THE NEW ZEALAND SOCIO-ECONOMIC INDEX (NZSEI) FOR THE 2018 CENSUS

Presented by Barry Milne and Natalia Boven
Other project contributors: Nichola Shackleton, Liza Bolton, Andrew Sporle, Inny Kang

IDI Disclaimer

The results in this presentation are not official statistics, they have been created for research purposes from the Integrated Data Infrastructure (IDI) managed by Statistics New Zealand.

The opinions, findings, recommendations and conclusions expressed in this presentation are those of the author(s) not Statistics NZ, or the University of Auckland.

Access to the anonymised data used in this study was provided by Statistics NZ in accordance with security and confidentiality provisions of the Statistics Act 1975. Only people authorised by the Statistics Act 1975 are allowed to see data about a particular person, household, business or organisation and the results in this presentation have been confidentialised to protect these groups from identification.

Careful consideration has been given to the privacy, security and confidentiality issues associated with using administrative and survey data in the IDI. Further detail can be found in the Privacy impact assessment for the Integrated Data Infrastructure available from www.stats.govt.nz.

Presentation Outline

- Introduction to the NZSEI – history, uses, construction
- Description of 2018 Census challenges and possible consequences
- Aims of 2018 NZSEI project
- Methodology
- Results
  - comparison with previous NZSEI scales
  - assessing impact of use of alternative data sources
  - subgroup analyses
  - validation against outcomes
- Conclusions

Introduction to the NZSEI

- The New Zealand Socio-economic Index (NZSEI) is a measure of socio-economic status/position (SES/SEP) for individuals, based on their occupation.
- SES is a multidimensional construct which captures (or attempts to capture) the social and material resources individuals, families and households have access to.
- SES often used interchangeably with ‘social class’ or ‘social status’, though they are not the same concept (though are clearly related)
  - *No implied relationship to the labour market or work conditions (social class)*
  - *No implied perceived social superiority (social status).*

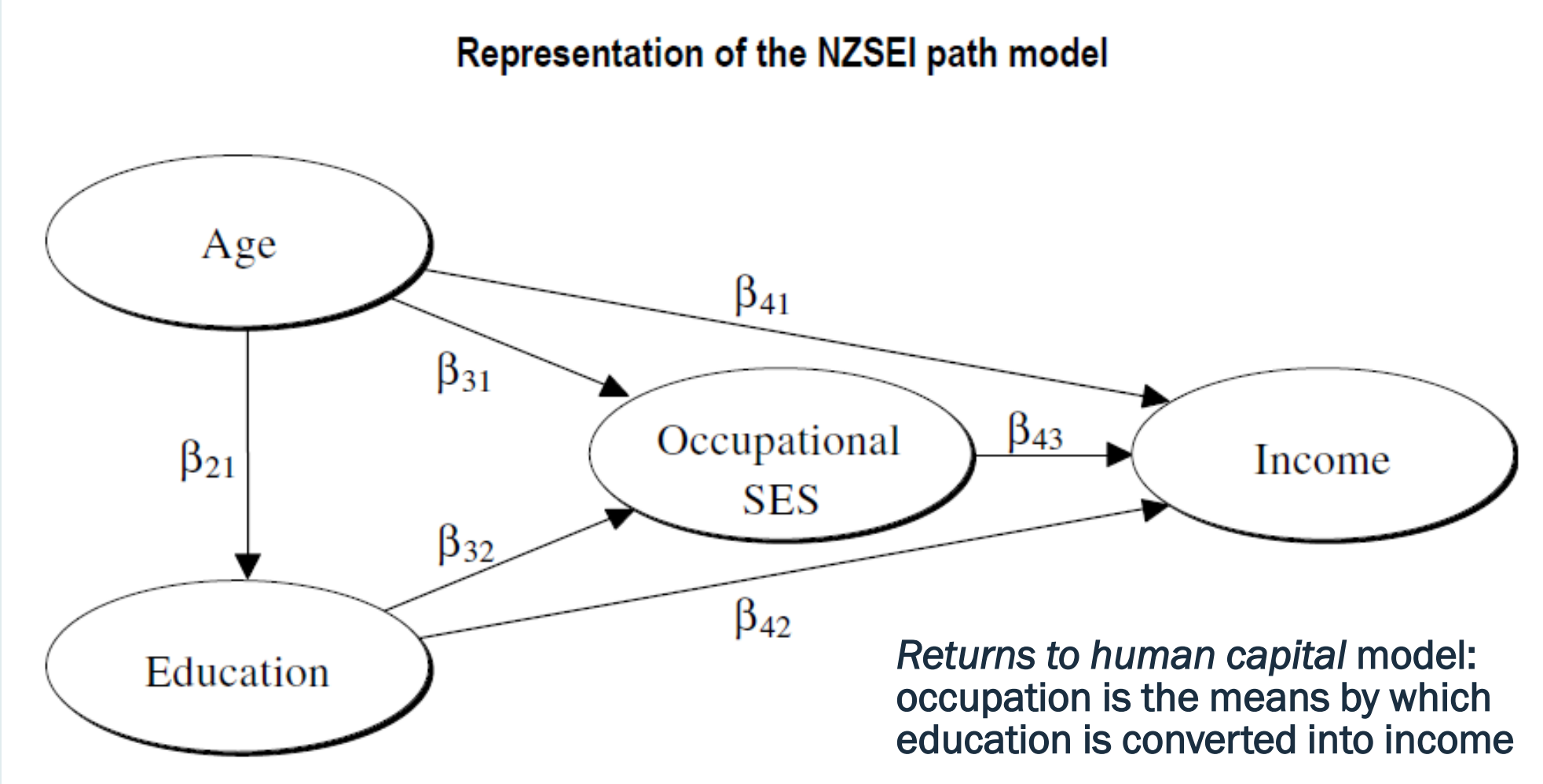
Why measure SES?

