PROTECT
A COVID-19 National Core Study

Understanding community level influences on COVID-19 prevalence in England

New insights from comparison over time and space

medRxiv
THE PREPRINT SERVER FOR HEALTH SCIENCES
CSH Cold Spring Harbor Laboratory BMJ Yale

Chaitanya Joshi[1,a], Arif Ali[1,+], Thomas ÓConnor[1,+], Li Chen[1,b], and Kaveh Jahanshahi[1,c]

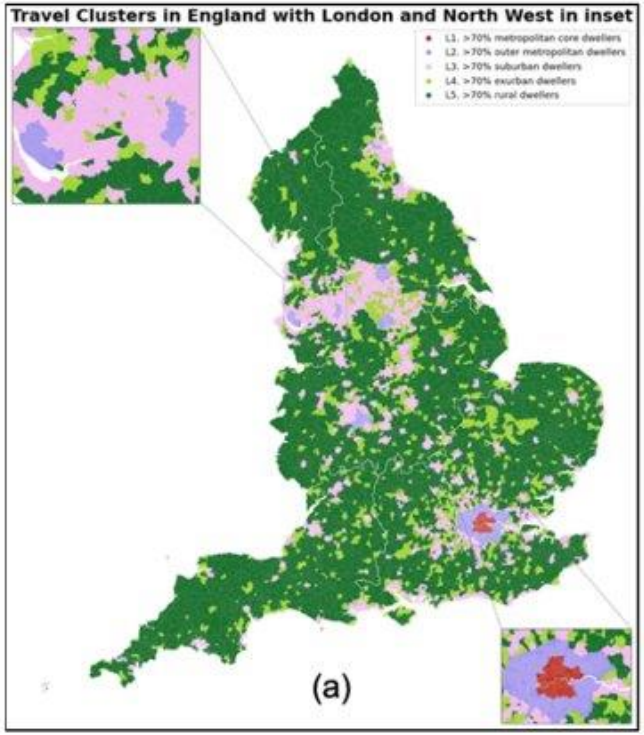
[1]Data Science Campus, Office for National Statistics, UK

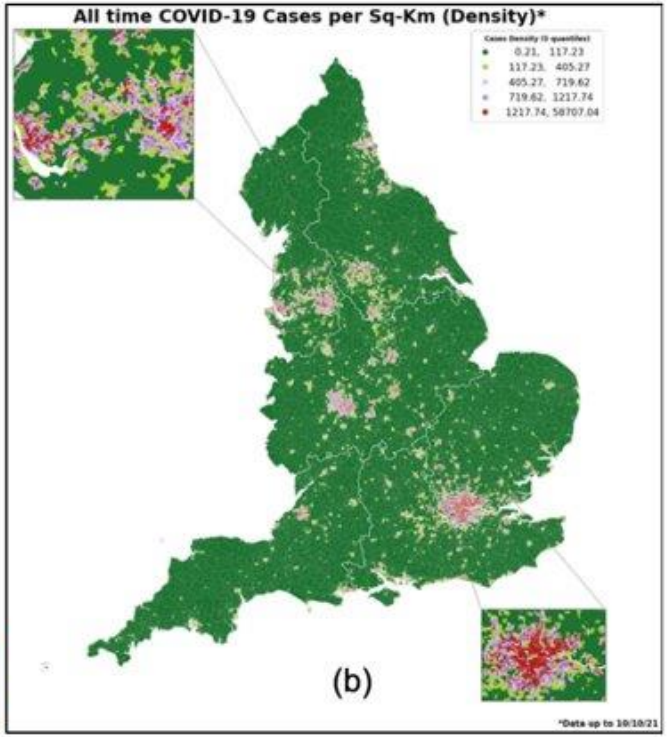
Census 2011, 2019 Mid-Year population estimates, Inter-Department Business Register (IDBR) 2019 dataset on workplace and industrial sectors, the national travel survey (NTS) 2002 to 2015 to segment travel patterns and clusters, mobile phone data to extract real-time mobility indicators, uptake of first and second doses of COVID-19 vaccination and test and trace data for gathering dynamic information on COVID-19 incidents.

