

Data, Analytics & Better Decision Making in the NI Public Sector

Dr. Martin Spollen
Chief Investment Officer, SIB &
Visiting Research Fellow, Inst. of Mathematics, QUB

04 May 2017
Chief Executives' Forum, Clare House

Contents

- Data analytics in context
- Role of data analytics to support reform
- Some applications to date
- Future potential and challenges

Data Science

Analytics

Big Data

Statistics

Why?

Sharing data to save time, paper and money

Learning about business performance in real time

Personalising services for a better user experience

Solving problems and predicting the future

Innovating for enhanced productivity and growth

Where data science fits in

User interface & data
visualisation important

Output

| | |
|----------------------|---------------------------------------|
| Static | <i>Interactive</i> |
| Table of Statistics | <i>Decision-Support Models</i> |
| AQs, Historic trends | <i>AI-based Predictive capability</i> |

Data analysis:

Largely descriptive - retrospective
Cross-tabulations; Freq. counts
Summary data; tabular formats

(Big) Data analytics:

Largely predictive - prospective
Model underlying patterns
Operational algorithms

Pre-analysis QA

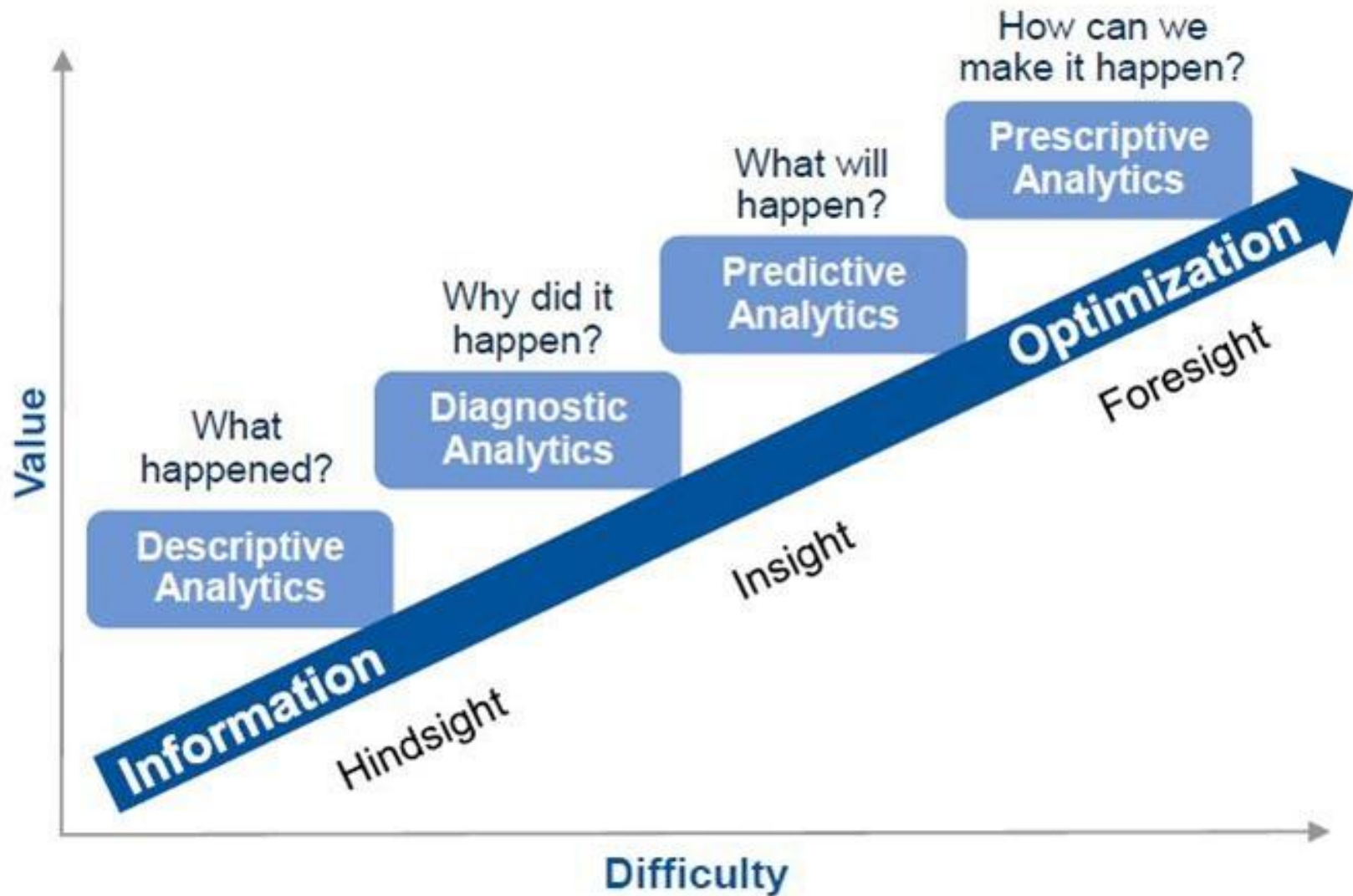
Data cleaning, error detection, data matching, meta labelling

Data collection, archiving & quality issues:

Channels: manual data entry, automation

Period: Continuous, periodic, episodic

Format: .csv, .mat, .xlsx, .sav, native



Strategic

Leadership issues (infrequent)
Long term direction & planning
Complex, unstructured problems
Whole organisation scope

Tactical

Operational

Strategic

Tactical

Senior managerial (C-level) issues
Medium term planning & control
More structured problems
Directorate scope

Operational

Strategic

Tactical

Operational

Junior managerial; operational staff
Short term; day-to day planning & control
Structured & routine problems (frequent)
Narrow functional scope

Economy

Efficiency

Effectiveness

Equity

Careful use of resources to
minimise effort, time and
cost

Economy **Efficiency** **Effectiveness** **Equity**

Obtaining maximum benefit
from a given level of input

Economy Efficiency Effectiveness Equity

Delivering a successful outcome and
meeting objectives as fully as possible

Economy Efficiency Effectiveness Equity

Delivering services and using
resources in a way that is fair to all

Economy Efficiency Effectiveness Equity

Strategic

Tactical

Operational

| | | | |
|--|-------------------------|----------------------------|------------------------------|
| | Library Services (DCAL) | School Planning (DENI) | Resource Allocation (DHSSPS) |
| | | Predictive Policing (PSNI) | Ambulance Response (DHSSPS) |
| | | | |

Strategic Infra Planning Models (SIPMs)

SIPMs are computerised models that help departments to plan the **right infrastructure** in the **right place** to meet needs and to ensure that individual capital schemes **improve overall productivity**.

Data collection & Mapping

- Mapping demands
- Mapping supply infra.

Demand Pattern

Financial Model

Infra. optimisation

Investment Return

Decision Support Tool

Data collection & Mapping

Demand Pattern

Financial Model

Infra. optimisation

Investment Return

Decision Support Tool

- Who goes where? Why?
- Choice analysis
- Influence of distance
- Influence of other factors
- Characteristics of users
- Attributes of services

Data collection & Mapping

Demand Pattern

Financial Model

Infra. optimisation

Investment Return

Decision Support Tool

- Revenue costs analysis
- Unit cost v size of facility
- Econ. of Scale curve

Data collection & Mapping

Demand Pattern

Financial Model

Infra. optimisation

Investment Return

Decision Support Tool

- Reconfiguration options
- Lowest recurrent costs
- but recognising CONSTRAINTS

Data collection & Mapping

Demand Pattern

Financial Model

Infra. optimisation

Investment Return

Decision Support Tool

- Cost to move (C)
- Saving after move (S)
- $ROI = S/C$
- Invest rule $ROI > x\%$

Systems that help to forecast rapidly the expected impacts of public services **opening/closing, relocation** or **transformation** on:

- demand & budget impacts
- viability and capacity issues
- relative accessibility in affected communities
- external costs (traffic)

...and gain a **whole-system perspective** on longer run issues such as:

- impact of local demographic changes over time
- potential opportunities for mergers and co-location
- support cross-sectoral Area-Based Planning
- other potential changes (e.g. academic selection)

...and **speed up** business case development with ready access to relevant information:

- official data from DENI, NISRA etc. mustered in one place
- user-friendly interface
- evaluate alternative policy/investment scenarios in real time
- reduce reliance on external consultants



DENI & Education Authority

- Inform needs assessment for new/merged schools
- Pilot SIPM for DENI - launched in Executive's ISNI 2011-21
- Statistical analysis of school enrolment records (04-09)
- Strategic level Decision Support System – “What if...?”

Post-Primary Schools Mapping & Profiling Tool

Version 1.5 May 2012

Use the selector at the top left of each map screen to select the required school and Area Learning Community (ALC)

[Map of school catchments](#)

[Map of Area Learning Communities](#)

Notes:

- (1) Based on anonymised data from the Annual Schools Census
- (2) Data on catchments based on Year 8 admissions from 2004-05 to 2008-09. Data for more recent can be included if required.
- (3) Mileages calculated based on the actual road distances (not crow-fly) from OA centroids to school gates.
- (4) Small area ALC catchments generated based on ALC designation of the nearest school.
- (5) Counts less than 3 are suppressed on maps
- (6) Admission & Enrolment caps, info to be supplied by DE for years after 2008-09

Feedback and queries should be sent to:

Martin Spollen
Strategic Investment Board
07966 982363

martin.spollen@sibs.org

Grammar, Voluntary
Other managed
Girls

BELB

Trend in Enrolments & Enrolment Number (Yrs 8-15)



Trend in First Pref. Applications (FPA), Admissions to Year 8 and Admission Number

