Exploring trajectories of crime at a Local Authority level: comparing and combining latent class and multi-level approaches

Interim results

Dr Ellie Bates and Dr Rebecca Pillinger
AQMEN, University of Edinburgh
ellie.bates@ed.ac.uk
Research Questions

- Given the crime drop in the last decade is there variation between local authorities in the amount of crime fall they’ve seen?
  - This is investigated for two crime types: - Violence and Burglary / Housebreaking
  - Is there also variation between the two crime types?

- Does the type of trajectory or growth curve model chosen to investigate this impact on results?

- [Are there differences between Scotland and England and Wales?]
England & Wales
Community Safety Partnerships (CSPs) (302)
Usually Local Authorities
2004/5 to 2014/15
Police Recorded Crime; Source: Home Office

For these models in England and Wales CSPs are excluded where there are not data for all years; or there is not population data; as well as the City of London*. Reasons for missing data can be boundary changes, mergers or that CSP boundaries do not reflect local authority areas.

*The City of London is an extreme outlier with very high crime for resident population (potentially reflecting that the resident population estimate is a poor indicator of population level in the City of London area).
Crime Definitions
Violence

• Attempted murder, serious and common assaults and woundings, with and without injury
  • includes assaults occasioning grievous AND actual bodily harm (difference in intent is not considered)
  • includes racially motivated assaults
  • excludes murder and other forms of homicide
  • England and Wales Home Office Crime Recording Standard Codes: 2;5;5A;5B;5C;5D;5E;8A;8D;8F;8G;8H;8J;8K;8N;8P;104;105A;105B
  • Scottish Crime Recording Standard Codes: 002000; 004000; 047001
Violence

• This definition is used because:

  • It is arguably less sensitive to crime code definition changes in violence

  The England and Wales national crime recording standard was introduced 2002-3 and violence codes were amended in 2008-09 and 2012-13.

  • It allows for comparison with Scottish data

  The more commonly used England and Wales violence with injury definition could be used for E&W data only but is not comparable with Scotland as Scottish crime recording does not split less serious assaults into violence with and without injury.
Burglary / Housebreaking

• All Burglaries and Attempted Burglaries
  • In England and Wales you must enter as a trespassers you do not actually have to break-in. There is no equivalent to the aggravated burglary crime code in Scotland.

• All Housebreaking and Attempted Housebreaking
  • In Scotland you must break-in – defined as overcoming the properties security

• Both home and business premises (domestic and non-domestic) are included because Scotland and England have different definitions of what counts as a dwelling.
  • E&W codes: 28;28A;28B;28C;29;30;30A;30B;31
  • Scottish codes: 19004;19007;19010;19005;19008;19011;19006;19009;19012
The Models
A model for trajectories

\[ y_{ti} = \beta_0 + \beta_1 t_i + \beta_2 t_i^2 + \beta_3 t_i^3 \]

\[ + e_{(3+t)i} \]

\[
\begin{bmatrix}
    e_{4i} \\
    e_{5i} \\
    \vdots \\
    e_{(3+T)i}
\end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e4}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{e5}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{e(3+T)}^2 \end{bmatrix} \right)
\]
A model for trajectories

\[ y_{ti} = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + e_{(3+t)i} \]

\[
\begin{bmatrix}
  e_{4i} \\
  e_{5i} \\
  \vdots \\
  e_{(3+T)i}
\end{bmatrix}
\sim N\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e_4} & 0 & \cdots & 0 \\ 0 & \sigma^2_{e_5} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2_{e_{(3+T)}} \end{bmatrix} \right)
\]
Latent Class Growth Analysis (LGCA)

\[ y_{ti} = \beta_0^c + \beta_1^c \text{time}_t + \beta_2^c \text{time}^2_t + \beta_3^c \text{time}^3_t \]
\[ + e_{(3+t)i} \]

\[ c_i \sim \text{Multinomial} \left( 1, P(c_i|y_{i1}, \ldots, y_{iT}) \right) \]

