Economics Bulletin

Volume 44, Issue 1

A summary measure of answers to statements (SMAS) in the case of ordered rating scales

Joseph Deutsch
Bar-Ilan University

Jacques Silber
Bar-Ilan University

Abstract

This paper suggests that an index recently introduced in the literature on the measurement of health achievement when only ordinal variables are available, could be used as a summary measure of answers to statements in the case of ordered rating scales. An empirical illustration based on data from the International Social Survey Program shows the usefulness of such an index.

Paper #2 in Special issue "In memory of Pr. Michel Terraza" One of the authors, Jacques Silber, was several times in the jury of PhD students or HDR (Habilitation to Direct Research) at the University of Montpellier. He had hence, more than once, the opportunity to meet the late Professor Michel Terraza. Jacques Silber remembers him not only as a very competent economist and econometrician but also as a charming person with a lot of humour. The present paper has been written in his memory.
Citation: Joseph Deutsch and Jacques Silber, (2024) "A summary measure of answers to statements (SMAS) in the case of ordered rating scales", *Economics Bulletin*, Volume 44, Issue 1, pages 283-294
Contact: Joseph Deutsch - jdeutsch@biu.ac.il, Jacques Silber - jsilber_2000@yahoo.com.
Submitted: September 22, 2021. **Published:** March 30, 2024.

Economics Bulletin

Picture credits: Virginie Terraza

Special issue “In memory of Professor Michel Terraza”

1. Introduction

There is quite a vast literature on the measurement of qualitative variation, that is, on the measurement of the dispersion of answers when the categories distinguished are unordered (see, for example, Simpson's, 1949; Mueller and Schuessler, 1961; Theil, 1967; Reynolds, 1977; Agresti and Agresti, 1978; Kvålseth, 1988; Berry and Mielke, 1992).

There have been also quite a few propositions to measure the dispersion of answers when the latter can be ordered (see, Leik, 1966; Berry and Mielke, 1992; Blair and Lacy, 1996 and 2000; Van der Cees, 2001; Allison and Foster, 2004; Tastle and Wierman, 2007; Apouey, 2007; Abul Naga and Yalcin, 2008; Kalmijn and Arends, 2010; Giudici and Raffinetti, 2011; Kobus and Milos, 2012; Lazar and Silber, 2013; Lv et al., 2015; Chakravarty and Marahaj, 2015; Peñalosa 2016; Cowell and Flachaire, 2017, and the section on Inequality in the survey of Silber and Yaloneztky, 2021). The indices introduced by these authors for the case or ordered answers could therefore be used to measure the extent of consensus between the answers provided by respondents to questions where they were asked whether they agree with a given statement. None of the measures proposed would however tell us to what extent on average people agree with such a statement. The purpose of this paper is to show that an index recently introduced (see, Apouey et al., 2020) in the literature on the measurement of health achievement when only ordinal variables are available, could be used as a summary measure of answers to statements (*SMAS*) in the case of ordered rating scales. Section 2 defines this index and describes its main features while Section 3 gives an empirical illustration based on data from the International Social Survey Program.

2. Applying achievement measures to derive a summary index of ordered rating scales

In a recent paper, Apouey et al. (2020) derived axiomatically a measure of achievement for the case of ordinal variables, that can take into account both the average level of achievement (e.g. average health or educational achievement) as well the inequality of the distribution of such achievements. In this section, we will argue that such an index can be used in the more general case when the available information concerns answers to questions where the respondents are asked to select a statement among various possible answers that are ordered. Individuals, for

example, could be asked to state how important they think having ambition or working hard is, if they want to succeed in life.

Assume therefore that there are n individuals and that a_i represents the answer selected by individual i with $1 \leq a_i \leq K$ where K refers to the number of possible answers. These answers denote the level of agreement with a given statement and they are assumed to be ranked by decreasing level of agreement with the statement under study. Let a be the vector indicating the answers given by the various individuals so that $a = \{a_1, \dots, a_i, \dots, a_n\}$. Let *SMAS* (summary measure of agreement with the statement) refer to the index summarizing the distribution of answers to this statement with $SMAS = h(a)$. Call now f_k the relative frequency of answer k $(k = 1 \text{ to } K)$ and let α be a parameter such that $0 < \alpha < 1$.

Using the achievement index introduced by Apouey et al. (2019) we suggest to summarize the distribution of answers to this statement via the index *SMAS* where

$$SMAS = h(a) = \sum_{k=1}^{K} f_k \frac{1-\alpha^{K-k}}{1-\alpha^{K-1}} \quad (1)$$

Apouey et al. (2020) have proven that such a summary measure satisfies Normalization, Independence, Weak Pareto Principle, Anonymity, Equity Principle, and Proportional Equality.

