transport for quality of life

CWIS Active Travel Investment Models: Model structure and evidence base

Technical appendix 7: Factors affecting walking and cycling levels, and model scaling factors

Sally Cairns

Transport for Quality of Life

September 2019

Document name:	Technical appendix 7: Factors affecting walking and cycling levels, and model scaling factors
Project name:	CWIS Investment Model – Additional analysis and reports WP1-817
Paper number:	Technical appendix 7
Client:	Department for Transport
Date:	September 2019

This report was delivered through the SPaTS Framework Lot 1, by the Arup supplier group.

Contact:
Transport for Quality of Life Ltd
T: 01654 781358
E: info@transportforqualityoflife.com
W: www.transportforqualityoflife.com

Contents

1. Potential factors explaining variation in levels of active travel at local authority district (LAD) level

A key feature of the Investment Models is that the effectiveness of each intervention package is varied according to the characteristics of the LAD in which it is applied. In other words, the models recognise that some areas may be more fertile territory for investment than others. In order to model this, it was necessary to generate a series of scaling factors that indicate the relative effectiveness of investment in different LADs.

The available evidence does not lend itself to direct assessment of how 'place' characteristics influence the cost of generating an active travel trip. There are few examples of comparable interventions being delivered in different places with associated cost data, although there was one small dataset, for cycling, for which this was the case. In addition, in practice, observed variation will be due to both the efficiency and effectiveness with which any individual measure has been implemented (i.e. how well it is designed, whether it is a roll-out of an existing programme or a new initiative requiring 'start-up' costs etc.), as well as the suitability of the area in which it takes place (including the receptiveness of the local population and the conditions in which they are making travel choices).

Therefore, for the purposes of these models, our assumption was that the underlying or 'intrinsic' characteristics in each LAD that partly influence baseline levels of active travel, are also likely to influence the impact of future investment (i.e. the cost per additional stage generated).

Considerable exploratory analysis was carried out to identify characteristics that predict baseline levels of active travel, using datasets relating to population structure, urban structure, deprivation, car ownership, traffic speeds, propensity to cycle, rainfall and measures of accessibility, including access to primary schools.

Much of the exploratory analysis took place before a decision had been made as to what data to use for the model baseline – and included comparisons with data from different years including Census statistics on commuting mode share, Active People Survey and Active Lives Survey data. Consequently, the dates for the datasets vary somewhat between 2010 and 2015. Any changes in relative levels of active travel (and underlying characteristics) of LADs over the five-year period that the different datasets span are likely to be small.

Most analyses were conducted using 324 LADs, excluding the Isles of Scilly and the City of London, because their small population sizes (and associated characteristics) had the potential to distort the results.

The analysis led to the development of scaling factors that could be used in each model. These scaling factors modify the calculated impact of investment, according to the characteristics of the area where investment is applied. For example, in a LAD with a scaling factor of 1, an intervention package costing £1 per trip will produce 1 trip for every £1 of investment. If the scaling factor is 2, 2 trips are produced, and if the scaling factor is 0.5, only 0.5 trips are produced.

The scaling factors developed should be considered indicative only, and values for specific LADs are not intended to be used or compared on an individual basis.

Table 1 lists the main data sources used for identifying characteristics that predict baseline levels of active travel.

Table 1: Potential explanatory factors for levels of active travel

Measure	Source
Population structure	
% 0-15	Data from Census table KS102. Data relates to the percentage of the usual resident population in different age bands.
% 16-64	
% 65+	
% female	Data from Census table KS101. Data relates to the percentage of the usual resident population that is female.
Urban structure	
Population density	Data from Census table KS101. Data relates to the number of usual residents per hectare.
% in terraced housing, flat, maisonette or apartment	Data from Census table QS401. Data relates to the percentage of the usual resident population in households living in terraced housing (including end terraces), flats, maisonettes or apartments. (Those living in communal establishments are excluded from the figures.)
% in rental accommodation	Data from Census table KS402. Data relates to the percentage of households that are in private or social rental accommodation. (Those living in communal establishments are excluded from the figures.)
Deprivation	
Index of multiple deprivation	LAD scores for the Index of Multiple Deprivation, generated by averaging 2010 LSOA values[1]
Car ownership and traffic speeds	
% households without cars	Data from Census table KS404. Data relates to the percentage of households with no cars. (Those living in communal establishments are excluded from the figures.)
Cars per person	Data on the number of private cars in each LAD in 2011 kindly supplied by DfT[2], which was then divided by the usual resident population.
Journey times on A-roads	Data from DfT table CGN0201b[3]. Data were for weekday morning peak (7-10am) average journey times on locally-managed A-roads, in minutes per mile, in 2011/12 for all traffic.
Propensity to cycle tool indicators[4]	
Measure of hilliness	The average fast route gradient (%) of commute trips in the relevant zone with fast route distance <10km.
Measure of commuters within 10km	Percentage of commuters in zone with fast route commute distance <10km (calculated excluding trips with no fixed work place).
% anticipated to cycle	Potential number of extra commuter cyclists compared to 2011 Census, based on PCT calculations (the 'Govtarget_sic' figure), converted to a percentage of the usual resident population.

[1] Indices of deprivation 2010. Note that the local authorities summaries file and/or more recent data could have been used.

[2] Data received 12th June 2018, personal correspondence in relation to CWIS model.

[3] DfT Congestion Statistics

[4] Data downloaded 8th May 2018 from PCT. Note that current values downloadable from the PCT website differ from the ones used here.

Measure	Source
Rainfall	
Rainfall	Data generated from Met Office information about average April rainfall between 1981-2010. MSOA values produced by Ian Philips at the University of Leeds as part of the MOT research project[5]. For this work, LAD values were generated as the mean of the relevant MSOA values.
Accessibility statistics	
Time to all 8 key services	Accessibility statistics are published by DfT. These give travel times to eight key services (employment centres, primary schools, secondary schools, further education institutions, GPs, hospitals, food stores and town centres). 2015 data were analysed[6]. Times are given for: (a) Cycling (b) Public transport / walking (where the quicker of the two modes is chosen, albeit that walking is always assumed to take at least 5 minutes and travel times are capped at 120 minutes). Travel times (in minutes) to the relevant services have been added together to produce an overall metric. Travel times to two sub-clusters of services were also considered, having analysed the datasets to investigate variance in travel time between different (sub-)groups of services.
Time to the most localised services (i.e. food shops, primary schools and employers with up to 100 employees)	
Time to intermediate services (i.e. secondary schools, GPs and employers with up to 500 employees)	
Travel time (in minutes) by PT/walk to nearest primary school (2011 and 2014)	Specific accessibility datasets for travel to primary schools were downloaded for both 2011 (Table ACS0402, revised 2011 statistics, variables psPTtime, psPT15 and psPTcont) and 2014 (Table JTS0402, variables PS101, PS102 and PS106), given changes in methodology might mean that one was more powerful than the other. For 2011, a new variable was created, which gives the proportion of 5-10-year-olds with access to a primary school within a 'reasonable time' by PT/walk (using indicators PSCHO005 and PSCHO032). Note that it has not been possible to find a definition of 'reasonable time' as the technical note associated with the table could not be located on the website. However, as a rough rule, the number seems to be about half of the equivalent metric for the proportion of 5-10-year-olds within a 15 minute walk, suggesting 5-10 minutes must have been used as the cut-off. This metric is not given in the 2014 journey travel times dataset.
Number of primary schools within 15 minutes travel by PT/walk (2011 and 2014)	
Number of primary schools 'accessible' by PT/walk (2011)	
% 5-10yr olds within 15 mins of primary school by PT/walk (2014)	
% 5-10 year olds within a 'reasonable time' of primary schools by PT/walk (2011)	