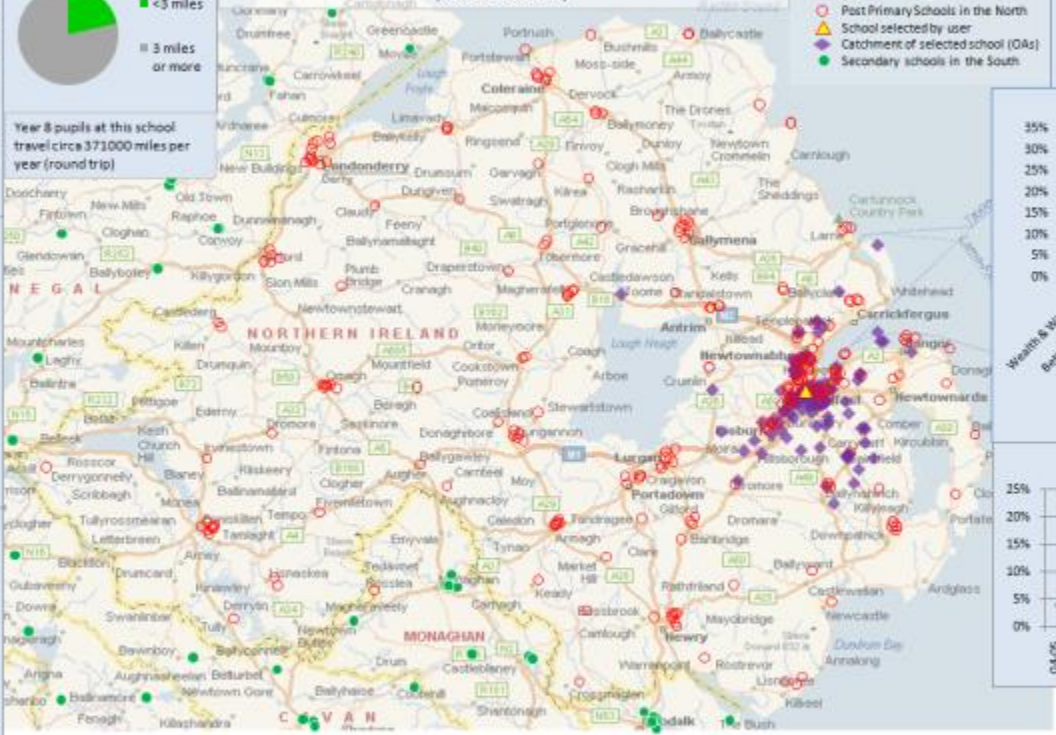


% 5 GCSEs at A*-C



Year 8 pupils at this school travel circa 371000 miles per year (round trip)

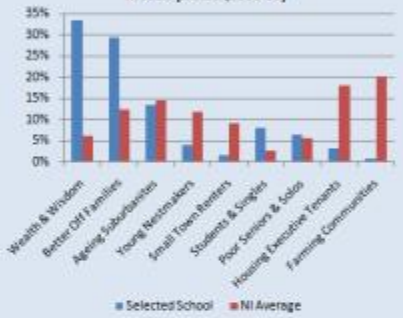
Map of Year 8 admissions (2004-05 to 2008-09)



KEY:
 ○ Post-Primary Schools in the North
 ▲ School selected by user
 ◆ Catchment of selected school (OAs)
 ● Secondary schools in the South



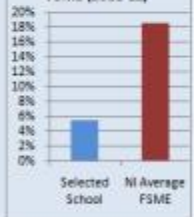
Mosaic profile (2008-09)



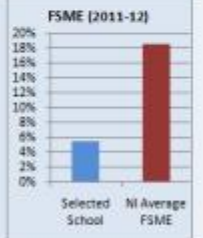
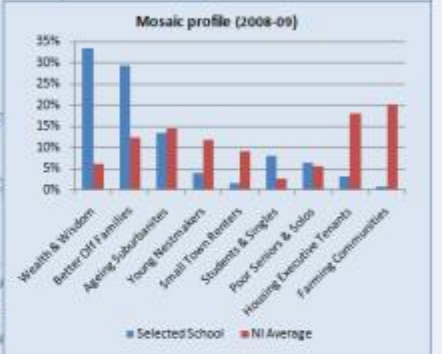
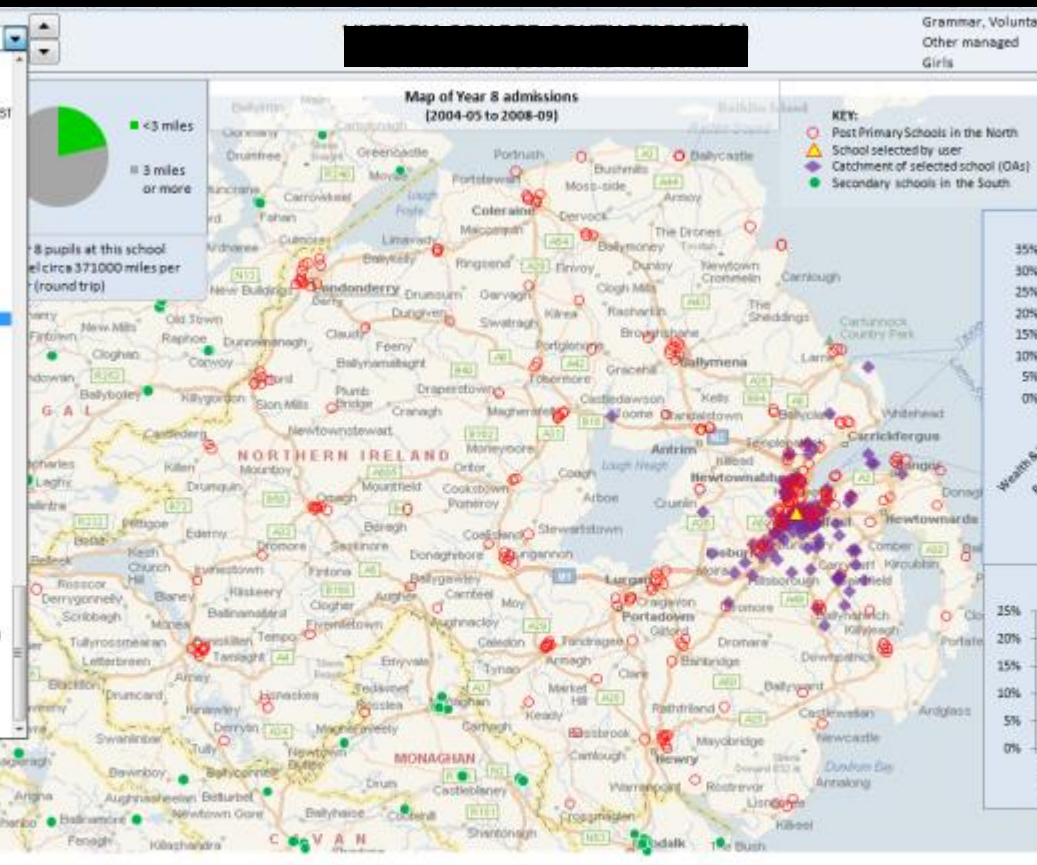
SEN (1-5)



FSME (2011-12)



- 1
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10
- 11
- 12
- 13
- 14
- 15
- 16
- 17
- 18
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26
- 27
- 28
- 29
- 30
- 31
- 32
- 33
- 34
- 35
- 36
- 37
- 38
- 39
- 40
- 41
- 42
- 43
- 44
- 45
- 46
- 47
- 48
- 49



Secondary, Catholic Maintained
RC managed
Co-educational
NEFLB

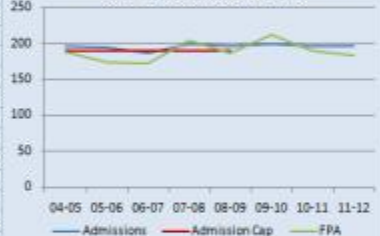


Trend in Enrolments & Enrolment Number (Yrs 8-15)

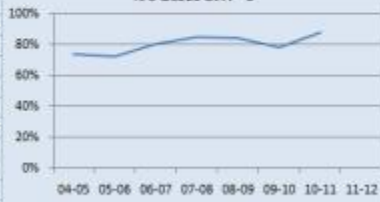


Year 8 pupils at this school travel circa 601000 miles per year (round trip)

Trend in First Pref. Applications (FPA), Admissions to Year 8 and Admission Number



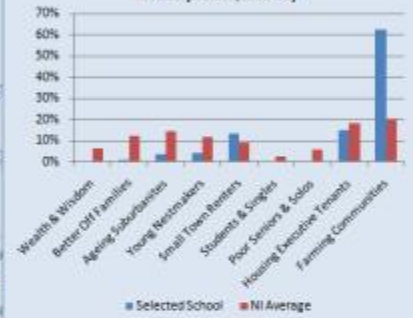
% 5 GCSEs at A*-C



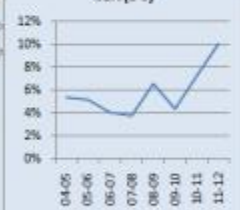
Map of Year 8 admissions (2004-05 to 2008-09)



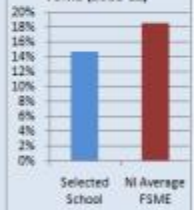
Mosaic profile (2008-09)



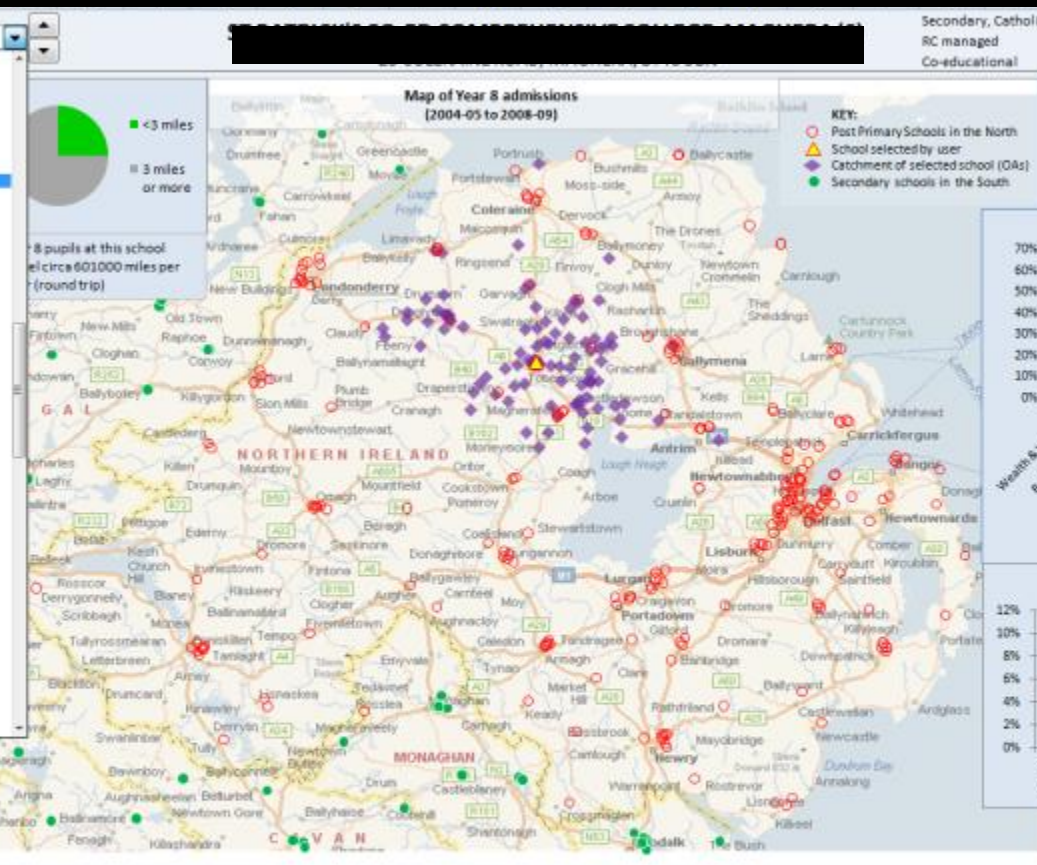
SEN (1-5)



