduction of 12 %, both for the actual bicycle superhighway links and for the planned bicycle superhighway links. This suggests that the routes included in the bicycle superhighway network were ante attractive and that the transformation from planned to actual bicycle superhighway does not yield any additional cost reductions beyond those already accounted for at the link level.

The model also includes parameters accounting for interactions between the type of infrastructure and the neighboring land use. The cost rate is much reduced for cycleways in industrial areas (48 %) or green areas (53 %). It makes intuitive sense that cycleways in green areas may be pleasant. Another potential explanation which also applies for industrial areas is the attractiveness of isolation from heavy traffic.

In summary, provision of bicycle-friendly infrastructure has a substantial effect on route choice. We shall see below that this translates into a substantial effect on the volume of bicycle trips.

Bicycle travel demand

We now set up a gravity model to measure the association between the volume of bicycle trips in each OD pair and the characteristics of the bicycle network. The route choice model parameter estimates show that the characteristics of the network have a large impact on the subjective cost of using the links of the network. It is therefore natural to use the route choice model to compute the subjective cost of travelling by bicycle in any given OD pair, thereby aggregating the network information in a model-consistent manner.

Let \( \hat{c}_{od} = \sum_{e \in E} l_e \left( c_e \hat{\gamma}_{e,od} + F(\hat{\gamma}_{e,od}) \right) \) be the subjective cost associated with the cost-minimizing flow \( \hat{\gamma}_{e,od} \) connecting origin \( o \) to destination \( d \) and let \( Y_{od} \) be the observed number of trips in OD-relation \( od \). We assume that \( Y_{od} \) follows a Poisson distribution with expectation given as a loglinear function of the cost:

\[
\ln E \left[ Y_{od} \right] = D \left( \hat{c}_{od} \right) + \delta + \eta_o + \gamma_d,
\]

where \( \delta \) is a constant, and \( \eta_o \) and \( \gamma_d \) are constants for each origin and destination except one. The constants account for the total bicycle traffic volume out of each origin and the total volume into each destination. The demand function \( D \) is expected to be downward sloping, such that higher cost implies less volume. We specify \( D \) to be continuous and piecewise linear with the number of pieces chosen by eye-balling.

Figure 2 shows the estimated demand function. The figure also shows a bootstrapped pointwise confidence band for the estimate of the demand function. The confidence band is very tight at small values of \( \hat{c}_{od} \), while the difference to medium roads (roads with at most one lane in each direction) is small and statistically insignificant. From the estimated demand function, we can compute the implied demand elasticity as a function of the cost. We find that the demand elasticity decreases close to linearly from 0 to about -6 as the cost increases from 0 to 8. A 10 % increase in the cost of a short trip thus has very little impact on the volume of bicycle trips while it reduces the volume of trips by up to 60 % for the longest trips.

Figure 2: Piecewise linear estimate of \( D \) using 10 knots. \( n = 39,172 \), \( R^2 = 0.41 \), \( RMSE \) (in levels) = 10.7. Red lines trace pointwise confidence intervals at various quantiles computed using 1,000 bootstraps. Black dashed lines indicate the estimated 95% pointwise confidence band.
Table 1: Key figures of three counterfactuals. Numbers in square brackets indicate 95% bootstrap confidence intervals based on 1,000 full sample repetitions. *Measured as route distance rather than network distance.

**Counterfactuals**

We now exploit our model to simulate the impact of general counterfactual changes to the bicycle-relevant network, chosen to illustrate how much bicycle infrastructure has contributed to encourage bicycling in Copenhagen. These results may be of interest for other cities, where it is considered to improve or expand the bicycle-relevant network. Table 1 summarizes the results for the counterfactual scenarios. We compute the change in consumer surplus for bicyclists as well as the change in external cost due to health and accidents. We have scaled the gravity model output such that the base scenario matches the total volume of bicycle traffic in the region. A full economic evaluation of constructing bicycle infrastructure would also need to take into account construction costs, the effects on travel times by car, and the induced effects on climate, accidents, noise, and air pollution.

The first counterfactual simulates a situation where all 1,428.2 km of bicycle lanes have been removed as well as all cycle superhighway classifications. On average, this increases the subjective cost of bicycling by 20.9% per km, which induces a reduction of 27.9% in the number of bicycle trips and 37.4% in the total distance travelled by bicycles. The bootstrapped confidence intervals indicate that these numbers are quite precisely determined. Relative to the situation without bicycle lanes, the simulation thus suggests that the provision of bicycle lanes has induced an increase in bicycle use of 39% (trips) and 60% (total km).

To measure the loss to bicyclists in the counterfactual scenario compared to the base scenario, we have computed the change in the consumer surplus [11]. To convert this number from subjective cost units to monetary values, we apply the sample average speed to convert the subjective cost to time units, then we apply the official Danish value of travel time [12] to convert from time to monetary units. Our numbers suggest a decrease in consumer surplus of 161.3 M € per year. Bicycling is associated with both health benefits and accident risk [e.g., 13, 14]. The official Danish guidelines for cost-benefit analysis suggest a net external benefit due to health and accidents of 0.91 EUR per bicycle km [12]. Applying this figure, we estimate the welfare loss induced by removing the bicycle lane network through health and accidents to be 348.5 M € per year. In total we find a loss if all bicycle lanes were removed of 509.8 M € per year or 0.342 M € per km of bicycle lane.

In the second counterfactual, we convert the 407.0 km of protected lanes on large roads to just painted lanes, while maintaining the cycle superhighway classifications. This increases the subjective cost of bicycling by 4.5% on average, which in turn induces 7.2% less bicycle trips and 10.5% less bicycle km. We find that the removal of protected lanes leads to an annual loss of 41.0 M € of consumer surplus and an annual loss due to health and accidents of 16.0 M € in total, we compute an annual loss of 139.1 M € or 0.342 M € per km of protected bicycle lane.

The third counterfactual removes the existing and planned cycle superhighway classifications. We interpret this as representing the effect of no longer having long connected bicycle routes. The change involves 333.7 km of cycle superhighways routes and leads to an average increase of 7.6% in the subjective cost, which induces a reduction of 11.9% in the number of trips and 17.2% in the number of bicycle km. Removing the route level features that constitute the cycle superhighways is associated with a annual total loss of 227.0 M € or 0.680 M € per lane km.

We find in all three counterfactuals that the total distance travelled by bicycle responds relatively more than the number of trips. This means that the number of long bicycle trips responds more than the number of short trips, in alignment with our observation that the demand elasticity increases with the trip cost.

**Discussion**

We find substantial impact of the provision of bicycle-relevant infrastructure on the subjective cost and the volume of cycling. We work at a very fine level of resolution, which allows us to distinguish between a large number of infrastructure types. This is first order important as we find a difference in the subjective cost of cycling of more than a factor 8 between the best and the worst infrastructure types. Thus it matters very much which infrastructure is provided and where.

We have carried out counterfactual simulations, illustrating the effect of broad changes to the bicycle network. These results may be of interest when considering the consequences of expanding the bicycle network in cities with less bicycle infrastructure than Copenhagen. We find that the existing bicycle network in Copenhagen has led to a substantial increase in the volume of bicycle traffic of about 60%. These changes can be interpreted as representing short-term effects, as they hold constant the fixed effects associated with origins and destinations. In the longer term, location patterns can be expected to adapt to an improvement in the bicycle network, making the long-term effect of a network improvement
larger than the short-term effect. Previous research supports the broad conclusion that bicycle infrastructure induces more bicycle traffic [e.g., 2, 15, 16, 17].

From the counterfactual scenarios, we have calculated the net benefit of bicycle lane provision associated with the change in subjective cost, health and accidents to be 340–370 k€ per lane km per year. According to [18], construction costs are in the range 0.5 – 1.5 M€ per lane km.\textsuperscript{2} The estimated benefit associated with cycle superhighway status is greater, 680 k€ per km per year, even though it relates only to route level features, holding link level features constant. As construction are incurred once while benefits accrue year by year, these results indicate that provision of well located and high quality bicycle infrastructure can easily generate a positive net present value in a standard cost-benefit analysis.

Copenhagen already has a lot of bicycle infrastructure, so the effect of additional infrastructure may be smaller. On the other hand, we find a large net benefit of cycle superhighways, which may arise from having long and connected bicycle routes. This means that limited investments can potentially lead to large net benefits by improving overall connectivity of the bicycle network. Maps like Figure 1b can be used to identify candidate locations for such investments.

We have combined a very large database of observed bicycle trajectories and a very fine-grained representation of the bicycle-relevant network with a modelling approach that allows us to take the whole network into account. Our model can be applied to predict the effect of providing specific infrastructure in specific places. Similar analyses can be undertaken for other cities. The main obstacle is obtaining sufficient data on observed route choices similar to the Hövding dataset that we have used.

This is the first study of its kind, so there is much scope for future research. On the bicycle front, it is of interest to estimate similar models on datasets from other cities, in order to consolidate and extend the conclusions regarding the impact of bicycle infrastructure on bicycle demand. Similarly, it is of interest to investigate datasets sampled by different means or from different kinds of users to check the robustness of our conclusions.

On the methodological front, a general research agenda can be formulated for the perturbed utility route choice model, with a view to applications to bicycle traffic or other traffic through complex networks. The most important point here, we think, is to develop approaches that allow the estimation of the models at the level of individual trajectories. This would make it possible to avoid the data loss associated with the aggregation of data to the OD level and allows the inclusion of individual level information. A related point is to develop solution methods for the cost minimization problem (1) that make it feasible to work with meta-networks where link-pairs take the place of links. This would allow turn movements to be represented and hence allow to take into account, e.g., the cost of left-turns, crossing roads with car traffic.

