C. PCT methodology

In this Section of the Manual we summarise the method used to create the PCT. Full details can be found in the technical appendix of our publication Lovelace et al (2016).

1. PCT input datasets

i. Core input dataset: 'travel to work' origin-destination dataset from the Census 2011

To estimate cycling potential, the PCT was designed to use the best available geographically disaggregated data sources on travel patterns. Currently for England this is the 2011 Census data on main mode of travel to work. For this reason, the commuting layer has been the first layer added to PCT-England. The 2011 English Census was conducted on 27th March 2011 and covered an estimated 94% of the population. All individuals aged 16 or over with a current job were asked "How do you usually travel to work? (Tick one box only, for the longest part, by distance, of your usual journey to work)". The commuting layer of PCT-England is based on the 22,676,958 commuters living in England in 2011, with adults who reported that their home address was also their place of work being treated as non-commuters.

The core input dataset contains origin-destination (OD) pairs that link each commuter's usual place of residence to the workplace location of their main job, and disaggregate these OD pairs by commute mode. Equivalent datasets are available at two geographical levels:

- Middle Layer Super Output Area (MSOA): MSOAs are administrative regions designed to contain a population of around 7500 individuals (average 3325 commuters). Usual place of residence is identified at the level of the MSOA, as are workplace locations for those with a fixed workplace. This dataset contains N= 2,339,535 OD pairs for commuters living in England, and is available as an open-access dataset from https://wicid.ukdataservice.ac.uk/.
- Lower Layer Super Output Area (LSOA): LSOAs are administrative regions designed to contain a population of around 1560 individuals (average 690 commuters). Usual place of residence is identified at the level of the LSOA, as are workplace locations for those with a fixed workplace. This dataset contains N= 7,433,540 OD pairs for commuters living in England, and is available as an safeguarded dataset from https://wicid.ukdataservice.ac.uk/.

We enhanced these OD datasets by merging in other Census and route characteristic data, including the number of male and female commuters and the number of male and female cyclists in each OD pair and the background mortality rate for existing and new cyclists under different scenarios. We also merged into each OD pair the distance and gradient of the 'fastest' routes. This was done using a routing algorithm 'developed for cyclists by cyclists' by the not-for-profit organisation CycleStreets (www.CycleStreets.net). Gradient was measured as the average slope experienced along the course of the route as a percentage. For example, a gradient of 2% indicates that for every 100m travelled horizontally the route involves a total change in vertical distance of 2m. This change of 2m could potentially reflect a rise of 2m or a fall of 2m or, for example, a rise of 1m followed by a fall of 1m.

Complementary analyses of national travel surveys, to parameterise scenarios

In addition, some of our analysis decisions and model parameterisation drew on analyses of the National Travel Surveys (NTS) in England (2008-2014, accessed from http://discover.ukdataservice.ac.uk/), the Netherlands (2010-2014, accessed from https://easy.dans.knaw.nl/ui/home) and Switzerland (2010, obtained from the Swiss Federal Statistical Office, Neuchâtel (Bundesamt für Statistik 2012), with data processing by Thomas Götschi).

All three are nationally-representative surveys that include a travel diary, of duration 1 week in England and 1 day in the Netherlands and Switzerland.

ii. Who is included in the propensity to cycle models?

In our analysis, we distinguish between 4 types of OD pairs as shown in Table 1 with reference to the MSOA layer. As this table shows, all commuters are included in our counts of the number of cyclists at baseline. However, we do not model cycling as increasing for OD pairs that have fast route distance of >30 km, or where the workplaces outside England. All types of OD pairs are included in our zone-level summaries on the PCT. Only some OD pairs are represented as lines in the PCT interface. Specifically, each region only shows lines that a) have a fast-route distance less than 20km, and b) contain more than a certain number of commuters (usually 10 for the MSOA layer and 5 for the LSOA layer) by any mode, counting commuters in both directions. In addition, the Route Network (MSOA) only includes commuters who start and end in the PCT region. The Region Stats tab gives details of the criteria used in each region.

Table 1: Summary of how different types of OD pairs are modelled and represented in PCT, for the MSOA layer*

Type of OD pair	% of OD pairs	% of comm-uters	% of cyclists at baseline	Included in count of cyclists at baseline?	Modelled as increasing in scenarios?	Included in zone-level summaries in the PCT interface?	Represented as lines in the PCT interface?	Included in route network estimates in the PCT interface?
Type 1: <30km, between MSOAs	44.6%	69.9%	78.2%	Yes	Yes	Yes	Sometimes, see Region Stats tab	Sometimes, see Region Stats tab
Type 2: within MSOAs	0.3%	9.1%	13.1%	Yes	Yes	Yes	No, represented as centroids	No
Type 3: No fixed workplace	0.3%	9.1%	4.9%	Yes	Yes	Yes	No	No
Type 4: >30km within England or workplace outside England	54.9%	12.0%	3.9%	Yes	No	Yes	No	No

* Results for the LSOA layer are similar except that I am proportion of commuters are in Type 1 as opposed to Type 2 flows

2. Modelling baseline propensity to cycle

iii. Plain language overview

In order to generate 'what if' scenarios regarding possible future levels of cycling, we first sought to model current propensity to cycle – i.e. the current proportion of commuters who cycle to work. We did this using OD data from the 2011 Census, and modelling cycling commuting as a function of route distance and route hilliness. We modelled cycling at baseline so using logistic regression applied at the individual level, modelling the relationship between the proportion of commuters cycling (the dependent variable) and the fastest route distance and route gradient (the two explanatory variables). Our equations included squared and square-root terms for distance to capture 'distance decay' – the non-linear impact of distance on the likelihood of cycling – and included 'interaction' terms to capture the fact that the impact of trip distance varies according to the level of hilliness. We also developed equations to estimate commuting mode share among groups with no fixed workplace.

