Article:

Parkin, John, Wardman, Mark and Page, Matthew (2008) Estimation of the determinents of bicycle mode share for the journey to work using census data. Transporation, 35 (1). pp. 93-109. ISSN 1572-9435

https://doi.org/10.1007/s11116-007-9137-5

Published paper
Parkin, J., Wardman, M., and Page, M. (2008) Estimation of the Determinants of Bicycle Mode Share for the Journey to Work using Census Data. Transportation, 35, pp.93-109.

Estimation of the determinants of bicycle mode share for the journey to work using census data

John Parkin[1], Mark Wardman[2] & Matthew Page[3]

1 The University of Bolton, Deane Road, Bolton, BL3 5AB, UK, j.parkin@bolton.ac.uk +44 (0)1204 903027, +44 (0)1204 399074. Corresponding Author.

2 University of Leeds Institute for Transport Studies, Leeds, LS2 9JT, UK, m.r.wardman@its.leeds.ac.uk

3 University of Leeds Institute for Transport Studies, Leeds, LS2 9JT, UK, m.page@its.leeds.ac.uk

Keywords

Bicycle, journey to work, logistic regression model, census, travel demand modelling

Abstract

A model is presented that relates the proportion of bicycle journeys to work for English and Welsh electoral wards to relevant socio-economic, transport and physical variables. A number of previous studies have exploited existing disaggregate data sets. This study uses UK 2001 census data, is based on a logistic regression model and provides complementary evidence based on aggregate data for the determinants of cycle choice. It suggests a saturation level for bicycle use of 43%.

Smaller proportions cycle in wards with more females and higher car ownership. The physical condition of the highway, rainfall and temperature each have an effect on the proportion that cycles to work, but the most significant physical variable is hilliness. The proportion of bicycle route that is off-road is shown to be significant, although it displays a low elasticity (+0.049) and this contrasts with more significant changes usually forecast by models constructed from stated preference based data.

Forecasting shows the trend in car ownership has a significant effect on cycle use and offsets the positive effect of the provision of off-road routes for cycle traffic but only in districts that are moderately hilly or hilly. The provision of infrastructure alone appears insufficient to engender higher levels of cycling.

1.0 Policy and trends

The UK National Cycling Strategy (NCS) (DfT, 1996) set a target of quadrupling the number of cycle trips from a 1996 base by the year 2012. The strategy has been superseded by The Transport White Paper of Summer 2004 (DfT, 2004), which contains a policy aim over the ensuing two to three decades of increasing cycling by making it more convenient, attractive and realistic for short journeys, especially those to work and school. The "one size fits all" NCS target has been abandoned, and local authorities must set their own targets. Local target setting will demand a realistic estimation of potential increases in cycle use in an area and is required by government to be monitored. The estimation of realistic targets for increasing cycle use requires knowledge of the determinants of cycle use, and this paper makes a contribution.

The overall percentage of people that use the bicycle for the journey to work in England, Wales and Scotland from the UK 2001 census is 2.89%. This compares with 2.97% in 1991 and 3.76% in 1981. There appears to have been an arrest in the decline of bicycle use for the journey to work that took place during the 1980s, and this may be because a residual level of bicycle use has been reached or because of the success of promotion measures which have prevented further decline. Parkin (2003) provides a full discussion of the pattern of changes in cycle use using census data for the years 1981, 1991 and 2001 and shows that fourteen districts[1] out of the 376 in England and Wales in the 2001 census have percentage point increases in cycling greater than 1% compared with the 1991 census. Seven of these are London boroughs and this suggests that dense urban areas may have more potential for growth in cycling, perhaps because of shorter trip lengths and suppressed car ownership, parking problems and congestion. The historically higher levels of cycling in

[1] A district comprises a local authority area, and may be predominantly rural or urban in nature.

the drier, flatter eastern regions[2] of England demonstrated the largest declines in the 1980s, but cycling levels in these areas were more stable across the 1990s.

There remains a significant variation in use of the bicycle for the journey to work across England and Wales. Table 1 presents the distribution of the proportion cycling to work from the 2001 census.

Table 1 Inserted here

There are twenty-nine districts (7.7%) with a journey to work proportion by bicycle greater than 6.00%. Of these, seventeen are located in the East of England. The ancient university cities of Oxford (16.22%) and Cambridge (28.34%) have notably large proportions that cycle to work. The warmer districts of the South coast (Gosport,11.44%, Portsmouth, 7.59%) and of the South West (Isles of Scilly, 15.59%, Cheltenham, 7.55%, Taunton Deane, 7.45%, Sedgemoor, 7.05% and Gloucester, 6.52%) are well represented. Other districts with in excess of 6.00% include Vale of White Horse in Oxfordshire (7.52%), the northern districts of Crewe and Nantwich (7.58%) and Barrow-in-Furness (6.35%) and the London Borough of Hackney (6.83%). While the East of England may display geography most conducive to cycling, the variation in the level of cycling is not fully explained by topography and climate and most certainly merits further investigation.

This paper presents an aggregate model that explains the proportion of journeys to work in the 8800 English and Welsh electoral wards[3] that are by bicycle in terms of relevant

---

[2] England comprises nine Government regions. Parts of Yorkshire and Humberside, the East Midlands and the East Region are generally very flat.

[3] Wards are electoral units within a district with mean size of 17 hectares. The 50th percentile ward population aged 16 to 74 is 3469, with the 10th percentile and 90th percentiles being 1402 and 8660. The inter-quartile range is from 1940 to 5582. The census data for Scotland and Northern Ireland is collected and stored in different ways than for England and Wales using different geographical units. Data for the journey to work in Scotland include journeys for education for those aged 16 and over. Some of the explanatory variables are not available in the same form in Scotland and Northern Ireland as in England and Wales. For these reasons the study has been limited to England and Wales.

socio-economic, transport and other physical determinants in order to forecast the potential level of bicycle use for the journey to work based on policy interventions and changed socio-economic factors.

