

Incorporating social norms into a configurable agent-based model of the decision to perform commuting behaviour

Robert Greener^{1,3}, Daniel Lewis¹, Jon Reades², Simon Miles³ and Steven Cummins¹

¹Population Health Innovation Lab, Department of Public Health, Environments & Society, London School of Hygiene & Tropical Medicine, 15–17 Tavistock Place, London, UK

²Bartlett Centre for Advanced Spatial Analysis, Bartlett School of Planning, University College London, London, UK

³Centre for Urban Science & Progress London, Department of Informatics, King's College London, London, UK

Abstract

Interventions to increase active commuting have been recommended as a method to increase population physical activity, but evidence is mixed. Social norms related to travel behaviour may influence the uptake of active commuting interventions but are rarely considered in their design and evaluation. In this study we develop an agent-based model that incorporates social norms related to travel behaviour and demonstrate the utility of this through implementing car-free Wednesdays. A synthetic population of Waltham Forest, London, UK was generated using a microsimulation approach with data from the UK Census 2011 and UK HLS datasets. An agent-based model was created using this synthetic population which modelled how the actions of peers and neighbours, subculture, habit, weather, bicycle ownership, car ownership, environmental supportiveness, and congestion (all configurable parameters) affect the decision to travel between four modes: walking, cycling, driving, and public transport. The developed model (MOTIVATE) is a configurable agent-based model where social norms related to travel behaviour are used to provide a more realistic representation of the socio-ecological systems in which active commuting interventions may be deployed. The utility of this model is demonstrated using car-free days as a hypothetical intervention. In the control scenario, the odds of active travel were plausible at 0.091 (89% HPDI: [0.091, 0.091]). Compared to the control scenario, the odds of active travel were increased by 70.3% (89% HPDI: [70.3%, 70.3%]), in the intervention scenario, on non-car-free days; the effect is sustained to non-car-free days. While these results demonstrate the utility of our agent-based model, rather than aim to make accurate predictions, they do suggest that by there being a 'nudge' of car-free days, there may be a sustained change in active commuting behaviour. The model is a useful tool for investigating the effect of how social networks and social norms influence the effectiveness of various interventions. If configured using real-world built environment data, it may be useful for investigating how social norms interact with the built environment to cause the emergence of commuting conventions.

Keywords

agent-based modelling, active travel, physical activity, car-free days

Funding: This research was funded by KCL–LSHTM seed funding. Robert Greener is supported by a Medical Research Council Studentship [grant number: MR/N013663/1]. Steven Cummins is funded by Health Data Research UK (HDR-UK). HDR-UK is an initiative funded by the UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities.

ATT'22: Workshop Agents in Traffic and Transportation, July 25, 2022, Vienna, Austria

✉ Robert.Greener@lshtm.ac.uk (R. Greener); Daniel.Lewis@lshtm.ac.uk (D. Lewis); j.reades@ucl.ac.uk (J. Reades); simon.miles@kcl.ac.uk (S. Miles); Steven.Cummins@lshtm.ac.uk (S. Cummins)

ORCID 0000-0003-2118-2360 (R. Greener); 0000-0002-2111-4256 (D. Lewis); 0000-0002-1443-9263 (J. Reades); 0000-0001-9956-4217 (S. Miles); 0000-0002-3957-4357 (S. Cummins)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

1. Introduction

Recently, policies have been introduced which aim to encourage more physically active modes of transport, particularly walking and cycling, to incorporate more physical activity into everyday life [1]. Reduced car dependency is likely to mitigate other risk factors known to impact health, including noise and air pollution, the risk of road traffic injury [2, 3, 4], all-cause mortality and cancer incidence [5, 6, 7], and obesity [8, 9]. Governments are therefore investing to improve the environment to remove active-travel barriers. However, intervention evaluations are difficult, time-consuming, and expensive, and few have produced robust impact evidence [10].

Interventions assume that environment changes will automatically result in individuals adapting their behaviour [11]. However, this ignores the role of norms of travel behaviour in influencing the uptake of such interventions. A social norm is an understanding held by members of a population as to what a ‘proper’ behaviour is under a given set of circumstances. By norms, we refer to emergent social norms, i.e., behaviour which becomes prevalent through interaction, rather than prescriptive norms (obligation, prohibition, and permission). Mobility-related social norms may relate to the desirability owning or driving a car, or the appropriateness of riding a bicycle. Social norms may differ according to population groups (age, gender, etc.) and may be conditioned by environment or culture and are often socially and spatially patterned.

A social norm encompasses two strongly related ideas: first, that individual behaviour – and the willingness to change that behaviour – is strongly influenced by both previous performances of that behaviour (‘habit’) and the social context within which the behaviour unfolds; and second, that the observation of other people’s behaviour (peers, neighbours, the public, etc.) can either shift or reinforce a behaviour. Therefore, the behaviour of individuals cannot be easily separated from the broader social and community systems in which they are embedded. Environmental interventions to promote active commuting are complex primarily due to the social and physical systems within which these interventions occur, the contextually contingent nature of impacts, and the agency of groups and individuals whose behaviours they aim to influence [12]. This suggests that to generate better representations of the systems into which active commuting interventions are deployed, incorporation of social norms related to travel behaviour is required. Using norms related to travel when evaluating active commuting interventions may help reduce uncertainty over effectiveness and support practitioners in making more informed decisions.

In this paper, we describe the development of an agent-based model that incorporates norms related to travel behaviour and then, to demonstrate its utility, we simulate the impact of introducing car-free days as a hypothetical use case. We present MOTIVATE (Modelling Normative Change in Active Travel), an open-source, agent-based model, which has configurable built environments and populations, that simulates the shift to active commuting as a result of interventions. We briefly review the related work and its relation to ours in Section 1.1. We describe in Section 2 how we generate a synthetic population using a microsimulation approach, before describing the design and development of MOTIVATE. Then, we describe how we implemented an intervention of car-free Wednesdays, followed by the statistical methods used. We then provide the results of this intervention – car-free days increased the odds of an active journey by 77.7% (HPDI: [77.7%, 77.7%]). Finally, we discuss our results in comparison to existing transport models, while recognising the strengths and limitations.

