

Classifying consumers: Adding more to the Census

John Rae



Agenda



- Basic principles
- What were we thinking about?
- Why is interaction of variables critical?
- Measuring improvement
- Requests we turned down



Basic principles



- People in similar areas have the same needs and lifestyles
- You can find these people
- Without ever having met or heard of somebody we aim to predict
 - Likely spending
 - Likely ownership of goods
 - Likely readership
 - Likely reactions to offers
- And get it substantially right



Our thinking



- We've done this since the 70's
- One key change in the census changed our thinking
- OAs are more homogeneous



Initial approach (1970's)



Initially data came in big blocks.



Census output areas were pre-defined and you had to classify whatever mix of people occurred.

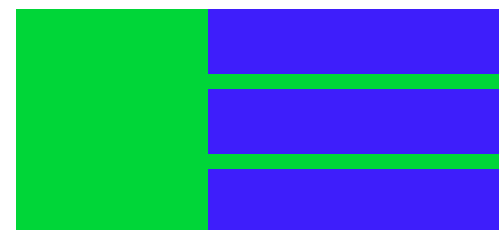
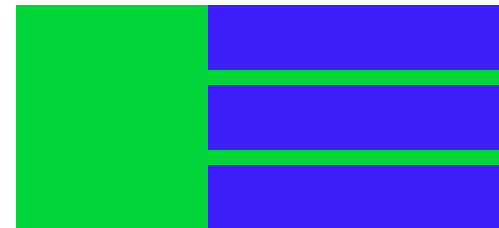
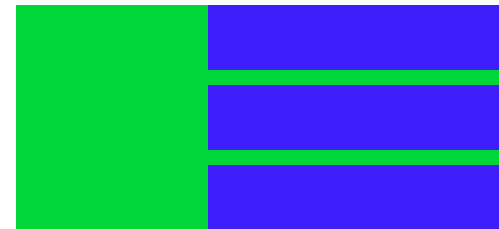


Adding detail

Census areas could contain different people so postcode data .

Random noise and sampling issues were not considered serious. The "Shaw test"

Assumed the balance of proof favoured the postcode



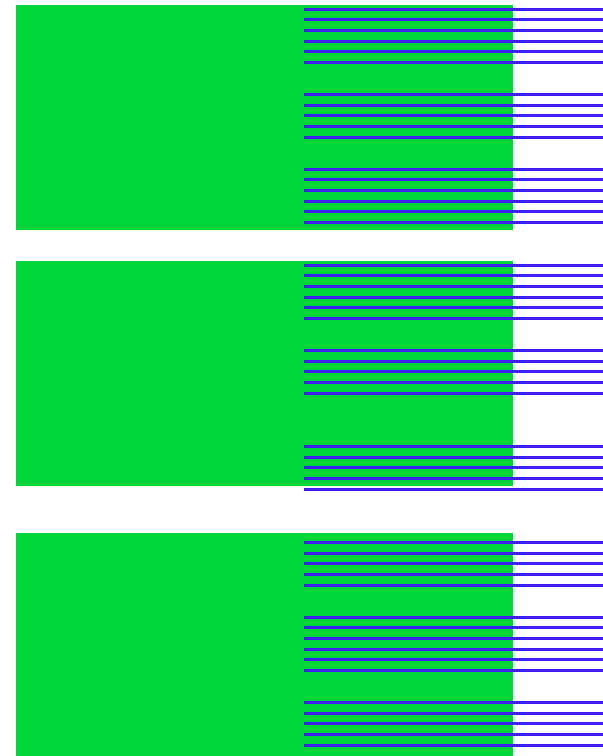
Approach turns on its head!



Now the OA's are **designed** to be homogeneous.

So in theory all the postcodes are the same.

Balance of proof has shifted. It is no longer sufficient to be as casual about the incompleteness of postcode data.