- *Part of DSC(ONS) support in response to COVID-19*
- *Developing understanding of risks at community level*
- *Workplace risks after controlling for residential characteristics, NPIs, travel patterns, vaccination, etc.*
- *Early warning system*
- *A new framework for augmenting data from different sources*

PROTECT Researcher Symposium 5-6 May 2022

PROTECT
A COVID-19 National Core Study



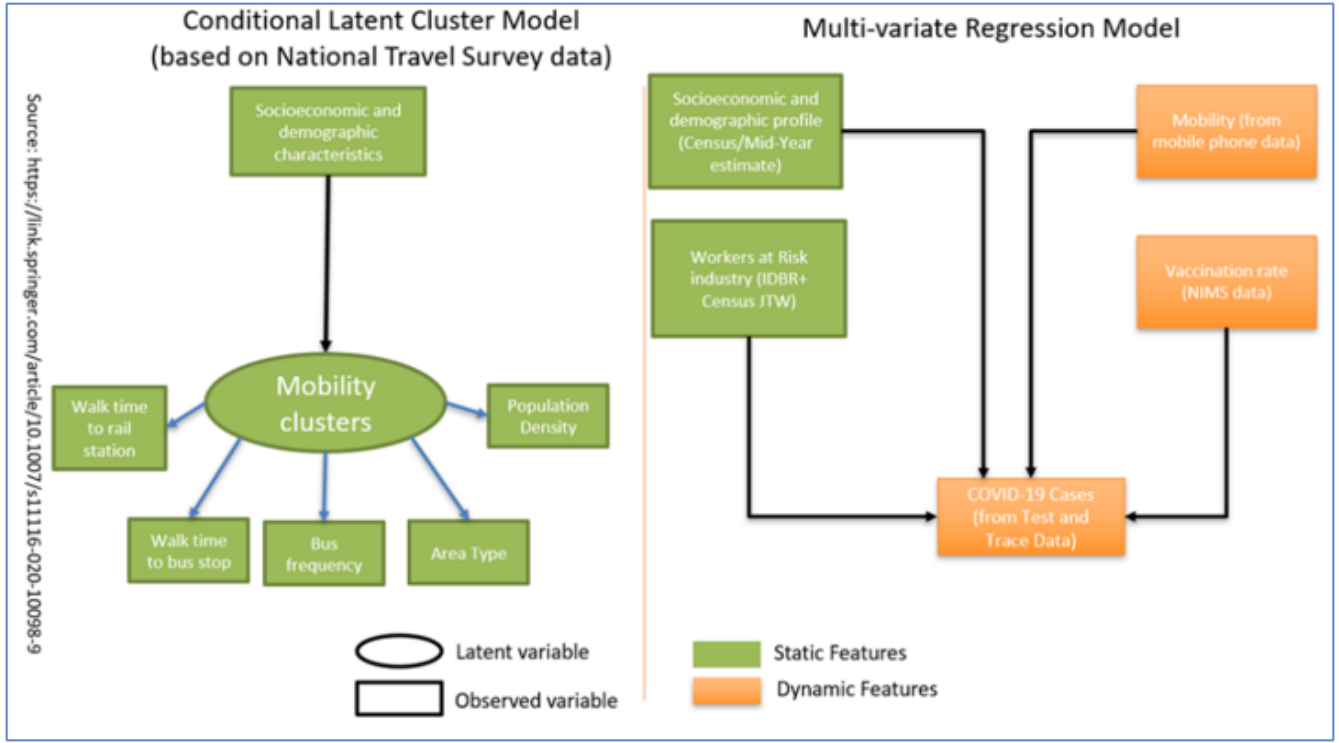


Interactions between socioeconomic profile, mobility patterns, and land use features of neighbourhood and settlements- Travel clusters