1.1. Related work

Agent-based models (ABMs) are a tool that allow the simulation of the dynamic processes of behaviour change at the population level (the model) by accounting for interactions between heterogeneous individuals (the agents) and their environment [13]. While there exist ABMs exploring transport systems at a low-level [14], there do not exist models that investigate the impact of social norms on transport behaviour. The existing work is focussed mainly on traffic flow and route decisions [15, 16, 17]; our work focusses on the decision for the method of commuting and is not focussed on modelling the flow of people through an environment. We aim to produce a configurable agent-based model of the decision process which takes a wide range of inputs resulting in commuting behaviour. This differs from existing simulations which focus on how agents flow through cities, with built environment and socio-demographic attributes being a key determinant of commuting behaviour. By focussing on the decision, we can focus our efforts on how a diverse range of inputs, such as the actions of peers and neighbours, habit, subcultures, congestion, bicycle and car ownership, and dynamic characteristics such as the weather affect the decision to undertake different transport methods. This allows us to explore the parameters' effect on interventions, without being concerned about traffic flow.

Furthermore, there do not exist models of the decision to commute that incorporate social norms. In our model, social norms exist as the joint effects of friends, neighbours, and a mobility culture to which agents belong. Other work has modelled how the actions of friends influences norms around smoking cessation using agent-based modelling [18]. Further normative agent-based models have shown that cultural heterogeneity can cause heterogeneity in results [19]. Applied agent-based modelling work has shown how norms, the (food) environment, and peer influence interact to cause changes in fruit and vegetable consumption [20].

The methods for incorporating norms is discussed in greater detail elsewhere [21]; however, here we primarily focus on the emergence of norms through transmission [22, 23, 24]. We also model how the norms influence behaviour through *internally directed enforcement* [25, 26, 27, 28], where there is internal pressure to conform with the norm due to the inherent social desire to conform with the prevailing norm. The sanction for not conforming is implicit, as there is a benefit to fitting in that is lost if not conforming.

While there exist many studies investigating norms in health outcomes there do not exist ones that use norms to investigate active commuting behaviour. Our model addresses this by exploring the decision to perform active commuting by including an agent's norm in their decision-making. These norms emerge through observing the behaviours of friends and neighbours and through cultural influence, and they influence behaviour through the desire to fit-in.

2. Materials & Methods

To develop an ABM with relevance, we selected the London Borough of Waltham Forest (UK), who have been a leader in interventions to increase active travel [29], as a reference, which could be replaced by other users. Our approach involved three steps: (i) create a synthetic population, mirroring Waltham Forest; (ii) create an ABM of active commuting; and (iii) demonstrate the utility of the ABM through a hypothetical active travel intervention: car-free Wednesdays.

2.1. Synthetic population generation using microsimulation

To generate realistic agent behaviours, an Iterative Proportional Fitting (IPF) microsimulation approach [30, 31] was used to assign attributes (e.g., sex, ethnicity, distance to workplace) as a constraint on modal choice. IPF microsimulation is a well-established statistical method for commuting behaviour [30] which will seek to generate synthetic population based upon a series of constraints from observed data. Data from the UK Household Longitudinal Survey (UK HLS) (2014–16) [32] on active commuting behaviour for London (n=6 310) was combined with 2011 Census population data for Waltham Forest [33] using IPF to produce a realistic agent population whose socio-demographic attributes and behaviours were highly correlated (99.9%).

Following integerisation and expansion using the Truncate, Replicate and Sample (TRS) approach [31] we identified an absence of questions about bicycle access from the ‘ethnic boost’ component of the UK HLS dataset. To solve this, we modelled access to a bicycle using a logistic regression model conditioned on sex, age group, ethnic group, employment status and car usage. Inferring access in this way yielded a synthetic population in which 46.5% of people had access to a bicycle, similar to the National Travel Survey estimate of 42% [34].

Table 1: Modal distribution of commutes.

Mode	Local (%)	City (%)	Beyond (%)
Car	43.5	25.9	54.0
Cycle	3.5	2.8	0.8
Public Transport	31.5	71.1	45.2
Walk	21.5	0.2	0.0

was 9.36 km; and public transit distance was 10.72 km. These are the average distances that individuals in Waltham Forest commute on a daily basis. The empirical distributions informed the choice of either a log-normal or Gaussian Mixture Model to generate individual agent commute distances. Commute distances could not be negative, or greater than 10 km (walking), 40 km (cycling), or 80 km (public transit or private car) as these are infeasible. Values outside these bounds were ignored, with another value being drawn. Commutes were then classified as local (0 to 4 943 m; mean=2 588 m), city (4 944 to 20 059 m; mean=10 536 m), or beyond (>20 059 m; mean=31 096 m), yielding the modal distribution (Table 1).

Commute distances reported in the UK HLS data were often rounded (e.g., 5, 10, 15, etc.) and thus are subject to misclassification. To generate more accurate measures of commute distance we used data from a ‘safeguarded’ 2011 Census table [35]. For Waltham Forest, we observed that the mean straight line walking distance for an individual was 2.64 km; cycling distance was 7.72 km (with a bimodal distribution); car distance

2.2. Model development

The agent-based model was developed in Rust [36]. The source code for the model and for the statistical analysis, as well as the raw results, are publicly available to download [37].

In developing the model we deemed the following variables to be of importance: the weather, as it would likely impact walking and cycling; environmental supportiveness, as neighbourhoods better supporting active travel should have more active travel; congestion, as people living in neighbourhoods which have congested roads may be more likely to walk or cycle; subculture, there exist certain mobility subcultures affecting their propensity to adopt active travel modes [38]; the actions of neighbours and peers, as there is likely to be a peer influence; commute length,

Table 2
Global-level variables in the model.