[5] Met Office 5km – 5km gridded data, supplied through the MetOffice data portal, was assigned to MSOA centroids using the Extract Data to Points tool in ESRI ARCGIS 10.2. Grateful thanks to Ian Philips. Data processing completed as part of the MOT EPSRC project (EP/K000438/1).

[6] DfT accessibility statistics.

2. Considering characteristics that could explain differences in the costs of generating a cycle trip in different locations

For the Cycling Model, a subset of the variables in Table 1 were analysed in conjunction with the baseline dataset of cycle stages per person in each LAD (as defined in Appendix 1). Table 2 gives the correlation coefficients between baseline cycle stages per person and each of the variables, as calculated using data for 324 LADs. The correlation coefficients[7] provide an indication of the relative strength of the relationship between cycle stages per person and each of the explanatory variables – i.e. the larger the magnitude of the values (either positive or negative), the stronger the relationship. Many of these variables are also related to each other.

Table 2: The correlation coefficients between a range of variables and the number of cycle trip stages per person per annum (for 324 LADs)

	Cycle stages per person p.a.
PCT anticipated proportion of people who might cycle	0.49
PCT measure of commuters within 10km	0.32
PCT measure of hilliness	-0.32
Rainfall	-0.30
Density	0.31
% households without cars	0.25
Private cars per person	-0.26
% aged 65+	-0.22
% aged 0-15	-0.23
% aged 16-64	0.40
% in terraced housing, flat, maisonette or apartment	0.29
% female	-0.16
% living in rental accommodation	0.35
Time taken to eight key services by bike	-0.15
Index of multiple deprivation	-0.01
Journey times on A-roads	0.16

Using the regression tool within Excel, a series of regression runs were used to explore which *combinations* of variables provided the greatest explanatory power for the observed variation in cycle stages per person in the baseline dataset.

There is a trade-off between maximising explanatory power by using many variables, and minimising the number of variables for simplicity. We used the minimum number of variables that gave reasonable explanatory power, such that adding in more variables only provided a small additional improvement.

[7] Note that whilst correlation coefficients (R-values) provide an initial measure of relationship strength, determination coefficients (R^2-values) are needed to estimate how much the variation in one variable can be predicted from another.

Separately, estimates from the 18 Cycling Demonstration Towns and Cities of 'cost per cycle trip generated'[8] were compared with the variables in Table 2, to try to understand which characteristics seemed to be most useful for explaining variation in observed costs.

3. Scaling factors used in the Cycling Model

Two scaling factors were developed for the Cycling Model.

The first of these drew on our assessment of the combination of variables that best explained variation in cycle stages per person at baseline. This combination of variables gave an 'Intrinsic Cycling Potential' (ICP) score for each LAD.

The second scaling factor drew on our assessment of the variation in cost per cycle trip generated in the 18 Cycling Demonstration Towns and Cities.

3.1 Intrinsic Cycling Potential

The Intrinsic Cycling Potential score was created from:

(a) The PCT anticipated proportion of people who might cycle (which is itself a function of hilliness and length of cycle commute distances);
(b) The proportion of people aged 0-15;
(c) The proportion of people aged 65+; and
(d) The Index of Multiple Deprivation

using the following equation[9]:

ICP = 6.74*PCT – 0.88*%65plus – 2.33*%0-15 – 0.18*IMD + 71.91

Figure 1 shows the relationship between ICP and baseline cycling levels. Taken together, these 'intrinsic' characteristics of each LAD can predict about 46% of the variation in baseline levels of cycling in LADs.

As a check on the measure, the LADs were divided into 10 deciles, where decile 1 included the 10% of LADs with the lowest baseline cycling levels, and decile 10 included the 10% of LADs with the highest baseline cycling levels. Average ICPs were calculated for each decile. The results are shown in Figure 2.

Looking at the two graphs suggests that, for the majority of local authorities, cycling levels typically vary by a factor of about 3 (between 10 and 30 cycle stages per person per year), depending on the ICP, although those at the two ends of the distribution are lower or higher than this.

The ICP provides an explanation of levels of cycling at baseline. As explained in section 1, our hypothesis is that it can also be used to help predict the effectiveness of future cycling investment, since, in places with higher ICP, people might be expected to take advantage of improved cycle facilities or services to a greater degree than elsewhere, because:

- The area is less likely to feature steep hills or unfeasibly long commute distances;

[8] Data from unpublished analysis prepared for Sloman L, Taylor I, Wilson A, King N, Goodwin P, Anable J, Davison S, Crawford M, Cope A and Adcock S (2014) Finding the Optimum: Revenue / Capital Investment Balance for Sustainable Travel Report to Department for Transport

[9] p values of all coefficients <0.01; overall model F statistic p<0.01; sample 324 LADs.

- There is likely to be a higher proportion of adults of working age, who are more likely to cycle than older people or children (particularly for commuting, which accounts for 40% of all cycling trips according to NTS);
- A lower IMD may mean that a higher proportion of residents are likely to own (or be able to afford to buy) a bike, to have somewhere to store it, and/or to be economically active, and hence to make more trips (including for commuting).

All of these factors would mean that a local authority had to work less hard to get more people cycling in places with higher ICP, and so each £ spent could be expected to have a bigger impact.