- Research
  - *Can test hypotheses about the impact of unequal distribution of opportunities, advantages, resources and power on*
    - Health, wellbeing, life choices, use of services, crime
    - Confounding the impact of other risk factors
  - *Can investigate SES stability/mobility, within one's life and inter-generationally*
- Describing populations
- Funding allocation
  - *Social and health services are sometimes funded (in-part) based on the socio-economic characteristics of the areas that they serve.*

Measurement of SES

- All measures have their advantages and drawbacks
  - *Income – face validity, often recorded administratively; often reluctantly reported, known under-reporting (self-employed)*
  - *Education – stable past a certain age; but inversely associated with age*
  - *Deprivation/living standards measures*
    - Area-based – proven validity, easily coded, summarises multiple adversities; individuals within area may differ, address may mislead
    - Individual-based – proven validity, summarises multiple adversities; need specific questionnaire
  - *Occupation – readily recalled, often recorded, proven validity; coding not straightforward, individuals with the same occupation may differ; how to code those not in workforce?*
- Not the case that one 'best' captures SES; each might be seen as complementary to others. No reason to just focus on one (possible to combine)

NZSEI – Theory & Construction



Returns to human capital model: occupation is the means by which education is converted into income

NZSEI – Construction

- Derive SES scores which equate to an optimal weighting of education and income, corrected for age
- Scale scores to be from 10 (low SES) – 90 (high SES)

ANZSCO	Occupation	NZSEI
253	Medical Practitioners	90
134	Education, Health and Welfare Services Managers	78
212	Media Professionals	70
612	Real Estate Sales Agents	61
451	Personal Service and Travel Workers	49
334	Plumbers	42
732	Delivery Drivers	30
811	Cleaners and Laundry Workers	19
832	Packers and Product Assemblers	10

2018 Census challenges

- Big issue with constructing the NZSEI using the 2018 Census is the 2018 Census data…
- Around 1 in 6 New Zealanders didn't complete a census form, and this was differential across age, ethnicity, geography.

FIXES

1. Use the IDI to find the people who didn't fill out the census and add them in.
2. Where data for a variable wasn't available from Census 2018, get data from alternative data sources :
   - *Census 2013*
   - *Administrative data*
   - *Imputation*

Alternative data sources

- For key variables for the NZSEI, nearly 20% of data came from other sources, and if one key variable had data from other sources, the others also typically did.

Socioeconomic variable	2018 Census	Admin data	Imputation	2013 Census	No info
Occupation	80.4%		19.6%		
Total income	81.9%	16.6%	1.5%		
Secondary school qualifications	82.4%	4.1%		8.0%	5.5%
Post school qualifications	81.1%	7.3%		6.1%	5.4%

Figures are presented for the working population aged 21-69.

Aims

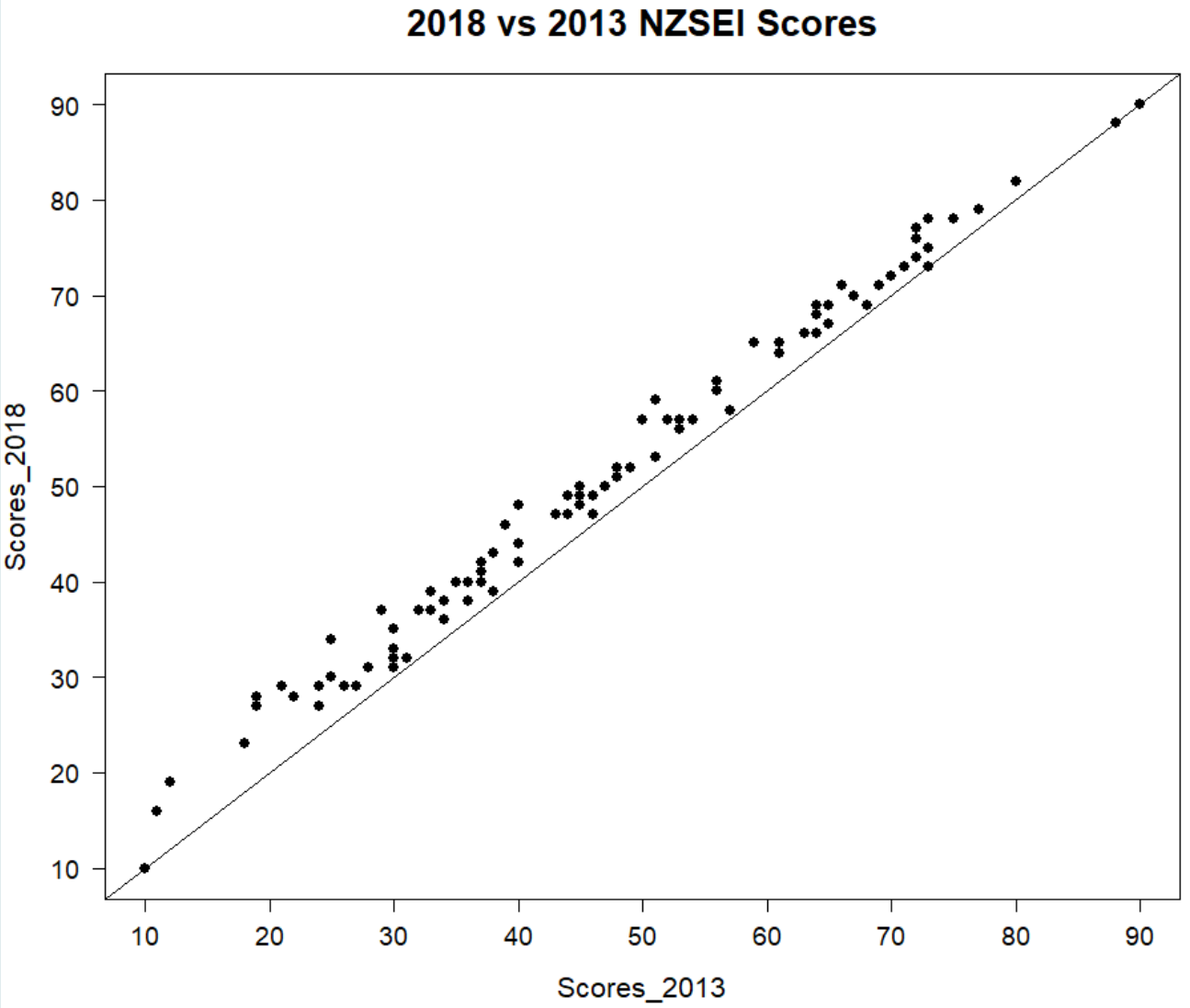
- Given the widespread use of alternative data sources for the first time, we wanted to find out