Parameter name	Values	Description
Weather pattern	$f : \text{Time} \rightarrow (\text{Wet}, \text{Dry})$	For a given time, is the weather Wet or Dry.
Weather modifier	$f : (\text{Weather}, \text{TransportMode}) \rightarrow [0, \infty)$	For a given weather and mode, how is the agent's desire to take a mode affected. This is applied multiplicatively, such that values < 1 reduce the likelihood of the mode being taken, and values > 1 increase the likelihood.

as at some distances, some modes are impractical; habit, as what you have done before influences what you will do in the future; bicycle ownership; and car ownership. The weather pattern, environmental supportiveness, capacity (which determines congestion), subculture, commute length, bicycle and car ownership are all configurable parameters to the model. Figure 1 gives the relationship between all the global-level, neighbourhood-level, and agent-level variables and travel behaviour, using arrows to denote the direction of influence. This shows the higher-level structure of the model, without focussing on implementation details (for this, see the class diagram in Appendix A). These parameters will be referred to in Tables 2, 3, and 4.

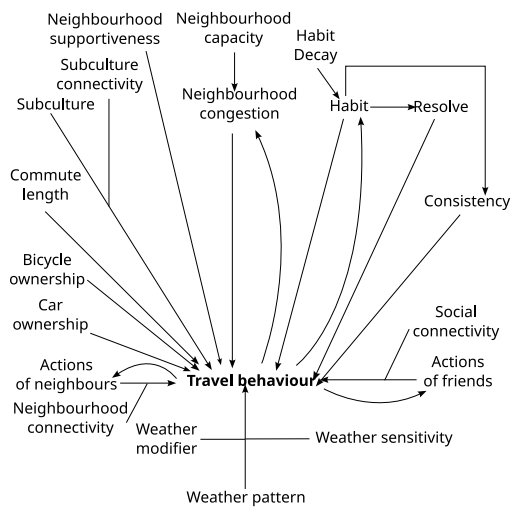


Figure 1: Relationship between global-level, neighbourhood-level, and agent-level variables and travel behaviour.

There are two global-level variables which both relate to the weather. These are given in Table 2. The first is the *weather pattern*. This was generated by taking a Markov chain model with two states: Wet and Dry. This was informed from historical daily weather data [39] from 2017; days with over 4.4 mm of rain were chosen as Wet days, other days were Dry. The threshold of 4.4 mm is the point at which daily hires from TfL's Cycle Hire scheme fell. The Markov chain was used as it is more likely to be bad weather if the previous day was bad weather. The second is the *weather modifier*, this states that for a given weather and mode how is the agent's desire to take that mode changed. In our parameterisation, when bad weather, there is a negative influence on walking and cycling.

We defined 20 neighbourhoods, based upon electoral wards in Waltham Forest in London, UK. Each of these neighbourhoods has a number of parameters, given in Table 3. The *supportiveness* value describes how much a given neighbourhood supports a given transport mode and, if parameterised in future works, could be taken as a proxy for features such as the number of

Table 3
Neighbourhood-level variables in the model.

Parameter name	Values	Description
ID	String (text)	The ID of the neighbourhood
Supportiveness	$f : \text{TransportMode} \rightarrow [0, 1]$	Supportiveness of neighbourhood for mode.
Capacity	$f : \text{TransportMode} \rightarrow [0, \infty)$	The capacity for a mode in a neighbourhood. Once exceeded there is congestion.
Congestion	$\begin{cases} 1 & \text{if capacity not exceeded} \\ 1 - \frac{\text{Journeys}(t,m) - \text{Capacity}(m)}{\text{Population}} & \text{otherwise} \end{cases}$	This is calculated every time step. It is a negative influence proportional to how much capacity for a mode m was exceeded at each time step t .

public transport stations, number of cycle lanes, width of pavements, quality of roads, etc. In our implementation, this means that the greater this value, the better the neighbourhood supports a given mode. This value is used, rather than individual components of the environment, to allow other users of the model to include parts of the environment of interest by converting their values to a value between 0 and 1. For each neighbourhood and transport mode, there is also a *capacity* beyond which there is *congestion*, which serves a negative influence on an agent’s decision-making. Congestion is proportional to the excess transport (i.e., that in excess of the capacity). The more congestion there is for a mode, the less likely it is an agent will take it.

We also define three subcultures to which agents can belong; these associate a *desirability* with each mode which expresses how ‘desirable’ it is to an agent belonging to that subculture. This also falls within the range 0 to 1. There was driven by a substantive body of work surrounding the role of cultural norms in mobility choices and policy formulation [40, 41]. Informed by this work, we developed three hypothetical subcultures. Consequently, the first subculture, *A*, in which modes are ordered cycling, driving, walking, and public transport. This group values active travel as part of a healthy lifestyle, but also sees it as a consumption choice in which money is spent on both bicycles and cars as signifiers of success, with public transit being seen as the least desirable in that sense. The second subculture, *B*, is a pro-driving subculture, in which modes are ordered driving, public transport, walking, and cycling. People belonging to this culture are averse to active commuting; this is reflected in the ordering. The final subculture, *C*, is a pro-active-travel subculture, with the ordering: walking and cycling (equal), then public transport, then driving. People belonging to this subculture value taking active journeys.

We defined 111 166 agents in the model – the number of commuters in Waltham Forest, as identified by the microsimulation approach. The microsimulation approach described previously was used to generate the synthetic population used in the model. The agents were randomly allocated to a neighbourhood with the probability determined by the microsimulation approach based upon car ownership, bicycle ownership, and population of the neighbourhood, and were randomly allocated to a subculture uniformly. Two social networks that agents belong to were also generated randomly. The first network is a global social network across all agents, representing social and work relationships. The second is a neighbourhood-wide ‘neighbour network’, representing the observation of other local (i.e., within the same neighbourhood / electoral ward) agent behaviours by each individual network. The global network is a Watts-Strogatz small-world network ($k = 3; \beta = 0.6$) [42] and the neighbour network is a scale-free

network generated using the Barabási-Albert model of preferential attachment ($m_0 = 10$) [43]. The small-world network was chosen as networks of friends were expected to exhibit small-world properties and the scale-free network was chosen as there are expected to be highly influential neighbours. The use of these two networks, as opposed to one global network, is because friends are expected to have greater influence, and because the properties of the two networks are likely to differ. Each agent also has a habit value for each mode which represents their recent actions. This is calculated as the exponential moving average of their choices. This was chosen as more recent actions are expected to have greater influence. The parameters used to configure the agents are described in Table 4.

