Is there a role for subjective health measures?

Denzil G. Fiebig

What's the issue?

▸ Subjective measures of health routinely collected

  ▸ Survival probabilities **What is the percent chance you will live to be 75 or more?**

  ▸ Self-assessed health (SAH) **In general how would you rate your overall health? Excellent, very good, good, fair, poor**

  ▸ Patient-reported outcomes (EQ-5D)

    ▸ Health states defined across multiple dimensions – mobility, pain, …

    ▸ Value set determined by respondents trading off quantity & quality of life

▸ Concerns include cognitive burden, survey design, justification bias, measurement error & reporting heterogeneity

  ▸ **"The fact that some individuals may be prone to mis-reporting their disability or health status is well documented…"** Gannon (2009)

What's the issue?…

▸ **If have objective health are subjective measures needed?**

  ▸ At 2020 summit NSW Premier Morris Iemma said the health system **"…collects so much data we all get dizzy"**

▸ **"Big data" is more than the amount of data**

  ▸ Opportunity to combine data from multiple sources is possibly a more important element of the data deluge

    ▸ **"Patient reported outcomes and preferences drawn from stated preference methods are especially useful sources of data but even more so if they are used to complement existing sources of information."** Fiebig (2017)

  ▸ **"Economists have long been hostile to subjective data. Caution is prudent, but hostility is not warranted."** Manski (2004)

The case of SAH

- SAH ubiquitous
  - **"Self-reported health ... has become the most commonly used measure of health in social sciences study."** Mu (2014)
- Does SAH health measure health?
  - **"I'm not sure what self-assessed health measures but I do know it's not health."** Anonymous (economist)
  - What is its health content as distinct from overall quality of life (QoL)?
  - If health what domains of health?

Dorion et al. (2015) – SAH does measure health

- Basic hypothesis: **If SAH measures health it should predict future health outcomes**
  - Outcomes are health care utilization – so "objective"
    - Good measure with Australia's universal health coverage
- Employs rich data from 45 & Up Study
  - Survey of 260,000 NSW residents aged 45+
  - 45+ population incurs about 60% of Australia's health care expenditure
  - Survey linked to administrative (panel) data providing hospital separations, ED presentations, MBS & PBS claims

UNSW
THE UNIVERSITY OF NEW SOUTH WALES

Econometric methods

▸ Use prospective models

- ▸ Predicting future health outcomes (12 months after survey)

▸ **Utilization:** Future health use measures

- ▸ Binary logits for hospital admission, > 6 GP visits, …, varying controls to gauge impact on SAH estimates
- ▸ SAH reduced to 3 level scale with excellent/very good the base
- ▸ Baseline controls are socio-economic, demographic, self-reported & family health variables then add QoL & past utilization