Section 2 sets out the research need and Section 3 details the sources and measurement of determining factors and describes the structure of the model. Section 4 presents the results from the analysis and Section 5 provides forecasts for potential levels of bicycle use for the journey to work. Section 6 provides a discussion and conclusion.

2.0 The research context

The ready availability of socio-economic and distance to work census data in machine readable format has allowed for an analysis to be undertaken at the level of the whole population, not a sample of the population, and at the relatively fine level of the 8800 wards in England and Wales.

Other transport and physical data that is now also available on a geographically comprehensive basis includes road condition and road length, hilliness and weather. These data represent factors that may influence cycle use and, at the aggregate level, are representative, may be re-measured and are valid and reliable.

The literature demonstrates wide use of disaggregate modelling techniques to explore the mode and route choice decisions relating to cycling (e.g. Bovy and Bradley, 1985; Hopkinson and Wardman, 1996; Wardman et al., 1997; Ortúzar et al., 2000; Stinson and Bhat, 2004; Stinson and Bhat, 2005; Moudon et al., 2005; Plaut, 2005; Tilahun et al., 2006; Wardman et al., 2007). These studies have revealed a comprehensive set of variables which are relevant to use of the bicycle and include socio-economic, geographic and transport related variables. Socio-economic variables which have been found to be relevant include: age; sex; car ownership; income; extent of higher or further education; ethnicity; household size and marital status; type of employment; and experience of cycling and engagement in other physical activity and exercise. Geographic variables include: journey distance; home location (urban versus rural); seasonality and weather variables (rain and wind); and type of neighbourhood (car oriented or not). Transport system variables include:

type of provision for cycle traffic; surface roughness; volume of motor traffic; parking facilities and other journey end facilities (changing rooms and lockers); and characteristics of public transport alternatives.

Disaggregate modelling is powerful for examining travel choices in detail because it links behaviour to a wide range of causal factors and, particularly, explores the intervening effects on demand responsiveness of a wide range of socio-economic features of decision makers and their trip characteristics. Disaggregate methods also facilitate the collection of data specifically for the purpose of modelling, rather than relying on secondary data. Notable in this respect are their ability to exploit Stated Preference (SP) data which has particular attractions when examining cycle facilities which do not yet exist and in controlling the experimental choice context. Researchers have been able to make good use of the capabilities of SP surveys to differentiate between types of provision for cycle traffic, for example provision segregated from motor traffic, and different degrees of provision and width available within general traffic streams.

While we do not wish to denigrate the significant contribution of disaggregate modelling, we note that it can be challenging to express to respondents the nature of a travel condition and a change in that condition through an SP survey. Even with a seemingly appropriate use of display material, measures of variation in the level of physical and transport factors may have little or no meaning. For example, surface roughness may be a variable of interest but there is no accurate way of knowing how a respondent has judged the difference between the scenarios presented. Similarly, the precise benefits of improved cycle facilities might have to be experienced to be appreciated.

A particular issue is that there is no guarantee that respondents behave in accordance with their stated preferences, not only because of incentives to response bias but also because SP does not reflect real-world habitual behaviour associated with information acquisition and behavioural change. As a complement to existing disaggregate modelling, we present here an aggregate model which uses objective measures for relevant variables. Aggregate modelling, however, brings with it its own deficiencies and these include the identification of correlations which may be spurious and which do not actually represent causal behaviour.

There are two readily available sources of revealed preference (RP) data in Great Britain. Data about individuals' actual choices for the journey to work has been routinely collected for many years as part of the UK government's National Travel Survey (NTS). In addition to spatial effects due to differences in person, trip and location characteristics, this extremely large data set can also explore inter-temporal variations and trend effects. NTS data, supplemented with a range of SP exercises, has been the subject of discrete choice modelling to explain cycle use for the journey to work (Wardman et al., 2007). The research reported here exploits the opportunities offered by secondary data available at an aggregate and geographically specific level through the census and other sources, supplemented by a survey based exercise relating to the perceived risks of cycling (Parkin et al., 2007), and therefore, in contrast, it re-visits aggregate models of cycling behaviour.

Previous studies using aggregate models to consider cycling have been constructed using United States census data for the 284 metropolitan statistical areas (Baltes, 1996), 18 cities (Nelson and Allen, 1997), and 43 large cities (Dill and Carr, 2003). The first aggregate study in the United Kingdom was undertaken by Waldman (1977) and covered 195 urban district areas. This was followed by a much smaller study by Ashley and Banister (1989) which considered three districts in Greater Manchester. Recent work in The Netherlands (Rietveld and Daniel, 2004) covered 103 Dutch Municipalities.

These studies have revealed some important attributes relating to cycle mode choice as being: sex, car ownership, age, proportion of students within the population, ethnicity, socio-economic class and income. In addition to these, other physical variables of relevance have been found to include journey distance, the degree of urban density and weather attributes, particularly mean temperature and rainfall and, very significantly, hilliness. Bicycle facilities in the models have been specified in various ways, including detailed surveys to determine, for example, the required stop frequency on a journey (Rietveld and

Daniel, 2004), miles of bicycle paths per person (Nelson and Allen, 1997) and state spending per capita on cycling (Dill and Carr, 2003).

The UK census provides information on the proportion who cycle to work. Data is available at the relatively fine level of the census ward and hence, even for just one census year, this constitutes a large amount of data. The model presented here adds to previous aggregate modelling by using a finer level of geographical detail and uses data for the whole of England and Wales, rather than a sample. It includes the refinement of a saturation level estimated by the modelling process and includes a wide range of objectively measured variables of some detail for person type, transport and physical factors. Tests for interactions between variables have been carried out.

3.0 Data sources and structure of models

Table 2 summarises the socio-economic, physical and transport system independent variables that were tested in the modelling. Each category is discussed in turn below. The dependent variable that we wish to explain is the proportion that cycled to work at the level of the 8800 wards in England and Wales as reported in the UK 2001 census. The independent variables may pertain to ward, district or regional level.

