Marriage matters
A practical guide to pricing contingent dependants

Andy Harding
Demographic Horizons™ team, Aon

Luke Davies
LexisNexis® Risk Solutions

18 November 2019

Contingent dependants

Proportion married (or with wider financial dependants) – why does it matter?

Increasingly material
- PV impact
- Pricing focus

Increasing sophistication required
- Data and definitions
- Segmentation

Contingent dependants

Proportion married (or with wider financial dependants) – why does it matter?

Increasingly material
- PV impact
- Pricing focus

Increasing sophistication required
- Data and definitions
- Segmentation

Potential impact ±3% of joint life PV

Contingent dependants

Proportion married (or with wider financial dependants) – why does it matter?

Increasingly material
- PV impact
- Pricing focus

Increasing sophistication required
- Data and definitions
- Segmentation

Potential impact ±3% of joint life PV

Accuracy is paramount
- over-pricing may lose deals
- under-pricing may impair profitability or weaken reserves

Agenda

1 **Traditional approaches**
- National statistics, experience data and surveys
- Limitations and how to deal with them

2 **Member tracing and LexisNexis® Risk MSP**
- Who's who? – linking data records
- Are they married? – identifying spouses

3 **Tracing in practice**
- How to interpret the codes
- Does it actually work? – performance testing

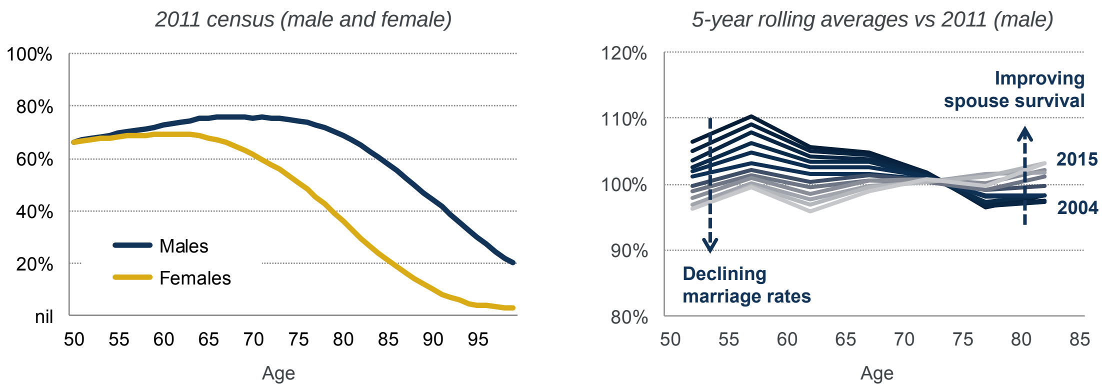
Institute and Faculty of Actuaries

LexisNexis® RISK SOLUTIONS

1. Traditional approaches

National statistics

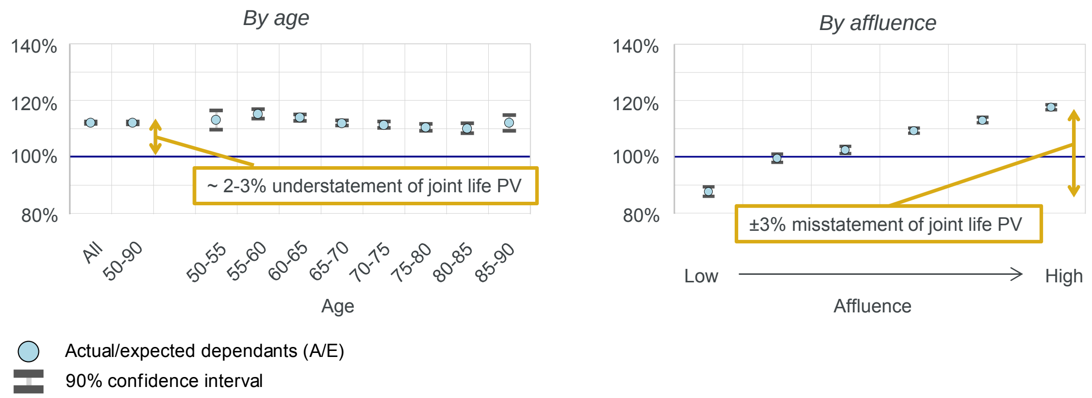
Proportion married – England & Wales (2011 census)