Materials and methods

Bicycle data

Figure 3 shows the raw data from our sample of GPS traces of bicycle trips, collected in Greater Copenhagen. Figure 1a presents corresponding views of the road and dedicated bicycle network, utilizing Open Street Map data\textsuperscript{3}. Copenhagen has an extensive bicycle network with many cycleways, especially outside central Copenhagen. In central Copenhagen there is a dense network of protected bicycle tracks and also many non-protected bicycle lanes.\textsuperscript{4}

SI Appendix 2.3-2.4 describes how we processed our data. In brief, the pre-processed dataset of 218,489 GPS trajectories collected from 8,588 individuals were map-matched to the Copenhagen bicycle network using software presented in [19]. Our estimator for the route choice model requires trips to be aggregated such that there are at least two distinct trajectories for each origin-destination (OD) pair. Therefore we applied an algorithm that trims individual trajectories at both ends such that the trimmed trajectories have a small number of origins and destinations in common. Our analyses are based on data with 200 origins and 200 destinations, which represents a compromise between including most of the observed bicycle travel and avoiding many OD pairs with only a small number of observations. Robustness check with 100 and 400 origins and destinations did not indicate problems, see SI Appendix 3.2.3. We use data for all OD pairs that are more than 1 km apart and have at least two distinct individual route choice observations per OD. Our estimation data comprises 152,323 trimmed trips from 7,672 individuals.

Route choice model

A directed network consisting of nodes and links (\(V, E\)) is described by the incidence matrix \(A\) with elements \(a_{ve} = 1\) if \(v\) is the origin node of link \(e\), \(-1\) if it is the destination node and 0 otherwise. A set of OD pairs \(B\) is given in terms of OD demand vectors \(b \in B \subset \mathbb{R}^{\mid V\mid}\), where \(b_{v} = 1\) indicates the origin node of trip \(b\), \(b_{v} = -1\) indicates the destination node and \(b_{0} = 0\) otherwise. The flow conservation constraint \(Ax = b\) ensures that a non-negative flow vector \(x \in \mathbb{R}_{+}^{\mid E\mid}\) is physically consistent with demand \(b\) through the network.

The PURC model holds that the flow vector \(\hat{x}\) for bicyclists with demand \(b\) minimizes the cost function (1) under the flow constraint \(Ax = b\). Fosgerau et al. [4] show that this model generates very reasonable substitution patterns. Moreover, the model directly applies to the complete network, without any need to specify a choice set of route alternatives. Given (noisy) observations of flow vectors, [4] transform the active first-order conditions for the cost minimization problem to a linear regression equation that leads directly to an estimate of \(\hat{\beta}\). There is an active first-order condition

\textsuperscript{2}Lower if the numbers in [18] pertain to route km.
\textsuperscript{3}https://www.openstreetmap.org
\textsuperscript{4}See SI Appendix 2.1 for definitions of the various infrastructure types.
for each link with positive observed flow. The transformation eliminates Lagrange multipliers corresponding to the flow conservation constraints at each node of the network. The data for the regression comprises many observations for each OD, which enables standard errors to be clustered by OD.

Figure 4 plots the total observed link flow \( (x_e = \sum_{o \in O} \sum_{d \in D} x_{od}^e) \) against the total predicted link flow \( (\hat{x}_e = \sum_{o \in O} \sum_{d \in D} \hat{x}_{od}^e) \) for each link \( e \in \mathcal{E} \) across all origins and destinations. A perfect prediction would follow the 45° line exactly. We find that a non-parametric regression line tracks the the 45° line quite closely. This is satisfactory, especially keeping in mind that the route choice model uses 28 parameters. The correlation (defined in the SI Appendix (2)) between \( x^e \) and \( \hat{x}^e \), is \( \rho(x^e, \hat{x}^e) = 0.9061 \).

Figure 4: Heatmap of total observed link flow \( (x_e = \sum_{o \in O} \sum_{d \in D} x_{od}^e) \) against the total predicted link flow \( (\hat{x}_e = \sum_{o \in O} \sum_{d \in D} \hat{x}_{od}^e) \) for each link \( e \in \mathcal{E} \). The color of each grid cell represents the number of links belonging to each cell. The thin green dotted line is the 45° line. The black line is a Nadaraya-Watson non-parametric regression [20, 21] with Gaussian kernel and bandwidth 12 chosen by eyeballing. The corresponding 95% pointwise confidence band is indicated by dashed red lines.

SI Appendix 3.2 comprises a range of validation tests of the route choice model.
References

[1] J. Pucher and R. Buehler. “Making Cycling Irresistible: Lessons from The Netherlands, Denmark and Germany”. In: Transport Reviews 28.4 (2008), pp. 495–528. DOI: 10.1080/01441640701806612. URL: https://www.tandfonline.com/doi/abs/10.1080/01441640701806612.

[2] R. Buehler and J. Dill. “Bikeway networks: A review of effects on cycling”. In: Transport Reviews 36.1 (2016), pp. 9–27. DOI: 10.1080/01441647.2015.1069908. URL: https://www.tandfonline.com/doi/full/10.1080/01441647.2015.1069908.

[3] T. Koglin, M. te Brömmelstroet, and B. van Wee. “Cycling in Copenhagen and Amsterdam”. In: Cycling for Sustainable Cities. Ed. by R. Buehler and J. Pucher. Cambridge, Massachusetts: The MIT Press, 2021. Chap. 18, pp. 347–370. DOI: 10.7551/MITPRESS/11963.003.0022. URL: https://direct.mit.edu/books/book/5000\chapter/2654063/Cycling-in-Copenhagen-and-Amsterdam.

[4] M. Fosgerau, M. Paulsen, and T. K. Rasmussen. “A perturbed utility route choice model”. In: Transportation Research Part C 136.103514 (2022). DOI: https://doi.org/10.1016/j.trc.2021.103514. URL: https://arxiv.org/abs/2103.13784.

[5] A. Wilson. “A statistical theory of spatial distribution models”. In: Transportation research 1.3 (Nov. 1967), pp. 253–269. DOI: 10.1016/0041-1647(67)90035-4. URL: http://scholar.google.com/scholar?q=intitle:A+statistical+theory+of+spatial+distribution+models&hl=en&btnG=Search&q=a+statistical+theory+of+spatial+distribution+models#0.

[6] M. Fosgerau and D. L. McFadden. “A theory of the perturbed consumer with general budgets”. In: NBER Working Paper (Mar. 2012), pp. 1–27. DOI: 10.3386/w17953. URL: http://www.nber.org/papers/w17953.pdf.

[7] D. Fudenberg, R. Iijima, and T. Strzalecki. “Stochastic Choice and Revealed Perturbed Utility”. In: Econometrica 83.6 (2015), pp. 2371–2409. DOI: 10.3982/ECTA12660. URL: http://scholar.harvard.edu/files/fudenberg/files/stochastic_choice_and_revealed.pdf%5Cnhttps://www.econometricssociety.org/doi/10.3982/ECTA12660.

[8] R. Allen and J. Rehbeck. “Identification With Additively Separable Heterogeneity”. In: Econometrica 87.3 (May 2019), pp. 1021–1054. DOI: 10.3982/ecta15867. URL: https://www.ssrn.com/abstract=3073945.

[9] G. Liu, M. te Brömmelstroet, S. Krishnamurthy, and P. van Wesemael. “Practitioners’ perspective on user experience and design of cycle highways”. In: Transportation Research Interdisciplinary Perspectives 1 (2019), p. 100010. DOI: 10.1016/j.trc.2019.100010. URL: https://www.sciencedirect.com/science/article/pii/S2590198219300107.

[10] Sekretariat for SuperCykelstier. Rateoversigt: Visionsplan 2017-2045. Technical report. 2017. URL: www.supercykelstier.dk.

[11] A. Mas-Colell, M. Whinston, and J. Green. Microeconomic Theory. Oxford: Oxford Press, 1995, p. 1008. DOI: 10.2307/135312.

[12] Technical University of Denmark. Transport Economic Unit Prices vers. 2.0. 2022. URL: https://www.cta.man.dtu.dk/modelbibliotek/teresa/transporteokonomiske-ehedspriser.

[13] J. J. de Hartog, H. Boogaard, H. Nijland, and G. Hoek. “Do the Health Benefits of Cycling Outweigh the Risks?” In: Environmental Health Perspectives 118.8 (Aug. 2010), pp. 1109–1116. DOI: 10.1289/ehp.0901747. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2920084/.

[14] P. Oja, S. Titze, A. Bauman, B. de Geus, P. Krenn, B. Reger-Nash, and T. Kohlberger. “Health benefits of cycling: A systematic review”. In: Scandinavian Journal of Medicine and Science in Sports 21.4 (2011), pp. 496–509. DOI: 10.1111/j.1600-0838.2011.01299.x.

[15] J. Pucher, J. Dill, and S. Handy. “Infrastructure, programs, and policies to increase bicycling: An international review”. In: Preventive Medicine 50 (Jan. 2010), S106–S125. DOI: 10.1016/j.ypmed.2009.07.028. URL: https://linkinghub.elsevier.com/retrieve/pii/S0091743509004344.

[16] R. Aldred. “Built Environment Interventions to Increase Active Travel: a Critical Review and Discussion”. In: Current Environmental Health Reports 6.4 (Nov. 2019), pp. 309–315. DOI: 10.1007/s40572-019-00254-4. URL: https://link.springer.com/article/10.1007/s40572-019-00254-4.