Table 2 inserted here

3.1 The Socio-economic variables

The socio-economic variables used in the modelling are those most expected to influence transport mode choice and include sex, ethnicity, socio-economic classification, age and level of qualification. Various measures for car ownership were constructed and the measure finally adopted is the number of cars per employee in the ward.

Indices of deprivation were used to proxy income, which is not covered in the census, and in England this is created from a basket comprising numbers of adults and children on income support, job seekers allowance, those with income less than 60% of the mean who are in receipt of Working Tax Credit and Disabled Person’s Tax Credit and asylum seekers in receipt of subsistence and accommodation support. The Welsh index has similar components but excludes asylum seeker numbers. The Welsh and English indices are constructed in different ways and hence their means and ranges are different.

3.2 Physical Factor Variables

Distance travelled to work place is taken from the 2001 census and calculated based on a straight line between the centroid of the postcode of residence and the centroid of the postcode of the workplace[4] as stated on the census return. 23.3% of journeys to work are less than 2km, 46.6% less than 5km and 67.8% are less than 10km. The total district population (all people aged 16-74) is used in combination with the area of the district to create a measure of the population density and is a good measure of the degree of urbanisation of a district.

Rainfall, temperature and sunshine data have been taken from the Meteorological Office web site (Meteorological Office, 2004). Wind speed has been estimated using the British Standard Wind Code (British Standards, 1997) for structural engineering design. All climate data has been aggregated to the six meteorological regions of the UK.

Topographical data is available through the Countryside Information System (Defra, 2003) for each 1km square of the UK. The downloadable software and complementary datasets can be interrogated using user-specified areas to determine the number of kilometre squares of a particular mean slope (to the nearest 1%)[5]. Two measures for hilliness were tested in the modelling: one based on the proportion of 1km squares in a district with a mean slope of 3% or greater, the other based on a mean slope of 4% or greater. The measure for a mean slope of 3% or greater provided the more significant correlation.

---

[4] A Large User Postcode is one that has been assigned to a single address due to the large volume of mail received at that address. A Small User Postcode identifies a group of delivery points. On average there are 15 delivery points per Postcode, however this can vary between 1 and 100. There are 1.71 million postcodes in the UK. The travel to work distance is calculated to the nearest 1km.

[5] The mean slope for a kilometre square is determined by passing a 3 x 3 operator (grid) over a 20 x 20 matrix within each kilometre square column by column and row by row. The 3 x 3 operator determines the slope at the centre point of the matrix by calculating the change in slope in both orthogonal directions for the surrounding matrix points and then averaging.

The measure for hilliness relates to the general topography of a district and not specifically to the hilliness of routes within the district. The measure adopted will, however, be related to the hilliness of routes and, perhaps more importantly, to the potential behavioural response to cycling. We found no correlations between hilliness and population density, which might have suggested a different effect of hilliness in rural and urban areas.

3.3 Transport system variables

The dependent variable is the proportion that cycle to work, and other modes do not explicitly appear in the model. It is important, however, not only to model the transport attributes of the mode being considered but also, so far as is possible, other relevant transport system and competing mode attributes.

The most heavily used competing mode is the car and this is represented, as discussed above, in terms of car ownership as a socio-economic variable. The level of use of the car is also an important variable to consider. A high level of car use will create busy roads and conditions in which it is potentially less desirable to cycle. A measure of “transport demand intensity” relevant to commuting has been derived from available aggregate data based on the number of workers in the district divided by the total road length in an area. This has not been adjusted to account for the level of public transport availability and cost in a district, as such variables are not available at a district level. Transport Demand Intensity provides a measure of the condition of the infrastructure for cycling and hence its potential effect on cycling. It is different than the previously defined variables for population density, which measures the degree of urbanisation, and car ownership, which measures the effect of availability of a competing mode.

Bicycles are generally un-sprung and hence reflect through to the rider in a direct way deficiencies in the carriageway. Measures for the quality of highways have therefore been adopted to account for the effect of roads on both the comfort and effort of cycling. These have been taken from Audit Commission Best Value indicators and comprise the proportion of road length with “negative residual life” or a defect score higher than 70. These are roads that are deemed to have failed. This proportion is a measure of the proportion of highway that needs some remedial treatment. Two measures are available for all districts at an aggregate level: one for principal roads and the other for non-principal roads.

The level of cycle use can be expected to be related to the level of infrastructure provided for cycling. Fifty three districts have data available to a reasonable standard of cartography detailing facilities for bicycle traffic, including the length of sign-posted bicycle route, length of traffic free path, length of traffic free path adjacent to a highway, and length of bicycle lane and length of bus lane on highway. These districts comprise the thirty-three London Boroughs and also Basildon, Blyth Valley, Bradford, Bristol, Cardiff, Chelmsford, Doncaster, Gateshead, Leeds, Leicester, Manchester, Newcastle-upon-Tyne, Newcastle-under-Lyme, North Tyneside, Rotherham, Sheffield, South Tyneside, Stoke-on-Trent, Wansbeck and York. The fifty-three districts represent 1,111 of the 8,800 wards in England and Wales and cover 9 million people (24%) aged 16-74. The lengths of different types of route from the mapping have been converted to proportions of the total route network in the district. A dummy variable is applied to account for districts which do not have mapping data[6] and the model is estimated for all 8800 wards.

Following from the recommendation of Waldman (1977) that further work be undertaken on measuring the perception of risk of cycling as a determining choice factor, Parkin et al. (2007) undertook analysis at the individual level to assess the perception of risk in different cycling circumstances and then applied this at a district level. A “probability of acceptability” of cycling within the district boundary for districts with data available from mapping on bicycle facilities was derived as a measure for risk.