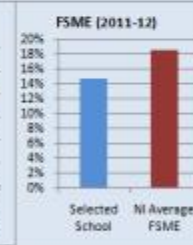
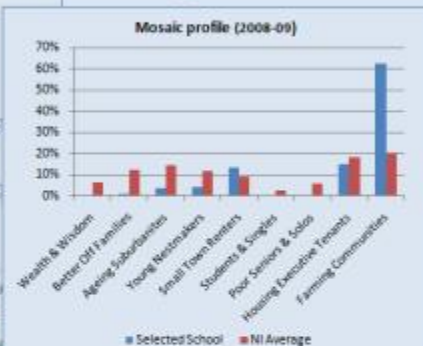
FSME (2011-12)



| | |
|----|---|
| 1 | |
| 2 | |
| 3 | LINAVADY HIGH SCHOOL (S) |
| 4 | LISMORE COMPREHENSIVE SCHOOL, CRAIGAVON (S) |
| 5 | LISNAGARVEY HIGH SCHOOL, LISBURN (S) |
| 6 | LISNASHARRAGH HIGH SCHOOL, EAST BELFAST (S) |
| 7 | LISNASKEA HIGH SCHOOL (S) |
| 8 | LISNALL COLLEGE, L DERRY (S) |
| 9 | LITTLE FLOWER GIRLS' SCHOOL, NORTH BELFAST (S) |
| 10 | LORETO COLLEGE, COLERAINE (S) |
| 11 | LORETO GRAMMAR SCHOOL, OMAGH (S) |
| 12 | LURGAN COLLEGE (S) |
| 13 | LURGAN JUNIOR HIGH (S) |
| 14 | MAGHERAFELT HIGH SCHOOL (S) |
| 15 | MALONE INTEGRATED COLLEGE, SOUTH BELFAST (S) |
| 16 | MARKETHILL HIGH SCHOOL (S) |
| 17 | MASSEERINE COMMUNITY COLLEGE, ANTRIM (S) |
| 18 | METHODIST COLLEGE, SOUTH BELFAST (S) |
| 19 | MONKSTOWN COMMUNITY SCHOOL, NEWTOWNABBEEY (S) |
| 20 | MOUNT GILBERT COMMUNITY COLLEGE, WEST BELFAST (S) |
| 21 | MOUNT LOURDES GRAMMAR SCHOOL, ENNISKENNEN (S) |
| 22 | MOVILLA HIGH SCHOOL, NEWTONWARDS (S) |
| 23 | NENDRUM COLLEGE, COMBER (S) |
| 24 | NEW BRIDGE INTEGRATED COLLEGE, LOUGHBRICKLAND (S) |
| 25 | NEWRY HIGH SCHOOL (S) |
| 26 | NEWTOWNABBEEY COMMUNITY HIGH SCHOOL (S) |
| 27 | NEWTOWNABBEEY HIGH SCHOOL, SOUTH BELFAST (S) |
| 28 | NEWTOWNHAMILTON HIGH SCHOOL (S) |
| 29 | NORTH COAST INTEGRATED COLLEGE, COLERAINE (S) |
| 30 | OMAGH GROVE INTEGRATED COLLEGE, L DERRY (S) |
| 31 | OMAGH ACADEMY (S) |
| 32 | OMAGH HIGH SCHOOL (S) |
| 33 | ORANGEFIELD HIGH SCHOOL, EAST BELFAST (S) |
| 34 | OUR LADY AND ST PATRICK'S COLLEGE, EAST BELFAST (S) |
| 35 | OUR LADY OF LOURDES HIGH SCHOOL, BALLYMONEY (S) |
| 36 | OUR LADY OF MERCY GIRLS' SCHOOL, NORTH BELFAST (S) |
| 37 | OUR LADY OF MERCY HIGH SCHOOL, STRABANE (S) |
| 38 | OUR LADY'S GRAMMAR SCHOOL, NEWRY (S) |
| 39 | PARKHALL HIGH SCHOOL, ANTRIM (S) |
| 40 | PARKHALL INTEGRATED COLLEGE, ANTRIM (S) |
| 41 | PORTADOWN COLLEGE (S) |
| 42 | PORTADOWN COLLEGE (S) |
| 43 | PRIORY COLLEGE, HOLYWOOD (S) |
| 44 | RAMEY ENDOVED SCHOOL, MAGHERAFELT (S) |
| 45 | RATHFRILAND HIGH SCHOOL (S) |
| 46 | RATHMORE GRAMMAR SCHOOL, FINAGHY (S) |
| 47 | REGENT HOUSE SCHOOL, NEWTONWARDS (S) |
| 48 | SACRED HEART COLLEGE, OMAGH (S) |
| 49 | SACRED HEART GRAMMAR SCHOOL, NEWRY (S) |
| 50 | SAINTFIELD HIGH SCHOOL (S) |



Secondary, Catholic Maintained
RC managed
Co-educational
NEELB



Grammar, Voluntary
RC managed
Co-educational

WELB

Trend in Enrolments & Enrolment Number (Yrs8-15)



Trend in First Pref. Applications (FPA), Admissions to Year 8 and Admission Number



% 5 GCSEs at A*-C

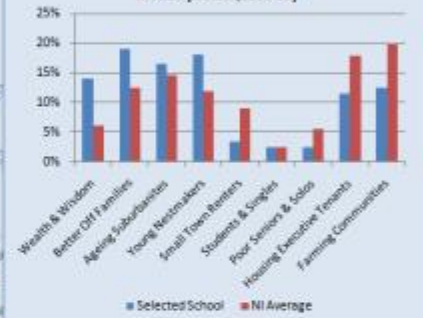


Year 8 pupils at this school travel circa 250000 miles per year (round trip)

Map of Year 8 admissions (2004-05 to 2008-09)



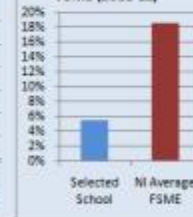
Mosaic profile (2008-09)



SEN (1-5)

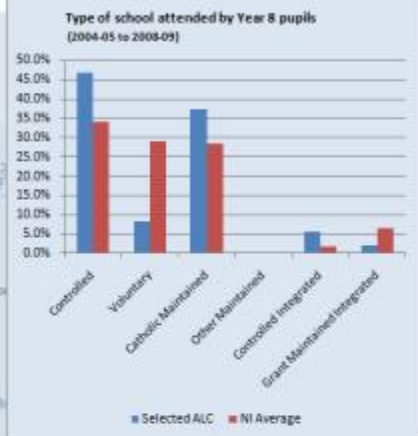
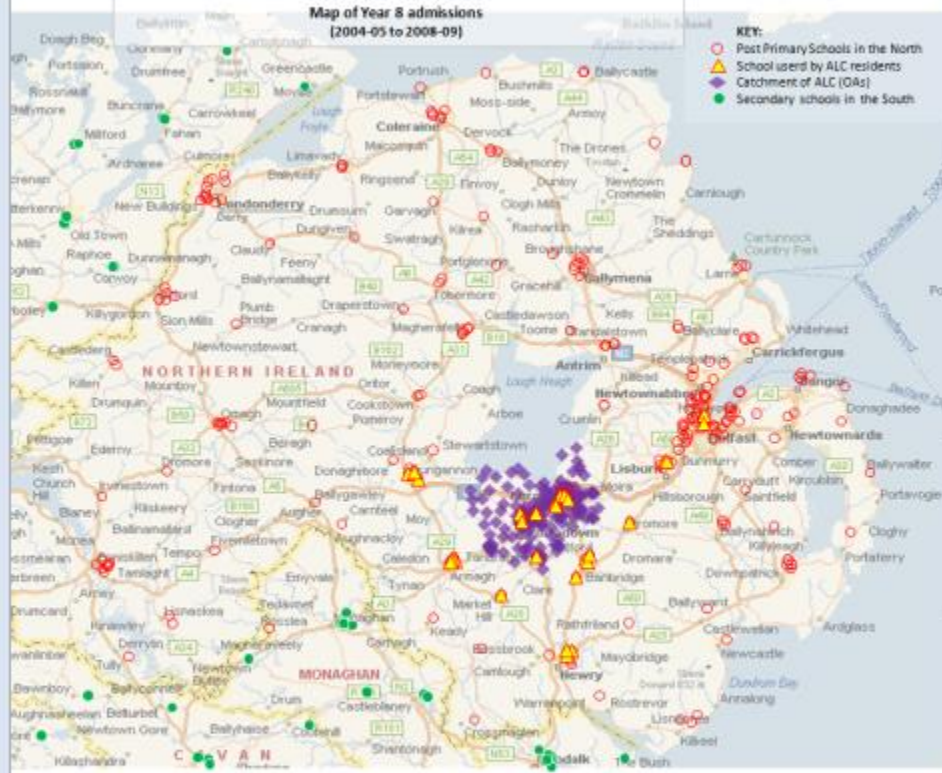
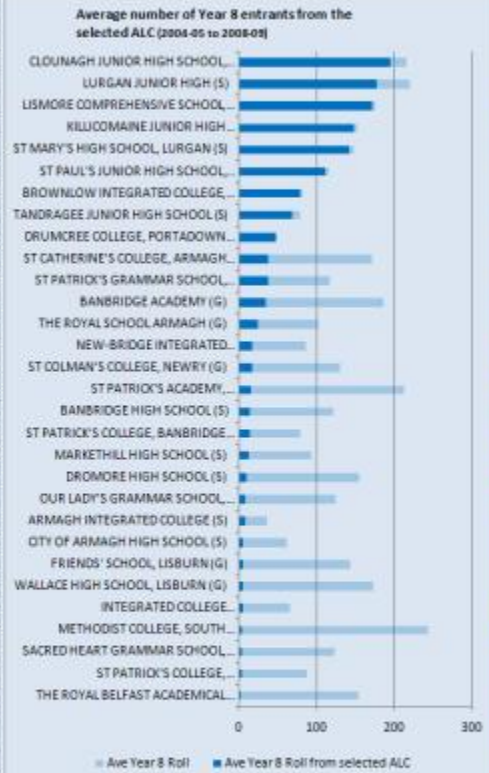


FSME (2011-12)