▸ **Illness:** Disaggregate use into illness indicators

- ▸ Binary logits for 14 illness groups – cancer, eye, …

Utilization results:
Fair/poor APEs across specifications
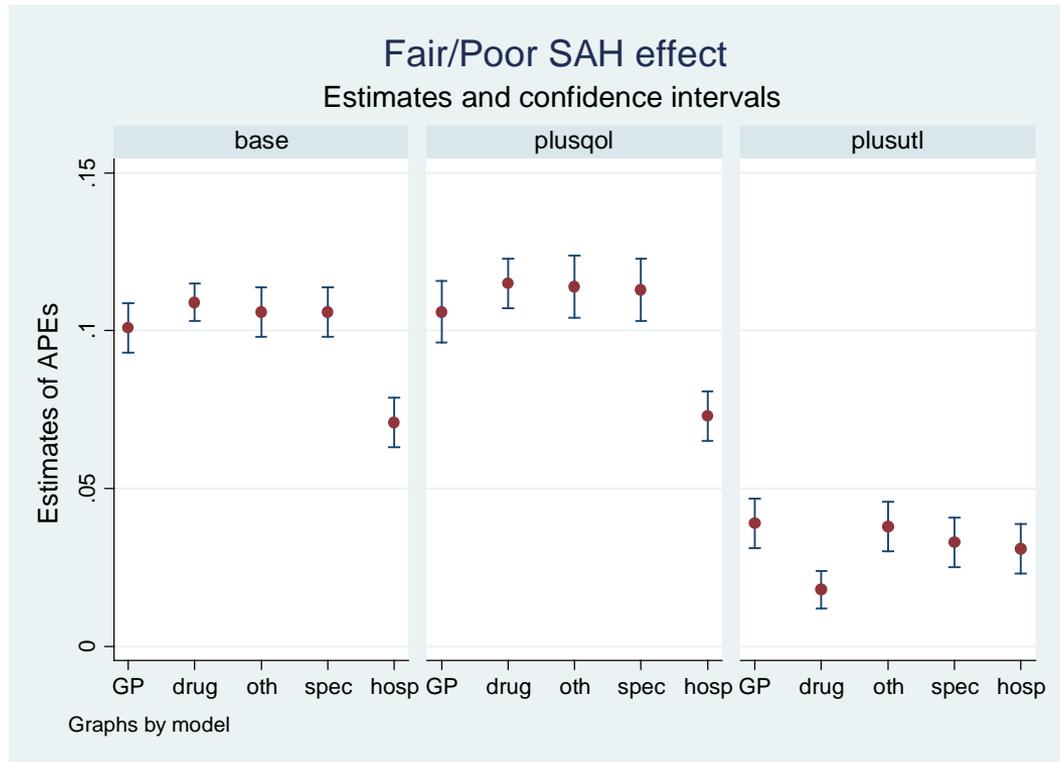
# Utilization results:
Good APEs across specifications