Source: ONS data with Aon calculations
Annual variation in proportion married based on ONS Labour Force Survey adjusted for mortality improvement

Pension scheme members

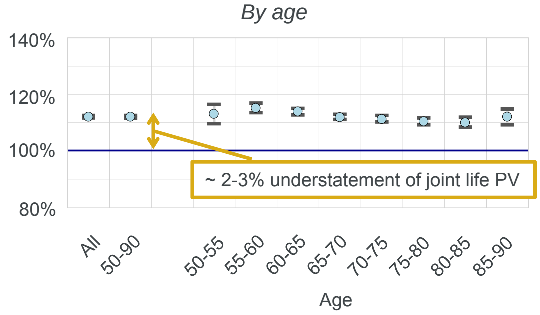
Proportion married – male pension scheme A/E vs England & Wales (amounts-weighted)*

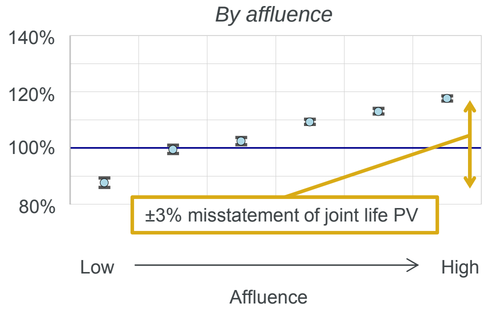


** Demographic Horizons pension scheme survey data (adjusted for respondent bias) vs ONS E&W 2011 census data (projected from 2011 using annual adjustments from ONS Labour Force Survey)*

Pension scheme members

Proportion married – male pension scheme A/E vs England & Wales (amounts-weighted)*





Actual/expected dependants (A/E)
90% confidence interval

Solution: Use postcode model calibrated to *pension scheme* data, with realistic age shape, time trends and socio-economic variation, and allowance for alternative eligibility definitions (e.g. legal spouse vs wider financial dependant)

** Demographic Horizons pension scheme survey data (adjusted for respondent bias) vs ONS E&W 2011 census data (projected from 2011 using annual adjustments from ONS Labour Force Survey)*

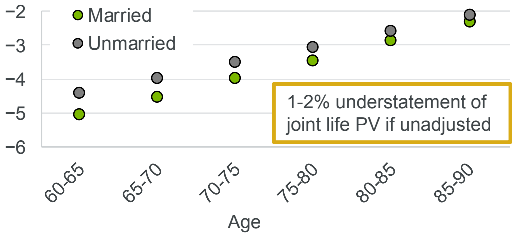
Experience data (i.e. deaths)

Deceased vs current members may differ in terms of

- age profile
- socio-economic profile
- effective date of information

And mortality rates are *lower* for married than unmarried individuals, even after controlling for these factors:

Log mortality rates* – England & Wales males (2011)



So care is needed when

- *fitting* a dependants model to deaths data and then
- *applying* it to value current lives

** Standardised by Index of Multiple Deprivation (IMD) 2015 decile*
Source: ONS data with Aon calculations

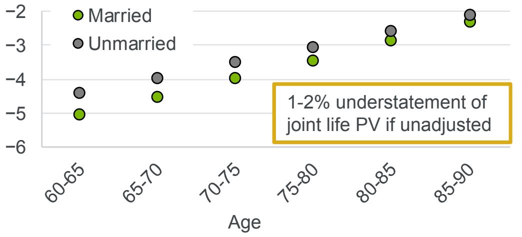
Experience data (i.e. deaths)

Deceased vs current members may differ in terms of

- age profile
- socio-economic profile
- effective date of information

And mortality rates are *lower* for married than unmarried individuals, even after controlling for these factors:

Log mortality rates* – England & Wales males (2011)



So care is needed when

- *fitting* a dependants model to deaths data and then
- *applying* it to value current lives

Solution

Fit to data using proportional odds model:

$$o_{it}(\beta) = o_{it}^{prior}\, exp(\beta^T \varphi_{it}), \quad \text{where} \quad o_{it} = \frac{p_{it}}{1-p_{it}}$$