3.4 The structure of the model

The choice between cycling and not cycling to work may be expressed in the form of the logit model as in Equation 1:

$$P_i = \frac{1}{(1+e^{-Z_i})} \quad \text{Equation 1}$$

---

[6] A value of unity is ascribed to districts with no mapping data and zero otherwise.

P_i is the proportion of individuals in ward *i* who cycle to work and Z_i represents the relative attractiveness of cycling. The most practical form of the model for estimation is the Berkson-Theil transformation into a model, as in Equation 2, that can be estimated by ordinary least squares.

$$\ln[\frac{P_i}{(1-P_i)}] = Z_i = \beta_1 X_1 + \beta_2 X_2 ..... + \beta_n X_n + \varepsilon \qquad \text{Equation 2}$$

The β's denote the relative importance of each of the explanatory variables *X*, with up to n such variables, and *ε* is an independently distributed random error with mean zero.

The logit function's upper limit is 100%. It is not realistic to expect an upper limit for the proportion cycling to work of 100% since there will be sections of the working population that will never cycle, either because they are unable to, or because of distance or other physical constraints such as having to carry items needed at work. Introducing a saturation level (S), to be directly estimated in the modelling process, to allow the upper level of cycling to be less than a 100% share, leads to an amended form, a logistic regression model, as follows:

$$P_i = \frac{S}{(1 + e^{-Z_i})} \qquad \text{Equation 3}$$

Equation 3, with Z_i as defined in Equation 2, is used for modelling the variation in bicycle use between wards and is estimated using non-linear least squares. The basic unit of analysis is the ward and these have unequal sizes in terms of worker population. As a consequence, and to address the issue of heteroscedasticity resulting from this unequal size, we have weighted the observations (Pindyck and Rubinfeld, 1991, page 262).

4.0 Results

The model presented in Table 3 is for all England and Wales. The R-squared value is high (0.816) and demonstrates that the model is explaining a good proportion of the variation on cycle use for the journey to work. The model has estimated a large number of significant (that is a t-statistic greater than 2, with larger values indicating a greater precision in the coefficient estimate) and right sign effects for a wide range of variables. There are no large correlations between the coefficient estimates of the variables in the model. The use of weights had only minor impact on the coefficient estimates and t-statistics. Although correlation amongst the error terms was identified, the degree of correlation was low (r=0.34).

Insert Table 3 here

The estimate for the saturation constant, *S*, indicates an upper limit to proportion that would cycle to work of 43%, and it is estimated extremely precisely. This is larger than the ward with the highest proportion of bicycle to work journeys in Cambridge of 35%, but is a level reached in some Dutch towns. The advantage of the use of a saturation level is that estimates of the proportion cycling are constrained to an appropriate upper bound and the value estimated in this model is an important numerical finding.

4.1 Socio-economic determinants

As is expected, based on monitoring evidence (for example McClintock and Cleary, 1996), wards with higher proportions of males demonstrate a greater level of cycling to work. A higher proportion of non-white residents is linked with a smaller proportion of employees cycling for the journey to work and demonstrates that cultural norms with respect to cycling may be different across ethnic groups within society. This is supported by survey work in London (LRC, 1997).

As the number of cars per employee rises, so the proportion that cycle for the journey to work falls and this finding accords with expectation and with observed trends in cycle use. Piecewise estimation showed monotonic increases in the negative value of the coefficients estimated to each discrete level of car ownership and this strongly confirmed not only the negative effect but also justified the linear formulation adopted.

The lowest socio-economic classes, 5 “lower supervisory and technical”, 6 “semi-routine occupations” and 7 “routine occupations”, form the base against which other socio-economic classes are compared. All other socio-economic classes bar 1.2 “Higher professional"[7] show a reducing effect on cycling to work for higher proportions in those classes.

Higher income deprivation scores for England and Wales are linked with lower proportions that cycle for the journey to work[8]. This finding indicates that lower income has the effect of lowering the proportion that cycles to work. Income deprivation might also be acting as a proxy for crime, safe storage, bicycle availability and image issues.

4.2 Physical and transport system determinants

Journey to work distance is represented in the model by the bands '2km to less than 5km' and '5km to less than 20km'; no other bands demonstrated significant coefficients and were omitted and form the base against which the bands which remain in the model may be judged. The increasingly negative coefficient with rising distance for the bands which remain in the model confirms the disutility, in terms of time and effort, of commuting longer distances. A variety of forms for representing distance were tested, including use of average distance and average distance for all journeys longer than 2km, but coefficients were not of correct sign.

The variable representing the intensity of transport demand has a negative coefficient which confirms the commonly held view that larger traffic volumes are linked with a lesser willingness to cycle. This aggregate finding confirms research at a disaggregate level (e.g. Landis et al., 1997; Guthrie et al., 2001; Parkin et al., 2007) that higher volumes of traffic reduce the level of service for cycling.

---

[7] Note that "higher professional occupations" cover all types of higher professional work , whether occupied by employers, the self-employed, or as an employee. The other sub-category within "higher managerial and professional occupations" is for large employers and higher managerial occupations, both of which are occupations usually within larger organisations than "higher professional occupations".

[8] Note that the difference in magnitude of the coefficient estimates results from the different scales used for England and Wales: when this is accounted for, income deprivation has approximately the same magnitude of effect in each country.

An increase in population density has the effect of increasing the likelihood of cycling for the journey to work. This result appears reasonable given that cycling can be expected to be more attractive in more tightly packed and localised neighbourhoods because of increased parking problems and because the finer grain will be less conducive to motorised travel, a finding confirmed by Cervero and Radisch (1996). This effect could stem from journey distance, but it was not highly correlated with population density. The coefficient is small but this is due to the large average size of the variable.

It is interesting to note the effect of the measure for the condition of principal and non-principal highway pavement: the higher the defects score, the lower the proportion that cycle for the journey to work. The indication from this model confirms a view that poorly maintained highways are a deterrent to cycling since they are both less pleasant to cycle along and also take a greater amount of energy to traverse than well maintained roads. This finding confirms disaggregate modelling undertaken by Bovy and Den Adel (1985) and survey work by Guthrie et al. (2001).