Craigavon ALC

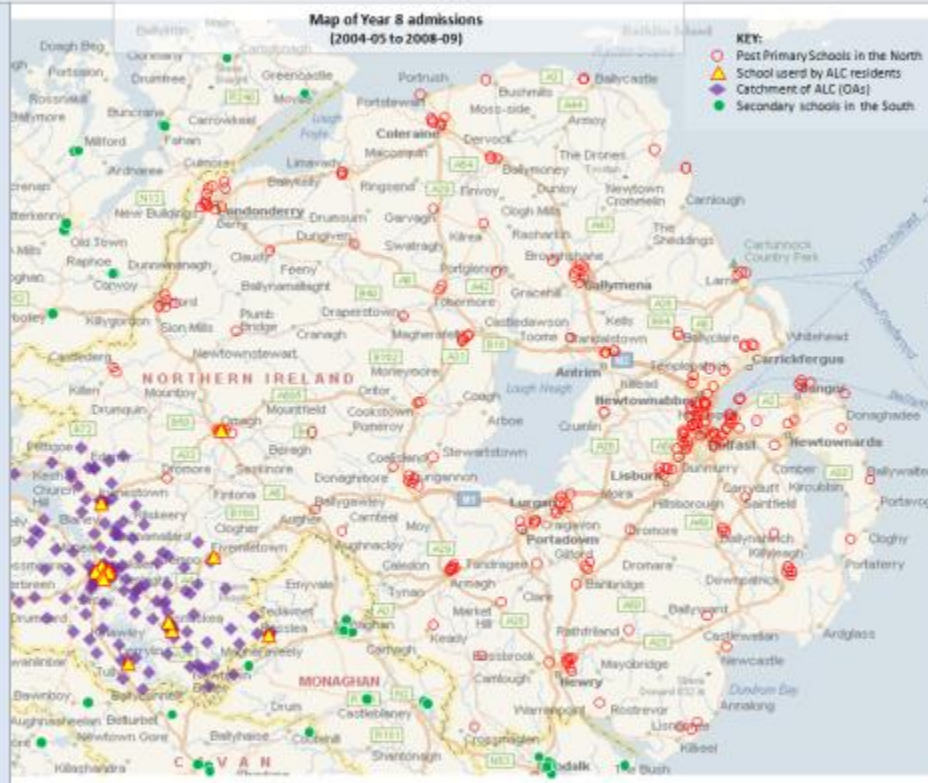
Craigavon ALC SELB



Fermanagh ALC

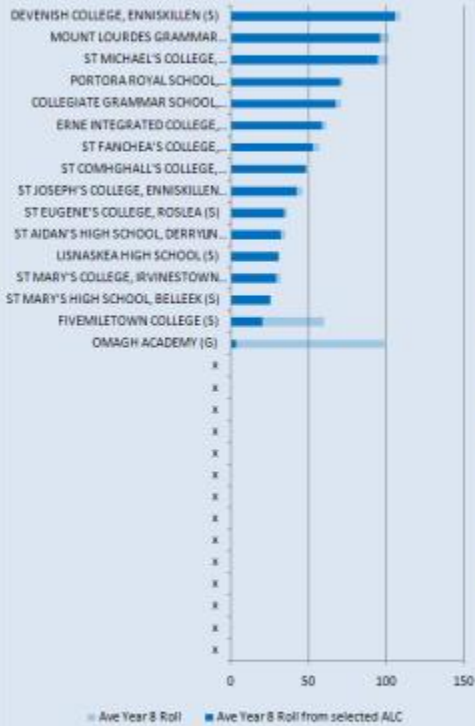
Fermanagh ALC WELB

Map of Year 8 admissions
(2004-05 to 2008-09)

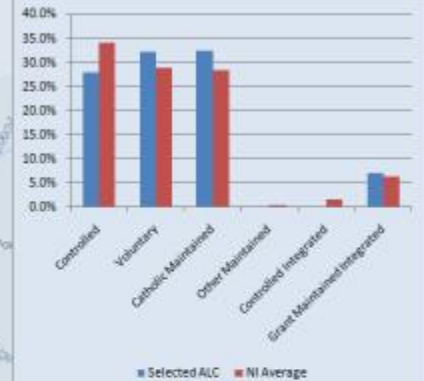


KEY:
 ○ Post Primary Schools in the North
 ▲ School used by ALC residents
 ▲ Catchment of ALC (OAs)
 ● Secondary schools in the South

Average number of Year 8 entrants from the selected ALC (2004-05 to 2008-09)



Type of school attended by Year 8 pupils
(2004-05 to 2008-09)



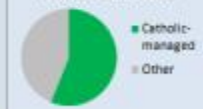
ALC (2004-05 to 2008-09)



NI Ave (2004-05 to 2008-09)



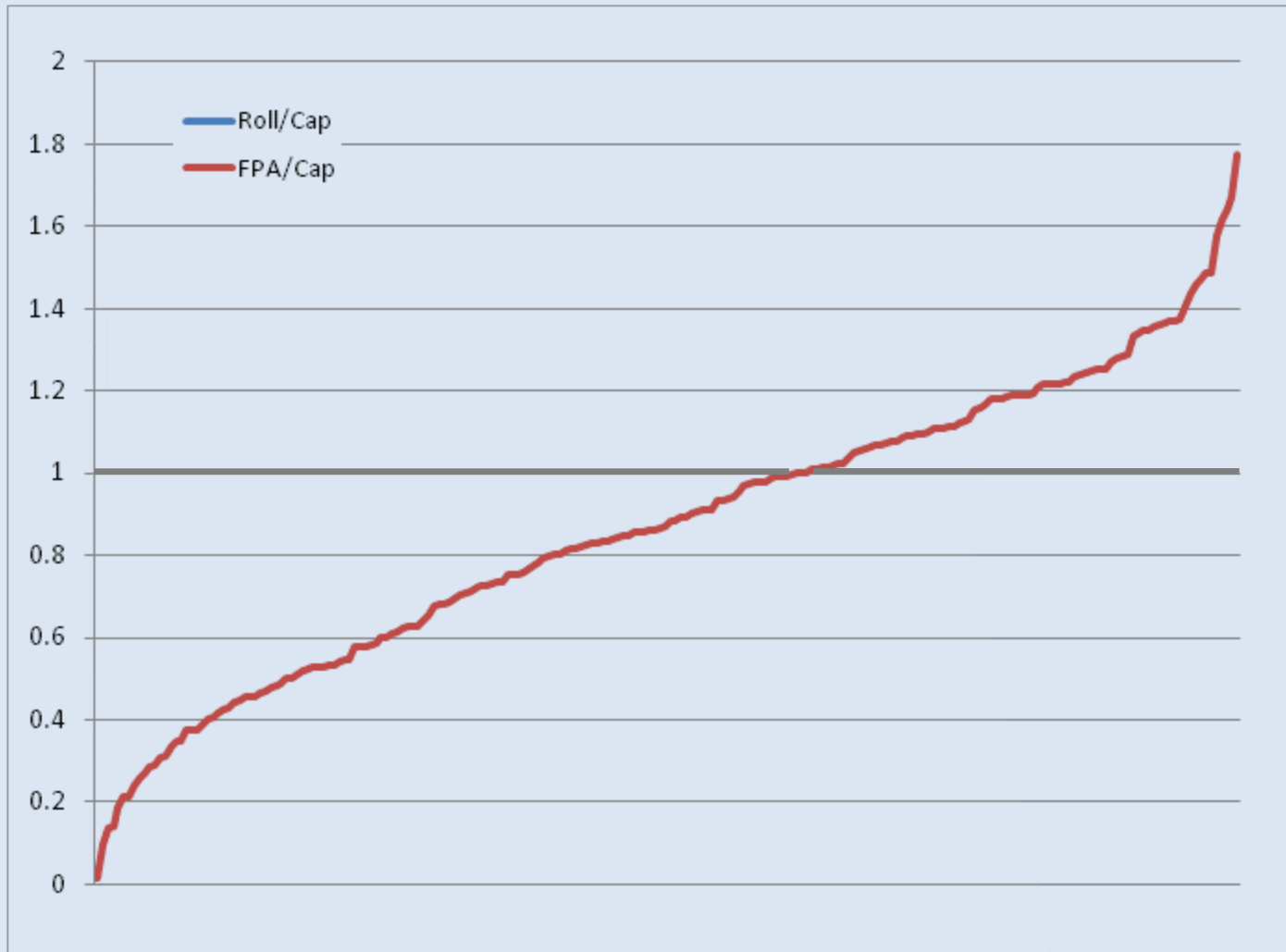
ALC (2004-05 to 2008-09)



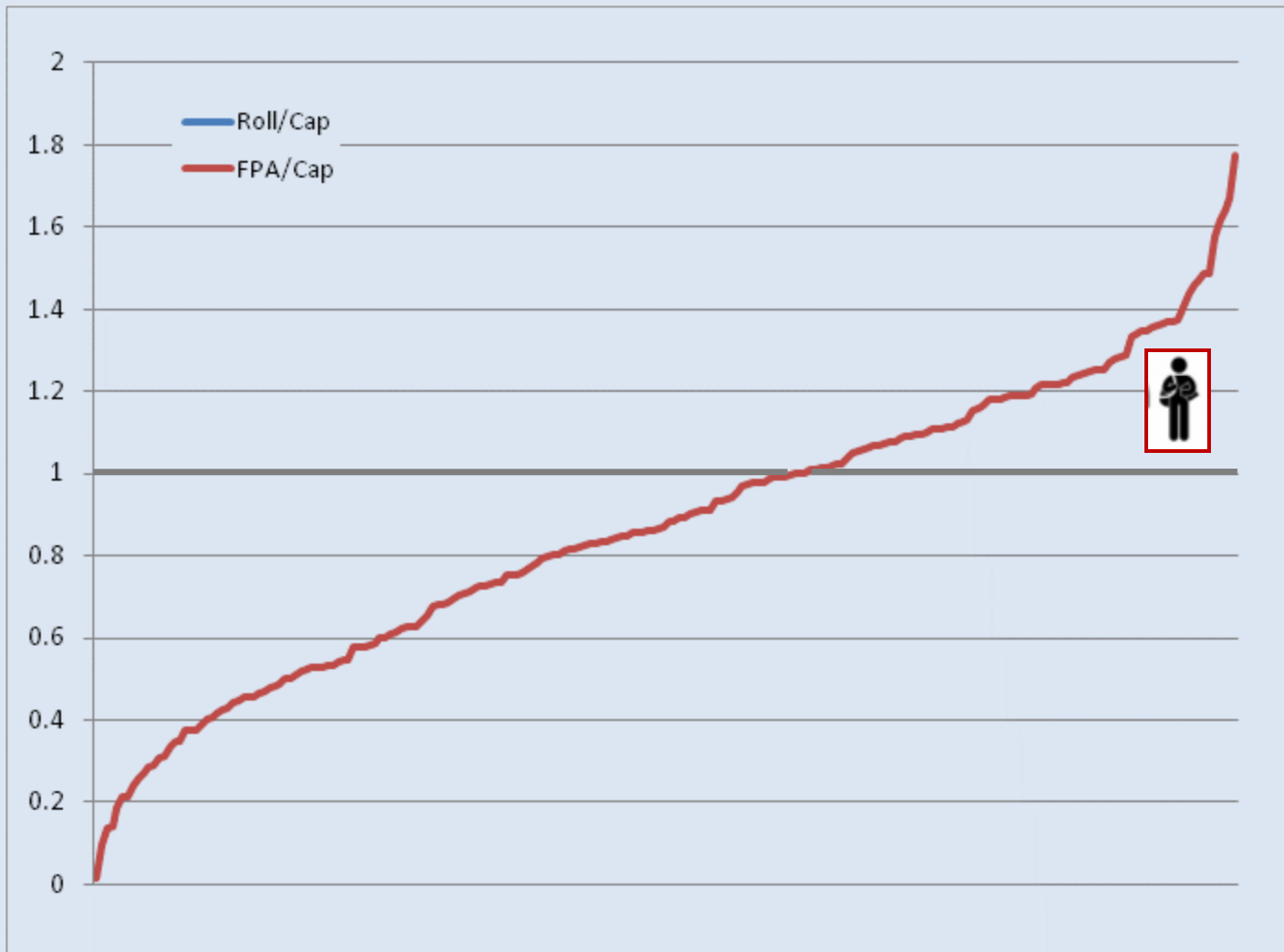
NI Ave (2004-05 to 2008-09)