This model of baseline propensity to cycle formed the basis of three of the four scenarios (Government Target, Go Dutch and Ebikes), as described in more detail in the next section.

iv. Why focus on distance and hilliness?

In modelling baseline propensity to cycle, we focused on the two characteristics of distance and hilliness as both are strong predictors of the probability of cycling a trip, and as both are likely to continue to have some effect on cycling propensity in all cycling futures. For example, even in high-cycling places like the Netherlands, people are much more likely to cycle a 2 km trip than a 10 km trip. By contrast, other possible predictors of current propensity to cycle, such as gender or age, may be more amenable to change. For example, although cycling in England is concentrated among younger males, in the Netherlands cycling is more common among women than among men, and is common across all age groups (see Section C3viii). Thus we did not want to include such individual-level characteristics as predictors of cycling, as we did not want to assume that future cyclists in England would have the same characteristics as current cyclists.

v. Why focus on more direct 'fast' routes?

In measuring trip distance and hilliness, we focused on the 'fastest' routes presented by CycleStreets. We did this despite the fact that many cyclists currently choose to take a quieter route at the cost of extra time because often the fast route involves sharing with motor traffic on busy roads. However, the aim of the PCT is not to predict exactly where people are currently cycling. Rather we are trying to prioritise where to put new infrastructure.

We believe that in general the 'fastest' route should be considered as the first choice for creating good cycling routes. This is particularly the case if one is seeking to encourage cycling among groups currently underrepresented, such as women and older people. This is important for 2 reasons. First, these groups are more likely to be put off cycling on direct routes in the absence of high quality infrastructure. A systematic review found that most people find cycling with busy traffic is hugely off-putting, and this is particularly true of women and probably also older people and those riding with children (Aldred, Elliott et al. In press). Second, these groups are also more likely to be put off by cycling longer distances, which alternative 'quiet' routes may involve. For example, analysis of the National Travel Survey indicates that if a quieter route creates a detour such that a 2-mile trip becomes effectively a 3-mile trip, younger men's propensity to cycle the route will decrease 11%. But for younger women, the decline is 19%, and for older adults (60+) the propensity would decrease by 35%.

Thus, for utility trips, improving direct routes will reduce safety and time disincentives to cycling. So while a good proportion of current cyclists may use the ‘quietest’ route, a big increase in capacity will likely necessitate substantial improvements to the ‘direct’ route, which will then carry many more riders from a wider demographic.

vi. Technical details

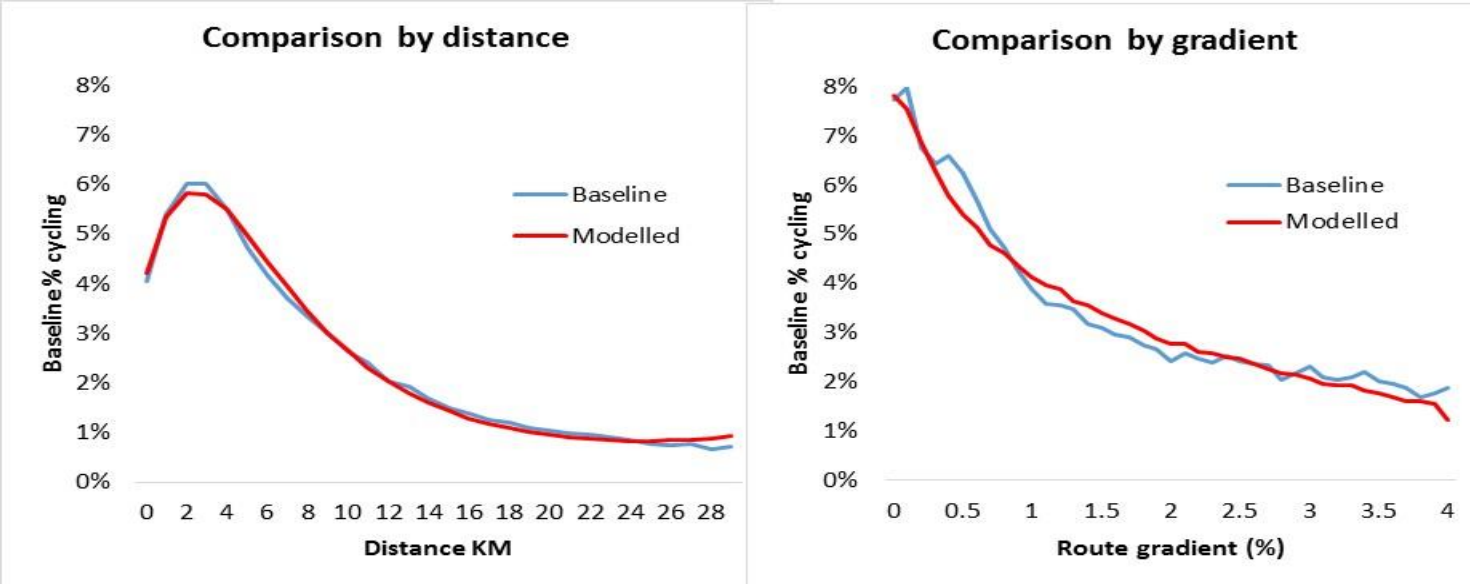
For all within-MSOA and between-MSOA OD pairs in England with a fastest-route distance of <30km, we modelled the relationship between the proportion of commuters cycling (the dependent variable) and the fastest route distance and route gradient (the two explanatory variables). We did this using an individual-level logit model, expanding the ~1 million OD pairs to their constituent ~18 million commuters. Distance decay was modelled using linear, square-root and square terms (Equation 1A). The ‘gradient’ variable was entered as the original gradient derived from CycleStreet.net minus 0.57%, which is the estimated average route gradient in the Netherlands. By centring our gradient measure on the estimated Dutch average in this way, we facilitated the subsequent addition of ‘Go Dutch’ parameters to the baseline equation (see Section C3viii). Interaction terms were included to capture the fact that the deterrent effect of a steeper slope appeared to be stronger for individuals travelling intermediate distances. The resulting equation for baseline propensity to cycle was:

Equation 1A: logit (pcycle) $= -3.894 + (-0.5872 * \text{distance}) + (1.832 * \text{distance}_{sqrt}) + (0.007956 * \text{distance}_{sq}) + (-0.2872 * \text{gradient}) + (0.01784 * \text{distance}*\text{gradient}) + (-0.09770 * \text{distance}_{sqrt} *\text{gradient})$

pcycle $= \exp([\text{logit (pcycle)}]) / (1 + (\exp([\text{logit(pcycle)}])$

where ‘pcycle’ is the proportion of cyclists expected; ‘distance’ is the fastest route distance in km, ‘distance_{sqrt}’ and ‘distance_{sq}’ are, respectively the square-root and square of distance; and ‘gradient’ is the fastest-route gradient (centred on 0.57%). Note that although this equation was derived at the individual level, it can be applied at the level of the OD pairs as distance and gradient are constant within OD pairs. Equation 1A showed good fit to the observed data with respect to both distance and hilliness (Figure 1).

Figure 1: Observed versus predicted prevalence of cycling to work among 17,896,135 English commuters travelling <30km to work, according to a) route distance and b) route gradient



For commuters with no fixed workplace, we modelled propensity to cycle as a function of the average propensity to cycle among commuters living in the same MSOA and commuting <30km. The resulting equation for baseline propensity to cycle among those with no fixed workplace was:

<u>Equation 2A</u>: logit (pcycle) $= -6.219 + (189.9 * \text{meanpropensity}_{sq}) + (9.275 * \text{meanpropensity}_{sqrt})$
pcycle $= \exp(\text{logit(pcycle)}) / (1 + (\exp(\text{logit(pcycle)}))$

where '$\text{meanpropensity}_{sq}$' is the square of the mean propensity to cycle among type 1 and type 2 OD pairs in the home MSOA in question, and '$\text{meanpropensity}_{sqrt}$' is the square root term. Both Equations 1A and 2A were derived using the MSOA dataset and subsequently applied to the LSOA dataset.

Finally, we did not model baseline propensity to cycle among individuals living more than 30km from their place of work or commuting outside England. Instead, given the considerable uncertainties about where the cycling reported by these individuals was taking place, we assumed no increase in cycling levels among these commuters in our scenarios.

3. Modelling cycling across scenarios

Four scenarios were developed to explore cycling futures in England. These can be framed in terms of the removal of different infrastructural, cultural and technological barriers that currently prevent cycling being the natural mode of choice for trips of short to medium distances.

The scenarios are not predictions of the future. They are snapshots indicating how the spatial distribution of cycling may shift as cycling grows based on current travel patterns. At a national level, the first two could be seen as shorter-term and the second two more ambitious. The choice of scenarios was informed by a government target to double the number of cycle trips and evidence from England overseas about which trips *could* be made by cycling.

Each scenario is described below, with both a plain language overview and an account of the technical details. The accounts of the technical details can be complemented by the summary of the scenario generation rules presented in Table 2.

Table 2: Summary of scenario generation rules

Scenario	Baseline no. cyclists (A)	Initial estimation of scenario no. cyclists (B1)	Additional processing of scenario no. cyclists (B2)	Scenario increase in no. cyclists (C)
Government Target	Recorded no. in Census 2011, OD pair types 1-4.	Column A + (Baseline propensity to cycle [Equations 1A+2A]† in OD pair types 1-3 * no. commuters)	• Cap Column B1 at 100%.	Column B2 – Column A
Go Dutch	Recorded no. in Census 2011, OD pair types 1-4.	'Go Dutch' propensity to cycle [Equations 1B+2B, with 'dutch'=1 and 'ebike'=0] in OD pair types 1-3 * no. commuters.	• Set Column B1 as equal to Column A if B1 is less than A.	Column B2 – Column A
Ebikes	Recorded no. in Census 2011, OD pair types 1-4.	'Ebikes' propensity to cycle [Equations 1B+2B, with 'dutch'=1 and 'ebike'=1] in OD pair types 1-3 * no. commuters.	• Set Column B1 as equal to Column A if B1 is less than A.	Column B2 – Column A
Gender Equity	Recorded no. in Census 2011, OD pair types 1-4.	Apply Equation 3 in OD pair types 1-3.	• Set Column B1 as equal to Column A if number of males in the OD pair is zero, or if B1 is less than A.	Column B2 – Column A

† Or, equivalently, using equations 1B + 2B in Section A1.3.3, 'dutch'=0 and 'ebike'=0

vii. Government target scenario

Plain language overview

The ‘Government Target’ scenario models a doubling of cycling nationally, corresponding to the proposed target in the Department for Transport’s draft Cycling Delivery Plan to double cycling in England between 2013 to 2025 (Department for Transport 2014). Although substantial in relative terms, the rate of cycling under this scenario (rising from 3% to 6% of commuters) remains low compared with countries such as the Netherlands and Denmark.

