Assessing the impact of the National Cycle Network and physical activity lifestyle on cycling behaviour in England

Assessing the impact of the National Cycle Network and physical activity lifestyle on cycling behaviour in England

Dr Paul Downward

School of Sport, Exercise and Health Sciences

Loughborough University

Loughborough

LE11 3TU

UK

p.downward@lboro.ac.uk

+44 (0)1509 226365

Dr Simona Rasciute

School of Business and Economics

Loughborough University

Loughborough

LE11 3TU

UK

s.rasciute@lboro.ac.uk

+44 (0)1509 223620

Assessing the impact of the National Cycle Network and physical activity lifestyle on cycling behaviour in England

Abstract

This paper examines the association between access to National Cycle Network (NCN) routes in England and an individual's cycling behaviour whilst accounting for their broader physical activity lifestyle and controlling for their socio-economic circumstances. It identifies a positive association between access to these routes and the total minutes of any form of cycling, and the number of days that cycling takes place primarily for recreational purposes. The broader physical activity of individuals also has a positive association with cycling. Walking appears most likely to be complementary to non-recreational cycling, whilst participation in sport with all forms of cycling, but not with longer duration utilitarian trips. The research also indicates that access to NCN routes has the potential to increase such cycling further, with the exception of longer utilitarian trips, as does a more physically active lifestyle, particularly walking. The main policy implications of the research are to recognise that cycling is intrinsically linked to other physical activity, notably, walking, but that the NCN routes measured in this study primarily support longer duration recreational activity, which is also affected by sporting activity. This suggests that one avenue for achieving the health benefits of cycling may be through promoting NCN routes to harness a more generally active lifestyle and particularly in leisure, whilst sustainability may be further promoted through being linked more to other active travel such as walking. There is a therefore a need to exploit the potential of such NCN route provision as part of this promotion.

Key words

Active Travel, cycling, walking, sport, Zero-inflated negative binomial model.

1. Introduction

Active travel through cycling is generally seen to be an important way of both reducing congestion and also contributing to a healthier and fitter population (Department for Transport, 2012). This has led to recent UK government support of £148m to improve infrastructure in 8 major cities (The Guardian, 2013). Nonetheless, significant investment in cycle routes has already taken place in the UK, notably through the efforts of active transport charity, Sustrans, which began in Bristol in the UK in 1977. Up until the 1990s Sustrans focussed on developing and improving transport conditions for walking and cycling with route development in specific localities. Beginning in 1995 and underpinned by National Lottery funding, Sustrans has since developed the National Cycle Network (NCN) to try to link both new and existing traffic free paths to quiet and traffic-calmed roads. The NCN currently stretches to 14,500 miles across the UK and passes within a mile of 55 per cent of UK homes (Sustrans, undated). Whilst the NCN development and maintenance is a considerable achievement, involving in excess of 400 partners that include planning authorities, private and public landowners and other local groups, and also requiring considerable localised voluntary support, Aldred (2012) notes that it should be recognised that the NCN has emerged in a context in which cycling is viewed primarily as a local transport issue and not one of national strategic transport planning. NCN route provision has had to develop, therefore, without adversely affecting other modes of transport. This naturally brings with it compromises and limitations in terms of what the NCN can deliver with routes often being coincident with traffic or blocked by parking (Aldred, 2012).

Nonetheless, based on the Active People Survey (APS), which is a large-scale survey of physical activity in England, including cycling, this paper examines the association between the presence of NCN routes on the total minutes of any form of cycling, and the frequency in days, and the intensity of the effort on those days, in which recreational and utilitarian cycling takes place. The latter is distinguished as cycling for purposes such as commuting to work or undertaking errands, rather than for leisure. The research controls for the socio-economic circumstances of the individuals but, of more importance, for the first time in a large-scale national data analysis, the paper also examines the association of cycling with other physical activity from walking, and sport and recreation. This is a need that has been identified in the literature by Yang et al. (2010) and argued to be theoretically important because lifestyle physical activity will be an important mediating factor influencing cycling behaviour (Chatterjee et al. 2013).

The paper employs a zero-inflated Negative Binomial regression model (ZINB) to analyze the data. This is because the data comprises 'over dispersed' counts bounded below by zero, in measuring the minutes or number of days in which the respondent cycled in the last four weeks prior to the interview .[1] One reason for the overdispersion is because the data are also characterized by excess zeros. In survey data such as the APS, respondents are asked about their behavior over a particular time period, which in this case is four weeks. Zero responses to activity such as cycling could therefore reflect never having cycled at all as 'absolute' zeros, and 'relative' zeros that reflect not having cycled in the period asked of the survey but that the respondent may do so otherwise if the circumstances were different. It follows that not only can the ZINB model examine if access to NCN routes and participation in other physical activity are associated with the actual frequency of cycling behavior, but also their impact on potential cycling through a consideration of the factors associated with a reduction in the incidence of the absolute zeros in the data.

The paper proceeds as follows. The literature examining the benefits of active travel, particularly cycling, as well as the determinants of cycling behavior and evidence on the impact of policy initiatives on cycling behavior is briefly presented in Section 2. The data are described in Section 3, with the ZINB model presented in Section 4. Results are presented in Section 5, with a discussion of the implications of the research and its limitations in Section 6. Conclusions follow in Section 7.

2. Literature review

International literature suggests that active travel through walking and cycling has two main benefits for society. On the one hand, it can reduce congestion and environmental pollution (Olgivie et al. 2004). On the other hand, it can improve the health of participants (Oja et al. 2011). Targeting policy to promote such behaviour consequently relies on identifying the factors that encourage active travel generally and, in the context of this paper, cycling specifically. Despite relatively recent arguments that, compared to walking, active travel research in cycling was undeveloped (Pikora et al., 2003), there is now a large literature addressing the determinants of cycling, employing a variety of different research designs (For a review see, for example, Heinen et al. 2010).

International surveys indicate that it is generally the case that both individual characteristics as well as social and physical environmental features such as geographic and transport related characteristics can influence cycling (Pikora et al. 2003; Panter and Jones 2010). For example, Winters et al. (2007) in a study of commuter cycling in Canada identify that males and younger adults are more likely to cycle. The same is the case for those with a lower income, and higher education. Examining commuter cycling in England and Wales, Parkin et al. (2008) support these results and identify that with the exception of higher professionals, individuals in most socio-economic classifications cycle less than those of the lowest socio-economic classification. Less cycling is also identified for non-white ethnicities. Based on more recent Census data from 2011 for England and Wales, however, Goodman (2013a) argues that whilst active travel is more common for socio-economically disadvantaged groups, this may soon reverse for cycling overall and recreational cycling is more likely to be positively associated with affluence.

Overall, therefore, some of this literature can suggest that having lower income promotes cycling behaviour because it is a cheaper transport alternative. Not surprisingly, studies also identify that car availability reduces cycling in England and Wales (Parkin et al., 2008), and an increase in the cost of fuel (Buehler and Pucher 2012), or higher car-parking charges, increase commuter cycling (Rietveld and Daniel 2004) in US cities and Dutch municipalities respectively. Similarly, modest financial incentives, good parking and shower facilities at work have a positive effect on the level of cycling to work in England (Wardman et al., 2007). However, the results linking the highest incomes and white ethnicities to greater cycling and particularly recreational cycling also suggest that cycling can be driven by choice, cultural context and perhaps facilitating conditions rather than simply travel cost (Chatterjee et al., 2013).

In the context of England and Wales, Aldred and Jungnickel (2012) suggest that changing norms can be evident through emergent cultures challenging established ones. In the context of the above results this might be that lifestyle choices have changed for those with more discretion over their travel behaviour and particularly in the context of recreation. More generally, their research suggests that interventions to promote cycling will be mediated by the identities that people have with respect to cycling, and identity can be influenced by gender, access to travel opportunities, proximity of schools and access to social networks. So too, Sherwin et al. (2014) examine how social influences and cultural cues affects the decision to start cycling in England and three levels of influences are identified: direct (immediate family), less direct (peers and colleagues) and indirect (wider cultural context). There is, consequently, debate about the impacts, for example, of gender on cycling. Whilst Pucher and Buehler (2008) show that differences in cycling by gender is much less pronounced in some European Countries such as the Netherlands, Denmark and Germany in which cycling is more embedded in behaviour and has been consistently and actively championed in transport policy, Parkin et al. (2008) and Downward and Riordan (2007) show that both commuting and recreational cycling respectively are more common for adult males in the UK, a country in which cycling remains a marginal form of transport and is therefore a key target for policy and initiatives such as the NCN. In this respect, in the context of Australian commuter cycling, Garrard et al. (2008) indicate the importance of segregating cycling from motorised transport in order to increase female cycling. Moreover, Caulfield (2014) argues, that policymakers should tailor their strategies to target distinct groups to encourage them to take up cycling. For example, it is shown that lowering motorised vehicular speed limits has led to a significant increase in female cycling to work on a regular basis in Dublin (Eire). This is particularly the case for higher professional females. Finally, it has also been argued that pro-cycling strategies should be regionally differentiated, as the results by Vandelbulcke et al. (2011) suggest that bicycle use for commuting in Belgian municipalities is influenced by neighbouring municipalities, and that differences exist in the northern and southern parts of Belgium. It follows that some areas may require greater focus for policy than others.

