

RESEARCH ARTICLE

Open Access

# A novel nonparametric item response theory approach to measuring socioeconomic position: a comparison using household expenditure data from a Vietnam health survey, 2003

Daniel D Reidpath<sup>1,2\*</sup> and Keivan Ahmadi<sup>2,3</sup>

## Abstract

**Background:** Measures of household socio-economic position (SEP) are widely used in health research. There exist a number of approaches to their measurement, with Principal Components Analysis (PCA) applied to a basket of household assets being one of the most common. PCA, however, carries a number of assumptions about the distribution of the data which may be untenable, and alternative, non-parametric, approaches may be preferred. Mokken scale analysis is a non-parametric, item response theory approach to scale development which appears never to have been applied to household asset data. A Mokken scale can be used to rank order items (measures of wealth) as well as households. Using data on household asset ownership from a national sample of 4,154 consenting households in the World Health Survey from Vietnam, 2003, we construct two measures of household SEP. Seventeen items asking about assets, and utility and infrastructure use were used. Mokken Scaling and PCA were applied to the data. A single item measure of total household expenditure is used as a point of contrast.

**Results:** An 11 item scale, out of the 17 items, was identified that conformed to the assumptions of a Mokken Scale. All the items in the scale were identified as strong items ( $H_i > .5$ ). Two PCA measures of SEP were developed as a point of contrast. One PCA measure was developed using all 17 available asset items, the other used the reduced set of 11 items identified in the Mokken scale analysis. The Mokken Scale measure of SEP and the 17 item PCA measure had a very high correlation ( $r = .98$ ), and they both correlated moderately with total household expenditure:  $r = .59$  and  $r = .57$  respectively. In contrast the 11 item PCA measure correlated moderately with the Mokken scale ( $r = .68$ ), and weakly with the total household expenditure ( $r = .18$ ).

**Conclusion:** The Mokken scale measure of household SEP performed at least as well as PCA, and outperformed the PCA measure developed with the 11 items used in the Mokken scale. Unlike PCA, Mokken scaling carries no assumptions about the underlying shape of the distribution of the data, and can be used simultaneous to order household SEP and items. The approach, however, has not been tested with data from other countries and remains an interesting, but under researched approach.

**Keywords:** Mokken scale analysis (MSA), Principal component analysis (PCA)

\* Correspondence: daniel.reidpath@monash.edu

<sup>1</sup>South East Asia Community Observatory (SEACO), Monash University, Bandar Sunway, Malaysia

<sup>2</sup>Global Public Health, School of Medicine and Health Sciences, Monash University, Bandar Sunway, Malaysia

Full list of author information is available at the end of the article

## Background

Socioeconomic position (SEP) has played an important role in many health studies [1-5]. The relationship between SEP and health has been studied in its own right, [6-8] and it has been treated as a potential covariate/confounder in studies of other substantive causes of poor health [9,10]. Typically, the households in such studies are divided into quintiles according to their estimated SEP and then comparisons are made between fifths of the population [11]. Identifying valid, reliable, acceptable, and low cost methods of measuring SEP is an ongoing and important area of research in the health sciences [11-13].

Among the possible measures of SEP, household expenditure is associated with various health outcomes, [14-16] and tends to be preferred by economists [12]. However, in low and middle income countries expenditure data can be difficult to obtain, [11,13], and common alternatives have been asset-based indices of household wealth that may include access to utilities and infrastructure. A minority view holds that asset-based measures are in fact superior to expenditure based measures of SEP; [17] with others advocating a middle position. Somi et al., for instance, argued that both expenditure measures and asset-based indices should be treated as legitimate proxies of SEP, given that SEP is a latent variable that cannot be directly observed [18]. In a recent, significant (though not fully comprehensive (p.883)) review of socio-economic measures in low and middle income countries, the authors concluded that the research question, the setting, and the available resources needed to guide the choice of approach to the measurement of SEP [19]. While this is undoubtedly true, in many cases, particularly in the secondary analysis of household survey data there is a tendency to fall back on a small handful of techniques that can be readily applied to data over which the researcher had no control during the collection [20,21].

Various approaches exist for the construction of asset-based indices [13,22]. Ubiquitous among these, which is used in this article as a point of comparison with Mokken scaling, is a principal components analysis (PCA) of a parcel of household assets [17,18,20,21]. The PCA approach was famously, although not first described by Filmer and Pritchett [17] and was adopted by the HNP/Poverty Thematic Group of the World Bank as a standard technique in their poverty and equity analyses covering 44 countries; (cf[23-25]). PCA is the approach taken in the DHS Wealth Index; [26] and it remains a common tool in health research today [7,27,28]. The PCA approach in wealth measurement has its early development in the recognition that multiple measures of wealth create analytic problems associated with collinearity, and PCA offers an efficient data reduction technique to extract orthogonal dimensions [29].

In contrast, Mokken scales take as their conceptual starting point Guttman scales [30,31]. A Guttman scale is a set of ordered (increasingly “harder”) items. Without formally defining “harder” questions, those households with a higher SEP would respond positively to increasingly “harder” questions, leaving a SEP rank order of respondent households from low SEP to high SEP, with all but the highest ranked households eventually finding some questions “too hard” to respond to positively. For example, a question about car ownership is likely to be a “harder” question than a question about ownership of a bucket. Guttman scales, however, are strictly deterministic and do not allow for error in the measurement.