The result is that cycling overall doubles at the national level, but at the local level this growth is not uniform, in absolute or relative terms. Areas with many short, flat trips and a below-average current rate of cycling are projected to more than double. Conversely, areas with above-average levels of cycling and many long-distance hilly commuter routes will experience less than a doubling.

Technical details

The Government Target scenario was generated by adding together a) the observed number of cyclists in each OD pair in the 2011 Census, and b) the modelled number of cyclists in each OD pair, as estimated using the baseline propensity to cycle equations described in the previous section.

This is scenario is illustrated by the following example. Take an OD pair of 200 commuters containing 7 cyclists in the 2011 Census, and with a modelled propensity to cycle of 5.2%. Using Equation 1A the total number of cyclists in the Government Target scenario would be 7 + (200 * 0.052) = 7 + 10.4 = 17.4. Thus the Government Target scenario leads to a doubling of cyclists in England but not necessarily of each OD pair. Note the reported ‘baseline’ number of cyclists directly influences the total number of cyclists in the scenario (column B2 in Table 2), but does not influence the scenario increase in the number of cyclists (Column C).

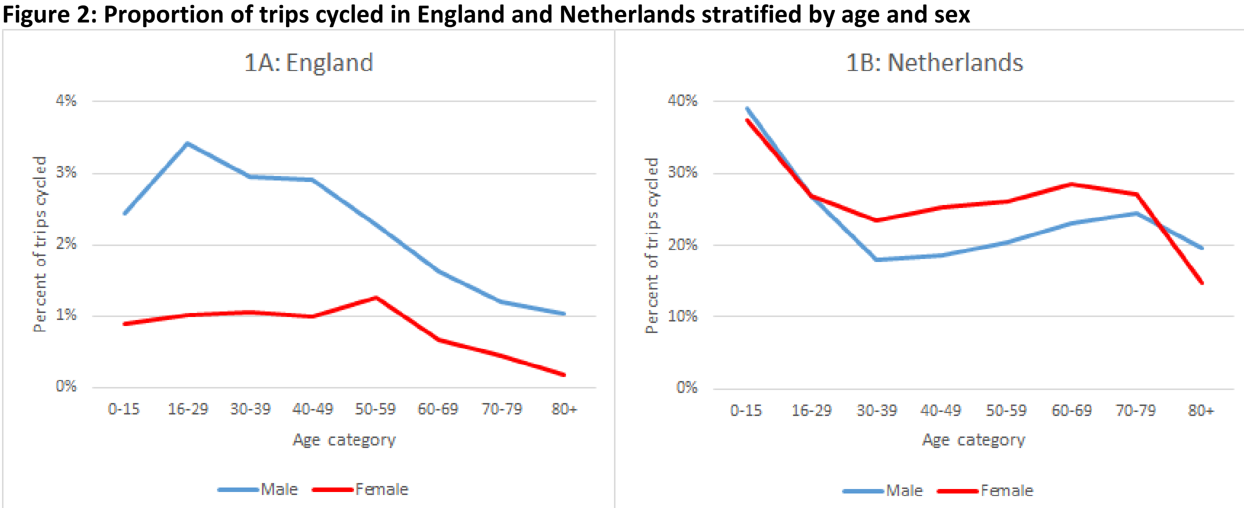
viii. Go Dutch and E-bike scenarios

Plain language overview

While the Government Target scenario models relatively modest increases in cycle commuting, the Go Dutch and E-bikes scenarios are an ambitious vision for what cycling in England look like. People in the Netherlands make 26.7% of trips by bicycle, fifteen times higher than the figure of 1.7% in England. In addition, cycling in England is skewed towards younger, male cyclists (Figure 2A). By contrast in the Netherlands cycling remains common into older age, and women are in fact slightly more likely to cycle than men (Figure 2B).

This means that the difference between England and the Netherlands is particularly large for women and older people. For example, whereas the cycle mode share is 'only' six times higher in the Netherlands than in England for men in their thirties, it is over 20 times higher for women in their thirties or men in their seventies and eighties. For women in their seventies and eighties, the cycle mode share is over 60 times higher in the Netherlands than in England.

Figure 2: Proportion of trips cycled in England and Netherlands stratified by age and sex



The Go Dutch scenario represents what would happen if English people were as likely as Dutch people to cycle a trip of a given distance and level of hilliness. This scenario thereby captures the proportion of commuters that would be expected to cycle if all areas of England had the same infrastructure and cycling culture as the Netherlands (but retained their hilliness and commute distance patterns). The scenario was generated by taking baseline propensity to cycle (see Section 30), and applying Dutch scaling factors calculated through analysis of the English and Dutch National Travel Surveys. The Go Dutch scaling factors comprised two parameters which boost the rate of cycling for each OD pair above the baseline model, with one fixed and one distant-dependent term - the latter takes into account the fact that the "Dutch multiplier" is greater for shorter trips compared to longer trips.

