# 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

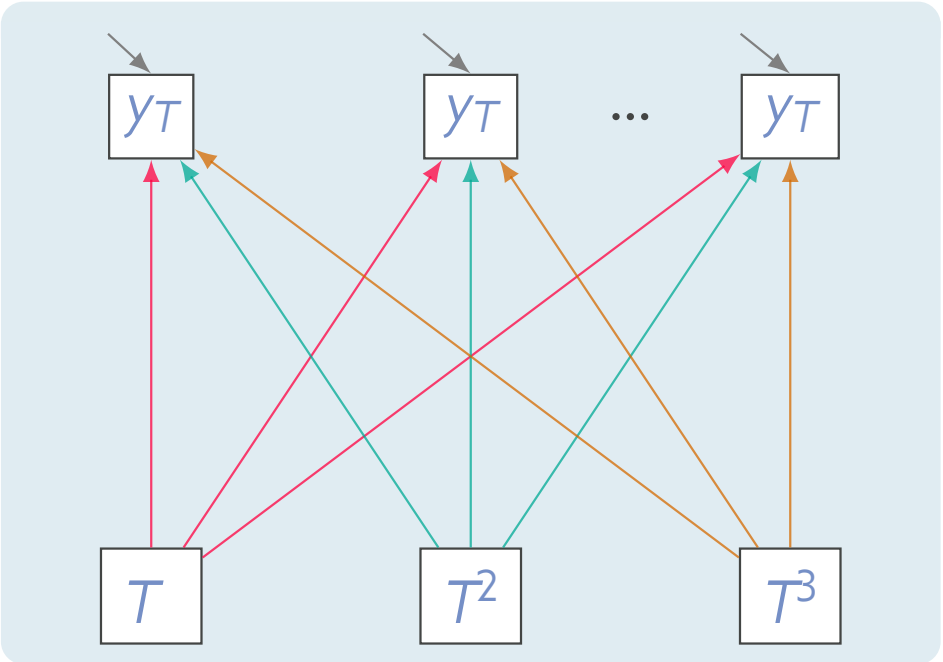
- All Burglaries and Attempted Burglaries
  - In England and Wales you must enter as a trespassers you do not actually have to break-in. There is no equivalent to the aggravated burglary crime code in Scotland.
- All Housebreaking  and Attempted Housebreaking
  - In Scotland you must break-in – defined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

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

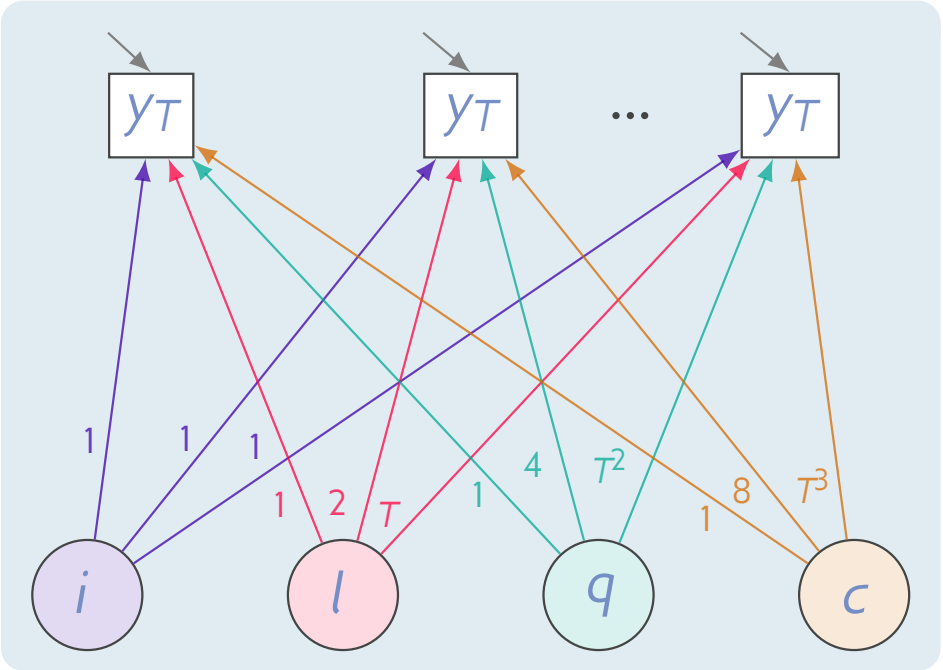
$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim \text{N}\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e4} & & & \\ 0 & \sigma^2_{e5} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \sigma^2_{e(3+T)} \end{bmatrix} \right)$$



# A model for trajectories

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

$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e4} & & & \\ 0 & \sigma^2_{e5} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \sigma^2_{e(3+T)} \end{bmatrix} \right)$$






AQMeN
Applied Quantitative Methods Network

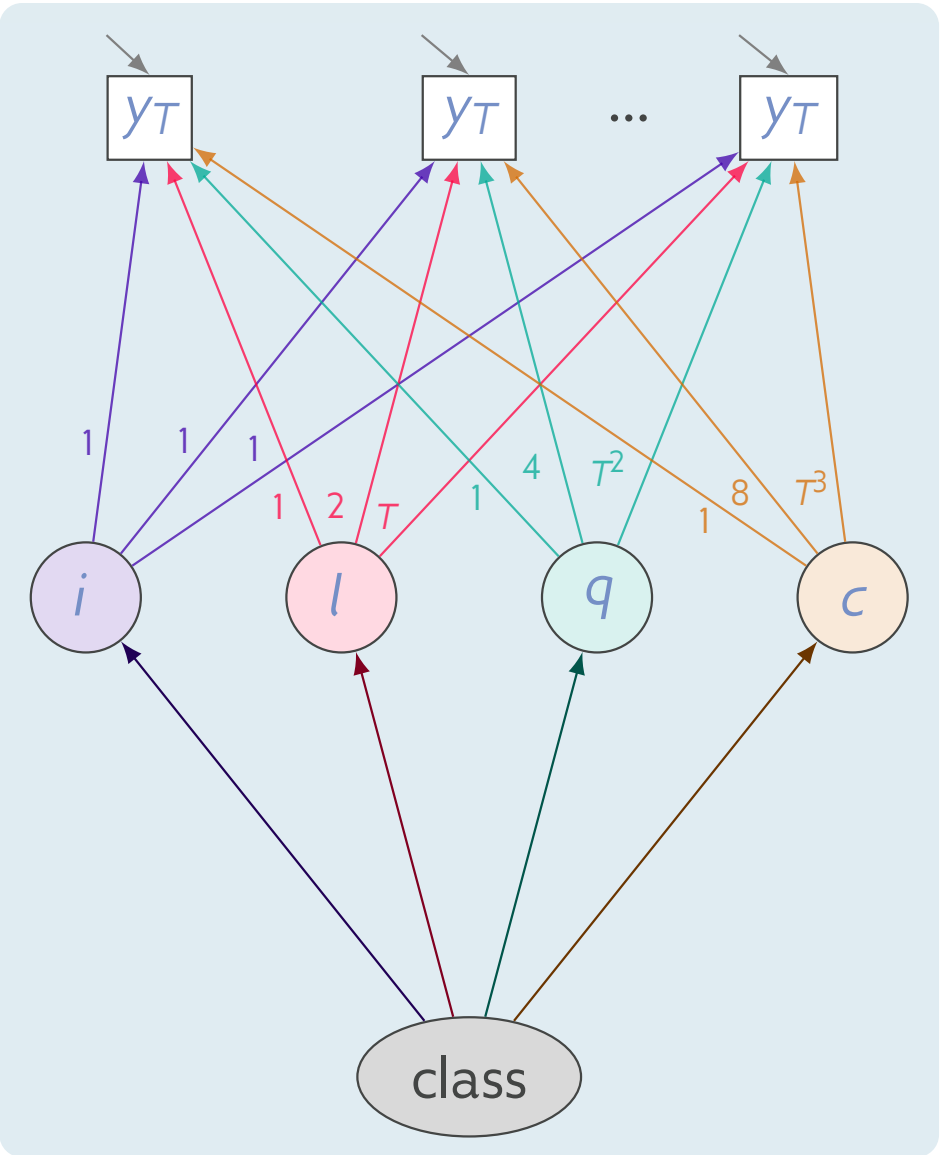
# Latent Class Growth Analysis (LGCA)

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

$$c_i \sim \text{Multinomial}\left(1, \text{P}(c_i|y_{1i}, \ldots, y_{Ti})\right)$$

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

$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e4} & & & \\ 0 & \sigma^2_{e5} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & \sigma^2_{e(3+T)} \end{bmatrix}\right)$$





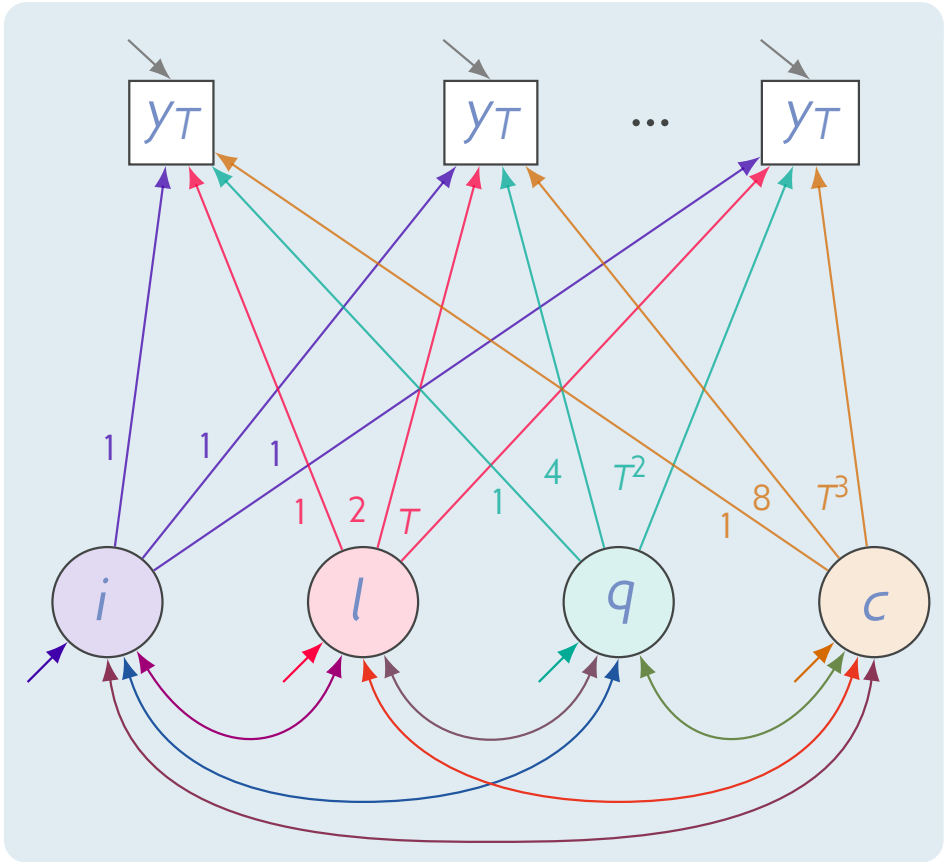

AQMeN
Applied Quantitative Methods Network

# Latent Growth Model (LGM)

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

$$\begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \\ u_{3i} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{u0} & & & \\ \sigma_{u01} & \sigma^2_{u1} & & \\ \sigma_{u02} & \sigma_{u12} & \sigma^2_{u2} & \\ \sigma_{03} & \sigma_{u13} & \sigma_{u23} & \sigma^2_{u3} \end{bmatrix} \right)$$

$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim N\left( \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{e4} & & & \\ 0 & \sigma^2_{e5} & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & \sigma^2_{e(3+T)} \end{bmatrix} \right)$$






AQMeN
Applied Quantitative Methods Network

# Growth Mixture Model with fixed slopes (GMM-FS)

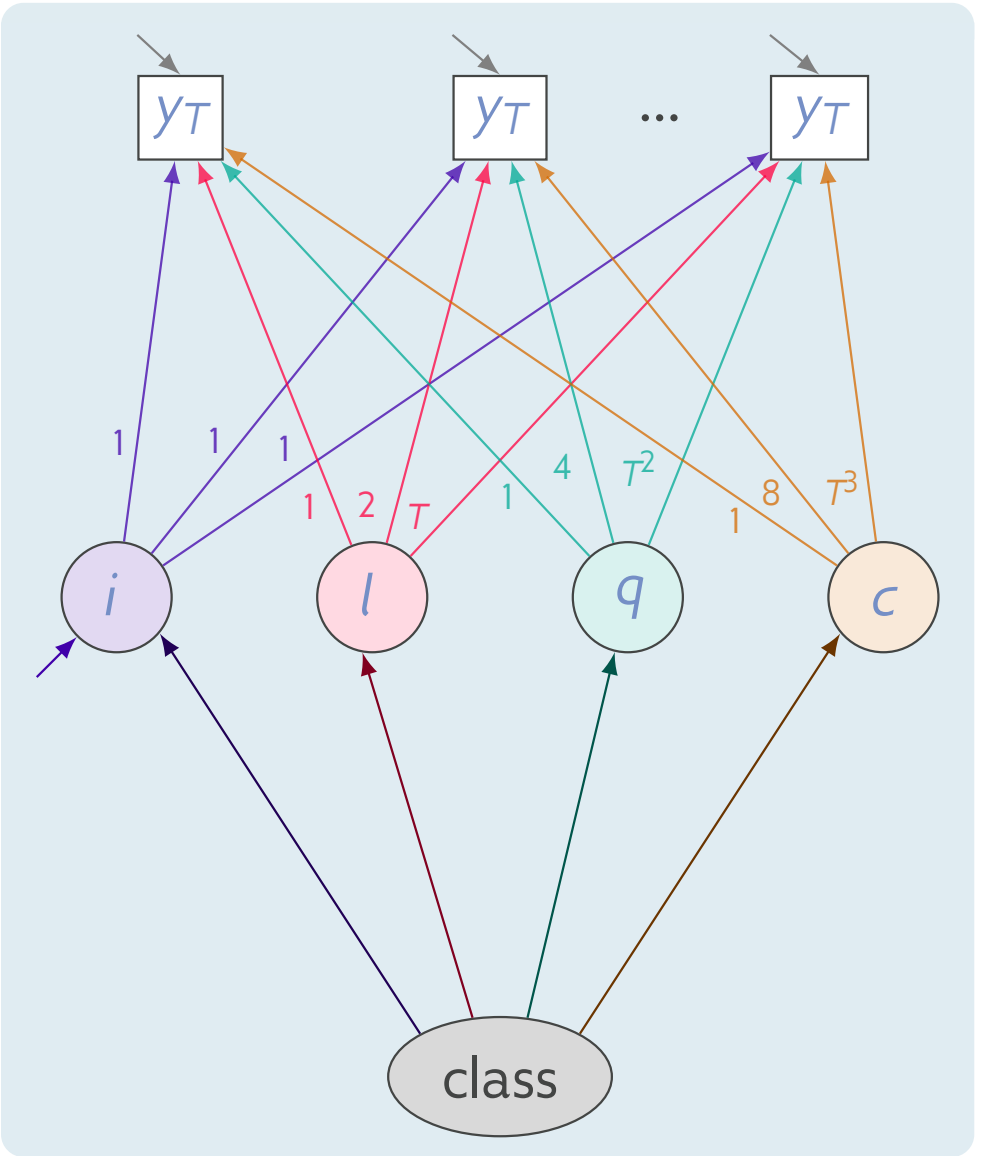
$$y_{ti} = \beta_0^{c_i} + \beta_1^{c_i}\text{time}_t + \beta_2^{c_i}\text{time}_t^2 + \beta_3^{c_i}\text{time}_t^3 + u_{0i} + e_{(3+t)i}$$

$$c_i \sim \text{Multinomial}\left(1, \text{P}(c_i|y_{1i}, \ldots, y_{Ti})\right)$$

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

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

$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e4}^2 & & & \\ 0 & \sigma_{e5}^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & \sigma_{e(3+T)}^2 \end{bmatrix}\right)$$