Normalization simply means that *SMAS* varies between 0 and 1.

According to the axiom of *Independence*, if the answer given by one individual changes without affecting the answers given by the other individuals, the variation in *SMAS* is independent of the initial answer given by the other individuals.

The *Weak Pareto Principle* implies that if every individual gives the same answer k, the index $SMAS$ will be higher than if every individual gives the same answer k', with $k' > k$.

The *Anonymity* axiom means that *SMAS* depends only on the answers given by the various individuals and not on any other characteristics of these individuals.

According to the *Equity principle*, changes in the answers given by two individuals from two further-apart levels to two "closer" levels will, ceteris paribus, increase *SMAS*.

Finally the axiom of *Proportional Equality* may be interpreted as follows. Suppose that originally everyone gave the same answer k. If now every individual selects the answer $(k+1)$ or if every individual chooses the answer $(k-1)$, the ratio of the changes in the index *SMAS* as a consequence of two such changes is independent of the initial answer k. In other words the ratio $\frac{h(k-1,\dots,k-1)-h(k,\dots,k)}{h(k,\dots,k)-h(k+1,\dots,k+1)}$ is independent of k.

Note that when the parameter α tends towards 1, the axioms of equity and proportional equality will not hold and it can be shown that *SMAS* may then be written as

$$SMAS = \sum_{k=1}^{K} f_k \frac{K-k}{K-1} \quad (2)$$

or, more simply, as

$$SMAS = \frac{1}{(K-1)} \sum_{k=1}^{K-1} F_k(s) \quad (3)$$

where $F_k(s)$ refers to the cumulative relative frequency of the different possible answers.

3. An empirical illustration

3.1. Data sources

The database for the present study is the 2009 International Social Survey Program which focused on Social Inequality. We looked at the answers given by respondents to questions on the importance of family wealth, parents' and own education.

3.2. Summarizing the answers to the various questions

For each question we summarized the distribution of the answers by using the index *SMAS* defined in (1). We took two scenarios into account. In the first case we assumed that the summary measure *SMAS* depended both on the location of the distribution and the dispersion of the answers, as explained in Section 2, and supposed that the parameter α was equal to 0.5. In the second case the value of α was assumed to be equal to 0.999 which corresponds more or less to the case $(\alpha \to 1)$ where the summary measure *SMAS* depends only on the location of the distribution.

In Table I we give the value of the measure *SMAS* to the following three questions:

- How important is coming from a wealthy family?
- How important is having well-educated parents?
- How important is having a good education yourself?

Note that for each questions the possible answers are: 1. Essential; 2. Very important; 3. Fairly important; 4. Not very important; 5. Not important at all.

Table I: Importance of coming from a rich family, of having well-educated parents and of having a good education