The prior model o^{prior} provides

- sensible age shape and rating factor variation, plus
- in-built allowance for time trends and
- adjustment for mortality bias (opposite)

** Standardised by Index of Multiple Deprivation (IMD) 2015 decile*
Source: ONS data with Aon calculations

Survey data

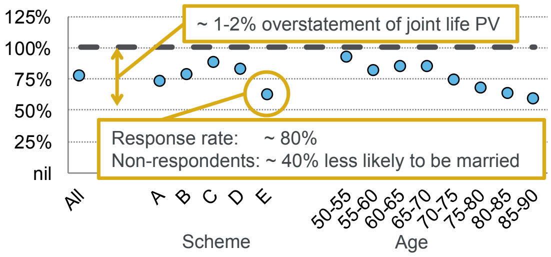
Married members are typically *more likely* to respond to a survey than unmarried members

This means that survey non-respondents

- may be biased toward not being married
- can't just be valued using the survey average

Survey non-respondent bias ratio

= *P(married* | non-respondent) ÷ P(married* | respondent)*



** According to trace status*

Survey data

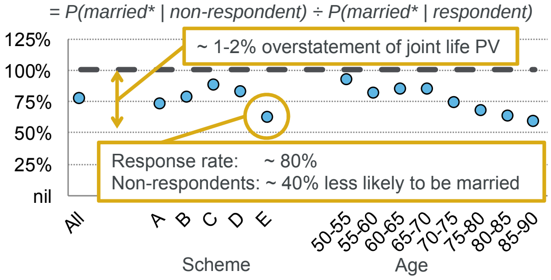
Married members are typically *more likely* to respond to a survey than unmarried members

This means that survey non-respondents

- may be biased toward not being married
- can't just be valued using the survey average

Survey non-respondent bias ratio

= P(married | non-respondent) ÷ P(married* | respondent)*

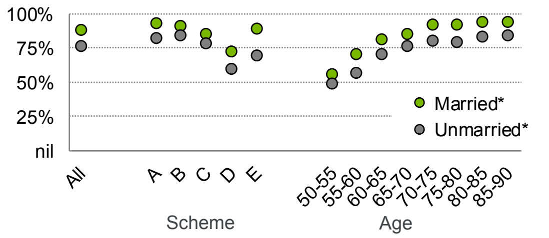


Solution

Estimate scheme- and exercise-specific non-respondent bias by

- modelling *relative* response rates (unmarried vs married) across scheme and survey characteristics
- based on large dataset of members who have been surveyed *and* traced

Survey response rate



** According to trace status*

An alternative approach?

Modelling dependant proportions is hard!

Ideal solution

- Scheme-specific (unlike national averages)
- Relates to current lives being valued (unlike experience data)
- Non-invasive (so we expect little 'non-respondent' bias)
- Identifies legal spouses vs wider dependants (to deal with alternative eligibility definitions)
- Objective standardised output (so easy to test and compare across schemes)
- Relatively quick and cheap

Does member tracing fit the bill?

2. Member tracing and LexisNexis® Risk MSP

18 November 2019

What we do

We leverage four main components to provide end-to-end solutions that help customers assess risk and opportunity associated with industry-specific problems.

Vast Data Resources

We maintain over six petabytes of content comprising billions of public and proprietary records.

Big Data Technology

We designed our own proprietary super-computing platform, HPCC Systems®, enabling us to process at very high speeds.

Linking & Analytics

We use our own unique identifier, LexID®, together with a proprietary linking technology. Our patented linking and clustering method is the engine behind many of our products.

Industry-Specific Expertise & Delivery

The people in our businesses have deep industry experience and expertise – we employ professionals that worked in the industries we serve, so they have walked in the shoes of our customers.

Customer-Focused Solutions

We connect the dots between billions of public records and transactions, resulting in actionable information our customers use to advance their goals.

How does LexisNexis® Risk MSP work?

Our proprietary logic identifies the scheme member within our consumer universe…

…and systematically profiles every co-habitant to identify a spouse

Should a spouse be identified, a name and date of birth (if available) will be output, enabling longevity to be calculated on both member and spouse. If the member is not married, the product will publish household composition segmentation, including for example 'living as married', 'living with family'.

How?