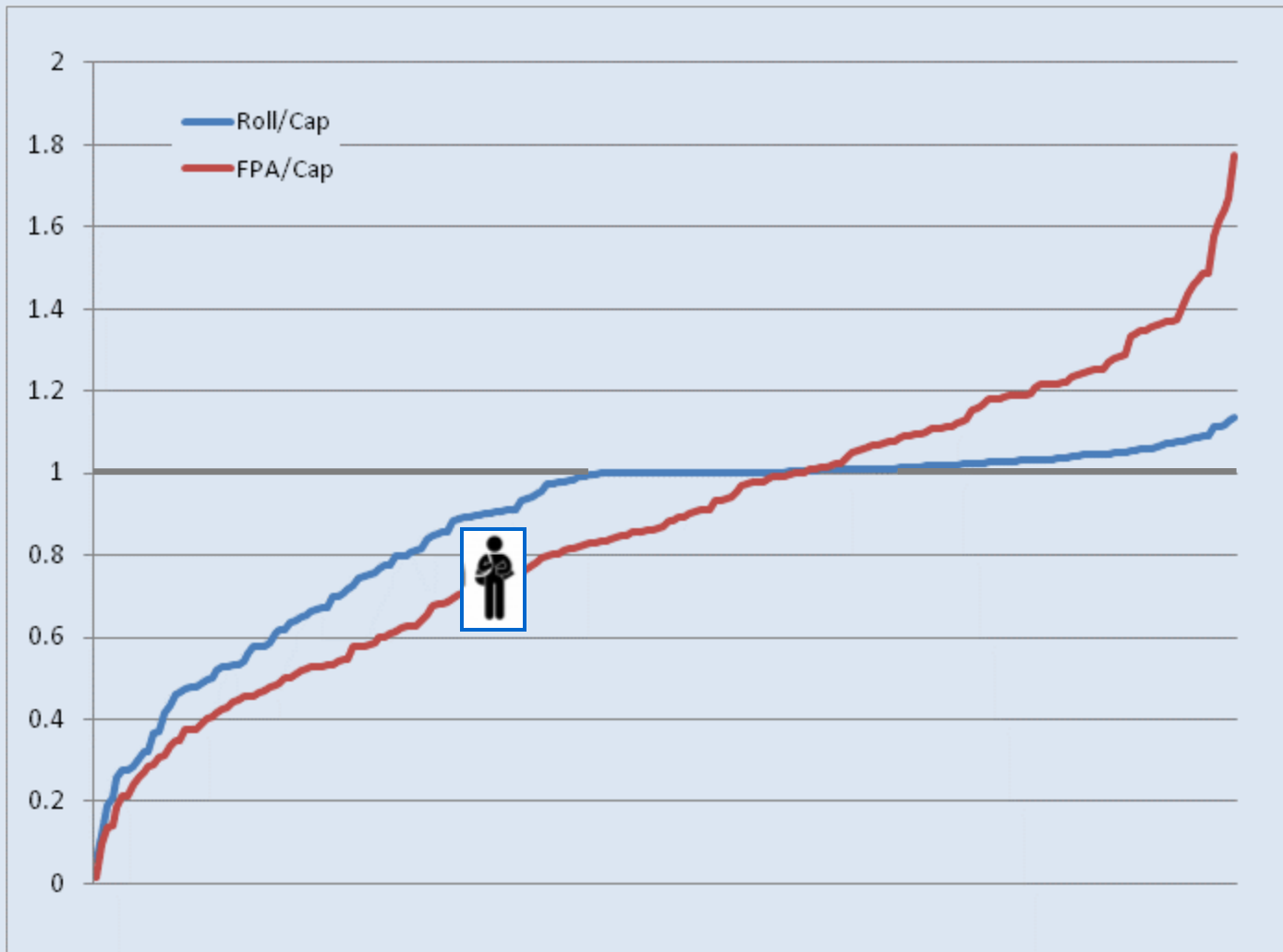
Gap between Demand & Supply



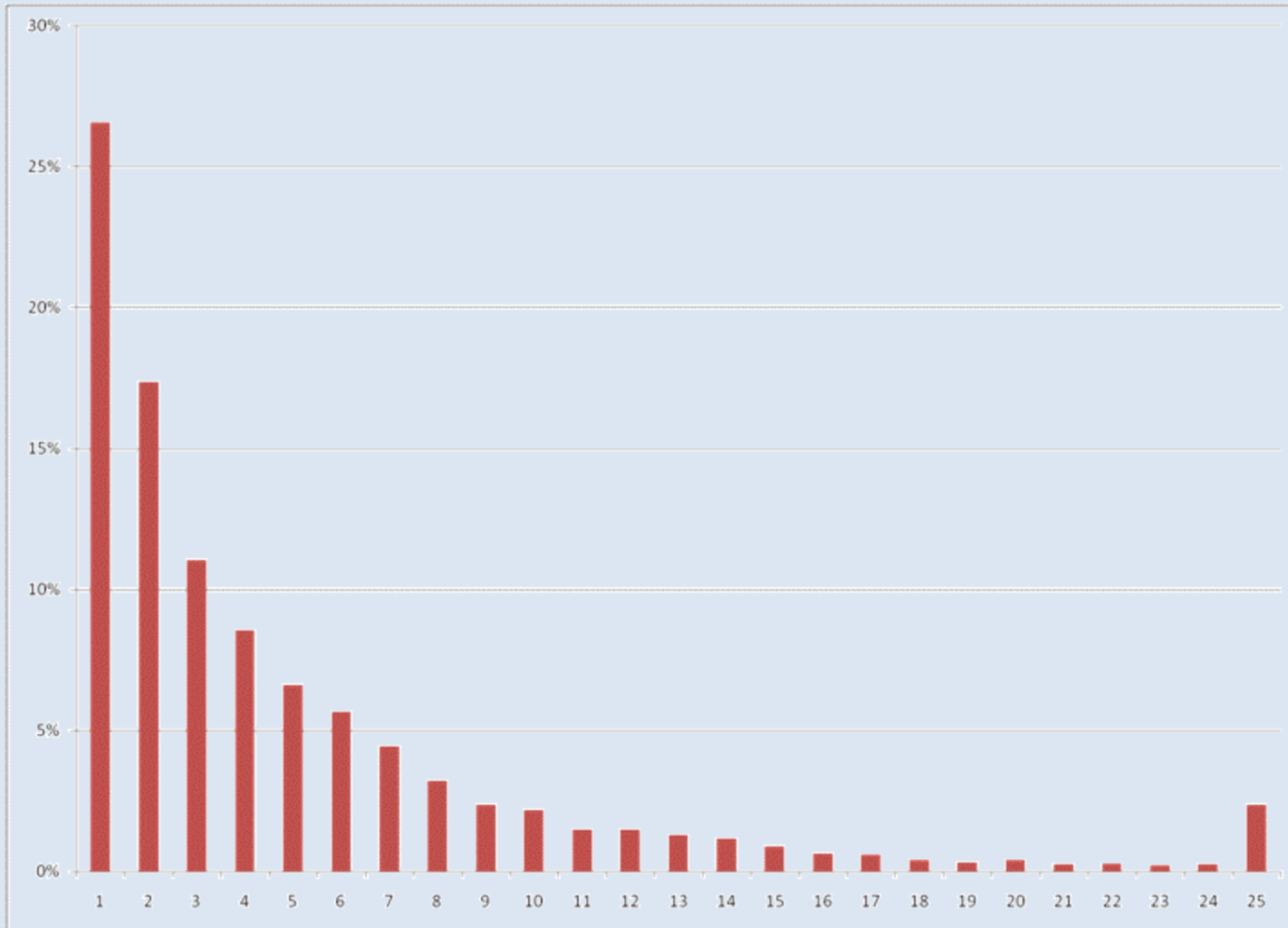
Gap between Demand & Supply



Gap between Demand & Supply



% attending the x^{th} nearest approp. school



Comparison of miles travelled to school attended & nearest school

| School type & ELB | Miles travelled to school attended | Ave miles per pupil | % going to nearest school | Miles to nearest school | Extra miles travelled |
|--------------------|------------------------------------|---------------------|---------------------------|-------------------------|-----------------------|
| Grammar | 394,656 | 6.3 | 11.7% | 131,647 | 263,009 |
| BELB | 91,176 | 5.6 | 5.1% | 20,847 | 70,329 |
| NEELB | 86,451 | 6.4 | 17.6% | 33,394 | 53,057 |
| SEELB | 64,150 | 6.2 | 8.9% | 20,329 | 43,821 |
| SELB | 77,564 | 6.9 | 15.2% | 27,842 | 49,722 |
| WELB | 75,314 | 6.9 | 13.1% | 29,234 | 46,080 |
| Secondary | 349,898 | 4.1 | 29.8% | 171,431 | 178,467 |
| BELB | 36,344 | 2.6 | 13.3% | 11,754 | 24,590 |
| NEELB | 86,550 | 4.6 | 31.7% | 42,079 | 44,471 |
| SEELB | 63,219 | 4.1 | 35.0% | 25,817 | 37,401 |
| SELB | 94,247 | 4.3 | 38.1% | 53,625 | 40,622 |
| WELB | 69,538 | 4.4 | 25.4% | 38,155 | 31,383 |
| Grand Total | 744,553 | 5.0 | 22.1% | 303,078 | 441,476 |
| Return journey | 1,489,107 | | | 606,156 | 882,951 |

Nearest school assumes no capacity constraints, and existing schools available to all communities

Miles to school attended measured from nearest OA centroid to home postcode (update pending)

Miles to nearest school (and identify of nearest school) measured from home postcode centroid

Based on road routed distances using Navteq 2008.