It is here that Mokken scales differ from Guttman scales. Mokken scales are probabilistic, belonging to the non-parametric item response theory model of scales, and allow for stochastic error in measurement [30]. For each item in an asset-based Mokken scale of SEP, the *probability* of a positive response to a question of asset ownership depends on two factors: the SEP latent trait characteristics of the household; and the item characteristics of the asset ownership question. The higher the SEP latent trait of a household, the greater the probability that the household will respond positively to any asset ownership question – without regard to the difficulty of the question itself. The “harder” the ownership question, the lower the probability that a household will respond positively to the question – without regard to the household. This means that a Mokken scale, asset-based index of SEP can rank order households according to their latent trait, and rank order items according to their probability of eliciting a positive response [31]. A Mokken scale analysis (MSA) identifies those items that can be used to rank respondent households according to their probability of a positive response (i.e. their position on the latent trait of SEP), and it orders items according to their probability of being answered positively. Mokken scales are also frequently shorter than scales developed using other procedures, holding out the promise of more concise measures of SEP. While item response theory approaches have been applied to the measurement of household SEP previously, the application has been in high income countries, and relied on parametric techniques (such as the Rasch model) with stronger underlying assumptions than the nonparametric approach of Mokken scales [32-34].

We illustrate the use of MSA in the development of a measure of household SEP in a low income country setting, and contrast Mokken scaling with an equivalent PCA measure of SEP and with a single item measure of household expenditure. The comparison is made using data from a nationally representative household survey conducted in Vietnam. The analysis of a single data set cannot stand as a robust comparison of a scaling

technique. It can however illustrate the use of a novel approach to SEP measurement; and in the spirit of an “emerging theme” it may motivate further interest and research.

### Methods

The World Health Survey was a household survey utilising a uniform methodology conducted in 70 countries between 2002 and 2004 [35,36]. Asset ownership questions were included in the survey at the household level. The household asset data analysed here were drawn from the Vietnam, World Health Survey 2003 [37]. Data were collected from 4,154 of 4,174 consenting households (a response rate of 99.5%) [37]. Approximately 23% of households were urban and 77% rural. Households with incomplete asset ownership data (8.7% of households) were excluded from the analysis, leaving a usable sample of 3,810 households (an effective response rate of 91.3%). The level of missing data was considered small enough not to warrant imputation [38]. The survey included 16 dichotomous questions on household asset ownership, access to utilities and infrastructure, and one continuous response question which was dichotomised for the analysis (Table 1).

Household expenditure was measured with the question: “In the last 4 weeks, how much did your household spend in total?”

The data from the World Health Survey are publicly available for analysis as anonymised, unit record files. Ethics Committee approval for the analysis presented here was neither sought nor required.

### Data analysis

PCA is a well described technique for developing asset-based indices of SEP and will not be described in detail here [11]. It is nonetheless worth noting that PCA is a statistical technique to reduce the dimensionality of data by identifying sets of weighted linear combinations (principal components) of the original asset measures, such that each new principal component accounts for a smaller proportion of the variance than the preceding components, and that each of the identified principal components are orthogonal [39]. It is the first principal component (accounting for the greatest proportion of the variance) that is typically used to construct an index of household SEP. An underlying assumption of PCA is that the data are continuous and drawn from a multivariate normal distribution [20]. This is not the case with a series of dichotomous asset-ownership questions. However, assuming an underlying continuous, normally distributed latent variable, the polychoric correlation between two observed dichotomous variables can be used as an estimate of the actual correlation [20]. It was the polychoric correlation matrix that was used in the PCA described here.

**Table 1 The parcel of goods/items used in the Mokken scale analysis**

Item	Proportion <sup>3</sup>	Mokken Scale	H <sub>i</sub> <sup>4</sup>
Does anyone in your household have a <b>bicycle</b> ?	83.96	-	-
Does anyone in your household have more than one <b>house/apartment</b> <sup>2</sup>	4.67	-	-
How many <b>cars</b> are there in your household? <sup>1</sup>	3.65	-	-
Does anyone in your household have a <b>clock</b>	90.47	1	.62 (.019)
Does your home have <b>electricity</b> ?	90.42	1	.56 (.021)
Does anyone in your household have a <b>television</b>	78.53	1	.74 (.013)
Does anyone in your household have a <b>motorbike</b>	49.79	1	.70 (.013)
Does anyone in your household have a <b>video/DVD</b> <sup>2</sup>	31.57	1	.68 (.014)
Does anyone in your household have a <b>telephone</b> , fixed-line	16.72	1	.69 (.013)
Does anyone in your household have a <b>refrigerator</b>	15.30	1	.69 (.014)
Does anyone in your household have a <b>magazine subscription</b> <sup>2</sup>	7.51	1	.46 (.023)
Does anyone in your household have a <b>washing-machine</b>	5.33	1	.59 (.021)
Does anyone in your household have a <b>computer</b>	5.04	1	.50 (.026)
Does anyone in your household have a <b>mobile telephone</b>	4.33	1	.64 (.019)
Does anyone in your household have a <b>dishwasher</b> <sup>2</sup>	0.34	1	.57 (.106)
Does anyone in your household have a <b>bucket</b>	84.62	2	.49 (.097)
Does anyone in your household have an <b>agricultural machine</b> <sup>2</sup>	8.06	2	.49 (.097)

<sup>1</sup>A continuous item dichotomised as none (0) and one or more (1).

