Big Data and Health:
From New Zealand to the UK

Dr Fran Darlington-Pollock

Geographic Data Science Lab | University of Liverpool

Using big data to tackle inequalities in society: University of Auckland, June 2018

Geographic Data Science Lab

UNIVERSITY OF LIVERPOOL

"The era of big data is upon us"

Davies and Green, 2018

- Explosion of (fundamentally different) new data: large datasets, richer temporal and spatial resolution (Cambridge Public Policy SRI, 2016)

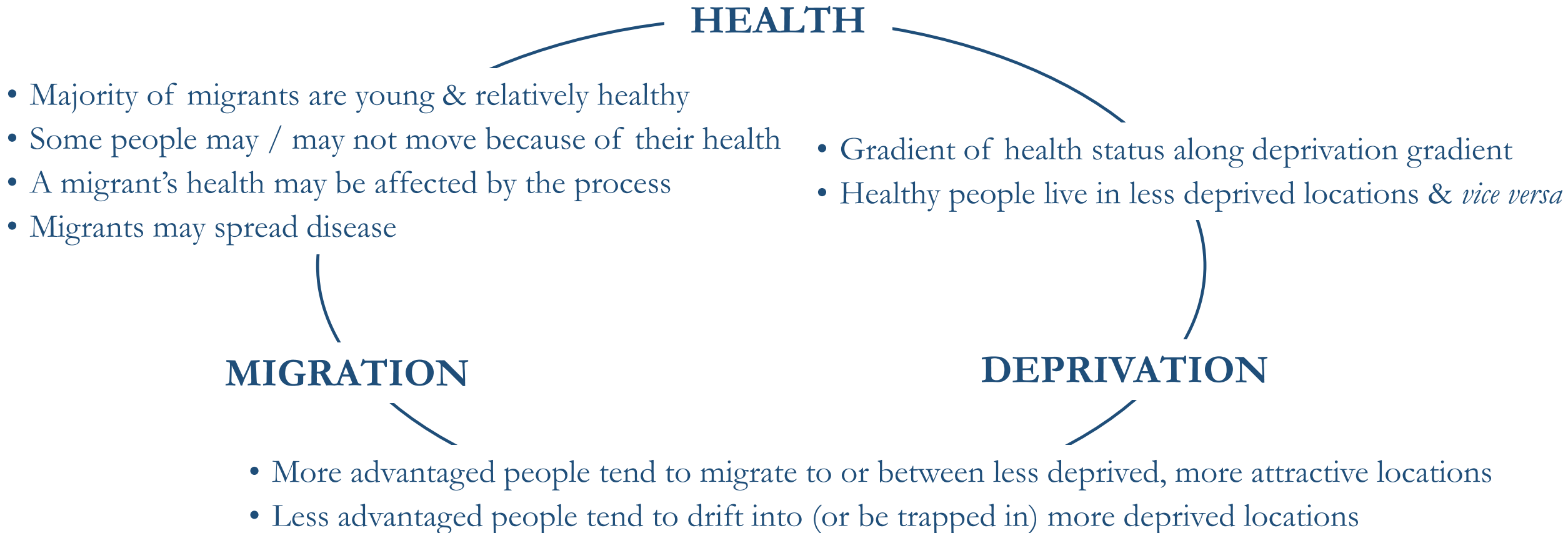- Implications for how we **measure** and **understand social phenomena**

Examples:

- Administrative data → linking **routine national health databases**;

- Social media data → georeferenced tweets;

- Consumer data → consumption patterns; **business and retail locations**

Example 1: Administrative Data

Risky moves and cardiovascular disease in New Zealand

Team includes: Nichola Shackleton, Dan Exeter (University of Auckland), Paul Norman (University of Leeds)

Research context (II)

HEALTH

- Majority of migrants are young & relatively healthy
- Some people may / may not move because of their health
- A migrant's health may be affected by the process
- Migrants may spread disease

- Gradient of health status along deprivation gradient
- Healthy people live in less deprived locations & *vice versa*

MIGRATION

DEPRIVATION

- More advantaged people tend to migrate to or between less deprived, more attractive locations
- Less advantaged people tend to drift into (or be trapped in) more deprived locations

Research Context (I)

- Complex health-migration inter-relationships;

- Importance of ***deprivation mobility/change*** for migration-health relationship;

- Residential mobility an important **determinant** of CVD in Auckland (Exeter et al., 2015);

- Cardiovascular disease (CVD) one of the leading causes of death globally, marked variations between ethnic groups;

- Differences in migration patterns between ethnic groups in New Zealand

- **Relationship with ethnic inequalities in CVD?**

(BIG!) Data

- Enrolment with Primary Health Organisation
- Pharmaceutical Dispensing Claims
- Hospital Discharges
- Mortality

→

Patient records anonymously linked with *National Health Index (NHI)* number

n = 94-97% population

→

VIEW Dataset

Outcomes (e.g.)
- Lipid testing
- Diabetes
- Hospitalisations
- Medication dispensing

Demographics
- Age
- Gender
- Ethnicity
- NZDep06

Geographies
- Meshblock
- Area Unit
- Electorate
- District Health Board

→

CVD and Migration Dataset

Eligible if…
- Aged 30-84
- Complete socio-demographic / geographic information
- No prior history of CVD

Study Period
- 36 calendar quarters 01.01.2006-30.06.2014

Methods

Cardiovascular Disease – Residential Mobility – Deprivation

ASSOCIATIONS

- Binary logistic regression– total population & stratified by ethnic group
- Compare risk of CVD for moves with that for stayers
- Ethnic differences?
- Differences by nature of the move?