AQMeN
Applied Quantitative Methods Network

# Growth Mixture Model (GMM)

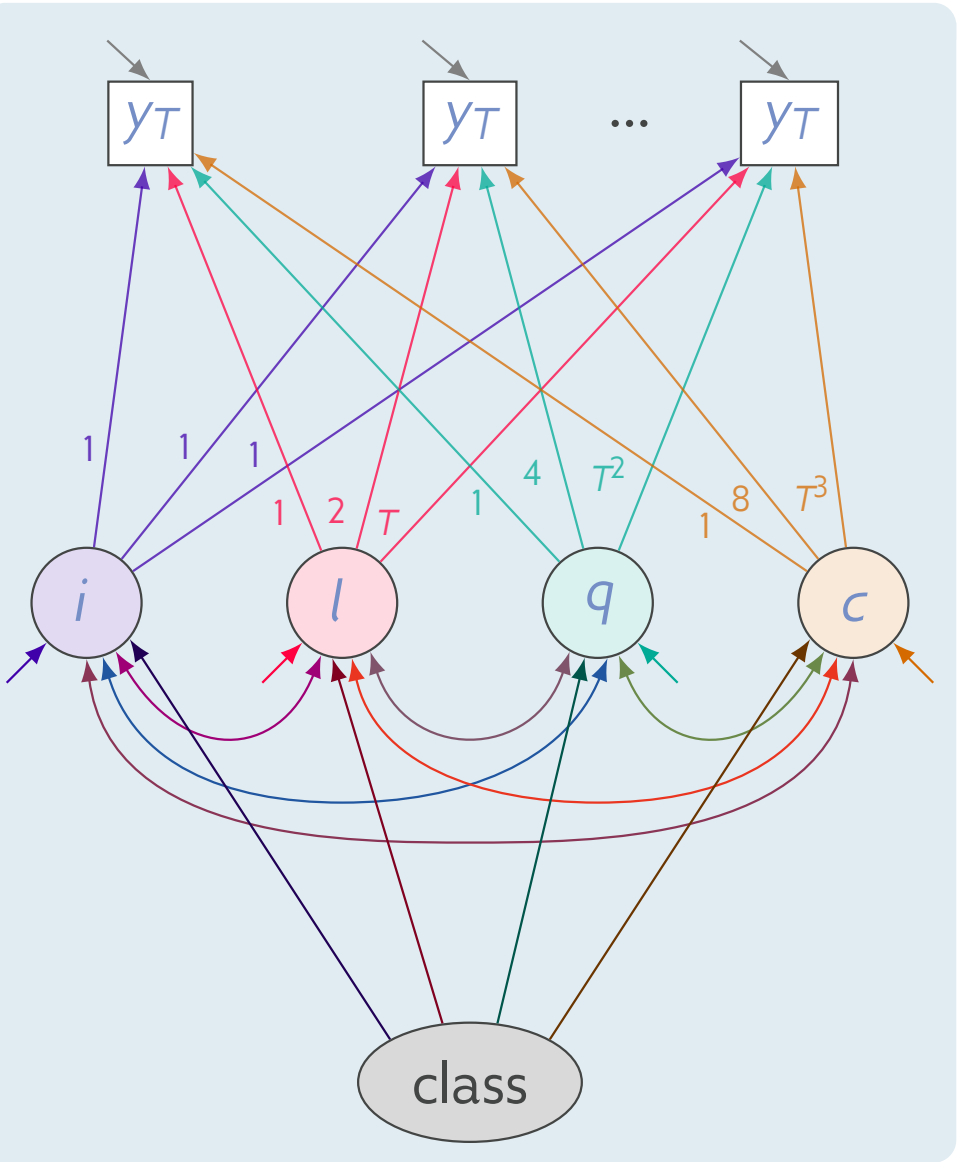
$$y_{ti} = \beta_0^{c_i} + \beta_1^{c_i}\text{time}_t + \beta_2^{c_i}\text{time}_t^2 + \beta_3^{c_i}\text{time}_t^3 + u_{0i} + u_{1i}\text{time} + u_{2i}\text{time}_t^2 + u_{3i}\text{time}_t^3 + e_{(3+t)i}$$

$$c_i \sim \text{Multinomial}\left(1, \text{P}(c_i|y_{1i}, \ldots, y_{Ti})\right)$$

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

$$\begin{bmatrix} u_{0i} \\ u_{1i} \\ u_{2i} \\ u_{3i} \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{u0}^2 & & & \\ \sigma_{u01} & \sigma_{u1}^2 & & \\ \sigma_{u02} & \sigma_{u12} & \sigma_{u2}^2 & \\ \sigma_{03} & \sigma_{u13} & \sigma_{u23} & \sigma_{u3}^2 \end{bmatrix}\right)$$

$$\begin{bmatrix} e_{4i} \\ e_{5i} \\ \vdots \\ e_{(3+T)i} \end{bmatrix} \sim \text{N}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{e4}^2 & & & \\ 0 & \sigma_{e5}^2 & & \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \ldots & \sigma_{e(3+T)}^2 \end{bmatrix}\right)$$



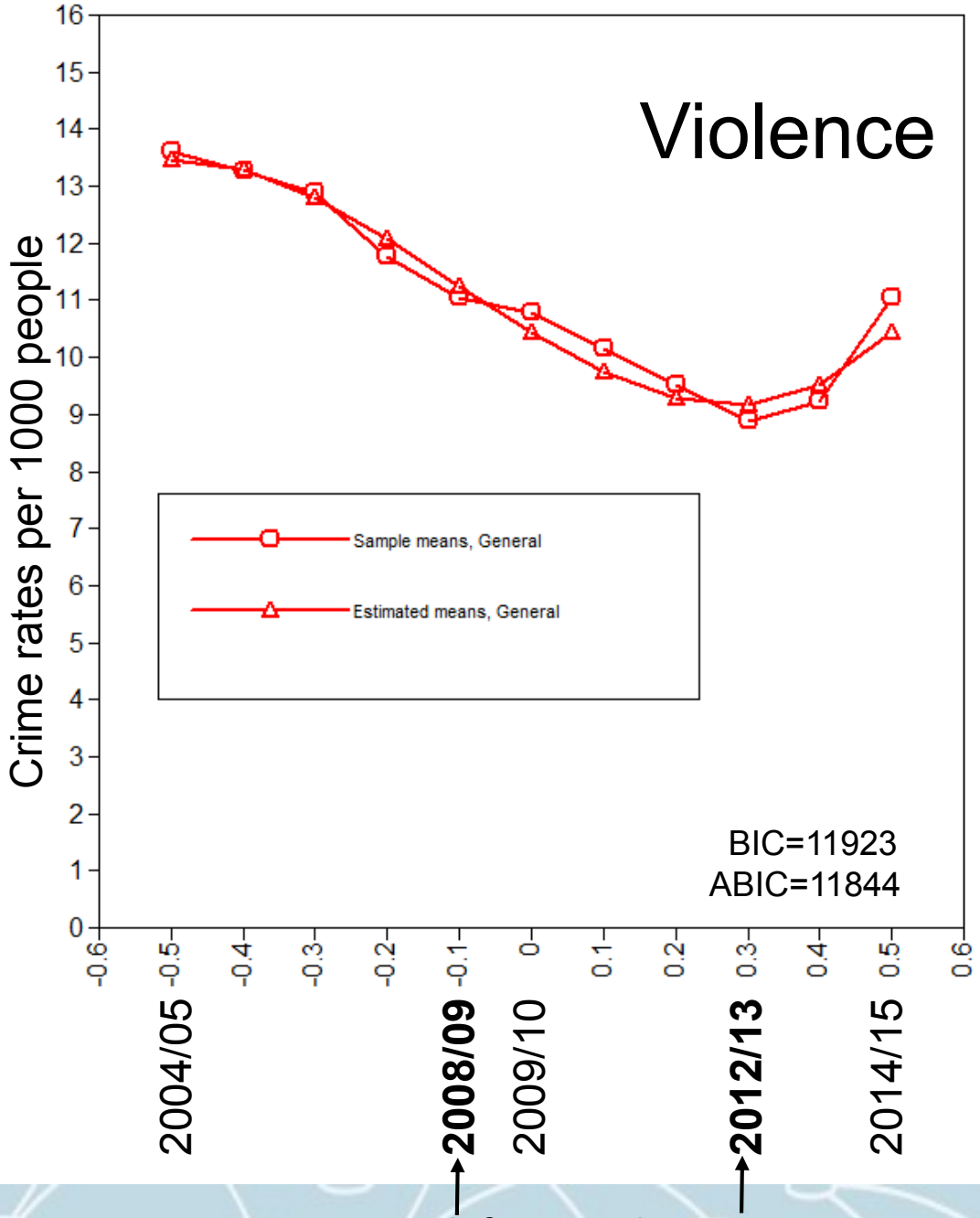



AQMeN
Applied Quantitative Methods Network

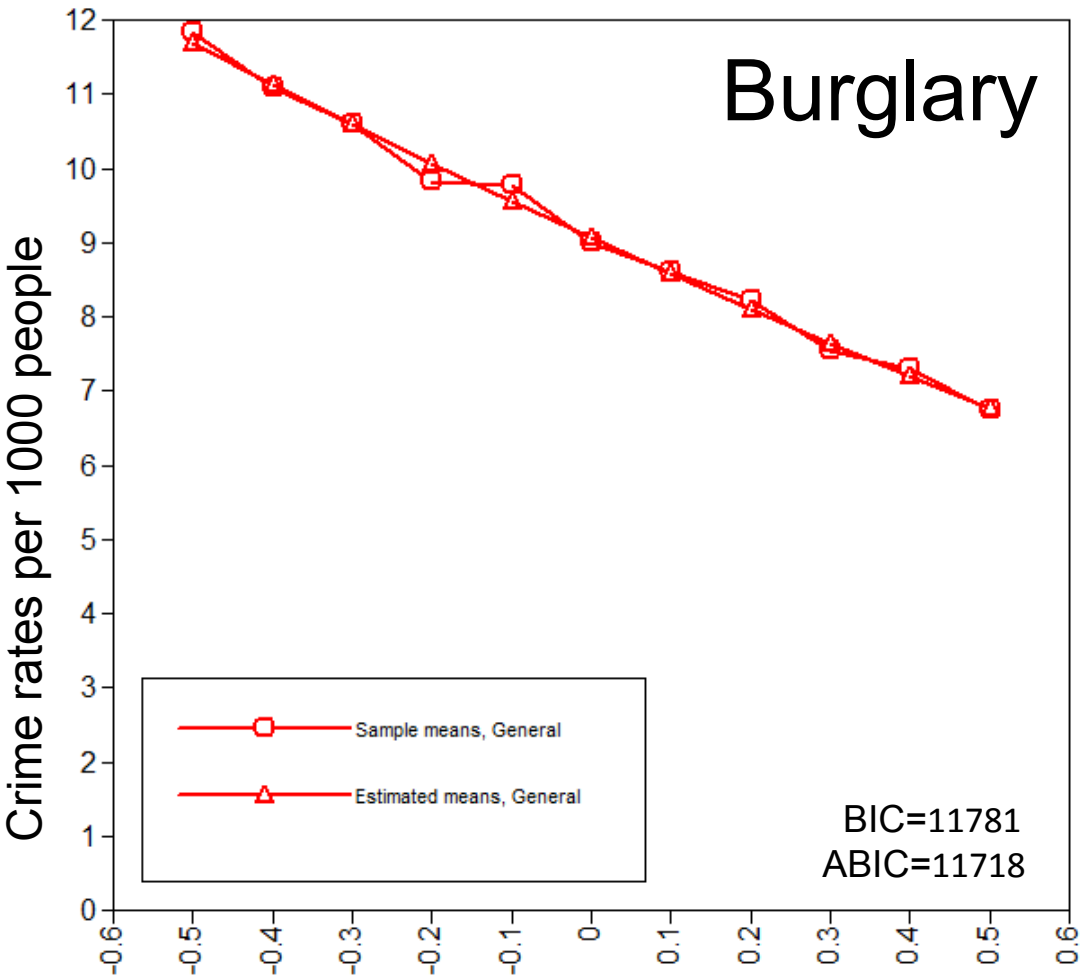
# Additional Model Information

- Software used: Mplus 7.3
- All models run with continuous data – crime rates per 1000 people
  [(crime count / resident population estimate)*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’




ENGLAND AND WALES

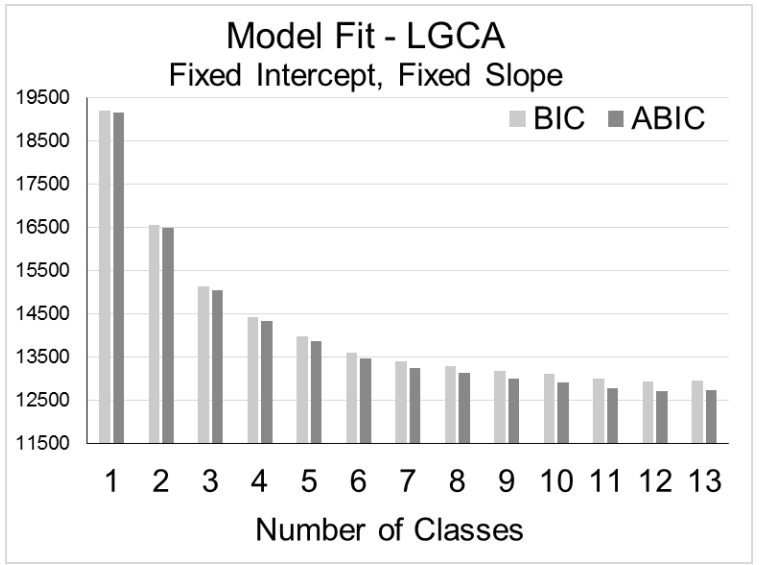


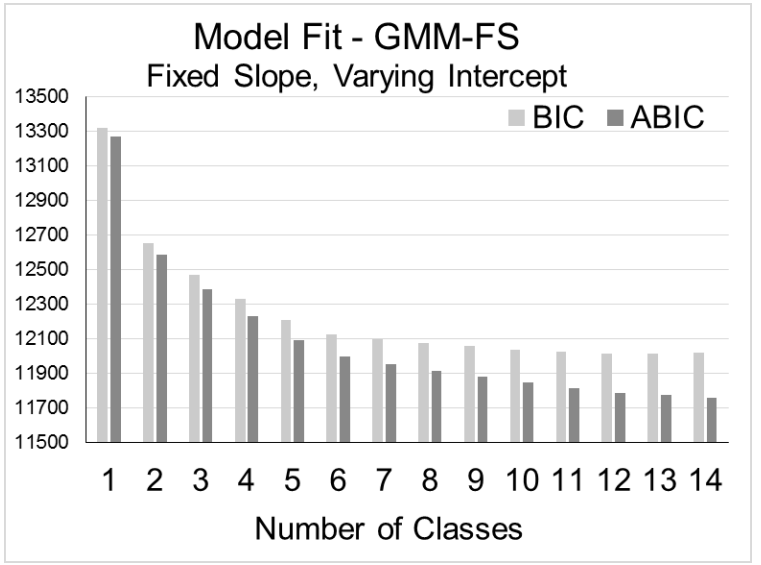

Significant variation in intercepts, slopes and co-variances for both crime types