1. Does the NZSEI using 2018 data pattern occupations similarly to previous versions of NZSEI?
2. Would the NZSEI look different if data from only those who completed a census were used?
3. Does the NZSEI pattern occupations similarly across men and women, and different major ethnic groups?
4. Does the NZSEI show evidence of 'construct validity'. E.g., are there NZSEI gradients across:
    a. *NZDep (an area-based socioeconomic measure)*
    b. *Smoking, a health behavior known to have a strong socio-economic gradient*
    c. *Hospitalizations, an objective health outcome known to have a strong socio-economic gradient*
    d. *Self-rated health, a subjective health outcome known to have a strong socio-economic gradient*
    e. *Life satisfaction, a subjective wellbeing outcome known to have a strong socio-economic gradient*

Methodology for constructing NZSEI-18

- Used 2018 Census data restricted to working adults aged 21-69 (n ≈ 2.2 million)
- Key variables: occupation, education, income and age
  - occupation coded to the minor group (3-digit) level of ANZSCO which has 97 occupations
  - highest qualification converted to years of education
  - total annual income inflated for part-time workers (< 30 hours per week), extreme values removed and applied log transformation
- NZSEI algorithm uses mean values for age, education and income at the occupational level
- Scores scaled between 10 and 90 and transformed to have mean ≈ 50 (raised to ^0.5 for 2013, ^0.55 for 2018).

Results for Aim 1: comparison with previous scales



Path	2018	2013	2006
β32 (education-SEP)	0.545	0.570	0.572
β43 (SEP-income)	0.306	0.313	0.299

Correlation with 2013 scores: r=0.99.

Aim 2: Effect of alternative data

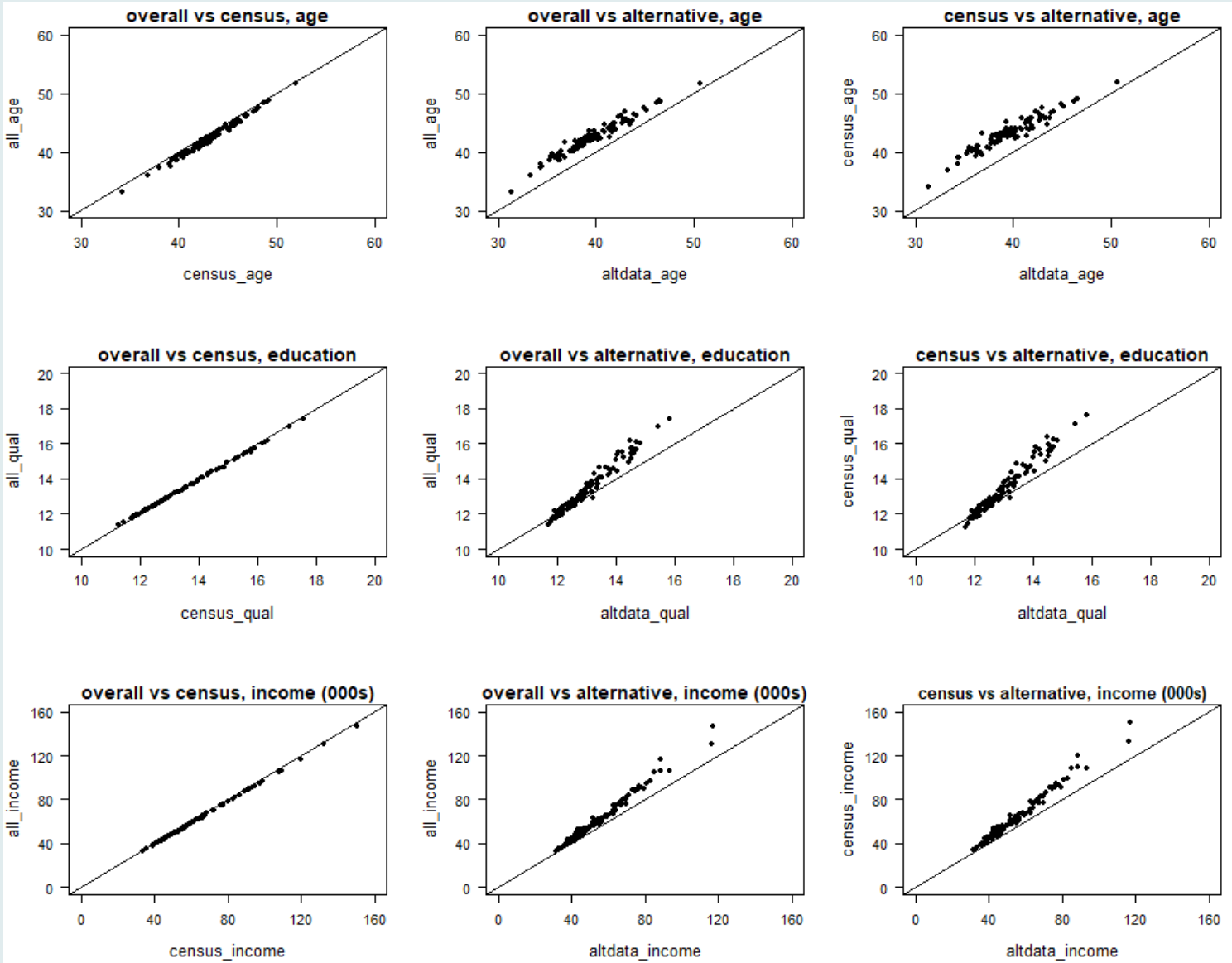
- Assigned records to one of two mutually exclusive cohorts:
  - occupation & income from 2018 individual Census forms (79%),
  - occupation &/or income from alternative data sources (21%).
- Checked compositional differences, betas, change in scaled scores and relationship with smoking and NZDep18.
- Also checked for patterns when broken down by Level 1 ethnic group.