<sup>2</sup>These are “Country Specific” items, which vary from country to country in which the World Health Survey was conducted.

<sup>3</sup>The proportion of households owning each asset.

<sup>4</sup>Standard errors are shown in parentheses.

The final items used in both the Mokken scale analysis and the 11 item PCA are identified as belonging to Mokken Scale 1.

MSA relies on an automated procedure for the selection of items that belong to one or more independent Mokken scales (or no scale at all), and a series of methods to investigate the extent to which the scales maintain the assumptions of a nonparametric item response theory model [40].

A Mokken scale of household SEP is based on four assumptions:

1. *Unidimensionality*. A scale of responses to questions of household asset ownership measures a dominant, single latent trait of household SEP.
2. *Local Independence*. Responses to an asset ownership question are not influenced by the responses to any other asset ownership question in the same scale.
3. *Monotonicity*. The probability of a positive response to an asset ownership question is a monotonically increasing function of the latent trait. This assumption would be violated, for instance, if both low and high SEP households had a low probability of owning asset  $a_i$ , but middle SEP households had a high probability of owning asset  $a_i$ .
4. *Non-intersection*. If the probability of households owning asset  $a_i$  is lower than probability of households owning asset  $a_k$ , for one level of the SEP latent trait (e.g., a low SEP household), then it will be lower for all levels of the latent trait (i.e., middle and high SEP households). This is referred to in the literature as *invariant item ordering* or (IIO), and means that the ordering of difficulty of the asset ownership question holds for all households without regard to their SEP [41-43].

In Mokken scaling, the *model of monotone homogeneity* (MMH) is based on the first three assumptions. In its practical application, a household SEP scale meeting the requirements of MMH allows for the ordering of households by the sum of the number of positive responses to the asset ownership questions. The more positive responses, the higher a household's SEP [41,42].

If a scale meets the requirements of the MMH and the scale meets the fourth assumption of non-intersection (or IIO), it also fulfills the requirements of the *double monotonicity model* (DMM). In its practical application it means that not only can households be ordered on the latent trait of household SEP, but the asset ownership questions (items) can be ordered according to how "hard" or "difficult" they are.

There are various methods for testing the assumptions of a Mokken scale. At the heart of the procedures are Loevinger's homogeneity coefficients [44,45]. These are three related coefficients, which are used to select items that contribute to a unidimensional (homogeneous) scale [30,44]. For details, readers are referred to a number of

articles and books written on the topic, where for brevity we focus on the conceptual and applied application. [40,43,46,47]. The main coefficients used are  $H_i$  and  $H$ . The  $H_i$  coefficient provides a measure of the scalability of each item  $i$  that makes up the potential scale, and the  $H$  coefficient provides a measure of the scalability of the whole scale (i.e., the degree to which the items always appear in the same relative order) [48]. Guidelines for the interpretation of the coefficients suggest that values of .3–.4 are indicative of a weak scale (or item), values of .4–.5 are indicative of a medium scale (or item) and values  $>.5$  are indicative of a strong scale (or item) [30]. When the  $H$  coefficient is calculated on the transpose matrix of dichotomous asset ownership responses ( $H^T$ ), one obtains a summary statistic of the accuracy of asset ordering within a scale.

The automated item selection procedure (AISP) partitions a set of items into zero or more Mokken scales and provides summary statistics for the items and scales [46]. This is a necessary but not sufficient procedure for establishing a Mokken scale. For the MMH one needs to establish monotonicity, and for the DMM one also needs to establish the non-intersection assumption, or IIO. There are a number of possible approaches to establishing non-intersection, and in this study, the *restscore method* was used, which compares all possible item pairs to establish whether significant violations of the non-intersection assumption occur [43,46].

All analyses were conducted in the R statistical environment [49]. The analyses were supported by the mokken package for the Mokken scale analysis, [47] and the polycor package for estimating polychoric correlations [50].

## Results

### Mokken Scale Analysis

Of the 17 items included in the MSA, the automated item selection procedure identified three items which could not be scaled, 12 items potentially belonged to one scale, and two items potentially belonging to a second scale (Table 1).

For the remainder of the paper, the focus is on the 12 items contributing to scale one. The single item scalability coefficients,  $H_i$  ranged in value from .46 to .79 with 10 of the 12 items having values greater than 0.5; i.e., potentially "strong" items [30]. The standard errors of each  $H_i$  were relatively small, with the exception of the item measuring dishwasher ownership. Indeed, of the 10 items with  $H_i$ 's indicating strong scalability, only the  $H_i$  of the item measuring dishwasher ownership had a 95% confidence interval to include a value less than .5 (i.e.,  $H_i - 1.96 \times SE = .480$ ) [47,51]. Importantly, there was no item for which the lower bound of the 95% confidence interval included .3, a conventionally used cut-off to reject potential items as un-scalable [51].

The overall scalability coefficient for scale one, H was .65. There were, furthermore, no violations of monotonicity. However the *rest scores*, used to test the IIO, indicated two critical violations associated with electricity availability and clock ownership. It is recommended that the worst offending item is removed from the preliminary scale and the rest scores re-examined – in this case the item related to electricity availability. Once electricity availability had been removed as an item in the scale, there were no further violations [46].

