EICC, Edinburgh

The finer points of postcode mortality modelling

Stephen Richards
26th November 2009

Plan of talk

1. Postcode anatomy
2. Geodemographics
3. Comparing geodemographic profilers
4. Enhancing geodemographic profilers
5. Conclusions

1. Postcode anatomy

1. Hierarchical structure of a UK postcode

1. Postcode modelling options — geographical

Postcode element	Example	Count
Region	EH	121
District	EH11	2,951
Sector	EH11 2	10,156

Source: Own calculations using 1,706,823 postcodes in Mosaic 2008 directory from Experian

1. Number of annuities in typical category

Postcode element	Annuities
Region	3,682
District	150
Sector	44

- Postal district and sector are *micro-regions*
- Micro-regions are swamped by random variation
- Micro-regions therefore unsuitable for own-portfolio modelling

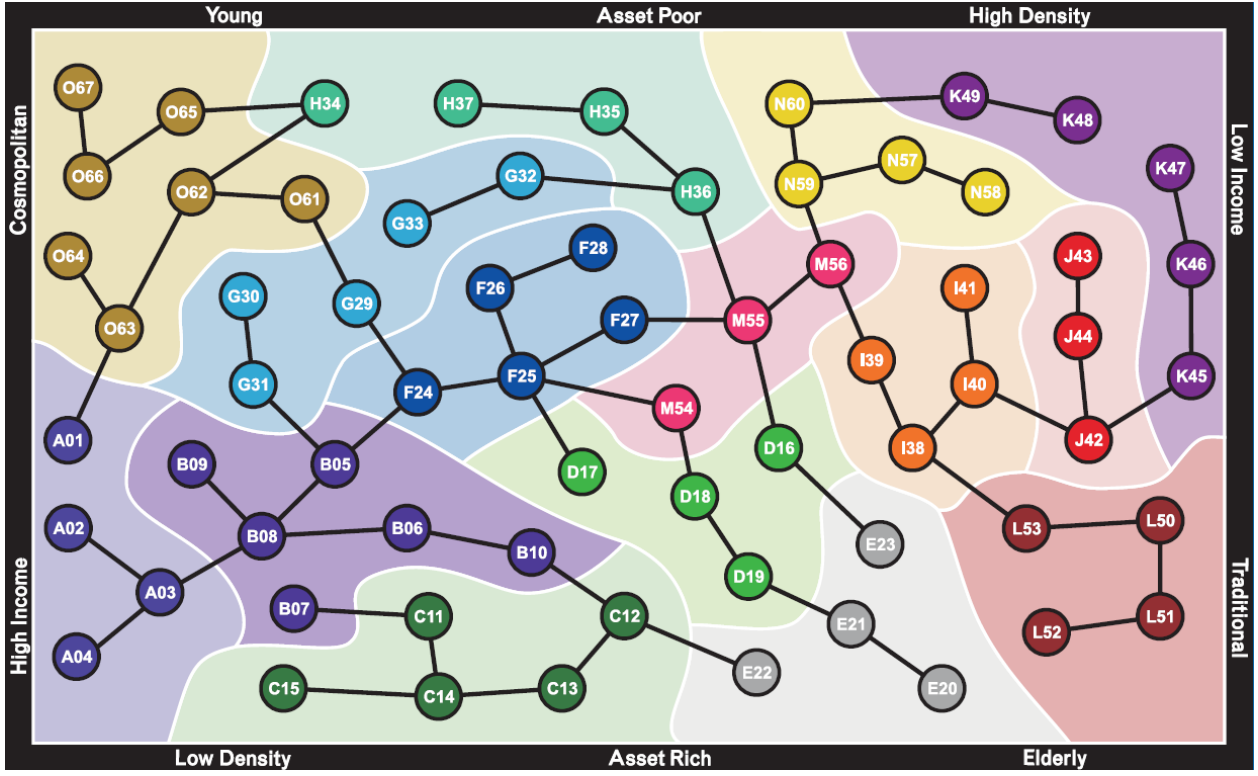
Source: Own calculations using large annuity portfolio

2. Geodemographics

2. Postcode modelling options — geodemographics

- 1.6 million residential postcodes
- Each maps to a *geodemographic group*

2. Geodemographic example — Mosaic



Source: Experian Ltd

2. Some postcode profilers in U.K.

- Mosaic (Experian)
- Acorn (CACI)
- P^2 (Beacon Dodsworth)
- Health Acorn (CACI)
- plus others such as FSS (Experian), Personicx (Acxiom) and CAMEO (Eurodirect)