Results for Aim 2: Effect of alternative data

Age

Education

Income



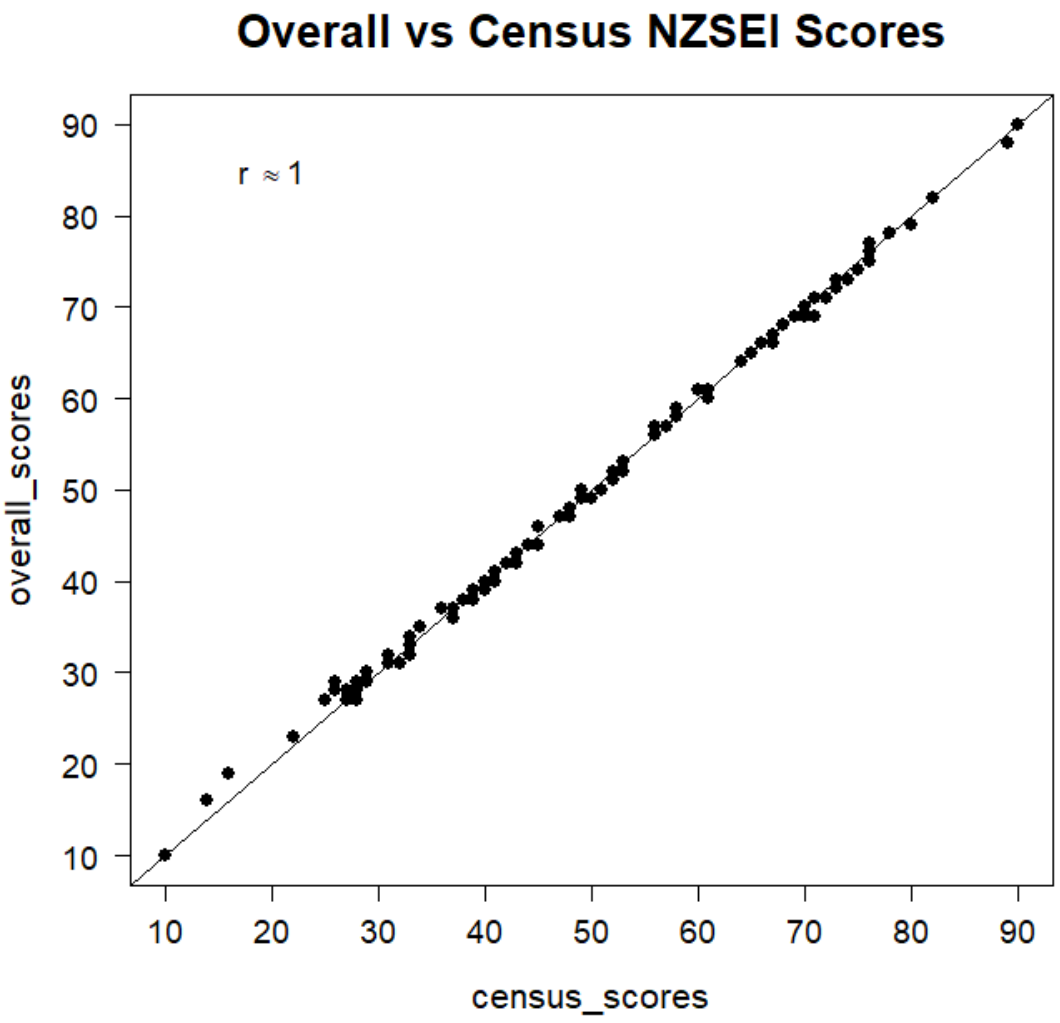
- Census and overall data set very similar
- Within the same occupational group, the alternative data sources cohort is younger, less educated, and earns less, on average.

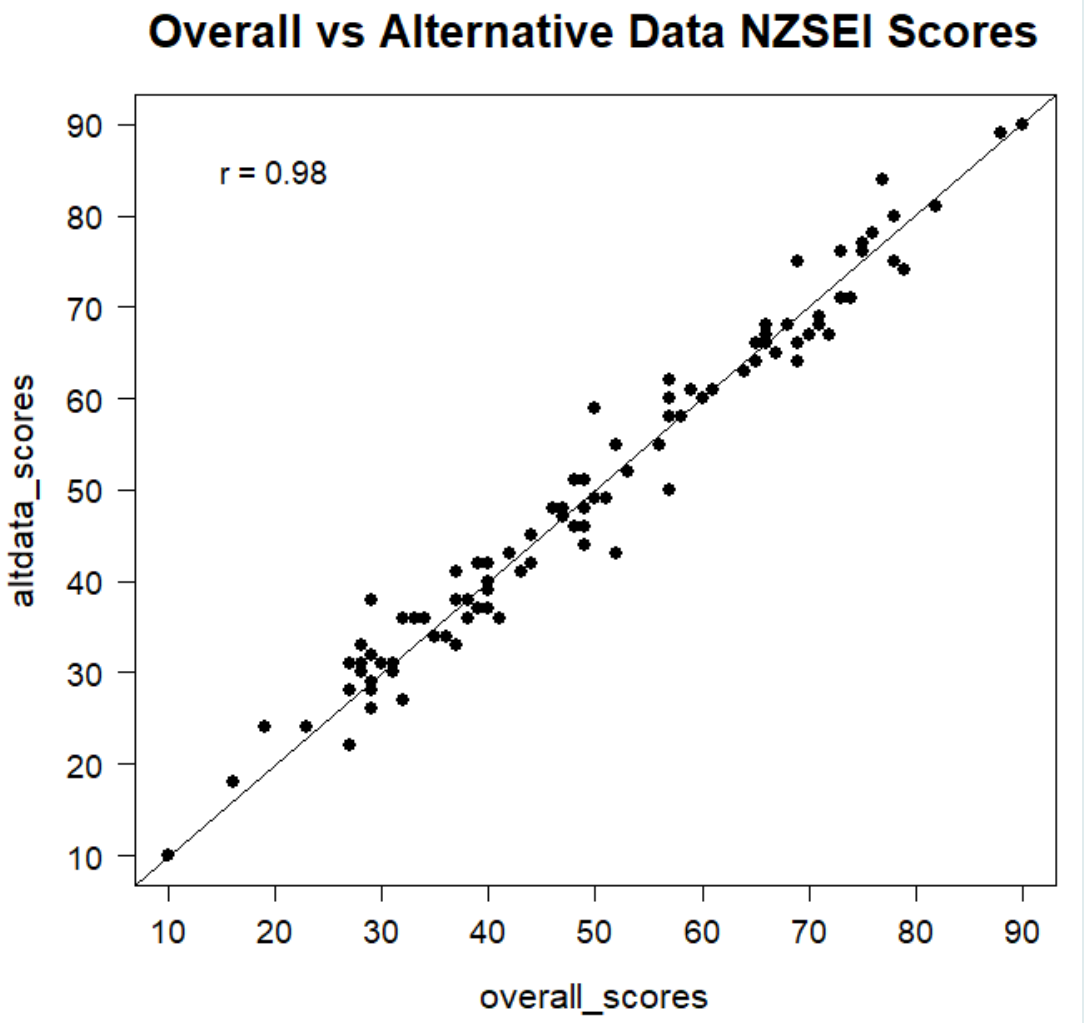
Results for Aim 2: Effect of alternative data