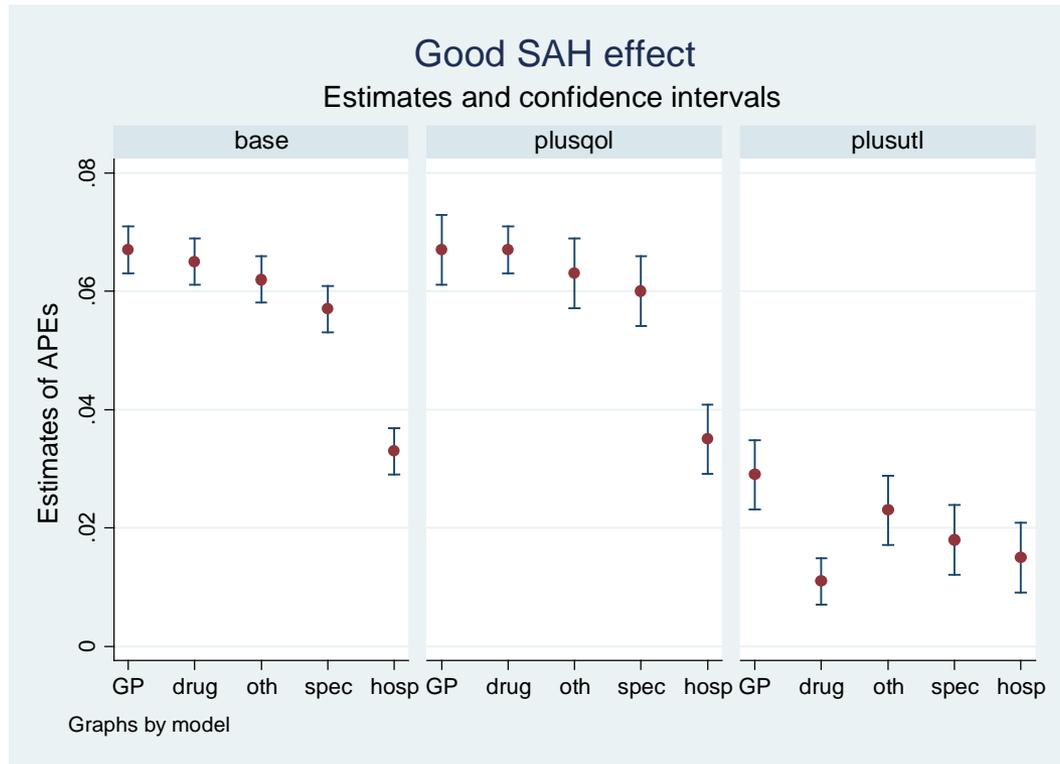
Utilization results: Heterogeneity by age
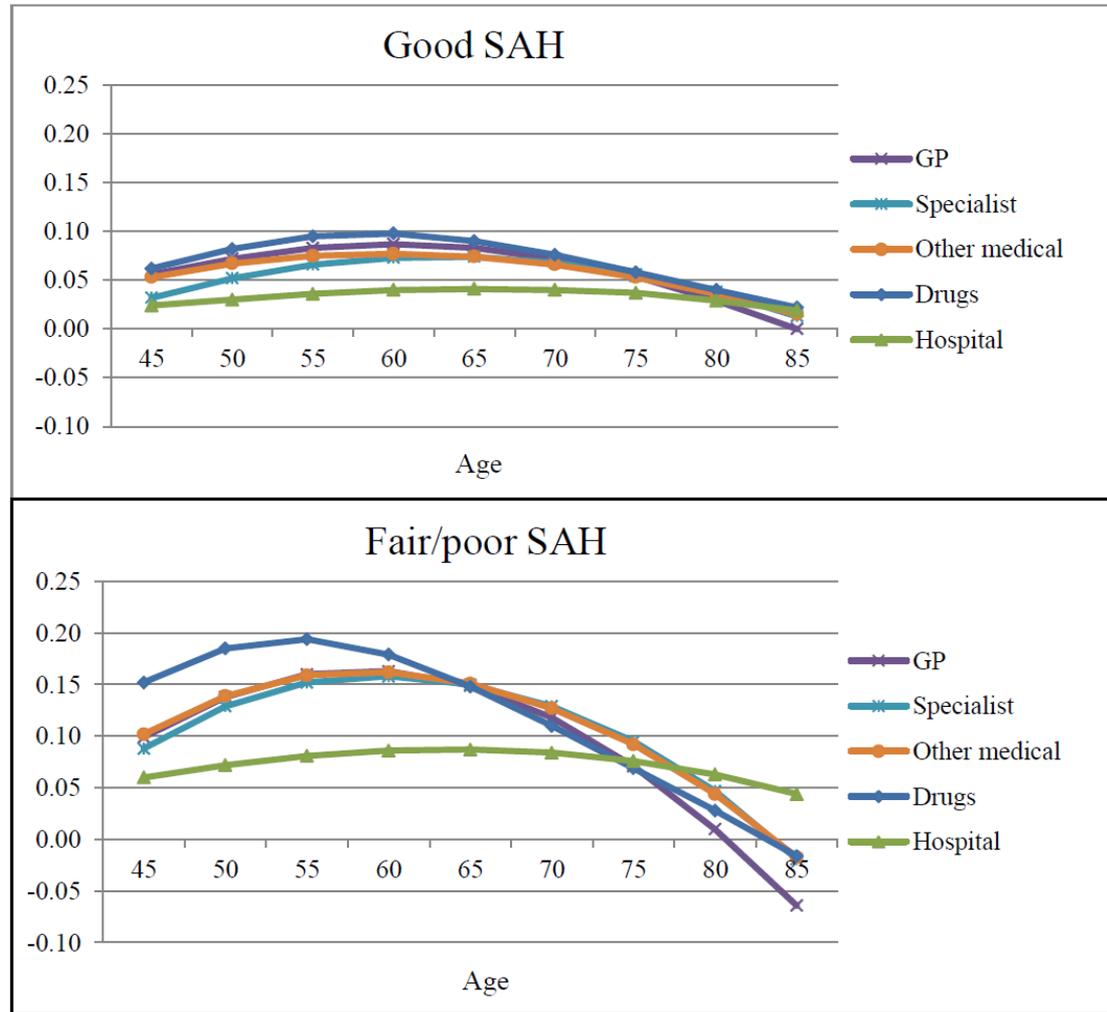
Illness results: Specification with QoL but no past utilization
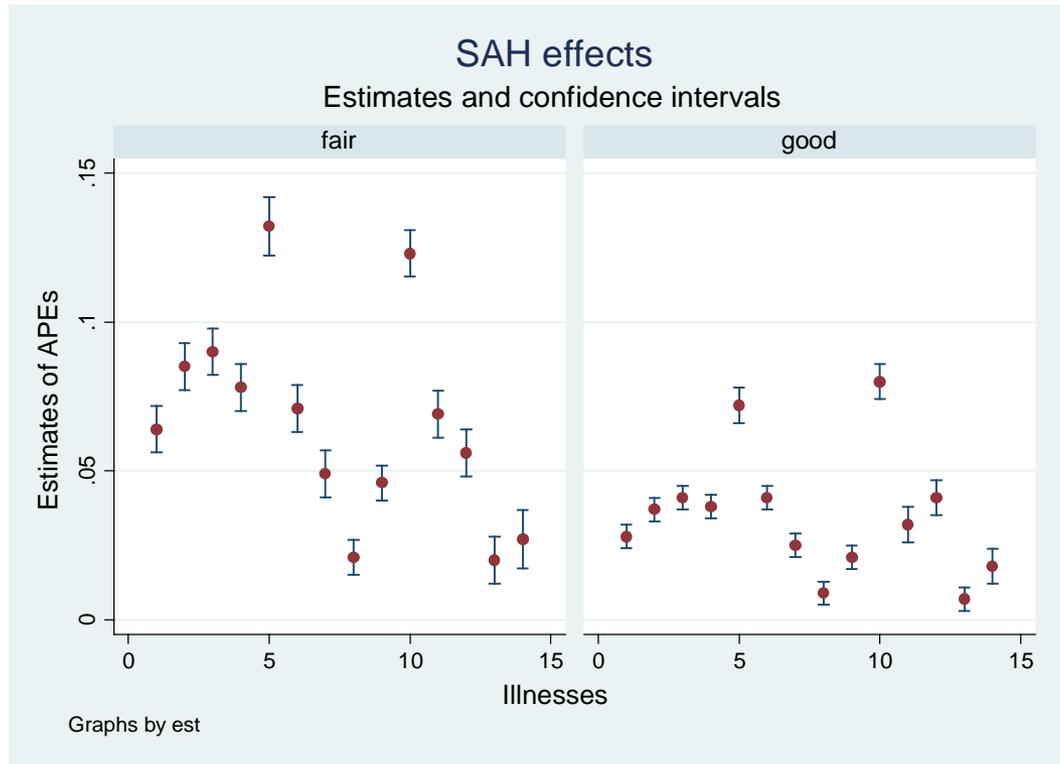
Illness results:
APEs scaled by incidence