Hilliness is, as expected, a very significant indicator of proportion that cycle to work. The elasticity for hilliness is high at -0.893 [9] and indicates that a 10% increase in the hilliness proportion is associated with an 8.93% reduction in proportion cycling for the journey to work[10]. This confirms the powerful effect of hilliness found by Waldman (1977) and detected even in modelling of variation in cycle use in the famously flat Netherlands by Rietveld and Daniel (2004). A significant advantage of modelling aggregate data for geographically specific areas means that hilliness may be appropriately assessed and creates a significant advance over disaggregate modelling where the hilliness effect has not been modelled successfully (e.g. Wardman et al., 2007).

---

[9] The elasticity is determined about the mean value for the variable with all other variables held at their mean values.

[10] The mean hilliness value is 0.67 and a 10% change about this mean represents only relatively subtle changes in hilliness, for example moving from Camden in inner London to Watford on the outskirts of London, or equivalently in the Midlands, moving from Coventry to North Warwickshire.

Rainfall has a relatively high negative impact (elasticity of -0.665) and this is despite its measurement at regional rather than ward level. Temperature also has a high elasticity (+0.703), with, as would be expected in a temperate climate, higher mean temperatures being linked with a greater proportion cycling to work.

The variable for the acceptability of cycling, which has values based on a logit curve determined from the proportion of route in an area that is traffic free, was not found to contribute to the powers of explanation of the model in the manner expected and this could be linked with the limited range of the variable (0.68 to 0.73) because it was determined for average district conditions. However, the variable for the proportion of route that is off-road had a positive coefficient indicating that a larger proportion of route that is off-road is linked with a higher proportion that cycle to work. The elasticity is, however, small (+0.049) and suggests that a large quantity of off-road cycle route building would be required to stimulate only a modest increase in use of the bicycle to travel to work. The proportion of route that has bicycle and bus lane did not have a significant coefficient and was eliminated from the model. The disaggregate modelling in Wardman et al. (2007) demonstrates a forecast 55% increase in cycling for the infeasible scenario of a complete network of segregated cycle routes. The elasticity from the aggregate modelling presented here challenges such high forecasts. It should be noted that if simultaneity bias were present then the aggregate forecasts themselves would be too high.

4.3 Implications for policy

An important issue, which aggregate modelling can provide valuable insights into, is whether the impact of measures to increase cycling are critically dependent upon factors largely outside the control of policy makers, such as hilliness and climate. It might be hypothesised that, for example, there is little point in investing in cycle facilities in hilly areas since the hilliness is such that only a ‘hard core’ of people would cycle, and improving facilities would not overcome this major deterrence to the general population. Interactions were modelled between hilliness and distance, highway condition, transport demand intensity, population density and provision of off-road routes, but none were found to be significant and this suggests that, for example, the effect of providing off-road routes is similar irrespective of hilliness.

Insofar as the physical variables are concerned, the model has estimated coefficients that reasonably confirm expectations. The mix of socio-economic variables that remain significant has painted an interesting and complex picture of ethnic origin, socio-economic class, car ownership and income all playing a part in the level of cycling to work. The model has identified the very significant physical effects of hilliness and indicates that modelling the impact of the consequences of hilliness, that is the additional effort and time required to cycle, needs to be taken seriously in mode choice modelling.

It is instructive for a moment to consider the variables that do not appear in the model. The variables for wind and sunshine were highly correlated with the saturation level, and were therefore eliminated, because they exhibit little variation across wards. The variable designed to measure risk and based on the proportion of traffic free cycle route was not significant, and neither were variables for other cycle infrastructure features including signed cycle routes, cycle routes adjacent to the carriageway and cycle and bus lanes on the carriageway. On the one hand the absence of these variables implies their lack of importance so far as providing reasonable infrastructure for cyclists is concerned; on the other hand, there could be a need for further data collection across a wider range of local authority areas to increase the proportion for which these data are available. We had appropriate mapping data covering 24% of the population.

5.0 Use of the model as a prediction tool

The model has been used to forecast changes in the proportion that would cycle to work based on the important background trend in car ownership and variations in the proportion of route length that is off-road and the percentage of the non-principal road length with a UK Pavement Management System defects score of 70 or higher. These variables have been selected as being those most readily affected by short-term policy interventions. In addition, the effect of changes in distance travelled to work has also been forecast.

Forecasts for four districts are presented. Bradford has been selected because it is hilly and has a low percentage cycling for the journey to work in the 2001 census (0.84%). Doncaster has been selected based on intermediate hilliness and cycling proportion for the journey to work (3.13%). York has been selected as the city with the largest proportion cycling for the journey to work in the 2001 census (13.06%) for the districts that have mapping data. It is very flat. The London Borough of Merton has been selected based on it having the average journey to work proportion for London of 2.55%.

The forecasts are presented in Table 4. A 25km increase in the length of off-road routes has been adopted and this could be thought to represent the introduction of a cycle network the equivalent of five radial routes each of 5km in length, or other equivalent appropriate network, within the District. The removal of all highway defects is forecast as well as 20% increases in the number of cars per employee[11] and proportion who travel between 5km and 20km to work.

Insert Table 4 here

The forecasts indicate that the provision of traffic free radial routes might produce an increase in cycling of between 17% and 101%, with the lowest increase being in the hilliest area. The increases in cycle use would materialise only if the facilities were along corridors that would benefit cycle traffic, and this realisation helps emphasise that the shape of the network, a feature not modelled, will be significant in determining use levels. Cycling design guidance reminds us that a bicycle network needs to be coherent, attractive, comfortable and direct (CROW, 1993).