\[ P(c_i = c|y_1, \ldots, y_{iT}) = \frac{\left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_i = c) \right) P(c)}{\sum_{c=1}^{C} \left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_i = c) \right) P(c)} \]

\[
\begin{bmatrix}
  e_{4i} \\
  e_{5i} \\
  \vdots \\
  e_{(3+T)i}
\end{bmatrix}
\sim N \left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e4}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{e5}^2 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \sigma_{e(3+T)}^2 \end{bmatrix} \right)
\]
Latent Growth Model (LGM)

\[ y_{ti} = \beta_0 + \beta_1 \text{time}_t + \beta_2 \text{time}_t^2 + \beta_3 \text{time}_t^3 \\
+ u_{0i} + u_{1i} \text{time} + u_{2i} \text{time}_t^2 + u_{3i} \text{time}_t^3 \\
+ e_{(3+T)i} \]

\[
\begin{bmatrix}
  u_{0i} \\
  u_{1i} \\
  u_{2i} \\
  u_{3i}
\end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix}
  \sigma^2_{u0} & \sigma_{u01} & \sigma^2_{u1} & \sigma^2_{u2} & \sigma^2_{u3} \\
  \sigma_{u01} & \sigma^2_{u02} & \sigma_{u12} & \sigma^2_{u2} & \sigma^2_{u3} \\
  \sigma_{u01} & \sigma_{u12} & \sigma^2_{u03} & \sigma^2_{u2} & \sigma^2_{u3} \\
  \sigma_{u01} & \sigma_{u12} & \sigma_{u03} & \sigma^2_{u2} & \sigma^2_{u3} \\
  \end{bmatrix} \right)
\]

\[
\begin{bmatrix}
  e_{4i} \\
  e_{5i} \\
  \vdots \\
  e_{(3+T)i}
\end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix}
  \sigma^2_{e4} & 0 & \cdots & 0 \\
  0 & \sigma^2_{e5} & \cdots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & \sigma^2_{e(3+T)} \\
  \end{bmatrix} \right)
\]
Growth Mixture Model with fixed slopes (GMM-FS)

\[
y_{ti} = \beta_{0}^{c_{i}} + \beta_{1}^{c_{i}} \text{time}_t + \beta_{2}^{c_{i}} \text{time}_t^2 + \beta_{3}^{c_{i}} \text{time}_t^3 + u_{0i} + e_{(3+t)i}
\]

\[
c_{i} \sim \text{Multinomial} \left( 1, P(c_{i}|y_{1i}, \ldots, y_{Ti}) \right)
\]

\[
P(c_{i} = c|y_{1i}, \ldots, y_{Ti}) = \frac{\left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_{i} = c) \right) P(c)}{\sum_{c=1}^{C} \left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_{i} = c) \right) P(c)}
\]

\[
u_{0i} \sim N\left( 0, \sigma_{u0}^2 \right)
\]

\[
\begin{bmatrix}
e_{4i} \\
e_{5i} \\
\vdots \\
e_{(3+T)i}
\end{bmatrix} \sim N\left( \begin{bmatrix}0 \\0 \\0 \\0 \end{bmatrix}, \begin{bmatrix} \sigma_{e4}^2 & 0 & \cdots & 0 \\
0 & \sigma_{e5}^2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \sigma_{e(3+T)}^2 \\
\end{bmatrix} \right)
\]
Growth Mixture Model (GMM)

\[ y_{ti} = \beta_0^{ci} + \beta_1^{ci} \text{time}_t + \beta_2^{ci} \text{time}_t^2 + \beta_3^{ci} \text{time}_t^3 + u_{0i} + u_{1i} \text{time} + u_{2i} \text{time}_t^2 + u_{3i} \text{time}_t^3 + e_{(3+t)i} \]

\[ c_i \sim \text{Multinomial}(1, P(c_i|y_1, \ldots, y_{Ti})) \]