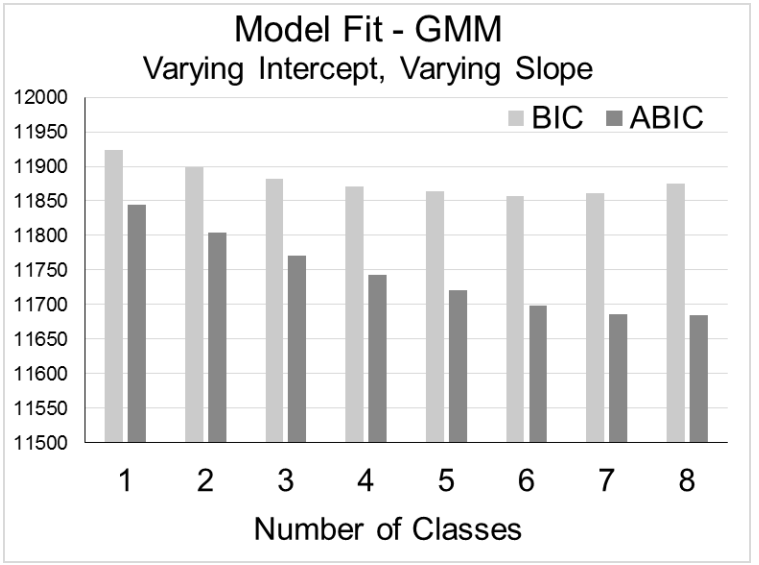


AQMeN
Applied Quantitative Methods Network

# Model Comparison




## LGCA

Fixed Intercept

Fixed Slope

12 classes

## GMM – FS

Varying Intercept

Fixed Slope

10 (or 12) classes

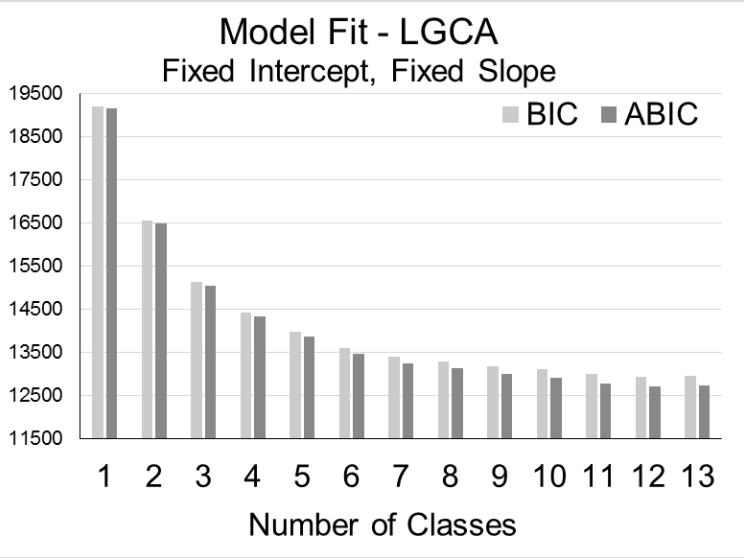
## GMM

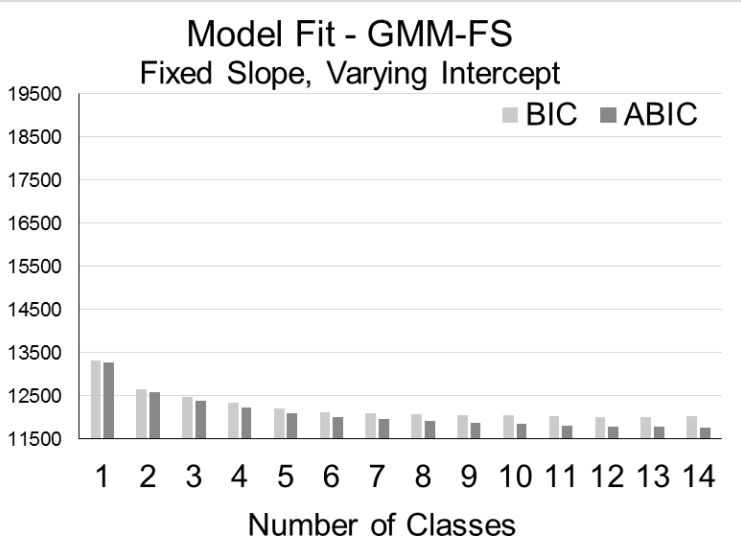
Varying Intercept

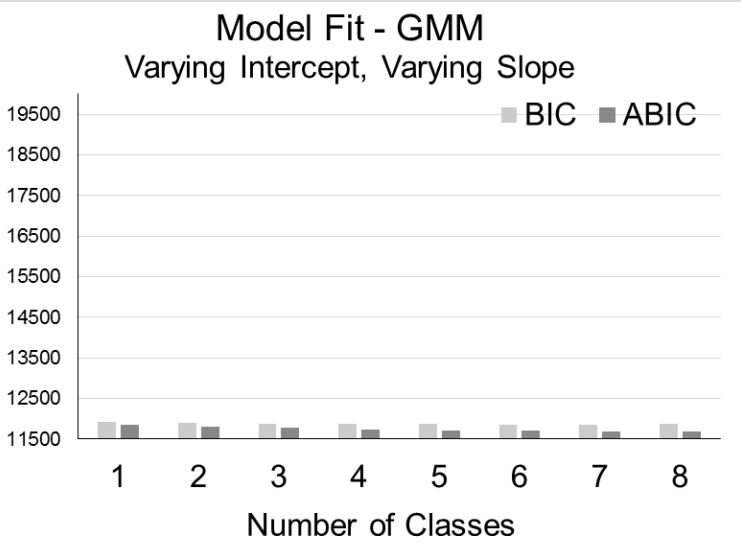
Varying Slope

6 (or 7) classes

# Model Fit BIC and ABIC




# LGCA
BIC=12935
ABIC=12713

Fixed Intercept
Fixed Slope
12 classes

# GMM – FS
BIC=12039
ABIC=11846

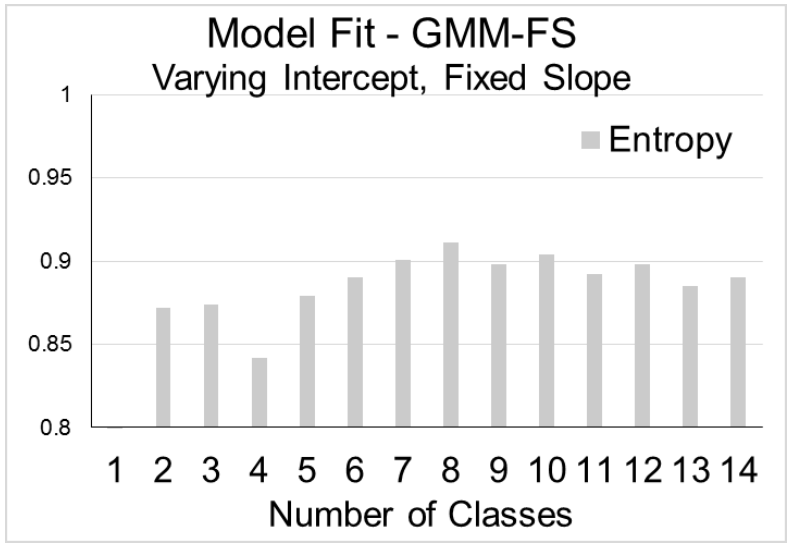
Varying Intercept
Fixed Slope
10 (or 12) classes

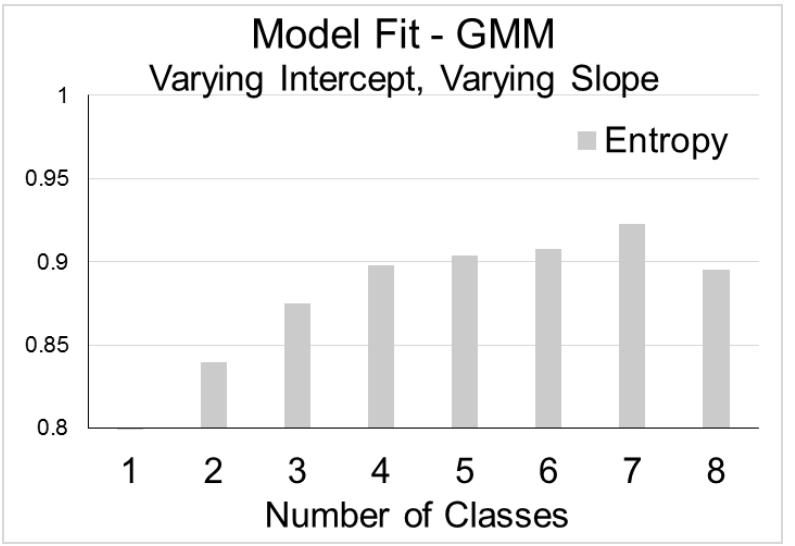
# GMM
BIC=11857
ABIC=11698

Varying Intercept
Varying Slope
6 (or 7) classes

## Model Fit BIC and ABIC – same y axis scale


ENGLAND AND WALES - Violence


# LGCA

Fixed Intercept

Fixed Slope

12 classes

# GMM – FS

Varying Intercept

Fixed Slope

10 (or 12) classes

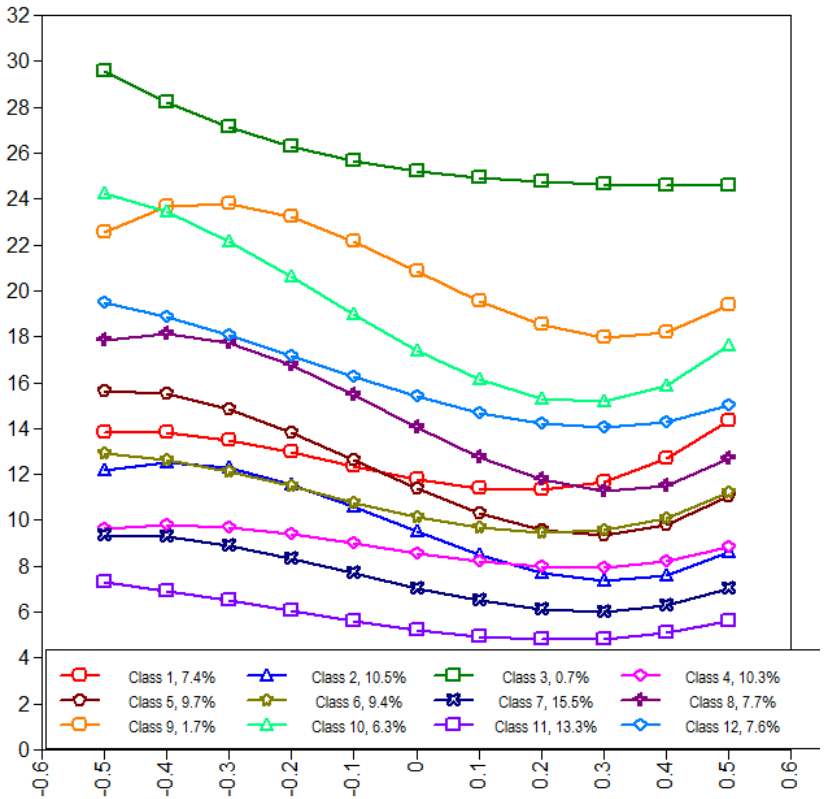
# GMM

Varying Intercept

Varying Slope

6 (or 7) classes

# Model Fit - Entropy

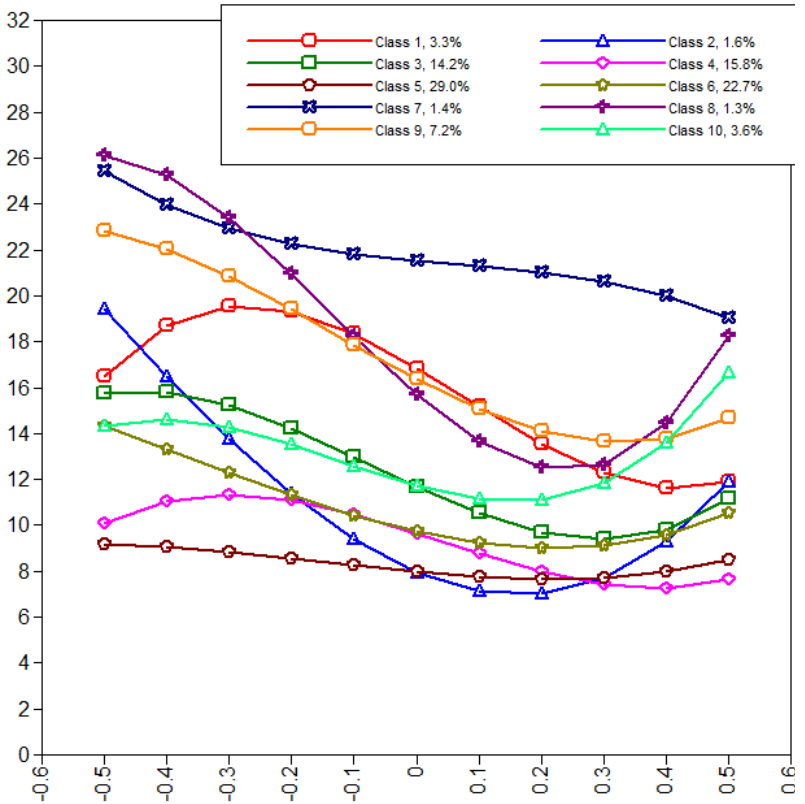



# LGCA

Fixed Intercept

Fixed Slope

12 classes

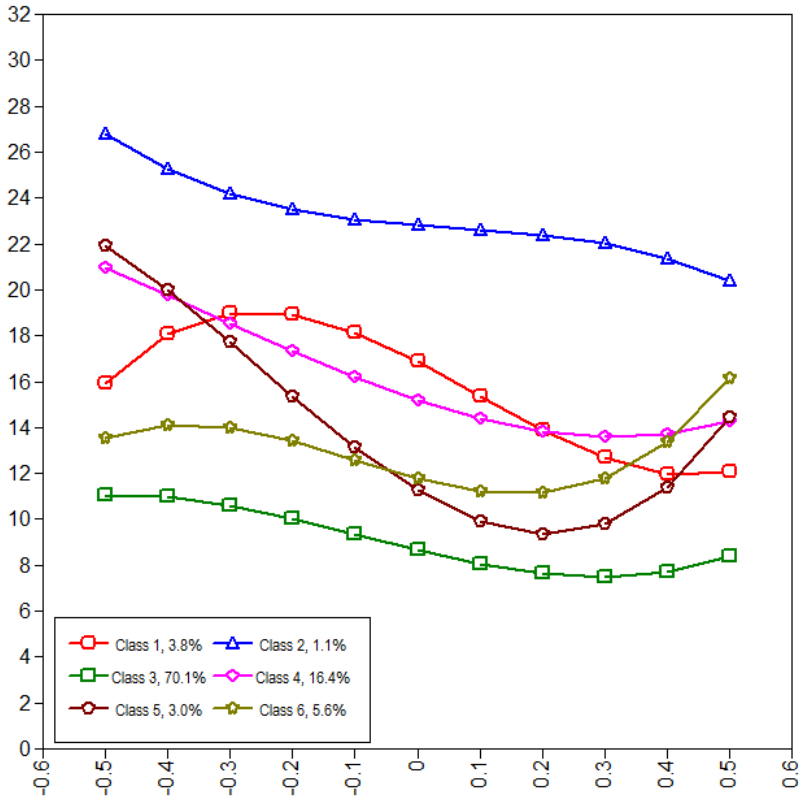



# GMM – FS

Varying Intercept

Fixed Slope

10 (or 12) classes




# GMM

Varying Intercept

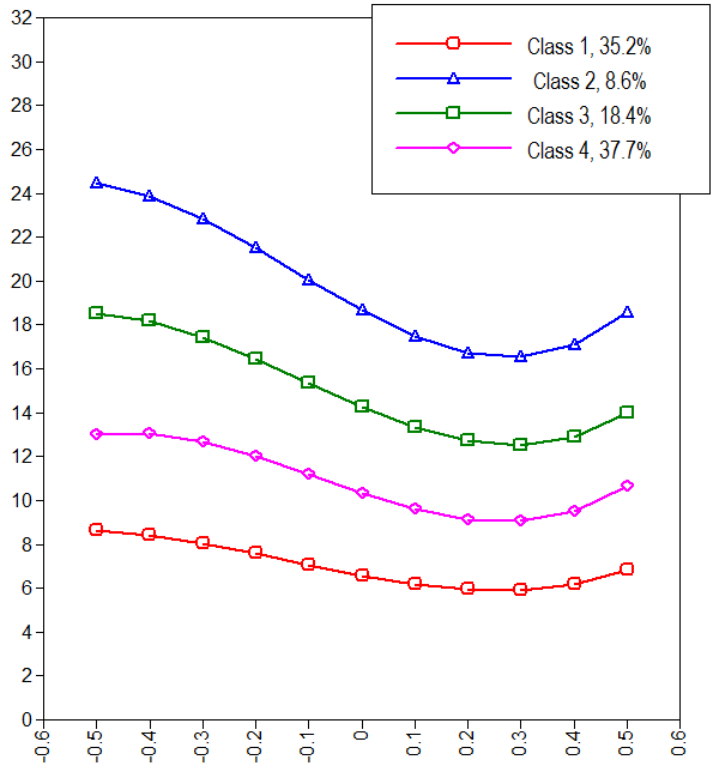
Varying Slope

6 (or 7) classes



E&W Violence - 'Best fit' Model Estimates

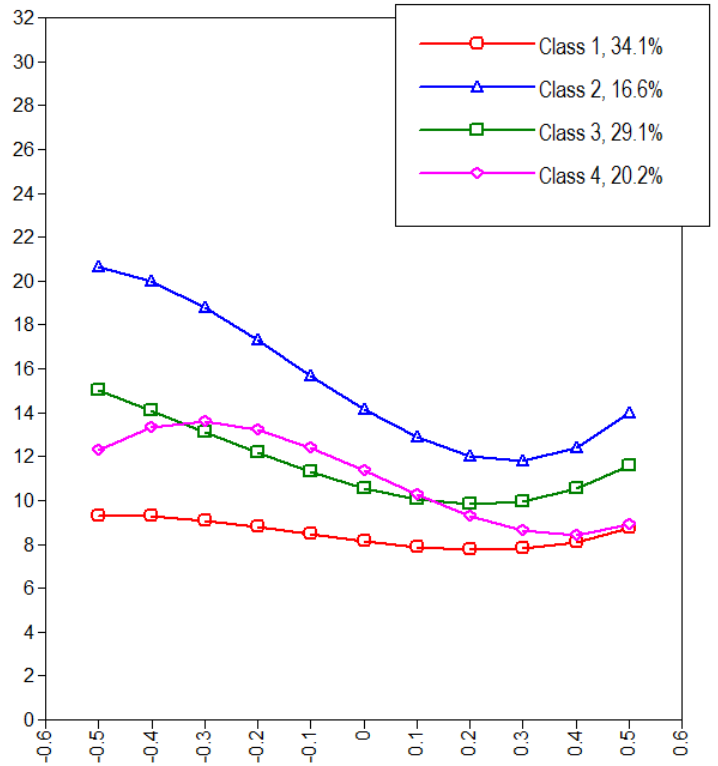
# At 4 Classes – an illustration of model differences