Note that the level of cycling under the Go Dutch scenario is unaffected by the current level of cycling, but is instead purely a function of trip distance and hilliness. Note also that this means that a few lines or areas show a decrease in cycling under the Go Dutch scenario as compared to baseline; this might happen in a very high-cycling area, where cycle commuting in the 2011 Census is similar to or even higher than the average for the Netherlands. For example, Cambridge, England's highest cycling region, shows only a modest overall increase under the Go Dutch scenario for this reason. Planners in

Cambridge might therefore want to consider creating a bespoke alternative scenario, e.g. “Go Groningen”, using cycling propensity from Groningen, the highest-cycling province in the Netherlands.

The E-bikes scenario models the additional increase in cycling that would be achieved through the widespread uptake of electric cycles ('e-bikes'). This scenario is built as an extension of the Go Dutch scenario, making the further assumption that all cyclists in the Go Dutch scenario own an e-bike. It builds on the Go Dutch scenario by applying three additional E-bike scaling factors to account for the increased willingness of E-bike users to cycle long distance, hilly and simultaneously long distance and hilly routes. These scaling factors were generated by analysing the impact of e-bike ownership based on the Swiss National Household Travel Survey and the Dutch National Travel Survey, weighted to be representative of English commuters. This scenario may be particularly suitable for examining cycling potential in hilly areas and/or where trip distances are longer (e.g. in rural areas).

Technical details

For the Go Dutch and E-bike scenarios, our approach was to start from the Equations estimating baseline propensity to cycle (Equation 1A and 2A) and add additional parameters. Here we provide an overview of the methods and input datasets used: full details can be found in Lovelace et al (In press).

The Go Dutch scenario required us to model the increase in propensity to cycle that would be observed if English commuters became as likely to cycle a given trip as Dutch commuters. We estimated this additional parameter using trip-level analysis of the English and Dutch National Travel Surveys, restricting the analysis to commute trips of less than 30km. In estimating the increased propensity to cycle among Dutch people, we included both a main effect term and an interaction term with distance (as a linear term). We introduced the interaction term to reflect the fact that Dutch propensities to cycle exceed English propensities by a greater amount for short distances (e.g. Dutch people are 5.4 times more likely to cycle a trip of 0-4.9km versus 3.6 times more likely to cycle a trip 10-14.9km). As hilliness data was not available in the Dutch survey, we weighted the data so that the English sample of commuters lived in areas with the same hilliness profile as the Dutch commuters.

The Ebike scenario builds on the Go Dutch scenario and models the further increase in propensity to cycle that would be observed if all Dutch cyclists acquired an e-bike. To generate the relevant parameters, we restricted our analysis to the Dutch NTS 2013-2014, the only years that measured e-bikes as a separate mode. We further restricted our analysis to the 26,807 commute trips made by 13,693 adults who owned a bicycle. We then compared propensity to cycle between the population of e-bike owner trips (N = 2175) with the full population of all bicycle-owner trips (N = 26,807). This analysis therefore takes into account the fact that some e-bike owners are already present in the ‘Go Dutch’ scenario, and captures only the extra cycling that would occur if *everyone* with a traditional bicycle acquired an e-bike.

In estimating the extent to which this would increase propensity to cycle in the E-bike scenario, we included interaction terms with distance (as a linear and squared term). We did this to capture the fact that owning an e-bike increases propensity to cycle more for long trips than for short trips (e.g. Dutch e-bike owners are 1.1 times more likely than all Dutch bicycle owners to cycle a trip 0-4.9km versus 2.3 times more likely to cycle a trip 10-14.9km). Because we did not have data on hilliness in the Dutch National Travel Survey we could not estimate the magnitude of any interaction between e-bike ownership and hilliness in this dataset. We therefore instead estimated the interaction term between e-bike use and average route gradient using data from the Swiss National Household Travel Survey 2010.

Adding these 'Go Dutch' and 'Ebikes' parameters together, we derived the following propensity to cycle equation:

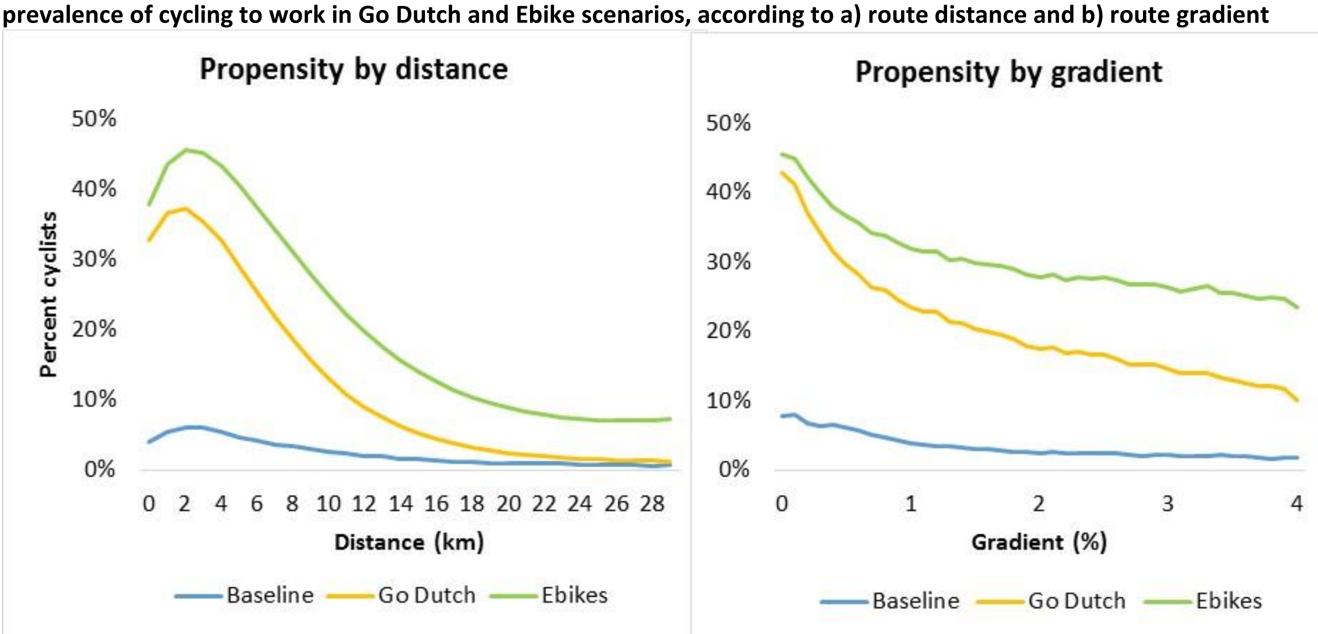
Equation 1B: logit(pcycle) = Equation 1A + Dutch parameters + Ebike parameters

logit(pcycle) = -3.894 + (-0.5872 * distance) + (1.832 * $distance_{sqrt}$) + (0.007956 * $distance_{sq}$) + (-0.2872 * gradient) + (0.01784 * distance*gradient) + (-0.09770 * $distance_{sqrt}$ *gradient) + (2.499 * dutch) + (-0.07384 * dutch * distance) + (0.05710 * ebike * distance) + (-0.0001087 * ebike * $distance_{sq}$) + (0.1924 * ebike * gradient).

where 'pcycle' is the proportion of cyclists expected; 'distance' is the fastest route distance in km, '$distance_{sqrt}$' and '$distance_{sq}$' are, respectively the square-root and square of distance; 'gradient' is the fastest-route gradient (centred on 0.57%); 'Dutch' is a binary variable that takes the value '0' for the Government Target scenario and '1' for the Go Dutch or the Ebike scenario; and 'ebike' is a binary variable that takes the value '0' for the Government Target and Go Dutch scenario and '1' for the Ebike scenario.

Figure 3 shows the distribution of cycling propensity generated by Equation 1B, according to distance and hilliness.

Figure 3: Prevalence of cycling to work at baseline among English commuters travelling <30km to work, and modelled prevalence of cycling to work in Go Dutch and Ebike scenarios, according to a) route distance and b) route gradient



For commuters with no fixed workplace, we similarly started with Equation 2A, and extended this as follows.

Equation 2B: logit(pcycle) = Equation 2A + mean Dutch parameters + mean Ebike parameters
= -6.218 + (189.9 * meanpropensitysq) + (9.275 * meanpropensitysqrt) + (dutch * meandutch) + (ebike * meanebike)

where '$\text{meanpropensity}_{sq}$' is the square of the mean propensity to cycle among type 1 and type 2 OD pairs in the home MSOA in question, and '$\text{meanpropensity}_{sqrt}$' is the square root term; 'meandutch' is the average value of the Equation 1B Dutch parameters for commuters living in the same home MSOA; and 'meanebike' is the average value of the Equation 1B Ebike parameters for commuters living in the same home MSOA. Again, these equations were derived using the MSOA dataset and subsequently applied to the LSOA dataset.

ix. Gender equality

Plain language overview

In the 2011 Census, women accounted for 48% of all English commuters but only 27% of all cycle commuters. This gender disparity is seen across the country, with no local authority having a proportion of female cyclists greater than 50%. However, in places such as the Netherlands where cycling accounts for a high proportion of personal travel, women cycle at least as much as men (Pucher, Dill et al. 2010). Places in England with higher overall levels of commuter cycling also tend to have smaller gender inequalities in commuter cycling (Aldred, Woodcock et al. 2016).

The 'Gender Equity' scenario seeks to capture a situation in which these gender disparities are eliminated. In this respect, it differs somewhat from the preceding three scenarios, as it does not use distance and hilliness data to model propensity to cycle. Instead it assumes that male propensity to cycle remains unchanged – i.e. there is no change in the number of male cycle commuters – and that female propensity to cycle rises to match male propensity. This scenario has the greatest relative impact in areas where the rate of cycling is highly gender-unequal.

Technical details

The Gender equality scenario assumes that male propensity to cycle remains unchanged – i.e. there is no change in the number of male cycle commuters – and that female propensity to cycle rises to match male propensity. We implemented this using the following equation:

Equation 3: $\text{SNcyclists} = \text{BNcyclists}_m * (1 + (\text{BNcommuters}_f / \text{BNcommuters}_m))$

where 'SNcyclists_f' is number of female cycle commuters in the gender equality scenario, 'BNcyclists_m' is the recorded number of male cycle commuters at baseline, and 'BNcommuters_f' and 'BNcommuters_m' are the total numbers of females and males in the OD pair respectively.