EFFECTS

- Cox proportional regression (survival analysis) – total population & stratified by ethnic group
- Compare risk of CVD for movers **who move before first CVD event** with stayers
- Ethnic differences?
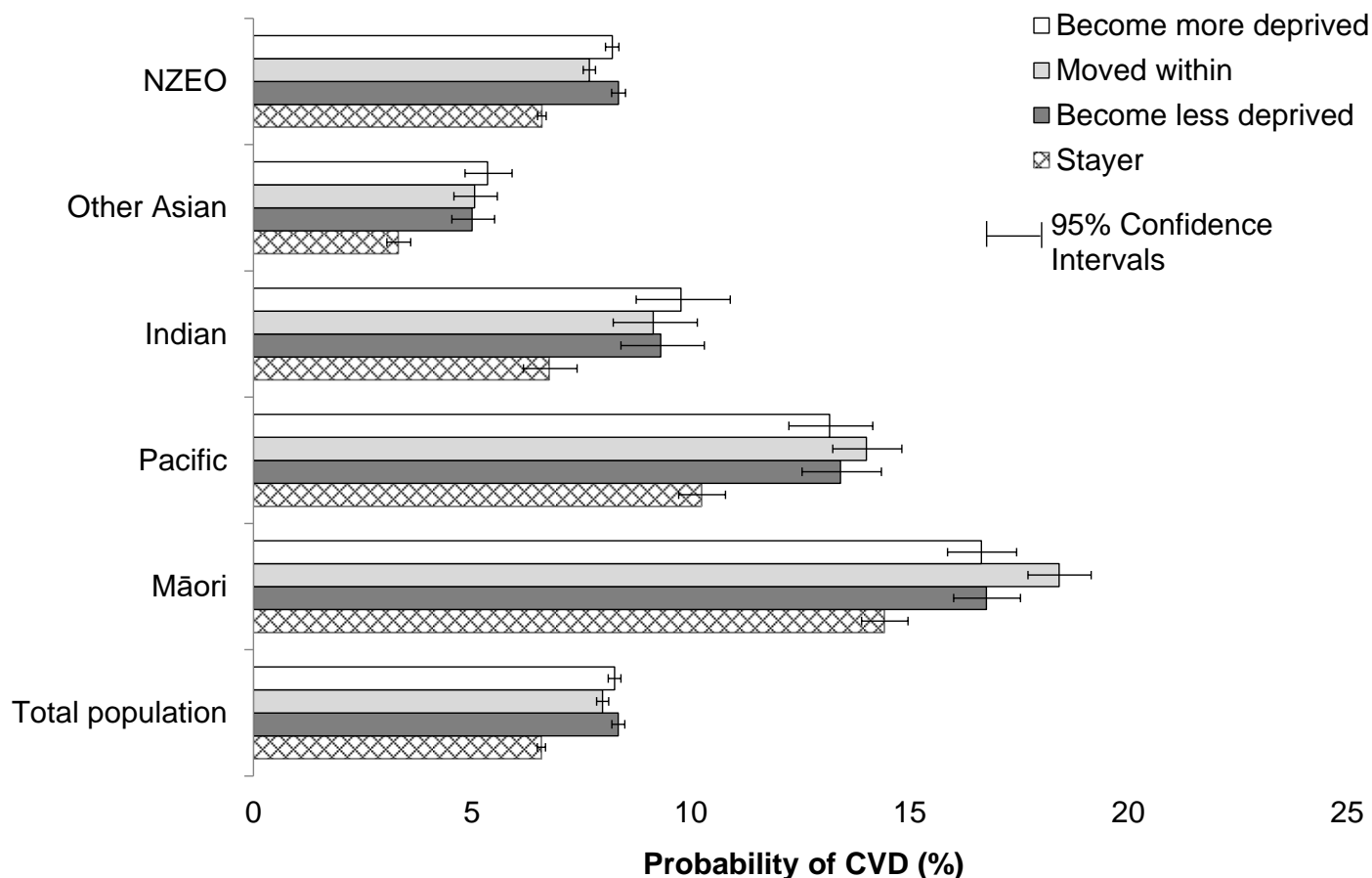- Differences by nature of the move?