The removal of highway defects has a minor positive effect on levels of cycling but this effect is smaller than the change brought about by increases in car ownership, which are as high as -23%. The effects of car ownership changes compare with Wardman et al. (2007),

[11] The forecast increases in car ownership in Bradford, Doncaster, York and Merton between 2001 and 2011 are respectively 22%, 18%, 22% and 14%. Average trip distance was 6.6 miles in 1998/00, rising to 6.8 miles in 2002 and 6.9 miles in 2003, but then declining to 6.8 miles in 2004 (National Statistics, 2005).

who forecast that a 1% increase in car ownership over the 10 year period from 1997 would reduce cycle more share for journeys under 7.5 miles from 5.8% to 4.8%, an 18% reduction. Increases in distance travelled to work also have an effect, albeit minor, and in any case it is not clear from recent trends that journey distances will increase in the near future. Combining the effect of a policy measure to create 25km of off-road route, but set against a background of increasing car ownership, has the net effect of reductions in cycle use in Bradford and Doncaster, a slight positive effect in York and a significant positive effect in the London Borough of Merton.

Hopkinson and Wardman (1996) show significant route switching away from trafficked routes to traffic free routes and Wardman et al. (1997) show that a trebling in cycle mode share could be achieved with wholly segregated facilities. In more recent work Wardman et al. (2007) suggest that a completely segregated route would result in a 55% increase in cycling. These forecasts appear large relative to the results of the aggregate modelling presented here.

6.0 Discussion and conclusion

A logistic regression model of the proportion that cycle to work in England and Wales has been estimated using relevant socio-economic variables and variables representing the transport system and the physical determinants of cycling. The modelling has the added feature of a saturation level estimated from the data, and this has been found to be 43%.

As we noted in Section 2, there are limitations of aggregate modelling particularly in relation to whether they are identifying a true causal effect. Recognising this limitation, we note, however, that the model confirms expectations from other aggregate and disaggregate studies, including the effects of being non-white (e.g. Rietveld and Daniel, 2004, but shown for the first time in a model using UK data), being male (e.g. Moudon et al., 2005), car ownership (e.g. Rietveld and Daniel, 2004), socio-economic classification, income and distance to work (e.g. Waldman, 1977), a measure for the presence of motor traffic (e.g.

Bovy and Bradley, 1985), population density (e.g. Plaut, 2005), highway condition (e.g. Bovy and Bradley, 1985), hilliness and rainfall (e.g. Waldman, 1977).

While the work by Waldman (1977) considered the “joint” effects of hilliness and danger, the analysis he performed was limited to consideration of the potential effect on cycling of combinations of high and low levels of these two variables. The work presented here tested for a range of interactions, including the interaction between hilliness and off-road infrastructure provision and for the first time demonstrates that, while hilliness has a significant effect, it does not have a detrimentally compounding effect when linked with policy variables that may be adjusted to increase cycle use. However, even in moderately hilly areas the provision of off-road routes would be offset by the predicted effect of car ownership growth over ten years. The economic evaluation of bicycle infrastructure schemes will be adversely affected in hilly areas where levels of take up of cycling may be low. Our work points to the central importance of estimating the effects of hilliness in mode choice modelling when cycle mode shares are being estimated.

The model has, for the first time in the UK context, demonstrated the effect of a warmer climate supporting higher levels of cycling. Also for the first time in aggregate modelling, the direct effects of sex, transport demand intensity, population density and highway pavement condition have been estimated from aggregate data.

Additionally, our model shows that, while higher car ownership is linked with a lower proportion cycling, some variables that might be deemed to be related to car ownership, namely socio-economic classification and income, show contrary effects which we have quantified. Wards with higher proportions of ‘higher professionals’ display higher levels of cycling to work, but wards that have higher income deprivation display lower levels.

The effects on the perception of the acceptability of cycling based on safety in different conditions have not proved significant and further work is recommended on modelling at an aggregate level to differentiate better between districts with respect to conditions for cycling.

Evidence from disaggregate analysis using stated preference techniques shows high values placed on facilities for bicycle traffic, particularly facilities segregated from other motor traffic, with potentially high switching to cycling as a result. The model presented here has, however, shown that reasonable increases in length of bicycle facilities would generate only modest increases in cycling to work. We must recognise that forecasts produced by different approaches will vary, and it is apparent that further work to reconcile forecasts between aggregate and disaggregate models and with actual experience is required.

Even small forecast increases in cycle use may be valuable for traffic management, parking management and the health of the individual cyclists. Significant increases in infrastructure length may be unrealistic and unwarranted and greater increases in bicycle trip numbers may be possible from a less ambitious programme of investment in facilities but coupled with appropriate promotional measures.

There would be value in inter-temporal models that measure the change in level of cycling relative to changes in relevant determinants of cycle use linked to decadal census. There would also be value in expanding the model to include Scotland and Northern Ireland, and to include other relevant variables not presently included in the model, in particular for the public transport alternative and also person type characteristics relating to physical activity. Other interesting aspects concerning bicycle mode choice may emerge from a similarly constructed pan-European model.

The work has contributed to official government guidance on estimating changes in levels of cycling use (Webtag, 2007), to assist in planning bicycle access for the 2012 London Olympic Games and in targeting travel planning in North Yorkshire. The UK Highways Agency has also expressed interest in the model for use in development control. The model provides a tool for policy makers to assist them in forecasting and demonstrates many sensible results that are not possible to achieve by other than aggregate modelling. The model is complementary to disaggregate modelling work that has considered the evaluation of facilities for cycling and there is value in more work which seeks to maximise the contribution of both types of model.

