

Analysis of Risk Perception about Energy Technologies on Indian

Yicheng Wang

Summary

Energy technologies play important roles of fueling the engines of today's world. Meanwhile, they are viewed by some people as threats to their livelihood. To better explain how people with different background respond to risk, benefits and acceptance brought by energy technologies, two modules of questionnaires are conducted in India to investigate the relationship between people's perception and their living condition as well as economic political worldviews. In the first module, statistical analysis is aimed at finding the best way to aggregate responses in sub-questions. In the second module, a newly developed scale is compared to the well-known Kahan's Scale.

Introduction

Risks are omnipresent in today's world. Energy technologies like nuclear energy, if not managed carefully, can have chance of causing physical and mental harms to human. To decide if the states build up more energy plants, it is important for decision makers to take into account the public's perception about risk. However, the response/perception to the risks can vary from people to people, related to their races, genders, economic status and many other cultural factors. Sustainable Living Index (SLI) is proven to be good indicator for assessing the livelihood security of the poor [[Kamaruddin and Samsudin, 2014](#)]. In this study, we test different ways of extracting SLT to test the assumption that if people in poor living conditions tend to have higher risk perception on energy technologies and be harder to recognize the benefits brought by those technologies.

Beside the livelihood insecurity, the economic-political factors can also affect risk perception. Kahan's scale [[Kahan et al., 2007](#)] is developed, by drawing on both cultural theory of risk and social psychology, to examine how cultural worldview interact with the impact of genders and races on risk perception. But Kahan's scale has some problems when applied to certain groups of people: In India, differentiation of the groups separated by it is not fine-grained to make traits of groups distinctive, which will undermine the effect of prediction on risk perception. Therefore, a new scale is developed by Prerna Gupta to overcome this drawback.

In this study, the questions of interest are the following:

1. Which one of scale, Kahan's or Prerna's, can better reflect the relation of risk perception and economic-political factors?

2. What is the best way to combine various items of a scale – arithmetic versus geometric mean?
3. Is linear regression appropriate for this study?
4. Is there any limitations on sampling methods in this study?

Proposed Statistical Method

Data Description

There are two data sets involved in this study each of which was collected in one module of questionnaire. The subjects selected in the surveys are representatives of people from all levels of Indian society. Each question in questionnaire adopts 2-point likert, 3-point likert or 5-point likert.

In the first data set, subjects' perception about risks, benefits and acceptance of different energy technologies are recorded in the columns starting with **Ben_**, **Risk_** and **N_**. For example, **Risk_Solar** means people's attitudes towards the risk brought by solar energy plants represented by 5-point likert, i.e, Extremely Risky, Very Risky, Moderately Risky, Slightly Risky and Not Risky at all. Details of these columns can be found in the questionnaire. Moreover, SLI indicators are also included in the questionnaires, which are encoded as column names after the column **N_Reject**. The meaning of each variables can be founded in [[Kamaruddin and Samsudin, 2014](#)].

In the second data set, the likerts of benefits, risks and acceptance are recorded as well with the same name encoding in data columns. Questions of Kahan's scales are also included with the column names **K_**. Columns after **K_ERADEQ2** all represent Prerna's scales, meaning of which is explained in her questionnaire.

Statistical Analysis

For the first data set, the focus is on the way to aggregate sub-questions in the questionnaire into the six indexes–Human Asset, Physical Asset, Social Asset, Financial Asset, Natural Asset and Livelihood Outcomes. To compare the geometric means and arithmetic means, each dependent variable from **Ben_**, **Risk_** and **N_** is regressed on the six indexes by arithmetic means and then on those by geometric means. So, at each time of regression, R Squares of two models by two ways of aggregation are obtained. Then, after all dependent variables are iterated, we get two list of R Squares—one is for arithmetic means and the other one is for geometric means. It turns out that arithmetic mean is better than geometric in term of the mean of R Square. Moreover, it is worth noting that there are too many 0-1 variables in aggregation resulting in too many zeros in aggregated indexes when calculating the product in the geometric means.

In the second data set, Kahan's Scale and Prerna's Scale are compared by the way as the first data set. When dependent variables are iterated over, each one of them is performed regression on independent variables in Kahan's scale first and then those in Prerna's Scale. R Squares are recorded in each time of regression. At last, two list of R squared are compared. Prerna's Scale wins out with average of R Square 0.5821 much more than Kahan's Scale(0.16226).

Also, since the likerts are 5-points values and can be viewed as categorical variables, the classification methods are more appropriate than regression methods. Logistic regression is performed to compare two scales. For all the dependent variables, Prerna's Scale achieves the average AIC of 827.2122 while Kahan's Scale has a much worse AIC average of 1752.381.

Potential Limitation

In the Kahan's Scale, there are much less missing data than that in Prerna Scale, which also accounts for the relatively small R Square values obtained in Kahan's Scale. Also, R Square can only describe the proportion of uncertainty explained by linear model without reflecting any other possible non-linear uncertainty. Although AIC is a good indicator for whether logistic regression fits the dataset or not, more classification models, for example decision tree and SVM, should join the sets of "trial tubes" to test their performance and find the best one.