To illustrate how this method works in practice, imagine an OD pair in which 50 from a total of 500 people commute by cycle, 35 males ($\text{BNcyclists}_m = 35$) and 15 females ($\text{BNcyclists}_m = 15$). 300 of the total trips in the OD pair are made by males ($\text{BNcommuters}_m = 200$) and 200 by females ($\text{BNcommuters}_f = 200$). Applying Equation 3:

$$\text{SNcyclists} = \text{BNcyclists}_m * (1 + (\text{BNcommuters}_f / \text{BNcommuters}_m))$$
$$\text{SNcyclists} = 35 * (1 + (200 / 300))$$
$$= 58.3$$

All of these extra 8.3 cyclists are female, giving a new total of 15 + 8.3 = 23.3 female cyclists (and still 35 male cyclists). Gender equality in cycling has been reached, such that 11.7% of commute trips are made by cycling among both men (35/300) and women (23.3 / 200). Equation 3 was applied to commuters with 'no fixed workplace' in the same way, and as in other scenarios we assumed no change among commuters travelling >30km or outside England.

4. Estimating mode shift, health impacts and reductions in carbon emissions

x. Modelling scenario mode shift in walking and car driving

To estimate the health impacts of our scenarios, we needed to estimate the number of new cyclists who had previously commuted on foot. Similarly, to estimate the carbon impacts of our scenarios, we needed to estimate the number of new cyclists who had previously commuted as car drivers. We assumed that within any given OD pair commuters were equally likely to shift to cycling from any baseline mode, and therefore the mode shift was proportional to mode share at baseline.

For example, take an OD pair containing 220 commuters at baseline, of whom 20 cycle, 80 walk, 50 are car drivers and 70 use other modes of transport. If the 'Government Target' scenario number of cyclists rose to 50 in this OD pair, this would mean that the number of non-cyclists decreased to 170, giving a ratio of change among non-cyclists of 170 / 200 = 0.85. We assume this 0.85 applies to all modes, and therefore the number of pedestrians in the scenario 80 * 0.85 = 68; the number of car drivers in the scenario is 50 *0.85 = 42.5; and the number of commuters using other modes in the scenario is 70 * 0.85 = 59.5.

For the purposes of estimating health and carbon impacts of the current level of cycling relative to a 'no cycling' counterfactual, we made the same assumption. For example, again take the OD pair containing 220 commuters at baseline, of whom 20 cycle, 80 walk, 50 are car drivers and 70 use other modes of transport. In a 'no cyclists' counterfactual, the number of non-cyclists would increase to 220, giving a ratio of change among non-cyclists of 220 / 200 = 1.1. Thus in the 'no cyclists' counterfactual, the number of pedestrians would be 80 * 1.1 = 88, and so on. When estimating mode split in the 'no cyclists' counterfactual in the 4552 OD pairs that at baseline consisted entirely of cyclists, we assumed a mode split of 31% walking, 35% car drivers and 34% other modes. These percentages correspond to the observed mode split among the 974 OD pairs in which 70-99% of individuals cycled in the 2011 Census.

xi. Estimating the physical activity health benefits

An approach based on the World Health Organization's Health Economic Assessment Tool (HEAT) was used to estimate the number of premature deaths avoided due to increased physical activity (Kahlmeier, Kelly et al. 2014). In this Manual we provide an overview of our methods: for full details see Lovelace et al (In press).

Trip duration was estimated as a function of the 'fast' route distance and average speed, with the latter being calculated using the National Travel Survey. For walking and cycling we applied the standard HEAT approach. E-bikes are not specifically covered in HEAT Cycling but enable faster travel and require less energy from the rider than traditional bikes. Thus we estimated new speeds and intensity values for this mode, giving a smaller benefit for every minute spent using E-bikes than conventional cycles.

The risk of death varies by gender and increases rapidly with age. This was accounted for using age and sex-specific mortality rates for each local authority in England. For the baseline and Government Target scenario the age distribution of cyclists recorded in the 2011 Census was used. New cyclists under Go Dutch and E-bikes were assumed to have the age-gender profile of commuter cyclists in the Netherlands. The health benefits per additional cyclist were thus particularly large in the latter two scenarios, since cycling at older ages is more common in the Netherlands (see section 3Cviii), and since the health benefits of additional physical activity are greatest among older people.

To allow for the fact that cycling would in some cases replace walking trips, HEAT estimates of the increase in premature deaths due to the reduction in walking were also calculated. For a trip of a given distance, walking involves more physical activity than cycling. This means that the observed health benefits can be negative if a high proportion of new cyclists previously walked. This is particularly common in very short trips, and in these cases health benefits are presented in red.

The net change in the number of deaths avoided for each OD pair was estimated as the number of deaths avoided due to cycle commuting minus the number of additional deaths due to reduced walking. Note that this approach means that for some OD pairs where walking made up a high proportion of trips, additional deaths were incurred. The monetary value of the mortality impact was calculated by drawing on the standard 'value of a statistical life' used by the Department for Transport (£1,855,315 in 2014 money).

xii. Estimating reductions in transport carbon dioxide emissions from car driving

When comparing each scenario to baseline, we estimated the reduction in transport carbon dioxide (CO_2) emissions as follows:

Change in CO_2-equivalent emissions (in kg) per year
= Change in no. car drivers * former distance travelled by former car drivers * mean cycle commute trips per cyclist per week * 52.2 * CO_2-equivalent emissions (in kg) per kilometre

The change in the number of car drivers was estimated using the mode shift calculations described in Section 3Dx. Note that we specifically focus on car drivers, not car passengers, as the standard practice in estimating transport CO_2 emissions is to attribute all emissions to the car driver, to avoid double-counting. Their average former distance was assumed to be equal to the new ‘fastest route’ distance travelled by the cycle commuters. The mean cycle commute trips per cyclist per week was estimated to be 5.24 from the National Travel Survey. The average CO_2-equivalent emission per kilometre car driving was taken as 0.186kg, which is the 2015 value for an ‘average’ car of ‘unknown’ size in the UK government’s carbon conversion factors (DEFRA 2015).