Mixing methods – two stages



Stage 1

Classification of Census
output areas

Stage 2

Introduce other data
Decision rules to re-classify
postcodes



Advantages of two stages



- Method allows easier annual updates
 - Address new housing
 - Address infill house build
 - Address migration
 - Focus on areas of change
- Comparison with direct classification showed sharper targeting in all tests.



Classification :

How variables interact



- The information implicit in key variables makes others redundant
- Throwing too many variables into the pot makes classifications worse



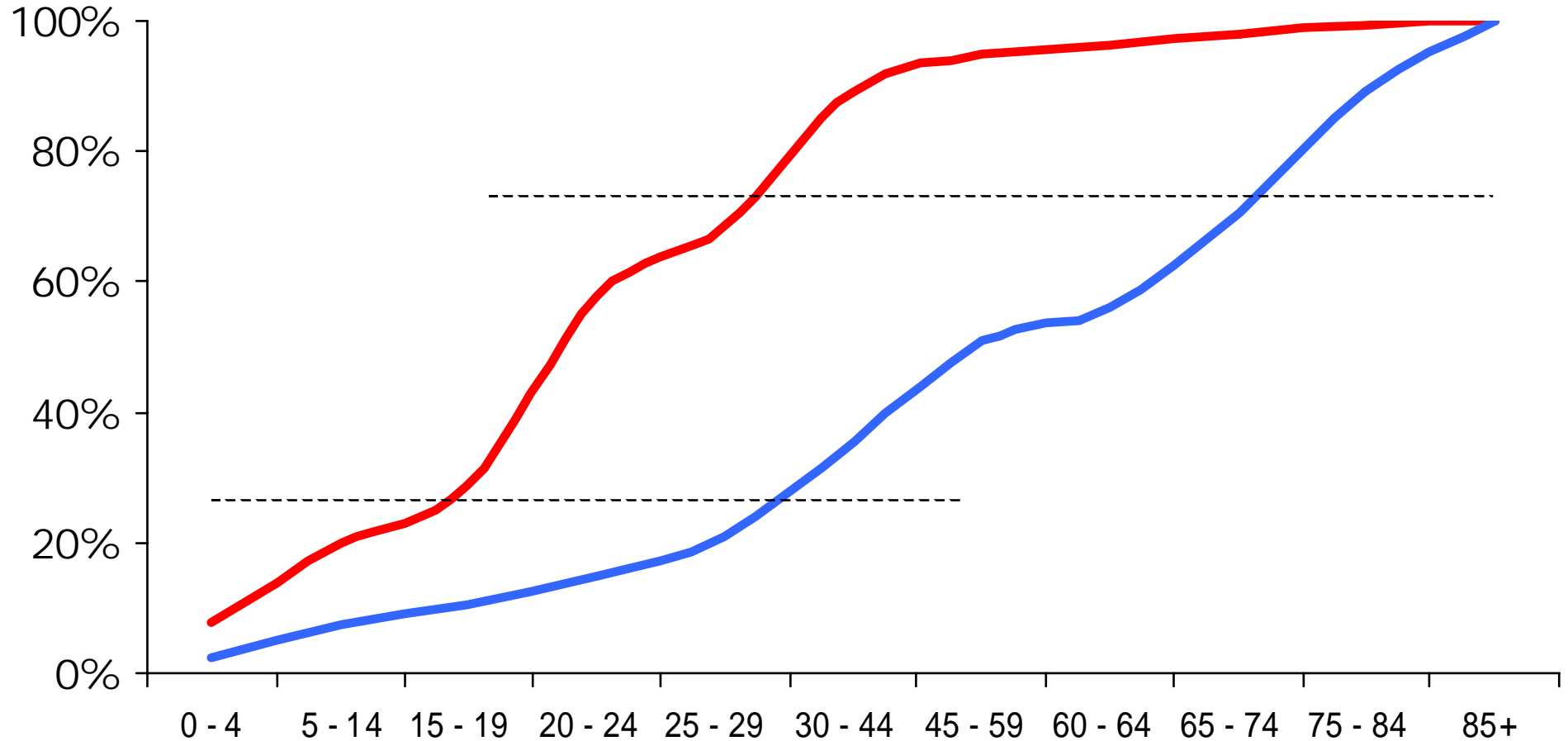
Types of variables



- Ranges
 - Household size
- Ratios
 - Rooms per hh
- Fractions
 - % dwellings with 1-2 rooms
- Incomplete sample
 - Bungabws
- Inverse variables
 - White
 - No religion
- Adapted ranges
 - Age quartiles