Accurate data linking is crucial for maintaining customer records

Rules Linking

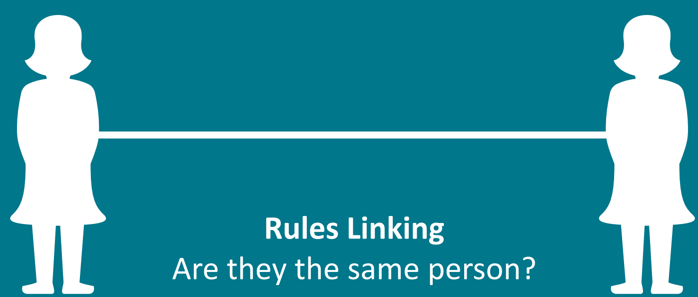
When comparing two records, a combination of individual rules are used.

If all rules are matched then a link is established between the two records.

Statistical Linking

When comparing two records a weight is assigned to each matched field value based on how statistically common that value is across the data universe.

The total combined weight of each matched field value determines whether there is a link between the two records.



Rules Linking

Are they the same person?

Rules Linking: Are they the same person?

Record 1	Record 2
Forename: Sarah	Forename: Sarah
Surname: Barker	Surname: Jones
Postcode: AB1 1BB	Postcode: IV1 2CC
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974
Telephone: 012 3344 5566	Telephone: 012 4455 6677

Rules Linking: Are they the same person?

Record 1	Match	Record 2	Match
Forename: Sarah	✔	Forename: Sarah	✔
Surname: Barker	✖	Surname: Jones	✖
Postcode: AB1 1BB	✖	Postcode: IV1 2CC	✖
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔
Telephone: 012 3344 5566	✖	Telephone: 012 4455 6677	✖

Linking Rule:
For this example a Forename, Surname and Date of Birth match were required to establish a link between both records

✖ **Not the same person**

Rules Linking: Are they the same person?

Record 1		Record 2	
Forename: Sarah	✔	Forename: Sarah	✔
Surname: Barker	✖	Surname: Jones	✖
Postcode: AB1 1BB	✖	Postcode: IV1 2CC	✖
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔
Telephone: 012 3344 5566	✖	Telephone: 012 4455 6677	✖

This method can work well but isn't perfect…

John Smith	John Smith
Aberdeen	Plymouth
22/03/1980	22/03/1980

Two individuals that happen to share the same common name and date of birth could be incorrectly identified as the same person

Rules Linking: Are they the same person?

Record 1		Record 2	
Forename: Sarah	✔	Forename: Sarah	✔
Surname: Barker	✖	Surname: Jones	✖
Postcode: AB1 1BB	✖	Postcode: IV1 2CC	✖
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔
Telephone: 012 3344 5566	✖	Telephone: 012 4455 6677	✖

We could add extra rules to do things like check for Electoral Roll overlap, ensure continuity in household composition or a plausible shift in socio-demographic attributes…

Before too long the rules can become very complicated

Statistical Linking

Are they the same person?

Statistical Linking: Are they the same person?

	Record 1		Record 2	
Forename:	Sarah	✔	Sarah	✔
Surname:	Barker	✖	Jones	✖
Postcode:	AB1 1BB	✖	IV1 2CC	✖
Date of Birth:	17/04/1974	✔	17/04/1974	✔
Telephone:	012 3344 5566	✖	012 4455 6677	✖

Forename = Sarah

Common Forename in UK

Match Weight:
Low

Weight
5

Approximate number of Sarah's in the UK:
387,000
Match Weight:
5

Approximate number of Veronica's in the UK:
33,000
Match Weight:
10

Statistical Linking: Are they the same person?

Record 1		Record 2	
Forename: Sarah	✔	Forename: Sarah	✔
Surname: Barker	✖	Surname: Jones	✖
Postcode: AB1 1BB	✖	Postcode: IV1 2CC	✖
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔
Telephone: 012 3344 5566	✖	Telephone: 012 4455 6677	✖

Surname Mismatch

Penalty applied

Penalty Weight:
Moderate

Weight
5
-7

Penalty weights are determined by the likelihood of a field value varying between records…

Postcode Mismatch = Low Penalty: **-3**
Surname Mismatch = Moderate Penalty: **-7**
Date of Birth Mismatch = High Penalty: **-16**

Statistical Linking: Are they the same person?