TRAJECTORIES

- Trajectory analysis
- Compare CVD risk for **movers** according to their **deprivation trajectory**
- Only movers **who move before first CVD event**

Associations (I)
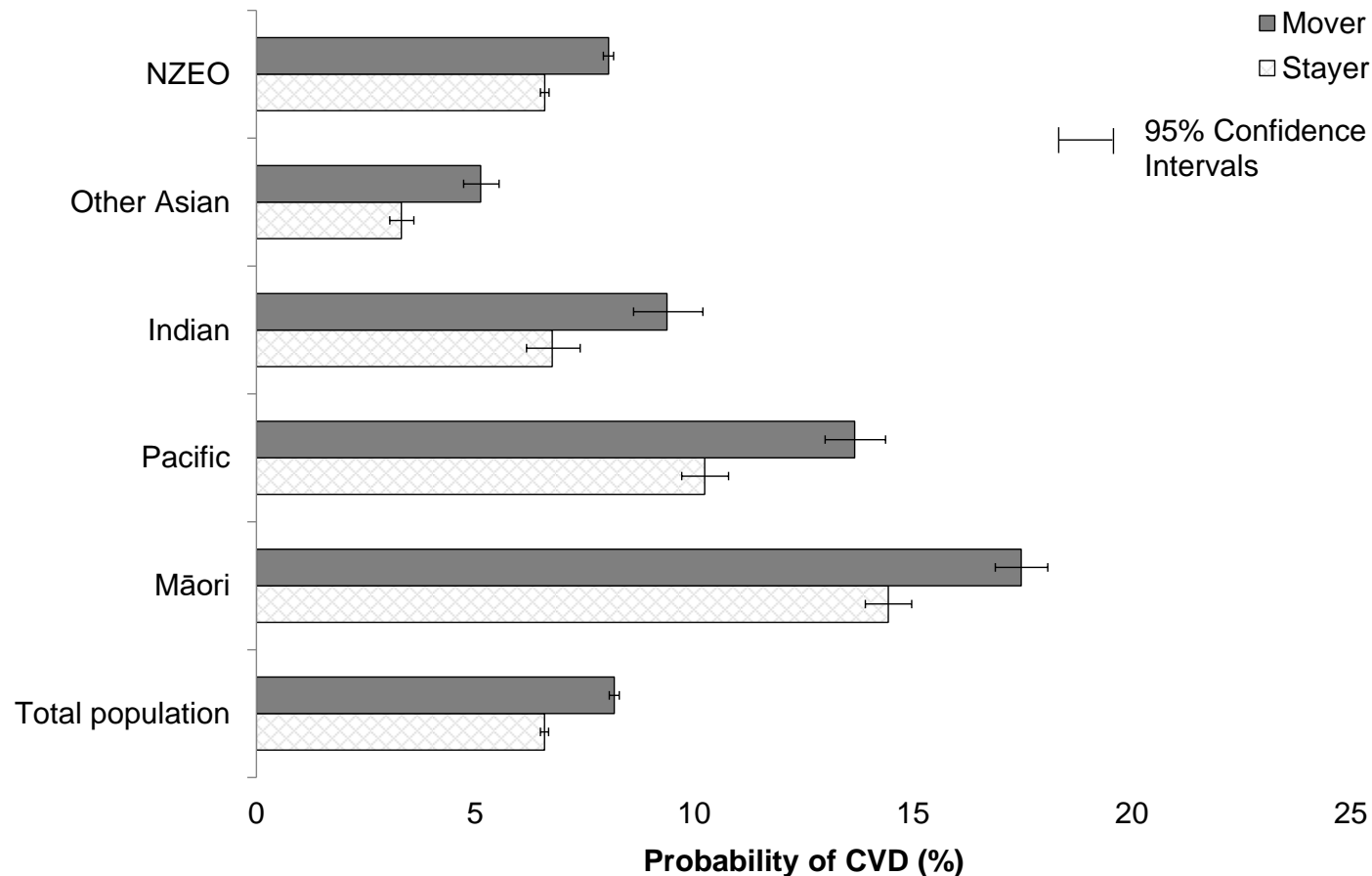
- Movers **significantly higher probability** of CVD compared to stayers **for all ethnic groups**

- Variation **between ethnic groups**

- **Māori and Pacific groups** higher probability of CVD than total population, also true for **Indian movers**

- Does the nature of the move matter?

Associations (II)

Legend:
- □ Become more deprived
- ☐ Moved within
- ▨ Become less deprived
- ⊠ Stayer
- ⊢⊣ 95% Confidence Intervals

Categories (top to bottom): NZEO, Other Asian, Indian, Pacific, Māori, Total population

X-axis: Probability of CVD (%) — 0, 5, 10, 15, 20, 25

- Moving within **same deprivation quintile** has different implications for different ethnic groups

- Māori and Pacific groups live in most deprived areas: moves within the same deprivation quintile = moves within the most deprived quintile

- Moving to a more deprived area not always associated with higher risk of CVD

- Is it the move, or is it the person?