\[ P(c_i = c|y_1, \ldots, y_{Ti}) = \frac{\left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_i = c) \right) P(c)}{\sum_{c=1}^{C} \left( \prod_{t=1}^{T} P(Y_t = y_{ti}|c_i = c) \right) P(c)} \]

\[
\begin{bmatrix}
  u_{0i} \\
  u_{1i} \\
  u_{2i} \\
  u_{3i}
\end{bmatrix}
\sim
N\left(
\begin{bmatrix}
  0 \\
  0 \\
  0 \\
  0
\end{bmatrix},
\begin{bmatrix}
  \sigma_{u0}^2 & \sigma_{u01} & \sigma_{u1}^2 \\
  \sigma_{u01} & \sigma_{u02} & \sigma_{u12} & \sigma_{u2}^2 \\
  \sigma_{u02} & \sigma_{u12} & \sigma_{u13} & \sigma_{u23} & \sigma_{u3}^2 \\
  \sigma_{u01} & \sigma_{u12} & \sigma_{u13} & \sigma_{u23} & \sigma_{u3}^2
\end{bmatrix}
\right)
\]

\[
\begin{bmatrix}
  e_{4i} \\
  e_{5i} \\
  \vdots \\
  e_{(3+T)i}
\end{bmatrix}
\sim
N\left(
\begin{bmatrix}
  0 \\
  0 \\
  \vdots \\
  0
\end{bmatrix},
\begin{bmatrix}
  \sigma_{e4}^2 \\
  0 \\
  \vdots \\
  0
\end{bmatrix}
\right)
\]

\[
\begin{bmatrix}
  e_{(3+T)i}
\end{bmatrix}
\sim
N\left(
\begin{bmatrix}
  0 \\
  0 \\
  \vdots \\
  0
\end{bmatrix},
\begin{bmatrix}
  \sigma_{e(3+T)}^2
\end{bmatrix}
\right)
\]
Additional Model Information

• Software used: Mplus 7.3
• All models run with continuous data – crime rates per 1000 people
  \[\text{crime rate per 1000 people} = \left(\frac{\text{crime count}}{\text{resident population estimate}}\right) \times 1000\]
• A Maximum Likelihood estimator robust for skew and non-independence is used
  [Mplus option ESTIMATOR = MLR]
• Models were centred at the mid-point for the 11 years 2009/10 (this is set as time 0 with other time points specified from -0.5 to +0.5 in order of years)
LGM
Latent Growth Model
‘Multi-level’
Significant variation in intercepts, slopes and co-variances for both crime types.

**Violence**

- Sample means, General
- Estimated means, General

BIC=11923
ABIC=11844

**Burglary**

- Sample means, General
- Estimated means, General

BIC=11781
ABIC=11718
Model Comparison
LGCA
Fixed Intercept
Fixed Slope
12 classes

GMM – FS
Varying Intercept
Fixed Slope
10 (or 12) classes

GMM
Varying Intercept
Varying Slope
6 (or 7) classes

Model Fit BIC and ABIC

ENGLAND AND WALES - Violence
Model Fit BIC and ABIC – same y axis scale

**ENGLAND AND WALES - Violence**

- **LGCA**
  - Fixed Intercept
  - Fixed Slope
  - 12 classes
  - BIC = 12935
  - ABIC = 12713

- **GMM – FS**
  - Varying Intercept
  - Fixed Slope
  - 10 (or 12) classes
  - BIC = 12039
  - ABIC = 11846

- **GMM**
  - Varying Intercept
  - Varying Slope
  - 6 (or 7) classes
  - BIC = 11857
  - ABIC = 11698
LGCA
Fixed Intercept
Fixed Slope
12 classes

GMM – FS
Varying Intercept
Fixed Slope
10 (or 12) classes

GMM
Varying Intercept
Varying Slope
6 (or 7) classes

Model Fit - Entropy
ENGLAND AND WALES - Violence
E&W Violence - ‘Best fit’ Model Estimates

LGCA
Fixed Intercept
Fixed Slope
12 classes

GMM – FS
Varying Intercept
Fixed Slope
10 (or 12) classes

GMM
Varying Intercept
Varying Slope
6 (or 7) classes
At 4 Classes – an illustration of model differences