Table 4
Agent-level variables in the model.

Parameter name	Values	Description
ID	0-∞	Each agent has a unique identifier.
Subculture	ID of the subculture	Each agent belongs to a subculture which has a desirability for each commute mode.
Neighbourhood	ID of the neighbourhood	Each agent resides in a neighbourhood which has a supportiveness, capacity, and congestion modifier for each commute mode.
Commute length	Local, City, or Distant	The commute length. Local means within the modelled borough; City means across boroughs; Distant means outside the city.
Weather sensitivity	0-1	How sensitive the agent is to the weather when making commuting decisions. Higher values mean more sensitive.
Consistency	0-1	How consistent the agent is; i.e., how much does their habit matter to them. Higher values mean more consistent.
Social connectivity	0-1	How important the actions of an agent's connections in their social network are to them. The greater the value, the greater the importance the agent attaches to the actions of their friends.
Subculture connectivity	0-1	How important the desirability of different commute modes from the subculture is. The greater the value, the more important the subculture is in their decision-making.
Neighbourhood connectivity	0-1	How important the actions of an agent's connections in their neighbour network are to them. The greater the value, the agent attaches more importance to the actions of their neighbours.
Habit decay	0-1	How far into the recent past should affect the agent's decisions. The greater the value, the longer in the past the agent will consider when determining their habit. At each time step, the previous exponential moving average is multiplied by the habit decay; small values cause it to decay faster.
Resolve	0.9, 1.0, or 1.1	This is calculated at every time step. This represents the resolve an agent to perform an active mode despite poor weather. 0.9 if it was wet the previous day and an active mode was taken. 1.0 if it was dry the previous day. 1.1 if it was wet the previous day and an inactive mode was taken.
Bicycle owner	True / False	Whether the agent owns a bicycle.
Car owner	True / False	Whether the agent owns a car.

In the simulation, we seek to explore how an individual's social and built environment influence decisions related to choice of commute mode. We assume that agents choose rationally, taking into account the benefits and cost of each mode. We calculate two constructs: a budget (representing the benefits) and a cost. The budget describes, in an ideal world, commuting preference, taking into account actions of friends, actions of neighbours, subculture, habit, and the previous congestion they have experienced. The cost describes how commuting distance, environmental supportiveness, and the weather prevents commuting from being undertaken.

$$\text{norm}(a, m, t) = \frac{\text{number of friends of } a \text{ performing } m \text{ at time } t-1}{\text{number of friends of } a} \cdot \text{social_connectivity}(a) + \frac{\text{number of neighbours of } a \text{ performing } m \text{ at time } t-1}{\text{number of neighbours of } a} \cdot \text{neighbourhood_connectivity}(a) + \text{subculture_desirability}(\text{subculture}(a), m) \cdot \text{subculture_connectivity}(a) \quad (1)$$

$$\text{budget}(a, m, t) = \text{congestion_mod}(\text{neighbourhood}(a), m, t) \cdot (\text{norm}(a, m, t) + \text{habit}(a, m, t) \cdot \text{consistency}(a)) \quad (2)$$

Every day (t), each agent (a), for each mode (m), calculates their norm (Equation 1). This is a combination of the actions of their friends, neighbours, and the pressure from their subculture.

By combining friends, neighbours, and subculture, it describes what there is social pressure for, and therefore describes “a standard or pattern of social behaviour that is accepted in or expected of a group” [44]. Behaviours of agents also influence the norms of others, resulting in feedback loops in the system. A *budget* for each mode is created (Equation 2); where if there was nothing preventing the use of the mode, this would be the willingness to perform it. Added to the norm is the *habit*, discussed previously, weighted by how *consistent* the agent is. This value is then multiplied by the *congestion modifier* (Table 3). Finally, these budgets are ranked for the four modes to create a preference. This is the *ranked travel budget*; where if there was nothing preventing the use of the mode, this would be the order of preference.

$$\text{cost}(a, m, t) = \frac{1}{2} \cdot (\text{commute_distance_cost}(\text{commute_distance}(a, m) + (1 - \text{neighbourhood_supportiveness}(\text{neighbourhood}(a, m)))) \cdot (\text{weather_sensitivity}(a, m) + \text{weather_modifier}(m, t) + \text{resolve}(a, m, t)) \quad (3)$$

The *cost* of taking each mode is then calculated (Equation 3). The commute cost due to the distance and the cost due to the neighbourhood supportiveness are averaged. Each agent’s weather sensitivity is multiplied by the *weather modifier* for the environment (Table 2) and by the agent’s *resolve*. *Resolve* is a calculated value which reduces the cost of performing an active mode if there was wet weather the previous day and an active mode was taken; it increases the cost if there was wet weather and an active mode was not taken. There is no effect if the weather was dry the previous day. *Resolve* was implemented as it was hypothesised that those who use active modes on a poor weather day are more likely to on a following poor weather day, and those who did not use active modes are less likely to on a following poor weather day. This weighted weather modifier is multiplied by the average cost. This is ranked in reverse – modes with a higher cost will have a lower rank – and called the *reverse ranked travel cost*.

Agents will then add the ranked travel budget and reverse ranked travel cost together, and will then choose to select the mode with the highest combined rank which is available to them. In other words, they cannot drive if they do not own a car, and they cannot cycle if they do not own a bicycle. This means that agents choose the commute mode with the greatest gap between budget and cost. This is primarily an economic calculation where the cost describes the (not fiscal) cost to an individual, taking into account the weather and the environment, and the budget describes the willingness to perform an action, based upon perceived benefit.

The approach was chosen as it corresponds well to work identifying the (not just fiscal) cost and benefits (such as health benefits) of commuting as key drivers of commuting behaviour [45]. While alternative, more domain-agnostic, models of behaviour, such as BDI [46] could have been used, this approach fits the existing public health literature around active travel better. We presented this method to domain stakeholders, including professionals working in local government, who agreed that it seemed a reasonable reflection of the decision process.

The parameters are configurable – a future model could use this as a base. However, depending on the criticality of results, calibration and validation against real-world data would be required.