Effects: Hazard ratios for Mobile groups relative to Stayers → Risk of CVD

	Total	Māori	Pacific	Indian	Other Asian	NZEO
Mover	0.64 (0.63–0.64)	0.59 (0.58–0.61)	0.66 (0.63–0.69)	0.65 (0.61–0.70)	0.63 (0.60–0.68)	0.64 (0.63–0.65)
Becomes less deprived	0.64 (0.63–0.65)	0.59 (0.56–0.63)	0.68 (0.53–0.73)	0.64 (0.59–0.72)	0.64 (0.58–0.71)	0.64 (0.53–0.66)
Churns (moves within)	0.63 (0.62–0.64)	0.60 (0.58–0.63)	0.64 (0.60–0.68)	0.67 (0.60–0.75)	0.63 (0.57–0.69)	0.64 (0.63–0.65)
Becomes more deprived	0.63 (0.62–0.64)	0.58 (0.55–0.61)	0.69 (0.64–0.75)	0.63 (0.56–0.71)	0.65 (0.59–0.72)	0.63 (0.62–0.65)

Source: Darlington-Pollock et al., 2017

Effects: Hazard ratios for Mobile groups relative to Stayers → Risk of CVD

	Total	Māori	Pacific	Indian	Other Asian	NZEO
Mover	0.64 (0.63–0.64)	0.59 (0.58–0.61)	0.66 (0.63–0.69)	0.65 (0.61–0.70)	0.63 (0.60–0.68)	0.64 (0.63–0.65)
Becomes less deprived	0.64 (0.63–0.65)	0.59 (0.56–0.63)	0.68 (0.53–0.73)	0.64 (0.59–0.72)	0.64 (0.58–0.71)	0.64 (0.53–0.66)
Churns (moves within)	0.63 (0.62–0.64)	0.60 (0.58–0.63)	0.64 (0.60–0.68)	0.67 (0.60–0.75)	0.63 (0.57–0.69)	0.64 (0.63–0.65)
Becomes more deprived	0.63 (0.62–0.64)	0.58 (0.55–0.61)	0.69 (0.64–0.75)	0.63 (0.56–0.71)	0.65 (0.59–0.72)	0.63 (0.62–0.65)

- Movers significantly **lower risk** of CVD than stayers

@F_Darlington | @geodatascience

Source: Darlington-Pollock et al., 2017

Effects: Hazard ratios for Mobile groups relative to Stayers → Risk of CVD

|	Total	Māori	Pacific	Indian	Other Asian	NZEO
Mover	0.64 (0.63–0.64)	0.59 (0.58–0.61)	0.66 (0.63–0.69)	0.65 (0.61–0.70)	0.63 (0.60–0.68)	0.64 (0.63–0.65)
Becomes less deprived	0.64 (0.63–0.65)	0.59 (0.56–0.63)	0.68 (0.53–0.73)	0.64 (0.59–0.72)	0.64 (0.58–0.71)	0.64 (0.53–0.66)
Churns (moves within)	0.63 (0.62–0.64)	0.60 (0.58–0.63)	0.64 (0.60–0.68)	0.67 (0.60–0.75)	0.63 (0.57–0.69)	0.64 (0.63–0.65)
Becomes more deprived	0.63 (0.62–0.64)	0.58 (0.55–0.61)	0.69 (0.64–0.75)	0.63 (0.56–0.71)	0.65 (0.59–0.72)	0.63 (0.62–0.65)

- Some variation between ethnic groups, Māori movers have the lowest risk of CVD relative to their immobile peers
- Similar risks across the other ethnic groups

Effects: Hazard ratios for Mobile groups relative to Stayers → Risk of CVD

	Total	Māori	Pacific	Indian	Other Asian	NZEO
Mover	0.64 (0.63–0.64)	0.59 (0.58–0.61)	0.66 (0.63–0.69)	0.65 (0.61–0.70)	0.63 (0.60–0.68)	0.64 (0.63–0.65)
Becomes less deprived	0.64 (0.63–0.65)	0.59 (0.56–0.63)	0.68 (0.53–0.73)	0.64 (0.59–0.72)	0.64 (0.58–0.71)	0.64 (0.53–0.66)
Churns (moves within)	0.63 (0.62–0.64)	0.60 (0.58–0.63)	0.64 (0.60–0.68)	0.67 (0.60–0.75)	0.63 (0.57–0.69)	0.64 (0.63–0.65)
Becomes more deprived	0.63 (0.62–0.64)	0.58 (0.55–0.61)	0.69 (0.64–0.75)	0.63 (0.56–0.71)	0.65 (0.59–0.72)	0.63 (0.62–0.65)

- Some variation in the size of the HR, but CIs overlap – deprivation change does not differentiate risk of CVD for these mobile groups relative to their immobile peers