SAH effects as percentage of incidence
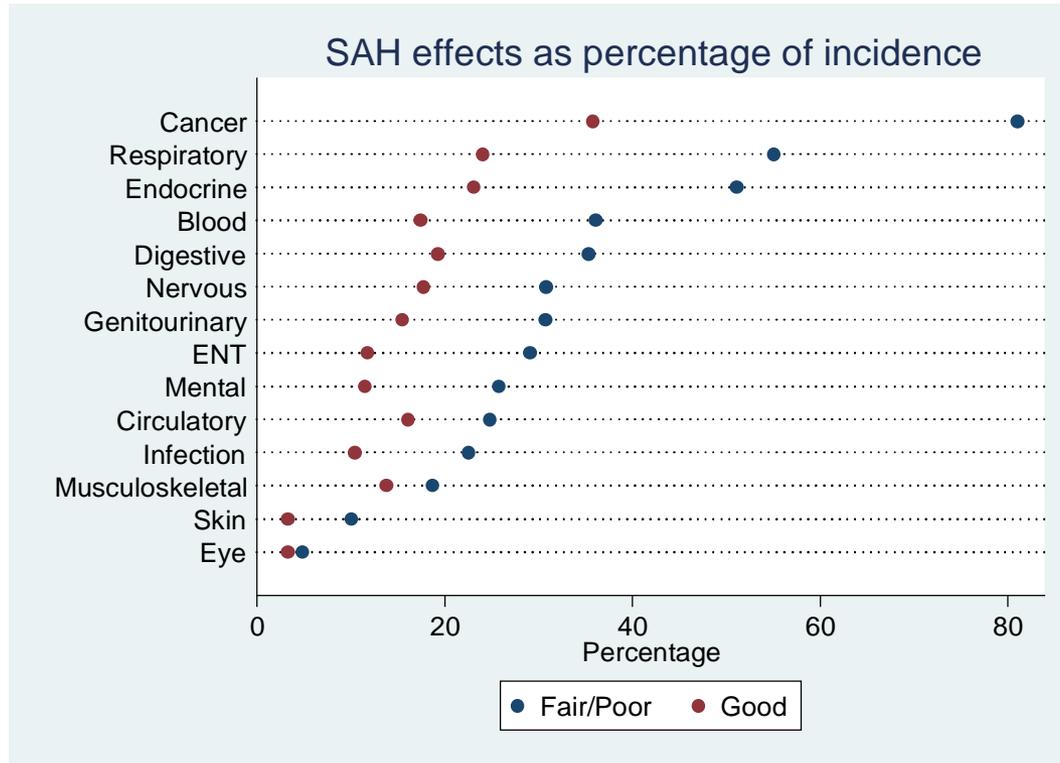
Yes ... but

