NatCen
Social Research that works for society

Transitioning a F2F survey to an online panel

Findings from a parallel run

Contents

- Background
- Differences in estimates
- Trying to understand the differences
  - Measurement (mode)
  - Sample composition
- Summary & conclusions

Background

1

Context

- Department for Transport (DfT) interested in transitioning survey questions from large F2F surveys (BSA/NTS) to online (panel) surveys
  - Speed, cost, ‘future-proofing’
- Aim of understanding what impact (if any) it might have on survey estimates/trend data
- 2017 field parallel run on BSA and NatCen panel
- Mostly attitudinal, but some behavioural questions on travel
  - Congestion, use of cars/bikes, willingness to use alterative transport to lower emissions, policies to reduce air travel, use of mobile phones while driving, speed cameras

Questionnaire content

- Mostly attitudinal, but some behavioural questions on travel
  - Impact of congestion
  - Use of cars/bikes
  - Willingness to use alterative transport to lower emissions
  - Policies to reduce air travel
  - Use of mobile phones while driving
  - Drink-driving
  - Speed cameras

Fieldwork summary

	BSA 2017	NCPanel Aug17
Fieldwork dates	July-October 2017	August-September 2017
Mode	Mixed: Face-to-face (F2F) & Paper self-completion (PAPI)	Sequential-mixed: Web & Telephone (Tel)
Sample size	2,963 (F2F) 783-848 (PAPI)	2,159
Response rate	45% (F2F) 36%-38% (PAPI)	15%