	Coming from a rich family	Coming from a rich family	Coming from a rich family	Coming from a rich family	Having well-educated parents	Having well-educated parents	Having well-educated parents	Having well-educated parents	Having a good education	Having a good education	Having a good education	Having a good education
Country	**Index** $\alpha = 0.5$	**Rank** $\alpha = 0.5$	**Index** $\alpha = 0.999$	**Rank** $\alpha = 0.999$	**Index** $\alpha = 0.5$	**Rank** $\alpha = 0.5$	**Index** $\alpha = 0.999$	**Rank** $\alpha = 0.999$	**Index** $\alpha = 0.5$	**Rank** $\alpha = 0.5$	**Index** $\alpha = 0.999$	**Rank** $\alpha = 0.999$
Argentina	0.508	41	0.339	38	0.719	31	0.511	28	0.880	28	0.695	29
Australia	0.629	27	0.409	27	0.781	16	0.555	16	0.920	7	0.776	9
Austria	0.722	14	0.501	16	0.775	17	0.544	21	0.911	12	0.755	12
Belgium	0.641	26	0.395	28	0.791	12	0.554	17	0.895	21	0.708	27
Bulgaria	0.756	9	0.569	8	0.804	10	0.613	6	0.914	9	0.777	8
Chile	0.673	24	0.480	22	0.787	13	0.596	10	0.912	11	0.762	10
China	0.868	1	0.709	1	0.916	1	0.776	1	0.942	2	0.828	2
Taiwan	0.674	23	0.475	23	0.693	35	0.480	34	0.848	35	0.668	33
Croatia	0.777	3	0.586	4	0.787	14	0.582	13	0.902	17	0.750	16
Cyprus	0.683	21	0.486	20	0.696	34	0.493	33	0.915	8	0.792	6
Czech Republic	0.622	29	0.420	26	0.628	40	0.407	40	0.795	41	0.601	40
Denmark	0.578	34	0.340	37	0.699	33	0.446	36	0.868	33	0.662	34
Estonia	0.724	13	0.490	18	0.809	9	0.592	12	0.914	10	0.761	11
Finland	0.514	40	0.295	41	0.553	41	0.333	41	0.846	37	0.632	38
France	0.531	37	0.323	39	0.804	11	0.596	11	0.888	24	0.716	24
Germany	0.704	18	0.481	21	0.820	6	0.606	8	0.940	4	0.807	5
Hungary	0.767	6	0.570	7	0.725	30	0.515	27	0.826	39	0.639	37
Iceland	0.577	35	0.348	35	0.749	22	0.507	30	0.893	22	0.713	26
Israel	0.722	15	0.510	14	0.747	27	0.545	20	0.879	29	0.718	22
Italy	0.763	8	0.539	9	0.748	25	0.518	26	0.906	14	0.755	13
Japan	0.626	28	0.388	30	0.667	38	0.429	38	0.796	40	0.567	41
South Korea	0.804	2	0.596	3	0.786	15	0.565	15	0.850	34	0.654	36
Latvia	0.743	10	0.538	10	0.766	19	0.552	18	0.908	13	0.752	15
Lithuania	0.594	32	0.389	29	0.690	36	0.466	35	0.870	31	0.680	31
New Zealand	0.521	39	0.312	40	0.730	29	0.494	32	0.902	18	0.732	20
Norway	0.595	31	0.359	32	0.660	39	0.414	39	0.844	38	0.621	39
Philippines	0.688	20	0.489	19	0.844	3	0.671	3	0.940	5	0.828	3
Poland	0.776	4	0.581	5	0.831	5	0.635	5	0.941	3	0.814	4
Portugal	0.733	11	0.518	12	0.717	32	0.498	31	0.847	36	0.660	35

Table I (cont.)

	Coming from a rich family	Coming from a rich family	Coming from a rich family	Coming from a rich family	Having well-educated parents	Having well-educated parents	Having well-educated parents	Having well-educated parents	Having a good education	Having a good education	Having a good education	Having a good education
Country	**Index $\alpha = 0.5$**	**Rank $\alpha = 0.5$**	**Index $\alpha = 0.999$**	**Rank $\alpha = 0.999$**	**Index $\alpha = 0.5$**	**Rank $\alpha = 0.5$**	**Index $\alpha = 0.999$**	**Rank $\alpha = 0.999$**	**Index $\alpha = 0.5$**	**Rank $\alpha = 0.5$**	**Index $\alpha = 0.999$**	**Rank $\alpha = 0.999$**
Russia	0.727	12	0.529	11	0.751	20	0.536	23	0.883	27	0.717	23
Slovakia	0.708	16	0.510	15	0.749	23	0.537	22	0.891	23	0.730	21
Slovenia	0.706	17	0.491	17	0.748	26	0.524	24	0.885	25	0.714	25
South Africa	0.765	7	0.599	2	0.853	2	0.690	2	0.949	1	0.844	1
Spain	0.671	25	0.470	24	0.811	8	0.607	7	0.885	26	0.706	28
Sweden	0.596	30	0.370	31	0.684	37	0.444	37	0.876	30	0.674	32
Switzerland	0.580	33	0.350	34	0.743	28	0.511	29	0.905	16	0.733	19
Turkey	0.703	19	0.518	13	0.836	4	0.649	4	0.906	15	0.749	17
Ukraine	0.769	5	0.581	6	0.767	18	0.566	14	0.898	20	0.754	14
United Kingdom	0.553	36	0.343	36	0.749	24	0.520	25	0.902	19	0.734	18
United States	0.681	22	0.466	25	0.817	7	0.600	9	0.932	6	0.792	7
Venezuela	0.525	38	0.357	33	0.750	21	0.550	19	0.870	32	0.687	30

It appears that coming from a rich family is the most important in China, South Korea, Croatia, Poland and Ukraine and South Africa, the ranking of these countries being slightly different when the parameter α is equal to 0.5 or 0.99, although in both cases China has the first rank. We thus observe that South Africa which has the second rank when ignoring the dispersion in the answers (when α is equal to 0.99) gets the seventh rank once such an inequality is taken into account (that is, when α is equal to 0.5). The countries which have the lowest rank as far as coming from a rich family is concerned are Argentina, Finland, New Zealand, Venezuela and France and Denmark, here also the ranking depending on whether α is equal to 0.5 or to 0.999. Here we can observe that Argentina has the 38^{th} rank when ignoring inequality but the 41^{st} when taking the dispersion of the distribution of the answer into account.