- ## SAH does measure health
  - Actual health main source of predictive power of SAH
  - Also captures other unobserved health factors
- ## Evidence supports role of SAH as good **proxy** for health
  - BUT still measurement concerns; see Black et al. (2017)

SAH reporting inconsistencies in HILDA
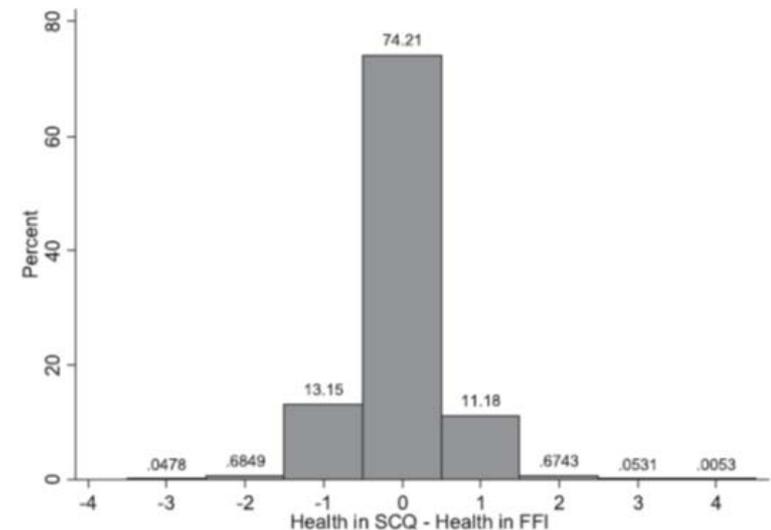
Fig. 2. Distribution of the differences in self-assessed health from the same survey wave.
Notes: Sample consists of wave 9 and 13 HILDA respondents aged 25 years or more. Sample size equals 18,834 individual-year observations.

# Need more than passive use of SAH

▸ Measurement issues need to be addressed in other modelling roles

- ▸ Causal interpretation of health as a determinant
- ▸ Health is the outcome in say studies of inequalities

▸ Adjustments for measurement problems can matter

- ▸ Mu (2014) uses CHARLS to explore regional differences in SAH of elderly in China
- ▸ Molina (2017) uses HRS, ELSA, IFLS & CHARLS to study cross-country differences in health self reports over 6 domains
- ▸ Both use vignettes to control for heterogenous response thresholds

Ranking of countries by predicted proportion in healthiest category

USA
UK
Indonesia
China
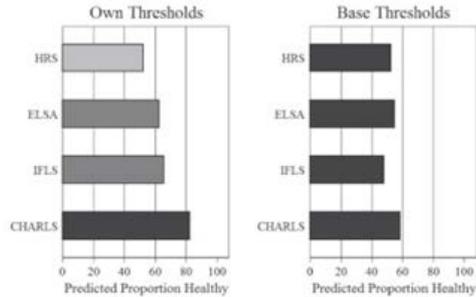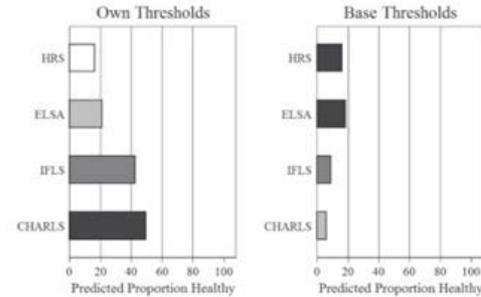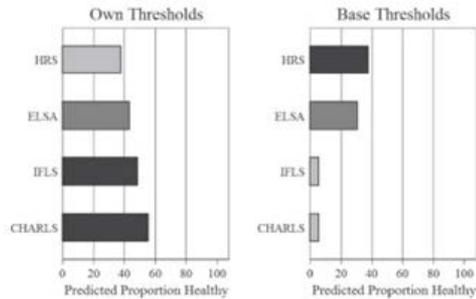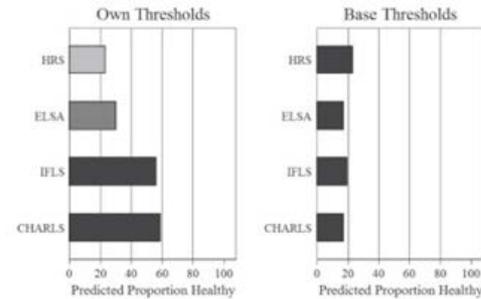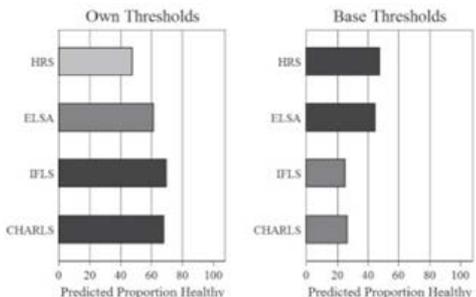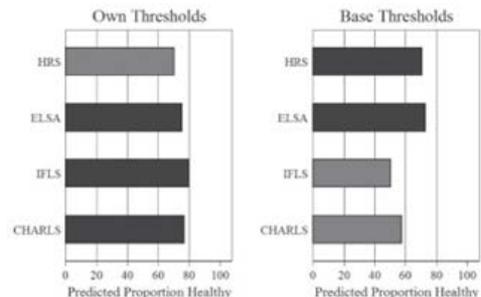
Case of EQ-5D