As far as environmental factors are concerned, Winters et al. (2007) find that greater precipitation and lower temperatures reduce commuter cycling in Canada, as do long-term seasonal variations, where commuter cycling is shown to be more frequent in warmer months in Australia (Nankervis, 1999). These results are supported for commuter cycling in US cities by Buehler and Pucher (2012). Nosal and Miranda-Moreno (2014) extend the analysis of the impact of weather on the use of urban bicycle facilities in North America and disaggregate weather conditions- temperature, humidity and precipitation- by hour as well as differentiate the analysis between weekends and weekdays. Temperature is found to be positively associated, and humidity and precipitation are found to be negatively associated, with utilitarian and recreational cycling, with strong non-linear effects for temperature and lagged effects for rain.

The topography of localities is also investigated, with the presence of hills reducing cycling in England and Wales (Parkin et al. 2008) and the Netherlands (Rietveld and Daniel 2004) as do the number of stops or hindrances (Rietveld and Daniel, 2004). Perhaps naturally it is found that greater trip distances also reduce the incidence of commuter cycling in the UK (Dickinson et al. 2003; Parkin et al. 2008) and Australia (Timperio et al., 2006) but the density of population per locality increases it's likelihood as, for example, indicated by cycling in urban areas in Denmark, Germany and the Netherlands (Pucher and Buehler 2008) and in the U.S. (Zahran et al. 2008).

Much research has focussed on the provision of cycling routes.[2] For example, Dill and Carr (2003) identify that an increase in the linear miles of cycle lane per square mile of U.S. city area increased commuter cycling. These results are supported in the U.S. by Tilahun et al. (2007). Moreover, Barnes et al. (2005) also found that commuter cycling is more prevalent in the U.S. in proximity to the provision of cycle lanes and paths. Mixed evidence is found over preferences for cycling routes, however. Vernez-Moudon et al. (2005) identify in the U.S. that proximity to bike paths encourages commuter cycling, but not the proximity to lanes, whereas Krizek and Johnson (2006) find the opposite case.

Finally, and recently in the UK, a number of longitudinal studies have identified the *causal* importance of policy and infrastructure to active travel by focussing upon selected sites and comparing behaviours before and after various policy and infrastructure interventions. Making use of comparative census data from 2001 and 2011, Goodman et al. (2013b) identify that increases in cycling and walking to work and decreases in car borne commuting took place across six Cycling Demonstration towns (funded between 2005-2011) and 12 Cycling Cities and Towns (funded between 2008-2011) as a result of a variety of initiatives including capital infrastructure investment in, for example, cycle lanes along with investment in promotional and training activity. Moreover, making use of primary data, results from the iConnect study in three selected intervention towns and cities, when compared to comparator groups, found that past-week walking and cycling for transport increased in proximity to the infrastructure after 2 years (Sahlqyist et al. 2013; Goodman et al. 2014). The interventions were the installation of traffic free bridges in Cardiff and Kenilworth, and the development of an informal riverside path into a boardwalk in Southampton.[3]

There are three important features of these studies. The first is that they examine active travel through cycling alongside walking. The second is that they also take account of the recreational context of these activities as well as commuting, whereas much of the above reported literature focusses upon commuting behaviour. Finally, the studies also take account of other leisure time physical activity. The studies show that this latter behaviour also increases more in proximity to the interventions. This is suggestive of complementarity between commuter walking and cycling, walking and cycling of a recreational nature as well as other leisure time activity.

Implicit in the above discussion of the literature are different research designs. Some studies such as Parkin et al. (2008) and Rietveld and Daniel (2004) focus on comparisons of the proportion of cyclists in different aggregate geographical areas. They are necessarily based on secondary data. Other research focusses on the individual (Goodman et al. 2013b; Handy et al. 2010; Sloman et al. 2009) through analysis of secondary data, or data collected through primary survey or the monitoring of users. The research is more localised in nature either through the survey design or the focus on specific interventions.[4] It has been argued, therefore, that future research should focus on changes in behaviour according to environmental and infrastructure changes, *and*, as noted earlier, the relationship between cycling, walking and overall physical activity (See also Yang et al. 2010; Saelens at al. 2003). The former case was first noted by Lawlor et al. (2003) who also argued that research needed to move beyond the specific context, even if this context was essential for a randomised controlled trial, to assess impacts at the *population* level. In the latter case, in a population level study Rasciute and Downward (2010) show that cycling behaviour is conditional on other physical activity. This is not surprising given the insights noted above and, as Litvin et al. (2013) show, that non-travellers tend to have more sedentary lifestyles compared to those who travel, *regardless of the form of travel*. To address these issues, therefore, this paper extends the current research by examining the relationship between individuals' cycling behaviour and access to NCN routes using a nationally representative survey, in the context of their walking and other physical activity, whilst controlling for the typically identified socio-economic characteristics noted above. It should be emphasised, however, that the study is cross-sectional and thus causal insight is limited. Nonetheless, the aim of the study is to offer some generality to the emergent findings from the longitudinal studies in the UK noted above.

3. Data and variables

This research primarily draws upon the Active People Survey (APS). Commissioned by Sport England, with data being collected through random sampling on a rolling monthly basis, and published in annual waves since 2005/6, the APS measures cycling, walking and sports participation, as well as the individual's socio-economic circumstances, for representative samples of individuals for each local authority in England. The current research focusses upon Wave 6, which was collected between mid October 2011 and mid October 2012.[5] As the local authority level is the lowest level of disaggregation that localities can be identified in the APS, this data were then matched at local authority level to the miles of NCN route for each local authority for the commensurate period from data provided by Sustrans[6], as well as local authority population data from the 2011 Census, and the geographic area of the local authority from the Local Government Boundary Commission.[7]

Table 1 provides an outline of the variables employed in the study. The total sample size upon which the alternative models are estimated is n= 22, 845 including both cyclists and non-cyclists of any type. The original sample size for the APS is n = 163,420. However, for the purposes of the current multivariate analysis of cycling, the sample size is reduced for various reasons. The first is because income is only measured for a random sample of 50% of the original targeted sample size in the survey. Consequently, the sampling strategy for the survey involved only asking about income of a random sample of n= 78,807 respondents. In addition, amendments were made to the questions on walking and cycling in this wave of the survey such that subsequent to the first three months of data the actual length in minutes of any form of walking and cycling were collected for the first time unlike previously when the APS had a sole measurement emphasis on trips of at least 30 minutes duration (TNS/BMRB, 2013). This previous situation reflected the fact that the main role of the APS for the commissioner, Sport England, is to provide information on the '1 x 30' indicator of physical activity, which 'measures the percentage of the adult population participating in sport, at a moderate intensity, for at least 30 minutes on at least four days out of the last four weeks equivalent to 30 minutes on one or more day a week' (Sport England, 2013; See also Sloman, et al. 2009). Focussing on data for which this general form of walking and cycling were available, therefore, meant that in excess of 46,000 cases from the total original sample, for example for walking, were also ineligible. Genuine missing values were also present across some of the variables. For example, there were in excess of 10,000 cases for education, and in excess of 4000 cases for work status etc. The very large absolute size of the remaining 'core' data for the modelling across the covariates, coupled with the rolling random monthly sampling and sub-sampling, suggests that the results retain statistical reliability for aggregate inferences.[8]

For each variable three sets of means and standard deviations are presented covering the total sample, and sub samples for non-cyclists and cyclists. For economy of presentation, means are also presented for binary as well as scale variables. In the former case the means should be viewed as providing an indication of the sample proportions of the measured characteristic. Because of the large preponderance of zeros in the cycling data, which as discussed in the introduction is due to the low frequency of *population-level* behaviour in activities such as cycling, a more direct and meaningful comparison across the dependent variables can be made with reference to the sub-sample for cyclists.