Intervention scenario: We evaluate the implementation of a hypothetical intervention ‘car-free Wednesdays’ which prohibits the use of cars on that day. This was compared to a control scenario of no intervention. In the car-free days (CFD) scenario, for the first year (day <

365), there were no car-free days. From the first year onwards (day ≥ 365), there was a car-free day every Wednesday (day $\bmod 7 = 2$). The model was then run for four further years.

2.3. Statistical analysis

For each intervention tested, a one-year burn-in period after the intervention is excluded. This was chosen to ensure that the effect of the intervention had become stable. For each day the model outputs the count of active journeys over all, as well as separated by subculture and neighbourhood. These can be converted to a proportion by dividing the value by the total count or the count by subculture and neighbourhood, as appropriate. This is outputted as a CSV file. A Bayesian binomial regression model (Model 4) was fitted predicting the number of active journeys made as a proportion of total journeys between 1 and 4 years post-intervention.

$$\begin{aligned}
 y_{ij} &\sim \text{Binomial}(n_{ij}, \pi_j) \\
 \text{logit}(\pi_j) &= \alpha_j \\
 \alpha_j &\sim \text{Student-T}(3, 0, 1)
 \end{aligned}
 \tag{4}$$

y_{ij} is the number of active journeys (i.e., those made by walking or cycling) in the i th scenario (1=control; 2=CFD) for the j th population, which varies only by the 200 randomly generated social network structures. n_{ij} is the total number of journeys (population multiplied by number of days). A Student-T(3, 0, 1) prior was used for the intercept (α_j). This weakly informative prior has heavy tails to allow for extreme values. The odds ratio (OR) of CFD vs. control is $\exp(\alpha_2 - \alpha_1)$.

This estimates the number of active journeys as a function of the total number of agents in the population and scenario (control vs. CFD) (n_{ij}) and the probability of any given journey being active (π_j). The logit of the probability ($\ln(\pi_j / (1 - \pi_j))$) is defined as α_j which changes only based upon the scenario. This statistical model assumes that the probability of an active journey depends on whether the simulation is in the control or car-free days scenario. The *prior distribution* Student-T(3, 0, 1) constrains the potential values of α_j to what could be reasonably expected. This is a Bayesian binomial regression model [47], which, in layman's terms, predicts a combination of binary outcomes, active travel (1) or not (0), based upon an independent variable (the scenario). A property of this is that the exponentiated α_j coefficient is the odds of an active journey – i.e., for an inactive journey there are $\exp(\alpha_j)$ active journeys.

$$\begin{aligned}
 y_{ijk} &\sim \text{Binomial}(n_{ijk}, \pi_{jk}) \\
 \text{logit}(\pi_{jk}) &= \alpha_j + \beta_j \cdot \text{Wednesday?}_{jk} \\
 \alpha_j, \beta_j &\sim \text{Student-T}(3, 0, 1)
 \end{aligned}
 \tag{5}$$

Model 5 extends Model 4, where k is 1 if the count of active journeys is not from Wednesdays, 2 if it is. The value of Wednesday?_{jk} is 1 if the observation is from a Wednesday (i.e., $k = 2$), 0 otherwise. This separates the results for the car-free days from non-car free days. Everything else remains the same. The OR of Wednesdays vs. non-Wednesdays in the control is $\exp(\beta_1)$; in the CFD scenario it is $\exp(\beta_2)$. The OR for Wednesdays in the CFD

vs. control is $\exp((\alpha_2 + \beta_2) - (\alpha_1 + \beta_1))$; for non-Wednesdays it is $\exp(\alpha_2 - \alpha_1)$.

Model 5 extends Model 4 by separating the effect of Wednesdays vs. non-Wednesdays. In this model, α_j is the (logit) probability of an active journey in scenario j on a non-Wednesday and β_j is the increase in the (logit) probability of an active journey in scenario j because it is a Wednesday. A property of this model is that by exponentiating parameters, as described in the preceding paragraph, odds ratios (OR) can be obtained. These describe the percentage change in

the odds as a result of (for example) it being a Wednesday vs. a non-Wednesday in the car-free days scenario ($\exp(\beta_1)$). These models impose no structure upon the data, only assuming that the probability is as a result of the scenario, and (for Model 5) whether the day is a Wednesday.

Models were fit using Julia [48] and Turing [49] (using DynamicNUTS [50, 51]). Four chains were used with 1 000 warm-up and 1 000 sampling iterations per chain. Trace plots assessed convergence. Prior predictive checks [47, 52] assessed the suitability of the prior distributions. Posterior predictive checks [47, 52] assessed the fit. The distribution was summarised using the posterior mean and the 89% highest posterior density interval (HPDI) [47, 52]. 89% gives greater numerical stability [53, 52], and is an interval for which the parameter is expected to lie.

3. Results

Since we observe the entire population in each scenario ($n = 111\,166$ per simulation – the commuting population in Waltham Forest) and obtain daily results ($n = 783$ – every weekday for 3 years [from start of Y3 to end of Y5]) the generated credible intervals are narrow. For the first model (Model 4), in the control the odds of an active journey (walking or cycling) (α_1) were 0.091 (89% HPDI: [0.091, 0.091]); i.e., for every inactive journey, in the control there were 0.091 active journeys. CFD increased active journeys by 77.7% (OR 1.777; 89% HPDI: [1.777, 1.777]) compared to the control scenario; i.e., there was a 77.7% relative increase to the odds of 0.091.

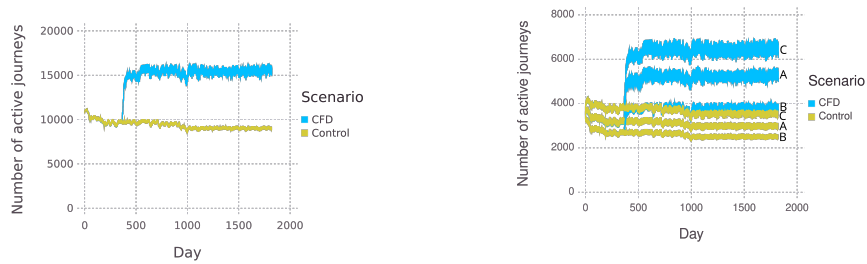
For Model 5, in the control scenario, the odds of an active journey on a non-Wednesday (α_1) were 0.091 (89% HPDI: [0.091, 0.091]). In the CFD scenario, the odds of an active journey on non-Wednesdays were 70.3% (OR: 1.703; 89% HPDI: [1.703, 1.703]) compared to the control. In the control, the odds of an active journey were 0.1% (OR: 1.001; 89% HPDI: [1.001, 1.001]) greater on Wednesdays than non-Wednesdays. In the CFD scenario, the odds of an active journey were 22.4% (OR: 1.224; 89% HPDI: [1.224, 1.224]) greater on a Wednesday compared to a non-Wednesday. Finally, the odds of an active journey were 108.2% (OR: 2.082; 89% HPDI [2.082, 2.082]) greater on Wednesdays in the CFD scenario compared to the control.