In keeping with the Mokken scaling approach, the Mokken SEP score was calculated as the unweighted sum of the 11 remaining dichotomous items. Figure 1 shows the distribution of Mokken SEP scores with approximate quintile boundaries. Twenty percent of households had a score of 1 or less, 45% of households had a score of 2 or less, 67% of households had a score of 3 or less, and 82% of households had a score of 4 or less. The households in the top 20% had scores of 5 or more.

### Principal components analysis

Two separate PCAs were conducted using different item pools. The first PCA was based on the 17 item asset ownership questions from the World Health Survey. The second PCA was based on the reduced, 11 item pool, identified in the Mokken analysis.

In the first analysis of all 17 items, the first principal component accounted for 51.5% of the variance, the second accounted for 20.8% of the variance, and the third accounted for 9.7% of the variance. After the fourth component, eigen values fell below 1. A scree plot (un-shown) indicated no obvious discontinuity or ‘elbow’ in

the declining eigen values. In the second PCA, using the 11-item pool, the first principal component accounted for 24.9% of the variance, the second accounted for 21.6% of the variance, and the third accounted for 14.4% of the variance. After the fourth component, eigen values again fell below 1. The scree plot (un-shown) indicated no obvious discontinuity or ‘elbow’ in the declining eigen values.

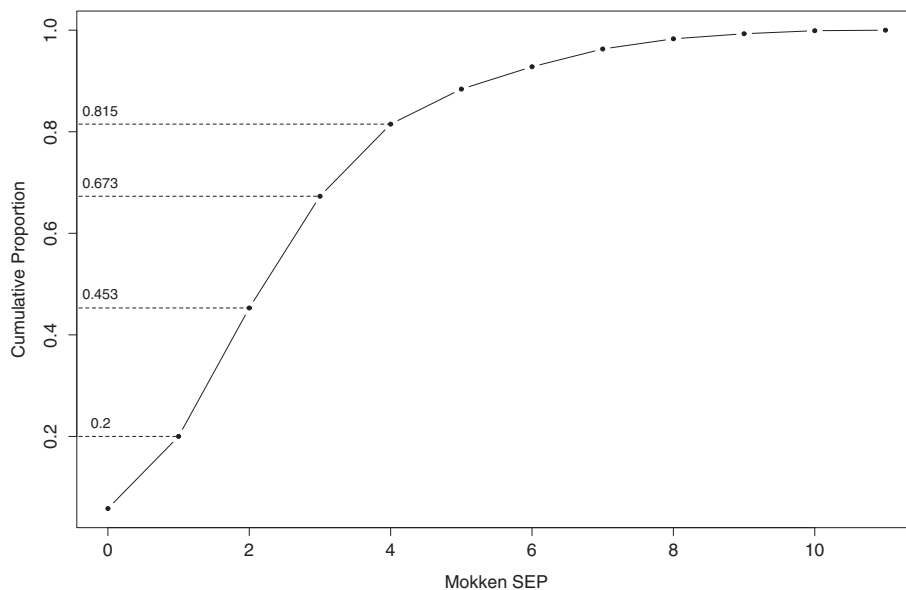
For both PCA analyses, asset scores based on the first principal component were used to create a continuous SEP score, and quintiles.

The Pearson’s product moment correlation between the Mokken measure and the 17 item PCA measure of SEP was very high,  $r = .98$ , and for the quintiles of wealth, Spearman’s rank order correlation was  $r = .96$ . The relationship, however, was weaker for the reduced, 11 item PCA measure. The Pearson’s product moment correlation between the Mokken-based measure and the PCA-based measure of SEP was moderate,  $r = .68$ , and the Spearman’s rank order correlation for quintiles of wealth was very low  $r = .11$ .

### Reliability

Cronbach’s alpha was used as a measure of the reliability three SEP scales. The items were weighted prior to the calculation of Cronbach’s alpha. They were weighted to ensure that each SEP scale was evaluated based on its adjusted item scores.

In the case of the Mokken scale, the items were unit weighted, because the scale is a simple sum of the asset-based items. In the case of the PCA scales, the items were weighted by the PCA loadings from the first principal



**Figure 1** An ogive of the distribution of Mokken SEP scores with approximate quintile boundaries.



component. Cronbach's alpha for the unit weighted 17 item scale was included as a point of contrast.

The reliability of the unit weighted, 17 item scale was .73 (95% CI: .71 – .74). The reliability of the PCA weighted, 17 item scale was .52 (95% CI: .49 – .54). The reliability of the unit weighted, 11 item Mokken scale was .76 (95% CI: .75 – .78). The reliability of the PCA weighted, 11 item scale was .19 (95% CI: .15 – .23). The Mokken scale had a significantly larger Cronbach's alpha than the other potential SEP scales.