Figure 1: Relationship between Intrinsic Cycling Potential and baseline cycling levels

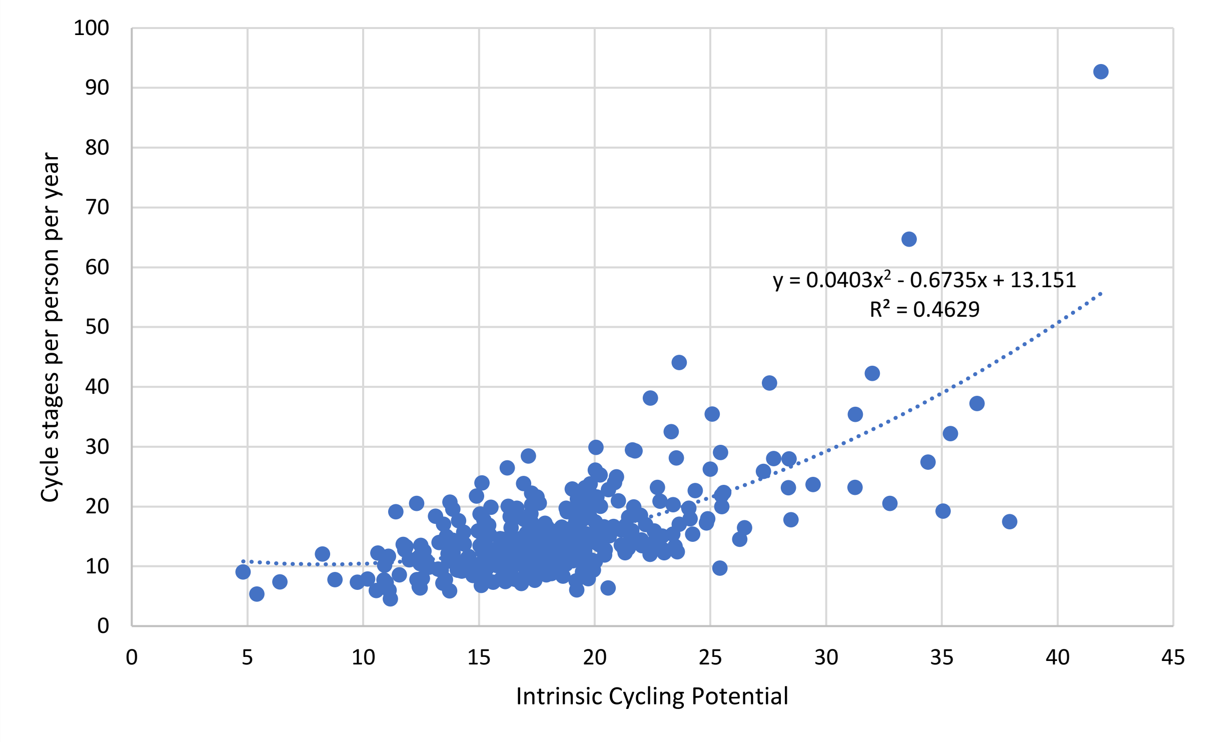
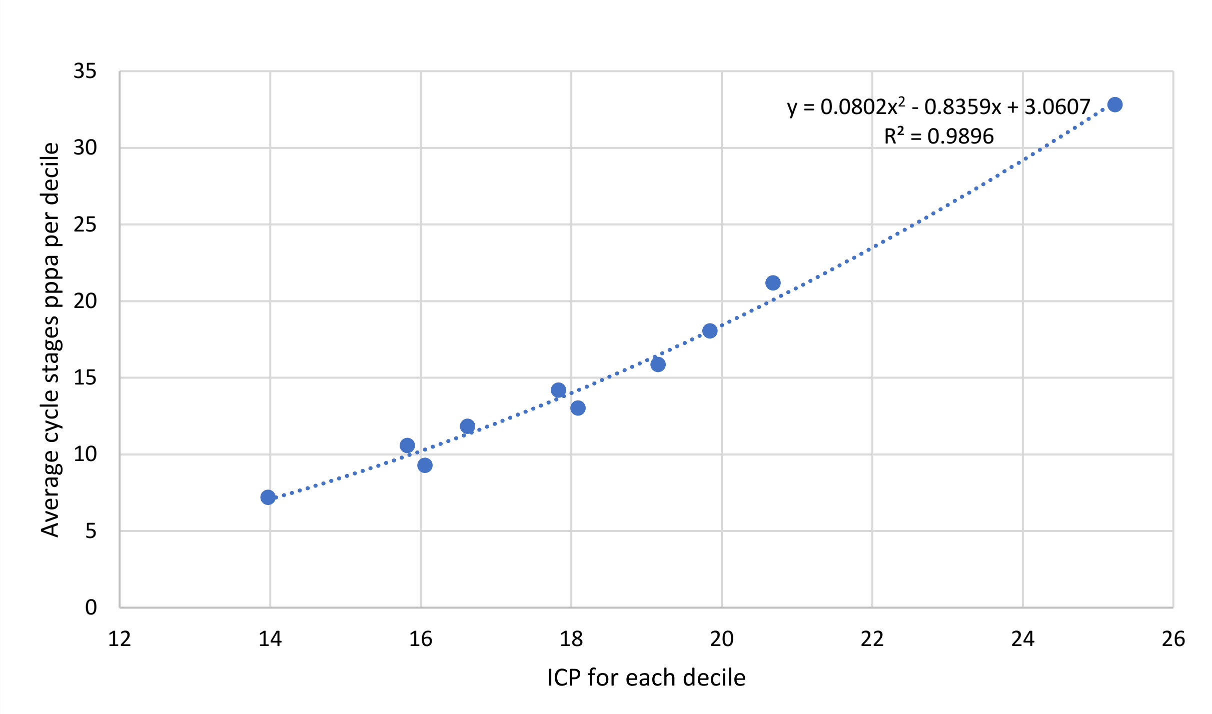


Figure 2: Change in Intrinsic Cycling Potential with each decile of baseline cycling levels