Source: Darlington-Pollock et al., 2017

Effects: Hazard ratios for Mobile groups relative to Stayers → Risk of CVD

	Total	Māori	Pacific	Indian	Other Asian	NZEO
Mover	0.64 (0.63–0.64)	0.59 (0.58–0.61)	0.66 (0.63–0.69)	0.65 (0.61–0.70)	0.63 (0.60–0.68)	0.64 (0.63–0.65)
Becomes less deprived	0.64 (0.63–0.65)	0.59 (0.56–0.63)	0.68 (0.53–0.73)	0.64 (0.59–0.72)	0.64 (0.58–0.71)	0.64 (0.53–0.66)
Churns (moves within)	0.63 (0.62–0.64)	0.60 (0.58–0.63)	0.64 (0.60–0.68)	0.67 (0.60–0.75)	0.63 (0.57–0.69)	0.64 (0.63–0.65)
Becomes more deprived	0.63 (0.62–0.64)	0.58 (0.55–0.61)	0.69 (0.64–0.75)	0.63 (0.56–0.71)	0.65 (0.59–0.72)	0.63 (0.62–0.65)

- Variation of a similar magnitude for the different ethnic groups

Source: Darlington-Pollock et al., 2017
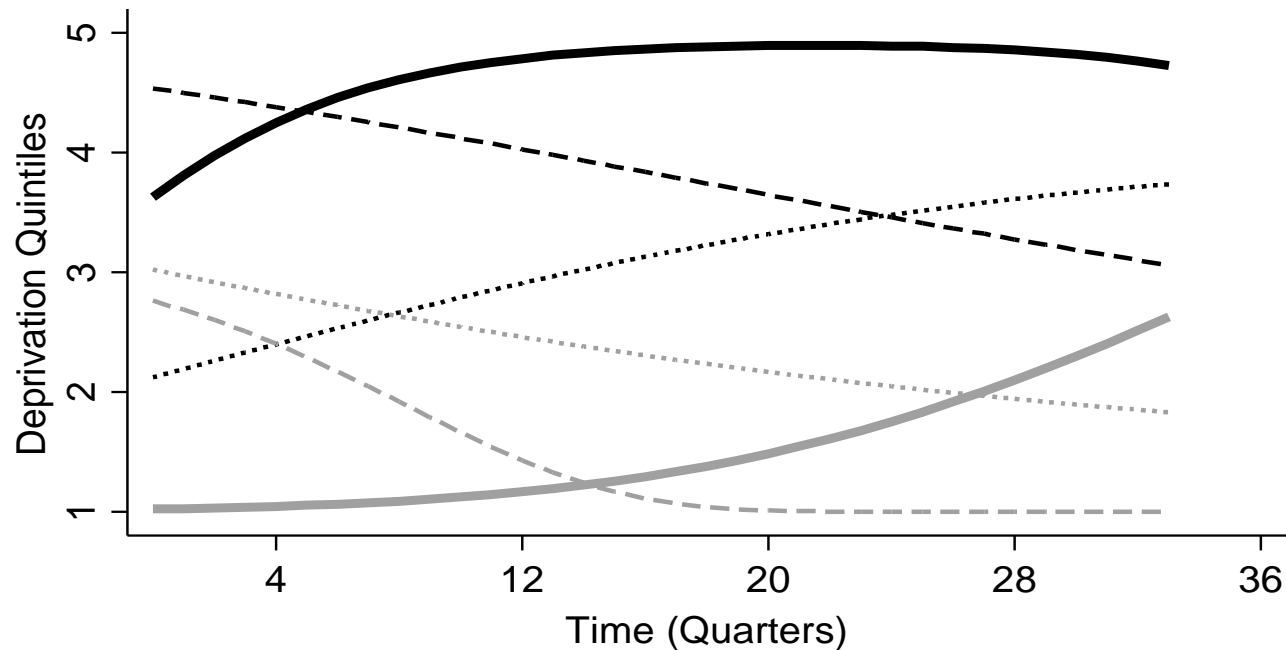
Trajectories (I)

- Incorporates repeated measures of deprivation, rather than simplify deprivation trajectories to first and last recorded observation

- Computationally intensive, difficult to implement in large datasets

- Movers have **lower risk** of CVD than stayers

- Deprivation characteristics of a move have larger impact on relative risk of CVD for younger movers than older movers

@F_Darlington | @geodatascience

Discussion: Strengths and Conclusions

- Invaluable longitudinal dataset covering 94% of NZ's adult population

- Good statistical power

- Temporal detail: sequencing of moves and health event

- Spatial detail: nature of moves

- Sequencing of the move important: movers **more likely to be in good health** than stayers

- Limitations of traditional measures of deprivation change: mis-classification of movers when define by difference between first and last recorded observation…

- Scope of data: rich.. but focussed

Example 2: Retail Data (and more!)