INSERT TABLE 1 HERE

The dependent variables each capture an aspect of cycling behaviour. Based on the newly available data in the APS, 'cyctot' measures the total minutes undertaking cycling of *any* sort and *any* duration (i.e. this includes both recreational and utilitarian cycling) in the last four weeks prior to the survey.[9] This was calculated as the product of three variables. A binary variable indicating if *any* cycling had been undertaken in the last four weeks, a variable measuring the number of days in which cycling takes place, and the typical duration in minutes of a cycling session. The remainder of the variables, focus on the number of days in which cycling takes place of at least 30 minutes duration, which was the traditional measurement emphasis of the APS, and which is still investigated. Consequently, the variables 'cyc30days' and 'cyc30daysMI' measure the number of days that any cycling has taken place in the last four weeks of at least 30 minutes duration and either of any intensity or moderate intensity respectively. The intensity of cycling is identified by respondents indicating if the activity raised their breathing rate and if it made them out of breath or sweat. The former indicates moderate intensity and the latter vigorous intensity. The moderate intensity variable is important as it is indicative of whether the cycling activity is contributing to certain typical expressions of health enhancing physical activity guidelines and embedded in the traditional measurement emphasis.[10] The survey also allows these latter variables to be disaggregated by purpose. The equivalent recreation versions of the above variables are 'cyc30daysr' and 'cyc30daysMIr' respectively. Finally, based on the differences between these two sets of variables, utilitarian cycling variables were calculated to measure non-recreational cycling of at least 30 minutes duration and also of at least 30 minutes duration of moderate intensity. The variables are named 'cyc30ut' and 'cyc30utMI' respectively.

The data reveal that on average the total duration of any form of cycling is 512 minutes over the four week period.[11] The data also reveal that cycling of at least 30 minutes duration takes place on approximately 6 days every four weeks with utilitarian cycling of this duration but of moderate intensity on only one day. Recreational cycling of at least 30 minutes takes place on approximately 3.5 days and similarly for such activity at a moderate intensity. Utilitarian cycling of at least 30 minutes duration, but not at moderate intensity, also takes place approximately 3 days every four weeks. This suggests that recreational cycling may be more likely to be of a longer duration and at this duration of a greater intensity. Consequently, it may have more potential to yield health benefits than the equivalent utilitarian travel. Of course the data are inconclusive with respect to shorter utilitarian trips, but this insight does qualify some of the more general pronouncements linking active travel to health and indicates how important recreational cycling may be to health. This is an important point that has recently been identified by Sahlqvist et al. (2015) in which it is argued that recent infrastructure investments in cycling in the UK have had a greater recreational impact. It is important to note, however, as the authors point out, that this could reflect the specific goals of the projects and potential lack of connection to feeder routes for the interventions to allow for effective commuting.

Three other sets of variables are presented in Table 1. These are locality variables, other physical activity variables and socio-economic (control) variables. Comparison of the means of these variables allows some initial analysis of the cycling context and socio economic characteristics of cyclists and non-cyclists, which provides some expectations about possible behaviour to be captured in the ZINB model.

The mean values of the locality variables suggest that cyclists tend to come from local authorities with more miles of NCN route, but also those that are larger in area and which are less populated. This could be related to the potentially greater recreational use of the routes. In the analysis that follows therefore a composite variable measuring NCN route miles per population density is used to capture the likely interacting influences of the locality variables. The miles of available NCN route in the local authority clearly represents the absolute opportunity to cycle for the individual. Population density captures two factors that combine to influence this opportunity. Population size is an indicator of the potential total demand for the NCN routes whilst the geographical area indicates the confines within which this demand can operate. Population density thus gives an indication of the degree of concentration of demand and consequently the potential congestion facing an individual cyclist. This means that NCN route miles per population density represents the opportunity to cycle for the individual constrained by potential use of the route by others. An alternative scaling approach in which NCN route miles and population are both expressed in terms of the area of the locality, would in contrast, not fully capture this congestion effect but focus on the impacts that NCN route miles and population *per se* have on the individual because of the shared denominator. It is not clear that the latter variable in 'isolation' captures a likely influence on an individual. There is also the potential problem that NCN route miles are likely to be highly correlated to the area of the local authority, given that the NCN is a planned intervention to meet the needs of an identified population within its geographic space. This is suggestive of potential problems with multicollinearity in estimating such a model and this was indeed found to be the case. Some experimentation with the modelling, therefore, also supported the choice of the variable and the reasoning above.[12]

Cyclists are also far more active than non-cyclists undertaking almost 30 percent more minutes of walking on average for at least 10 minutes duration and almost 80 percent more minutes of sports participation on average relative to non-cyclists. Cyclists are also more likely to be of white British ethnicity, have an income in excess of the income category £31,999 ≥ £20,800, be in full or part time work or a student, have a higher education and be younger and male. There is less likelihood that cyclists have a long-term illness, but more that they come from households with more adults and children (i.e. families) and have access to more vehicles.

4. Statistical Model: Zero-Inflated Negative Binomial Model.

In order to analyse the relationships between access to cycle routes, physical activity lifestyle and the other variables on the dependent variables noted above, a ZINB model is employed. This model belongs to a class of models that analyse count data.[13] Count data comprises non-negative integers. Clearly the dependent variables measuring the numbers of days of cycling correspond to this categorisation. Whilst the variable measuring the total minutes of any cycling is bounded below by zero, it might be argued that minutes are less meaningfully viewed as integers. However, Wooldridge (2002) has argued that count models can be applied to non-negative continuous variables. Because of the importance of examining general cycling, as distinct from solely that of a fixed duration of at least 30 minutes on different days for the first time in a national survey, this is the approach adopted in this paper.

The most simple count regression is based on the Poisson distribution. The Poisson distribution of a variable 'Y' i.e. for the number of days and minutes of cycling, can be described according to the density,

$$\Pr[Y = y] = \frac{e^{-\lambda}\lambda^{y}}{y!} \qquad (1)$$

Where λ is known as the rate parameter. It is equal to the expected value of the Poisson random variable Y, which can be approximated by the average of the observed values of this denoted as y. A limit of this model is that the mean will be equal to the variance. Consequently, $E[Y] = var[Y] = \lambda$. This is the equidispersion property of the Poisson distribution.[14] The Poisson regression model is then derived from the Poisson distribution by using the exponential mean parameterisation of the relation between the mean parameter λ and the regressors x (Cameron and Trivedi, 2005) as:

$$\lambda_i = e^{x_i' \beta}, i = 1, \dots, N. \qquad (2)$$

Equation (2) shows that the expected counts associated with cycling are conditional on a set of variables , x_i, that is a 1 x k row vector, describing the access to cycling, other physical activity and control variables used in the analysis. The corresponding parameter vector to be estimated is β.Values of particular parameters thus indicate how a unit change in a relevant variable becomes reflected in changes in (the log of) expected counts.[15]

The Poisson regression model is usually too restrictive for count data. In count data the variance usually exceeds the mean, which leads to overdispersion. Overdispersion can arise, firstly, due to unobserved heterogeneity, that is some variability of cycling behaviour other than that observed and measured in the data on the regressors. For example, this could include particular tastes or dispositions for travel to be green, or socially responsible. In this case the model of the rate parameter in the Poisson case is not correctly specified (Cameron and Trivedi 2005). Secondly, overdispersion, may arise because the process generating a decision to cycle, as a first event, may differ from that determining how much to cycle, as a later event. For example, the decision to cycle may be due to supply-side factors such as access to routes or membership of a cycling club, while the frequency of cycling may be due to other factors such as the availability of time because of work and family commitments. Finally, overdispersion in count data may be due to the violation of the assumption of independence of events, that is, that cycle trips depend on previous trips. This is linked to the first context in which heterogeneity is present. Significantly, the first and the third problems may be amended by the Negative Binomial model, while the second problem suggests that a two-part model is needed to capture both aspects of behaviour. It is to account for all of these three problems in the data that the ZINB model is used. Nonetheless, tests of its applicability are also undertaken.