## LGCA

Fixed Intercept

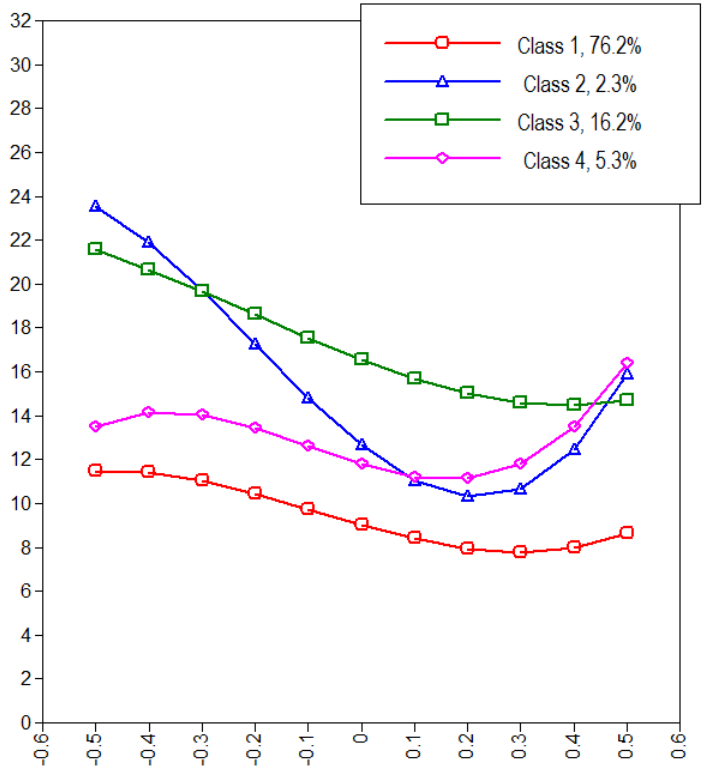
Fixed Slope




## GMM – FS

Varying Intercept

Fixed Slope




## GMM

Varying Intercept

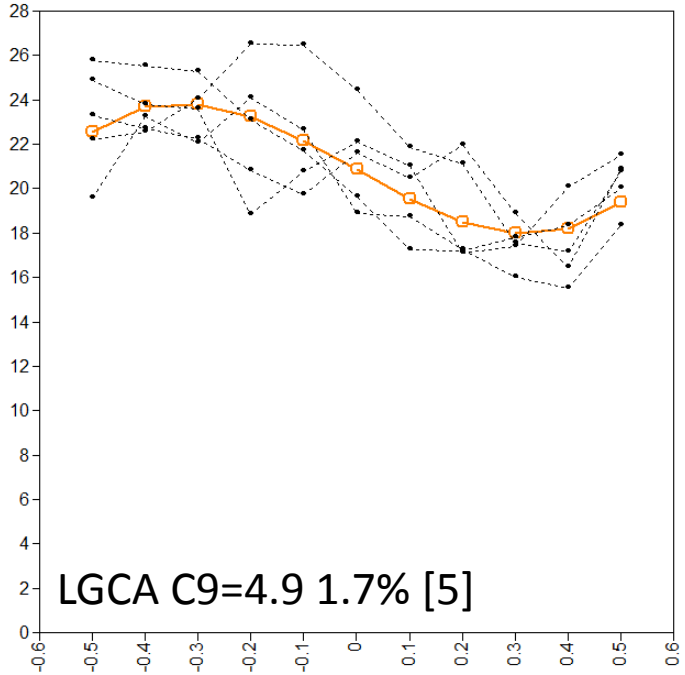
Varying Slope

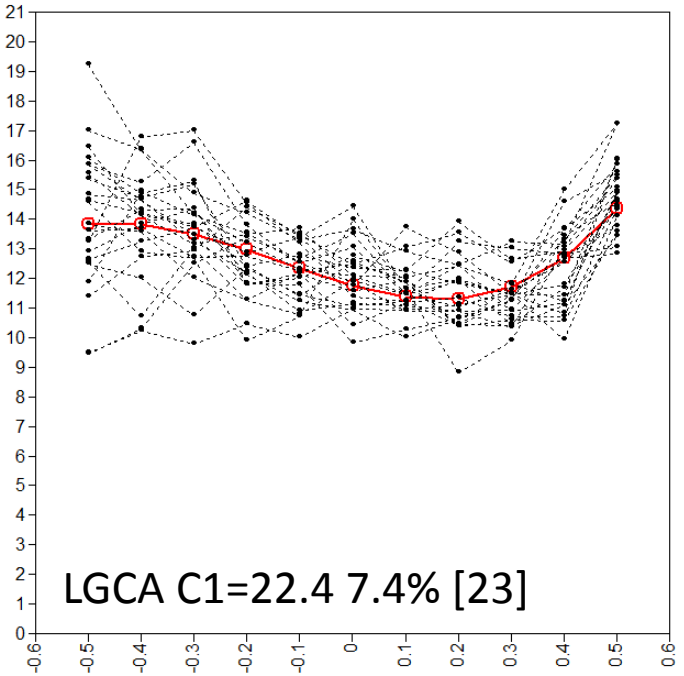


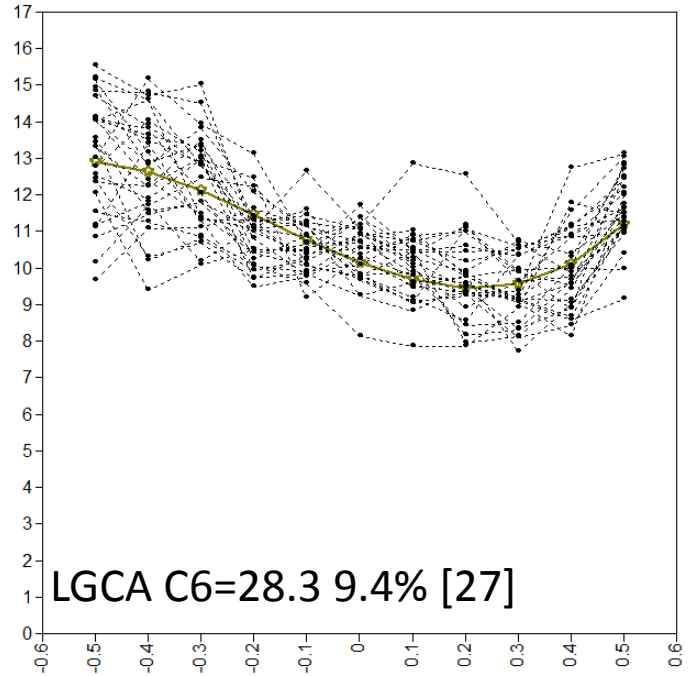
ENGLAND AND WALES - Violence

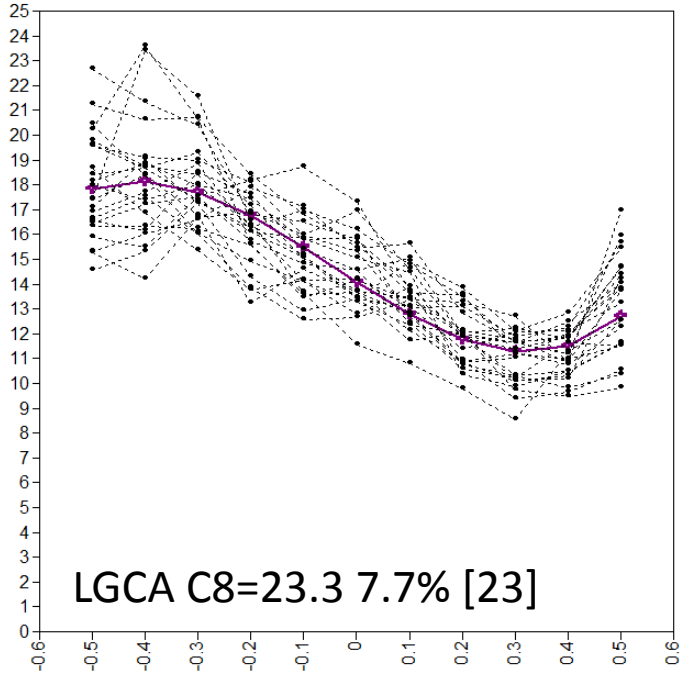
# Estimated Means and Individual Observed Trajectories of Most Likely Class Members

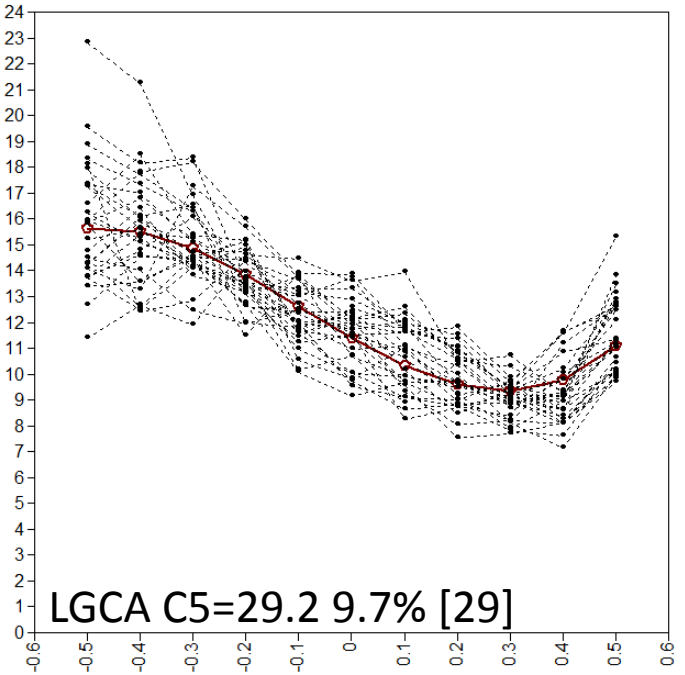
Labels:- Model Type; Class (C); Probability based class membership N and %; [N most likely class membership]

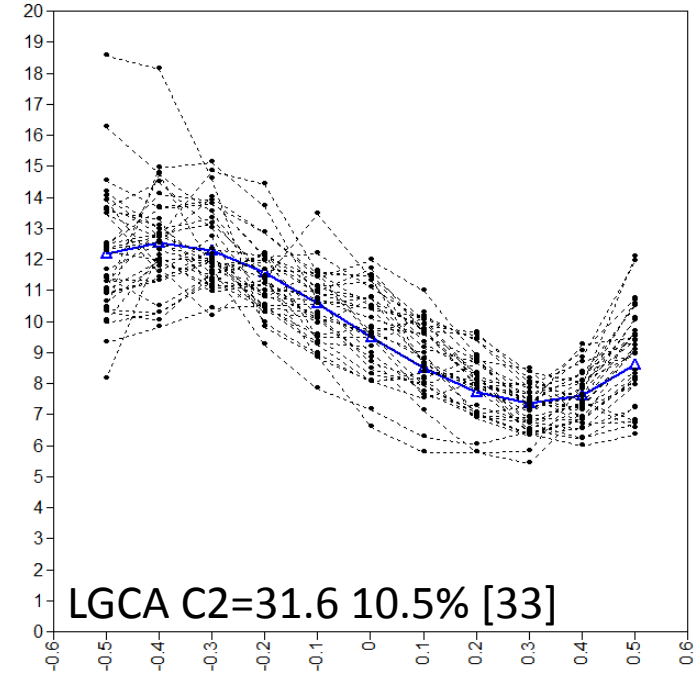

LGCA C9=4.9 1.7% [5]

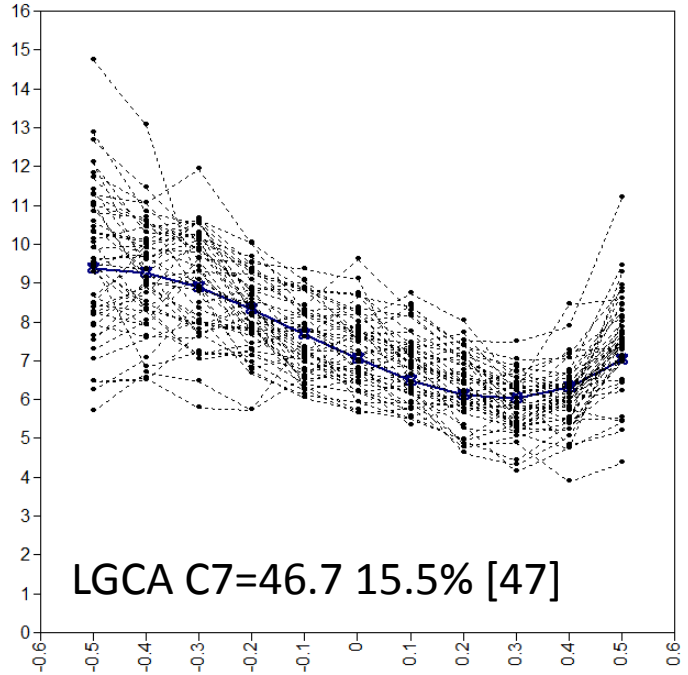

LGCA C1=22.4 7.4% [23]

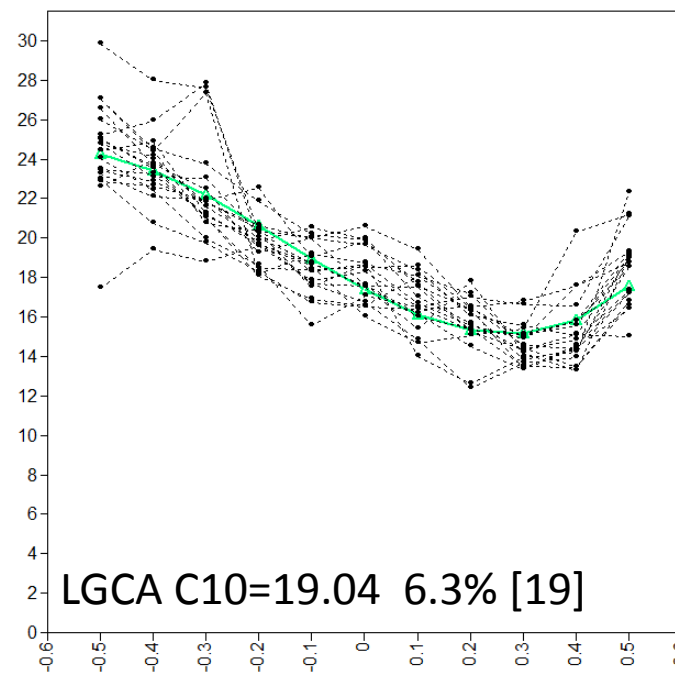

LGCA C6=28.3 9.4% [27]