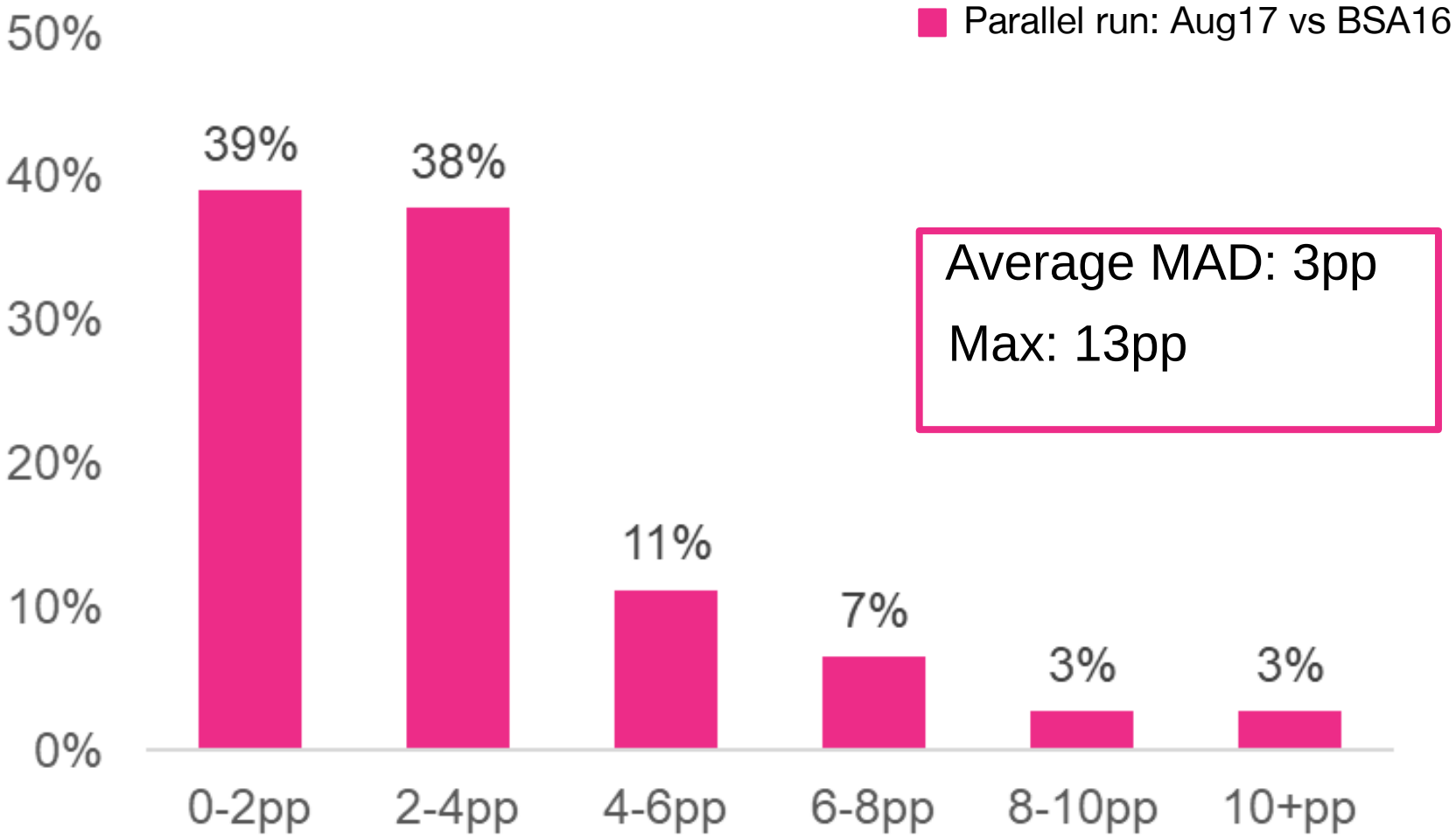
Differences in estimates

2

Analysis approach

- Compare weighted estimates for BSA and NCPanel
- 44 variables/questions, 259 categories
- Look at differences in percentage-point estimates at category-level
- Lots of different variables so will use some summary variables
  - Mean absolute difference (MAD)
  - Maximum absolute difference
  - % categories with differences 0-2, 2-4, 4-6, 6-8, 10+ pp

Size of difference in point estimates



Patterns in differences?

- *Broad* trend of NCPanel more concerned about environmental impact of travel, but less support for enforcement of road rules (using mobile phones/speed cameras)
  - But varies from variable to variable in direction and strength
- Less difference in ‘behavioural’ questions
  - E.g. Whether drive, number of cars owned, whether fly
- Greater difference with PAPI than F2F
  - F2F: 16% differences >4pp
  - PAPI: 26% differences >4pp

Understanding differences: Measurement (mode)

3

Changes in design of questions

- Spontaneous, hidden, and prompted answer categories
  - Cannot use ‘spontaneous’ answers in PAPI/web
  - Cannot used ‘hidden’ answers in PAPI
- In web we randomly reverse answer scales to reduce primacy effects
  - Not possible (easy) in F2F/PAPI where use printed showcards/questionnaires
  - Interaction between question type & device type

Do they explain the differences? (1)

DK/REF options

- F2F vs Web/CATI
  - No differences of 1+pp in % selecting DK/REF options
- PAPI vs Web/CATI
  - Web: consistently <1% selecting DK/REF
  - PAPI: average 3% ‘Can’t choose’ and 3% left blank

Spontaneous options

- F2F vs Web/CATI
  - Average of 3% selected spontaneous categories that were dropped
  - Average 3pp difference when shown up-front

Do they explain the differences? (2)

Randomising answer scale direction

- Saw expected primacy effect in web survey
- No indication ‘reversed’ categories show more or less difference
  - No differential impact for F2F/PAPI

Measurement: summary

- ‘Designed’ elements of measurement do have some impacts, *but…*
  - Impact varies between questions
  - Don’t explain level of difference
- Rest may be explained by…
  - Other elements of measurement (e.g. impact of social desirability bias, panel conditioning, etc.)
  - Sample composition…