Access to Health Assets and Hazards (AHAH)

Dr Mark Green, University of Liverpool

Context

- Better understand how features of the local environment contribute to health inequalities

- Context matters: but difficulties measuring features of it
  - Requires heavy data manipulation
  - Lack of national level data
  - Accessibility of existing data

- Support policy development at small area (neighbourhood) level that can improve health

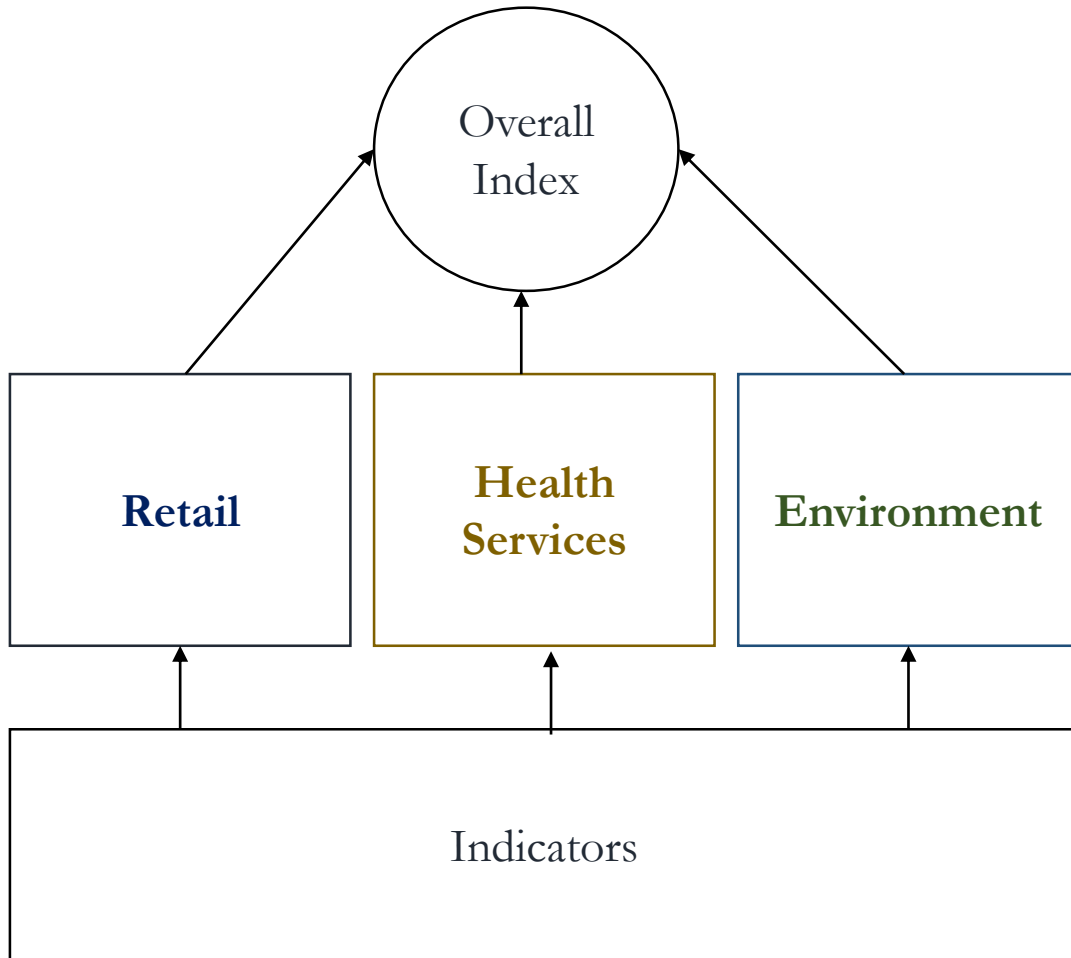- Open access, interactive outputs for policymakers, academics, public health professionals, etc…