2. Postcode modelling options

- Geodemographics proven in independent research:
  - → Richards (2008) used Mosaic with survival models
  - → Madrigal et al (2009) used Acorn with GLMs
- Little doubt remains as to usefulness of geodemographics

3. Comparing geodemographic profilers

3. How to compare profilers?

Profiler	Groups	Types
Mosaic	15	67
Acorn	17	56
P^2	13	40
Health Acorn	4	25

- More groups require more parameters
- Need to balance model fit against number of parameters

Source: Own calculations excluding groups or type codes representing unclassified, unmatched or crown dependencies

3. How to compare profilers?

- An *information criterion* balances fit against number of parameters
- A better model has a lower value
- Example: Akaike's Information Criterion — AIC — defined as:

$$\text{AIC} = -2\ell + 2n$$

where ℓ is the log-likelihood function and n is the number of parameters.

- Other examples include the BIC and GCV statistics

3. Comparing profilers for mortality modelling

- Explanatory power of given group:

Profiler	Explanatory power†
Mosaic	1,282
Acorn	1,118
P^2	909
Health Acorn	652

- Mosaic or Acorn group around twice as powerful as Health Acorn

Source: Own calculations using Perks survival model for ages 60–95 between 2000 and end-2006.

† "Explanatory power" is the drop in AIC for a model Age*Gender+Time+Group compared with a model for Age*Gender+Time.

3. Comparing profilers for mortality modelling

- Profilers have widely differing numbers of groups: from 4 to 17
- AIC perhaps too forgiving of larger number of parameters?

3. Comparing profilers for mortality modelling

- Map each type code into one of three lifestyle groups
- Number of parameters thus *same* for each profiler

3. Comparing profilers for mortality modelling

Profiler	Explanatory power†
Mosaic	1,588
Acorn	1,322
Health Acorn	1,163
P^2	1,052

- Lifestyles defined by Mosaic and Acorn still the best of the bunch
- Order above is unchanged using a five-level lifestyle (not shown)

Source: Own calculations using Perks survival model for ages 60–95 between 2000 and end-2006.

† "Explanatory power" is drop in AIC from Age*Gender+Time to Age*(Gender+Lifestyle)+Time, where Lifestyle is an optimised mapping of type code to a simpler three-level classification.

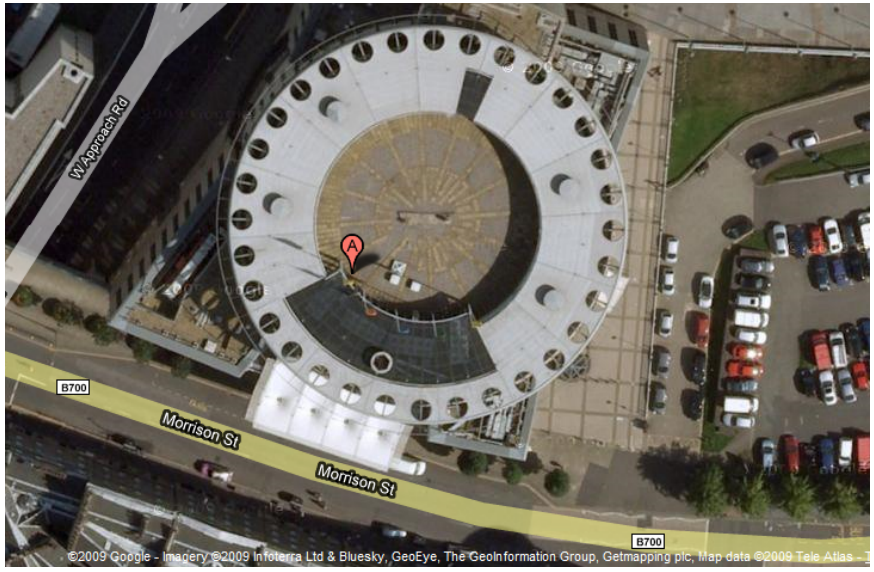
4. Enhancing geodemographic profilers

4. Enhancing profilers for mortality modelling

- Remove profiles for "large-user" postcodes
- Add discontinued postcodes
- Add postcodes for crown dependencies