A) Flow modelling - understanding individuals' behaviour in current schools policy and supply environment

Hypothesis: An individual's likelihood of choosing School *i* from the set of *j* schools in the region is functionally related to:

- inverse of distance (power) from home to School *i*, relative to all other schools
- size / physical capacity of School *i*, relative to all other schools
- individual's socio-economic circumstance (mosaic/Nobel) and school type (i.e. grammar / secondary)
- individual's religion and schools' management type (RC managed, integrated, controlled)
- individual's gender and schools' gender intake (girls only, boys only, co-educational)
- other factors affecting the "attractiveness" of School *i* compared to the set

School
Planning
(DENI)

- Best fit algorithm (11 var., MaxL): McFadden $\rho^2 = 65\%$
- Challenge to account for latency in observed choices

$$\tilde{\mathbb{P}}_{nj}(\mathbf{X}_n, \Omega) = \mathbb{P}_{nj}(\mathbf{X}_n, \Omega) \cdot \frac{F_j}{D_j} + \left(1 - \sum_{j \in \mathcal{S}} \left[\mathbb{P}_{nj}(\mathbf{X}_n, \Omega) \cdot \frac{F_j}{D_j} \right] \right) \cdot \frac{\mathbb{P}_{nj}(\mathbf{X}_n, \Omega) \cdot L_j}{\sum_{j \in \mathcal{S}} \mathbb{P}_{nj}(\mathbf{X}_n, \Omega) \cdot L_j} \quad (6.7)$$

where the prior probability of first preference, $\mathbb{P}_{nj}(\mathbf{X}_n, \Omega)$, is defined by the MNL kernel term:

$$\mathbb{P}_{nj}(\mathbf{X}_n, \Omega) = \frac{X_{nj} \cdot \beta}{\sum_{j \in \mathcal{S}} X_{nj} \cdot \beta} \quad (6.8)$$

$$\mathbb{P}^{nj}(\mathbf{X}^{nj}, \Omega) = \frac{\sum_{j \in \mathcal{S}} X^{nj} \cdot \beta}{X^{nj} \cdot \beta} \quad (6.9)$$

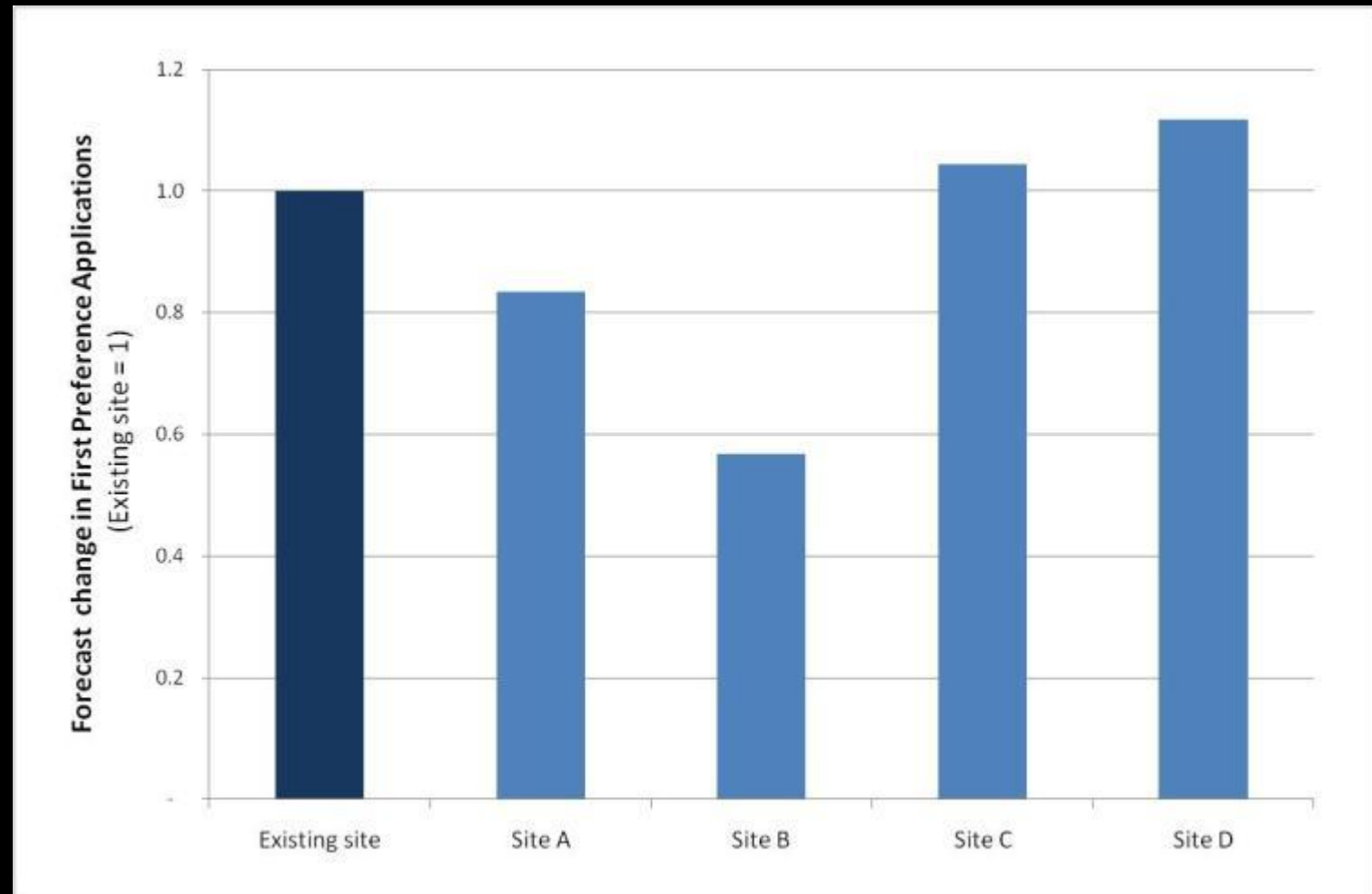
Initial model



Latest research



Results for exemplar school relocation:





School Planning (DENI)

School Access Simulator

Evidence to the School Transport Review Panel

Dr. Martin M. Spollen

Head of Investment Strategy
Strategic Investment Board (SIB)

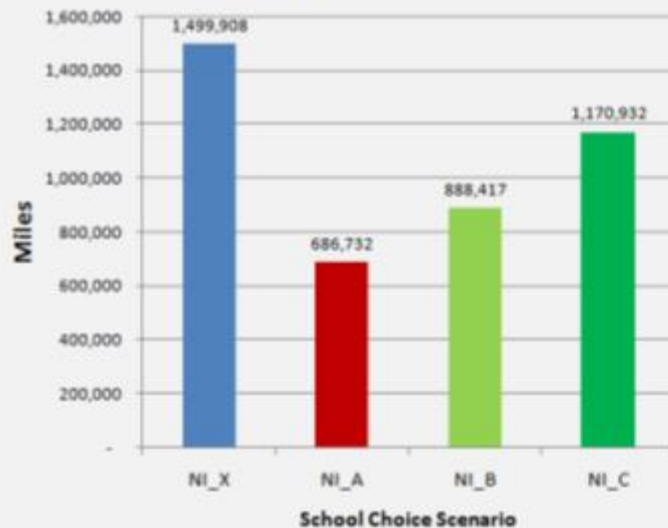
Centre for Statistical Science
School of Mathematics

Mob: +44 (0)7966 982363

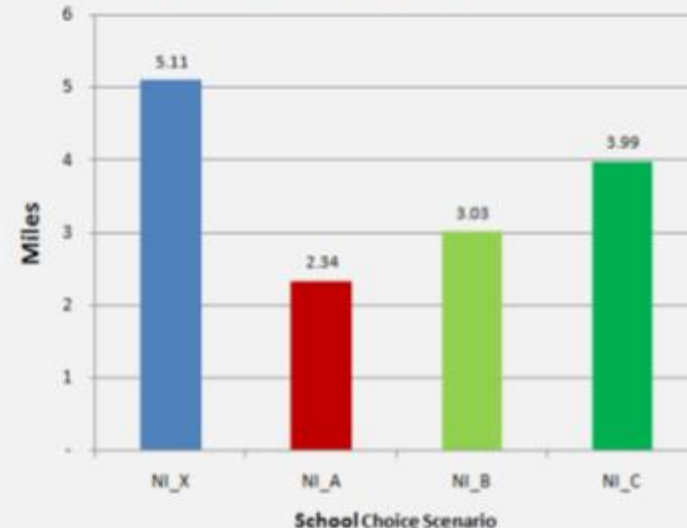
martin.spollen@sibni.org

msspollen01@qub.ac.uk

NI : Total daily round trip



NI : Average daily one way distance



- X - Current school
- A - Nearest school (Open, Age, Gender)
- B - Nearest school (Open, Age, Gender, Nest_3)
- C - Nearest school (Open, Age, Gender, Nest_5)

- C - Nearest school (Open, Age, Gender, Nest_5)
- B - Nearest school (Open, Age, Gender, Nest_3)
- A - Nearest school (Open, Age, Gender)
- X - Current school

EA IT environment

Existing client layer



SIPM

[1] user interface (UI) driver layer

desktop UI mobile

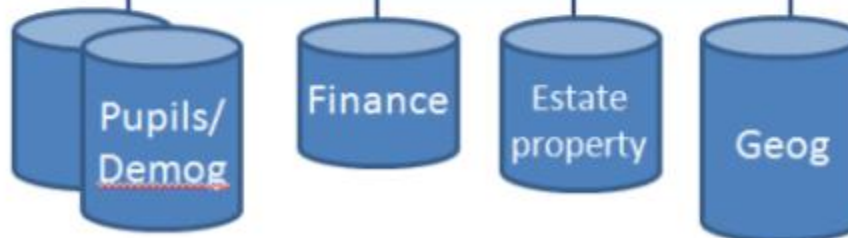
[2] applications * ('app') layer



[3] middleware layer

middleware interface

base layer (existing databases)



external web

web interface (input feed ONLY)



* Apps will be deployed on client side devices and/or server as is deemed most technically appropriate to maximise overall performance.