5. Aggregate estimates to provide zone-level estimates and to form the route network

xiii. Aggregating OD pairs to give zone-level results, and to give bidirectional lines

The OD pairs provided by the Census are directional, with one OD pair for travel from origin A to destination B, and another for travel from origin B to destination A. After performing the modelling stages described above, we aggregated the values for individual OD pairs to the zone level by summing our outcome variables across all OD pairs with the same home MSOA. This gave us MSOA-level estimates of the total number of cycle, foot and car commuters living in each MSOA in each scenario, plus the total change in mortality and in CO_2 emissions resulting from behaviour change among residents of that MSOA. Equivalent calculations were done for each LSOA layer when using the LSOA dataset.

In addition, we aggregated directional OD pairs to be bidirectional by adding up the values in both directions between a given pair of locations (e.g. adding the values for the A-to-B OD pair with the values of the B-to-A OD pair). These bidirectional totals are what we present in our visualisation tool.

xiv. The MSOA-level route network layer

Information about the *aggregate cycling potential* on the road network is shown at the MSOA level in the Route Network (MSOA) layer. This layer was generated by aggregating overlapping MSOA-level 'fast' routes, and summing the level of cycling for each scenario. This layer therefore relates to the *capacity* that infrastructure may need to handle.

Note that more confidence can be placed in the relative rather than the absolute size of the numbers presented for the route network: i.e. one can say with more confidence that “the number of commuters increases approximately 5-fold under this scenario” than that “there are 1200 cycle commuters using this route under this scenario”. The absolute numbers need to be treated with some caution because they are underestimates as the Route Network (MSOA) layer excludes within-zone commuters, commuters with no fixed workplace, and commuters travelling across regional boundaries (see Region Stats tab for details).

Route Network (MSOA) values also omit routes due to the adjustable selection criteria: maximum distance and minimum total numbers of all-mode commuters per OD pair. Typically, these are set to 20 km Euclidean distance and 10 commuters respectively. Nationally, the Route Network layer under these settings accounts for around two thirds of cycle commuters. Of course, in reality the total number of cyclists would also include people travelling for non-commuting purposes.

xv. The LSOA-level route network layer

The Route Network (LSOA) layer is generated at LSOA level. Because the LSOA-level dataset is so much larger than the MSOA-level dataset, it was not feasible to create a click-on route network at the LSOA level that was equivalent to that at the MSOA level. Instead, the solution to visualizing the LSOA level route network data involved rasterising the routes and pre-rendering the data as a raster tile set. This reduced the complexity of the in-browser visualisation but with some compromise in functionality, as the raster tiles lack the interactivity of the vector lines.

The Route Network (LSOA) layer provides a more detailed geographical resolution and can model the route for many short trips that were ‘within zone’ in the MSOA layer. This combination means that for some routes (e.g. busy main roads in London) the LSOA layer estimates a substantial uplift in modelled cycling flows. While the MSOA-level route network excludes commuters who cross regions, who travel on routes with relatively few commuters, and can only visualise up to 50% of the network, the LSOA-level route network includes all who commute within England, all LSOA-level OD pairs, and visualises the whole network.

6. References

Aldred, R., B. Elliott, et al. (2017). "Cycling provision separated from motor traffic: a systematic review exploring whether stated preferences vary by gender and age." Transport Reviews 37(1).

Aldred, R., J. Woodcock, et al. (2016). "Does more cycling mean more diversity in cycling? An analysis of English and Welsh Census data on cycle commuting, exploring shifts in gender and age composition." Transport Reviews **36**(1): 28-44.

Bundesamt für Statistik, B. f. R. (2012). Mobilität in der Schweiz: Ergebnisse des Mikrozensus Mobilität und Verkehr 2010. Neuchâtel, BfS, ARE.

DEFRA (2015). DEFRA Carbon Factors: UK Government conversion factors for Company Reporting, 2015, V2.0 Department for the Environment, Food and Rural Affairs [Accessed 18/2/2016 from http://www.ukconversionfactorscarbonsmart.co.uk/].

Department for Transport (2014). Cycling Delivery Plan. [Accessed 15/02/2016 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/364791/141015_Cycling_Delivery_Plan.pdf].

Kahlmeier, S., P. Kelly, et al. (2014). Health economic assessment tools (HEAT) for walking and for cycling: economic assessment of transport infrastructure and policies. Methods and user guide, 2014 update. Copenhagen, World Health Organisation.

Lovelace, R., A. Goodman, et al. (2017). "The Propensity to Cycle Tool: An open source online system for sustainable transport planning." Journal of Transport and Land Use 10(1).

Pucher, J., J. Dill, et al. (2010). "Infrastructure, programs, and policies to increase bicycling: an international review." Prev Med **50 Suppl 1**: S106-125.