In the first instance, the relevance of the Poisson or the Negative Binomial model can be assessed by testing for the presence of overdispersion. As discussed in the next section, the data support the use of a negative binomial model. The test can be understood by considering the distribution describing the probability of counts according to the Negative Binomial model (Greene 2008; Cameron and Trivedi 2005).

$$p(y_i) = \frac{\Gamma(\theta_i + y_i)}{\Gamma(y_i + 1)\Gamma(\theta_i)} \left(\frac{\lambda_i}{\lambda_i + \theta_i}\right)^{y_i} \left(\frac{\theta_i}{\lambda_i + \theta_i}\right)^{\theta_i} \qquad (3)$$

In equation 3, y_i, as in equation 1, refers to the observed minutes or days of cycling, Γ is the gamma function and θ_i is a parameter that determines the degree of dispersion. For identification it is often assumed to be the same for all individuals and that: $\theta_i = \alpha^{-1}$.

Under this assumption the mean is $E(y_i) = \lambda_i$ as with the Poisson distribution but the variance is $Var(y_i) = \lambda_i \left(1 + \frac{1}{\alpha^{-1}} \lambda_i\right)$. Testing $\alpha = 0$ thus determines the appropriateness of the Negative Binomial model over the Poisson model.

A problem with both the Poisson and Negative Binomial models occurs if the dependent variable is characterised by excess zeros. Failure to account for this would lead to an over-prediction of the number of zeros in the data by the model. As discussed earlier, in survey data such as the APS which has collected data on cycling in the last four weeks, there is a possibility that some zeros are 'absolute', implying that the respondent has never cycled, and some zeros are 'relative', in the sense that that they have not cycled in the last four weeks, but may have done so previously, or will do so if circumstances change. In this context a zero-inflated model is used, which considers the existence of two latent groups within the population: one group has zero counts – but with the two forms of zeros distinguished - and the other group has strictly positive counts. Consequently, estimation proceeds in two parts where a count density $f_2(\cdot)$ is supplemented with a binary process with density $f_1(\cdot)$. If the binary process takes value 0, with probability $f_1(0)$, then y=0. If the binary process takes value 1, with probability $f_1(1)$, then y takes count values 0,1,2,... from the count density $f_2(\cdot)$ (Cameron and Trivedi 2005). The density can then be written as :

$$g(y) = \begin{cases} f_1(0) + \left(1 - f_1(0)\right)f_2(0) & if\ y = 0 \\ \left(1 - f_1(0)\right)f_2(y) & if\ y \geq 1 \end{cases} \quad (4)$$

In our model $f_1(\cdot)$ is a logit model and $f_2(\cdot)$ is a negative binomial density. The first term of the model indicates the probability of observing a zero, that is no cycling, from a binary decision to cycle or not and the second term, which is the joint probability that an individual chooses to cycle rather than not, but currently chooses a zero frequency. The final term of the model then describes that positive counts have a probability given by the joint distribution of an individual choosing to cycle and choosing a non-zero frequency. The first part of the model, thus describes the excess zeros. The zero-inflation in the data can be tested following Vuong (1989) and, as shown in the next section, supports the use of the zero-inflated model.

5. Results and Discussion

Table 2 presents the results of the ZINB model. The heading of each column in the table describes the dependent variable. Two sets of coefficient estimates are then presented, one above the other. The first refer to the relationship between the independent variables and the counts or frequencies of cycling behaviour. The second refer to the relationship between the independent variables and the zero inflation of the dependent variables; in other words capturing the 'shift' from being an 'absolute' to a 'relative' zero in the data for the dependent variable. These latter results yield important information concerning the potential increases in cycling following from a change in each of the independent variables in the sense that they describe how cycling behaviour may vary from being considered never to occur to currently chosen not to occur. As well as the coefficient estimates the asymptotic z-values are also presented, with statistical significance indicated at the levels identified by the asterisks. Each set of results is based on robust estimation of the variances that are clustered around the local authorities as this is the unit in which NCN routes are measured and also the basis of sampling of individuals.

INSERT TABLE 2 HERE

As implied in the previous section, the last rows of Table 2 show that first $\alpha \neq 0$, and hence the variance of the distributions are not equal to the mean i.e. λ. This indicates that the Negative Binomial regression is preferred to the Poisson model. Second, the

Vuong test results indicate that the Zero-inflated Negative Binomial model is preferred to the standard Negative Binomial model.

The results indicate a positive association between miles of NCN route per population density and minutes cycled of any type but also a positive association with the number of days in which cycling is of at least 30 minutes duration for recreational purposes, and recreational cycling of a moderate intensity. The effect sizes are relatively small. A one unit change in the miles of NCN route per population density generates an extra 0.055 minutes of any cycling; an extra 0.0005 days in which cycling of at least 30 minutes duration for recreation purposes is undertaken, and an extra 0.0005 days in which such cycling takes place of moderate intensity.[16] There is no association of the NCN route variable with utilitarian cycling or non-specific cycling that occurs of at least 30 minutes duration of any intensity. This result adds some support to Sloman et al.'s (2009) argument that the NCN routes have a potentially twofold function; they facilitate shorter duration utilitarian travel and longer duration recreational activity. These results, consequently, indicate that the NCN routes may be more likely to contribute to healthy activity in a recreational context, if the focus for health centres upon activity that is at least 30 minutes duration. The results also suggest that the NCN routes could enhance these potential behaviours in that for the same dependent variables there is a negative sign on the zero inflation coefficients. The coefficients are, moreover, significant for all cases except utilitarian cycling of at least 30 minutes duration. Overall the results suggest that increased access to NCN routes is associated with a reduction in the likelihood of individuals never cycling at all, but particularly so in a recreational context.

In the case of the other physical activity variables, the results show that walking is most likely to be associated with changes in cycling of a non-specifically recreational nature, with some evidence that it can encourage the intensity of non-recreational cycling. These results suggest that utilitarian cycling is more closely complementary to walking, and that some health benefits may emerge from the joint behaviours through cycling if this is measured as a moderate intensity activity of at least 30 minutes duration. In contrast, greater minutes of sports participation are associated positively with all cycling activity measures other than utilitarian cycling of at least 30 minutes duration. This suggests a primary complementarity with longer duration, higher intensity recreational cycling, though connection as well with shorter utilitarian cycling. The reasons behind such behaviours could, of course, be linked to the built environment in which mixed use land in more urban areas could help to facilitate walking and cycling, whilst those inclined to participate in cycling for a recreational purpose may also prioritise fitness for their sports activities. Likewise, the zero-inflation results indicate a strong suggestion that walking activity has the potential to have a more general impact on cycling, other than for higher intensity longer duration utilitarian trips, and participation in sport likewise with the exception of any longer duration utilitarian trips. These results provide evidence that cycling is positively associated with, that is complementary to, a more generally active lifestyle. However, there are differences in emphasis. Health enhancing activity, at least as measured by 30 minutes of activity, is more likely to be linked to recreational activity to which NCN routes can service, and be associated with a more sporting set of behaviours. Walking is more likely to be associated with utilitarian trips. There is some evidence that NCN routes can support this activity. It is clear that policy that wishes to promote more active travel and health could focus more on these relatively distinct lifestyle complementarities. It remains, however, that the results also demonstrate in a general large-scale survey context, and as noted in Section 2 for more specific intervention studies in the UK, the potential of linking these three physical activity behaviours together. It remains that more work needs to be done on researching the links between NCN and other route availability and longer-run utilitarian cycling, which is a finding also supported from these intervention studies (Sahlqvist et al., 2015).

Other general results are that having a higher income, working, being a student, having higher education and being male reduce the zero-inflation in the data and are also factors that can potentially increase the probability of cycling. Being male is also associated with an increase in actual counts or frequencies of all types of cycling. Whilst the latter is a common result in the literature for the UK, the general thrust of the results, because of the use of the zero-inflated model, suggests that the potential for additional cycling is connected to lifestyle choice, rather than, say, just concerns with the cost of transport. Nonetheless, as access to vehicles is negatively associated with the number of days in which utilitarian cycling takes place of at least 30 minutes duration, whilst it increases the zero-inflation for such travel but not of moderate intensity, these results suggest that longer-term cycle commuting and its potential is a relatively distinct active travel segment and one which is more likely to be linked to transport cost.

The results also suggest that having a longer term illness is associated with a positive effect on zero-inflation suggesting, perhaps obviously, less potential for cycling. Finally, there are complex household compositional effects detected in the results. Whilst the presence of more children in the household is associated with a reduction in the duration of any cycling, it also decreases the zero-inflation of any cycling and that for a recreational purpose but not utilitarian cycling. This is indicative of children adding constraints to behaviour, but nonetheless retaining the potential to cycle for recreation. The number of adults in the household, however, increases the zero-inflation for cycling of a general and recreational nature, if it is of at least 30 minutes duration, but reduces it for non-moderate intensity utilitarian cycling of at least 30 minutes duration. As there is also some evidence that additional adults in the household are associated with a decrease in the incidence of utilitarian cycling of at least 30 minutes of a moderate intensity, the results suggests that family life could reduce utilitarian cycling, though retain its potential. Overall these results suggest that families as indicated by more adults being present could tend to reduce cycling, though the presence of children can increase the potential to cycle recreationally, whilst partners alone can preserve the potential for utilitarian cycling.