ENGLAND AND WALES - Violence
Estimated Means and Individual Observed Trajectories of Most Likely Class Members

Labels:- Model Type; Class (C); Probability based class membership N and %; [N most likely class membership]
GMM-FS C2=4.8 1.6% [5]

GMM – FS C8=4.0 1.3% [4]

GMM-FS C10=10.8 3.6% [10]

GMM-FS C7=4.2 1.4% [4]

GMM-FS C5=87.5 29.0% [91]
Similar class membership between models?
Yes and No…
LGCA Class 3, GMM Class 2 and GMM-FS Class 7 have similarities.
Some overlap only between LGCA Class 9 and other models Class 1

GMM-FS and GMM Class 1 both very similar
No exact LGCA Comparison
C10 appears to have some similarities

LGCA C10=19.04  6.3% [19]

GMM-FS C8 and C3 members (or areas with very similar trajectories) appear in GMM Class 5

GMM – FS C8=4.0  1.3% [4]
GMM-FS C2=4.8  1.6% [5]
GMM - Class 5 9.2  3.0% [9]
Appears to be some overlap between LGCA C10 and C8 and GMM –FS C10 and GMM – 6.
GMM-FS C10 and GMM C6 again appear very similar.

LGCA C10=19.04 6.3% [19]
LGCA C8=23.3 7.7% [23]

GMM-FS C10=10.8 3.6% [10]
GMM - Class 6 = 16.8 5.6% [16]
E&W Violence - ‘Best fit’ Model Estimates

GMM
Varying Intercept
Varying Slope
6 (or 7) classes

GMM – FS
Varying Intercept
Fixed Slope
10 (or 12) classes

LGCA
Fixed Intercept
Fixed Slope
12 classes

GMM
Varying Intercept
Varying Slope
6 (or 7) classes
Model Comparison
Burglary

A very brief note…
LGCA
Fixed Intercept
Fixed Slope
14 (or 18?) classes

GMM – FS
Varying Intercept
Fixed Slope
10 (or 14) classes

GMM
Varying Intercept
Varying Slope
5 (or 8) classes

ENGLAND AND WALES - Burglary

BIC=12390
ABIC=12180

BIC=11781
ABIC=11718

BIC=11855
ABIC=11693

BIC=11711
ABIC=11597

BIC=11781
ABIC=11718
Conclusions

• There are differing trajectories at the ‘regional’ CSP / Local Authority level between local CSP areas.

• Model choice has a clear effect on the of crime trajectories found.

• Violence and burglary have differing crime trajectory patterns, across time; for both crime types, the model choice impacts on the number of groups of areas with distinct modelled trajectories, and to a lesser extent the trajectory ‘shape’.

• If substantive findings here are replicated with (potentially) better specified / more robust models, this may suggest that there may be inequality in the crime trajectories of violence and burglary between Community Safety Partnerships.

• If a national crime fall is being experienced differently in different local areas - is it time to start thinking about crime as an inequality issue?
Next Steps

• Investigating models with logged data
• Investigating models with count data
• Looking at additional measures to compare model fit between classes and between models
• Looking further at variations in class membership, numbers of groups and types of trajectory found between models
• Further consideration of how to handle ‘missing’ data for England and Wales
• Further investigation of differences between the two crime types (Violence and Burglary)
• Investigating whether there does appear to be a difference between Scotland and England and Wales in crime trajectories – especially for violence (as suggested by initial results not shown here)
  • Investigating issues of power with Scottish Models
  • Combining English, Welsh and Scottish data into one model
Acknowledgements

Data Sources
Under the People and Society: Population and Migration Information on People and Society: Population and Migration tab; Selecting option - Resident Population Estimates, All Persons and downloading data for each year 2001 to 2014

Software
Mplus 7.3

Key reference

Funding
This research was supported by funding from the Economic and Social Research Council, Grant Reference Number ES/K006460/1