Understanding differences: Sample

4

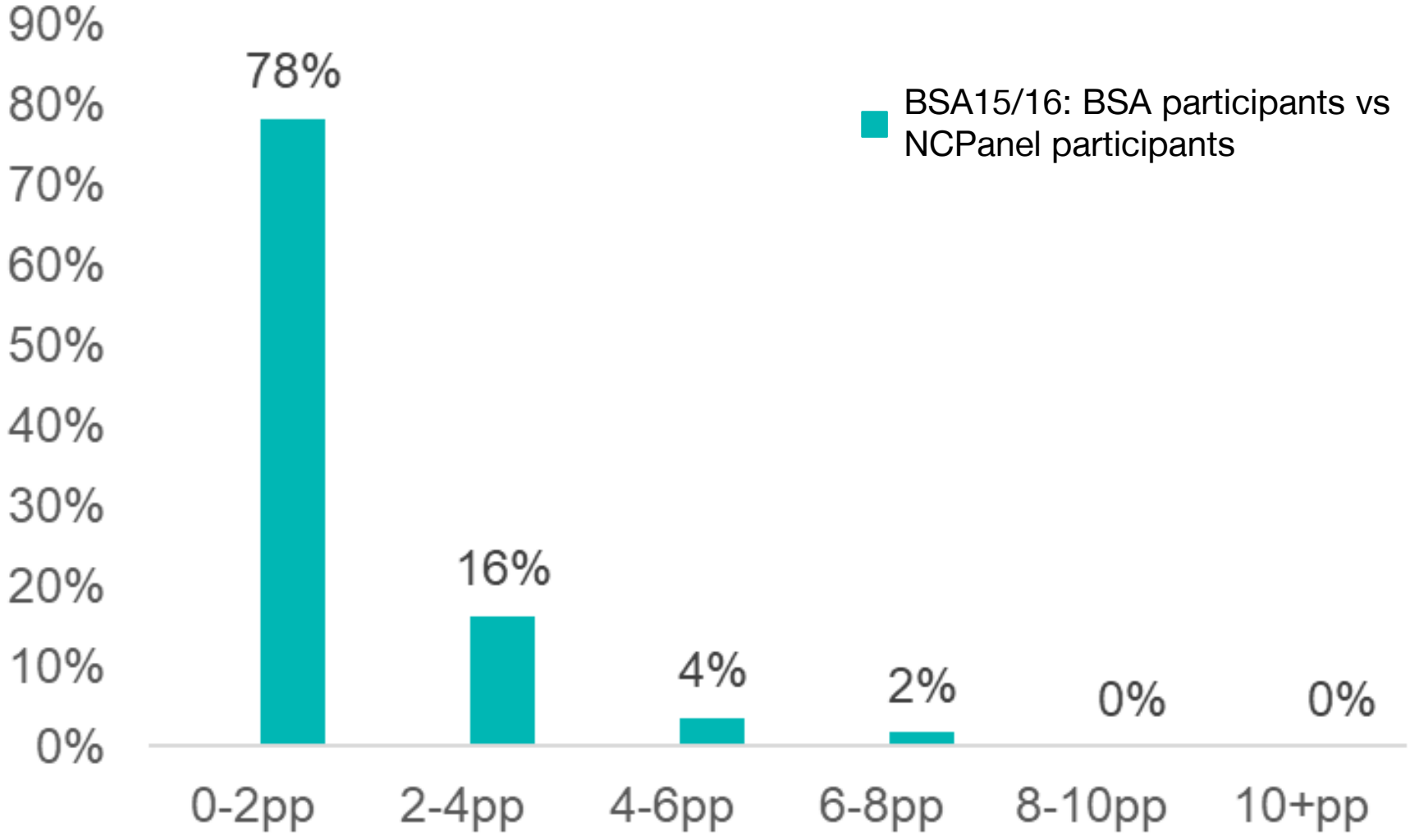
Differences in sample composition

- Non-response may introduce bias to survey samples if some types of people are more likely to take part than others
- Participating NCPanel sample has a different profile to BSA sample (Jessop, 2018).
  - E.g. more educated, higher social grade
- Most of this is addressed in non-response weights
- *However* – impacts of ‘hidden’ bias/bias within weighting categories?

