# Harnessing datasets in England to understand determinants of car ownership


**The Sixth Research Data Alliance (RDA) plenary meeting, Paris - France, 23rd – 25th September 2015**

**Dr. Godwin Yeboah and Prof. Jillian Anable**

CENTRE FOR Transport Research


## Introduction

This poster presents an exploratory study using novel datasets from UK Government Agencies to understand the relationship between car ownership in English households and potential factors that might influence it. As part of an Engineering and Physical Sciences Research Council (EPSRC) sponsored project, we explore, in this poster, variety of selected/derived variables that might explain car ownership at a fine geographical scale (Lower-layer Super Output Areas - LSOAs) accounting for socio-economic and demographic factors, transport accessibility factors, built-environment factors, and natural factors *ceteris paribus*. These variables were derived from the following data sources: 2011 UK Census, 2011 Experian data, travel time data from UK Department for Transport, bus service departure information from GBRail.info feed, national public transport access nodes from UK National Public Transport Data Repository, and UK Ordnance Survey Terrain 50. The LSOAs are geographical areas classified as having minimum of 1000 population and maximum of 3000 population or minimum of 400 households and maximum of 1200 households respectively.

## Background of EPSRC Sponsored MOT Project

The work presented in this poster is part of a project which is sponsored by EPSRC and it is about the Motoring and vehicle Ownership Trends in the UK (MOT) with the following objectives:

1. **Combine new sources of data** to give a spatially and temporally disaggregated understanding of car ownership and use, and associated energy demand and emissions;
2. **Develop new methodologies**, datasets and research capability to understand the relationships between energy demand and emissions from car use, and a wide range of structural and social factors;
3. **Describe and explain the linkages** between different fuel uses, energy end uses and energy service demands at the domestic level;
4. **Develop a baseline** of spatially disaggregated energy demand from car use from which future scenarios can be developed and modelled;
5. **Assessment of social and environmental justice issues** in relation to income, fuel use and price, emissions of pollutants and exposure to impacts; and,
6. **Track changes over time and space** in order to evaluate the scale and distribution of the impacts of local transport policy interventions.

## Methodological approach

Identification and selection of explanatory variables from available literature and associated data sources. Also, authors added additional variables to test the strength of their relationship with car ownership given data availability. See table and introduction sections for data sources. Software used are ArcGIS, QGIS, SPSS, R Programming tools, Microsoft Office Applications.

Socio-economic & demographic factors | Transport accessibility factors | Built-environment factors | Natural environment factors

Statistical analysis using Pearson Correlation technique

## Examining linear dependencies - Bivariate analysis of car ownership and selected explanatory variables

| Variable Category | Variable description (mean/standard deviation) | Pearson's r | Sig. (2-tailed) | LSOA | Data source |
|---|---|---|---|---|---|
| Dependent variable | Car ownership as number of cars or vans per household (1.18/0.39) | 1.0 | | 32844 | 2011 UK Census via Infuse Service (infuse.mimas.ac.uk) |
| Socio-economic and demographic factors | Non drivers 0-15 age group (305.17/102.16) | -.1** | .000 | 32844 | |
| | **Young drivers 16-29 age group (302.51/194.70)** | **-.4**** | .000 | 32844 | |
| | Mid-age drivers 30-49 age group (451.35/119.62) | -.2** | .000 | 32844 | |
| | **Old drivers 50-69 age group (367.72/108.53)** | **.6**** | .000 | 32844 | |
| | **Very old drivers 70+ age group (187.32/88.37)** | **.3**** | .000 | 32844 | |
| | **Population density (42.63/42.26)** | **-.6**** | .000 | 32844 | |
| | Gender: Male (793.73/161.15) | -.1** | .000 | 32844 | |
| | Gender: Female (820.34/150.48) | -.1** | .000 | 32844 | |
| | Average household size - Persons per household (2.39/0.30) | .2** | .000 | 32844 | |
| | **Household Median Income (34302.41/12902.92)** | **.5**** | .000 | 32840 | 2011 Experian data |
| Transport accessibility factors | Car driving time to employment centres (5.34/3.91) | .0** | .000 | 32840 | Department for Transport (DfT): 2011 Travel time, destination and origin indicators to key sites and services, by Lower Super Output Area (LSOA). Time is in minutes. This data covers only England. |
| | Car driving time to Primary schools (5.20/3.83) | -.0 | .079 | 32840 | |
| | Car driving time to Secondary schools (5.56/4.03) | .1** | .000 | 32840 | |
| | Car driving time to Further Education institutions (6.14/4.48) | .1** | .000 | 32840 | |
| | Car driving time to GPs (5.27/3.83) | .0** | .009 | 32840 | |
| | Car driving time to Hospitals (9.06/6.07) | .2** | .000 | 32840 | |
| | Car driving time to Food stores (5.26/3.86) | .0* | .028 | 32840 | |
| | Car driving time to Town Centres (6.73/4.79) | .2** | .000 | 32840 | |
| | **Public Transport/walk time to employment centres (8.58/4.69)** | **.4**** | .000 | 32840 | |
| | **Public Transport/walk time to Primary schools (6.68/2.21)** | **.4**** | .000 | 32840 | |
| | **Public Transport/walk time to Secondary schools (11.98/7.50)** | **.4**** | .000 | 32840 | |
| | **Public Transport/walk time to Further Education institutions (14.64/10.10)** | **.4**** | .000 | 32840 | |
| | **Public Transport/walk time to GPs (8.22/3.89)** | **.5**** | .000 | 32840 | |
| | **Public Transport/walk time to Hospitals (28.89/35.77)** | **.3**** | .000 | 32840 | |
| | **Public Transport/walk time to Food stores (6.98/3.59)** | **.5**** | .000 | 32840 | |
| | **Public Transport/walk time to Town Centres (15.23/10.03)** | **.4**** | .000 | 32840 | |
| | **Cycling time to employment centres (6.49/4.70)** | **.3**** | .000 | 32840 | |
| | **Cycling time to Primary schools (5.26/1.26)** | **.3**** | .000 | 32840 | |
| | **Cycling time to Secondary schools (7.88/6.77)** | **.4**** | .000 | 32840 | |
| | **Cycling time to Further Education institutions (10.42/11.17)** | **.4**** | .000 | 32840 | |
| | **Cycling time to GPs (5.95/3.18)** | **.4**** | .000 | 32840 | |
| | **Cycling time to Hospitals (20.86/22.46)** | **.3**** | .000 | 32840 | |
| | **Cycling time to Food stores (5.89/3.56)** | **.3**** | .000 | 32840 | |
| | **Cycling time to Town Centres (11.60/11.39)** | **.4**** | .000 | 32840 | |
| | **Frequency score reflecting the availability of public bus services providing the travel times to employment centres (92.49/17.78)** | **-.5**** | .000 | 32840 | |
| Built-environment factors | Density of departures of off-street rail service (285.34/3041.1) | -.1** | .000 | 32844 | GBRail.info GTFS 1st August 2015 feed |
| | Density of off-street bus/coach stops (0.14/2.09) – stops per sqkm | -.1** | .000 | 32844 | National Public Transport Data Repository (NPTDR) – 2011 version of National Public Transport Access Nodes (NaPTAN) |
| | Density of off-street rail/metro/tram stations (0.18/1.04) – stops per sqkm | -.2** | .000 | 32844 | |
| | **Density of on-street bus/coach stops (15.10/13.32) – stops per sqkm** | **-.5**** | .000 | 32844 | |
| Natural environment factors | Degree of hilliness using the average slope per LSOA (2.16/1.67) | .1** | .000 | 32844 | Ordnance Survey (OS) Terrain 50 |
| | Degree of hilliness using the average altitude per LSOA (67.30/54.69) | .2** | .000 | 32844 | |