[17] S. Kraus and N. Koch. “Provisional COVID-19 infrastructure induces large, rapid increases in cycling”. In: Proceedings of the National Academy of Sciences of the United States of America 118.15 (Apr. 2021). DOI: 10.1073/PNAS.2024399118/SUPPL\_FILE/PNAS.2024399118.SAPP.PDF.

[18] A. Buczyński. The Costs of Cycling Infrastructure. Tech. rep. European Cyclists’ Federation, 2021, pp. 1–8. URL: https://ecf.com/system/files/The%7B%5C_%7DCosts%7B%5C_%7DOF%7B%5C_%7DCycling%7B%5C_\%D7Infrastructure%7B%5C_%7DFactsheet.pdf.

[19] J. H. Haunert and B. Budig. “An algorithm for map matching given incomplete road data”. In: GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems (2012), pp. 510–513. DOI: 10.1145/2424321.2424402.
[20] E. A. Nadaraya. “On Estimating Regression”. In: *Theory of Probability & Its Applications* 9.1 (1964), pp. 141–142. DOI: 10.1137/1109020. URL: http://epubs.siam.org/doi/10.1137/1109020.

[21] G. S. Watson. “Smooth Regression Analysis”. In: *Sankhya: The Indian Journal of Statistics, Series A* 26.15 (1964), pp. 175–184.
Bikeability and the induced demand for cycling

Fosgerau, Lukawska, Paulsen and Rasmussen

October 7, 2022

1 Literature review

1.1 Perturbed utility

In a general perturbed utility model, a consumer chooses a consumption vector \( x \) from some budget set \( B \) that solves a utility maximization problem of the form

\[
\hat{x} = \arg \max_{x \in B} (a^T x - F(x)),
\]

i.e. where the utility function is a linear function “perturbed” by subtracting a convex function [1–3]. Perturbed utility models are firmly rooted in modern microeconomic theory and can be interpreted as representing a population of optimizing agents whose individual behavior is described by one of a wide range of models [4]. The additive random utility discrete choice model [5] belongs to the set of perturbed utility models in which the budget set \( B \) is the set of probability vectors.

[6] introduced the perturbed utility route choice (PURC) model. The model represents traveller behaviour as a utility maximizing flow across an entire network. The perturbed utility budget now requires that flow is conserved through all network nodes from origin to destination. It is a partial specification of behaviour as it does not specify exactly which route a given traveller will take but only the probabilities that are implied by the optimally chosen flow. A special perturbation function is specified that induces a tendency to distribute flow on more routes, while allowing optimal flow to be zero on most links in the network. [6] estimate and validate the model using 1,337,096 car trips in a large road network.

1.2 Bicycle route choice models

The bicycle route choice model literature has relied almost exclusively on path-based models. These are random utility discrete choice models, viewing the route choice as a discrete choice among a predefined set of path alternatives [7–19]. However, pre-defining the choice set is problematic as it leads to bias in the parameter estimates [20] and risks excluding the actually chosen alternatives. This is a problem in particular for bicycle route choice, since bicycle networks tend to be very fine-grained [11]. Researchers have been forced to discard almost half of the observed trips due to inadequate similarities of observed trips with any of the predefined alternatives [21].
The recursive logit model [22] and the nested recursive logit model [23] view the choice of a path as a Markov chain of link choices, where the traveler at each step anticipates the expected utility associated with reaching the next node. The recursive logit models incorporate the whole network and thereby avoid pre-defining choice sets. However, a downside of this feature is that the models distribute positive flow on all network links, whereas in reality most of the network will be unused for any given OD. [24] have applied recursive logit models to bicycle route choice, but experienced calculation times of 15 days (nested recursive logit) and 43 hours (recursive logit). [21] were not able to obtain identification with recursive logit models. An inherent limitation of the recursive logit models is that computation times are strongly affected by the size of the network. The network in [24] has 42,000 links, so it would most likely not be feasible to estimate a recursive logit model for the network with 420,000 links that we use in the present paper.

In this paper, we instead model bicycle route choices using the PURC model, which has a range of attractive features. The PURC model does not require choice sets to be pre-specified, it simply incorporates the complete network. In contrast to the recursive and nested recursive logit models, it leaves most of the network unused for any OD. The PURC model allows very fast estimation of link cost parameters using linear regression, even when taking a large number of link-level variables into account. It also makes it possible to estimate the model on much larger datasets (we estimate our model on more than 150,000 trips within a minute). The way we estimate the PURC model requires pre-processing of the data, which leads to some loss of data - nevertheless, we are able to retain 70\% of the observed route choices (see Section 2.4), which is more than has been the case with path-based approaches.

1.3 Bicycle infrastructure and the demand for bicycling

[25–27] review the empirical evidence on the effect of bicycle infrastructure on bicycle demand. The evidence supports that the infrastructure mostly does affect cycling levels, and that the effect is heterogeneous across infrastructure types [25]. [27] points out that studies are mostly just case-based, while [26] calls for studies that use individual-level data to assess the influence of the whole network on bicycle demand.

Going into the individual studies, one group reviewed by [28] and [29] carries out case-based before-and-after analyses of the demand effects of specific new bicycle infrastructure projects, using traffic count data or travel surveys [e.g., 30–35]. This kind of study is directly aimed at identifying causal effects of new infrastructure. They find demand increases up to 140\% on the specific new pieces of infrastructure. Due to the research design, they are, however, not able to distinguish between traffic actually induced by the new infrastructure and existing traffic that is attracted from elsewhere [29]. In the present paper, we are clearly to distinguish these two mechanisms. They are also less well suited for identifying the effects of the specific attributes of the new infrastructure, which the present paper finds to be very important.
A second group of studies relates bicycle demand to attributes characterizing entire networks at the macro-level [36–45]. The studies are cross-sectional, either at the city or country level [36–38, 40, 44] or at the level of areas within a region [39, 45, 41–43]. A general finding from these studies is that overall bicycle demand is positively correlated with the length of the bicycle network [36–39, 44, 41, 42], and negatively correlated with the share of large roads [41, 44]. Both findings are supported by our study, in which we furthermore find that the preference against large roads more than disappears when the roads are equipped with protected bicycle lanes.

[46] considers the effects of provisional pop-up bicycle lanes on travel demand during the COVID-19 pandemic in a cross-section of European cities. They find that the pop-up bike lanes induced an average increase in cycling of 41.6%, some of which may be new bicycle traffic while some may also just be diverted. The effect of pop-up lanes was lower in cities with a larger pre-existing bicycle network per capita. For comparison, we find in our counterfactual simulation which removes all bicycle lanes, that the number of bicycle trips decreases by 27.9%. The research design in [46] arguably allows identification of causal effects, but cannot account for network-wide effects or explicitly for the characteristics of the existing infrastructure.

A third group of studies relate the bicycle modal share at the level of origin-destination pairs to corresponding distances based on shortest paths [47, 48]. Their results broadly agree with ours, but in contrast to the present study, they cannot take the quality of the bicycle relevant network into account.

Finally, a fourth group of studies uses traditional transport models [49–51] or agent-based transport simulation [52] to model bicycle flows in a network. This allows them to analyze detailed counterfactual scenarios without having access to observed route choice data. [50, 52, 51] found an increase in bicycle use of 18-35% (Patna, India) and 4-9% (Copenhagen, Denmark), resulting from building/expanding the network of cycle superhighways. In our study, we find a similar effect in a counterfactual where the number of bicycle trips drops by 11.9% when removing the existing and planned cycle superhighways. [49] finds a generalized cost elasticity of demand of $-0.7$, where the present paper finds a higher elasticity in the range ($-1.63, -1.34$). Equipping all links with bicycle lanes leads demand to increase by 20% in [49], whereas we find demand to decrease by 28% when we remove all bicycle lanes.

In this paper, we estimate a combined bicycle route choice model and demand model. Our model is firmly rooted in data, incorporating a very large dataset comprising more than 150,000 observed bicycle trips across a large network and including an extensive set of explanatory variables. From the observed route choices, we back out a generalized cost measure that is used to predict bicycle demand across the network. This improves on the previous studies on a number of points. Specifically, we are able to incorporate all bicycle infrastructure in the network, and not just a few specific cases of new infrastructure. We are able to assess the effects of detailed and network-wide counterfactual changes to the bicycle network, distinguishing between new and diverted bicycle trips. Our generalized cost measure integrates all our observed infrastructure.
attributes across the network to the extent that they affect route choice. The model can thus take into account both the quality of bicycle infrastructure and where it is located.

2 Data and data processing

2.1 Network data

The network representation is based on Open Street Map (OSM\textsuperscript{1}) and includes the bicycle-relevant infrastructure, i.e., all network links where it is possible to ride or carry a bicycle, including the elements listed in Table 1. In the representation, bicycling in both directions is allowed on all network links, while keeping track of the direction. The resulting network representation of the Copenhagen Metropolitan Area contains a total of 420,973 directed links and 324,492 nodes.

We define infrastructure types by combining three infrastructure attributes: road type (based on OSM tags), road size (based on the number of car lanes), and type of bicycle infrastructure (whether it is present and if so, whether it is a protected or a painted bicycle lane). This creates 16 distinct infrastructure types, see Table 1.