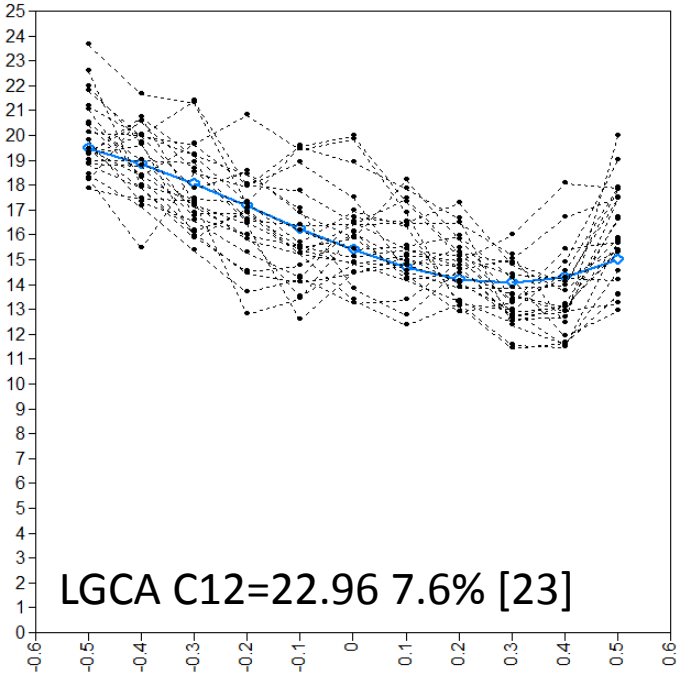

LGCA C8=23.3 7.7% [23]

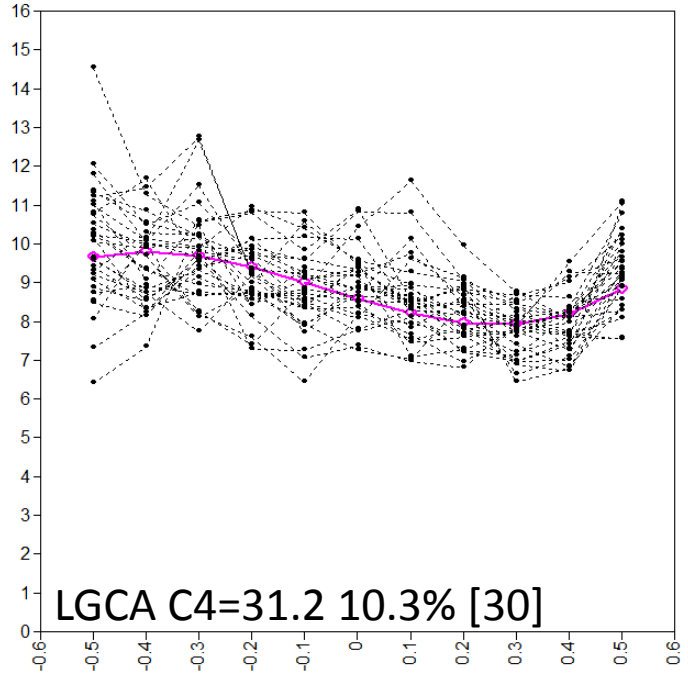

LGCA C5=29.2 9.7% [29]

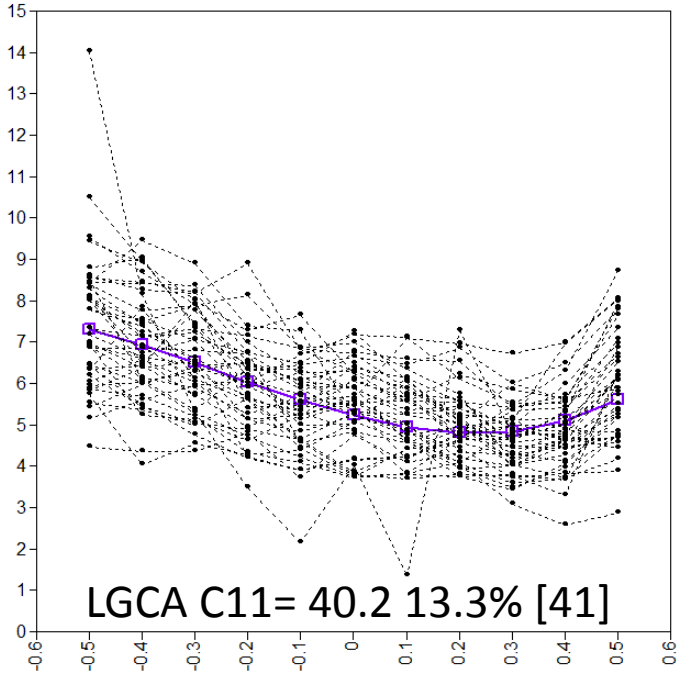

LGCA C2=31.6 10.5% [33]

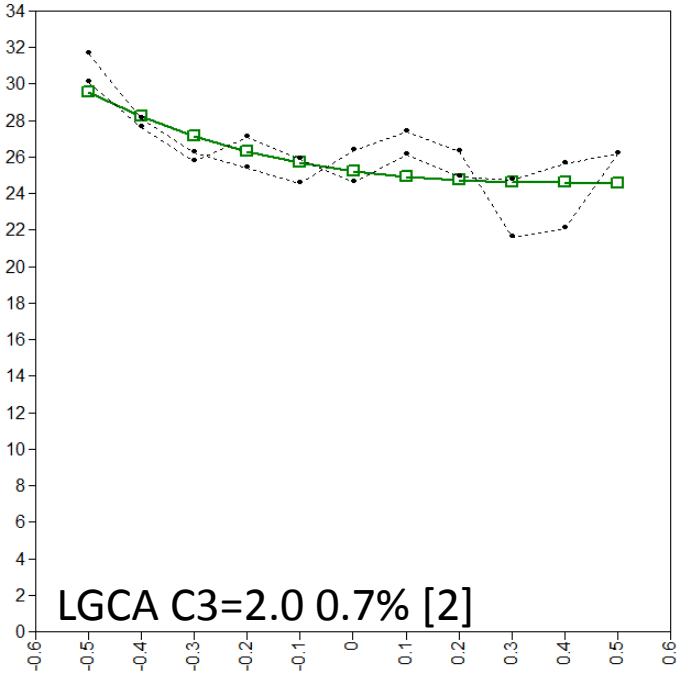

LGCA C7=46.7 15.5% [47]

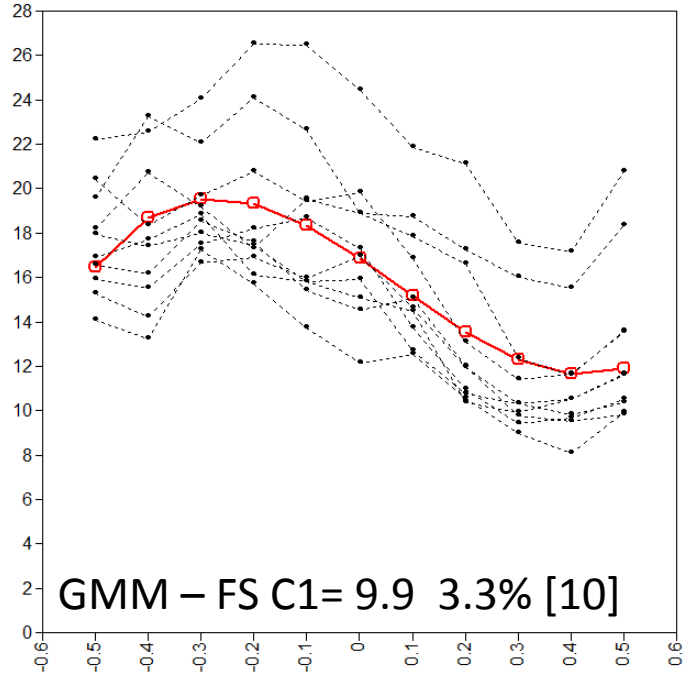

LGCA C10=19.04 6.3% [19]

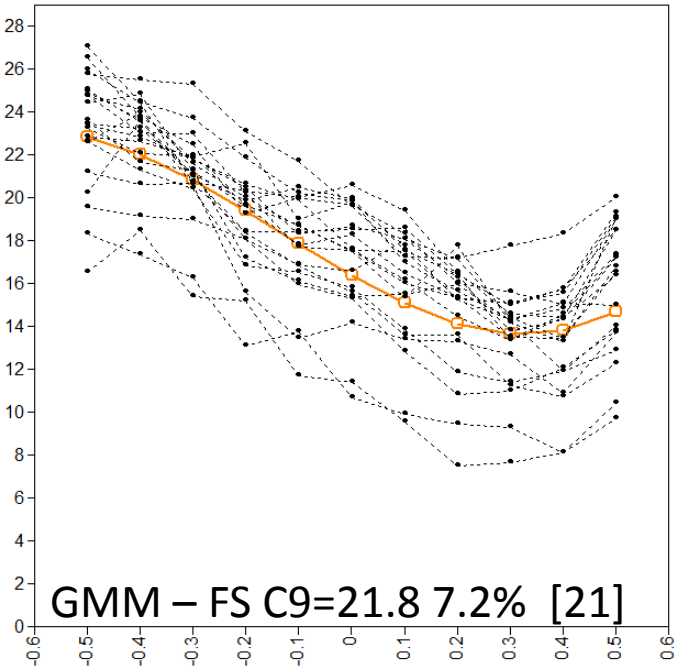

LGCA C12=22.96 7.6% [23]

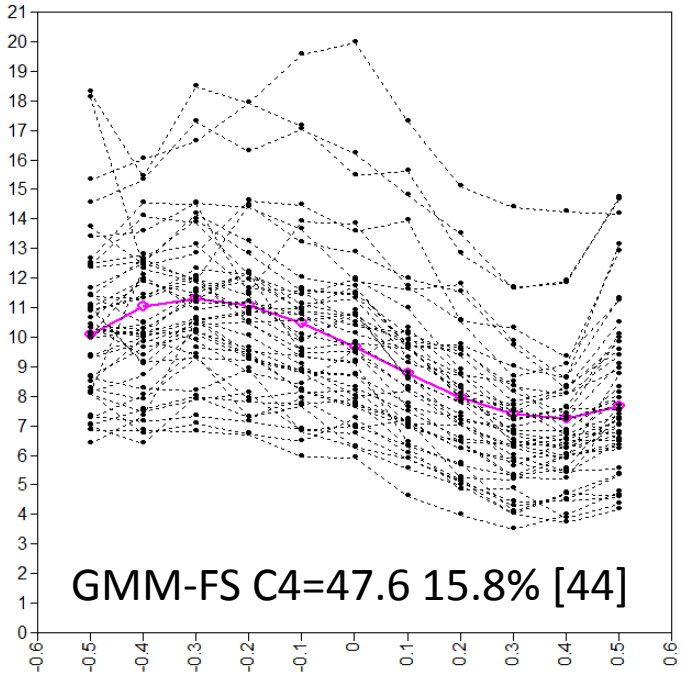

LGCA C4=31.2 10.3% [30]

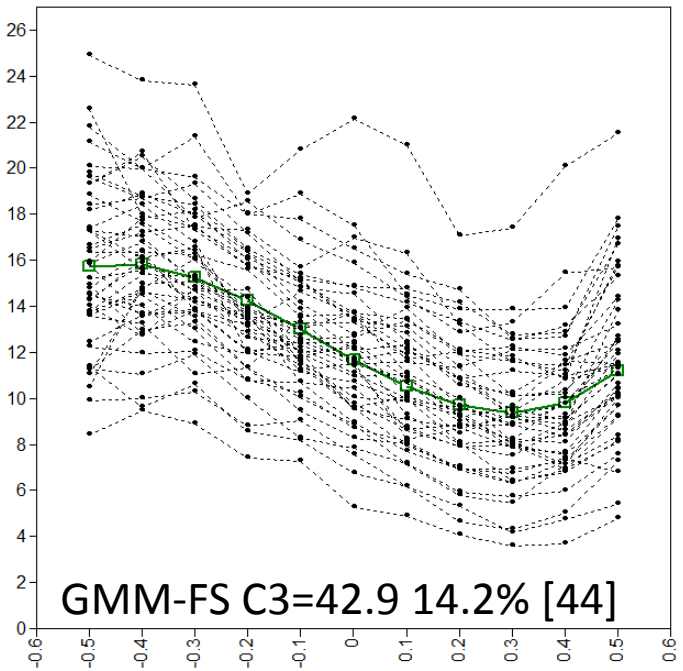

LGCA C11= 40.2 13.3% [41]

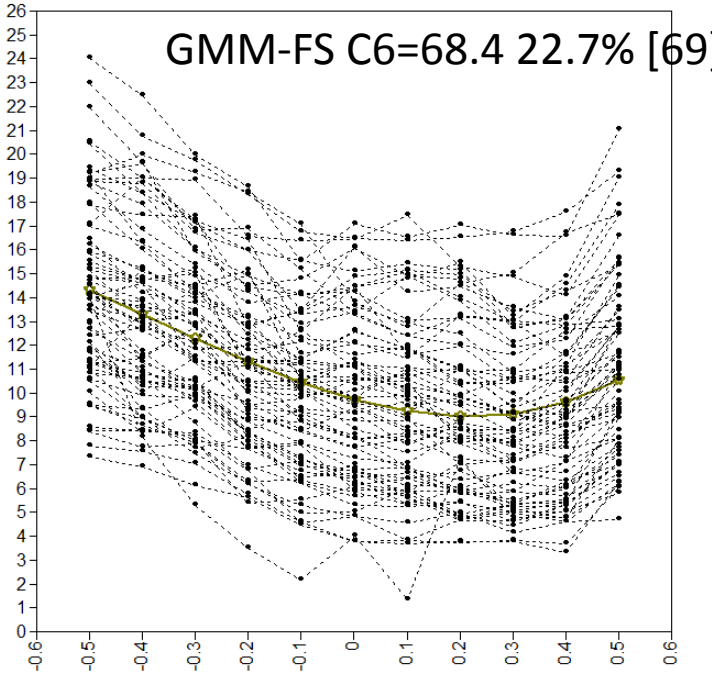

LGCA C3=2.0 0.7% [2]

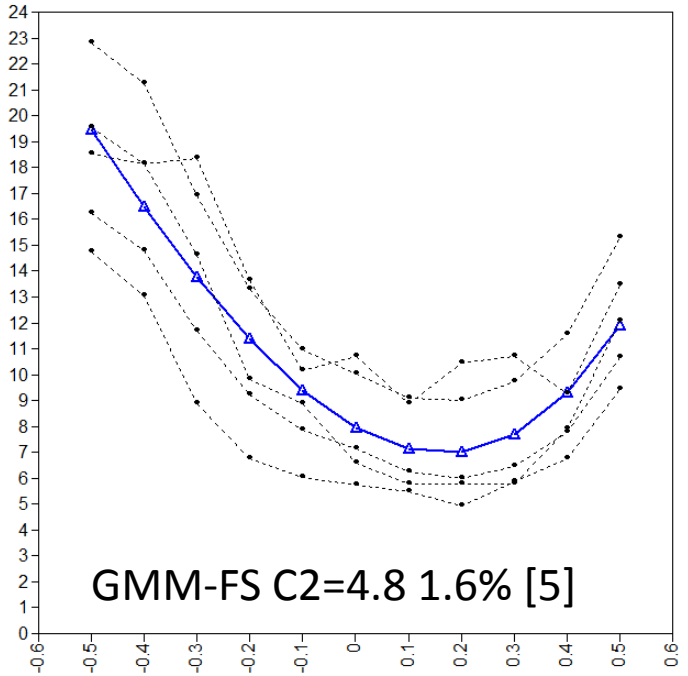

GMM – FS C1= 9.9 3.3% [10]

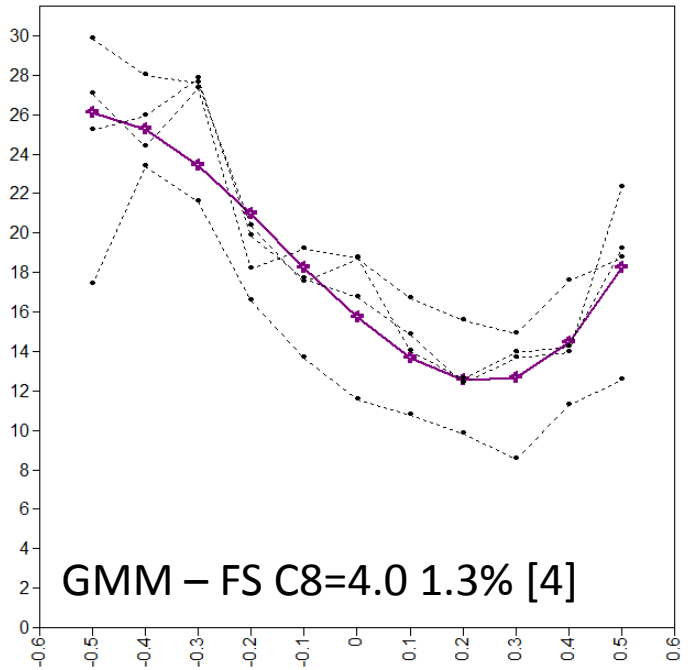

GMM – FS C9=21.8 7.2% [21]