Figure 2a shows the CFD scenario, with the moving average of active commutes separated by run and scenario. A run in the same scenario differs only by the social networks, which are generated randomly. Focussing on the long-run, we observe that car-free days causes a step change active commutes. Also, it appears to introduce more instability, as shown in the figure.

Figure 2b shows the CFD scenario's effect across the 3 subcultures. Focussing on the long-run trend, in the control scenario, as expected, the pro-driving subculture (B) has the least number of active journeys. This is followed by the pro-cycling subculture (A) and then the pro-active-travel subculture (C). The CFD intervention causes the pro-driving subculture (B) to perform more active journeys than the pro-active-travel subculture (C) in the control scenario. The others are greater in turn (C is greater than A which is greater than B). The gap between the subcultures is also widened with the intervention having most effect in the pro-active-travel group.

4. Discussion

This paper presents MOTIVATE, an agent-based model which incorporates social norms related to travel in simulations of active commuting interventions. We demonstrated the model through



(a) Separated by scenario and run (b) Separated by scenario, run, and subculture (A / B / C)

Figure 2: 14 day moving average of active commutes. The scenarios are CFD and Control. For 200 networks each with the same population, each scenario was run. The weather pattern was constant. The dashed vertical line indicates the date at which the intervention was introduced. $n = 111\,166$.

a scenario introducing car-free days vs. a control. Introducing car-free days increased the odds of active travel by 77.7% (89% HPDI: [77.7%, 77.7%]) in the four years following intervention implementation. This is a large relative increase; however, absolute effects are modest, as the odds in the control were 0.091 (89% HPDI: [0.091, 0.091]) – a 77.7% increase on the initial rate of 8.3% (the percentage is odds / 1 + odds) is an increase of ≈ 6.4 percentage points.

Car-free days force an individual to make a change to commuting behaviour once per week. This constraint not only reduces the use of inactive modes on that particular day but may also act as a ‘nudge’ to change overall commuting preferences by exposing individuals to active commuting and shifting norms around active commuting on non-car-free days. This is shown in our model, as on non-Wednesdays, in the car-free days intervention scenario, the odds of active travel were 70.3% (OR: 1.703; 89% HPDI: [1.703, 1.703]) greater than the control scenario.

A ‘nudge’ of car-free days, in our model, causes a lasting change in behaviour. Unless agents need to divert from their habit, they will not. By forcing a change in habit, this results in a change in norm. The short-term change in habit due to car-free Wednesdays results in the influence given as peers and neighbours changing, resulting in a change in norm for those observing the new non-car-free behaviour – the forced change in habit destabilises the norm.

Our model differs from existing models of transport [14, 15, 16, 17]; it is not a traffic simulation – we do not aim to model traffic flow through a city. The intended purpose is not to make accurate predictions of how people commute to work. Rather, it is to explore how a range of inputs, including multiple social networks capturing the effect of peers, as well as neighbours, habit, and variable environment parameters, such as the weather, affect the decision to perform behaviour. For example, in Figure 2b we show how the CFD intervention may affect members of different psychological subcultures differently. We also see that habit causes the CFD intervention to be sustained on non-car-free days, where there were 70.3% (OR: 1.703; 89% HPDI: [1.703, 1.703]) more active journeys than the control scenario. This model can be used to explore the effect of interventions on various aspects of the decision-making process, such as habit, social networks, congestion, and the actions of peers and neighbours. This is a complement to models that more accurately model the spread of individuals through urban environments. By combining the evidence generated by real-world studies, traffic simulations, and simulations of the decision-

making process (such as ours), a greater understanding of transport behaviour may be attained.

In common with all simulations, the results presented in this paper are not necessarily an accurate prediction of real-world effects; the only way to obtain this is to perform the intervention in the real-world. However, the utility of the model is in allowing the comparison of the effectiveness of interventions under consistent conditions in order to aid decision-making. In this study, car-free days are effective. Such information, accounting for model assumptions, may help better select and prioritise active commuting interventions in the absence of strong evidence on effectiveness. It may also help funders prioritise evaluations of promising interventions identified by simulations [54]. This model may be helpful to understand currently radical policies and how the underlying social structures impact the effectiveness of these interventions.

4.1. Strengths and limitations

The strengths in this approach have been the use of microsimulation to generate a realistic synthetic population for use in the model. This has been grounded in data from official statistics. The open-source nature of the model allows for adaptation and use of the model by others.

Future work could address some weaknesses. The environmental data is not calibrated – geographic data could do this. Personal journeys and the processes changing car and bicycle ownership could be considered. Encouragement to purchase bicycles or give-up cars may work, though we cannot currently assess this. The fiscal cost of travel is also disregarded; this is a major factor. Future work could consider the effect of price given your social and built environment. There is room to include variables as part of the travel budget or cost as needed.

The HPDIs cannot be interpreted as intervals of what would happen in the reality. Normally, statistical models directly model real-world data, as such the statistical model is one of the real-world data-generating process, and assuming that the model is not mis-specified intervals can be interpreted in the terms of the real world. This is not the case in this statistical model. Here, the model it is directly modelling results from the agent-based model, and it can only be interpreted in terms of the simulated model ‘world’ – the agent-based model is the data-generating process. In order to interpret it in terms of the real world, we would need to strengthen the link between the simulated model world and reality through further calibration and validation.