The OSM network attributes have been enriched with information on land-use and elevation. The land-use information is obtained from an external GIS layer\textsuperscript{2}, and includes categories: green areas (including green restricted areas, parks, and forests), areas near water, industrial areas, open landscape areas, low-rise urban areas, and high-rise urban areas (merged with the city centre). For each directed link, the land use on the immediate right-hand side of the link is determined, tracking the length of each land-use category. The elevation gradient is computed with 10 meter splits of the network. Using overlay analysis, elevation information per 10 meters is attributed to each link. Based on this, the slope and difference in elevation is attained per 10 meters on each link, and the total vertical meters gained when the slope is greater than 3.5 % per direction per link.

2.2 Trajectory data

The data is collected in Greater Copenhagen (the study area framed in Figure 1) between 16\textsuperscript{th} September 2019 and 31\textsuperscript{st} May 2021 from 9,564 individuals using a Hövding head protection airbag helmet designed for cyclists\textsuperscript{2}. Positional data is collected passively among users that have given consent to share their data, and is transmitted to a database server through the users smartphone that is connected to the airbag helmet by Bluetooth. The dataset of observed trajectories contains 347,430 trips starting and/or ending in Greater Copenhagen covering a total of 939,711.8 kilometres traveled by bicycle. Each trip

\textsuperscript{1}https://www.openstreetmap.org
\textsuperscript{2}https://hovding.com
connects an origin-destination (OD) pair, represented by the starting point and the endpoint of the trip, respectively.

2.3 Map matching

Each of the observed trajectories has been map matched to the bicycle network using the hidden Markov algorithm proposed in [54]. The algorithm allows for off-road parts in the matched route, which is often necessary for bicycle trips as bicyclists do not always stick to formal roads and paths. However, our network has a high resolution, and we found that only 35 trips were matched with off-road segments. We discarded these trips from the subsequent analysis.

2.4 OD data and trip trimming

All trips shorter than 1 kilometer were discarded. Furthermore, we discarded circuitous routes that were more than $\frac{\pi}{7}$ times longer than the crow-fly distance. Trips with loops, where a part of the route was repeated, or where the same network node was visited twice, were also discarded. Finally, we discarded also trips where the map matching algorithm failed to match the entirety of the trip. The resulting dataset after the filtration steps consisted of 218,489 trips from 8,588 individuals covering a total of 762,791.8 kilometres.

Our estimator for the PURC model requires multiple observations in each OD pair. As common ODs are very rare in a large network, we follow [6] and trim the observed trips such that the trimmed trips share common ODs. Our algorithm selects first a set of origins and then a set of destinations. A trip is included in the estimation data if it passes first a selected origin and then a selected destination and only the part of the trip between the selected origin and destination is included.

More specifically, we include origins one by one, choosing in each step the origin that maximizes the total length of trips that it allows to include, while trimming the additional included trips to begin in that origin. After a list of origins have been compiled, we find in a similar way a list of destinations. The final output is a long list of origins and destinations. Our main results use 200 origins and 200 destinations.

In order to ensure that the generated origins and destinations are all found within Greater Copenhagen (see Figure 1), in this step we only consider trips that both start and end within Greater Copenhagen. This makes the dataset used for this task slightly smaller than the final estimation dataset, where we require only that either the origin or the destination is within Greater Copenhagen (see Section 2.2). The dataset used for finding origins and destinations consists of 208,410 trips (703,837.5 kilometres) across 8,456 individuals.

Having selected the list of origins and destinations, we identify, for each trip, the first origin and the last destination that are on the list. Only the trips that meet first an origin and then a destination from the list are included. Included trips are trimmed to begin and end at these points.
This process increases the likelihood of the included trips having origins and destinations in common with other included trips. Data is lost if the number of origins and destinations is small, which speaks for including many origins and destinations. However, the estimator that we use combines trips that are matched using the same OD pair into an observed average flow vector for that OD pair. Increasing the number of origins and destinations means that the observed average flow will be based on fewer matched trips per active OD pair, which implies more noise. It also means that there will be more unmatched trips, trips that are alone in using an OD pair, and these trips cannot be used for estimation. We choose the number of origins and destinations to balance these concerns.

At this stage, there are trips in the included data that are not matched to another trip with the same OD. Therefore we add a part to the algorithm, seeking to reduce the number of such trips. The part begins by identifying the longest unmatched trip. That trip is gradually shortened by trying combinations of later origins (from the list of candidates) and prior destinations (from the list of candidates), until it is found to travel between an origin and a destination that matches another trimmed trip, matched or unmatched. Trips that fail to find a match are discarded. The algorithm continues with the longest remaining unmatched trip until all unmatched trips have been either matched or discarded.

We choose the number of origins and destinations on which to base the trip trimming to maximize the number of origin-destination pairs that have at least ten observed trips (after recovering unmatched trips). It was found to be the case when using 200 origins and 200 destinations. As a check, we report also estimation results from the route choice model with 100 and 400 origins and destinations (Table 5). The parameter estimates are not very sensitive to this change, as we shall see in Section 33.2.3.

Table 2 summarizes the size of the data sets after the main steps of data processing.

In conclusion, we are able to retain 70% of the observed trips for the estimation. This is much higher than seen in traditional path-based route choice studies [55, 56]. Figure 1 shows heat maps of the trajectory data after the initial data cleaning (a), trips connecting candidate Os and Ds (b) and trimmed trips used for estimation. We observe that the trimmed trips preserve a good coverage of the network.

2.5 Computing predictions

The predicted flow for a trip starting in $o \in O$ and ending in $d \in D$ is the flow vector $\hat{x}^{od}$ that minimizes the cost

$$C(x) = \sum_{e \in E} l_e \left( c_e x_e + F(x_e) \right),$$

subject to the flow conservation constraint. The generalized cost association with this OD pair is $C(\hat{x}^{od})$.  

6
The edge cost rates \( c_e, e \in \mathcal{E} \) are found by multiplying the corresponding row in the link attribute matrix \( Z_e \) with the \( \hat{\beta} \) parameters estimates reported in Table 3, i.e. \( c_e = Z_e \hat{\beta} \). The minimization problems are solved using conic optimization in the software Mosek Fusion [57]. We define the average length between \( o \) and \( d \) corresponding to the predicted flows as \( \hat{l}^{od} = \sum_e \hat{x}^{od}_e l_e \). Finally, we obtain the predicted average generalized cost between \( o \) and \( d \), omitting the perturbation term, as \( \hat{c}^{od} = \sum_e c_e \hat{x}^{od}_e \).

3 Route choice model

3.1 Estimation results

Table 3 shows the estimated parameters for the preferred model specification, along with clustered heteroscedasticity consistent standard errors. To aid interpretation, the last column of the table shows the parameters divided by the parameter for the constant, such that the scaled parameters express the subjective cost rate in terms of metres travelled on the reference category road. We discuss the results in terms of the scaled parameters.

The reference category is residential roads without specific bicycle infrastructure in low-rise urban areas. The parameter for the constant thus represents the subjective cost of travelling one metre by bicycle on the reference category.

The next set of parameters measures the impact of various mutually exclusive infrastructure categories on the link cost rate. Links with stairs (intended for pedestrians) are classified as a separate category and incur a penalty of 76%. The cost rate is up to 27% larger for infrastructure types that are shared with pedestrians. The additional cost rate for ‘living streets’ is not significantly different from zero.\(^3\)

Compared to the reference, bicyclists have some preference against large roads (roads with at least two lanes in one direction, 11%), while the difference to medium roads (roads with at most one lane in each direction) is small and statistically insignificant.

Provision of dedicated bicycle infrastructure reduces the subjective cost of bicycling quite substantially. Cycleways (bicycle paths in own tracé) have 20% lower cost rate than the reference. On residential and medium roads, bicycle lanes, whether protected or just painted, reduce the cost rate by 14% and 22%, respectively. The type of bicycle lane matters for the large roads category: painted bicycle lanes have only a small and statistically insignificant effect, while protected bicycle lanes reduce the cost rate by 34%. It makes clear intuitive sense that the impact of bicycle lanes is larger the larger the road is, and that only protected lanes have impact on the largest roads where car traffic is heavier.

Provision of bicycle-friendly infrastructure thus has a substantial effect on route choice. As evident from Section Bicycle travel demand in the main text, this translates into a substantial effect on the volume of bicycle trips.

---
\(^3\)Small residential streets with parked cars and no bicycle infrastructure, where there may be children playing etc.
A number of routes are marketed as so-called bicycle superhighways. This is a label that is applied to high-quality, continuous bicycle routes, built to cater to commuter cyclists. The cost rate on the links of these routes is 12% lower than on similar links without the bicycle superhighway label. It is possible that the bicycle superhighway label and associated infrastructural changes are the cause of the reduction in subjective cost. It is, however, also possible that the bicycle superhighway label has just been attached to routes that were already attractive. To check this we have included a variable indicating routes that are proposed to become bicycle superhighways in the future. We find that the cost reduction associated with these links is almost exactly the same as the cost reduction found for the actual bicycle superhighway links. The model already accounts for a range of link characteristics including upgrades to the bicycle infrastructure that take place in the process of creating a bicycle superhighway. The attraction of the actual and proposed bicycle superhighways could therefore be due to route-level and not link-level features, perhaps the feature that these routes are high-quality, continuous bicycle routes [58, 59]. The similarity of the parameters suggests that it is not the labelling that makes the difference, but rather that it is routes that were already attractive that have been selected to receive the bicycle superhighway label.