### Comparisons with household expenditure

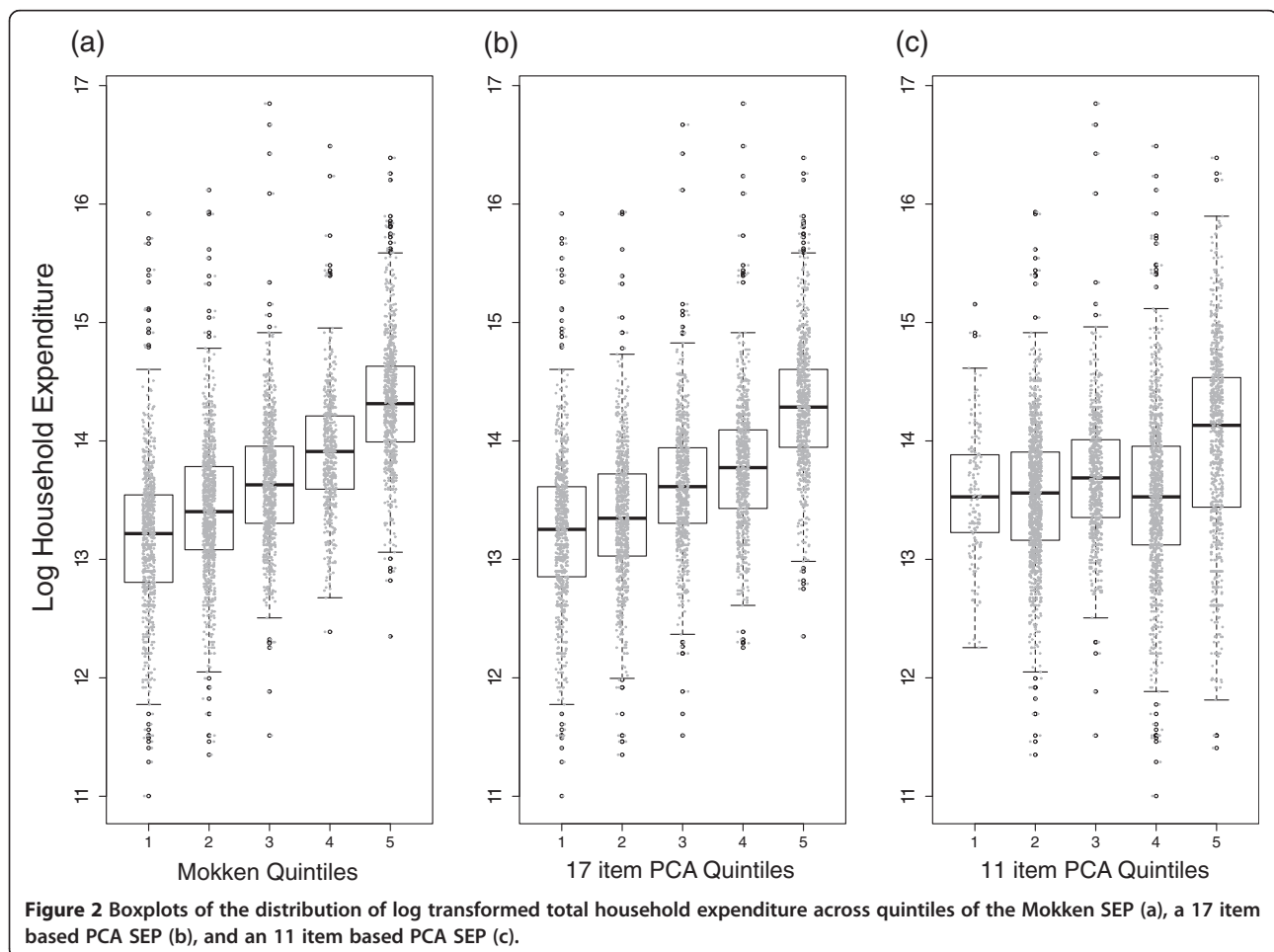
The responses to the single household expenditure question from the World Health Survey were log transformed because of the long tail of the distribution that is typical of expenditure and income data. Figure 2 shows a box plot of household expenditure data over the quintiles of household SEP estimated by the Mokken scale analysis (2a), the PCA using all 17-items (2b), and the PCA using the 11 items identified by the Mokken analysis (2c). The distribution of actual household expenditure values in each quintile was overlaid as grey coloured points.

A visual inspection suggests, as might be expected from the close to perfect correlation between the 17 item PCA quintiles and Mokken scale quintiles, that they perform very similarly (Figure 2a and b). In contrast the boxplot of the 11 item PCA quintiles against household expenditure shows no obvious systematic relationship (Figure 2c).

The Pearson's product moment correlation between the household expenditure data and the Mokken SEP continuous scores was  $r = .59$ , for the 17 item PCA SEP continuous scores it was  $r = .57$ , and for the 11 item PCA SEP continuous scores it was  $r = .41$ . The correlations between the household expenditure and the quintiles data for the Mokken ( $r = .57$ ) and the 17 item PCA SEP ( $r = .53$ ) showed similar levels of performance. The correlation between the household expenditure data and the 11 item PCA SEP quintiles was  $r = .18$ .

### Discussion

Mokken scaling appears to be a promising approach to the development of an asset-based measure of household SEP. The Mokken scale was strongly correlated with the



17 item PCA measure of SEP, and with a significantly better Cronbach's alpha. Both measures performed similarly with respect to the measure of total household expenditure. That measures of asset based SEP and expenditure were moderately correlated supports the general notion of an underlying latent wealth construct, [12,18] and it supports the use, of asset based measures if expenditure measures are desirable but unobtainable.

In sharp contrast to the 17 item PCA, the correlation was much weaker between the 11 item Mokken scale and the 11 item PCA measure of SEP. Furthermore, the 11 item PCA SEP quintiles showed no practical relationship with the household expenditure; and the Cronbach's alpha for the PCA measure was very weak.