4.2. Conclusion

The MOTIVATE model and the findings presented are an initial step in helping to improve understanding of active commuting interventions which seek to increase population physical activity. The presented agent-based model incorporates social norms related to travel behaviour in order to provide a more realistic representation of the socio-ecological systems in which active commuting interventions are deployed. The utility of the model has been demonstrated by simulating introducing car-free Wednesdays. Utilising *in silico* studies, such as MOTIVATE and more traditional traffic simulations, could be a cost-effective way of aiding public health decision-making and setting research priorities. This ABM is representative of new classes of model that draw on both social and environmental effects that point towards insights across applications not only in health, but in any area where place and network-based effects overlap.

References

- [1] M. Dinu, G. Pagliai, C. Macchi, F. Sofi, Active Commuting and Multiple Health Outcomes: A Systematic Review and Meta-Analysis, *Sports Med* 49 (2018) 437–542. doi:10.1007/s40279-018-1023-0.
- [2] L. B. Andersen, Active commuting is beneficial for health, *BMJ* 357 (2017) j1740. doi:10.1136/bmj.j1740.
- [3] E. Flint, E. Webb, S. Cummins, Change in commute mode and body-mass index: Prospective, longitudinal evidence from UK Biobank, *Lancet Public Health* 1 (2016) e46–e45. doi:10.1016/S2468-2667(16)30006-8.
- [4] J. Jarrett, J. Woodcock, U. K. Griffiths, Z. Chalabi, P. Edwards, I. Roberts, Effect of increasing active travel in urban England and Wales on costs to the National Health Service, *Lancet* 379 (2012) 2198–2205. doi:10.1016/S0140-6736(12)60766-1.
- [5] C. A. Celis-Morales, D. M. Lyall, P. Welsh, J. Anaderson, L. Steell, Y. Guo, R. Maldonado, D. F. Mackay, J. P. Pell, N. Sattar, J. M. R. Gill, Association between active commuting and incident cardiovascular disease, cancer, and mortality: Prospective cohort study, *BMJ* 357 (2017) j1456. doi:10.1136/bmj.j1456.
- [6] J. Panter, O. Mytton, S. Sharp, S. Brage, S. Cummins, A. A. Laverty, K. Wijndaele, D. Ogilvie, Using alternatives to the car and risk of all-cause, cardiovascular and cancer mortality, *Heart* 104 (2018) 1749–1755. doi:10.1136/heartjnl-2017-312699.
- [7] R. Patterson, J. Panter, E. P. Vamos, S. Cummins, C. Millett, A. A. Laverty, Associations between commute mode and cardiovascular disease, cancer, and all-cause mortality, and cancer incidence, using linked Census data over 25 years in England and Wales: A cohort study, *Lancet Planet Health* 4 (2020) e186–e194. doi:10.1016/S2542-5196(20)30079-6.
- [8] E. Flint, S. Cummins, A. Sacker, Associations between active commuting, body fat, and body mass index: Population based, cross sectional study in the United Kingdom, *BMJ* 349 (2014) g4887. doi:10.1136/bmj.g4887.
- [9] E. Flint, S. Cummins, Active commuting and obesity in mid-life: Cross-sectional, observational evidence from UK Biobank, *Lancet Diabetes Endocrinol* 4 (2016) 420–435. doi:10.1016/S2213-8587(16)00053-X.
- [10] J. Panter, C. Guell, D. Humphreys, D. Ogilvie, Title: Can changing the physical environment promote walking and cycling? A systematic review of what works and how, *Health Place* 58 (2019) 102161. doi:10.1016/j.healthplace.2019.102161.
- [11] S. Cummins, S. Curtis, A. V. Diez-Roux, S. Macintyre, Understanding and representing ‘place’ in health research: A relational approach, *Soc Sci Med* 65 (2007) 1825–1838. doi:10.1016/j.socscimed.2007.05.036.
- [12] G. F. Moore, R. E. Evans, J. Hawkins, H. Littlecott, G. J. Melendez-Torres, C. Bonell, S. Murphy, From complex social interventions to interventions in complex social systems: Future directions and unresolved questions for intervention development and evaluation, *Evaluation* 25 (2019) 23–45. doi:10.1177/1356389018803219.
- [13] T. Filatova, P. H. Verburg, D. C. Parker, C. A. Stannard, Spatial agent-based models for socio-ecological systems: Challenges and prospects, *Environ Model Softw* 45 (2013) 1–7. doi:10.1016/j.envsoft.2013.03.017.
- [14] K. W. Axhausen, ETH Zürich, The Multi-Agent Transport Simulation MATSim, Ubiquity

Press, 2016. doi:10.5334/baw.