6. Limitations and Implications

The current research has some limitations. The first is that potential omitted variable bias could be present in the results. The current research focusses on the NCN and does not account for the influence on behaviour of other potential cycle routes and opportunities. This means that the impacts of NCN route access could be either overstated or understated. The former would be the case if the (unmeasured) effects of the other routes on cycling would be positive if measured in the analysis and there was a positive relationship between the NCN and other cycle routes. The latter could be the case if there was, for example, a negative association between the NCN and the provision of other cycle routes. Likewise, topographical data along with actual road traffic and congestion statistics, or meteorological measurements were not explicitly accommodated in the analysis. This was because of either lack of availability of some of the data or inability to match this to measurement at the level of at least local authorities. Nonetheless, to the extent that variations in other cycle route availability and traffic and congestion are likely to be relatively constant or enduring features across local authorities, then basing inference on clustered standard errors at the local authority level could ameliorate variations in these effects. Moreover, as the data are randomly collected on a rolling monthly basis, then it is also possible that effects of seasonal and meteorological changes are also controlled for to an extent.

It should also be emphasised that the current analysis cannot claim to be a causal investigation, and this recognition is identified in the commentary on the results above, in which associations are central to the discussions. The potential lack of causality could apply in two senses. The first is that feedback between cycling behaviour and NCN route provision is not formally accounted for in the analysis. However, as the access to NCN routes is an explicitly 'supply side' influence , and it is also measured at a more aggregate level than individual behaviour, then there is some theoretical and statistical basis for cautious claims to identify cause as far as this variable is concerned. The second sense in which causality may not apply is that because this is a large-scale correlational study and not one that monitors actual route usage by individuals, it might be possible that the NCN routes can act as signals to cycle more generally, rather than being the main or only specific geographical conduit for activity, as noted above. It remains however, that such ambiguity over the causal mechanisms of changed behaviour does not undermine the fact that evidence is present in the research that access to NCN routes is associated with more cycling, and that other aspects of the physical activity of individuals influences this behaviour.

With these caveats in mind, the main policy conclusions to draw from the research are that the evidence suggests that NCN route provision can play a part in promoting active travel upon which health and congestion benefits for society can accrue. The current research suggests that these health benefits, being achieved through longer duration and intensity cycling, are more likely with recreational cycling as part of a more sporting leisure lifestyle. Importantly, capitalising on these benefits suggests linking the promotion of interventions like the NCN to broader sport and physical activity promotion as the behaviours are highly complementary.

The research also illustrates that more utilitarian travel is a relatively distinct segment of cycling and that routes are more likely to be linked to shorter trips of lower intensity that are in part driven by transport costs and more likely to be complementary to walking activity. This behaviour can also contribute to health guidelines as indicated by WHO (2010), but of which the current research cannot formally address. Overall, therefore, the results suggests that NCN marketing needs to be nuanced to appeal to these potential segments and, more importantly, that general claims that active travel through cycling *per se* contributes to health subject to further investigation by the large-scale analysis of shorter duration utilitarian cycling. The current research suggests that utilitarian cycling would have largely indirect effects on health in the sense of this being derived from being a part of longer duration physical activity for more generally physical active individuals, though it naturally has an impact on sustainability. This is also an issue that requires further research.

7. Conclusion

This research has used unique large-scale data to examine the impact of access to NCN routes on a variety of cycling behavior in England. Making use of an application of a ZINB model to the data, the research provides evidence that the presence of cycle routes in local authorities is associated with increases in both the total minutes of any form of cycling as well as increases in the number of days on which cycling is undertaken for at least 30 minutes duration, both of a moderate and lower intensity, for recreation purposes. No significant associations are identified for utilitarian cycling of a least 30 minutes duration. The results suggest that the routes could help to facilitate shorter utilitarian cycling as well as longer recreational activity. The analysis of the excess zeros in the data facilitated by the ZINB model also suggests that, with the exception of the measures of utilitarian cycling of at least 30 minutes duration, the presence of NCN routes reduces the number of absolute zeros in the data. It is argued that this reveals their impact on potential cycling behavior.

Differences in the broader physical activity of individuals are also shown to have an association with cycling. Walking is most likely to be complementary to non-recreational cycling, whilst participation in sport is associated with all forms of cycling but not for longer utilitarian trips. This suggests that harnessing the benefits of cycling and the promotion of NCN routes needs to be nuanced more to meet the needs of users and their lifestyles as part of meeting both sustainability and health objectives.

Table 1. Variables

Variable	Definition	Total Mean n=22,845	Total SD	Non-Cyclist Mean n=19,530	Non-Cyclist SD	Cyclist Mean n=3,315	Cyclist SD
Cycling							
cyctot*	Total minutes cycling	74.24	347.00	0.00	0.00	511.62	778.56
cyc30days*	Number of days cycled for 30 minutes	0.93	3.74	0.00	0.00	6.43	7.82
cyc30daysr*	Number of days cycled for 30 minutes for recreation	0.52	2.46	0.00	0.00	3.55	5.55
cyc30daysMI*	Number of days cycled for 30 minutes at moderate intensity	0.61	2.97	0.00	0.00	4.23	6.76
cyc30daysMIr*	Number of days cycled for 30 minutes at moderate intensity for recreation	0.49	2.38	0.00	0.00	3.34	5.44
cyc30ut*	Number of days cycled for 30 minutes not for recreation	0.44	2.70	0.00	0.00	3.02	6.53
cyc30utMI*	Number of days cycled for 30 minutes not for recreation at moderate intensity	0.15	1.51	0.00	0.00	1.02	3.86
Locality							
cycmile	Miles of cycle route in local authority	34.17	48.49	34.01	48.42	35.09	48.89
pop	Population of local authority	156,912	102,729	157,465	102,619	153,652	103,332
area	Area of local authority (square miles)	44,593	58,663	44,334	58,591	46,116	59,068
cycmpa	Miles of cycle route per population density	25.13	78.27	24.86	77.28	26.71	83.84
Physical Activity							
walk10tot*	Total time walking of at least 10 minutes duration	985	1,706	944	1,674	1,224	1,864
Stim*	Total time undertaking sport activity	399	879	358	829	644	1,097
Control							
White	White British or not	0.94	0.25	0.93	0.25	0.95	0.23
Income1[a]	Individual annual income < £10,399	0.12	0.32	0.13	0.33	0.05	0.21
Income2	Individual annual income £20,799 ≥ £10,399	0.19	0.39	0.19	0.40	0.13	0.34
Income3	Individual annual income £31,999 ≥ £20,800	0.14	0.35	0.14	0.34	0.15	0.36
Income4	Individual annual income £41,,599 ≥ £31200	0.13	0.33	0.12	0.32	0.16	0.37
Income5	Individual annual income £51,999 ≥ £41,600	0.08	0.26	0.07	0.25	0.11	0.31

Income6	Individual annual income > £52,000	0.12	0.32	0.11	0.31	0.20	0.40
Working	Full time or part time working or not	0.52	0.50	0.49	0.50	0.69	0.46
Student	Fulltime or part time student	0.05	0.22	0.04	0.20	0.09	0.29
Keephouse	Keeps house or not	0.03	0.18	0.04	0.19	0.03	0.16
Retired	Retired or not	0.31	0.46	0.35	0.48	0.13	0.34
otherwk[a]	Other work status or not	0.08	0.27	0.08	0.28	0.06	0.23
he	Higher education or not	0.38	0.49	0.36	0.48	0.48	0.50
sex	Male or female	0.42	0.49	0.39	0.49	0.60	0.49
age	Age in years	51.41	19.02	52.89	19.08	42.74	16.13
Longill	Long term illness or not	0.31	0.46	0.33	0.47	0.17	0.37
Numadults	Number of adults in respondent household	1.92	0.93	1.89	0.92	2.11	0.93
Numchild	Number of children in respondent household	0.46	0.87	0.42	0.84	0.70	0.99
Vehicle	Number of cars available to the household	1.30	1.10	1.26	1.05	1.51	1.34