Occupation (major group)	% alternative data sources	NZSEI score
Managers	15.2	58
Professionals	14.1	74
Technicians and Trades Workers	22.0	41
Community and Personal Service Workers	23.3	43
Clerical and Administrative Workers	17.1	50
Sales Workers	22.5	43
Machinery Operators and Drivers	27.7	31
Labourers	30.0	23

- Proportion of records from alternative data sources varies by occupation.
- Ranges from 14%-30% to major group level and 10%-30% for minor group occupations.

Results for Aim 2: Effect of alternative data



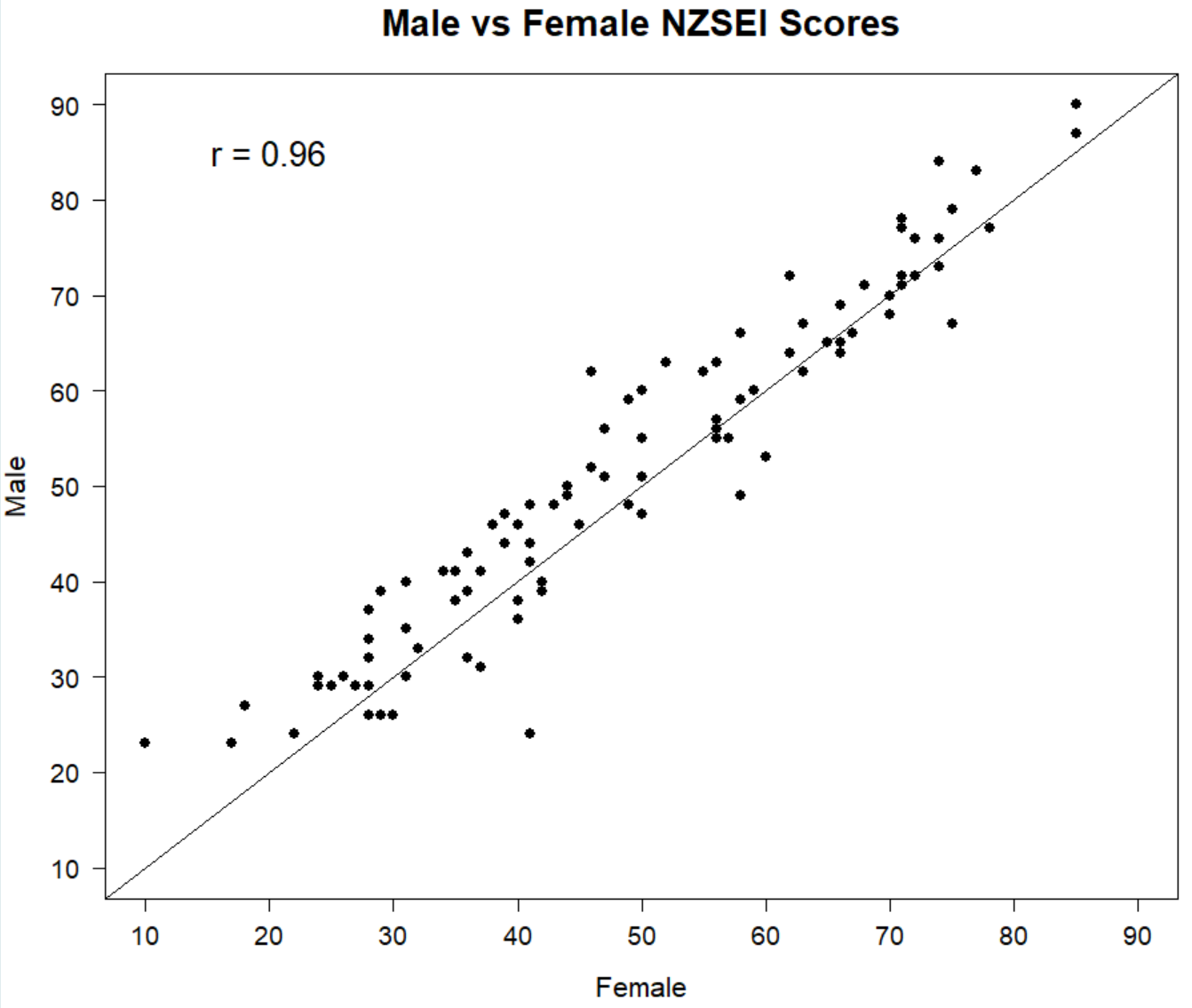


Results for Aim 2: Effect of alternative data

Path	Overall	Census only	Alternative data sources
β32 (education-SEP)	0.545	0.570	0.368
β43 (SEP-income)	0.306	0.309	0.252
Ratio β32/β43	1.8	1.8	1.5

- Paths for alternative data sources are weaker, especially for education – SEP path.
- Census and alternative data sources scales, including those broken down into Level 1 ethnic groups, are significantly related to regular smoking behaviour and NZDep2018 score.