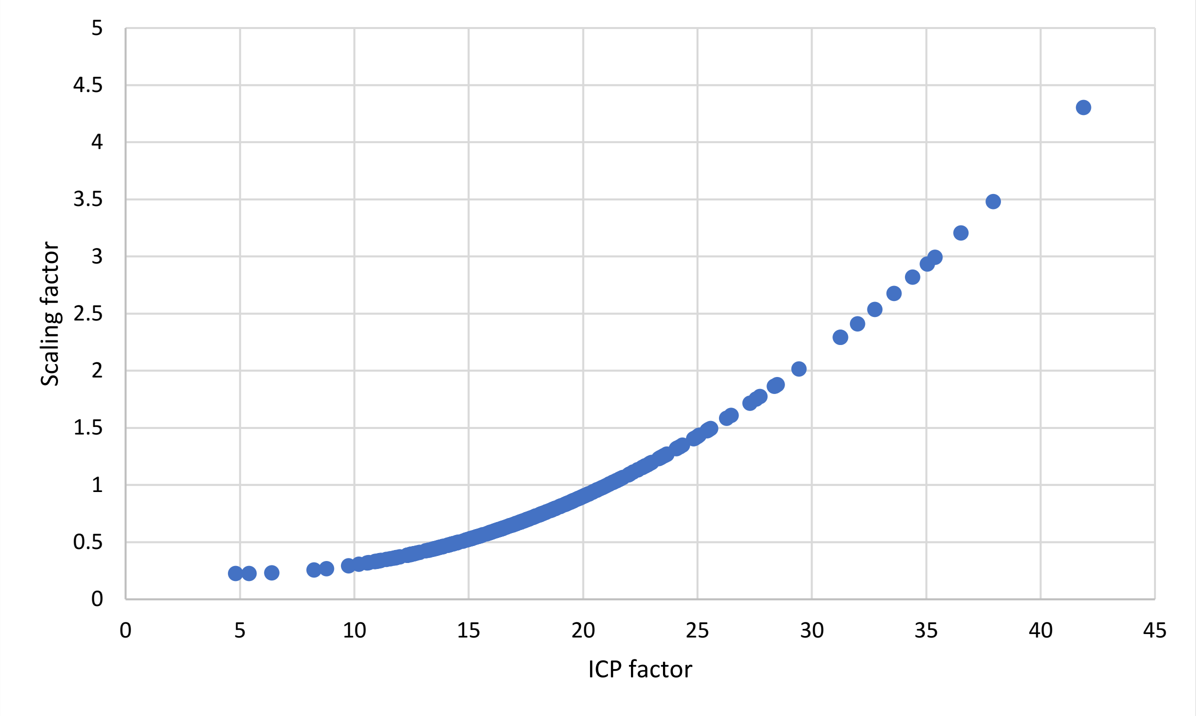


We hypothesised that variation in the effectiveness of investment *due to the underlying characteristics of different LADs* is likely to be of a similar order to the variation in baseline levels. This would mean that ICP might cause variation in the effectiveness of investment of about a factor of 3 between LADs. We scaled the ICP accordingly, using the following equation:

Scaling factor = 0.003 (ICP^2) – 0.03 (ICP) + 0.3.

This produces the plot shown in Figure 3. The range of the scaling factor for 80% of LADs (those between the 10^{th} and 90^{th} percentile) is 0.4 to 1.3, albeit that lower and higher factors are used at the extremes of the distribution.

Figure 3: Scaling factor suggested for the ICP



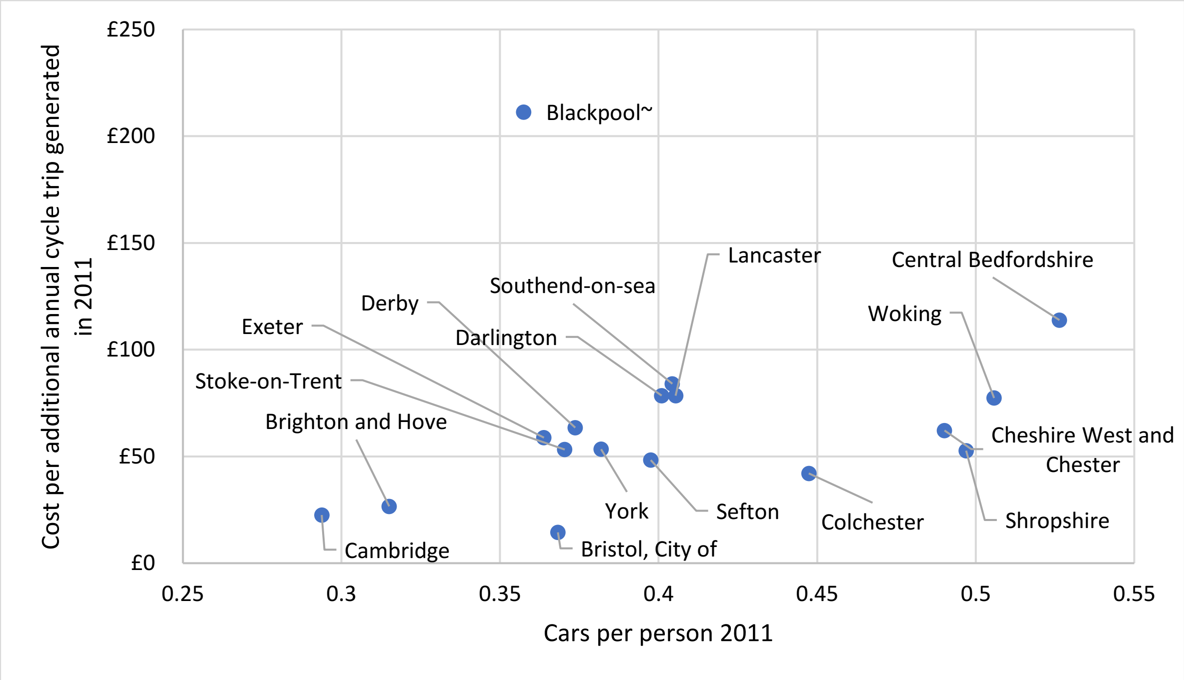
3.2 Traffic conditions factor

From the data about average costs per trip for different initiatives, it is clear that the variation may often be more than a factor of 3. For cycling, we had one other dataset, which enabled an assessment of how the costs of generating a cycle trip may vary in different locations. Specifically, data was available from the 18 Cycling Demonstration Towns and Cities (CDT/CCT) for the cost per trip generated.

Analysis of this dataset did not show a strong relationship between baseline levels of cycling and the costs of generating an additional trip – though the fact that there were only 18 data points (with Cambridge being an outlier in terms of cycling levels and Aylesbury and Blackpool being outliers in terms of cost), means that this did not provide a strong reason not to use the ICP in the model. In addition, differences in the efficiency and effectiveness of implementation between towns will have caused costs to vary, as well as any differences in the suitability of different locations for promoting cycling.

However, exploratory analysis suggested that measures of car use – in particular, the number of private cars per person (as defined in Table 1) – did show some relationship with costs, as shown in Figure 4. Specifically, if data for Blackpool and Aylesbury were excluded (due to their unusually high cost values), the relationship between cars per person and cost per trip was statistically significant, even given the small number of data points ($p<0.01$, $n=16$).

Figure 4: Relationship between private cars per person and cost per cycle trip generated (according to data from the Cycling Demonstration Towns and Cities)



~Both Blackpool (shown) and Aylesbury (outside range shown in graph, with a cost per additional annual trip of around £600) are outliers, recording smaller increases in cycling from automatic counter data than other CDTs/CCTs, and hence higher costs per additional annual cycle trip.