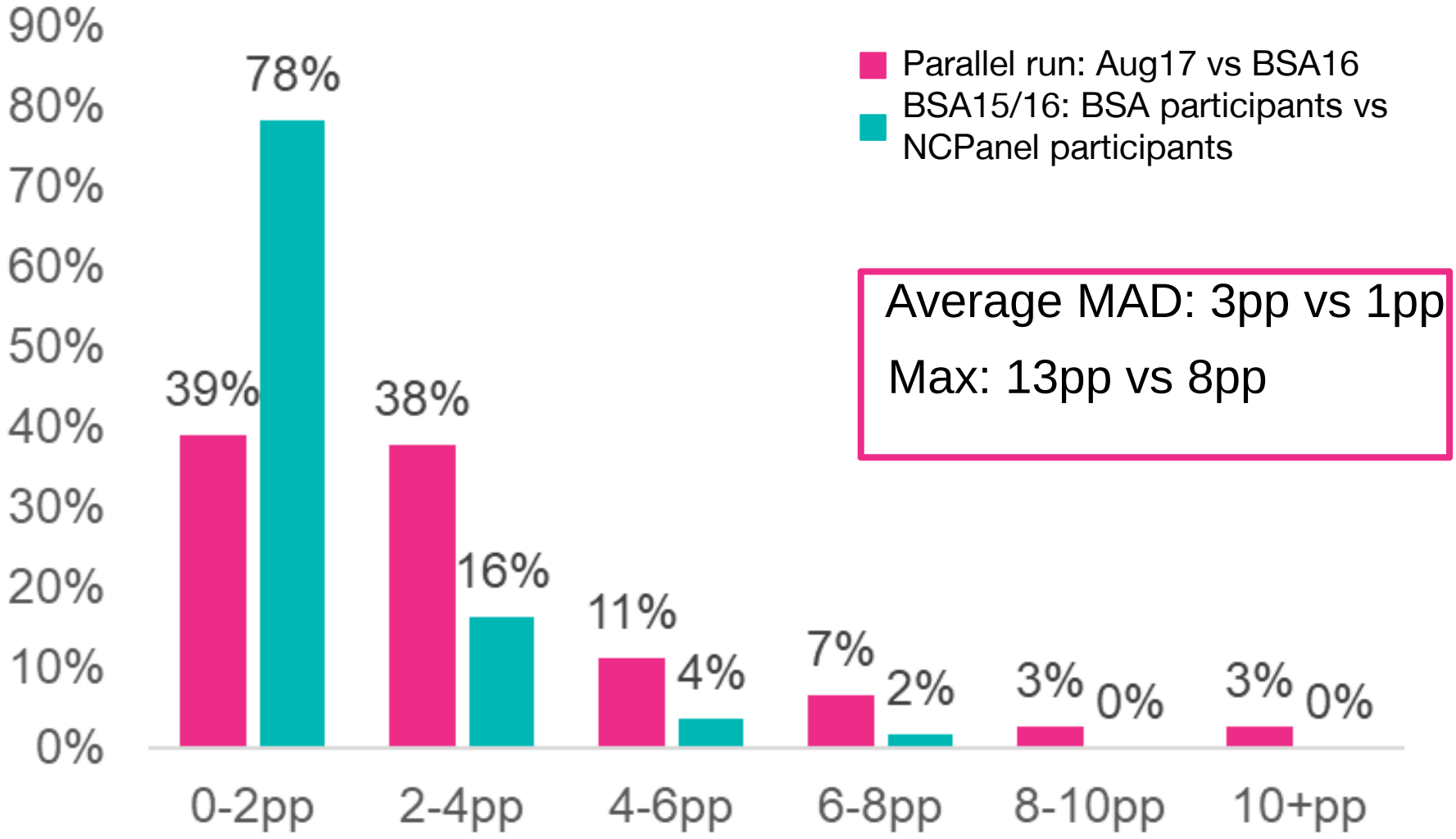
Analysis approach

- NCPanel members were asked the same questions in their original BSA interviews (2015, 2016)
- Compare weighted BSA 2015/2016 estimates for full sample to weighted estimates using just those who took part in Aug17 NCPanel
  - Allows separation of sample composition from issues of measurement

Size of difference in point estimates



Size of difference in point estimates



Average MAD: 3pp vs 1pp

Max: 13pp vs 8pp

Summary & conclusions

5

Summary

- Parallel run found that while many estimates were comparable, there were some substantial differences
- Marginal non-response bias does not explain these differences
- ‘Designed’ changes in measurement sometimes make some impact, on some questions
- Much of the difference remains unexplained
  - Social desirability? Panel conditioning? Panel age? …?

Conclusions

- Parallel runs important when transitioning surveys to understand impact of design change on estimates
- Importance of taking the time to properly consider the impacts & how to mitigate them
  - Tension with goals of time/cost savings
- Should evaluation of impacts of changes happen at question-level?
  - Intersection of many traits
  - Challenge for developing ‘best practice’

Thanks!

Curtis Jessop

Research Director

curtis.jessop@natcen.ac.uk

@CurtisJessop

NatCen
Social Research that works for society