As far as having educated parents is concerned, we observe that this is considered the most important in the following countries: China, South Africa, the Philippines, Turkey and Poland, the ranking of these countries being a bit different when α is equal to 0.5 or 0.99, although in both cases China has the first rank. The countries in which having well-educated parents is the least important are respectively Finland, the Czech Republic, Norway, Japan and Sweden, the ranking of these five countries being identical when α is equal to 0.5 or 0.99. Note that for this questions there is also not much difference between the case where the dispersion of answers is taken into account in the ranking of the countries with the lowest rank.

Finally the countries where having a good education is considered the most important are South Africa, China, Poland, Germany and the Philippines , the ranking of these five countries varying slightly when α is equal to 0.5 or 0.99. Note however that the Philippines have the third rank when ignoring the dispersion of the answers but the fifth when taking into account this dispersion. The countries with the lowest ranking, as far as having a good education is considered, are respectively the Czech Republic, Japan, Hungary, Norway and Finland. Here also the ranking is a bit different when α is equal to 0.5 or 0.99. Here also we may note that Hungary has the 37^{th} rank when ignoring the dispersion of the answers but the 39^{th} rank when taking this dispersion into account.

In Table II we computed various correlations. In the upper triangle (above the diagonal where the correlation is evidently 1) we gave the correlation between the various indices computed while in the lower triangle we indicate the rank correlations between the various indices. Note first that, for

a given index, the correlations between the case where the parameter α is equal to 0.5 and that where it is equal to 0.999 are very high (above 0.97), whether we look at the correlation between the indices or at that between the ranking of the indices.

Table II: Correlations between the answers to the three questions.

	Coming from a rich family ($\alpha = 0.50$)	**Coming from a rich family ($\alpha = 0.999$)**	**Having well-educated parents ($\alpha = 0.50$)**	**Having well-educated parents ($\alpha = 0.999$)**	**Having a good education ($\alpha = 0.50$)**	**Having a good education ($\alpha = 0.999$)**
Coming from a rich family ($\alpha = 0.50$)	1	0.978	0.555	0.589	0.312	0.433
Coming from a rich family ($\alpha = 0.999$)	0.982	1	0.589	0.652	0.340	0.476
Having well-educated parents ($\alpha = 0.50$)	0.476	0.483	1	0.977	0.765	0.787
Having well-educated parents ($\alpha = 0.999$)	0.527	0.550	0.976	1	0.747	0.798
Having a good education ($\alpha = 0.50$)	0.364	0.359	0.765	0.731	1	0.973
Having a good education ($\alpha = 0.999$)	0.430	0.433	0.745	0.738	0.980	1

Note: Numbers above the diagonal refer to correlations between the values of the index *SMAS* in the different countries while the numbers below the diagonal refer to correlations between the rankings of the countries.

When looking at the correlation between the ranking of the countries as far as the value of the index for the question on the importance of coming from a rich family is concerned and the ranking of the countries for the index measuring the importance of having well educated parents, we see that the correlation is quite high but not very high (between 0.55 and 0.65, depending on the value of the parameter α). The corresponding correlations based only on the ranking of the countries are somewhat smaller (between 0.47 and 0.55).

For the correlation between the ranking of the countries as far as the value of the index for the question on the importance of coming from a rich family is concerned and the ranking of the countries for the index measuring the importance of being well educated, we observe an even lower correlation: between 0.31 and 0.48 for the correlation between the indices and between 0.36 and 0.43 for the correlation based on only the ranking.

Finally for the correlation between the ranking of the countries as far as the value of the index for the question on the importance of having well educated parents is concerned and the ranking of the countries for the index measuring the importance of being well educated, we observe much higher correlations (above 0.75 for the correlation between the values of the indices and above 0.73 for the correlation between the rankings).

References

Abul Naga, R. H., Yalcin, T. (2008) "Inequality Measurement for Ordered Response Health Data" *Journal of Health Economics* (6): 1614-1625.

Agresti, A. and B. F. Agresti (1978) "Statistical Analysis of Qualitative Variation" *Sociological Methodology* 9: 204-237.

Allison, R. A. and J. E. Foster (2004) "Measuring Health Inequality Using Qualitative Data" *Journal of Health Economics* 23(3): S. 505 – 524.

Apouey, B. (2007) "Measuring Health Polarization with Self-Assessed Health Data" *Health Economics* 16(9): S. 875–894.

Apouey, B., J. Silber and Y. Xu (2020) "On inequality-sensitive and additive achievement measures based on ordinal data" *Review of Income and Wealth* 66(2): 267-286.