Record 1		Record 2		Weight
Forename: Sarah	✔	Forename: Sarah	✔	5
Surname: Barker	✖	Surname: Jones	✖	-7
Postcode: AB1 1BB	✖	Postcode: IV1 2CC	✖	-3
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔	21
Telephone: 012 3344 5566	✖	Telephone: 012 4455 6677	✖	-2

Total Weight:
14

✖ Not the Same Person

*(Minimum score of **40** is required to match both records)*

Statistical Linking: Are they the same person?

Record 1	Record 2		Record 3		Weight
Forename: Sarah	Forename: Sarah	✔	Forename: Sarah	✔	5
Surname: Barker	Surname: Jones	✖	Surname: Barker	✖	-7
Postcode: AB1 1BB	Postcode: IV1 2CC	✔	Postcode: IV1 2CC	✔	23
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974	✔	21
Telephone: 012 3344 5566	Telephone: 012 4455 6677	✖	Telephone: 012 3344 5566	✖	-2

Total Weight:

40

✔

Same Person

*(Minimum score of **40** is required to match both records)*

Statistical Linking: Are they the same person?

Record 1	Record 2	Record 3
	Consolidated Record	
Forename: Sarah	Forename: Sarah	Forename: Sarah
Surname: Barker	Surname: Jones	Surname: Barker
Postcode: AB1 1BB	Postcode: IV1 2CC	Postcode: IV1 2CC
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974	Date of Birth: 17/04/1974
Telephone: 012 3344 5566	Telephone: 012 4455 6677	Telephone: 012 3344 5566

Statistical Linking: Are they the same person?

Record 1	Consolidated Record
Forename: Sarah	Forename: Sarah
Surname: Barker	Surname: Jones
Postcode: AB1 1BB	Postcode: IV1 2CC
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974
Telephone: 012 3344 5566	Telephone: 012 4455 6677

Statistical Linking: Are they the same person?

Record 1	Consolidated Record
Forename: Sarah	Forename: Sarah
Surname: Barker	Surname: Barker / Maiden Name: Jones
Postcode: AB1 1BB	Postcode: IV1 2CC
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974
Telephone: 012 3344 5566	Telephone (Most Recent): 012 3344 5566

Statistical Linking: Are they the same person?

Record 1		Consolidated Record			Weight
Forename: Sarah	✔	Forename: Sarah		✔	5
Surname: Barker	✔	Surname: Barker	Maiden Name: Jones	✔	10
Postcode: AB1 1BB	✖	Postcode: IV1 2CC		✖	-3
Date of Birth: 17/04/1974	✔	Date of Birth: 17/04/1974		✔	21
Telephone: 012 3344 5566	✔	Telephone (Most Recent): 012 3344 5566		✔	24

Total Weight: **57**

✔ Same Person

*(Minimum score of **40** is required to match both records)*

Statistical Linking: Are they the same person?

Record 1	Consolidated Record		
Forename: Sarah	Forename: Sarah		✔
Surname: Barker	Surname: Barker	Maiden Name: Jones	✔
Postcode: AB1 1BB	Postcode: IV1 2CC		↻
Date of Birth: 17/04/1974	Date of Birth: 17/04/1974		✔
Telephone: 012 3344 5566	Telephone (Most Recent): 012 3344 5566		✔

Linked records are consolidated to form a single, constantly evolving, customer profile

Consolidated Record

Forename: Sarah	
Surname: Barker	Maiden Name: Jones
Postcode (Most Recent): AB1 1BB	Postcode (Previous): IV1 2CC
Date of Birth: 17/04/1974	
Telephone (Most Recent): 012 3344 5566	

Consolidated customer profiles are assigned a unique identifier called a LexID

LexID: 9794931

Forename: Sarah	
Surname: Barker	Maiden Name: Jones
Postcode (Most Recent): AB1 1BB	Postcode (Previous): IV1 2CC
Date of Birth: 17/04/1974	
Telephone (Most Recent): 012 3344 5566	