The next set of parameters accounts for the land-use near cycleways. We treat cycleways separately, as they turned out to act differently from the other infrastructure types. The cost rate is much reduced for cycleways in industrial areas (48%) or green areas (53%). It makes intuitive sense that cycleways in green areas may be pleasant. Another potential explanation which also applies for industrial areas is the attractiveness of isolation from heavy traffic. The parameters for cycleways near water or open landscape are not statistically significant.

For the other infrastructure types, the subjective cost is lower near all other land-uses than low-rise urban areas, but the differences are not statistically significant for all land-use types. The largest cost reduction is found for links near water (26%).

The last set of parameters concerns some special link characteristics. The elevation gain variable measures the total elevation gain on a link that is 3.5 percent or more. It aggregates the vertical distance on the parts of the links where the slope is at least +3.5 percent. The scaled parameter is estimated to be 16.6, which means that an elevation gain of 0.05 metre per metre implies an increase in the cost rate of 83%, which seems reasonable. If the surface is gravel, the cost rate increases by 19%, while the parameter for cobblestones is small and not statistically significant. Finally, going against (car) traffic in one-way streets increases the cost rate by 59%.
3.2 Model validation

3.2.1 Comparison of observed and predicted flows by infrastructure type

Using the estimated parameters, we have computed the predicted flows for each OD. Table 4 compares the observed and predicted flows, showing the percentage of flow that occurs on each link type. For comparison, the first column of the table shows the corresponding shares weighted just by link lengths, which is what would be the result if traffic was distributed on the network at random. We find, supporting the model, that the observed and predicted flow shares are very similar. They are both very different from the network shares, especially since cyclists use roads with bicycle infrastructure much more than would have been the case if traffic was distributed randomly on the network.

3.2.2 Comparison of predicted and observed flows by OD pair

To compare the observed and predicted flow vectors for a given OD, we introduce a correlation measure as follows. Define for convenience

\[
E_l(x) = \sum_{e \in E} \frac{l_e}{\sum_{e' \in E} l_{e'}} x_e.
\]

We compute then the correlation between the observed and the predicted flow on a random km of road (where \(\circ\) is the Hadamard product):

\[
\rho_{\text{od}} = \rho(x^{\text{od}}, \hat{x}^{\text{od}}) = \frac{E_l(x^{\text{od}} \circ \hat{x}^{\text{od}}) - E_l(x^{\text{od}})E_l(\hat{x}^{\text{od}})}{\sqrt{E_l(x^{\text{od}} \circ x^{\text{od}}) - (E_l(x^{\text{od}}))^2}\sqrt{E_l(\hat{x}^{\text{od}} \circ \hat{x}^{\text{od}}) - (E_l(\hat{x}^{\text{od}}))^2}}.
\]

The correlation bounded between \(-1\) and \(1\), with \(1\) corresponding to perfect correlation.

We have computed the correlation for every OD combination in the data. Figure 2 plots the correlation against the number of observed trips in each OD combination and against the expected route length in each OD combination. We find that the mean correlation ranges from more than 0.5 to more than 0.75. The correlation increases with the number of observed trips, which makes sense since sampling noise would cause the observed flow to differ from the predicted and hence reduce the correlation. The correlation decreases with route length, which may also be due to sampling noise since the number of observed trips decreases with trip length.

3.2.3 Changing the number of ODs

As mentioned in Section 2.4, it is necessary to choose the number of Os and Ds prior to the estimation. As a robustness check, we have estimated the model using 100 and 400 Os and Ds in addition to the 200 used for the main result. Table 5 shows the estimates, which are broadly similar across models.
4 Gravity model

4.1 Specification of the demand function

The gravity model assumes that the demand is driven by the cost \( c_{od} \) from the route choice model and not just by the OD distance. To test whether this is the case, we split the cost into a length component and a residual quality component that is orthogonal to length. In this decomposition, we omit the perturbation term as this term is expected to increase with reduced OD distance due to the fact that there are fewer relevant routes for shorter trips. For the test we therefore compute the length component as the predicted value in a linear regression of the cost \( c_{od} \) excluding the perturbation term against the predicted trip length. The quality component is the residual from this regression. We have then estimated the gravity model in (3) with two demand functions, one for length and one for quality. As before, both are specified as piece-wise linear.

\[
\ln E[Y_{od}] = D_1 \text{length}^{od} + D_2 \text{quality}^{od} + \delta + \eta_o + \gamma_d
\]  

Figure 3 shows in blue the estimated demand function from the original model (2) in the main text along with the demand functions from the model (3), where the green dotted curve is the influence of length on demand and the red dotted curve is the influence of quality on demand. We observe that both curves have about the same negative slope. We conclude, in support of our model, that quality has about the same influence on demand as length.

4.2 Demand elasticity

The elasticity of demand is the relative change in demand per relative change in cost. Thus an elasticity of -1 implies that a 10% increase in cost leads to a 10% decrease in demand. Figure 4 shows the elasticities calculated along the estimated demand curve. The result indicates that the elasticity decreases about monotonically with the cost.

5 Counterfactuals

5.1 Method

Using the average predicted generalized OD costs \( \hat{c}_{od} \), the estimated function \( D \) and values of \( \delta, \eta_o, \) and \( \gamma_d \), we apply the estimated gravity model to compute \( \hat{Y}_{od} \), the predicted number of trips between any \( o \in O \) and \( d \in D \). By weighting each OD flow prediction with the corresponding average predicted average length between \( o \) and \( d \), we find the total sum of predicted ridden kilometers in our sample. This number corresponds to the size of our sample of bicycle trips. We scale it to the actual annual number of ridden kilometers in the study area. We can estimate this roughly to be 1,026 mio km, computed using [60] as the total kilometers in Frederiksberg and Copenhagen municipalities, and 50%
of the kilometers in the suburbs. In this way we find that each predicted trip represents \( \zeta = 1,926 \) annual trips.

When simulating a counterfactual \( s \), we first adjust the link attribute matrix \( Z \) according to the counterfactual scenario, and obtain a new link attribute matrix \( Z^* \), and corresponding link-specific costs rates \( \tilde{c}_e^* = Z^* \beta \) (visualized for our three counterfactuals in Figure 5). Using these new link cost rates, we recompute the cost minimizing flows, which we denote \( \hat{x}^{od,s} \). We compute also the average predicted generalized OD costs \( \hat{c}^{od,s} = \sum_{e \in E} \tilde{c}_e \hat{x}^{od,s}_e \). We keep the estimated \( \delta, \eta_o, \) and \( \gamma_d \) from the base scenario, and scale each trip with the same factor \( \zeta \) as in the base case.

The relative cost increase in counterfactual \( s \) compared to the base is found as

\[
\frac{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od} \frac{\tilde{c}^{od,s} - \tilde{c}^{od}}{\tilde{c}^{od}}}{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od}}. \tag{4}
\]

The relative decrease in the number of trips in counterfactual \( s \) relative to the base scenario is

\[
1 - \frac{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od,s}}{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od}}. \tag{5}
\]

The corresponding measure for the volume of trip kilometers is

\[
1 - \frac{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od,s} \hat{l}^{od,s}}{\sum_{o \in O} \sum_{d \in D} \hat{Y}^{od} \hat{l}^{od}}. \tag{6}
\]

In order to compute the consumer surplus [61], we first convert from generalized cost to length units by dividing the generalized cost with the constant term in \( \beta \) (0.4548). Second, we use the average speed in our sample (14.84 km/h) to convert the lengths into travel time units. The change in travel time then becomes

\[
\Delta T^{od,s} = \frac{1}{14.84} \frac{\hat{c}^{od,s} - \hat{c}^{od}}{0.4548}. \tag{7}
\]

The value of time for cyclists in Denmark is 16.13€ per hour according to the official guidelines [62]. Combining the change in travel time, the value of time, and the change in demand, we apply the rule of a half to compute the change in consumer surplus for each counterfactual scenario

\[
16.13 \cdot \zeta \cdot \sum_{o \in O} \sum_{d \in D} \Delta T^{od,s} \frac{\hat{y}^{od} + \hat{y}^{od,s}}{2}. \tag{8}
\]
Health and accident benefits are computed multiplying the combined unit price of 0.91€ (1.17€ for health and -0.26€ for accidents [62]) by the scaled change in ridden kilometer to get the monetary annual benefit

\[ 0.91 \cdot \zeta \cdot \sum_{o \in O} \sum_{d \in D} \hat{Y}_{od,s} \hat{f}_{od,s} - \hat{Y}_{od} \hat{f}_{od}. \]  

The societal benefit due to cyclists’ travel time, health and accidents is the sum of the consumer surplus benefits and the health/accident benefits. The societal benefit per changed network length is found by dividing with the sum of (directed) link lengths of the links that was changed to create the counterfactual s. However, for the third counterfactual on cycle superhighways, the corresponding route distances were used instead of the network length, since the expense related to this counterfactual occurs at the route level.
(a) After initial data cleaning  
(b) Trips connecting candidate O's and D's  
(c) Trimmed trips used for estimation