References

Ashley, C.A. and Banister, C.: Cycling to work from wards in a metropolitan area. 1 Factors influencing cycling to work. Traff. Eng. and Con. **30(6)**, 297-302 (1989)

Baltes, M. Factors influencing nondiscretionary work trips by bicycle determined from 1990 US census Metropolitan Statistical Area Data. Transportation research record 1538 TRB, National research council, Washington DC, pp96-101 (1996)

Bovy, P.H.L, and Den Adel, D.N.: Routekeuzegedrag van fietsers: een analyse met de funktionele meetmethode. Delftse Universitaire Pers, Delft (1985)

Bovy, P.H.L. and Bradley, M.A.: Route choice analysed with stated preference approaches. Trans. Res. Rec. **1037**, 11-20 (1985)

British Standards: 6399 Part II Loading for buildings. Code of practice for wind loads (1997)

Cervero, R. and Radisch, C.: Travel choices in pedestrian versus automobile oriented neighbourhoods. Trans. Pol. **3(3)**, 127-141 (1996)

CROW: Sign up for the bike – design manual for a cycle friendly infrastructure. Centre for research and contract standardisation in civil engineering, The Netherlands (1993)

Defra: Department of environment food and rural affairs, Countryside Information System http://www.cis-web.org.uk/home/ (Cited 2003)

DfT: National Cycling Strategy. Department for Transport, London www.dft.gov.uk (1996)

DfT: The future of transport - a network for 2030. Transport White Paper, July 2004. Department for Transport, London. www.dft.gov.uk (2004)

Dill, J. and Carr, T.: Bicycle commuting and facilities in major US cities. Trans. Res. Rec. **1828** (2003)

Guthrie, N., Davies, D.G. and Gardner, G.: Cyclists' assessments of road and traffic conditions: the development of a cyclability index. TRL Report 490. Transport Research Laboratory, Crowthorne (2001)

Hopkinson, P. and Wardman, M.: Evaluating the demand for new cycle facilities. Trans. Pol. **3(4)**, 241-249 (1996)

Landis, B.W., Vattikuti, V.R. and Brannick, M.T.: Real-time human perceptions toward a bicycle level of service. Trans. Res. Rec. **1578** (1997)

LRC: The travel behaviour of different ethnic groups in London. London Research Centre, London (1997)

Meteorological Office: UK climate and weather statistics http://www.met-office.gov.uk/climate/uk/ (cited 2004)

McClintock, H. and Cleary, J.: Cycle facilities and cyclists' safety. Trans. Pol. **3(1/2)**, 67-77 (1996)

Moudon, A. V., Lee, C., Cheadle, A. D., Collier, C. W., Johnson, D., Schmid, T. L., Weather, R. D.: Cycling and the built environment, a US perspective. Trans. Res. D **11**, 245-261 (2005)
National Statistics: Transport statistics bulletin, national travel survey: 2004 SB (05) 30. Department for Transport, London (2005)
Nelson, A. and Allen, D. If you build them, commuters will use them: association between bicycle facilites and bicycle commuting. Transportation research record 1578, TRB, National Research Council, Washington DC, pp79-83 (1997).
Ortúzar, J. de D., Iacobelli, A.and Valeze, C.: Estimating demand for a cycleway network. Trans. Res. A **34**, 353-373 (2000)
Parkin, J.: Comparisons of cycle use for the journey to work from the '81, '91 and 2001 censuses. Traff. Eng. and Con. **44(8)**, 299-302 (2003)
Parkin, J., Wardman, M. and Page, M. Models of perceived cycling risk and route acceptability. *Accident Analysis and Prevention* Vol. 39 pp364-371 (2007). doi:10.1016/j.aap.2006.08.007
Pindyck, R.S. and Rubinfeld, D.L. (1991) Econometric models and economic forecasts. 3rd Edition. McGraw-Hill. New York.
Plaut, P.O.: Non-motorized commuting in the US. Trans. Res. D **10**, 347-356 (2005)
Rietveld, P. and Daniel, V.: Determinants of bicycle use: do municipal policies matter? Trans. Res. A **38**, 531-550 (2004)
Stinson, M. A. and Bhat, C. R. Frequency of bicycle commuting: internet-based survey analysis. Transportation Research Board No. 04-3493 (2004)
Stinson, M. A. and Bhat C. R. A comparison of the route preferences of experienced and inexperienced bicycle commuters. Transportation Research Board Paper No. 05-1434 (2005)
Tilahun, N. Y., Levinson, D. M. and Krizek, K. J.: Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Trans. Res. A (2006)
Waldman, J.A.: Cycling in towns: a quantitative investigation. LTR1 Working paper 3. Department for Transport, London (1977) (Out of print)
Wardman, M., Hatfield, R. and Page, M.: The UK national cycling strategy: can improved facilities meet the targets? Trans. Pol. **4(2)**, 123-133 (1997)
Wardman, M., Tight, M. and Page, M.: Factors influencing the propensity to cycle to work. Trans. Res. A **41**, 339-350 (2007)
Webtag. Guidance on the Appraisal of Walking and Cycling Schemes TAG Unit 3.14.1. Consultation, March. www.webtag.org.uk (2007)

Table 1 Cycle mode shares for the journey to work in England and Wales

Cycle mode share	Number of Districts	Cycle mode share	Number of Districts
10% + [1]	7	4.00 – 4.99%	45
9.00 – 9.99%	2	3.00 – 3.99%	64
8.00 – 8.99%	2	2.00 – 2.99%	78
7.00 – 7.99%	9	1.00 – 1.99%	120
6.00 – 6.99%	9	0.00 – 0.99%	20
5.00 – 5.99%	20		
		Total	376

Note: The district with the highest percentage is Cambridge at 28.34%.

Table 2 Data used to determine the propensity to cycle for the journey to work.