* *All in the last four weeks*

a *Base category*

Table 2. Zero-inflated Negative Binomial Analyses: All Sample

	cyctot	cyc30days	cyc30daysr	cyc30daysMI	cyc30daysMIr	cyc30ut	cyc30utMI
cycmpa	0.000468**	-0.0000272	0.000512*	0.000107	0.000544*	0.000204	0.000240
	(3.14)	(-0.11)	(2.18)	(0.57)	(2.16)	(0.59)	(0.30)
walk10tot	0.0000755***	0.0000411***	0.00000625	0.0000446*	0.0000142	0.0000384***	0.0000385
	(5.26)	(3.65)	(0.22)	(2.05)	(0.53)	(3.64)	(0.70)
Stim	0.0000932***	0.0000553**	0.000125***	0.0000762**	0.000113***	-0.000000194	-0.0000387
	(3.80)	(2.88)	(4.23)	(3.17)	(3.91)	(-0.01)	(-0.69)
White	0.0791	0.000433	-0.143	0.0959	-0.125	0.157	1.422**
	(0.70)	(0.00)	(-1.08)	(0.71)	(-0.89)	(1.10)	(2.85)
Income2	-0.0889	0.0484	0.0161	0.0585	-0.00553	-0.0672	-0.244
	(-0.98)	(0.59)	(0.15)	(0.55)	(-0.05)	(-0.67)	(-0.79)
Income3	-0.0611	0.0163	-0.00708	0.0581	-0.0205	-0.0792	-0.131
	(-0.76)	(0.21)	(-0.07)	(0.58)	(-0.20)	(-0.82)	(-0.45)
Income4	-0.112	-0.103	-0.0970	-0.0708	-0.0931	-0.146	-0.179
	(-1.32)	(-1.39)	(-0.98)	(-0.75)	(-0.95)	(-1.52)	(-0.64)
Income5	-0.135	-0.131	-0.133	-0.0709	-0.144	-0.0404	0.125
	(-1.53)	(-1.46)	(-1.17)	(-0.63)	(-1.27)	(-0.35)	(0.49)
Income6	-0.0392	-0.152	-0.123	-0.0821	-0.129	0.0546	0.195
	(-0.45)	(-1.84)	(-1.36)	(-0.87)	(-1.42)	(0.53)	(0.64)
Working	-0.165	-0.0266	-0.169	-0.183	-0.159	0.337*	0.180
	(-1.08)	(-0.25)	(-1.21)	(-1.52)	(-1.17)	(2.19)	(0.57)
Student	-0.679***	-0.258	-0.637***	-0.351*	-0.603**	0.146	0.390
	(-3.77)	(-1.78)	(-3.31)	(-2.04)	(-3.22)	(0.85)	(1.04)
Keephouse	-0.254	-0.0930	-0.00922	-0.0727	-0.0486	-0.187	0.118
	(-1.22)	(-0.46)	(-0.04)	(-0.36)	(-0.24)	(-0.61)	(0.23)
Retired	-0.104	-0.130	0.0908	-0.0675	0.103	-0.0183	-0.0342
	(-0.58)	(-0.91)	(0.51)	(-0.42)	(0.58)	(-0.09)	(-0.06)
he	-0.0270	-0.0402	-0.196**	-0.0520	-0.193**	-0.0745	-0.308
	(-0.48)	(-0.69)	(-2.99)	(-0.81)	(-2.97)	(-1.02)	(-1.20)

	cyctot	cyc30days	cyc30daysr	cyc30daysMI	cyc30daysMIr	cyc30ut	cyc30utMI
sex	0.515***	0.453***	0.546***	0.569***	0.531***	0.173*	0.540**
	(10.42)	(9.45)	(9.76)	(10.40)	(9.47)	(2.50)	(2.78)
age	-0.00708**	-0.00156	-0.00292	-0.00462	-0.00397	0.000285	-0.00581
	(-2.92)	(-0.73)	(-0.93)	(-1.78)	(-1.27)	(0.10)	(-0.83)
Longill	0.0226	0.0679	0.115	0.0878	0.136	-0.00746	0.00965
	(0.32)	(0.98)	(1.48)	(1.19)	(1.78)	(-0.08)	(0.04)
Numadults	-0.0332	-0.00795	-0.00724	-0.0232	-0.0194	-0.0285	-0.251*
	(-1.29)	(-0.31)	(-0.23)	(-0.66)	(-0.60)	(-0.73)	(-2.03)
Numchild	-0.0977***	-0.0847**	-0.0494	-0.0836*	-0.0468	0.00221	0.0424
	(-3.36)	(-3.14)	(-1.45)	(-2.45)	(-1.37)	(0.06)	(0.49)
Vehicle	-0.0278	-0.0543*	-0.0314*	-0.0452*	-0.0258	-0.198***	-0.411***
	(-1.82)	(-2.07)	(-1.98)	(-1.96)	(-1.59)	(-4.09)	(-4.04)
constant	6.377***	1.847***	1.636***	1.790***	1.723***	2.017***	-1.365
	(26.57)	(9.94)	(6.45)	(7.65)	(6.66)	(7.99)	(-1.62)
Zero inflation							
cycmpa	-0.000331*	-0.000537**	-0.000567**	-0.000720***	-0.000580***	0.000449	-0.000558
	(-2.04)	(-3.04)	(-3.23)	(-3.45)	(-3.55)	(1.06)	(-0.27)
walk10tot	-0.0000453***	-0.0000563***	-0.0000400*	-0.0000265*	-0.0000335*	-0.0000711***	-0.000132
	(-4.70)	(-4.80)	(-2.28)	(-2.05)	(-2.13)	(-5.28)	(-0.61)
Stim	-0.000111***	-0.000103***	-0.000106***	-0.000123***	-0.000114***	-0.0000250	-0.000605
	(-5.74)	(-4.46)	(-4.41)	(-5.33)	(-4.78)	(-0.74)	(-0.58)
White	-0.710***	-0.696***	-0.761***	-0.703***	-0.781***	-0.654***	-0.811
	(-7.38)	(-6.29)	(-6.07)	(-5.38)	(-5.88)	(-4.40)	(-1.02)
Income2	-0.0497	0.0166	-0.0212	-0.0311	-0.0508	0.0696	0.118
	(-0.74)	(0.21)	(-0.22)	(-0.32)	(-0.52)	(0.62)	(0.31)
Income3	-0.214***	-0.176*	-0.244**	-0.226*	-0.280**	-0.180	-0.714
	(-3.40)	(-2.42)	(-2.73)	(-2.53)	(-3.08)	(-1.62)	(-1.05)
Income4	-0.271***	-0.322***	-0.431***	-0.404***	-0.463***	-0.129	-0.178
	(-4.12)	(-4.09)	(-4.80)	(-4.59)	(-5.16)	(-1.15)	(-0.41)
Income5	-0.322***	-0.388***	-0.453***	-0.472***	-0.512***	-0.220	-0.424
	(-4.15)	(-4.31)	(-4.18)	(-4.44)	(-4.71)	(-1.83)	(-0.60)