A novel feature of Mokken scaling is that it orders items as well as households. Outside the direct value of SEP to health research, it may potentially be used to track changes in items indicative of wealth over time. Hard items today, (i.e., items to which only the wealthiest have a high probability of responding positively) may become easy items in the future, or visa-versa. In 2003 when the World Health Survey data for Vietnam was collected, the market penetration of the mobile telephone was less than 5%, supporting the apparent "hardness" of the item identified in the Mokken scale analysis (Table 1). Mobile phones were a rare and expensive commodity in 2003. In 2012 the market penetration of the mobile phone in Vietnam had exceeded 100%, making it an "easy" item that would not today readily separate the wealth quintiles [52]. The mobile phone as an asset item was highlighted for very similar reasons in a recent Rasch analysis of poverty in Europe (p.69) [32].

Given the growth of interest in item response theory approaches to modelling SEP, [33,34], the results of the Mokken scaling presented here should pique some interest. Unlike parametric item response theory models, there are fewer assumptions associated with Mokken scaling which can broaden its application. This has been found in other areas of health research where Mokken scaling has been used successfully in its own right, and used to support or check parametric item response theory approaches [53-55].

### Limitations

One of the limitations of this analysis relates to the generalisability of the approach; specifically whether Mokken scaling will always perform comparably well or outperform PCA; and whether it will perform as well as other approaches [56]. This limitation, however, needs to be placed in the context of at least one of the paper's goals, which was to illustrate the use of Mokken scale analysis in the context of SEP measurement.

For some the apparent complexity of Mokken scaling over PCA maybe seen as a limitation; indeed, one of the

Reviewers of a draft raised this very possibility. We would argue that the apparent complexity is a function of familiarity. Understanding PCA and the underlying eigen values is not trivial. Exposure to a technique creates familiarity. This is the first paper we know of that uses Mokken scaling in the development of an SEP measure. Furthermore, given the emergence of item response theory approaches in SEP measurement, this is well timed and should add another technique to the quiver of methods available to epidemiologists [56].

The use of a single global measure of household expenditure as a comparative measure of SEP is also a limitation. While additional questions could undoubtedly have improved the measure of household expenditure, these were not available, and as so often happens in secondary data analysis, one is constrained by the choices made by the original researchers. It is also known that expenditure data in low income settings can be hard to obtain, which motivated the creation of asset-based measures in the first place. The real problem with single question measures is that they are often very noisy (in a stochastic sense). The fact that both the 17 item PCA measure of SEP and the Mokken scale measure of SEP correlated moderately with the single item measure of household expenditure, however, suggests that the choice was not misguided in this context.

### Conclusion

The Mokken scale measure of household SEP performed at least as well as PCA, and outperformed the PCA measure developed with the 11 items used in the Mokken scale. Unlike PCA, Mokken scaling carries no assumptions about the underlying shape of the distribution of the data, and can be used simultaneous to order household SEP and item difficulty. The approach, however, has not been tested with data from other countries and remains an interesting, but under researched approach.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

DDR and KA jointly developed the idea of applying Mokken scaling to the problem of measuring SEP. KA provided technical input on Mokken scaling, DDR conducted the preliminary analysis. DDR and KA jointly drafted and edited the manuscript. All authors read and approved the final manuscript.

### Acknowledgment

This paper uses data from the Vietnam WHO World Health Survey, 2003.

### Author details

<sup>1</sup>South East Asia Community Observatory (SEACO), Monash University, Bandar Sunway, Malaysia. <sup>2</sup>Global Public Health, School of Medicine and Health Sciences, Monash University, Bandar Sunway, Malaysia. <sup>3</sup>School of Pharmacy, International Medical University, Kuala Lumpur, Malaysia.