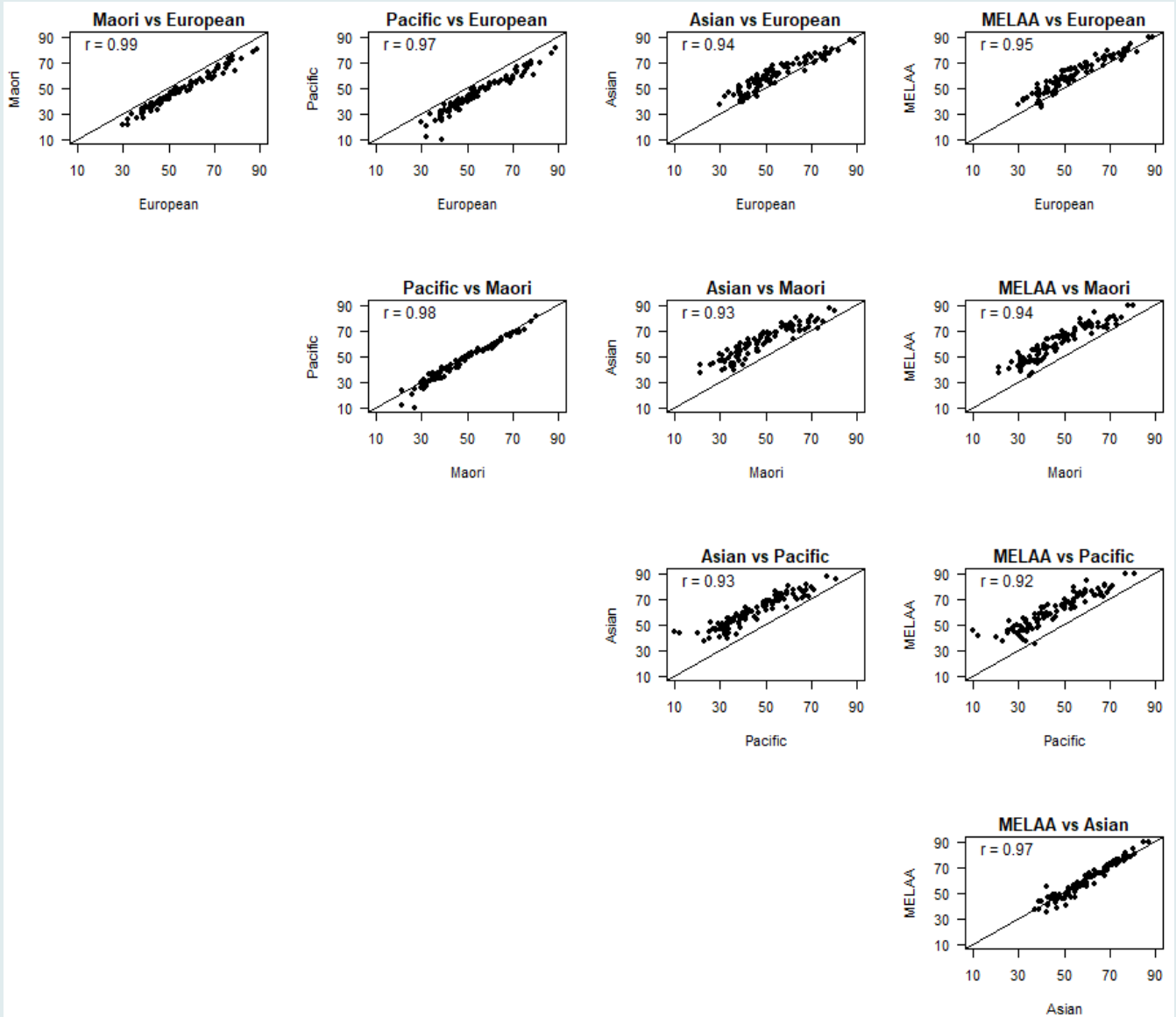
Results for Aim 3: subgroup analyses



Mean scores
Men – 50.9
Women – 48.0

Note: scores have a transformation applied so the mean ≈ 50 across subgroups (not within subgroups).