Age quartiles



A frequent misunderstanding



- The type exists before any name or description
- Names can only provide a superficial picture
- Names may describe the majority behaviour, or highlight special features



Special circumstances?



- If it is labeled an Asian type must Asian people live in the postcode?
- If it is labeled a student type must students live in the postcode?
- Is manual intervention appropriate?
- Which makes the classification better?

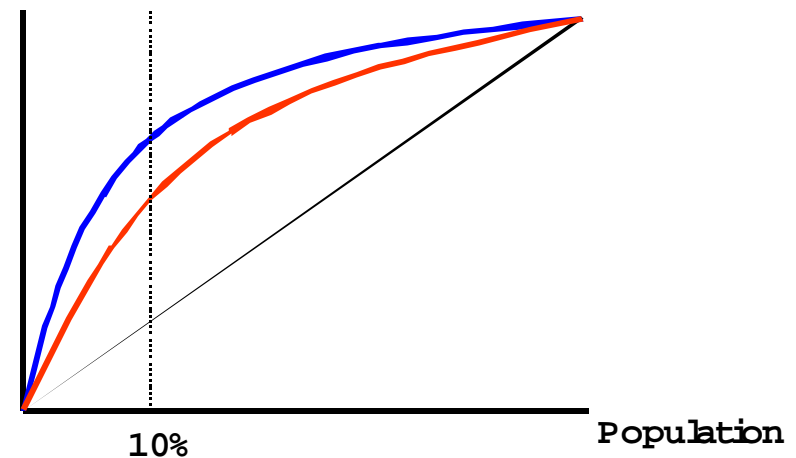


The acid test: Measuring improvement



- Randomly split consumer file
- Profile first half
- Identify **acorn** types
- Apply blind to other half
- Measure 10% point of gains chart
- Repeat frequently to assess standard deviation
- Measure statistical significance of improvement

Target Market



Better discrimination



- Family income
- House price
- Mortgage
- Medical insurance
- Pensions
- Savings
- Investments
- Credit cards
- Newspapers
- Media
- Cars
- White goods
- Food spend
- Eating out
- Weekend breaks
- Holidays
- Music/Opera/Art
- Computers
- Internet usage