Figure 1: Heat map of anonymized GPS traces after different filtration phases. Study area shown in blue. (a): All GPS points after initial filtration; (b): GPS points of the trips used for model estimation (untrimmed filtered trips); (c): GPS points of the trips used for model estimation (trimmed filtered trips). For visualization, trips have been anonymized by removing a random number of points at their start and end.
Figure 2: Correlation between predicted and observed flow as defined in (2) plotted against the number of observed trips per OD (panel a) and the expected route length per OD (panel b). Black line indicates mean correlation, red lines indicate corresponding binned deciles.
Figure 3: Estimated piecewise linear specifications of $D$ (from (2) in the main text) for three different explanatory variables. Blue: Demand function $D$ from the original model ( (2) in the main text). Green & Orange: Demand functions $D_1$ and $D_2$ from the model in (2), respectively.
Figure 4: Demand elasticities ($\text{elast}(\hat{c}^{od})$) calculated from the estimated demand function using $\text{elast}(\hat{c}^{od}) = D'(\hat{c}^{od})\hat{c}^{od}$. The demand elasticity at a point $\hat{c}^{od}$ is the relative change in demand per relative change in cost.
Figure 5: The estimated cost rates of links in the three counterfactual scenarios. Grey marks the reference case, i.e. residential roads with no bicycle infrastructure in low-rise urban areas, scaled to a value of 1. Shades of green correspond to increasingly more attractive links, and shades of orange and red correspond to increasingly less attractive links.
Table 1: Network attributes related to infrastructure type and their associated OSM tags

| Infrastructure                                      | Description                                                                 |
|-----------------------------------------------------|-----------------------------------------------------------------------------|
| Residential roads w/o bicycle infrastructure        | Residential roads: Roads with OSM tag **highway**='residential'             |
| Residential roads w/ painted bicycle lanes          |                                                                             |
| Residential roads w/ protected bicycle lanes        |                                                                             |
| Medium roads w/o bicycle infrastructure             | Medium roads: Roads with OSM tags **highway** in **{primary, secondary, tertiary, unclassified}**, that have at most one car lane per direction. |
| Medium roads w/ painted bicycle lanes               |                                                                             |
| Medium roads w/ protected bicycle lanes             |                                                                             |
| Large roads w/o bicycle infrastructure              | Large roads: Roads with OSM tags **highway** in **{primary, secondary, tertiary, unclassified}**, that have at least two car lanes in at least one direction. |
| Large roads w/ painted bicycle lanes                |                                                                             |
| Large roads w/ protected bicycle lanes              |                                                                             |
| Cycleways                                           | OSM tag **highway**='cycleway'                                             |
| Footways                                            | OSM tag **highway**='footway'                                              |
| Living streets                                      | OSM tag **highway**='living_street'                                        |
| Shared paths                                        | OSM tags **highway** in **{path, track, service}**                          |
| Pedestrian zones                                    | OSM tag **highway**='pedestrian'                                           |
| Stairs                                              | OSM tag **highway**='steps'                                                |
Table 2: Summary of the dataset size after different filtration subprocesses. Subprocesses correspond to panels (a), (b), and (c) in Figure 1, respectively.

| Subprocess                                      | Trips  | Users | Kilometres   |
|-------------------------------------------------|--------|-------|--------------|
| After initial data cleaning                     | 218,489| 8,588 | 762,791.8    |
| Trips connecting candidate Os and Ds            | 152,323| 7,672 | 614,909.7    |
| Trimmed trips used for estimation               | 152,323| 7,672 | 417,358.0    |
Table 3: The estimated parameters for the bicycle route choice model in the main text (200 origins and 200 destinations). Standard errors are clustered (per OD) and heteroscedasticity-consistent. Scaled values are scaled such that the parameter ‘Constant’ has a value of 1.

| Parameter                                      | Coef.  | Std. err. | P-val. | Scaled |
|------------------------------------------------|--------|-----------|--------|--------|
| **Constant**                                   | 0.455  | 0.027     | ***    | 1      |
| **Infrastructure**                             |        |           |        |        |
| Stairs                                         | 0.345  | 0.096     | ***    | 0.758  |
| Pedestrian zones                                | 0.121  | 0.025     | ***    | 0.266  |
| Footways                                        | 0.090  | 0.012     | ***    | 0.198  |
| Shared paths                                    | 0.026  | 0.016     |        | 0.057  |
| Living streets                                  | 0.053  | 0.039     |        | 0.116  |
| Cycleways                                       | -0.090 | 0.037     | *      | -0.198 |
| Residential roads                               |        |           |        |        |
| No bicycle infrastructure                       | —      | —         | —      | —      |
| W/ bicycle infrastructure                       | -0.065 | 0.019     | ***    | -0.143 |
| Medium roads                                    | 0.004  | 0.018     |        | 0.009  |
| No bicycle infrastructure                       | —      | —         | —      | —      |
| W/ bicycle infrastructure                       | -0.101 | 0.018     | ***    | -0.222 |
| Large roads                                     | 0.050  | 0.015     | ***    | 0.110  |
| No bicycle infrastructure                       | —      | —         | —      | —      |
| W/ painted bicycle lanes                        | -0.014 | 0.024     |        | -0.031 |
| W/ protected bicycle lanes                      | -0.155 | 0.016     | ***    | -0.341 |
| **Bicycle route classification**                |        |           |        |        |
| No classification                               | —      | —         | —      | —      |
| Cycle superhighway                              | -0.056 | 0.014     | ***    | -0.123 |
| Proposed cycle superhighway                    | -0.055 | 0.012     | ***    | -0.121 |
| **Land-use, cycleways**                         |        |           |        |        |
| High-rise urban areas                           | -0.119 | 0.041     | **     | -0.262 |
| Low-rise urban areas                            | —      | —         | —      | —      |
| Industrial areas                                | -0.220 | 0.052     | ***    | -0.484 |
| Green areas                                     | -0.243 | 0.054     | ***    | -0.534 |
| Areas near water                                | -0.043 | 0.052     |        | -0.095 |
| Open landscape                                  | 0.036  | 0.122     |        | 0.079  |
| **Land-use, other infrastructure**              |        |           |        |        |
| High-rise urban areas                           | -0.088 | 0.024     | ***    | -0.193 |
| Low-rise urban areas                            | —      | —         | —      | —      |
| Industrial areas                                | -0.071 | 0.038     |        | -0.156 |
| Green areas                                     | -0.055 | 0.037     |        | -0.121 |
| Areas near water                                | -0.120 | 0.030     | ***    | -0.264 |
| Open landscape                                  | -0.086 | 0.095     |        | -0.180 |
| **Elevation gain, > 35 m/km**                   | 7.553  | 3.617     | *      | 16.600 |
| **Surface type**                                |        |           |        |        |
| Asphalt                                        | —      | —         | —      | —      |
| Cobblestones                                    | 0.014  | 0.019     |        | 0.031  |
| Gravel                                         | 0.086  | 0.015     | ***    | 0.189  |
| **Wrong way**                                   | 0.266  | 0.007     | ***    | 0.585  |

***: P-value ≤ 0.001, **: P-value ≤ 0.01, *: P-value ≤ 0.05.
Table 4: Distribution of length shares for various link characteristics. Network indicates the share of the network, Observed use indicates the share of the observed trips, and Predicted use indicates the share of the predicted flow.

|                                      | Network [%] | Observed use [%] | Predicted use [%] |
|--------------------------------------|-------------|------------------|-------------------|
| **Constant**                         | 100.00      | 100.00           | 100.00            |
| **Infrastructure**                   |             |                  |                   |
| Stairs                               | 0.12        | 0.07             | 0.07              |
| Pedestrian zones                     | 0.19        | 0.72             | 0.34              |
| Footways                             | 8.24        | 6.80             | 2.43              |
| Shared paths                         | 38.54       | 6.21             | 4.95              |
| Living streets                       | 0.55        | 0.34             | 0.35              |
| Cycleways                            | 8.95        | 10.02            | 15.67             |
| Residential roads                    | 22.88       | 13.50            | 16.57             |
| No bicycle infrastructure            | 22.61       | 10.22            | 12.74             |
| W/ bicycle infrastructure            | 22.61       | 10.22            | 12.74             |
| Medium roads                         | 18.37       | 41.06            | 40.93             |
| No bicycle infrastructure            | 14.34       | 3.07             | 1.90              |
| W/ bicycle infrastructure            | 4.03        | 37.99            | 39.03             |
| Large roads                          | 2.16        | 21.27            | 18.68             |
| No bicycle infrastructure            | 2.16        | 21.27            | 18.68             |
| W/ bicycle lanes                     | 1.00        | 3.63             | 0.75              |
| W/ protected bicycle tracks          | 1.08        | 16.32            | 16.74             |
| **Bicycle route classification**     |             |                  |                   |
| No classification                    | 92.01       | 39.39            | 39.51             |
| Cycle superhighway                   | 2.33        | 23.50            | 23.19             |
| Proposed cycle superhighway          | 5.66        | 37.11            | 37.31             |
| **Land-use, cycleways**              |             |                  |                   |
| High-rise urban areas                | 0.58        | 4.05             | 5.27              |
| Low-rise urban areas                 | 3.49        | 1.20             | 1.37              |
| Industrial areas                     | 0.64        | 1.45             | 2.62              |
| Green areas                          | 1.09        | 2.09             | 4.86              |
| Areas near water                     | 0.11        | 0.76             | 0.87              |
| Open landscape                       | 3.04        | 0.48             | 0.69              |
| **Land-use, other infrastructure**   |             |                  |                   |
| High-rise urban areas                | 8.25        | 60.07            | 56.60             |
| Low-rise urban areas                 | 36.54       | 9.47             | 10.25             |
| Industrial areas                     | 7.92        | 5.59             | 5.55              |
| Green areas                          | 16.05       | 8.87             | 7.41              |
| Areas near water                     | 1.02        | 4.93             | 3.24              |
| Open landscape                       | 21.27       | 1.05             | 1.28              |
| **Elevation gain, > 35 m/km**        | 0.16        | 0.01             | 0.01              |
| **Surface type**                     |             |                  |                   |
| Asphalt                              | 86.25       | 95.85            | 97.62             |
| Cobblestones                         | 0.96        | 2.06             | 1.72              |
| Gravel                               | 12.79       | 2.09             | 0.65              |
| **Wrong way**                        | 3.54        | 9.15             | 2.17              |
Table 5: The estimated parameters for the bicycle route choice model with 100, 200 and 400 Os and Ds. Standard errors are robust. Scaled values are scaled such that the parameter 'Constant' has a value of 1.

