MAYERFELD CONSULTING PRACTICUM

Natsal survey case study

Did the change differ by sex?

Two independent survey waves were compared.
The analysis tests the survey-by-sex interaction
while accounting for age and survey weights.

DATA

Natsal-2 and Natsal-3

MODEL

Negative Binomial regression

RESULT

IRR 0.802 | p < 0.001

DATA PREPARATION

Comparable inputs before modelling

01

Clean

Retain exact numeric counts; treat special responses as missing.

02

Align

Restrict both waves to the common age range of 16–44.

03

Standardise

Harmonise labels and combine the two survey waves.

04

Weight

Rescale respondent weights within each survey wave.

Final modelling sample: 21,465 respondents

MODEL SELECTION

Why Negative Binomial regression

POISSON REFERENCE

Count outcome, but equidispersion failed.
Dispersion ratio: 40.85
AIC: 387,935

NEGATIVE BINOMIAL

Handles overdispersed count data.
Substantially better fit.
AIC: 138,582

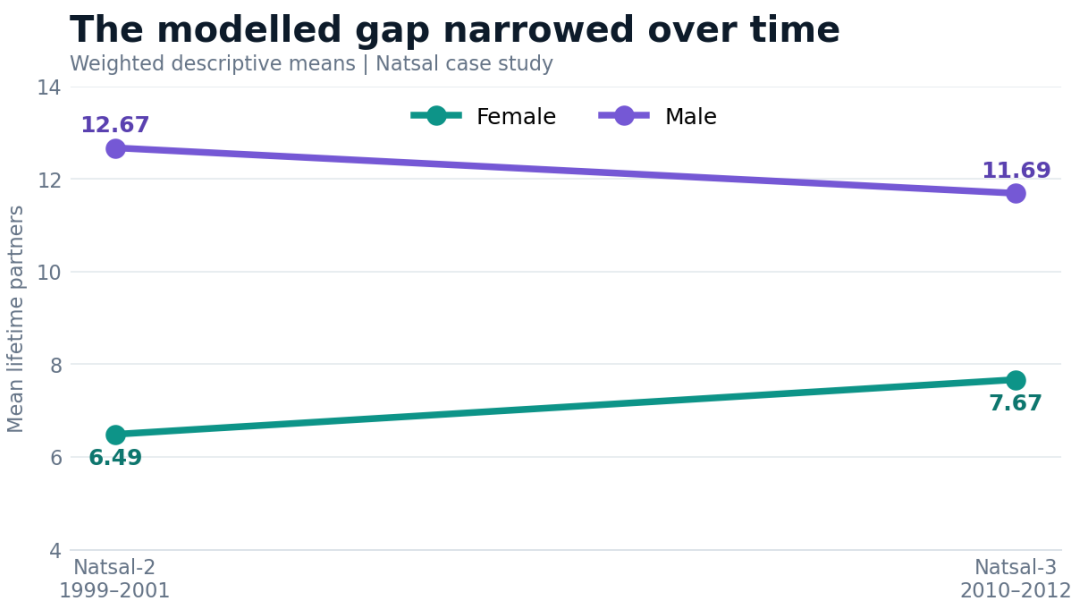
Model formula

```
count ~ survey + sex + survey:sex + age
```

The interaction term is the main test of whether change differs between groups.

RESULT AND INTERPRETATION

The gap narrowed between survey periods



KEY RESULT

Interaction IRR: 0.802
95% CI: 0.754–0.854
$p < 0.001$

INTERPRET WITH CARE

Association, not causation.
Weighted GLM approximation;
not a full complex-survey model.