Atkinson, A.B. (1970) "On the measurement of inequality" *Journal of Economic Theory* 2: 244–263.

Berry, K. J. and P. W. Mielke Jr. (1992) "Assessment of Variation in Ordinal Data" *Perceptual and Motor Skills* 74(1): S. 63–66.

Blair, J. and M. G. Lacy (1996) “Measures of Variation for Ordinal Data as Functions of the Cumulative Distribution” *Perceptual and Motor Skills* 82: 411-418.

Blair, J. and M. G. Lacy (2000) “Statistics of Ordinal Variation” *Sociological Methods and Research* 28(3): S. 251–280.

Cees van der Eijk (2001) “Measuring Agreement in Ordered Rating Scales” *Quality & Quantity* 35: 325–341.

Chakravarty, S. and B. Maharaj (2015) “Generalized Gini polarization indices for an ordinal dimension of human well-being” *International Journal of Economic Theory*, 11: 231-246.

Cowell, F. A. and E. Flachaire (2017) “Inequality with Ordinal Data” *Economica* 84(334): 290-321.

Duclos, J., Esteban, J.M. and Ray, D. (2004) “Polarization: Concepts, Measurement, Estimation” *Econometrica* 72(6): 1737-1772.

Esteban, J. and D. Ray (1994) "On the Measurement of Polarization" *Econometrica* 62(4): 819-851.

Giudici, P. and E. Raffinetti (2011) “A Gini Concentration Quality Measure for Ordinal Variable,” Serie Statistica 1/2011, Dipartimento di Economia, Statistica e Diritto, Università degli Studi di Pavia, Italy.

Kalmijn, W. M. and L. R. Arends (2010) “Measures of Inequality: Application to Happiness in Nations” *Social Indicators Research* 99:147–162

Kobus, M. and P. Milos (2012) “Inequality decomposition by population subgroups for ordinal data” *Journal of Health Economics* 31: 15– 21.

Kobus, M. and R. Kurek (2019) *“*Multidimensional polarization for ordinal data*”* *Journal of Economic Inequality* 17(3): 301-317.

Kvålseth T. O. (1988) “Measuring variation for nominal data” *Bulletin of the Psychonomic Society* 26 (5): 433-436.

Lazar, A. and J. Silber (2013) "On the cardinal measurement of health inequality when only ordinal information is available on individual health status" *Health Economics* 22: 106-113.

Leik, R. K. (1966) “A Measure of Ordinal Consensus” *The Pacific Sociological Review* 9(2): 85-90.

Lv, G, Y Wang and Y Xu (2015) “On a new class of measures for health inequality based on ordinal data” *Journal of Economic Inequality* 13(3): 465–477.

Mendelson, H. (1987) “Quantile-preserving spread” *Journal of Economic Theory* 42: 334-351.

Mueller, J. H., K. F. Schuessler and (1961) *Statistical Reasoning in Sociology*, Boston: Houghton Mifflin.

Peñaloza, R. (2016) "Gini coefficient for ordinal categorical data" Mimeo, Department of Economics, University of Brasília, Brazil.

Reardon, S. F. (2009) "Measures of ordinal segregation" *Research on Economic Inequality*, vol. 17. Emerald: Bingley, UK; pp. 129–155.

Reynolds, H. T. (1977). The *analysis of cross-classifications.* New York: Free Press.

Sarkar, S. and S. Santra (2020) "Extending the approaches to polarization ordering of ordinal variables" *Journal of Economic Inequality* 18: 421-440.

Silber, J. and G. Yalonetzky (2021) "Measuring Welfare, Inequality and Poverty with Ordinal Variables" in *Handbook of Labor, Human Resources and Population Economics*, K. F. Zimmermann K.F. (eds), Springer, Cham. Available at. https://doi.org/10.1007/978-3-319-57365-6_152-1

Simpson, E.H. (1949) "Measurement of diversity" *Nature* 163: 688.

Tastle, W. J. and M. J. Wierman (2007) "Consensus and dissention: A measure of ordinal dispersion" *International Journal of Approximate Reasoning* 45(3): S. 531–545.

Theil, H. (1967) *Economics and Information Theory*. Amsterdam: North-Holland.

Veenhoven, R. (2009) "International scale interval study: Improving the comparability of responses to survey questions about happiness" in V. Moller & D. Huschka (Eds.), *Quality of life and the millennium challenge: Advances in quality-of-life studies, theory and research*. Social Indicators Research Series 35, Springer, pp. 45–58.

Wolfson, M.C. (1994) "When Inequalities Diverge" *American Economic Review* 84(2): 353-358.