18 November 2019 39

Institute and Faculty of Actuaries

AON Empower Results

LexisNexis RISK SOLUTIONS

3. Tracing in practice

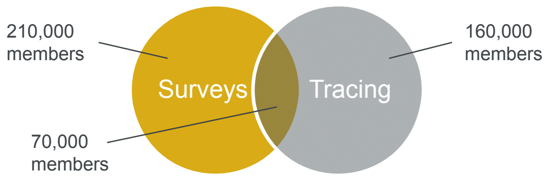
18 November 2019

Demographic Horizons dependants dataset

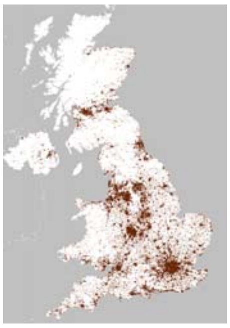
1. Over 300,000 members from 30 pension schemes spanning 2011-2019

2. Wide coverage of UK by
- geographic region
- age, sex and pension amount
- current vs future pensioners
- legal spouses vs wider financial dependants

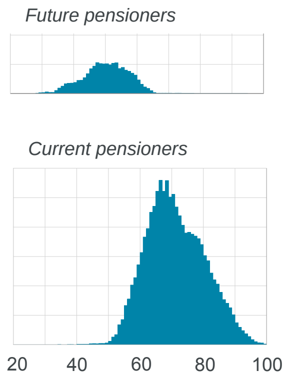
3. Multiple data sources for robust inference:



Geographic distribution



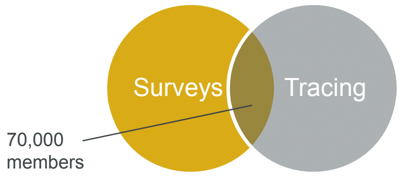
Pension distribution by age



Aon’s testing of MSP

Huge volume of recent survey data *with contemporaneous tracing* – enables statistically credible testing of MSP

High level results:
- over 95% of those with a 'Married' trace code *were actually married*
- over 95% of those with a 'Living alone' trace code *were unmarried and had no partner*
- over 90% of those recorded as married in our survey data also had a 'Married' trace code



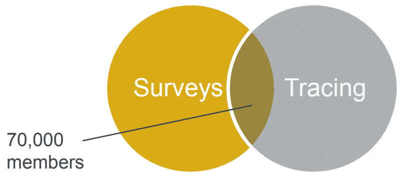
Aon’s testing of MSP

Huge volume of recent survey data *with contemporaneous tracing* – enables statistically credible testing of MSP

High level results:

- over 95% of those with a 'Married' trace code *were actually married*
- over 95% of those with a 'Living alone' trace code *were unmarried and had no partner*
- over 90% of those recorded as married in our survey data also had a 'Married' trace code

The tracing results from MSP correlate strongly with true marital status



Aon’s testing of MSP

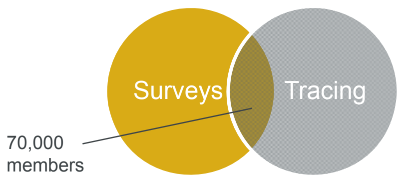
Huge volume of recent survey data *with contemporaneous tracing* – enables statistically credible testing of MSP

High level results:

- over 95% of those with a 'Married' trace code *were actually married*
- over 95% of those with a 'Living alone' trace code *were unmarried and had no partner*
- over 90% of those recorded as married in our survey data also had a 'Married' trace code

The tracing results from MSP correlate strongly with true marital status

But we still need to interpret the codes in order to apply them…

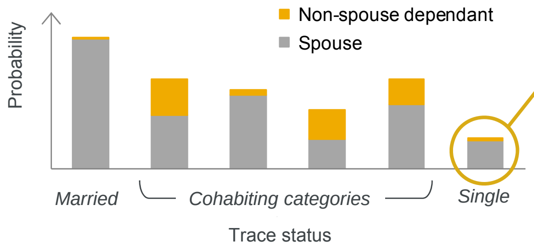


Interpreting the trace codes

Aon has calibrated a 'mapping' (which depends on age and sex) to estimate the probability of an individual being married, or having an unmarried partner, based on the trace code returned by MSP.

This mapping matrix is used within the Demographic Horizons model for members who have been traced.