| Number of Os and Ds | 100 | 200 | 400 |
|---------------------|-----|-----|-----|
|                      | Coef. | Std. err. | P-val. | Scaled Coef. | Std. err. | P-val. | Scaled Coef. | Std. err. | P-val. | Scaled |
| Infrastructure       |      |       |      |            |           |       |            |           |       |       |
| Stairs               | 0.257 | 0.035 |    * | 0.521 | 0.345 | 0.086 |    *** | 0.758 | 0.555 | 0.090 |    *** | 1.542 |
| Footways             | 0.072 | 0.014 |    *** | 0.148 | 0.090 | 0.012 |    *** | 0.182 | 0.099 | 0.012 |    *** | 0.275 |
| Skewed path          | 0.041 | 0.019 |    * | 0.083 | 0.020 | 0.026 |    | 0.040 | 0.019 | 0.016 |    * | 0.018 |
| Living streets       | 0.049 | 0.054 | 0.089 | 0.053 | 0.039 | 0.118 | 0.075 | 0.039 |    * | 0.021 |
| Cycleways            | -0.105 | 0.037 |    ** | -0.211 | 0.090 | 0.037 |    | -0.198 | 0.050 | 0.033 |    -0.164 |
| Residential roads    |      |       |      |            |           |       |            |           |       |       |
| No bicycle infrastructure |      |       |      |            |           |       |            |           |       |       |
| W/ bicycle infrastructure |      |       |      |            |           |       |            |           |       |       |
| Medium roads         | 0.028 | 0.021 | 0.057 | 0.004 | 0.028 | 0.059 | 0.033 | 0.018 | 0.0092 |
| No bicycle infrastructure |      |       |      |            |           |       |            |           |       |       |
| Large roads          | 0.005 | 0.018 | 0.021 | 0.008 | 0.015 | 0.020 | 0.010 | 0.0015 |    *** | 0.167 |
| No bicycle infrastructure |      |       |      |            |           |       |            |           |       |       |
| Residential roads    |      |       |      |            |           |       |            |           |       |       |
| W/ painted bicycle lanes |      |       |      |            |           |       |            |           |       |       |
| W/ protected bicycle lanes |      |       |      |            |           |       |            |           |       |       |
| Bicycle route classification |      |       |      |            |           |       |            |           |       |       |
| No classification    |      |       |      |            |           |       |            |           |       |       |
| Cycle superhighway   | -0.042 | 0.036 |    ** | -0.065 | -0.056 | 0.014 |    *** | -0.123 | -0.040 | 0.014 |    ** | -0.111 |
| Proposed cycle superhighway | -0.038 | 0.015 | -0.041 | 0.005 | 0.012 | -0.045 | 0.036 | 0.012 |    ** | -0.106 |
| Land-use, cycleways  |      |       |      |            |           |       |            |           |       |       |
| High-rise urban areas | 0.146 | 0.044 |    *** | -0.216 | 0.119 | 0.041 |    * | -0.202 | 0.100 | 0.038 |    * | -0.278 |
| Low-rise urban areas  |      |       |      |            |           |       |            |           |       |       |
| Industrial areas     | 0.279 | 0.080 |    *** | -0.564 | 0.228 | 0.052 |    | -0.489 | 0.193 | 0.049 |    ** | -0.588 |
| Green areas          | 0.220 | 0.060 |    *** | -0.507 | 0.283 | 0.054 |    *** | -0.524 | 0.199 | 0.053 |    *** | -0.530 |
| Area near water      | 0.105 | 0.057 | 0.213 | 0.043 | 0.052 | -0.095 | 0.040 | 0.048 | 0.111 |
| Open landscape       | 0.131 | 0.105 | 0.208 | 0.090 | 0.122 | 0.079 | 0.045 | 0.082 | 0.125 |
| Land-use, other infrastructure |      |       |      |            |           |       |            |           |       |       |
| High-rise urban areas | 0.118 | 0.031 |    *** | -0.219 | 0.089 | 0.024 |    *** | -0.183 | 0.057 | 0.025 |    * | -0.138 |
| Low-rise urban areas  |      |       |      |            |           |       |            |           |       |       |
| Industrial areas     | 0.080 | 0.051 | -0.140 | 0.071 | 0.028 | -0.128 | 0.066 | 0.029 | -0.183 |
| Green areas          | 0.052 | 0.046 | -0.105 | 0.053 | 0.037 | -0.123 | 0.041 | 0.040 | -0.114 |
| Area near water      | 0.102 | 0.039 |    *** | -0.120 | 0.120 | 0.030 |    *** | -0.264 | 0.081 | 0.031 |    * | -0.231 |
| Open landscape       | 0.130 | 0.127 | 0.323 | 0.086 | 0.095 | -0.149 | 0.036 | 0.078 | -0.180 |
| Elevation gain, > 35 m/km | 4.141 | 6.142 | 8.100 | 7.531 | 6.317 | * | 16.009 | 10.532 | 3.194 | ** | 29.101 |
| Surface type         |      |       |      |            |           |       |            |           |       |       |
| Asphalt              |      |       |      |            |           |       |            |           |       |       |
| Cobblestone          | 0.019 | 0.025 | 0.039 | 0.014 | 0.049 | 0.031 | 0.024 | 0.016 | 0.007 |
| Grass                | 0.097 | 0.036 |    *** | 0.197 | 0.080 | 0.035 |    *** | 0.189 | 0.096 | 0.015 |    *** | 0.207 |
| Wrong way            | 0.233 | 0.040 |    *** | 0.512 | 0.266 | 0.067 |    *** | 0.534 | 0.251 | 0.067 |    *** | 0.507 |

* P-value ≤ 0.05, ** P-value ≤ 0.01, ***: P-value ≤ 0.001.
References

[1] Mogens Fosgerau and Daniel L. McFadden. A theory of the perturbed consumer with general budgets. *NBER Working Paper*, pages 1–27, 3 2012.

[2] Drew Fudenberg, Ryota Iijima, and Tomasz Strzalecki. Stochastic Choice and Revealed Perturbed Utility. *Econometrica*, 83(6):2371–2409, 2015.

[3] Roy Allen and John Rehbeck. Identification With Additively Separable Heterogeneity. *Econometrica*, 87(3):1021–1054, 5 2019.

[4] Roy Allen and John Rehbeck. Revealed Stochastic Choice with Attributes. 2019.

[5] Daniel McFadden. Econometric Models of Probabilistic Choice. In C. Manski and D. McFadden, editors, *Structural Analysis of Discrete Data with Econometric Applications*, pages 198–272. MIT Press, Cambridge, MA, USA, 1981.

[6] Mogens Fosgerau, Mads Paulsen, and Thomas Kjaer Rasmussen. A perturbed utility route choice model. *Transportation Research Part C*, 136(103514), 2022.

[7] Gianluca Menghini, Nelson Carrasco, Nadine Schüssler, and Kay W Axhausen. Route choice of cyclists in zurich. *Transportation Research Part A: Policy and Practice*, 44(9):754–765, 2010.

[8] Jeffrey Hood, Elizabeth Sall, and Billy Charlton. A gps-based bicycle route choice model for san francisco, california. *Transportation letters*, 3(1):63–75, 2011.

[9] Joseph Broach, Jennifer Dill, and John Gliebe. Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, 46(10):1730–1740, 2012.

[10] Ranjit Khatri, Christopher R Cherry, Shashi S Nambisan, and Lee D Han. Modeling route choice of utilitarian bikeshare users with gps data. *Transportation Research Record*, 2587(1):141–149, 2016.

[11] Danique Ton, Oded Cats, Dorine Duives, and Serge Hoogendoorn. How do people cycle in amsterdaml, netherlands?: Estimating cyclists’ route choice determinants with gps data from an urban areaa. *Transportation research record*, 2662(1):75–82, 2017.

[12] Jeffrey M Casello and Vladimir Usyukov. Modeling cyclists’ route choice based on gps data. *Transportation Research Record*, 2430(1):155–161, 2014.

[13] Peng Chen, Qing Shen, and Suzanne Childress. A gps data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*, 12(3):218–231, 2018.
[14] Muhammad Ghanayim and Shlomo Bekhor. Modelling bicycle route choice using data from a GPS-assisted household survey. *European Journal of Transport and Infrastructure Research*, 18(2):158–177, 2018.

[15] Carlo Giacomo Prato, Katrín Halldórsdóttir, and Otto Anker Nielsen. Evaluation of land-use and transport network effects on cyclists’ route choices in the Copenhagen Region in value-of-distance space. *International Journal of Sustainable Transportation*, 12(10):770–781, 2018.

[16] Anae Sobhani, Hamzeh Alizadeh Aliabadi, and Bilal Farooq. Metropolis-hasting based expanded path size logit model for cyclists’ route choice using gps data. *International Journal of Transportation Science and Technology*, 8(2):161–175, 2019.

[17] Dillon T Fitch and Susan L Handy. Road environments and bicyclist route choice: The cases of davis and san francisco, ca. *Journal of Transport Geography*, 85:102705, 2020.