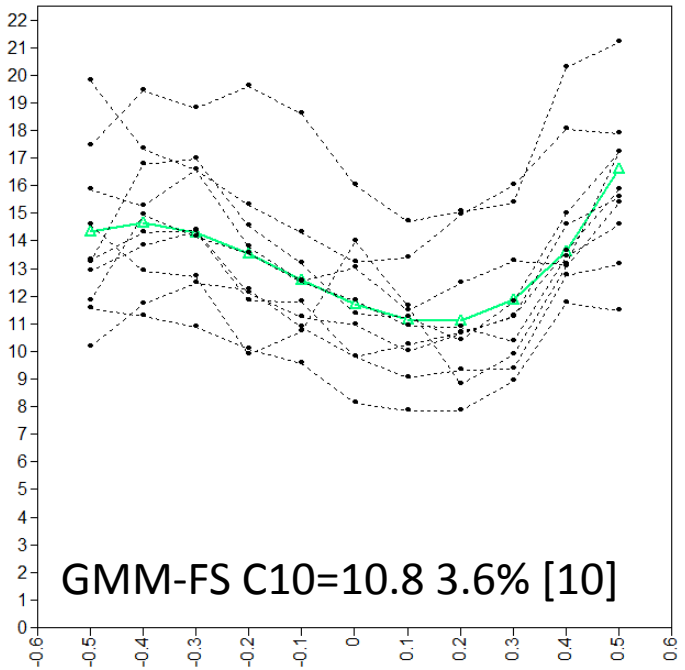

GMM-FS C4=47.6 15.8% [44]

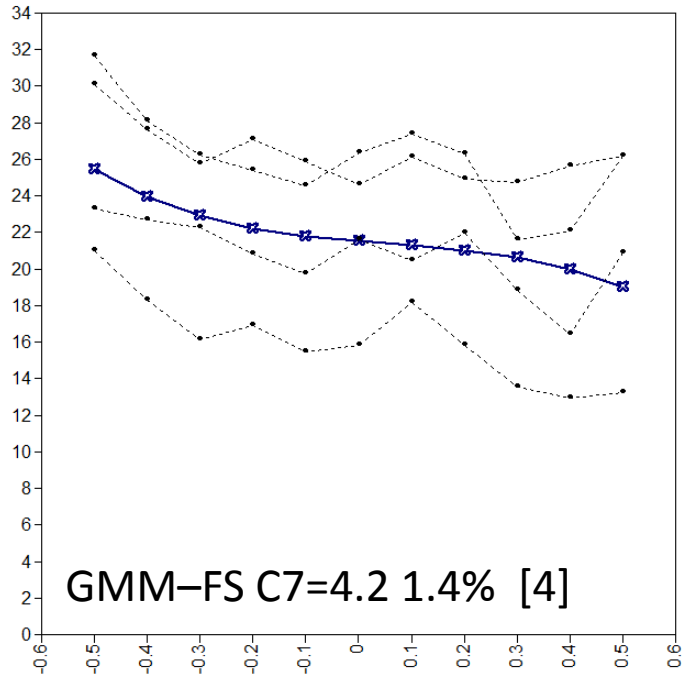

GMM-FS C3=42.9 14.2% [44]

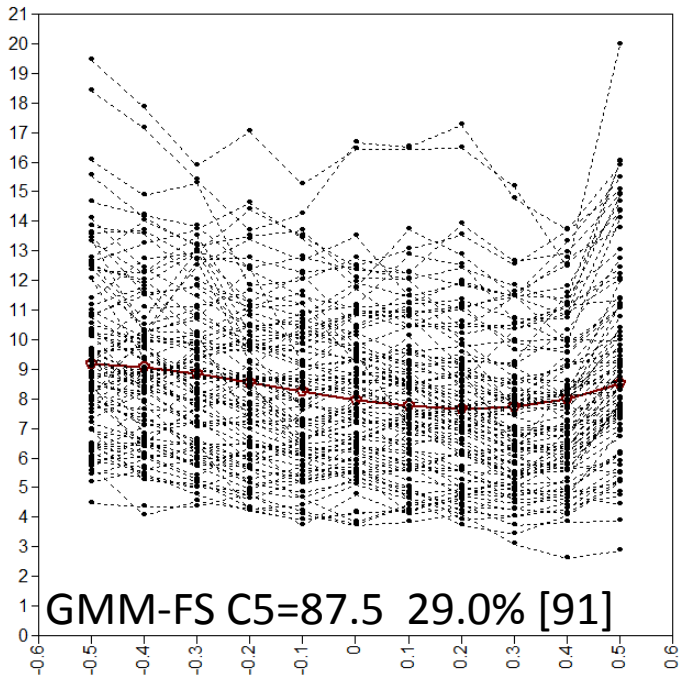

GMM-FS C6=68.4 22.7% [69]

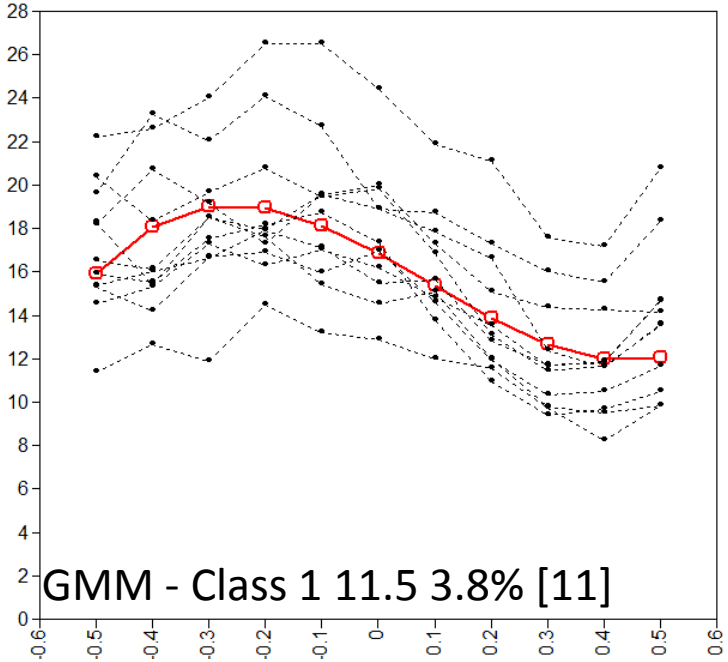


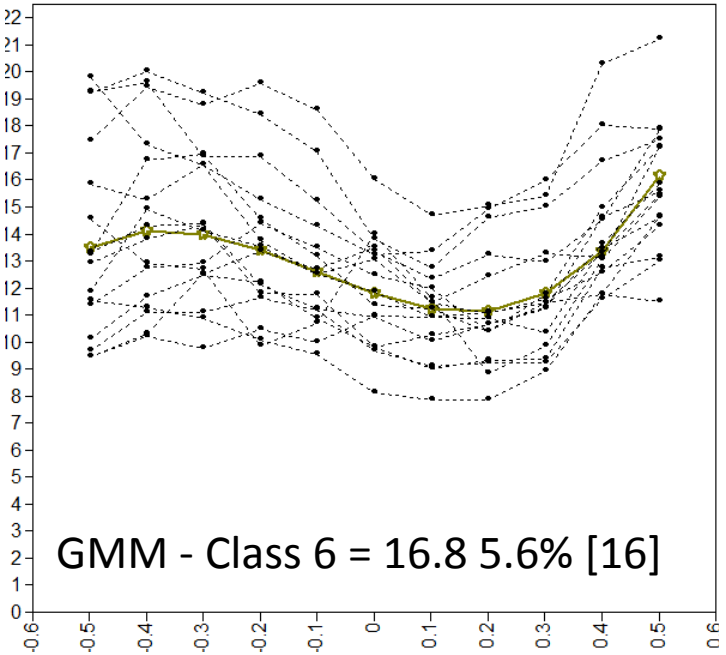



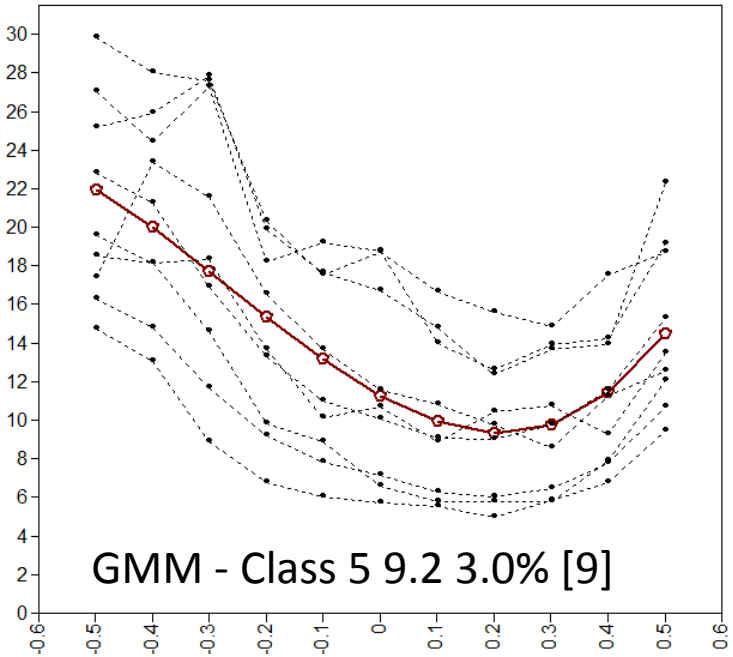



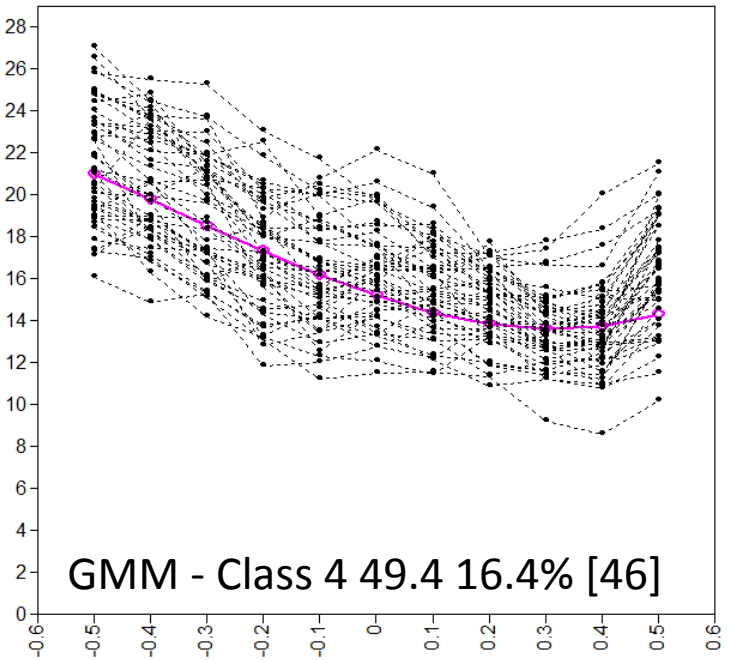



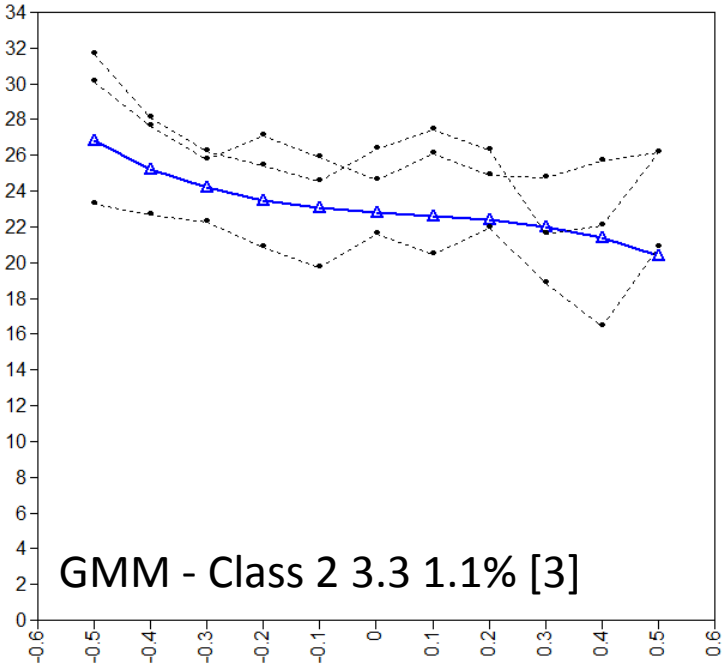



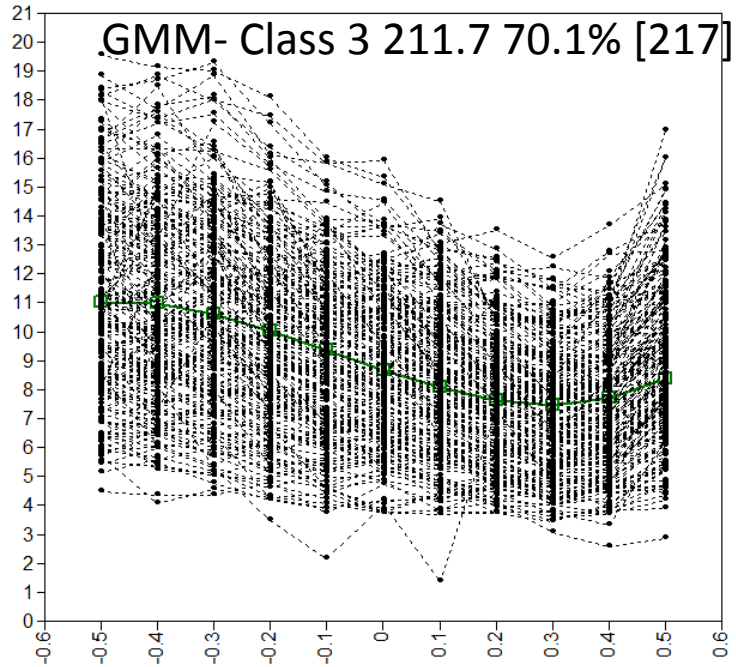

GMM - Class 1 11.5 3.8% [11]


GMM - Class 6 = 16.8 5.6% [16]


GMM - Class 5 9.2 3.0% [9]

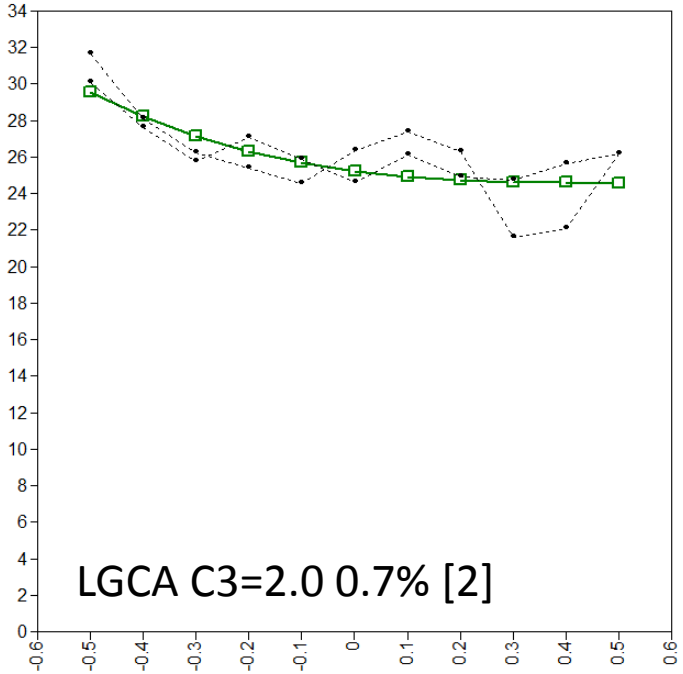

GMM - Class 4 49.4 16.4% [46]