Received: 3 March 2014 Accepted: 16 July 2014

Published: 14 August 2014

## References

1. Lebowitz MD, Malcolm JC: **Socioeconomic analysis of the Alameda County Health Department jurisdiction.** *Am J Public Health Nations Health* 1964, **54**:1876–1881.
2. Montgomery MR, Hewett PC: **Urban poverty and health in developing countries: Household and neighborhood Effects.** *Demography* 2005, **42**:397–425.
3. Graham H: *Unequal Lives: Health and Socioeconomic Inequalities.* Maidenhead: McGraw-Hill International; 2007.
4. Boccia D, Hargreaves J, De Stavola BL, Fielding K, Schaap A, Godfrey-Faussett P, Ayles H: **The association between household socioeconomic position and prevalent tuberculosis in Zambia: a case–control study.** *PLoS One* 2011, **6**:e20824.
5. Navalpotro L, Regidor E, Ortega P, Martinez D, Villanueva R, Astasio P: **Area-based socioeconomic environment, obesity risk behaviours, area facilities and childhood overweight and obesity: socioeconomic environment and childhood overweight.** *Prev Med* 2012, **55**:102–107.
6. Adams P, Hurd MD, McFadden D, Merrill A, Ribeiro T: **Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status.** *J Econom* 2003, **112**:3–56.
7. Okolo CO, Reidpath DD, Allotey P: **Socioeconomic inequalities in access to health care: examining the case of Burkina Faso.** *J Health Care Poor Underserved* 2011, **22**:663–682.
8. Patel SA, Murray-Kolb LE, LeClerq SC, Khatry SK, Tielsch JM, Katz J, Christian P: **Household wealth and neurocognitive development disparities among school-aged children in Nepal.** *Paediatr Perinat Epidemiol* 2013, **27**:575–586.
9. Van de Poel E, O'Donnell O, Van Doorslaer E: **Are urban children really healthier? Evidence from 47 developing countries.** *Soc Sci Med* 1982 2003, **65**:1986.
10. Rani M, Bonu S, Jha P, Nguyen SN, Jamjoum L: **Tobacco use in India: prevalence and predictors of smoking and chewing in a national cross sectional household survey.** *Tob Control* 2003, **12**:e4–e4.
11. Vyas S, Kumaranayake L: **Constructing socio-economic status indices: how to use principal components analysis.** *Health Policy Plan* 2006, **21**:459–468.
12. Howe LD, Hargreaves JR, Gabrysch S, Huttly SRA: **Is the wealth index a proxy for consumption expenditure? A systematic review.** *J Epidemiol Community Health* 2009, **63**:871–877.
13. Howe LD, Hargreaves JR, Huttly SRA: **Issues in the construction of wealth indices for the measurement of socio-economic position in low-income countries.** *Emerg Themes Epidemiol* 2008, **5**:3.
14. Sembra RD, Campbell AA, Sun K, de Pee S, Akhter N, Moench-Pfanner R, Rah JH, Badham J, Kraemer K, Bloem MW: **Paternal smoking is associated with greater food insecurity among poor families in rural Indonesia.** *Asia Pac J Clin Nutr* 2011, **20**:618–623.
15. Hoa NB, Tiemersma EW, Sy DN, Nhung NV, Gebhard A, Borgdorff MW, Cobelens FGJ: **Household expenditure and tuberculosis prevalence in VietNam: prediction by a set of household indicators.** *Int J Tuberc Lung Dis* 2011, **15**:32–37.
16. Tampubolon G, Hanandita W: **Poverty and mental health in Indonesia.** *Soc Sci Med* 1982 2014, **106**:20–27.
17. Filmer D, Pritchett LH: **Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of India.** *Demography* 2001, **38**:115–132.
18. Somi MF, Butler JR, Wahid F, Njau JD, Kachur SP, Abdulla S: **Use of proxy measures in estimating socioeconomic inequalities in malaria prevalence.** *Trop Med Int Health* 2008, **13**:354–364.
19. Howe LD, Galobardes B, Matijasevich A, Gordon D, Johnston D, Onwujekwe O, Patel R, Webb EA, Lawlor DA, Hargreaves JR: **Measuring socio-economic position for epidemiological studies in low- and middle-income countries: a methods of measurement in epidemiology paper.** *Int J Epidemiol* 2012, **41**:871–886.
20. Kolenikov S, Angeles G: **Socioeconomic Status Measurement with Discrete Proxy Variables: Is Principal Component Analysis a Reliable Answer?** *Rev Income Wealth* 2009, **55**:128–165.
21. O'Donnell O, Doorslaer EV, Wagstaff A, Lindelow M: *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and Their Implementation.* Washington, D.C: World Bank; 2008.
22. Ferguson BD, Tandon A, Gakidou E, Murray CJL: **Estimating permanent income using asset and indicator variable.** In *Health Syst Perform Assess Debates Methods Empiricism.* Edited by Murray CJL, Evans DB. Geneva: World Health Organization; 2003:747–760.
23. Gwatkin DR, Rutstein S, Johnson K, Suliman E, Wagstaff A, Amouzou A: *Socio-Economic Differences in Health, Nutrition, and Population: Malawi.* Washington, DC: World Bank; 2007 [Country Reports on HNP and Poverty].
24. Gwatkin DR, Rutstein S, Johnson K, Suliman E, Wagstaff A, Amouzou A: *Socio-Economic Differences in Health, Nutrition, and Population: An Overview.* Country Reports on HNP and Poverty. Washington, DC: World Bank; 2007.
25. Gwatkin DR, Rutstein S, Johnson K, Suliman E, Wagstaff A, Amouzou A: *Socio-Economic Differences in Health, Nutrition, and Population: Burkina Faso.* Washington, DC: World Bank; 2007 [Country Reports on HNP and Poverty].
26. Rutstein SO, Johnson K: *The DHS Wealth Index. DHS Comparative Reports.* Calverton, MD: ORC Macro; 2004.
27. Nwaru BI, Klemetti R, Kun H, Hong W, Yuan S, Wu Z, Hemminki E: **Maternal socio-economic indices for prenatal care research in rural China.** *Eur J Public Health* 2012, **22**:776–781.
28. Pariyo GW, Ekirapa-Kiracho E, Okui O, Rahman MH, Peterson S, Bishai DM, Lucas H, Peters DH: **Changes in utilization of health services among poor and rural residents in Uganda: are reforms benefitting the poor?** *Int J Equity Health* 2009, **8**:39.
29. Nicholson RJ, Topham N: **Step-wise Regression and Principal Components Analysis in Estimating a Relationship in an Econometric Model\*1.** *Manch Sch* 1973, **41**:187–205.
30. Mokken RJ: *A Theory and Procedure of Scale Analysis: With Applications in Political Research.* Mouton & Co: The Hague; 1971.
31. van Schuur WH: **Mokken Scale Analysis: Between the Guttman Scale and Parametric Item Response Theory.** *Polit Anal* 2003, **11**:139–163.
32. Guio A-C, Gordon D, Marlier E: *Measuring Material Deprivation in the EU Indicators for the Whole Population and Child-Specific Indicators.* Publications Office of the European Union: Luxembourg; 2012 [EuroStat Methodologies and Working Papers].
33. Szeles MR, Fusco A: **Item response theory and the measurement of deprivation: evidence from Luxembourg data.** *Qual Quant* 2013, **47**:1545–1560.
34. Martini MC, Vanin C: **A Measure of Poverty Based on the Rasch Model.** In *Adv Theor Appl Stat.* Edited by Nicola T, Fortunato P, Avner B-H. Heidelberg: Springer; 2013:327–337.
35. Andreotti A, Minicuci N, Kowal P, Chatterji S: **Multidimensional profiles of health status: an application of the grade of membership model to the world health survey.** *PLoS One* 2009, **4**:e4426.
36. Ustun TB, Chatterji S, Mechbal A, Murray CJL, WHS Collaborating Groups: **The World Health Surveys.** In *Health Syst Perform Assess Debates Methods Empiricism.* Edited by Murray CJL, Evans DB. Geneva, Switzerland: World Health Organization; 2003:797–808.
37. Vietnam: *World Health Survey; 2003* [http://apps.who.int/healthinfo/systems/surveydata/index.php/catalog/92]
38. Harrell FE: *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis.* New York, NY: Springer; 2001.
39. Jolliffe IT: *Principal Component Analysis.* New York: Springer; 2010.
40. Van der Ark LA: **Mokken Scale Analysis in R.** *J Stat Softw* 2011, **20**:1–19.
41. Stochl J, Jones PB, Croudace TJ: **Mokken scale analysis of mental health and well-being questionnaire item responses: a non-parametric IRT method in empirical research for applied health researchers.** *BMC Med Res Methodol* 2012, **12**:74.
42. Watson R, van der Ark LA, Lin L-C, Fieo R, Deary IJ, Meijer RR: **Item response theory: how Mokken scaling can be used in clinical practice.** *J Clin Nurs* 2012, **21**:2736–2746.
43. Sijtsma K, Meijer RR, Andries van der Ark L: **Mokken scale analysis as time goes by: An update for scaling practitioners.** *Personal Individ Differ* 2011, **50**:31–37.
44. Loevinger J: **A systematic approach to the construction and evaluation of tests of ability.** *Psychol Monogr* 1947, **61**:i–49. Ed.
45. Loevinger J: **The technic of homogenous tests compared with some aspects of "scale analysis" and factor analysis.** *Psychol Bull* 1948, **45**:507–529.
46. Sijtsma K, Molenaar IW: *Introduction to Nonparametric Item Response Theory.* Thousand Oaks, CA: Sage Publications, Inc; 2002.
47. Van der Ark LA: **New Developments in Mokken Scale Analysis in R.** *J Stat Softw* 2012, **48**:1–27.
48. Watson R, Deary I, Austin E: **Are personality trait items reliably more or less "difficult"? Mokken scaling of the NEO-FFI.** *Personal Individ Differ* 2007, **43**:1460–1469.
49. R Core Team: *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing; 2013.