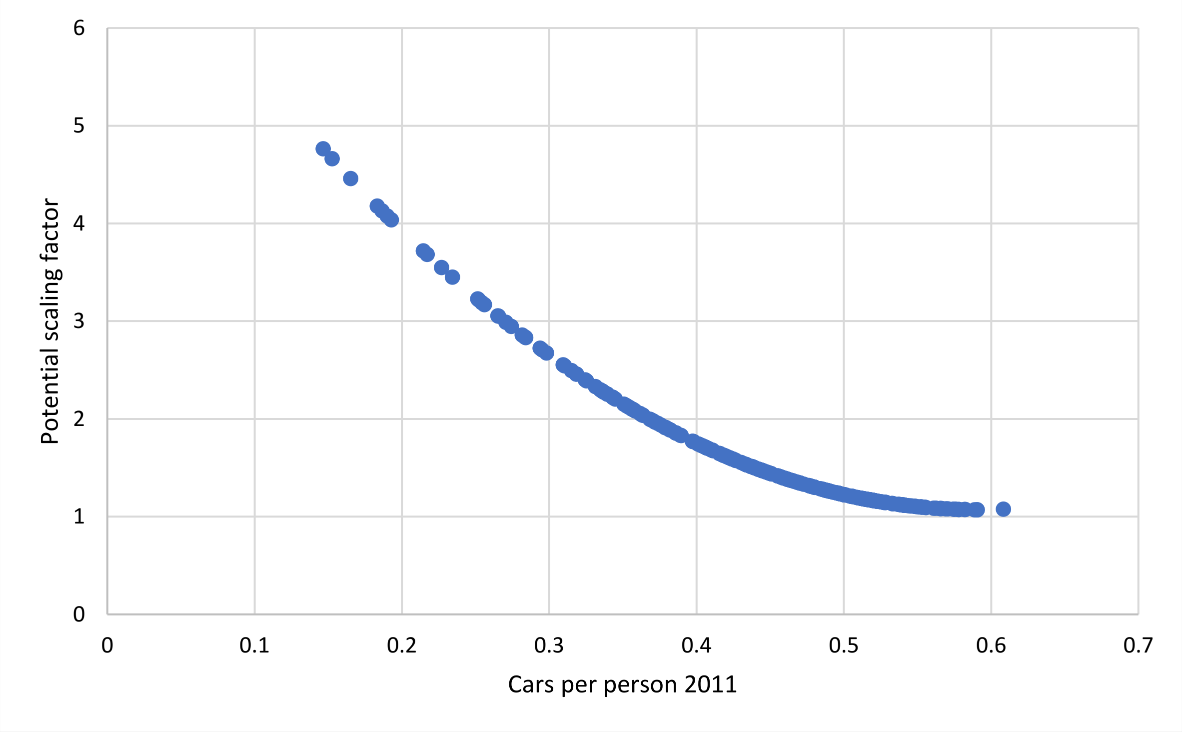
Consequently, a second scaling factor was developed[10]. In this instance, it was necessary to create an inverted scaling factor – since higher levels of car ownership mean the trips per pound spent in the local authority should be scaled down not up. Moreover, given the asymmetry in the underlying data (such that many places have lots of cars, but fewer places have few cars), this was scaled so that it made little difference to the majority of places, but provided an added boost to the effectiveness of investment in places with relatively low car ownership.

Using the following equation (roughly derived by inverting the relationship shown in Figure 4) produces the outcome shown in Figure 5.

Traffic conditions scaling factor = 18.7 (cars pp ^{2}) – 22.1 (cars pp) + 7.6

[10] As one alternative approach, we considered creating a revised ICP, based on regression modelling which included cars per person as one of the determinants of baseline cycling levels. However, doing so would have made relatively little difference to the final ICP values produced, improving the R^2 value of the model by only about 0.02. In order to reflect the potentially important effect of car ownership, use of two scaling factors was therefore adopted. Were more data to become available about how variation in costs is affected by underlying characteristics, the approach adopted here could be revisited in the future.

Figure 5: Scaling factor for the cars per person factor



3.3 Combined scaling factor

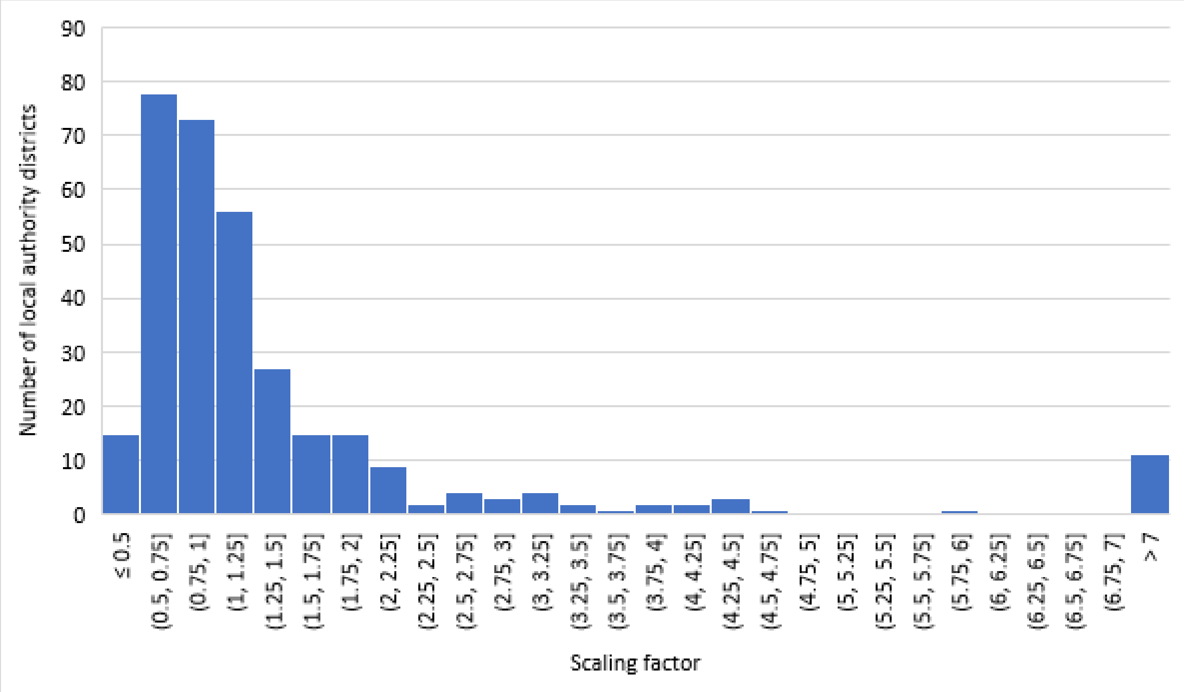
The model applies the ICP scaling factor and traffic conditions scaling factor to each LAD (with the same effect as if a single 'combined' factor were used that was the product of the ICP scaling factor and traffic conditions scaling factor). Figure 6 shows the distribution of values which is produced by applying both sets of scaling factors. Values for the 'combined' factor range from 0.34 to 13.34, although most LADs lie within a range that is considerably less than this (the range for the 5th to 95th percentile values is 0.51 to 4.22).

Were this scaling factor to be used without any modification, an intervention costing an average of £1 per trip would cost about £3 per trip in some places, whilst in other places, £1 would generate about 13 trips. The LADs with higher values are all places where it is plausible that it might be relatively cheap to generate additional cycle trips. However, in order to avoid the model being over-optimistic, it was decided to cap the higher values.

Specifically, for all places where combining the scaling factors generated a value of more than 5, the Traffic Conditions factor was reduced so that the combined effect was 5[11]. This affected the values for Oxford, Cambridge and ten London boroughs.

[11] A value of 5 was chosen based on the observed distribution of scaling factors produced.