*** and * mean correlation is significant at the 0.01 level (2-tailed) and 0.05 level (2-tailed) respectively. Pearson Correlation coefficient or Pearson's r, developed by Karl Pearson based on a related idea by Francis Galton in the 1880s, is a measure of the linear correlation (dependence) between two variables X and Y, giving a value between +1 and −1 inclusive, where 1 is total positive correlation, 0 is no correlation, and −1 is total negative correlation. The four missing LSOA areas were: Hartlepool 003G, Nottingham 028F, Leeds 067K, and Greenwich 035D.*

## Concluding remarks and future work

Findings from linear dependency analysis suggest that with the exception of car driving time to primary schools, the remaining 40 variables were significantly (at the 0.01 and 0.05 levels) correlated to car ownership (CO) with varying correlation strengths, albeit noting that this does not imply any causality. Only correlations equal to or greater than |.3| are highlighted (in bold) to suggest their potential strength in explaining CO. As expected from previous studies, population density and household median income are negatively (-.6) and positively (.5) correlated to CO respectively. However, our analysis goes much further than this to help to explain what it is about density in particular that might impact on CO and identifies several variables as being worthy of consideration in CO modelling in England. Future work will use advanced spatial statistical techniques to explore the strength of the relationships between car ownership and these variables using the location of the registered keeper of all vehicles in the UK Ministry of Transport (MOT) vehicle test records.

## Selected Reference


- Yeboah, G., Anable, J, Chatterton, T., Barnes, J., Wilson, E. R., Turnbull, O., Cairns, S. (2015), Understanding car ownership elasticities in England and Wales: Advancing the evidence base with new data sources. In: Malleson, Nick; Addis, Nicholas; Durham, Helen; Heppenstall, Alison; Lovelace, Robin; Norman, Paul; Oldroyd, Rachel (2015): GIS Research UK (GISRUK) 2015 Proceedings. figshare. http://dx.doi.org/10.6084/m9.figshare.1491375, pp.704-713. Retrieved 10:27, Jul 22, 2015 (GMT).
- Chatterton, T., Barnes, J., Wilson, R. E., Anable, J. and Cairns, S. (2015), Use of a novel dataset to explore spatial and social variations in car type, size, usage and emissions. Transportation Research Part D: Transport and Environment. ISSN 1361-9209 [In Press].


## Contact


Professor Jillian Anable, Chair of Transport and Energy Demand
Centre for Transport Research (CTR)
Tel:+44 (01224 273795 / Email: j.anable@abdn.ac.uk
Website: http://www.abdn.ac.uk/ctr


## Acknowledgement


The work has been undertaken under EPSRC Grant EP/K000438/1. We would particularly like to acknowledge other members of the MOT project team: Prof. Eddie Wilson, Dr Sally Cairns, Dr. Tim Chatterton, Dr. Jo Barnes, Dr Oliver Turnbull. Data used contains: National Statistics data © Crown copyright and database right 2014; Contains Ordnance Survey data © Crown copyright and database right 2014. Authors also acknowledge Experian data from UKDataService - SN: 5738 Experian Demographic Data, 2004-2005 and 2008-2011.
MOT Project website is http://www.abdn.ac.uk/ctr/research/currentbr-research-projects/MOT