Multi-group approach- separate model for different travel clusters

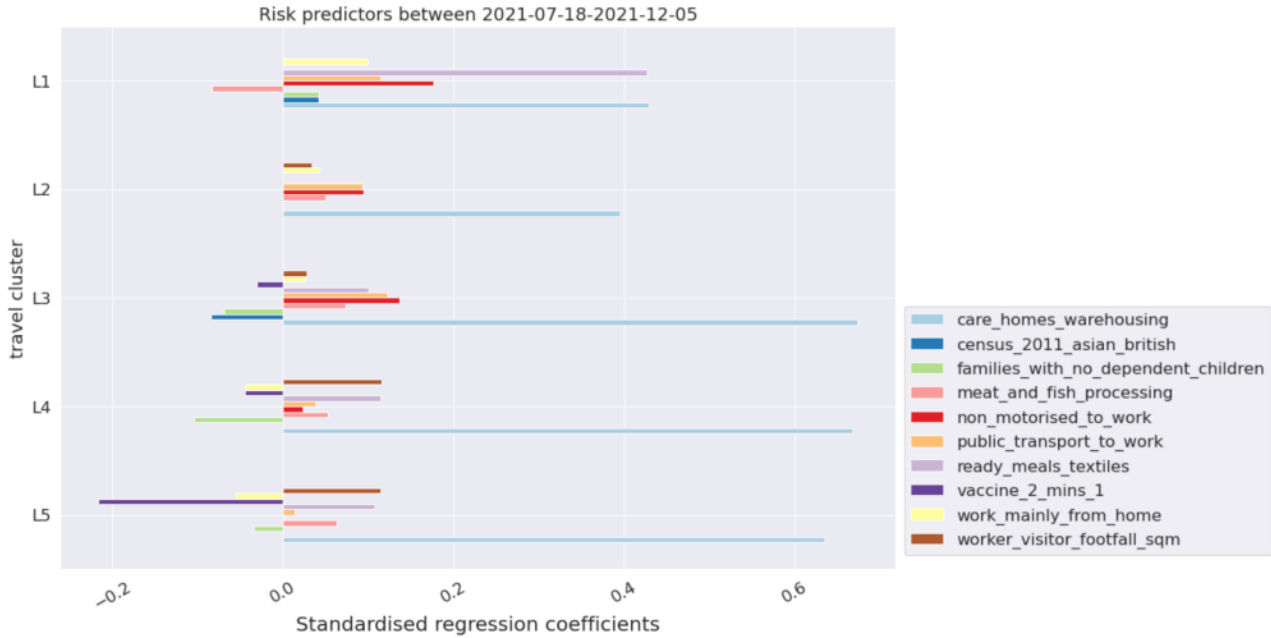
Exploratory Factor Analysis to account for spatial interactions

Multivariate regression to model risk influences


Conditional Latent Cluster Model
(based on National Travel Survey data)
Source: https://link.springer.com/article/10.1007/s11116-020-10098-9
Socioeconomic and demographic characteristics
Mobility clusters
Walk time to rail station
Population Density
Walk time to bus stop
Bus frequency
Area Type
Latent variable
Observed variable
Multi-variate Regression Model
Socioeconomic and demographic profile (Census/Mid-Year estimate)
Mobility (from mobile phone data)
Workers at Risk industry (IDBR+ Census JTW)
Vaccination rate (NIMS data)
COVID-19 Cases (from Test and Trace Data)
Static Features
Dynamic Features

PROTECT Researcher Symposium 5-6 May 2022

PROTECT
A COVID-19 National Core Study



- Analysis can be tailored to evaluate safety measures and regulations for high-risk workplaces.
- Continuous evaluation of community level risk can identify new threats and risk patterns at an early stage when there is a better chance to respond.

- *Areas with a larger proportion of residents working in care homes and warehouses are prone to higher risk of infection.*
- *The critical role of geographical variations in influences on COVID-19 Incidence – Bigger proportion of small families and fewer density of children are prone to lower risk of infection in medium and smaller urban and rural areas.*
- *Use of public transport has been identified as one of the main risk factors in smaller and bigger urban areas alike.*
- *Work towards explainable ML-open source code can be extended to more opaque models.*