[18] Nitesh R Shah and Christopher R Cherry. Different safety awareness and route choice between frequent and infrequent bicyclists: Findings from revealed preference study using bikeshare data. *Transportation Research Record*, 2675(11):269–279, 2021.

[19] Shin-Hyung Cho and DongHwa Shin. Estimation of route choice behaviors of bike-sharing users as first-and last-mile trips for introduction of mobility-as-a-service (maas). *KSCE Journal of Civil Engineering*, pages 1–12, 2022.

[20] Emma Frejinger, Michel Bierlaire, and Moshe Ben-Akiva. Sampling of alternatives for route choice modeling. *Transportation Research Part B: Methodological*, 43(10):984–994, 2009.

[21] Thomas Koch and Elenna Dugundji. Limitations of recursive logit for inverse reinforcement learning of bicycle route choice behavior in amsterdam. *Procedia Computer Science*, 184:492–499, 2021.

[22] Mogens Fosgerau and Kenneth A. Small. Hypercongestion in downtown metropolis. *Journal of Urban Economics*, 76(1):122–134, 2013.

[23] Tien Mai, Mogens Fosgerau, and Emma Frejinger. A nested recursive logit model for route choice analysis. *Transportation Research Part B: Methodological*, 75:100–112, 2015.

[24] Maëlle Zimmermann, Tien Mai, and Emma Frejinger. Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies*, 75:183–196, 2017.

[25] John Pucher, Jennifer Dill, and Susan Handy. Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, 50:S106–S125, jan 2010.

24
[26] Ralph Buehler and Jennifer Dill. Bikeway networks: A review of effects on cycling. *Transport Reviews*, 36(1):9–27, 2016.

[27] Rachel Aldred. Built Environment Interventions to Increase Active Travel: a Critical Review and Discussion. *Current Environmental Health Reports*, 6(4):309–315, Nov 2019.

[28] Kees van Goeverden, Thomas Sick Nielsen, Henrik Harder, and Rob van Nes. Interventions in Bicycle Infrastructure, Lessons from Dutch and Danish Cases. *Transportation Research Procedia*, 10:403–412, 2015.

[29] Famke J. M. Mölenberg, Jenna Panter, Alex Burdorf, and Frank J. van Lenthe. A systematic review of the effect of infrastructural interventions to promote cycling: strengthening causal inference from observational data. *International Journal of Behavioral Nutrition and Physical Activity*, 16(1):93, Dec 2019.

[30] Dafna Merom, Adrian Bauman, Philip Vita, and Glenn Close. An environmental intervention to promote walking and cycling—the impact of a newly constructed Rail Trail in Western Sydney. *Preventive Medicine*, 36(2):235–242, Feb 2003.

[31] Kelly R Evenson, Amy H Herring, and Sara L Huston. Evaluating change in physical activity with the building of a multi-use trail. *American Journal of Preventive Medicine*, 28(2):177–185, Feb 2005.

[32] Kevin J. Krizek, Gary Barnes, and Kristin Thompson. Analyzing the Effect of Bicycle Facilities on Commute Mode Share over Time. *Journal of Urban Planning and Development*, 135(2):66–73, 5 2009.

[33] Anna Goodman, Shannon Sahlqvist, and David Ogilvie. New Walking and Cycling Routes and Increased Physical Activity: One- and 2-Year Findings From the UK iConnect Study. *American Journal of Public Health*, 104(9):e38–e46, Sep 2014.

[34] Hans Skov-Petersen, Jette Bredahl Jacobsen, Suzanne Elizabeth Vedel, Sick Nielsen Thomas Alexander, and Simon Rask. Effects of upgrading to cycle highways—an analysis of demand induction, use patterns and satisfaction before and after. *Journal of Transport Geography*, 64:203–210, 2017.

[35] Jinhyun Hong, David Philip McArthur, and Mark Livingston. The evaluation of large cycling infrastructure investments in glasgow using crowd-sourced cycle data. *Transportation*, 47(6):2859–2872, 2020.

[36] Arthur C. Nelson and David Allen. If You Build Them, Commuters Will Use Them: Association Between Bicycle Facilities and Bicycle Commuting. *Transportation Research Record: Journal of the Transportation Research Board*, 1578(1):79–83, Jan 1997.
[37] Jennifer Dill and Theresa Carr. Bicycle Commuting and Facilities in Major U.S. Cities: If You Build Them, Commuters Will Use Them. Transportation Research Record, (1828):116–123, 2003.

[38] John Pucher and Ralph Buehler. Cycling trends & policies in Canadian cities. World Transport Policy & Practice, 11(1):43–61, 2005.

[39] John Parkin, Mark Wardman, and Matthew Page. Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation, 35(1):93–109, nov 2008.

[40] Jessica E Schoner and David M Levinson. The missing link: bicycle infrastructure networks and ridership in 74 us cities. Transportation, 41(6):1187–1204, 2014.

[41] Ahmed Osama, Tarek Sayed, and Alexander Y Bigazzi. Models for estimating zone-level bike kilometers traveled using bike network, land use, and road facility variables. Transportation Research Part A: Policy and Practice, 96:14–28, 2017.

[42] Thomas Alexander Sick Nielsen and Hans Skov-Petersen. Bikeability – Urban structures supporting cycling. Effects of local, urban and regional scale urban form factors on cycling from home and workplace locations in Denmark. Journal of Transport Geography, 69:36–44, 2018.

[43] Sherry Ryan, Ana Garate, Diane Foote, and Diane Foote. A Micro-Scale Analysis of Cycling Demand, Safety, and Network Quality. Mineta Transportation Institute Publications, 9 2020.

[44] Grégory Vandenbulcke, Claire Dujardin, Isabelle Thomas, Bas de Geus, Bart Degraeuwe, Romain Meeusen, and Luc Int Panis. Cycle commuting in Belgium: Spatial determinants and ‘re-cycling’ strategies. Transportation Research Part A: Policy and Practice, 45(2):118–137, 2 2011.

[45] Daniel J Fagnant and Kara Kockelman. A direct-demand model for bicycle counts: the impacts of level of service and other factors. Environment and Planning B: Planning and Design, 43(1):93–107, 2016.

[46] Sebastian Kraus and Nicolas Koch. Provisional COVID-19 infrastructure induces large, rapid increases in cycling. Proceedings of the National Academy of Sciences of the United States of America, 118(15), 4 2021.

[47] Michael Iacono, Kevin Krizek, and Ahmed El-Geneidy. Access to destinations: How close is close enough? estimating accurate distance decay functions for multiple modes and different purposes. Technical report, Hubert H. Humphrey Institute of Public Affairs, University of Minnesota, 2008.
[48] Robin Lovelace, Anna Goodman, Rachel Aldred, Nikolai Berkoff, Ali Abbas, and James Woodcock. The propensity to cycle tool: An open source online system for sustainable transport planning. *Journal of Transport and Land Use*, 10(1):505–528, 2017.

[49] Chengxi Liu, Andreas Tapani, Ida Kristoffersson, Clas Rydergren, and Daniel Jonsson. Development of a large-scale transport model with focus on cycling. *Transportation Research Part A: Policy and Practice*, 134:164–183, 2020.

[50] Martin Hallberg, Thomas Kjaer Rasmussen, and Jeppe Rich. Modelling the impact of cycle superhighways and electric bicycles. *Transportation Research Part A: Policy and Practice*, 149:397–418, 2021.

[51] Jeppe Rich, Anders Fjendbo Jensen, Ninette Pilegaard, and Martin Hallberg. Cost-benefit of bicycle infrastructure with e-bikes and cycle superhighways. *Case Studies on Transport Policy*, 9(2):608–615, 2021.

[52] Amit Agarwal, Dominik Ziemke, and Kai Nagel. Bicycle superhighway: An environmentally sustainable policy for urban transport. *Transportation Research Part A: Policy and Practice*, 137:519–540, Jul 2020.

[53] FOT-Kort10. FOT-kort10.

[54] Jan Henrik Haunert and Benedikt Budig. An algorithm for map matching given incomplete road data. *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, pages 510–513, 2012.

[55] Danique Ton, Dorine Duives, Oded Cats, and Serge Hoogendoorn. Evaluating a data-driven approach for choice set identification using GPS bicycle route choice data from Amsterdam. *Travel Behaviour and Society*, 13:105–117, 10 2018.

[56] Thomas Koch, Luk Knapen, and Elenna Dugundji. Path complexity and bicyclist route choice set quality assessment. *Personal and Ubiquitous Computing*, 25(1):63–75, Feb 2021.

[57] MOSEK ApS. *MOSEK Fusion API for Python 9.3.20*, 2019.

[58] Eva Heinen, Bert van Wee, and Kees Maat. Commuting by Bicycle: An Overview of the Literature. *Transport Reviews*, 30(1):59–96, Jan 2010.

[59] George Liu, Marco te Brömmelstroet, Sukanya Krishnamurthy, and Pieter van Wesemael. Practitioners’ perspective on user experience and design of cycle highways. *Transportation Research Interdisciplinary Perspectives*, 1:100010, 2019.

[60] Region Hovedstaden. Cykelregnskab 2020. Technical report, AFRY for Region Hovedstaden, 2020.

27
[61] Mogens Fosgerau and Ninette Pilegaard. The Rule-of-a-Half and Interpreting the Consumer Surplus as Accessibility. In Roger Vickerman, editor, *International Encyclopedia of Transportation*, pages 237–241. Elsevier, 1 2021.

[62] Technical University of Denmark. Transport Economic Unit Prices vers. 2.0, 2022.