- [15] A. Naiem, M. Reda, M. El-Beltagy, I. El-Khodary, An agent based approach for modeling traffic flow, in: 2010 The 7th international conference on informatics and systems (INFOS), IEEE, 2010, pp. 1–6.
- [16] K. Hager, J. Rauh, W. Rid, Agent-based modeling of traffic behavior in growing metropolitan areas, *Transportation Research Procedia* 10 (2015) 306–315.
- [17] K. Bernhardt, Agent-based modeling in transportation, *Artificial Intelligence in Transportation* 72 (2007).
- [18] R. Beheshti, G. Sukthankar, A normative agent-based model for predicting smoking cessation trends, 13th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2014 1 (2014) 557–564.
- [19] R. Axelrod, *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*, Princeton University Press, 1997. URL: <http://www.jstor.org/stable/j.ctt7s951>.
- [20] Y. Li, D. Zhang, J. A. Pagán, Social norms and the consumption of fruits and vegetables across new york city neighborhoods, *Journal of Urban Health* 93 (2016) 244–255.
- [21] C. D. Hollander, A. S. Wu, The current state of normative agent-based systems, *Journal of Artificial Societies and Social Simulation* 14 (2011) 6.
- [22] R. Boyd, P. J. Richerson, *Culture and the evolutionary process*, University of Chicago press, 1988.
- [23] R. Boyd, P. J. Richerson, *The origin and evolution of cultures*, Oxford University Press, 2005.
- [24] M. Nakamaru, S. A. Levin, Spread of two linked social norms on complex interaction networks, *Journal of theoretical biology* 230 (2004) 57–64.
- [25] C. von Scheve, D. Moldt, J. Fix, R. von Lüde, My agents love to conform: Emotions, norms, and social control in natural and artificial societies., in: *NORMAS*, Citeseer, 2005, pp. 73–84.
- [26] A. Staller, P. Petta, et al., Introducing emotions into the computational study of social norms: A first evaluation, *Journal of artificial societies and social simulation* 4 (2001) U27–U60.
- [27] R. Axelrod, An evolutionary approach to norms, *American political science review* 80 (1986) 1095–1111.
- [28] M. A. Mahmoud, M. S. Ahmad, M. Z. Mohd Yusoff, A. Mustapha, A review of norms and normative multiagent systems, *The Scientific World Journal* 2014 (2014).
- [29] R. Aldred, J. Croft, A. Goodman, Impacts of an active travel intervention with a cycling focus in a suburban context: One-year findings from an evaluation of London’s in-progress mini-Hollands programme, *Transp Res Part A Policy Pract* 123 (2019) 147–169. doi:10.1016/j.tra.2018.05.018.
- [30] R. Lovelace, D. Ballas, M. Watson, A spatial microsimulation approach for the analysis of commuter patterns: From individual to regional levels, *J Transp Geogr* 34 (2014) 282–296. doi:10.1016/j.jtrangeo.2013.07.008.
- [31] R. Lovelace, D. Ballas, ‘Truncate, replicate, sample’: A method for creating integer weights for spatial microsimulation, *Comput Environ Urban Syst* 41 (2013) 1–11. doi:10.1016/j.compenurbysys.2013.03.004.
- [32] University of Essex, Institute for Social and Economic Research, NatCen Social Research,

- Kantar Public, Understanding Society: Waves 1–10, 2009–2019 and Harmonised BHPS: Waves 1–18, 1991–2009, 2020. doi:10.5255/UKDA-SN-6614-14.
- [33] Office for National Statistics, National Records of Scotland, Northern Ireland Statistics and Research Agency, 2011 Census aggregate data, 2017. doi:10.5257/census/aggregate-2011-2.
- [34] Department for Transport, NTS0608: Bicycle ownership by age: England, 2020.
- [35] Office for National Statistics, Location of usual residence and place of work by method of travel to work, 2014.
- [36] N. D. Matsakis, F. S. Klock, The rust language, in: HILT '14: Proceedings of the 2014 ACM SIGAda Annual Conference on High Integrity Language Technology, Association for Computing Machinery, New York, NY, United States, 2014, pp. 103–104. doi:10.1145/2663171.2663188.
- [37] R. Greener, D. Lewis, J. Reades, S. Miles, S. Cummins, Software and results for: Incorporating social norms into a configurable agent-based model of commuting behaviour, 2022. doi:10.5281/zenodo.5993329.
- [38] R. Aldred, 'On the outside': Constructing cycling citizenship, *Soc Cult Geogr* 11 (2010) 35–52. doi:10.1080/14649360903414593.
- [39] Met Office, Met Office Hadley Centre Observation Data, 2021.
- [40] D. Balkmar, Towards an Intersectional Approach to Men, Masculinities and (Un)sustainable Mobility: The Case of Cycling and Modal Conflicts, in: C. L. Scholten, T. Joellsson (Eds.), *Integrating Gender into Transport Planning : From One to Many Tracks*, Palgrave Macmillan, 2019. doi:10.1007/978-3-030-05042-9_9.
- [41] M. Falcous, Why We Ride: Road Cyclists, Meaning, and Lifestyles, *J Sport Soc Issues* 41 (2017) 239–255. doi:10.1177/0193723517696968.
- [42] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* 393 (1998) 440–442. doi:10.1038/30918.
- [43] A.-L. Barabási, R. Albert, Emergence of Scaling in Random Networks, *Science* 286 (1999) 509–512. doi:10.1126/science.286.5439.509.
- [44] OED Online, norm, n.1, <https://www.oed.com/view/Entry/81855#eid302885565>, 2022.
- [45] G. Lyons, K. Chatterjee, A human perspective on the daily commute: Costs, benefits and trade-offs, *Transp Rev* 28 (2008) 181–198. doi:10.1080/01441640701559484.
- [46] M. Bratman, *Intention, plans, and practical reason*, 1987.
- [47] A. Gelman, J. Carlin, H. Stern, D. Dunson, A. Vehtari, D. Rubin, *Bayesian Data Analysis*, third ed., CRC Press, Boca Raton, FL, USA, 2014.
- [48] J. Bezanson, A. Edelman, S. Karpinski, V. B. Shah, Julia: A Fresh Approach to Numerical Computing, *SIAM Review* 59 (2017) 65–98. doi:10.1137/141000671.
- [49] H. Ge, K. Xu, Z. Ghahramani, Turing: A Language for Flexible Probabilistic Inference, in: A. Storkey, F. Perez-Cruz (Eds.), *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics*, volume 84 of *Proceedings of Machine Learning Research*, PMLR, 2018, pp. 1682–1690.
- [50] M. Betancourt, A Conceptual Introduction to Hamiltonian Monte Carlo, arXiv:1701.02434 [stat] (2018). arXiv:1701.02434.
- [51] T. K. Papp, D. Aluthge, JackRab, D. Widmann, J. TagBot, M. Piibelet, P. Monticone, St-, Tppapp/DynamicHMC.jl: V3.1.1, Zenodo, 2021. doi:10.5281/ZENODO.5554755.

- [52] R. McElreath, *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*, Texts in Statistical Science, second ed., CRC Press, Boca Raton, FL, USA, 2020.
- [53] J. Kruschke, a. O. M. C. Safari, *Doing Bayesian Data Analysis*, 2nd Edition, 2014.
- [54] D. Ogilvie, A. Bauman, L. Foley, C. Guell, D. Humphreys, J. Panter, Making sense of the evidence in population health intervention research: Building a dry stone wall, *BMJ Glob Health* 5 (2020) e004017. doi:10.1136/bmjgh-2020-004017.

A. Class diagram of MOTIVATE

Arrows between classes are omitted for space. See method signatures for the relationships.