Figure 6: Distribution of values produced by combining the scaling factors



4. Considering characteristics that could explain differences in the costs of generating a walk trip in different locations

For the Walking Model, a subset of the variables given in Table 1 were analysed in conjunction with the baseline dataset of walk stages per person in each LAD (as defined in Appendix 1). Table 3 gives the correlation coefficients between walk stages and each of the variables, as calculated using data for 324 LADs. The correlation coefficients[12] provide an indication of the relative strength of the relationship between walk stages and each of the explanatory variables – i.e. the larger the magnitude of the values (either positive or negative), the stronger the relationship.

[12] Note that whilst correlation coefficients (R-values) provide an initial measure of relationship strength, determination coefficients (R^2-values) are needed to estimate how much the variation in one factor can be predicted from another.

Table 3: The correlation coefficients between a range of variables and the number of walk trip stages per person per annum for travel, leisure, and all purposes (for 324 LADs)

	Walk stages for travel	Walk stages for leisure	All walk stages
Rainfall	-0.28	0.17	-0.18
PCT measure of hilliness	-0.22	0.25	-0.04
Density	0.79	-0.59	0.43
% households without cars	0.68	-0.55	0.34
Private cars per person	-0.68	0.63	-0.28
% aged 65+	-0.65	0.73	-0.14
% aged 0-15	0.12	-0.52	-0.29
% aged 16-64	0.74	-0.63	-0.34
% in terraced housing, flat, maisonette or apartment	0.79	-0.56	0.46
% female	-0.16	0.25	0.03
% living in rental accommodation	0.72	0.53	0.40
Time taken to the most local shops and services (PT/walk)	-0.52	0.68	-0.04
Time taken to intermediate shops and services (PT/walk)	-0.57	0.71	-0.07
Time taken to 8 shops and services (PT/walk)	-0.59	0.73	-0.08
Index of multiple deprivation	0.28	-0.40	-0.01
Journey times on A-roads	0.70	-0.57	0.35

One of the insights from this table is the relatively weak relationship between overall levels of walking, and many potential explanatory variables – partly because the determinants of walking for leisure and walking for travel are often acting in opposing directions.

This was also evident when looking at differences in walking levels between different types of area[13]. Average figures are given in Table 4, whilst Figure 7 shows the range of values.

Differences in the variables affecting walking for travel and walking for leisure were also evident when regression modelling was carried out on a similar basis to that conducted when investigating cycling (see section 2). Whilst such modelling (using many of the variables listed above) could explain about 80% of the variation in walking for travel and 70% of the variation in walking for leisure, it could only explain about 50% of the variation in 'all walking'.

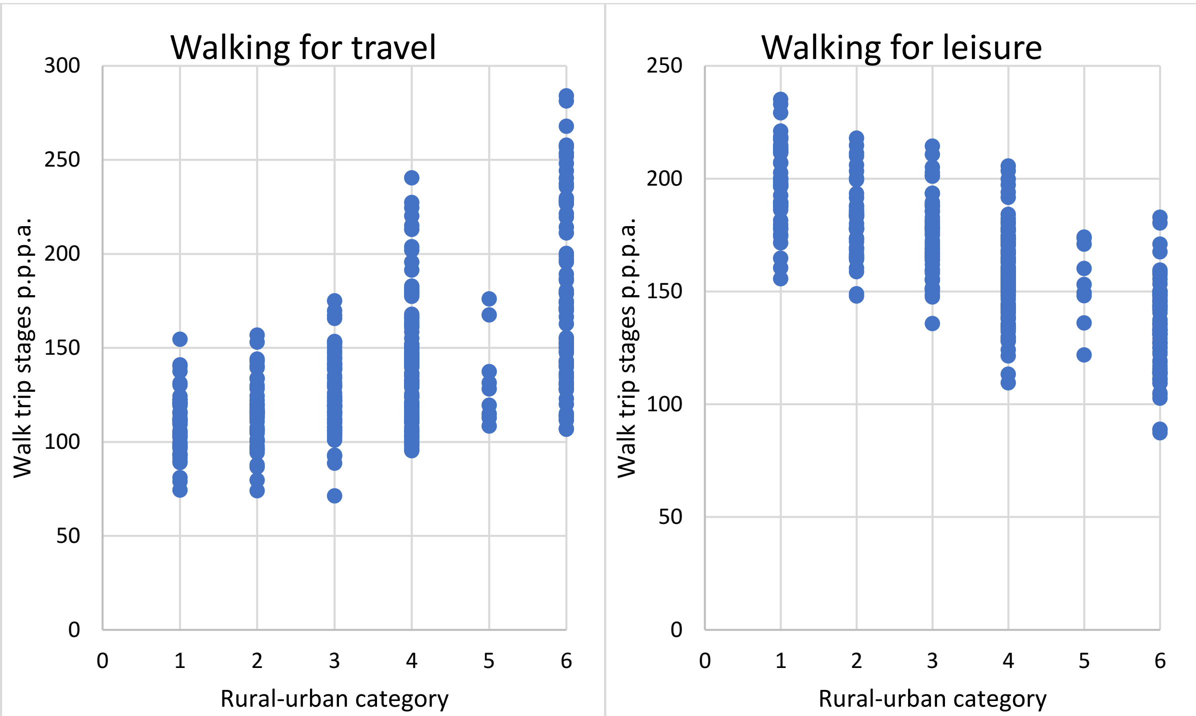
On this basis, modelling took place for the two types of walking separately.

[13] LADs classified according to the Rural-Urban classification produced by ONS

Table 4: Average walk trip stages per person per year in different types of area

		Walking for travel	Walking for leisure	Any walking
6	Urban with Major Conurbation	175	133	309
5	Urban with Minor Conurbation	133	154	287
4	Urban with City and Town	141	158	299
3	Urban with Significant Rural (rural including hub towns 26-49%)	126	174	300
2	Largely Rural (rural including hub towns 50-79%)	115	181	296
1	Mainly Rural (rural including hub towns >=80%)	110	198	308

Figure 7: Walk trip stages per person per year in different types of area



5. Scaling factor used in the Walking Model

5.1 Walking for travel

Using 9 of the factors listed above, it was possible to predict about 80% of the variation in levels of walking for travel. Measures used were:

- Measure of hilliness
- Density 2011
- % aged 65+
- % aged 0-15
- % in terraced housing, flat, maisonette or apartment
- % female
- Cars per person
- Rainfall
- Index of multiple deprivation

77% of the variance could be explained using just four factors:

- % in terraced housing, flats, maisonettes or apartments
- % female
- Cars per person
- Index of multiple deprivation

Given the importance of the rural-urban classification, the data were divided into rural areas and urban areas, and separate regression analysis was undertaken. This showed that the predictive power was stronger in urban areas, (where age 0-15, rainfall and the proportion of households without cars added significant additional explanatory power) compared with predictive power in rural areas. However, there were no major overall gains in predictive power from this approach.