Access to Healthy Assets and Hazards
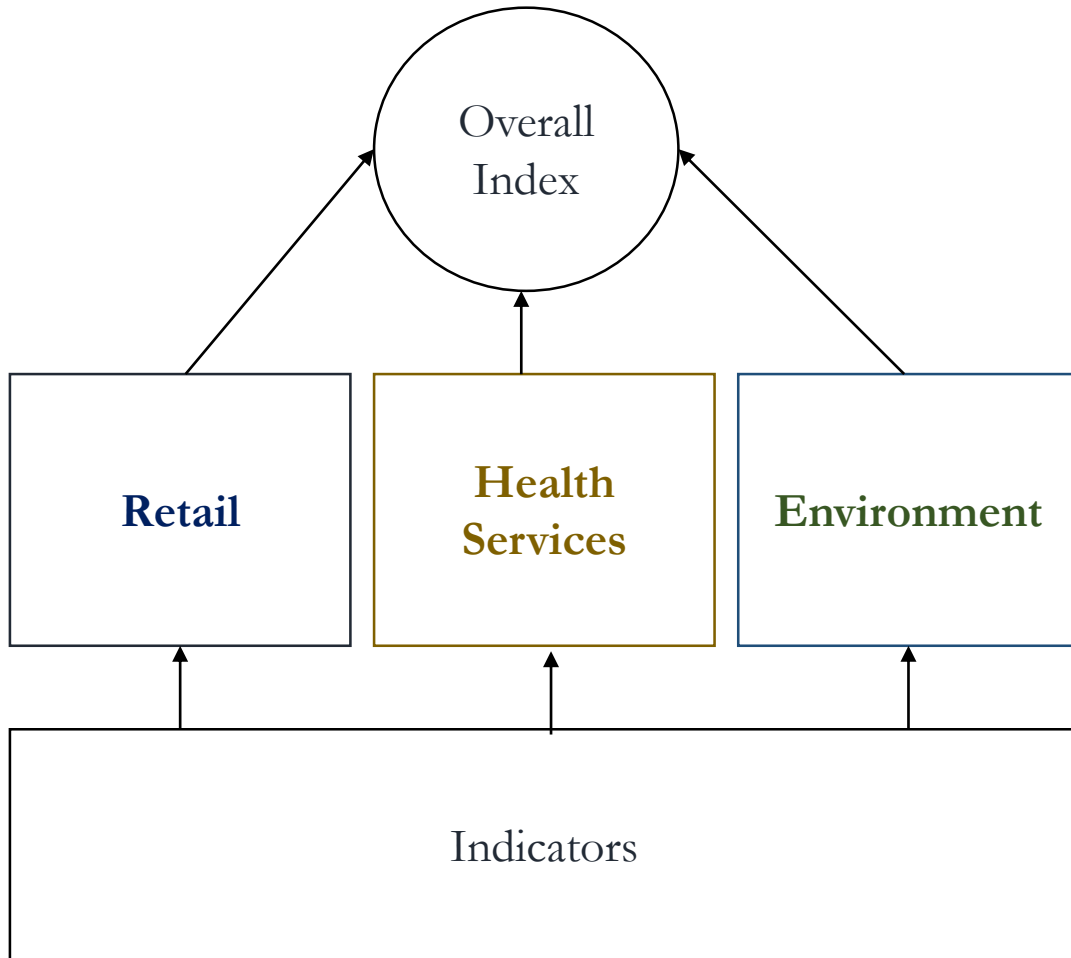
Median access of a postcode to its nearest…

- **Fast food outlet**
- **Gambling outlet**
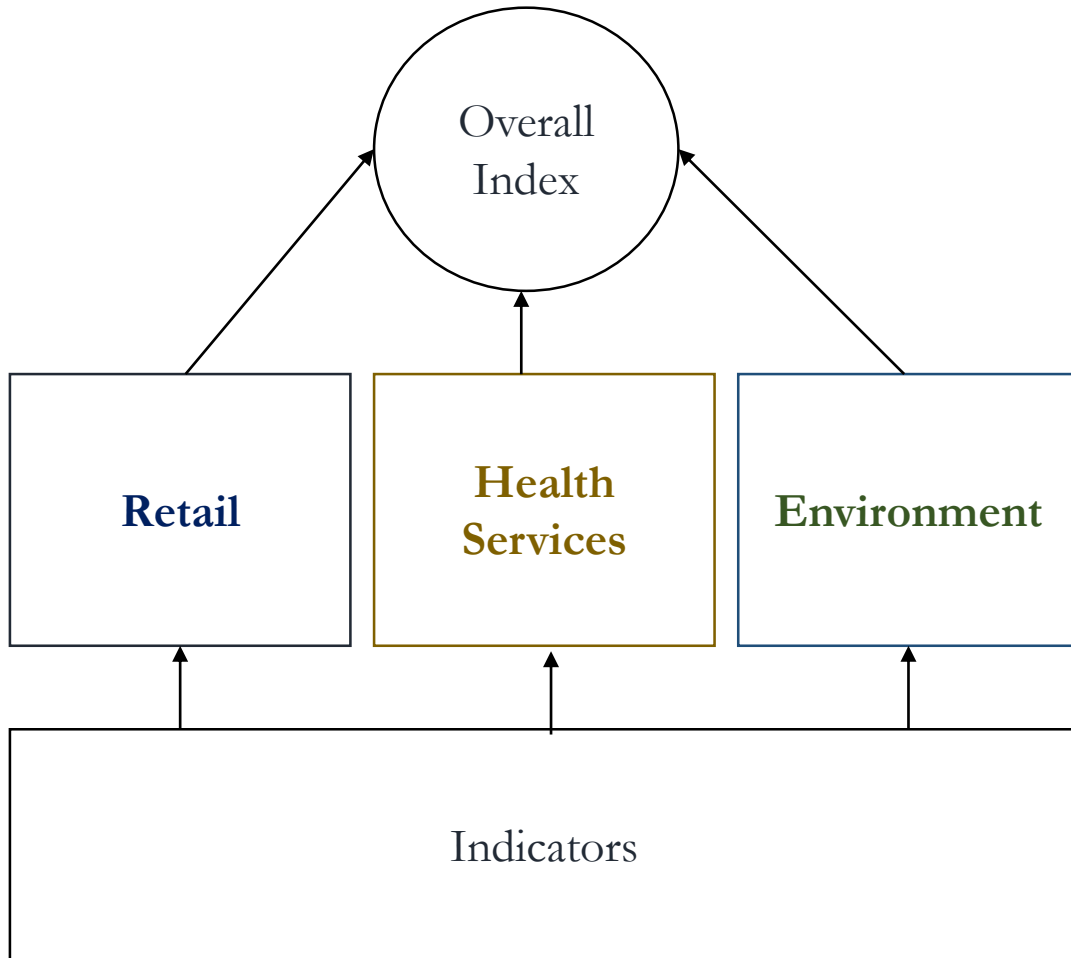- **Pubs, bars, nightclubs**
- **Off-licenses**
- **Tobacconists**