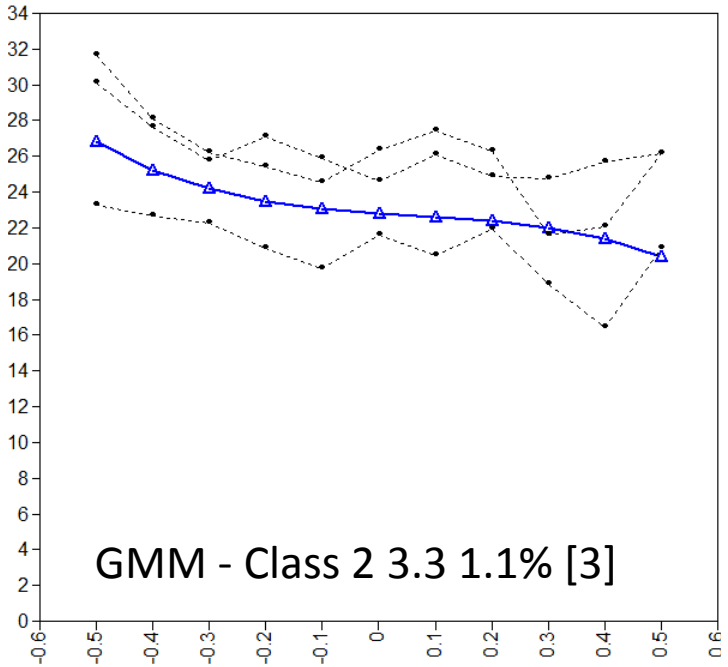

GMM - Class 2 3.3 1.1% [3]

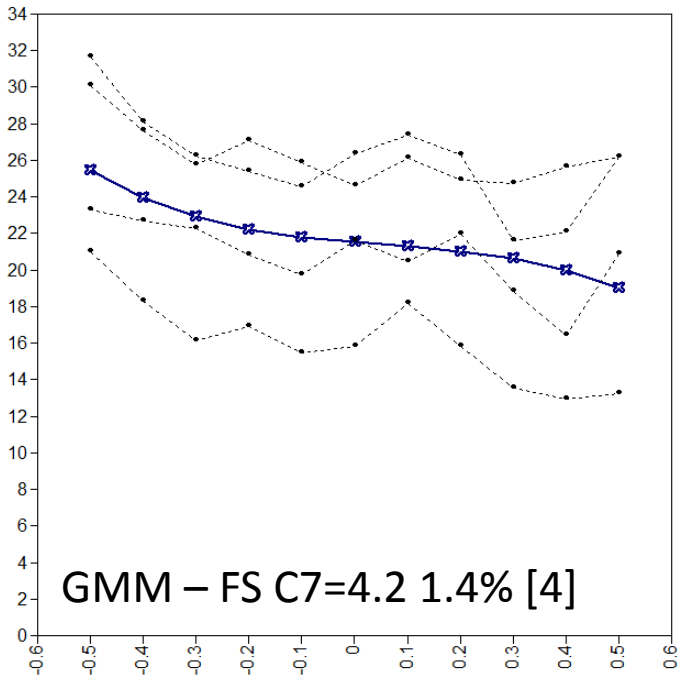

GMM- Class 3 211.7 70.1% [217]

# Similar class membership between models?

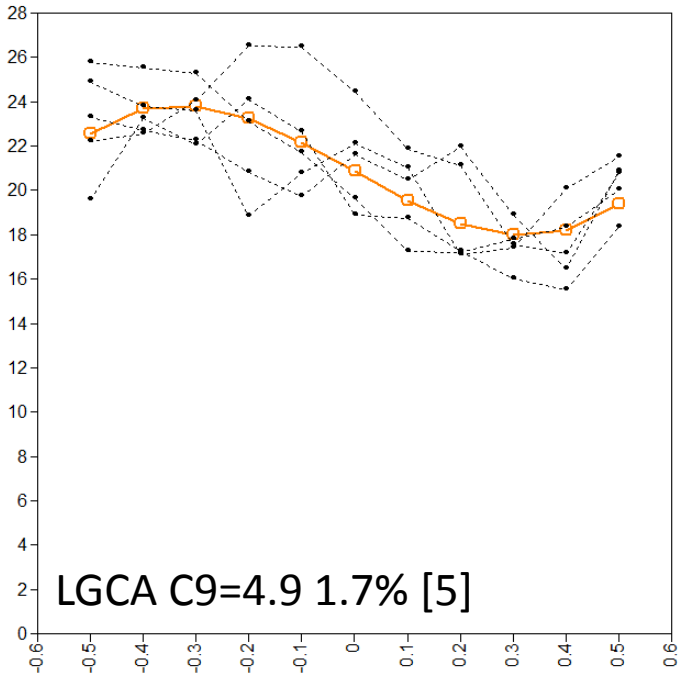
Yes and No…

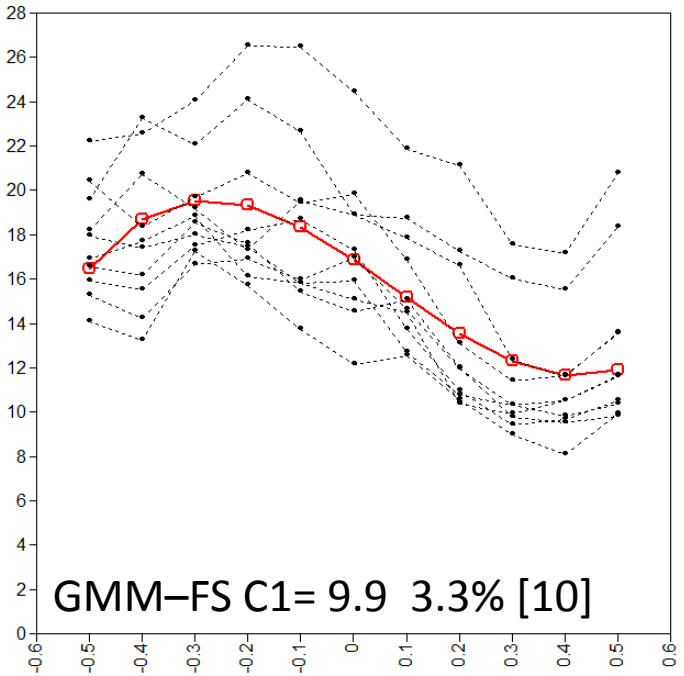


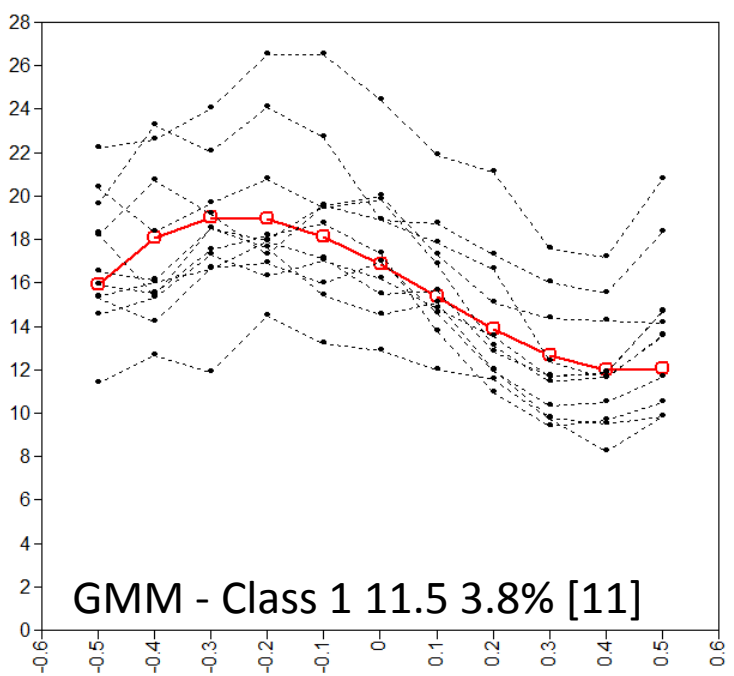




LGCA Class 3, GMM Class 2 and GMM-FS Class 7 have similarities



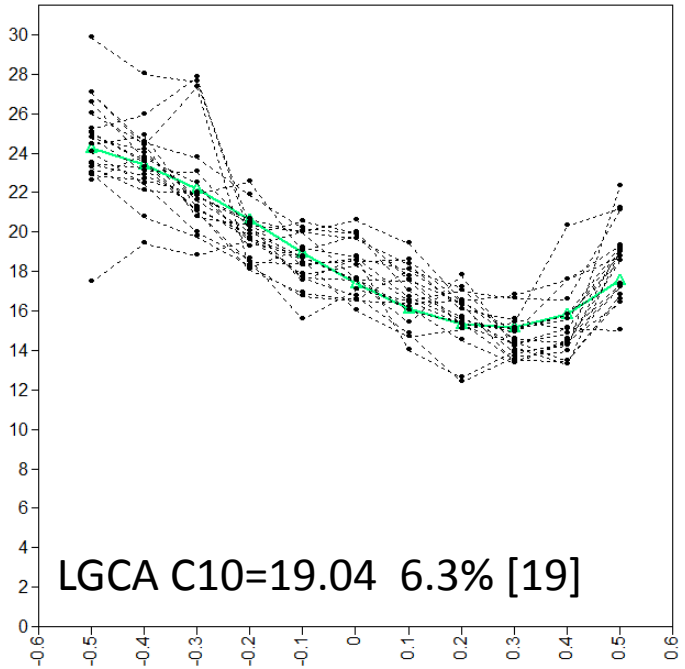



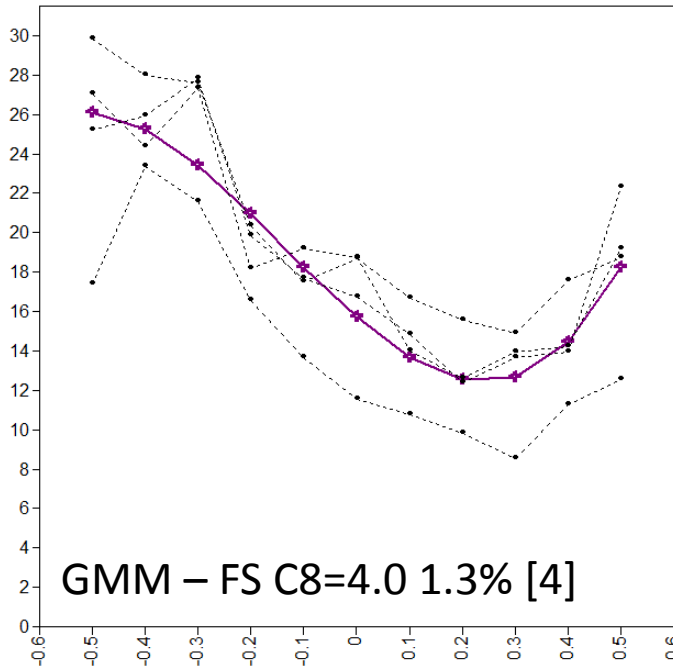

Some overlap only between LGCA Class 9 and other models Class 1

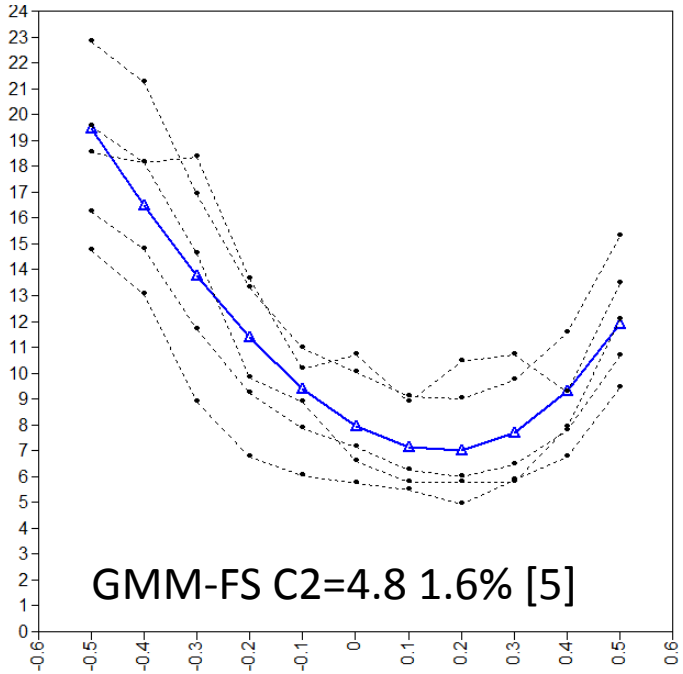
GMM-FS and GMM Class 1 both very similar

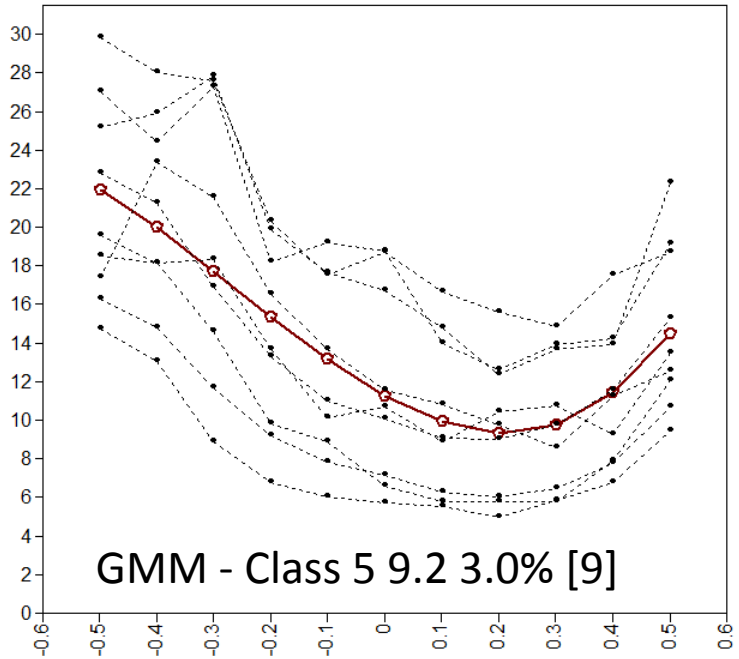
No exact LGCA Comparison C10 appears to have some similarities