	cyctot	cyc30days	cyc30daysr	cyc30daysMI	cyc30daysMIr	cyc30ut	cyc30utMI
Income6	-0.368***	-0.425***	-0.558***	-0.495***	-0.583***	-0.123	-0.206
	(-5.34)	(-5.34)	(-6.02)	(-5.55)	(-6.33)	(-1.03)	(-0.36)
Working	-0.354***	-0.334***	-0.310**	-0.322**	-0.341**	-0.368*	-0.293
	(-4.19)	(-3.44)	(-2.67)	(-2.82)	(-2.86)	(-2.29)	(-0.59)
Student	-0.671***	-0.545***	-0.587***	-0.456**	-0.607***	-0.712***	-1.433
	(-5.89)	(-3.76)	(-3.59)	(-2.86)	(-3.73)	(-3.57)	(-1.69)
Keephouse	0.00581	0.00279	-0.0407	-0.0796	-0.110	0.154	0.355
	(0.04)	(0.02)	(-0.20)	(-0.40)	(-0.54)	(0.52)	(0.43)
Retired	0.00678	0.0791	0.126	0.113	0.129	0.157	0.214
	(0.06)	(0.59)	(0.84)	(0.74)	(0.83)	(0.80)	(0.31)
he	-0.364***	-0.365***	-0.488***	-0.420***	-0.501***	-0.361***	-1.724**
	(-7.93)	(-7.13)	(-8.16)	(-7.39)	(-8.39)	(-4.24)	(-3.13)
sex	-0.760***	-0.645***	-0.533***	-0.552***	-0.549***	-0.814***	-1.630**
	(-18.57)	(-12.72)	(-9.33)	(-9.96)	(-9.40)	(-10.11)	(-3.08)
age	0.0158***	0.0183***	0.0172***	0.0172***	0.0173***	0.0170***	0.0351**
	(9.12)	(8.84)	(7.31)	(7.41)	(7.32)	(5.69)	(2.61)
Longill	0.492***	0.496***	0.428***	0.401***	0.404***	0.504***	0.848
	(9.41)	(8.26)	(6.16)	(5.99)	(5.74)	(5.10)	(1.75)
Numadults	0.00988	0.0663*	0.116**	0.0989**	0.109**	-0.0821*	-0.531
	(0.39)	(2.26)	(3.23)	(2.78)	(3.00)	(-2.14)	(-1.52)
Numchild	-0.178***	-0.180***	-0.216***	-0.190***	-0.192***	-0.00895	0.156
	(-8.62)	(-7.28)	(-7.25)	(-6.78)	(-6.36)	(-0.25)	(0.87)
Vehicle	0.00167	-0.0350	-0.135***	-0.121***	-0.122**	0.318***	0.340
	(0.08)	(-1.56)	(-3.86)	(-3.34)	(-3.21)	(6.80)	(1.47)
constant	2.698***	2.486***	2.886***	3.005***	3.082***	3.442***	2.019*
	(16.09)	(13.22)	(13.20)	(13.81)	(13.46)	(13.09)	(2.11)
N	22845	22845	22845	22845	22845	22845	22845
H0:alpha=0 χ bar^2(1)	7.6e+06***	7.4e+04***	4.2e+04***	5.5e+04***	4.1e+04***	5.0e+04***	2.2e+04***
Vuong Test (z)	15.79***	18.76***	15.45***	16.05***	15.23***	12.70***	7.95***

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Aldred, R. (2012). "Governing transport from welfare state to hollow state: The case of cycling in the UK." *Transport Policy*, 23: 95-102.

Aldred, Rachel and Katrina Jungnickel. (2014). "Why culture matters for transport policy: the case of cycling in the UK." *Journal of Transport Geography* 34, 78–87.

Barnes, Gary, Kristin Thompson and Kevin Krizek. (2005). "A Longitudinal Analysis of the Effect of Bicycle Facilities on Commute Mode Share." Minneapolis, MN, Active Communities /Transportation Research Group: 16. (Retrieved 15th February 2015 https://www.hhh.umn.edu/centers/slp/pdf/reports_papers/effect_bike_facilities_mode_share.pdf).

Buehler, Ralph and John Pucher. (2012). "Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes." *Transportation,* 39: 409-432.

Cameron, A. Colin. and Pravin K. Trivedi. (2005). Microeconometrics: Methods and Applications. New York: Cambridge University Press.

Caulfield, Brian. (2014). "Re-cycling a city- Examining the growth of cycling in Dublin." *Transportation Research Part A,* 61: 216-226.

Chatterjee, Kiron, Henrietta Shaw, and Juliet Jain. (2013) "Triggers for changes in cycling: the role of life events and modifications to the external environment." *Journal of Transport Geography*, 30, 183-193.

Department for Transport (2012). "Improving Local Transport." https://www.gov.uk/government/policies/improving-local-transport/supporting-pages/encouraging-people-to-cycle (retrieved October 9th 2014)

Dickinson, Janet E., Simon Kingham, Scott Copsey, and Deborah, J. Pearlman Hougie. (2003). "Employer travel plans, cycling and gender: will travel plan measures improve the outlook for cycling to work in the UK?" *Transportation Research Part D,* 8(1): 53–67.

Dill, Jennifer and Theresa Carr. (2003). "Bicycle commuting and facilities in major US cities: if you build them, commuters will use them – another look." *Transportation Research Record,* 1828: 116-123.

Downward, Paul M., Alistair Dawson and Trudo Dejonghe. (2009). Sports Economics: Theory, Evidence and Policy, Elsevier: London.

Downward, Paul M. and Joe Riordan. (2007). "Social Interactions and the Demand for Sport: An Economic Analysis." *Contemporary Economic Policy*, 25(4): 518-537.

Garrard, Jan., Geoffrey Rose and Sing Kai Lo. (2008). "Promoting transportation cycling for women: The role of bicycle infrastructure." *Preventative Medicine,* 46: 55-59.

Goodman, Anna. (2013a) Walking, Cycling and Driving to Work in the English and Welsh 2011 Census: Trends, Socio-Economic Patterning and Relevance to Travel Behaviour in General. PLoS ONE 8(8): e71790. doi:10.1371/journal.pone.0071790

Goodman, Anna, Jenna Panter, Stephen Sharp. and Olgivie, David. (2013b). "Effectiveness and equity impacts of town-wide cycling initiatives in England: A longitudinal, controlled natural experimental study", *Social Science and Medicine,* 97: 228–237.

Goodman, Anna, Shannon Sahlqvist, and David Olgivie. (2014). "New Walking and Cycling Routes and Increased Physical Activity: One- and Two-Year Findings From the UK iConnect Study". *American Journal of Public Health*, 104(9): e38-e46.

Greene, William. (2008). Econometric Analysis (6th ed). London: Prentice Hall.

Handy, Susan L, Yan Xing and Theodore J. Buehler. (2010). "Factors associated with bicycle ownership and use: a study of six small U.S. cities." *Transportation*, 37(6): 967-985.

Heinen, Eva, Bert van Wee, and Kees Maat. (2010). "Bicycle use for commuting: A literature review." Transport Review, 30(1): 105-132.

Krizek, Kevin J. and Pamela Jo Johnson. (2006). "Proximity to trails and retail: effects on urban cycling and walking." *Journal of the American Planning Association*,72(1): 33–42.

Lawlor, Debbie A., Andy R. Ness, Andy Cope, Phil Insall, Adrian Davis and Chris Riddoch. (2003). "The Challenges of Evaluating environmental interventions to increase population levels of physical activity: the case of the UK National Cycle Network." *Journal of Epidemiology and Community Health*, 57, 96-101.

Litvin, Stephen W., Wayne W. Smith and Robert E. Pitts. (2013). "Sedentary Behavior of the Nontravel Segment: A Research Note." *Journal of Travel Research,* 52(1): 131-136.

Moudon, Anne Vernez, Chanam Lee, Allen D. Cheadle, Cheza W. Collier, Donna Johnson, Thomas L. Schmid and Robert D. Weathers. (2005)."Cycling and the built environment, a U.S. perspective." Transportation Research D, 10(3): 245-261.

Nankervis, M. (1999). "The effect of weather and climate on bicycle commuting." *Transportation Research Part A,* 33: 417-431.

Nosal, Thomas and Luis F. Miranda-Moreno. (2014). "The effect o weather on the use of North American bicycle facilities: A multi-city analysis using automatic counts. *Transportation Research Part A,* 66: 213-225.

Oja, Pekka, Sylvia Titze, Adrian Bauman, Bas de Geus, Patricia Krenn P., Bill Reger-Nash B and T. Kohlberger. (2011). "Health benefits of cycling: A systematic review." *Scandinavian Journal of Medicine and Science in Sports*, 21: 496-509.

Ogilvie, David, Matt Egan, Val Hamilton and Mark Petticrew. (2004) Promoting walking and cycling as an alternative to using cars: systematic review, British Medical Journal, 329:763 _

Panter, Jenna and Andy Jones. (2010). "Attitudes and the Environment as Determinants of Active Travel in Adults: What do and Don't we know?" *Journal of Physical Activity and Health,* 7:551-561.

Parkin, John, Mark Wardman and Matthew Page. (2008). "Estimation of the determinants of bicycle mode share for the journey to work using census data." *Transportation,* 35:93-109.

Pikora, Terri, Billie Giles-Cortia, Fiona Bull, Konrad Jamrozik, and Rob Donovan. (2003). "Developing a framework for assessment of the environmental Determinants of walking and cycling." *Social Science & Medicine,* 56:1693–1703

Pucher, John and Ralph Buehler. (2008). "Making cycling irresistible: lessons from the Netherlands, Denmark and Germany." *Transport Reviews*, 28(1): 495-528.

Rasciute, Simona and Paul M. Downward. (2010). "Health or Happiness? What is the Impact of Physical Activity on the Individual." *Kyklos*, 63(2): 256–270.

Rietveld, Piet and Vanessa Daniel. (2004). "Determinants of bicycle use: do municipal policies matter?" *Transportation Research Part A,* 38: 531-550.