5.2 Walking for leisure

It was possible to predict about 70% of the variation in levels of walking for leisure using the following 7 factors:

- Density 2011
- % no car households 2011
- % aged 0-15
- % in terraced housing, flat, maisonette or apartment
- Cars per person
- Journey times on A-roads
- Time to 8 key shops and services

About 65% of the variation could be predicted using:

- % no car households 2011
- % aged 65+
- Cars per person
- Time to 8 key shops and services

5.3 Scaling factors for walking

Two scores were generated for each area, to represent intrinsic 'walking for travel' potential, and intrinsic 'walking for leisure' potential: IWP-travel and IWP-leisure.

The decision was made to create a scaling factor from IWP-travel for use in the Walking Model, because most of the interventions included in the model are likely to stimulate walking for travel rather than for leisure[14]. It was not possible to use different scaling factors for different interventions within the model, given the degree of complexity that would have been involved in operationalising this.

5.4 IWP-travel

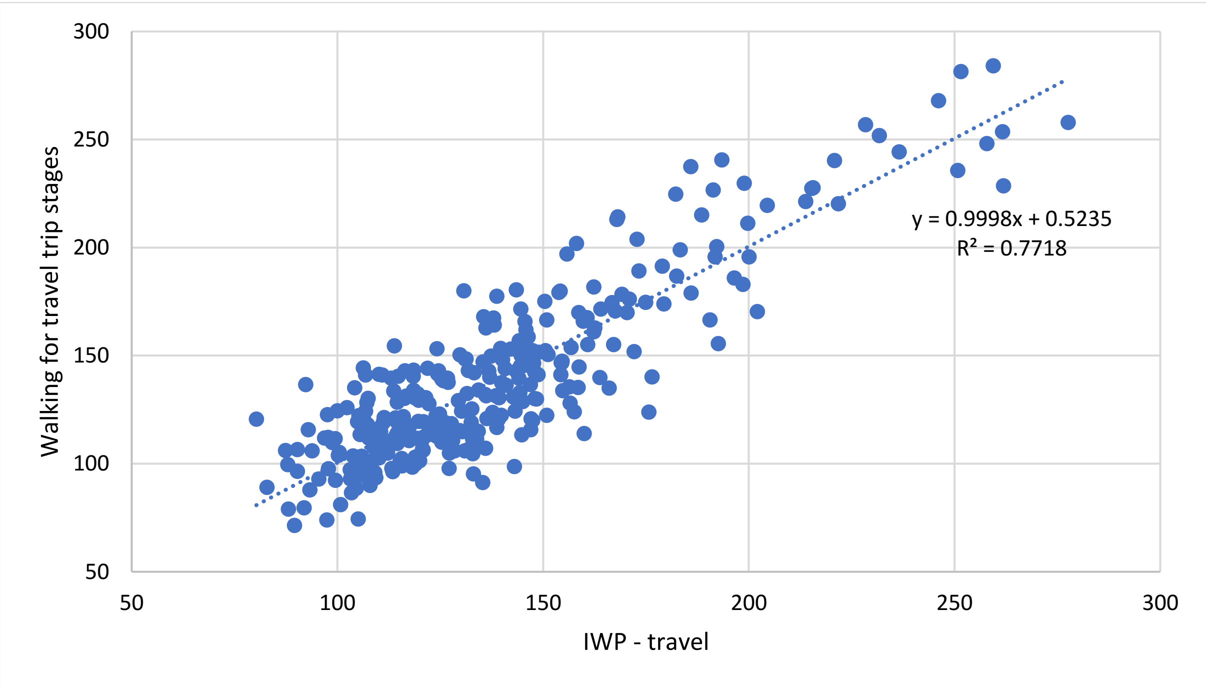
The IWP-travel score is based on the following equation[15]:

IWP-t = 0.985 (% in terraced housing etc.) + 5.789 (% female) – 420.979 (Cars pp) – 3.402 (IMD) + 54.57

This indicates that walking levels will typically be higher in areas with denser housing, and a higher proportion of women, and lower in areas with higher levels of car ownership and greater deprivation. (The relationship with deprivation is perhaps surprising, although it is notable that walking for leisure is higher in areas with greater deprivation.)

Figure 8 shows the relationship between IWP-t and baseline walking levels.

Figure 8: Relationship between Intrinsic Walking for Travel Potential and baseline walking levels



---

[14] For the intervention packages so far created, we estimated that around 70-80% of expenditure was on interventions that only affected walking for transport, 15-25% on interventions that affected both walking for transport and walking for leisure, and less than 5% on led walks, which was the only intervention that only affected walking for leisure. This means that in the majority of cases, IWP-t is the more relevant scaling factor.

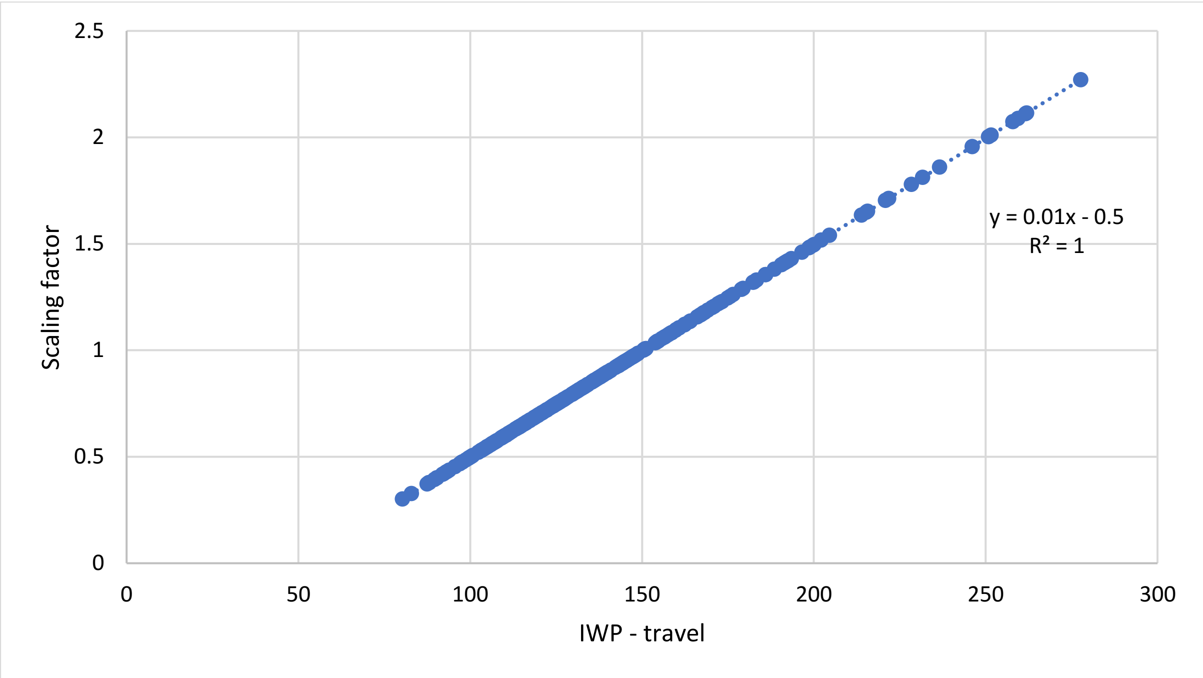
[15] p values of all coefficients <0.01; overall model F statistic p<0.01; sample 324 LADs.