Access to Healthy Assets and Hazards

Median access of a postcode to its nearest…

- **GP**
- **Hospital with A&E**
- **Pharmacy**
- **Dentist**
- **Leisure Service**

@F_Darlington | @geodatascience
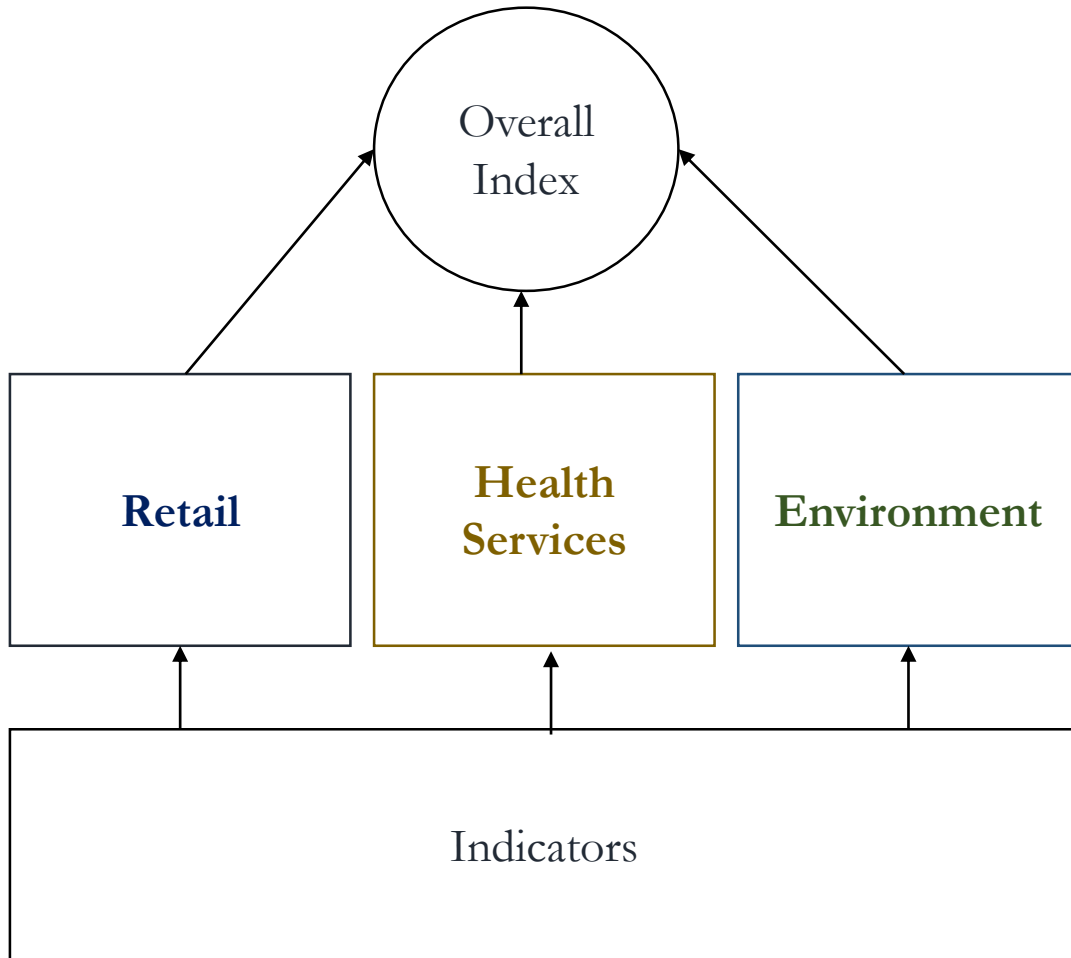
Access to Healthy Assets and Hazards

Median level of pollutants:

- **NO_2**
- **SO_2**
- **PM_{10}**

And:

- **Proportion of accessible green space within $900m^2$ buffer of a postcode**

Access to Healthy Assets and Hazards

Combine each domain into an overall index (equal weightings)

http://maps.cdrc.ac.uk

Thank you!

VIEW data provided by **Analytical Services** at the **New Zealand Ministry of Health**, Encryption of unique identifiers by www.enigma.co.nz

The VIEW programme thanks the **Health Research Council** of New Zealand for funding

Special thanks to **Dr Mark Green** and the team at the **Centre for Consumer Data Research**