*

Francisco Rowe [†1] and xxx[2]

[1]Geographic Data Science Lab, Department of Geography and Planning, University
of Liverpool, Liverpool, United Kingdom
[2]xxx

Model Specification

This note seeks to document the model specification we would ideally like to estimate and its key components and reference terms.

$$p_{ijt} = \alpha_{ij} + \gamma_t + FB_{ijt} + \varepsilon_{ijt}$$
$$\alpha_{ij} = \alpha_i + \alpha_j + \beta_1 d_{ij} + \beta_2 c_{ij} + u_{ij}$$
$$\alpha_i = \delta_i + \delta_{i1} pop_i + F_i + u_i$$
$$\alpha_j = \delta_j + \delta_{j1} pop_j + F_j + u_j \tag{1}$$
$$\gamma_t = \delta_t + \delta_{t1} day + \delta_{t2} ti + \delta_{t3} w + u_t$$
$$FB_{ijt} = fb_0 + fb_{ijt} z_{ijt} + fb_{ijt} au_{ijt} + u_{ijt}$$

p_{ijt}: proportion of people moving from tile i to tile j at time t;

α_{ij}: component relating to tiles of origin i and destionation j;

γ_t: component relating to temporal variation;

FB_{ijt}: component relating to variation in Facebook data quality;

α_i: factors relating to tiles of origin i;

α_j: factors relating to tiles of origin j;

β_1: coefficient capturing the deterring effect of distance d_{ij};

β_2: coefficient capturing the variation in mobility intensity across tiles in different functional areas c_{ij};

δ_i: average level of mobility related to tiles of origin i;

δ_{i1}: coefficient capturing the relationship between mobility intensity and population at origin tile pop_i;

F_i: coefficient capturing the relationship between mobility intensity and additional attributes at origin i eg. unemployment, housing affordability, etc.

δ_j: average level of mobility related to tiles of destination j;

δ_{j1}: coefficient capturing the relationship between mobility intensity and population at destination tile pop_j;

F_j: coefficient capturing the relationship between mobility intensity and additional attributes at destination j eg. unemployment, housing affordability, etc.

δ_t: average level of mobility related to time;

δ_{t1}: coefficient capturing the daily day variability in mobility;

δ_{t2}: coefficient capturing the variability in mobility according to the 8 hour time interval ti in which the Facebook data are organised; that is, 0am, 8am and 4pm;

$\delta_{t3} w$: coefficient capturing the variability in mobility during week days and weekends w;

ε and u correspond to the error terms;

Estimation

The plan is to estimate Equation 1 in a fractional regression framework (see Papke and Wooldridge, 1996). In this framework errors are assumed to follow a logistic distribution and quasi maximum likelihood procedure is used to estimate the set of regression coefficients. The variable identifying pair combinations of functional areas for tiles of origin and destination (i.e. c_{ij}) can enter the model as a dummy variable or as a random effect in a multilevel framework. Time can also be modelled in different ways (i.e. including autoregresive terms of p_{ijt}, autoregresive terms of errors and splines) and there may be different components which may need to be taken into consideration (i.e. seasonability and autocorrelation)

References

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401 (k) plan participation rates. *Journal of applied econometrics*, 11(6):619–632.