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

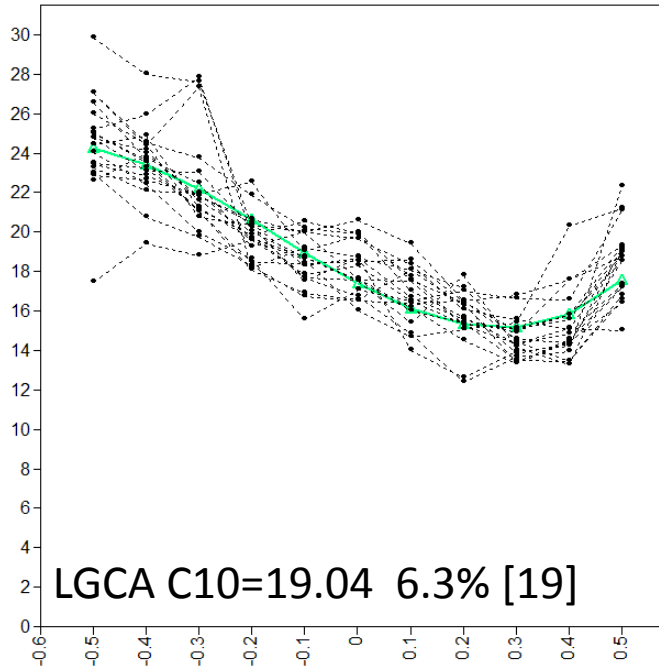


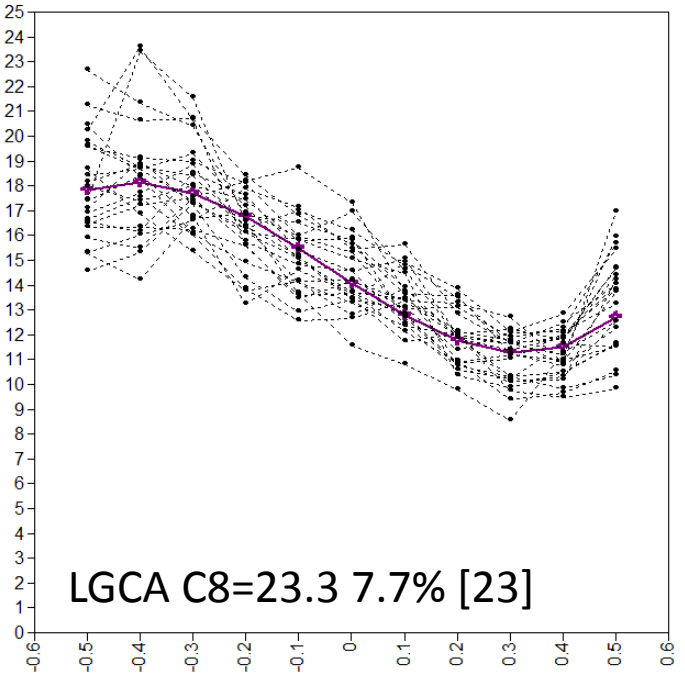



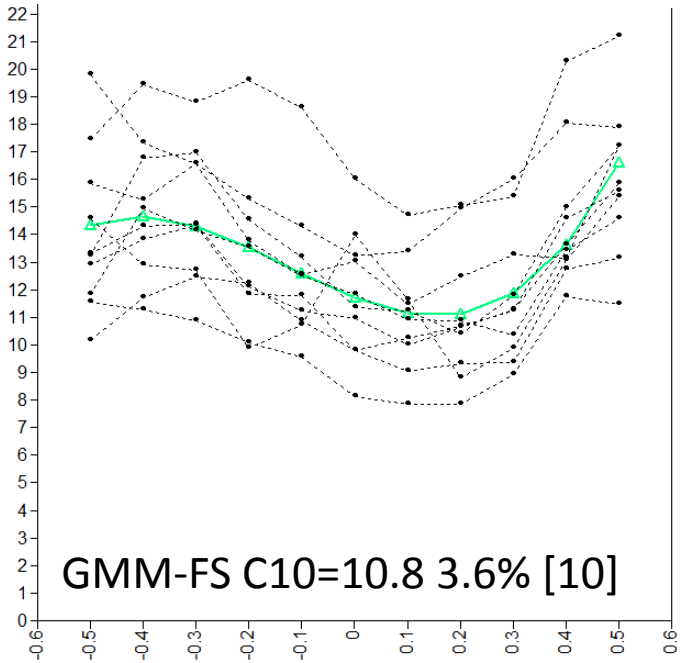



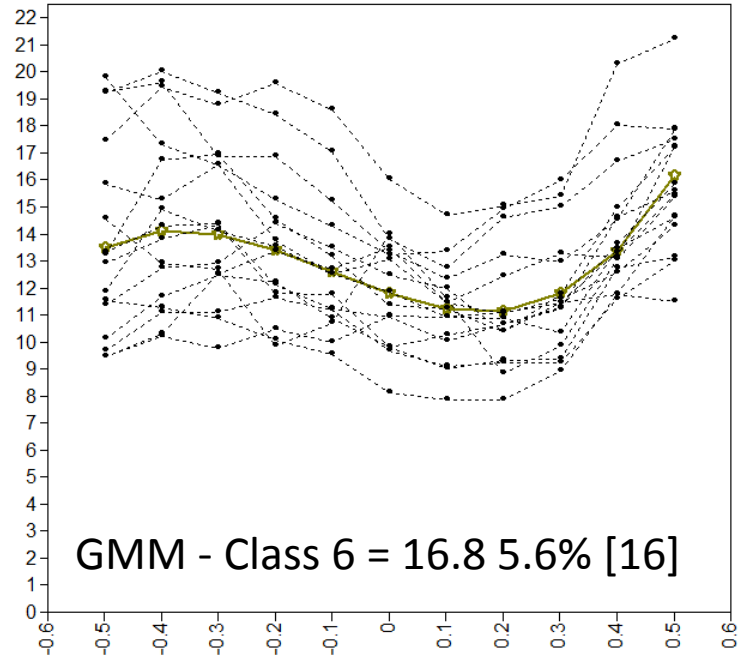
Appears to be some overlap between LGCA C10 and C8 and GMM –FS C10 and GMM – 6.

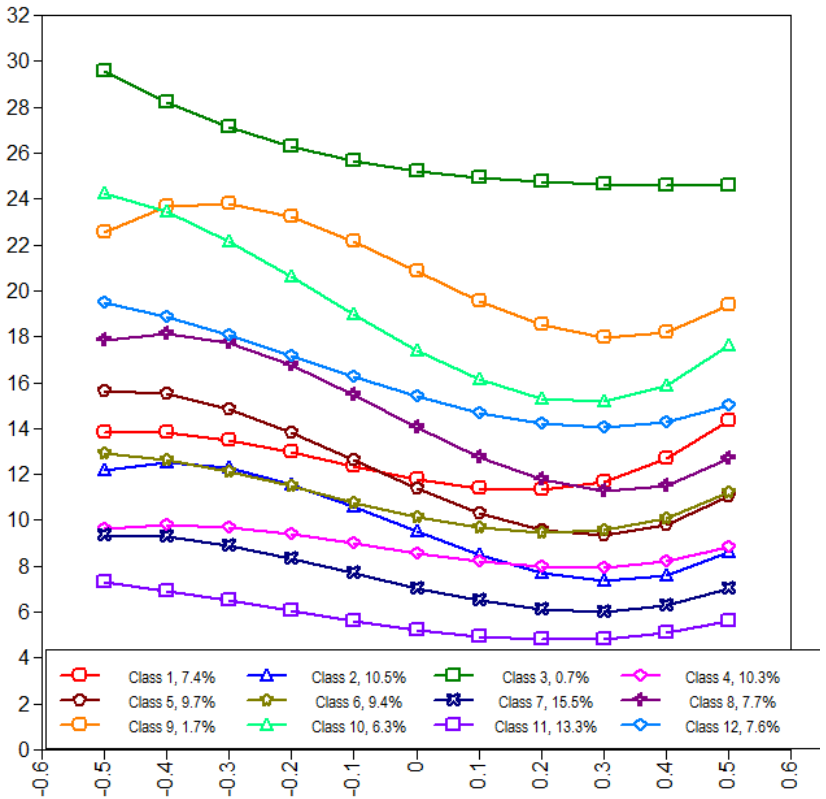
GMM-FS C10 and GMM C6 again appear very similar.



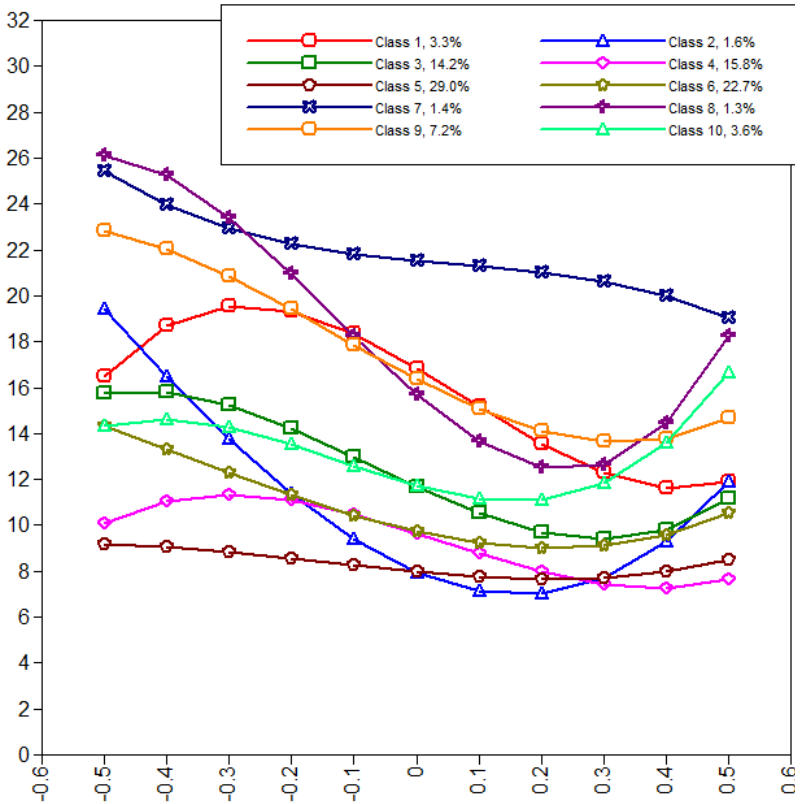



# LGCA

Fixed Intercept

Fixed Slope

12 classes

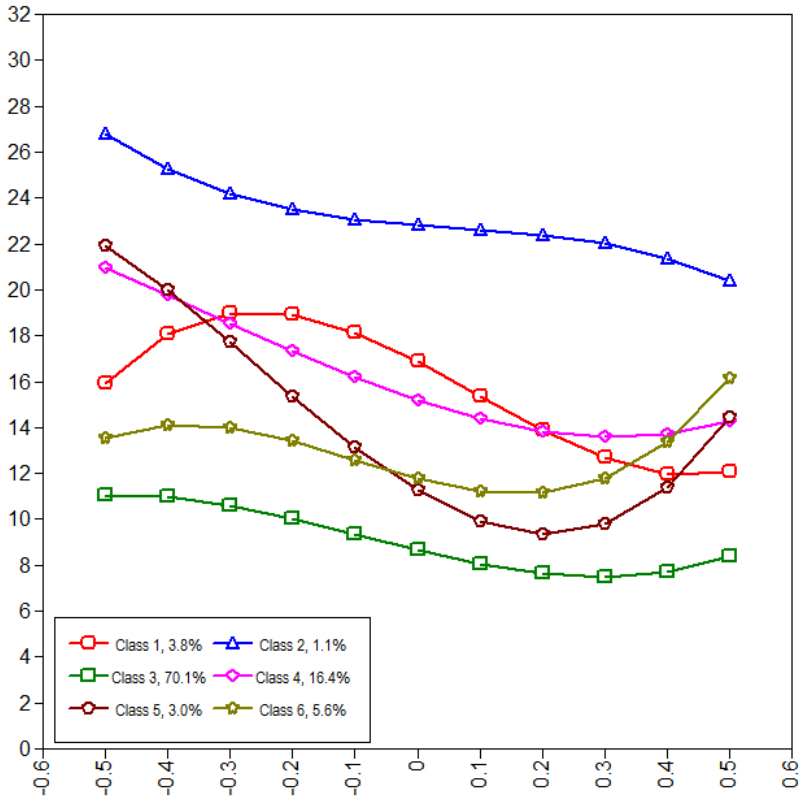



# GMM – FS

Varying Intercept

Fixed Slope

10 (or 12) classes




# GMM

Varying Intercept

Varying Slope

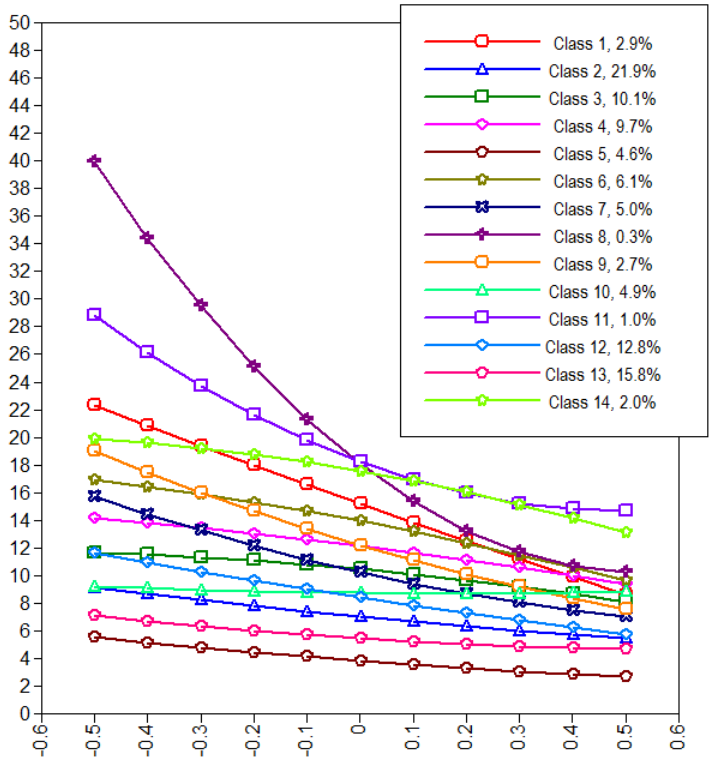
6 (or 7) classes



E&W Violence - 'Best fit' Model Estimates

# Model Comparison
## Burglary

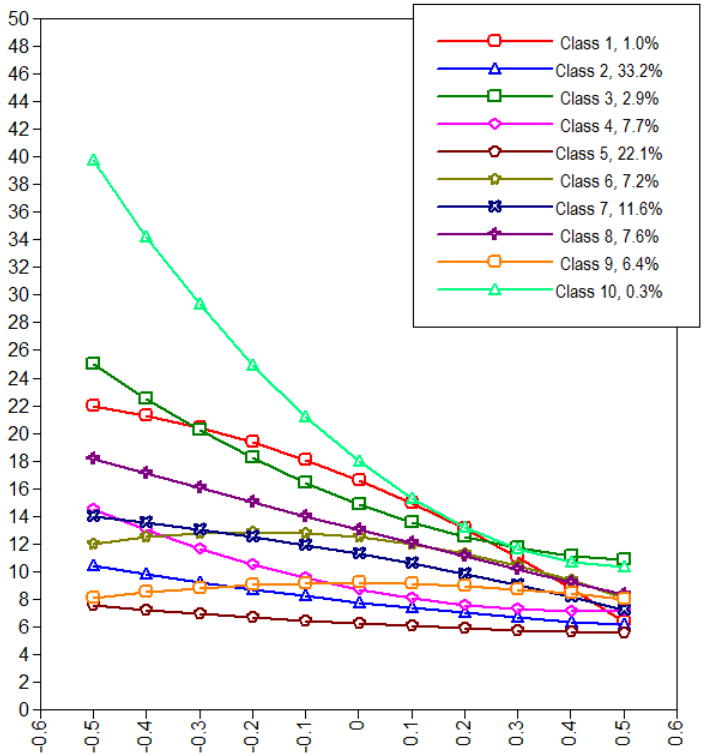
A very brief note…




# LGCA

BIC=12390
ABIC=12180

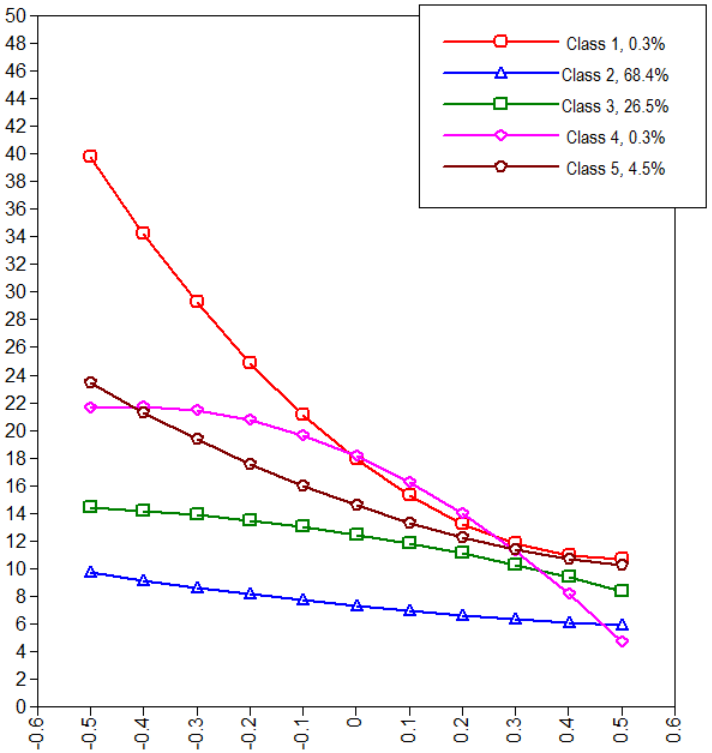
Fixed Intercept
Fixed Slope
14 (or 18?) classes




# GMM – FS

BIC=11855
ABIC=11693

Varying Intercept
Fixed Slope
10 (or 14) classes




# GMM

BIC=11711
ABIC=11597

Varying Intercept
Varying Slope
5 (or 8) classes

# ENGLAND AND WALES - Burglary

# LGM

BIC=11781
ABIC=11718



AQMeN
Applied Quantitative Methods Network

# 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**
**Police Recorded Crime – UK Home Office** Downloaded 07 August 2015.
https://www.gov.uk/government/statistics/police-recorded-crime-open-data-tables
**Resident Population Estimates All People – Office for National Statistics** for the 2011 Local Authority Geography hierarchy. Downloaded 23 December 2015
http://www.neighbourhood.statistics.gov.uk/dissemination/Download1.do
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
MUTHÉN, L.K. and MUTHÉN, B.O. (1998-2012), Mplus User's Guide. Seventh Edition. Muthén and Muthén.

**Key reference**
KREUTER, F., and MUTHÉN, B. (2008), 'Analyzing Criminal Trajectory Profiles: Bridging Multilevel and Group-based Approaches Using Growth Mixture Modeling', *Journal of Quantitative Criminology*, 24: 1-31.

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