Requests we turned down



- "Sexy" names



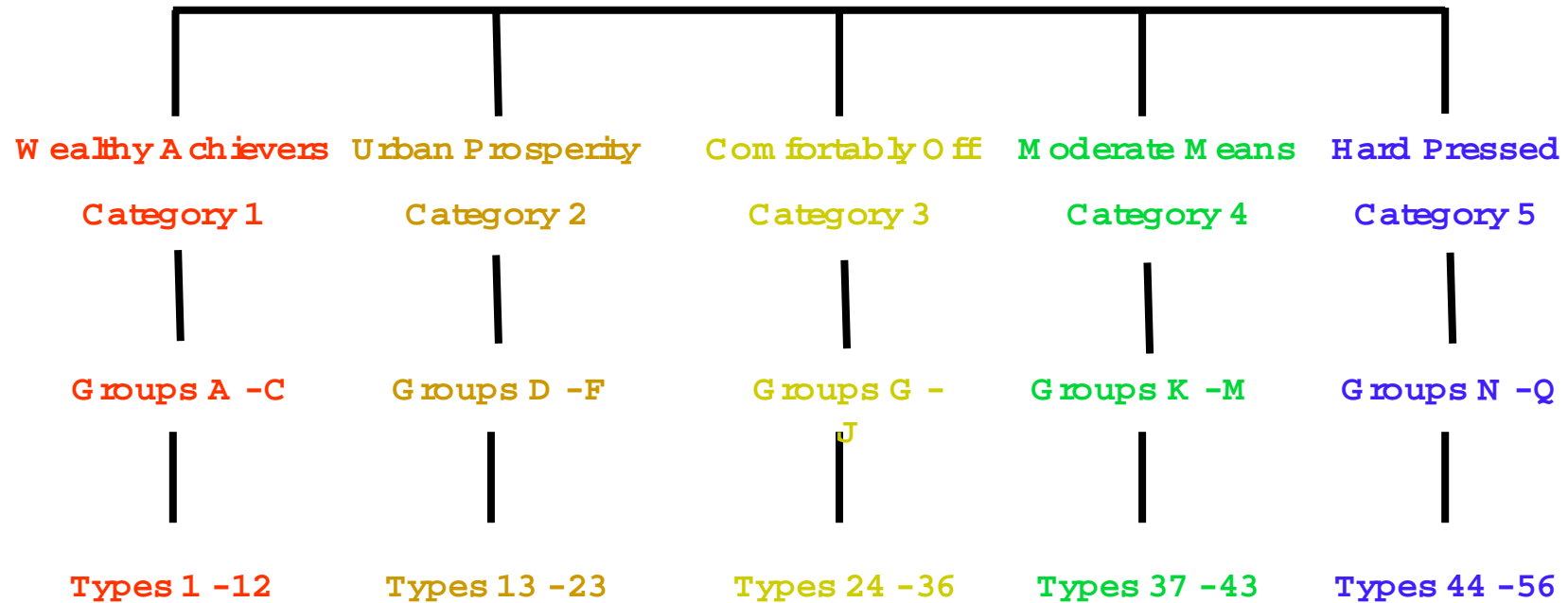
Names



- When asked over 95% of users opted for descriptive names
- Only advertising agencies were in favour of very evocative names
- 'Within our team we loved using funny names, but our MD banned us from talking about customers in this way.'



acorn - structure



1 : Affluent Achievers



- **A Wealthy Executives**
 - A.1 Wealthy mature professionals, large houses
 - A.2 Wealthy working families with mortgages
 - A.3 Villages with wealthy commuters
 - A.4 Well-off managers, larger houses
- **B Affluent Greys**
 - B.5 Older affluent professionals
 - B.6 Family communities
 - B.7 Old people, detached homes
 - B.8 Mature couples, smaller detached homes
- **C Flourishing Families**
 - C.9 Larger families, prosperous suburbs,
 - C.10 Well-off working families with mortgages
 - C.11 Well-off managers, detached houses
 - C.12 Large families & houses in rural areas



A 1 W ealthy m ature professionals, large houses

- Typically found in Kingston,
St Albans, Guildford
- Unit Trusts
- Food spend
- Earn £50,000+
- High Credit Card spend
- Big houses
- Company Car
- FT
- Detached houses



Requests we turned down



- "Sexy" names
- "Household" **acorn**



Mixing households and postcodes



- CACI have had an individual level "acorn" for over 4 years
- Branded as PeopleUK
- Marketing wanted to merge the brands
 - same names and types for both
 - "it sounds such a great idea"
 - elegant sales and PR story
 - coordinate support material



Well-off professionals, larger houses and converted flats



- Our Chief Executive
 - Wealthy, lives in a big house, is a senior manager and commutes by car
 - Like a type 1, Wealthy mature professionals, large houses
- A nearby neighbour
 - Affluent, rents a big flat, is a senior manager, and commutes by tube
 - Like type 15, Affluent urban professionals, flats
- This type is important in a postcode classification
- A postcode can be a mix
- A household can only be one thing or the other, it could never be this type



Summary



- Classification is a mix of skill and art
- Census changes prompted new thinking about our classification
- We supplement the census with additional data
- Decision rules check the accuracy of each postcode classification
- This facilitates updates