Dependant probability by trace status



Two key benefits:

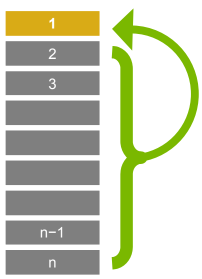
1. Corrects for noise in the tracing, eliminating any residual bias.

For example:
- Small – but non-negligible – proportion of married individuals traced as single (i.e. 'Living alone')
- Important to assign a small probability of being married to that code rather than assuming it is nil

2. Allows us to assess the proportion of members with wider financial dependants (i.e. married *and* unmarried partners), with allowance for the various cohabiting codes returned by MSP.

Performance testing – cross-validation

For each scheme in the dataset:
- re-fit the mapping *excluding* that scheme, then
- test how closely the mapped MSP results for the scheme agree with the *observed* survey data



How accurately is the dependant proportion for this scheme…

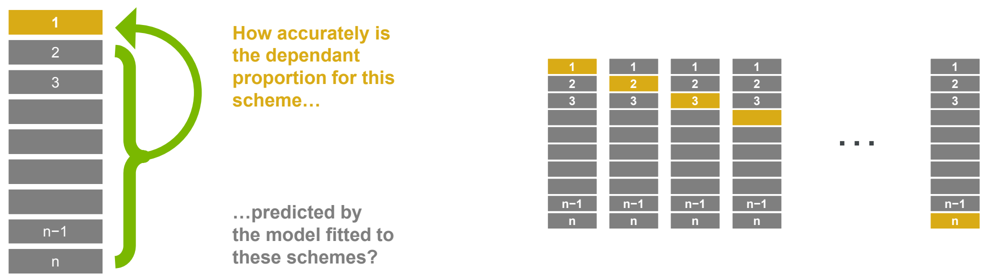
…predicted by the model fitted to these schemes?

Performance testing – cross-validation

For each scheme in the dataset:

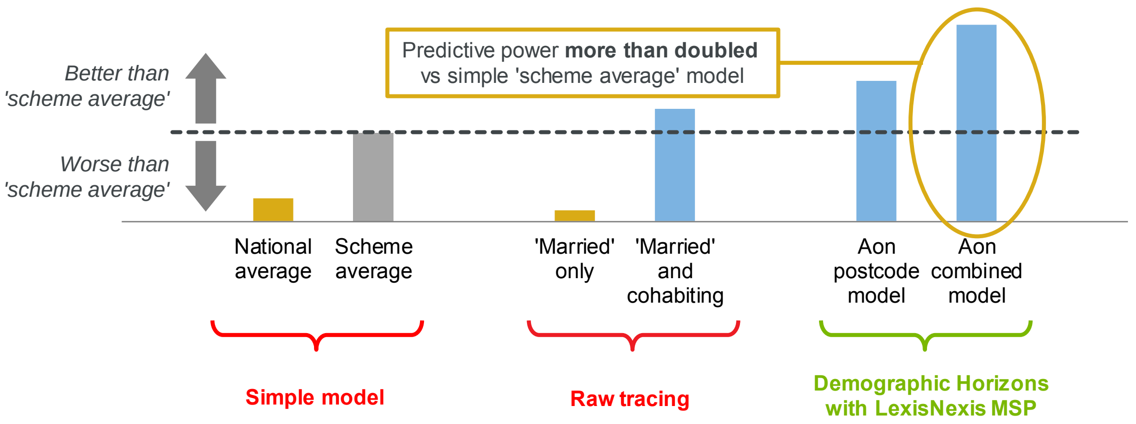
- re-fit the mapping *excluding* that scheme, then
- test how closely the mapped MSP results for the scheme agree with the *observed* survey data

Repeat across schemes to test actual *predictivity*, without cheating



Performance testing – results

Predictive power* relative to 'scheme average'



** Measured as the inverse mean square error of the model's predictions when tested across 40 datasets (by scheme and pension tranche)*

What makes a good tracing service?

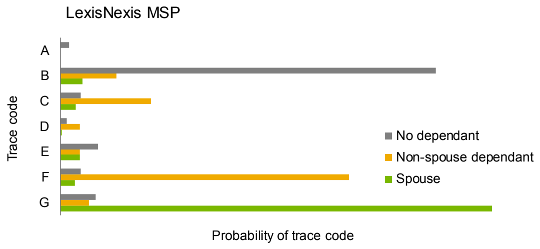
Key features:

- Degree of *differentiation* in codes
- *Information content* of tracing
- *Low bias* in 'Unknown' trace code
- *Stability* of the service

What makes a good tracing service?