4. Enhancements for mortality modelling — I

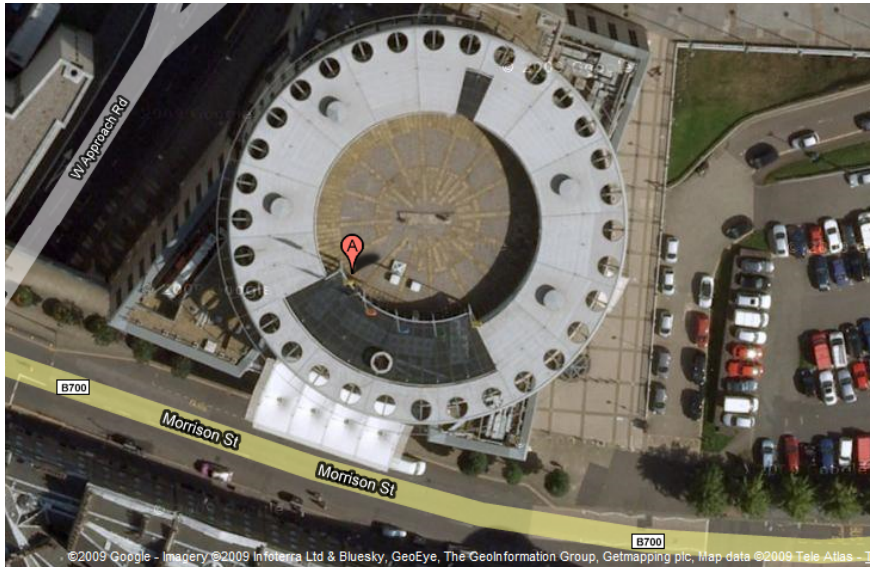
- Large-user postcodes assigned to non-residential buildings
- Sometimes spuriously assigned to a geodemographic type
- Example: EH3 8EE is EICC in Edinburgh



Source: Satellite image from Google Maps

4. Enhancements for mortality modelling — I

- Example: EH3 8EE is EICC in Edinburgh:



- EH3 8EE has Acorn type E18 "Multi-ethnic young, converted flats"
- EH3 8EE has P^2 type H22 "Students in the Community"

Source: Profiles for EH3 8EE from Acorn 07 directory and Beacon Dodsworth P^2 directory, satellite image from Google Maps

4. Enhancements for mortality modelling — II

- Around 1.7m current postcodes
- Around 2.4m postcodes including discontinued ones
- Standard marketing profilers often need old postcodes added back

4. Enhancements for mortality modelling — III

- Crown dependencies not part of United Kingdom
- No geodemographic data, so treated as unrecognised
- Add postcodes for Guernsey (GY), Jersey (JE) and Isle of Man (IM)

4. Enhancements for mortality modelling — III

- Surprising differences between crown dependencies:

Parameter	Estimate	Std. error	p-value	Significance
Age	0.122072	0.0006	0	***
Gender.M	0.465098	0.0117	0	***
Intercept	-13.1395	0.1229	0	***
Guernsey	0	n/a	n/a	n/a
Jersey	0.301669	0.1417	0.0333	*
Isle of Man	0.0699179	0.156	0.654	

Source: Longevitas Ltd. Selected parameters from a model of regional variation of U.K. and Crown Dependencies (Guernsey, Jersey and the Isle of Man). The baseline for the intercept is Guernsey, and a Perks survival model has been fitted for ages 60-95 over the years 2000-2006.

5. Conclusions

5. Conclusions

- Geodemographic profile powerful explanatory variable for mortality
- Not all geodemographic profilers are equal
- Health-based profilers not as good as marketing profilers
- Off-the-shelf systems can be (should be!) enhanced for mortality work

References

Madrigal, A. M., Matthews, F. E., Patel, D. D., Gaches, A. T. and Baxter, S. D. **2009** *What longevity predictors should be allowed for when valuing pension scheme liabilities?*, British Actuarial Journal (to appear)

Richards, S. J. **2008** *Applying survival models to pensioner mortality data*, British Actuarial Journal **14** (to appear)

Richards, S. J. and Currie, I. D. **2009** *Longevity risk and annuity pricing with the Lee-Carter model*, British Actuarial Journal (to appear)