Unit	Description	Mean	Range
	Socio-economic Variables (all at ward level)		
Proportion	of all people in employment aged 16-74 who are male	55%	47%-78%
Proportion	of all people aged 16-74 who are non-white.	5.2%	0%-88%
Proportion	of employees aged 16-74 with higher level qualifications	24%	4.4%-82%
Proportion	of employees in the bands "16-24", "25-34", "35-49", "50-59", "60-64" and "65-74"	-	0%-54%
Number	cars per employee	1.65	0.43-2.39
Proportion	of employees in each of eight Socio-Economic Classes[1]	-	0%-54%
Index	of deprivation income score	0.11 (English) 19 (Welsh)	0.01-0.62 (English) 0.12-84 (Welsh)
	Physical Variables (ward, district & weather region level)		
Proportion	of journeys to work in the distance bands "under 2km", "2-5km", "5-10km", "10-20km", "20-30km", "30-40km", "40-60km", "60km and over" at ward level	-	3.1%-23.3%
No/sq.km	Population Density (all people aged 16-74 divided by area of district in hectares)	825	16-10406
hours	Total annual hours of sunshine for the year May 2000 to April 2001 for the weather region	-	1377hrs-1463hrs
mm	Total annual millimetres of rainfall for the year May 2000 to April 2001 for the weather region	-	890mm-1643mm
^{0}C	Mean temperature for the year May 2000 to April 2001 for the weather region	-	8.6^{0}C-10.3^{0}C
m/s	Basic wind speed for structural design for the district	-	38-48m/sec
Proportion	of 1km squares in district with mean slope 3% or greater (and also 4% or greater was tested)	68%	0%-100%
	Transport Variables (all at district level)		
No/km	Transport demand intensity (workers aged 16-74 divided by district road length)	2.58	0.06 – 33.3
Proportion	of Principal road length deemed to have failed	11.6%	0%-51%
Proportion	of Non-principal road length deemed to have failed	9.4%	0%-56%
Proportion	of road and cycle route that is signed	-	0%-17%
Proportion	of cycle route that is off-road	-	0%-11%
Proportion	of cycle route that is adjacent to the road	-	0%-4.3%
Proportion	of road that has a bicycle or bus lane	-	0%-14.6%
Proportion	Probability of acceptability of cycling	0.70	0.68-0.73
	The Dependent Variable		
Proportion	of journeys to work by bicycle	2.9%	0%-35.4%

Notes

1 *The eight Socio-Economic Classifications (SECs) comprise Class 1.1 "higher managerial and professional occupations in large employers", Class 1.2 "higher professional occupations", Class 2 "lower managerial and professional occupations", Class 3 "intermediate occupations, Class 4 "small employers and own account workers", Class 5 "lower supervisory and technical occupations", Class 6 "semi-routine occupations" and Class 7 "routine occupations".*

Table 3 Model of the variation in use of the bicycle for England and Wales

Variable	Coefficient Estimate	t-statistic
S, saturation level	0.4278	37.23
Propn. Population non-white	-1.1708	-11.93
Propn. Employees who are male	2.8284	12.59
Number of cars per employee	-0.9758	-22.90
Propn. In socio-economic classifications (SECs) 5 “lower supervisory and technical”, 6 “ semi-routine occupations” and 7 “routine occupations”	Base	
Propn. In SEC 1.1 “Higher managerial & professional in larger organisation”	-4.7724	-10.14
Propn. In SEC 1.2 “Higher professional”	5.7281	24.44
Propn. In SEC 2 “Lower managerial and professional”	-2.5041	-11.32
Propn. In SEC 3 “Intermediate occupations”	-2.4663	-8.75
Propn. In SEC 4 “Small employers & own account workers”	-4.1446	-13.63
Index of deprivation Income Score – English	-2.2200	-16.49
Index of deprivation Income Score - Welsh	-0.0159	-5.39
Propn. in the distance band “less than 2km” and bands 20km & greater	Base	
Propn. In distance band “2km to less than 5km”	-0.6916	-8.53
Propn. In distance band “5km to less than 20km”	-1.6556	-20.49
Transport demand intensity (employees divided by road length)	-0.0373	-17.74
Population density (population divided by area)	0.0001	9.11
Propn. Principal Roads with negative residual life	-0.3493	-3.87
Propn. Non-principal roads with negative residual life	-0.7830	-8.25
Propn. 1km squares with slope 3% or steeper	-1.3920	-50.93
Total annual rainfall in millimetres	-0.0006	-17.40
Mean temperature in degrees centigrade	0.0782	7.87
Proportion of off-road route	12.5162	18.72
Dichotomous variable for non-mapped wards	0.9376	18.78
$\overline{R}^2$	0.816	

Table 4 Forecasts for cycling to work

	Bradford	**Doncaster**	**York**	**Merton**
Proportion of 1km squares with slope 3% or greater	100%	31%	5%	26%
Base				
2001 census proportion cycling to work	0.84%	3.1%	13.1%	2.6%
Modelled proportion cycling to work	**1.2%**	**4.1%**	**11.8%**	**3.7%**
Modelled number of trips by bicycle	2600	5800	11200	3900
25km increase in off-road route				
Existing length of route with bicycle facilities	55km	65km	78km	14km
Forecast proportion cycling to work	**1.4%**	**4.9%**	**14.9%**	**7.5%**
Percentage increase	17%	18%	26%	101%
Elimination of highway defects				
Existing percentage of principal /non-principal highway that has negative residual life	29%/5%	3%/3%	11%/11%	18%/18%
Forecast proportion cycling to work based on elimination of highway defects	**1.3%**	**4.3%**	**12.9%**	**4.5%**
Percentage increase	15%	3%	9%	20%
Car per employee increases by 20%				
Forecast proportion cycling to work	0.90%	3.2%	9.8%	3.0%
Percentage decrease	-23%	-22%	-17%	-20%
Proportion in distance band 5km to 20km increases by 20%[1]				
Forecast proportion cycling to work	1.0%	3.7%	11.2%	3.4%
Percentage decrease	-11%	-11%	-6%	-10%
Effects of 25km of off-road route plus 20% increase in car ownership				
Forecast proportion cycling to work	**1.1%**	**3.8%**	**12.5%**	**6.2%**
Percentage change	-10%	-8%	+6%	+65%

Notes

1 It is assumed that the increase in the percentage for the distance band 5km to 20km is drawn from other bands that form the base.