Key features:

- Degree of *differentiation* in codes
- *Information content* of tracing
- *Low bias* in 'Unknown' trace code
- *Stability* of the service



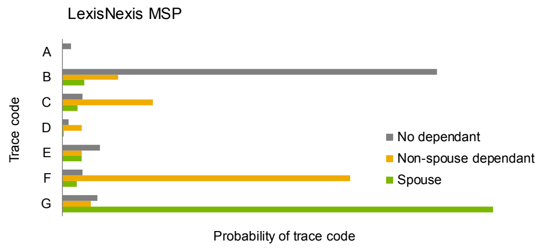
What makes a good tracing service?

Key features:

- Degree of *differentiation* in codes
- *Information content* of tracing
- *Low bias* in 'Unknown' trace code
- *Stability* of the service

High information content

- Highly polarised distribution
- Just one large bar for each dependant status



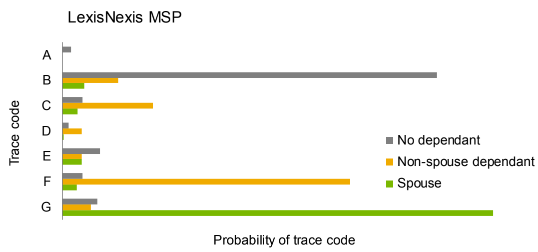
What makes a good tracing service?

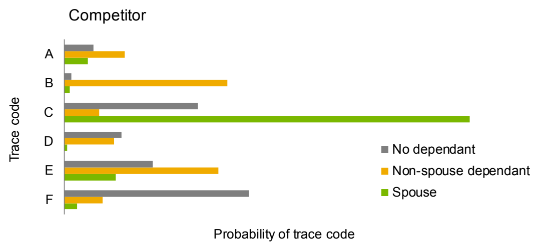
Key features:

- Degree of *differentiation* in codes
- *Information content* of tracing
- *Low bias* in 'Unknown' trace code
- *Stability* of the service

High information content

- Highly polarised distribution
- Just one large bar for each dependant status





What makes a good tracing service?

Key features:

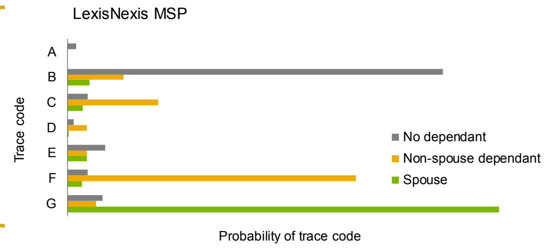
- Degree of *differentiation* in codes
- *Information content* of tracing
- *Low bias* in 'Unknown' trace code
- *Stability* of the service

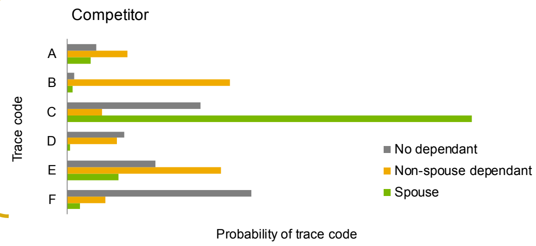
High information content

- Highly polarised distribution
- Just one large bar for each dependant status

Lower information content

- Less polarised distribution
- e.g. married individuals most likely to return code C (green bar)
  … but code C is also a fairly common outcome for individuals with no dependant (grey bar)





Summary

Dependant proportions *matter*

- Increasingly material
- Growing focus of price assessment
- Increasing sophistication required to deal with different data sources, eligibility scope and slicing approaches

A *robust* modelling framework is critical

- Calibrate to *actual* pension scheme / annuitant data
- Capture age shape, time trends and socio-economic variation
- Correct for survey non-respondent bias and mortality bias

When mapped correctly, tracing can be *highly predictive*

- But you do need to interpret the codes…
- … And not all tracing services are equal!

Questions

Comments

The views expressed in this presentation are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this presentation and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this presentation.

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this presentation be reproduced without the written permission of the authors.