LibrariesNI rationalisation programme:

Rationalisation from 110 to c.90 public libraries in NI

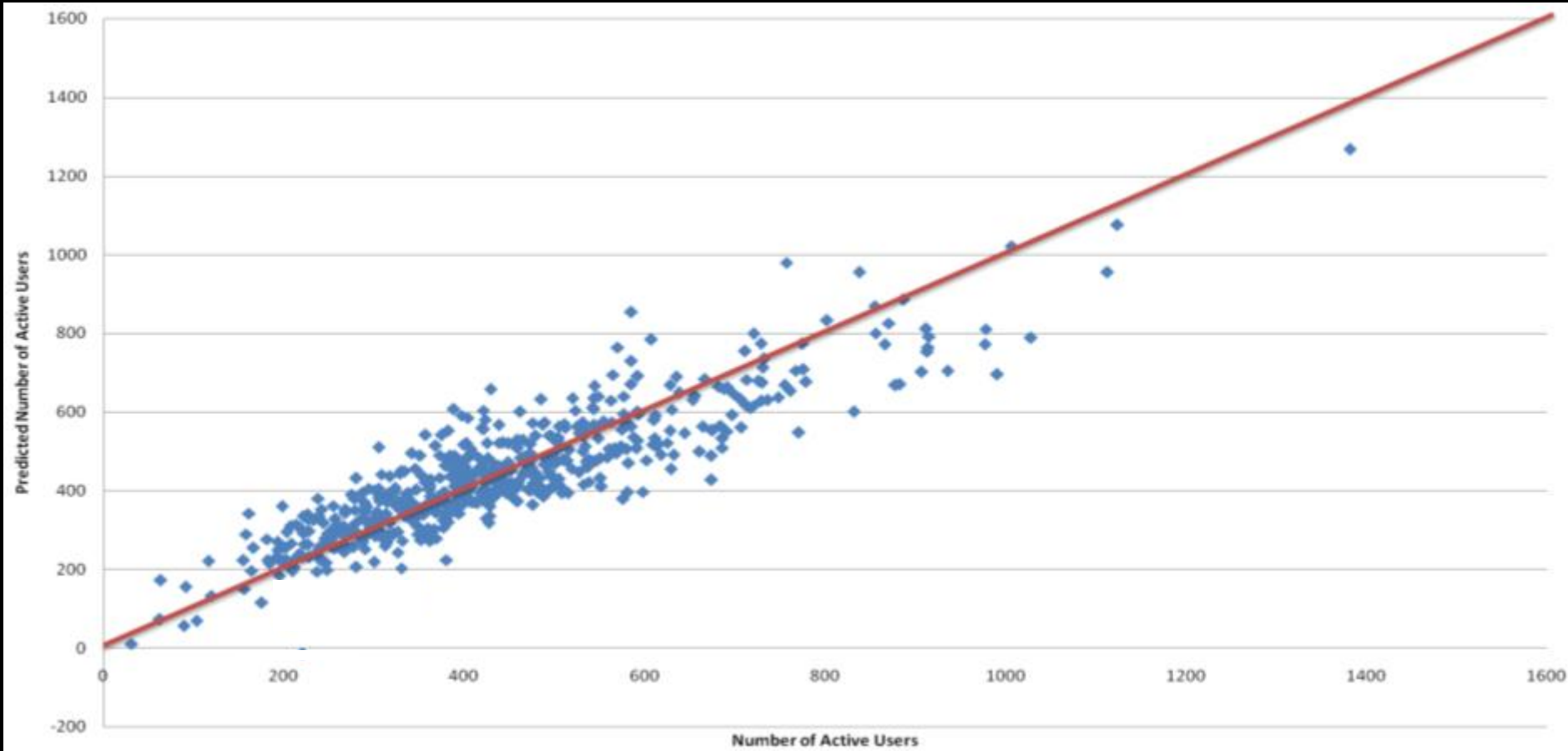
Optimal deployment of 30 mobile libraries across rural communities and smaller towns.

Analytics used for determine how population access could be maximised through choice of optimal locations



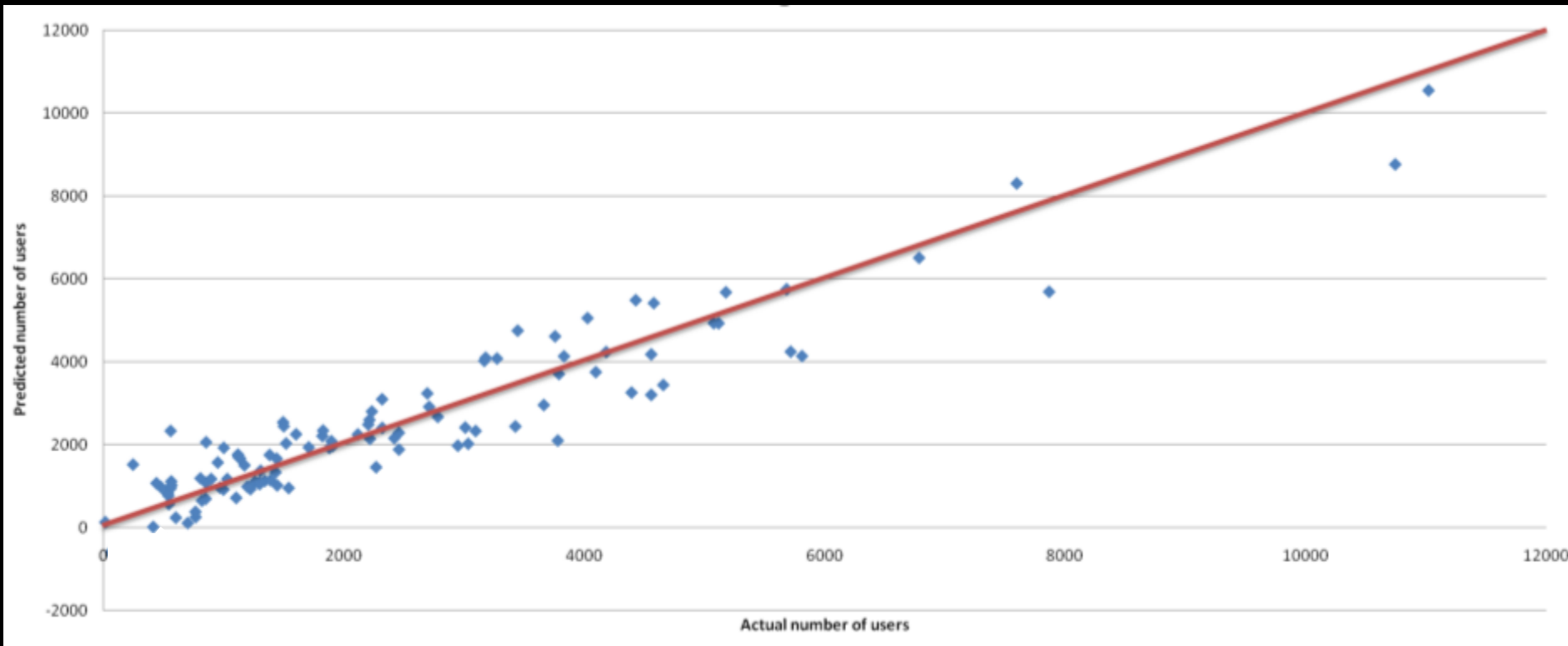
Library Services (DCAL)

Predicted v Actual number of active library users in each electoral ward (582 wards)





Predicted v Actual level of demand at each public library in NI (109 libraries)



Library Services (DCAL)

| Library branch | BASE SCENARIO | Closed: | | | |
|----------------------------------|---------------|------------|------------|------------|------------|
| | | 10 | 19 | 41 | - |
| | | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 |
| Andersonstown Library, Belfast | Open ✓ | Closed ✗ | Closed ✗ | Closed ✗ | Open ✓ |
| Antrim Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ardoyne Library, Belfast | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Armagh Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ballycastle Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ballyclare Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ballyhackamore Library, Belfast | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Ballymacarrett Library, Belfast | Open ✓ | Closed ✗ | Closed ✗ | Closed ✗ | Open ✓ |
| Ballymena Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ballymoney Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Ballynahinch Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Banbridge Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Bangor Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Belvoir Park Library, Belfast | Open ✓ | Closed ✗ | Closed ✗ | Closed ✗ | Open ✓ |
| Bessbrook Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Braniel Library, Belfast | Open ✓ | Closed ✗ | Closed ✗ | Closed ✗ | Open ✓ |
| Broughshane Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Brownlow Library, Craigavon | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Carnlough Library | Open ✓ | Open ✓ | Closed ✗ | Closed ✗ | Open ✓ |
| Carrickfergus Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Carryduff Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Castlederg Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Castlewellan Library | Open ✓ | Open ✓ | Open ✓ | Open ✓ | Open ✓ |
| Central Lending Library, Belfast | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Chichester Library, Belfast | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Cherry Library, Newtownabbey | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Cherry Library, Newtownabbey | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |
| Cherry Library, Newtownabbey | Open ✓ | Open ✓ | Open ✓ | Closed ✗ | Open ✓ |

Resource
Allocation
(DHSSPS)

DHSSPS(NI), DOH(Eng. & Wales)

- Allocate funding equitably between health trusts
- Take account of catchment populations':
 - relative **needs** (age & morbidity weighted); and
 - relative **unit costs** of provision (esp. rurality & EoS)
- Statistical analysis of PAS & costing data
- Now used extensively in GB (e.g. Adult PSS)



NIAS Emergency Ambulance deployment:

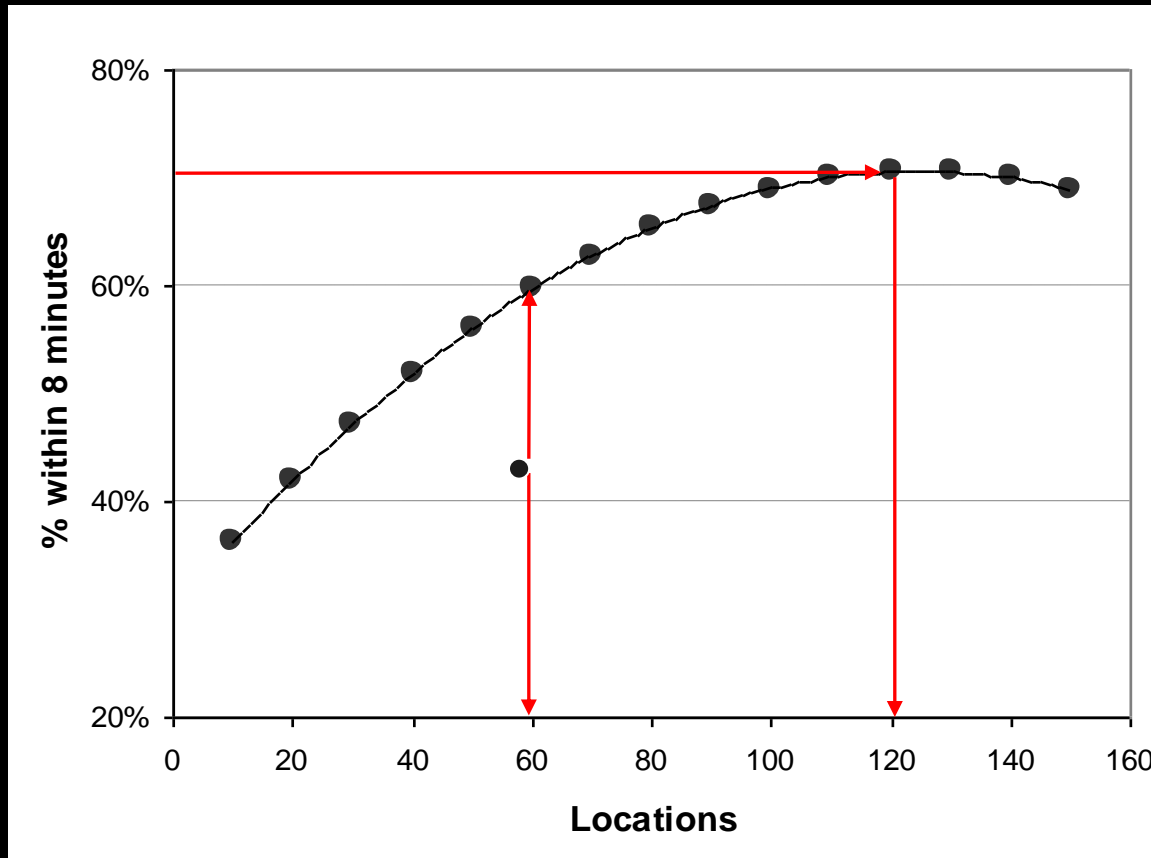
5,450 sq. miles & 21k emergency calls per year. 300 emergency vehicles (at time of orig. study)

Optimal deployment of emergency vehicles, first responder and defibrillation assets across NI.

Analytics used for determine 'dynamic' deployment points; move from traditional 'stations'



Analytics used to establish 'efficiency frontier'





Police Service for Northern Ireland:

Optimal deployment of 7,000 frontline police officers and mobile assets to meet publically accountable performance standards.

GPS tracking in-situ (what is/was where?)

Analytics used for demand look-ahead and strategic deployment (what needs to be where?)

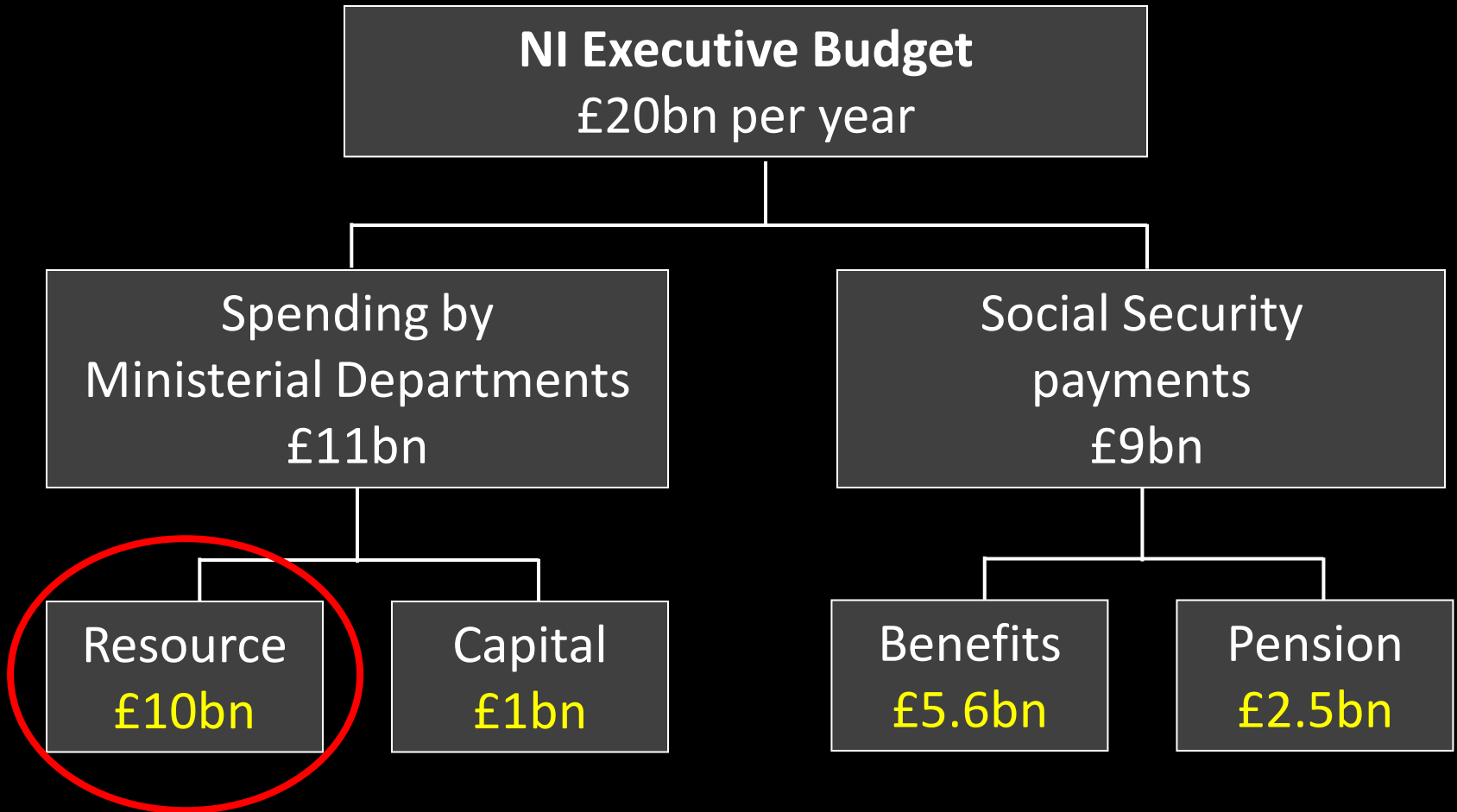
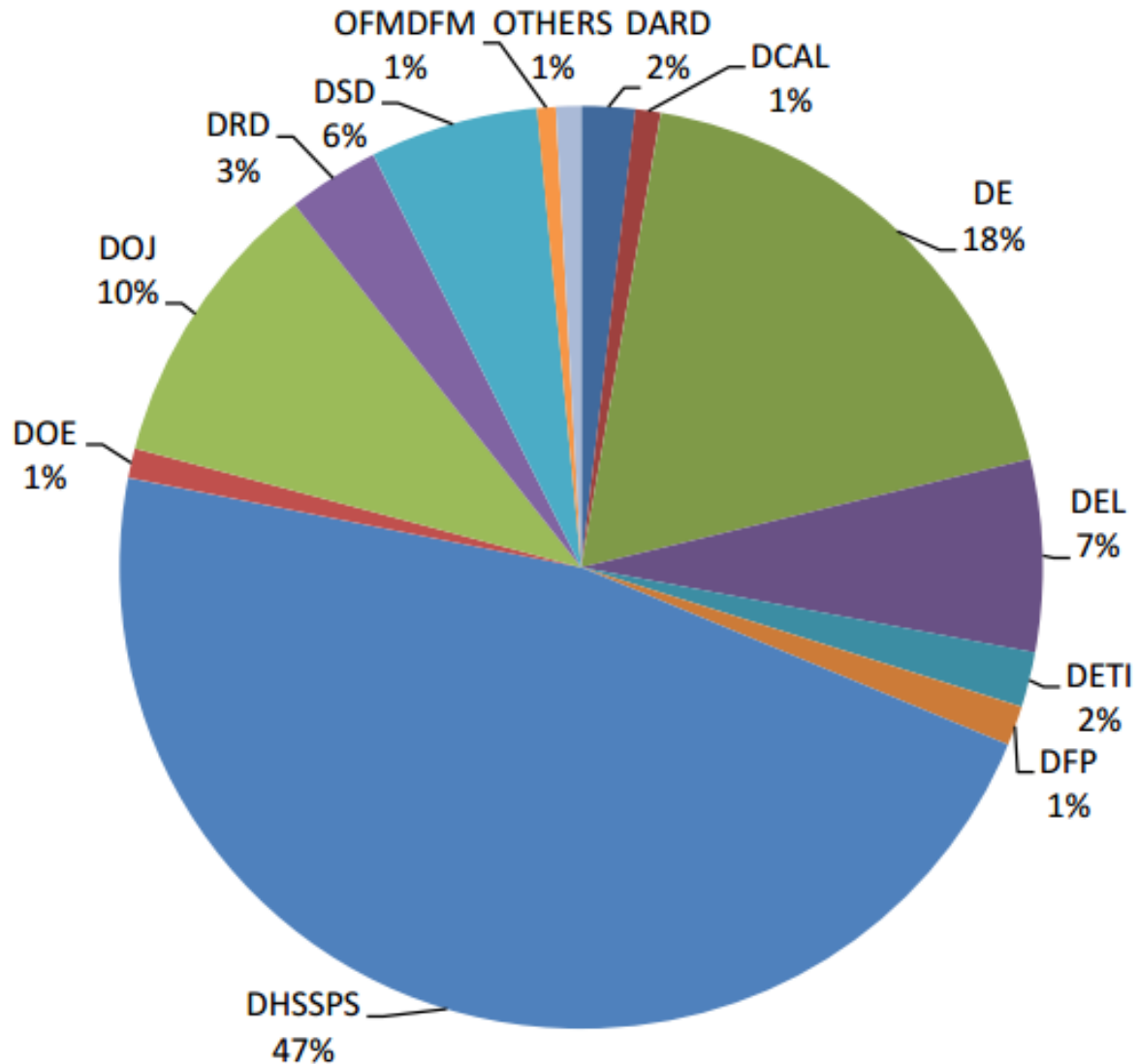


Chart 4.1: Breakdown of Non Ring-fenced Resource DEL Expenditure by Department 2015-16



Realising the potential: some challenges

- **Develop NICS data science skills base**
 - data collection/cleaning technologies
 - statistical techniques for data mining/analysis
 - big data software architecture (Hadoop, MR)
- **Agree consistent publication scheme:**
 - raw data please - avoid excessive aggregation
 - mandate use of meta data (EU Inspire)
 - longitudinal and cross-sectional identifiers
- **Data warehousing & accessibility:**
 - NICS Data Centre project promising

Realising the potential: some challenges

- **Redefine legal boundaries under Data Protection & statute:**
 - remove/relax barriers to inter-NICS data sharing
 - joining up datasets adds value/insight
 - protocol for import of third-party sets (e.g. Met Office, Experian)
- **Stimulate demand across NICS client base**
 - demonstrator projects needed at all 3 levels
 - facilitator role