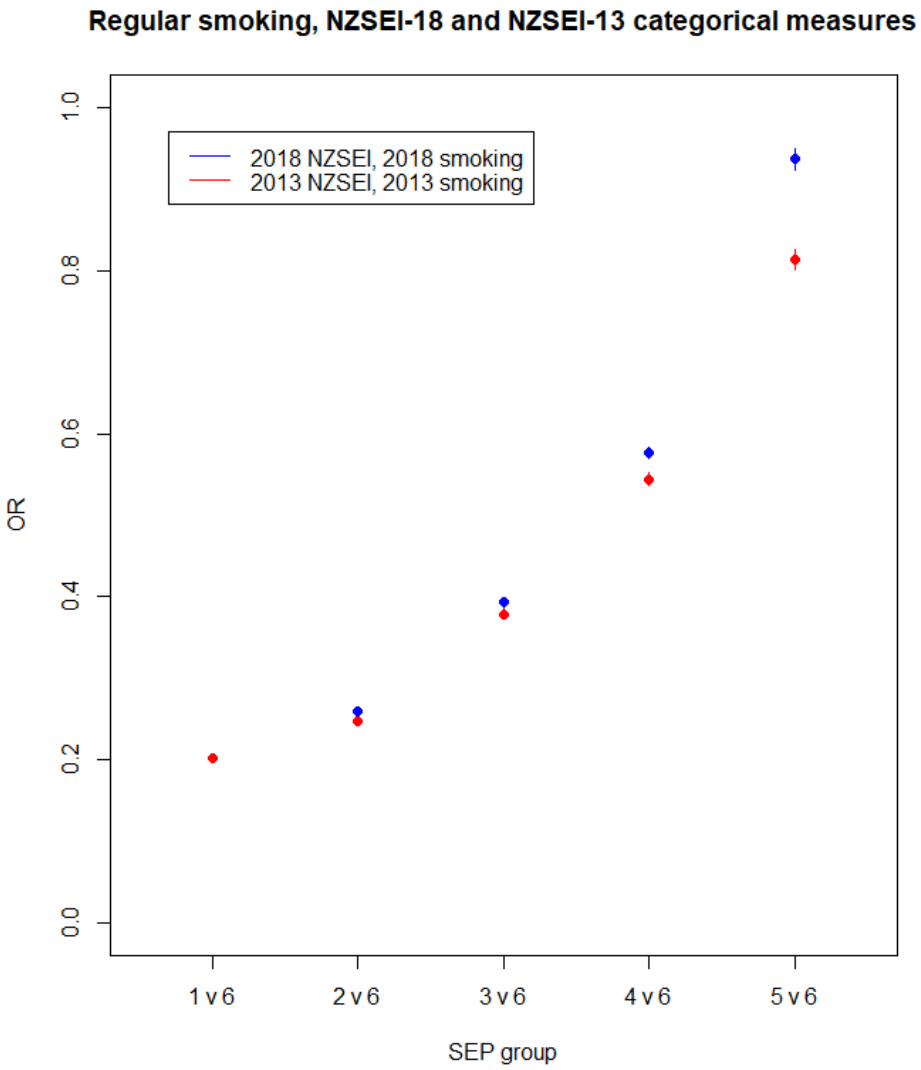
Results for Aim 3: subgroup analyses

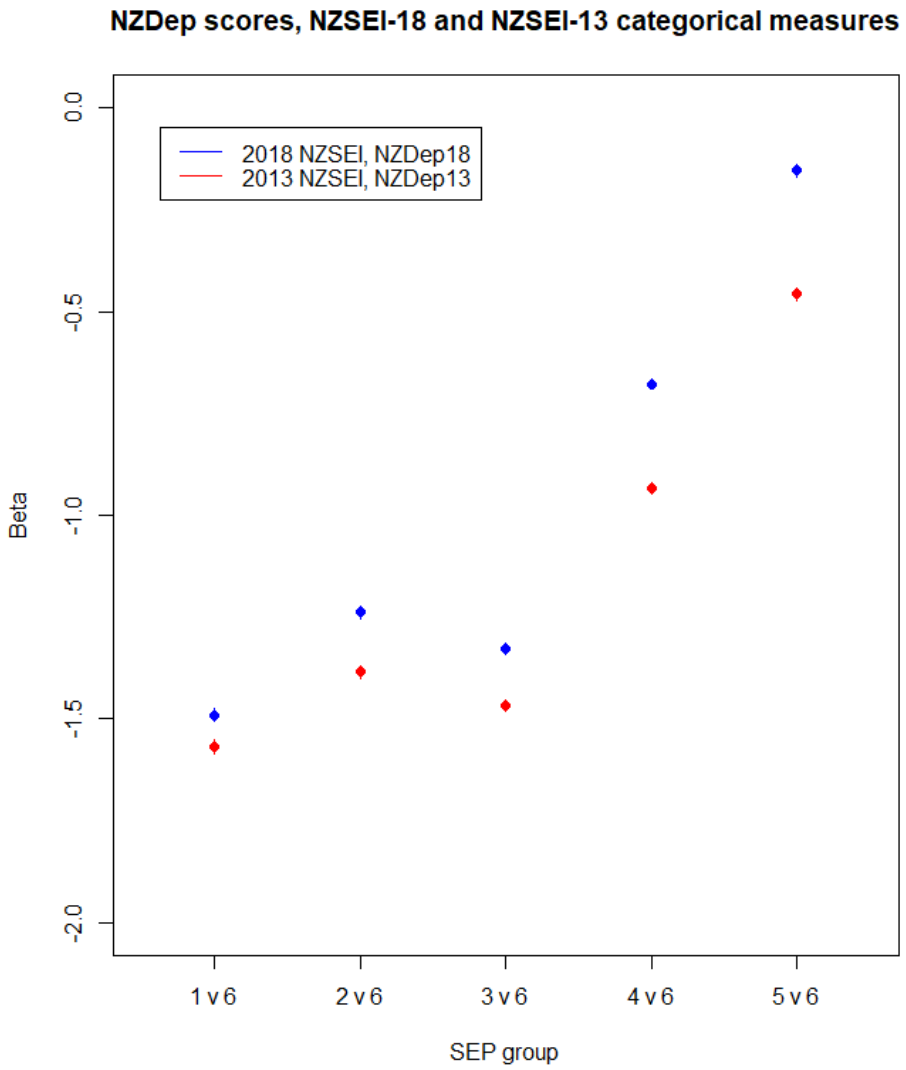


Mean scores:
Māori – 47.2
Pacific – 45.0
Asian – 60.1
MELAA – 60.0
European 54.4

Note: scores have a transformation applied so the mean ≈ 50 across subgroups (not within subgroups).