Saelens, Brian.E., James F. Sallis, and Lawrence, D. Frank.. (2003). "Environmental Correlates of Walking and Cycling: Findings from the transportation, Urban Design and Planning Literatures". *Annals of Behavioral Medicine*, 25(2): 80-91.

Sahlqvist, Shannon., Anna Goodman, Ashley R., Cooper, and David Olgivie. (2013). "Change in active travel and changes in recreational and total physical activity in adults: longitudinal findings from the iConnect study". *International Journal of Behavioural Nutrition and Physical Activity,* 10(28): 1-10

Sahlqvist, Shannon, Anna Goodman, Tim Jones, Jane Powell, Yena Song, and David Ogilvie. (2015). "Mechanisms underpinning use of new walking and cycling infrastructure in different contexts: mixed-method analysis," *International Journal of Behavioral Nutrition and Physical Activity*, DOI 10.1186/s12966-015-0185-5

Sherwin, Henrietta, Kiron Chatterjee and Juliet Jain. (2014). "An expolartion of the importance of social influence in the decision to start bicycling in England". *Transportation Research Part A,* forthcoming.

Sloman, Lynn, Nick Cavill, Andy Cope, Lisa Muller and Angela Kennedy. (2009). Analysis and synthesis of evidence on the effects of investment in six Cycling Demonstration Towns. Department for Transport, London.

Sport England (2013). A Summary of Sports participation Indicators http://www.sportengland.org/media/112857/summary-of-sport-participation-indicators.pdf (retrieved 29th September 2014).

Sustrans (undated). About the National Cycle Network, http://www.sustrans.org.uk/ncn/map/national-cycle-network/about (retrieved October 9th 2014)

The Guardian (2013). "David Cameron to announce largest ever investment in cycling", http://www.theguardian.com/lifeandstyle/2013/aug/12/david-cameron-largest-investment-cycling (retrieved October 9th 2014)

Tilahun, Y. Nebiyou, David M. Levinson and Kevin J. Krizek. (2007). "Trails, lanes, or traffic: Valuing bicyle facilities with an adaptive stated preference survey." *Transportation Research Part A,* 41: 287-301.

Timperio, Anna, Kylie Ball, Jo Salmon, Rebecca Roberts, Billie Giles-Corti, Louise A. Baur and David Crawford. (2006). "Personal, family, social, and environmental correlates of active commuting to school." *American Journal of Preventive Medicine,* 30(1): 45–51.

TNS BMRB (2013). "Active People Survey 5-7 Technical Report" UK Data Archive Study Number 7369.

Vandenbulcke, Gregory, Claire Dujardin, Isabelle Thomas, Bas de Geus, Bart Degraeuwe, Romain Meeusen, Luc Int Panis. (2011). " Cycle commuting in Belgium: Spatial Determinants and 're-cycling' strategies." *Transportation Research Part A,* 45: 118-137.

Vuong, Quang H. (1989). "Likelihood ratio tests for model selection and non-nested hypotheses." *Econometrica*, **57:** 307-33.

Wardman, Mark, Miles tight and Matthew Page. (2007). "Factors influencing the propensity to cycle to work." *Transportation Research Part A,* 41: 339-350.

Winters, Meghan, Melissa Friesen, Mieke Koehoorn and Kay Teschke. (2007). "Utilitarian bicycling: A multilevel analysis of climate and personal influences." *American Journal of Preventative Medicine*, 32:52-58.

Wooldridge, Jeffrey, M. (2002). Econometric Analysis of Cross Section and Panel Data. Boston: MIT Press.

Yang, Lin, Shannon Sahlqvist, Alison McMinn, Simon J. Griffin and David Ogilvie. (2010). "Interventions to promote cycling: A systematic review." *British Medical Journal* 341, c5293.

Zahran, Sammy, Samuel D. Brody, Praveen Maghelal, Andrew Prelog and Micahel Lacy. (2008)."Cycling and walking: explaining the spatial distribution of healthy modes of transportation in the United States." *Transportation Research. D,*13(7): 462–470

Endnotes

[1] The precise definition of over dispersion is discussed in Section 4.
[2] In this context cycle routes can be either a bike lane, which is a prioritised section of roadway in which motorised transport also travels, whilst a bike path is segregated and dedicated to cycling.
[3] For details of the iConnect study see http://www.iconnect.ac.uk/ (retrieved February 16th 2015).
[4] This literature focusses on quantitative evidence Chaterjee et al. (2013) review and provide qualitative evidence on cycling behaviour.
[5] The aim was to match as closely as possible population estimates.
[6] The data were provided by Dr Andy Cope, to whom we express our gratitude, and covered the operational length in miles of the NCN in each local authority in England.
[7] http://www.lgbce.org.uk/ (retrieved October 9th 2014)
[8] It was not attempted to test for the sensitivity of the results to sample size because omitting the variables that affected sample size failed exclusion restriction tests and would mean that the model suffered from omitted variable bias. The following examples of chi square (p value) tests of not including some variables like income $\chi2$ (10) = 44.89 (0.000) and education $\chi2$ (2)=65.03 (0.000) illustrate this.
[9] These data are collected by different sets of old-form and new-form questions. Consequently the actual minutes of recreational and utilitarian cycling cannot be identified. This distinction is only possible still for the questions that capture at least 30 minutes of activity.
[10] For example, The WHO recommends as a minimum that adults aged between "18–64 years should do at least 150 minutes of moderate-intensity aerobic physical activity throughout the week, or do at least 75 minutes of vigorous-intensity aerobic physical activity throughout the week, or an equivalent combination of moderate- and vigorous-intensity activity." (WHO 2010, p8). Thirty minute sessions are often recommended (See Sport England, 2013; and NHS guidelines, http://www.nhs.uk/Livewell/fitness/Pages/physical-activity-guidelines-for-adults.aspx) and this is what motivated the development of the APS. However it should be recognised that ultimately the benefits from aerobic activity can also be gained from bouts of activity of at least 10 minutes duration (WHO 2010). The walking variable included in the study recognises this (see Table 1).
[11] Clearly the cycling variables are indicative of skewed distributions, which help to explain the importance of the modelling strategy.
[12] A number of statistical factors also supported the use of the composite measure of NCN route miles per population density. Experimentation with including the locality variables separately suggested that this did not affect the interpretation of the results as given in the paper. For example an equivalent cyctot frequency regression including the variables separately yielded coefficient estimates (p values) of: cycmile 0.0039139 (0.000); area -3.05E-0.06(0.000); and pop -4.66E0.07(0.089). This is in keeping with the descriptive and regression insights noted in the text. These results suggest that a rise in NCN route miles increases cycling. A rise in area is associated with a reduction in cycling, but the same is also the case for a rise in population. These two latter results mean there could be a trade-off of between these influences overall if it is changes in population density that influence cycling. Moreover, in the zero-inflation regression only the population variable is individually significant with an estimate (p value) of 9.16E-07(0.000). However, in this model, in which all variables are included separately, a joint test of the significance of the three variables across both the frequency and zero-inflation equations rejects the null hypothesis that they are *jointly* insignificant with $\chi2$ (6) = 50.87 (0.000). This suggests that the variables do interact in their influence and that their separate use means that the regressions suffer from multicollinearity. Pairwise correlations suggest this with a correlation of 0.8423 between cycmile and area and 0.2474 between cycmile and population. This suggests strongly, as indicated in the text, that NCN route miles are closely linked to the area in which they are located. Estimating an alternatively scaled specification including cycmile/area and population density is thus also affected badly by multicollinearity because there is a pairwise correlation of 0.5402 between the two variables. Consequently only cycmile/area is statistically significant in the frequency equation with a highly inflated estimate (p value) of 89.95623 (0.001) and yet a joint test of the significance of this variable *and* population density across both equations rejects the null hypothesis that they are jointly insignificant with $\chi2$ (4) = 12.17 (0.0161).
[13] Like other limited or discrete dependent variable models, count models are nonlinear in nature.

[14] Consequently the homoscedasticity assumption is violated as any factor that affects the mean will affect the variance too. Some inefficiency relevant to simple linear estimators is thus a result. However, the key properties of the dependent variable are captured more fully.
[15] Assuming that the observations are independent, the maximum likelihood estimator is used to estimate the parameters.
[16] To put these effects into comparison the effects sizes for being male compared to female suggest approximately an additional 83 minutes of any cycling; an extra half a day of recreational cycling of at least 30 minutes duration, or of recreational cycling of this duration and also moderate intensity.