50. Fox J: *polycor: Polychoric and Polyserial Correlations. R package version 0.7-8*. 2010. [http://CRAN.R-project.org/package=polycor]
51. Kuijpers RE, Ark LAV der, Croon MA: **Standard Errors and Confidence Intervals for Scalability Coefficients in Mokken Scale Analysis Using Marginal Models**. *Sociol Methodol* 2013, **43**:42–69.
52. Do AM: *In Vietnam, For Every 100 People There are 145 Mobile Phones*. *Tech in Asia*. 2012 [http://www.techinasia.com/vietnam-100-people-145-mobile-phones/]
53. Adler M, Hetta J, Isacsson G, Brodin U: **An item response theory evaluation of three depression assessment instruments in a clinical sample**. *BMC Med Res Methodol* 2012, **12**:84.
54. Gerrard P, Goldstein R, Divita MA, Ryan CM, Mix J, Niewczyk P, Kazis L, Kowalske K, Zafonte R, Schneider JC: **Validity and reliability of the FIM instrument in the inpatient burn rehabilitation population**. *Arch Phys Med Rehabil* 2013, **94**:1521–1526.e4.
55. Galindo-Garre F, Hendriks SA, Volicer L, Smalbrugge M, Hertogh CPM, van der Steen JT: **The bedford Alzheimer nursing-severity scale to assess dementia severity in advanced dementia: a nonparametric item response analysis and a study of its psychometric characteristics**. *Am J Alzheimers Dis Other Demen* 2014, **29**:84–89.
56. Gordon D, Howe LD, Galobardes B, Matijasevich A, Johnston D, Onwujekwe O, Patel R, Webb EA, Lawlor DA, Hargreaves JR: **Authors' Response to: Alternatives to principal components analysis to derive asset-based indices to measure socio-economic position in low- and middle-income countries: the case for multiple correspondence analysis**. *Int J Epidemiol* 2012, **41**:1209–1210.

doi:10.1186/1742-7622-11-9

**Cite this article as:** Reidpath and Ahmadi: **A novel nonparametric item response theory approach to measuring socioeconomic position: a comparison using household expenditure data from a Vietnam health survey, 2003**. *Emerging Themes in Epidemiology* 2014 **11**:9.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