- Role of some subjective measures is mandated in HTA
  - NICE, PBAC, … accept public preferences over health states important in decisions on government resource allocation such as recommending medicines for PBS listing
  - PBAC "accepts" health-related quality of life measures including EQ-5D-3L (no, some or extreme problems) or EQ-5D-5L (no, slight, moderate, severe or extreme problems)
- Such measures contentious given role in health funding decisions under pressure from an ageing population
  - Accuracy needed for confidence in preference data

UNSW
THE UNIVERSITY OF NEW SOUTH WALES

An EQ-5D-5L controversy

▸ NICE's preferred measure EQ-5D-3L but now 5L

- ▸ Devlin et al. (2018) provided general public EQ-5D-5L value set
- ▸ Preferences obtained using subset of health states with other states predicted to form value set weights
- ▸ QALY=(length of life)x(weight)

▸ Updating value set & moving to 5L to provide more nuanced characterization of health both a priori sensible

- ▸ So what's the problem?

▸ Review of 5L value set critical of data quality & modelling

- ▸ Value set authors defended their work rebutting criticisms

Some thoughts on process

‣ Experts [Fiebig (2019), Manski (2019)…] reviewed the review

   ‣ **"Experts have recommended NICE should support a new study of how to value health-related quality of life in England."** Lovett & Cooper (2019)

‣ All data misbehave & applied work is difficult

   ‣ Easy to be critical

   ‣ But "everyone else does it" a poor response to criticism

   ‣ Evaluation difficult with no gold standard & subjective modelling decisions

‣ Cognitively demanding tasks (TTO & DCE)

   ‣ Respondents found process difficult but no significant age effects

   ‣ Need better understanding of respondent decision-making process

UNSW
THE UNIVERSITY OF NEW SOUTH WALES

Some thoughts on process…

▸ **Best way to undertake valuation given NICE & journal unlikely to have same objective function?**

  ▸ Should heterogeneity in population be modelled if only produce a single value set?

  ▸ Review team undertook substantial replication & exploratory analyses

▸ **Preferences matter so role for subjective health measures**

  ▸ Measurement process involved so needs to be done carefully & be fit for purpose

Acknowledgement

Core References

▸ Doiron, D., Fiebig, D.G., Johar, M. and Suziedelyte, A. (2015), "Does self assessed health measure health?", *Applied Economics*, 47, 180-194.

▸ Fiebig, D.G. (2017), "Big data: Will it improve Patient-Centered Care?", *The Patient: Patient–Centered Outcomes Research*, 10, 133-139.

▸ Fiebig, D.G. (2019), "Expert evaluation of EQ-5D-5L", available here

▸ Lovett, R. and Cooper, S. (2019) "NICE to support new valuation study for England for EQ-5D-5L questionnaire", available here

▸ Manski, C.F. (2019), "NICE Expert advice on the EQ-5D-5L valuation set for England", available here

UNSW
THE UNIVERSITY OF NEW SOUTH WALES