Results for Aim 4: construct validity

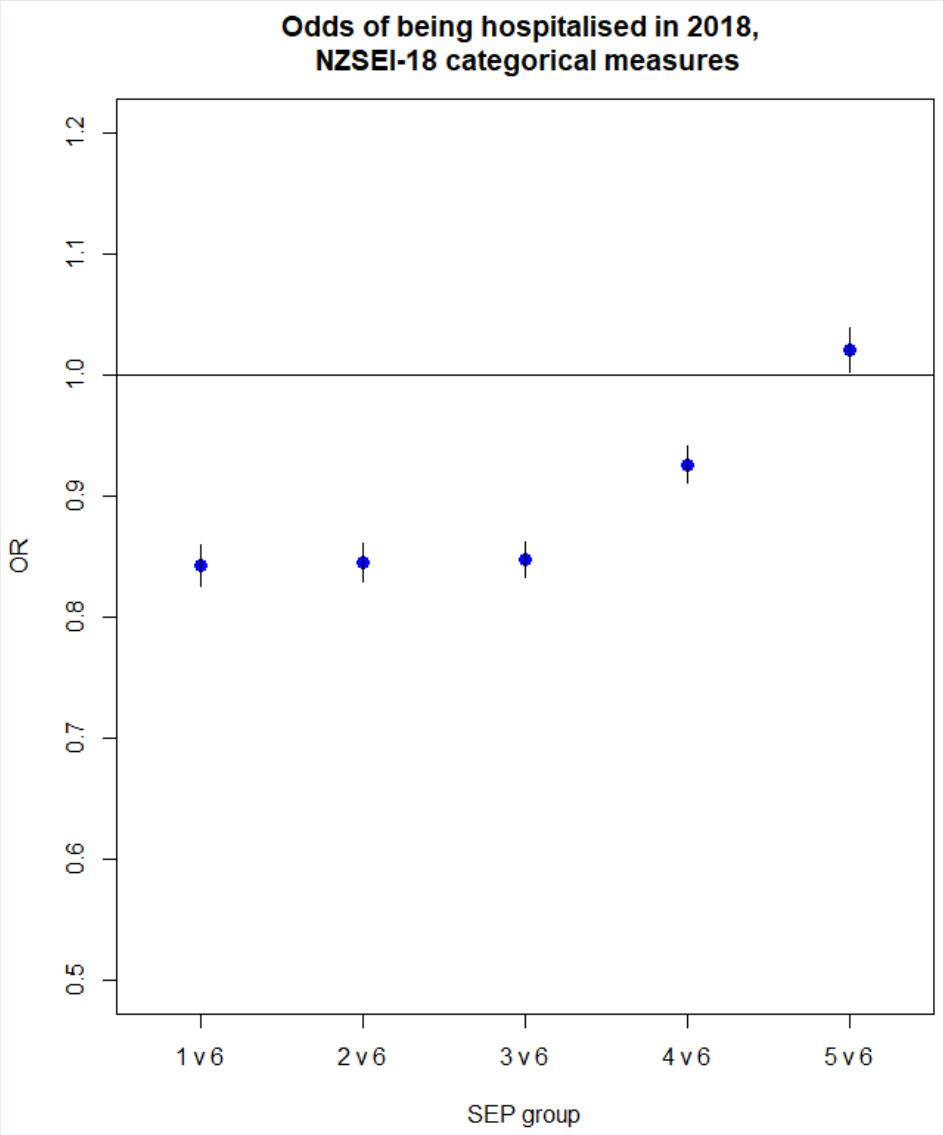


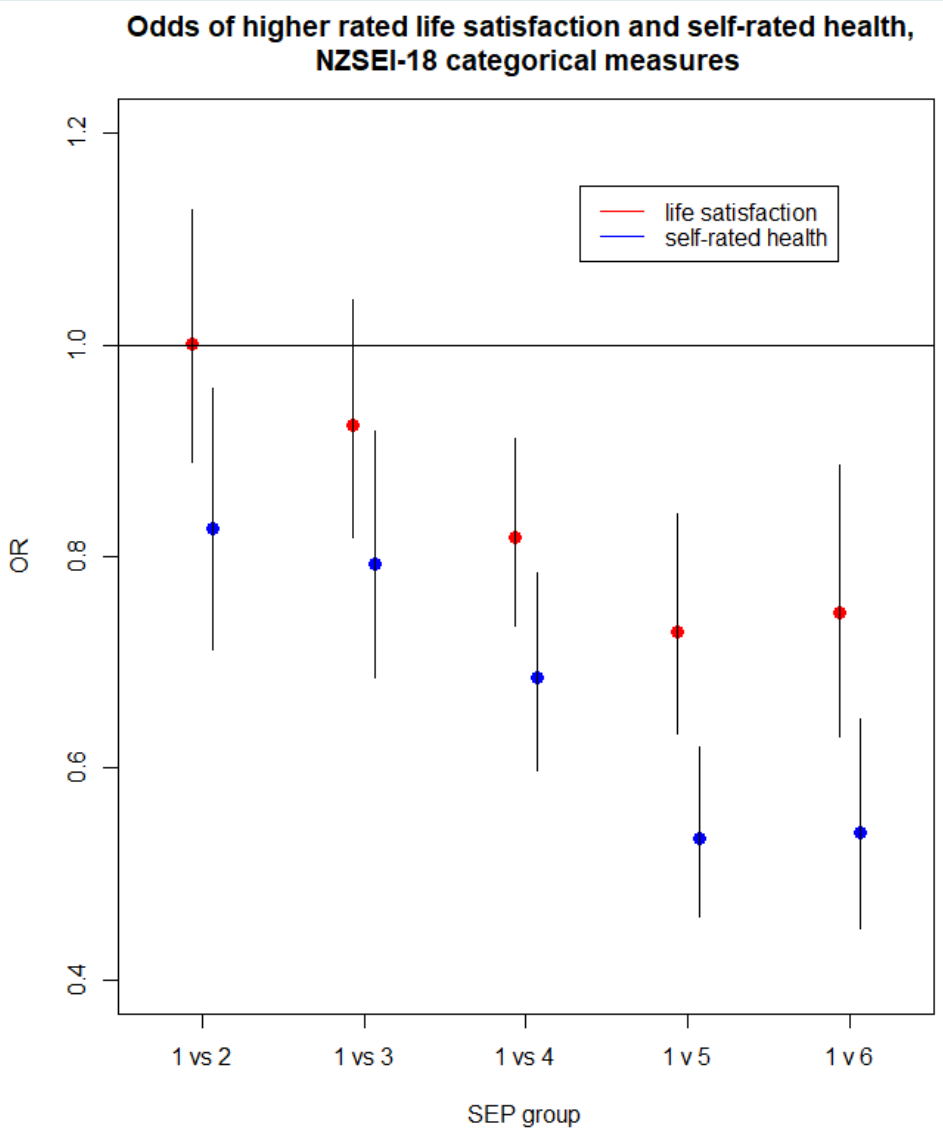


Adjusted for age group*, gender, and the following Level 1 ethnic groups; Māori, Pacific, Asian, MELAA, European.

* 2013 results adjusted for continuous age

Results for Aim 4: construct validity





Adjusted for age group, gender, and the following Level 1 ethnic groups; Māori, Pacific, Asian, MELAA, European.

Conclusions

- ❖ Despite the extensive use of alternative data sources, the 2018 NZSEI appears to be a valid measure of occupation-based socioeconomic position.
- The 2018 scale and path betas are similar to 2013
- The results for the overall and Census cohorts are similar despite the overall cohort having about 1/5 records with supplemented occupation and/or income data
- Results for subgroups are in line with expectations
- The 2018 NZSEI validates against smoking, NZDep, hospitalisations, life satisfaction and self-rated health

Acknowledgements

- We would like to thank Nichola Shackleton, Liza Bolton and Andrew Sporle for their help with analysis and interpretation of the results
- We would like to acknowledge Statistics New Zealand for providing access to this data, and the IDI support teams for their help throughout this project. We would especially like to thank Inny Kang for her help with analyses during lockdown!
- We would also like to acknowledge the Public Policy Institute for allowing us to use their data lab.

- Thank you for listening!

References

2018 Census External Data Quality Panel. (2019). 2018 Census External Data Quality Panel Assessment of Variables. Wellington: Statistics New Zealand.

Atkinson J, Salmond C, Crampton P (2019). NZDep2018 Index of Deprivation, Interim Research Report, December 2019. Wellington: University of Otago

Fahy, K. M., Lee, A., & Milne, B. J. (2017). New Zealand socio-economic index 2013. Wellington: Statistics New Zealand.

Stats NZ (2019). *Data sources, editing, and imputation in the 2018 Census*. Retrieved from www.stats.govt.nz