Looking at the graph suggests that, for the majority of local authorities, walking for travel levels typically vary by about 2.5, depending on the IWP-travel. Hence, on the assumption that variation in the effectiveness of investment (due to the underlying characteristics of different LADs) is of a similar order to the variation in baseline levels, a scaling factor was generated for the model using the following equation:

Scaling factor = 0.01 (IWP-t) – 0.5

This produces the plot shown in Figure 9.

Figure 9: Scaling factor suggested for the IWP-t



6. Considering characteristics that could explain differences in the costs of generating a walk-to-school trip in different locations

For the Walk to School Model, a subset of the variables given in Table 1 were analysed in conjunction with the baseline dataset of the proportion of 5-10-year-olds walking to school in each LAD (as defined in Appendix 1). Table 5 gives the correlation coefficients between the proportion walking to school and each of the variables, as calculated using data for 324 LADs. The correlation coefficients[16] provide an indication of the relative strength of the relationship between walk mode share and each of the explanatory variables – i.e. the larger the magnitude of the values (either positive or negative), the stronger the relationship. The strongest relationships are with the measures of car ownership, and the measures of primary school accessibility.

Table 5: The correlation coefficients between a range of variables and the proportion of primary-aged children walking to school in 2011 (for 324 LADs)

	Proportion of primary school children walking to school
Rainfall	-0.25
PCT measure of hilliness	-0.23
Density	0.65
% households without cars	0.69
Private cars per person	-0.73
% aged 65+	-0.67
% aged 0-15~	0.30
% aged 16-64	0.68
% in terraced housing, flat, maisonette or apartment	0.65
% female	-0.32
% living in rental accommodation	0.67
2014 Minutes by PT/walk to nearest primary	-0.69
2014 No. of primary schools within 15 mins walk/PT	0.68
2014 % 5-10-ear-olds within 15 mins of primary school by PT/walk	0.69
2011 Minutes by PT/walk to nearest primary	-0.66
2011 No. of primary schools within 15 min walk/PT	0.65
2011 No. of primary schools ‘accessible’ by walk/PT	0.72
2011 % 5-10-year-olds within ‘reasonable time’ of primary schools by PT/walk	0.74
Time taken to the most local shops and services	-0.61
Time taken to intermediate shops and services	-0.59
Time taken to 8 shops and services	-0.60
Index of multiple deprivation	0.51
Journey times on A-roads	0.51

~Note that this correlates very closely with the proportion of 5-10 year olds

[16] Note that whilst correlation coefficients (R-values) provide an initial measure of relationship strength, determination coefficients (R^2-values) are needed to estimate how much the variation in one factor can be predicted from another.

7. Scaling factor used in the Walk to School Model

Regression analysis was carried out in the same way as described for the Cycling Model in section 2. It is possible to predict 64% of the variation in levels of walking to school using the following five factors:

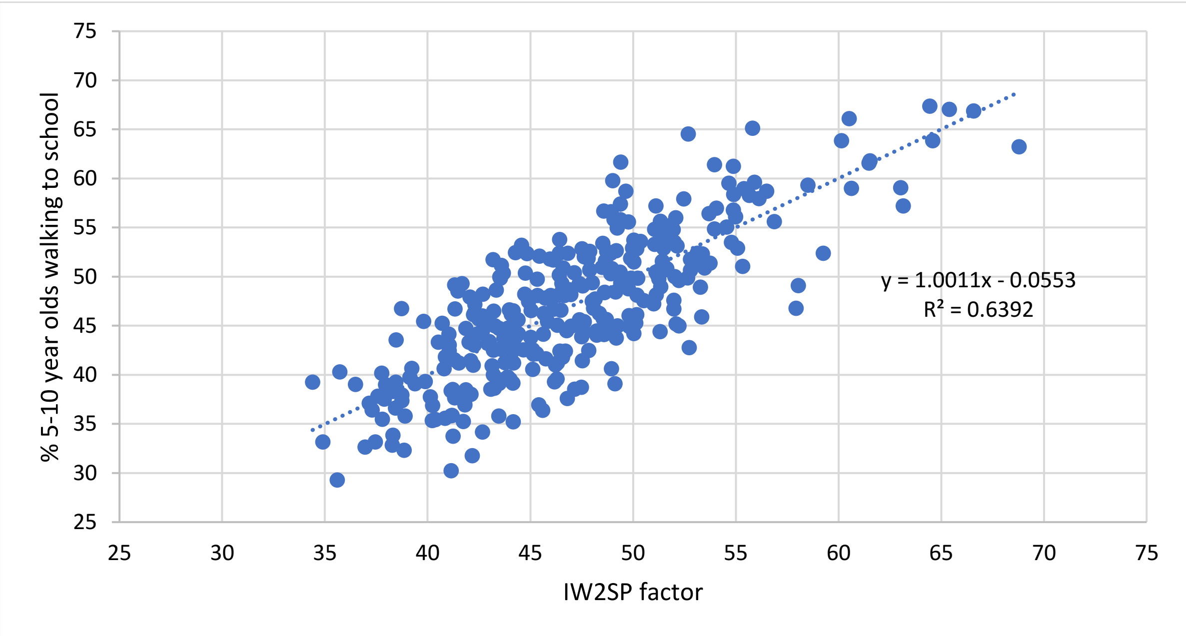
- Density
- Cars per person
- Journey times on A-roads
- Proportion of 5-10-year-olds with access to a primary school within a 'reasonable' time by public transport/walking
- Index of Multiple Deprivation

A measure of Intrinsic Walking to School Potential 'IW2SP' was created using the following equation[17]:

IW2SP = 0.862 (% with reasonable travel time) – 43.968 (Cars pp) + 0.087 (Density) - 2.488 (2011-12 journey times) – 0.173 (IMD) + 31.2

Figure 10 shows the relationship between IW2SP and baseline walking to school levels.

Figure 10: Relationship between Intrinsic Walking to School Potential and baseline walking to school levels



Looking at the graph suggests that, for the majority of local authorities, walking to school levels typically vary by a factor of about 2 (from 32-64% of pupils), depending on the IW2SP score. Hence, on the assumption that variation in the effectiveness of investment (due to the underlying characteristics of different LADs) is of a similar order to the variation in baseline levels, a scaling factor was generated for the model using the following equation:

Scaling factor = 0.021 (IW2SP)

This produces the plot shown in Figure 11.

[17] p values of all coefficients <0.01; overall model F statistic p<0.01; sample 324 LADs.

Figure 11: Scaling factor suggested for the IW2SP