Conclusion

For the four questions of interest, we provide the following interpretation:

1. Prerna's Scale is better than Kahan's in interpreting people's perception on risk, benefits and characteristics of risk.
2. Arithmetic mean is strongly recommended to aggregate the answers for sub-questions.
3. Linear model is not appropriate for Kahan's Scale to interpret people's perception while it is fairly fine for Prerna's Scale. But classification models like Logistic regression are highly recommended.
4. Too many "Rather not say" answers for some questions make the data set shrink a lot from its expected size. Subjects for these questions should be carefully considered and sampled to avoid so many NAs.

Appendix

Data Exploratory Analysis

The heat map about correlation of variables is drawn. Generally, benefits variables–**Ben__** have negative considerations with risk variables, implying that people who think energy technology will bring up benefits have less perception on risks caused by it. Also, it is a good sign for the question setting that benefits variables do not show much correlation between themselves because it means that questions do not overlap with each other. So do risk variables.

```
ggplot(data = melted_cormat, aes(x=Var1, y=Var2, fill=value)) +  
  geom_tile()+  
  theme_minimal()+  
  theme(axis.text.x = element_text(angle = 90, vjust = 1,  
    size = 6, hjust = 1), axis.text.y =element_text(angle = 0, vjust = 1,  
    size = 6, hjust = 1))
```

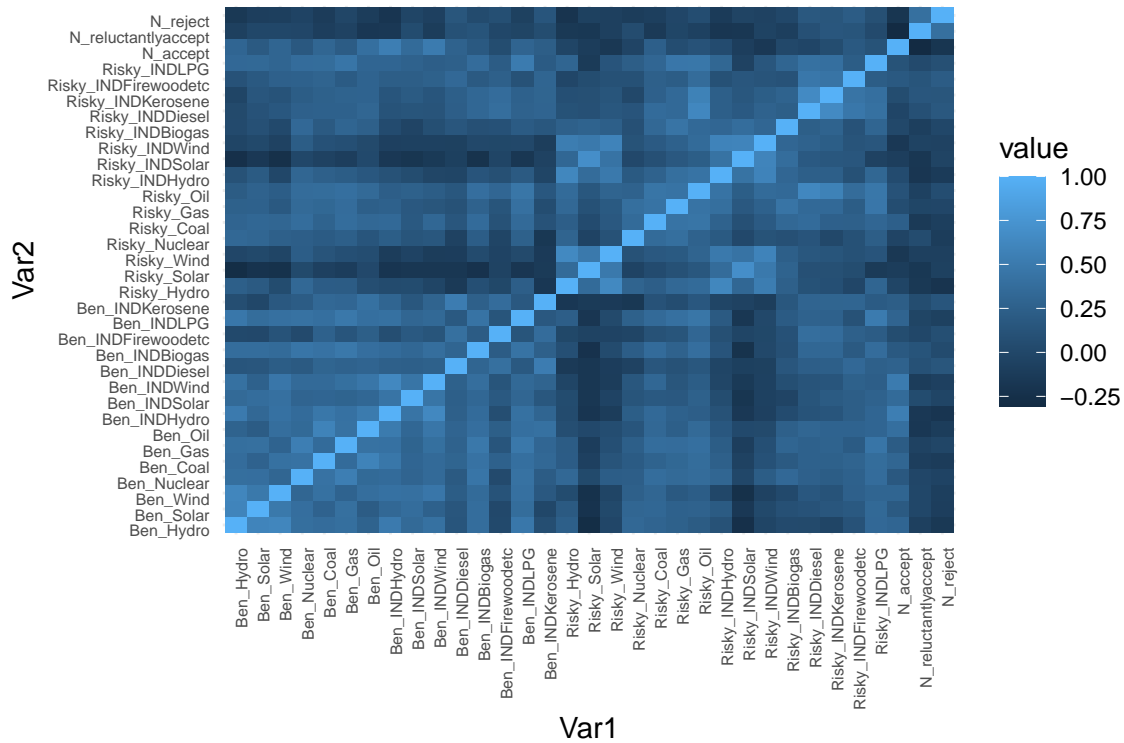


Figure 1: Heatmap of Correlation Coefficients of Risks, Benefits and Characteristics of Risk

Statistical Analysis

In the comparison of Kahan's Scale and Prerna's Scale, each dependent is used once to regress on the two scales as described in the Statistical Analysis of report, from which records of R Squares and AICs are

recorded. From Figure 2, we can see Prerna's Scale outperforms Kahan's Scale in achieving a much higher R Square value. But this may be attributed to non-linearity of relationship of Kahan's Scale to likerts. Therefore, logistic regression also is used to compare AIC of two scales. Figure 3 shows that Prerna's Scale obtain a much lower AIC than Kahan's Scale. So, it can be safe to say Prerna's Scale is better.

```
ggplot(data = rSquareResult , aes(y = R_Square, color =Scale)) + geom_boxplot() + theme(axis.ticks.x = element_b
```

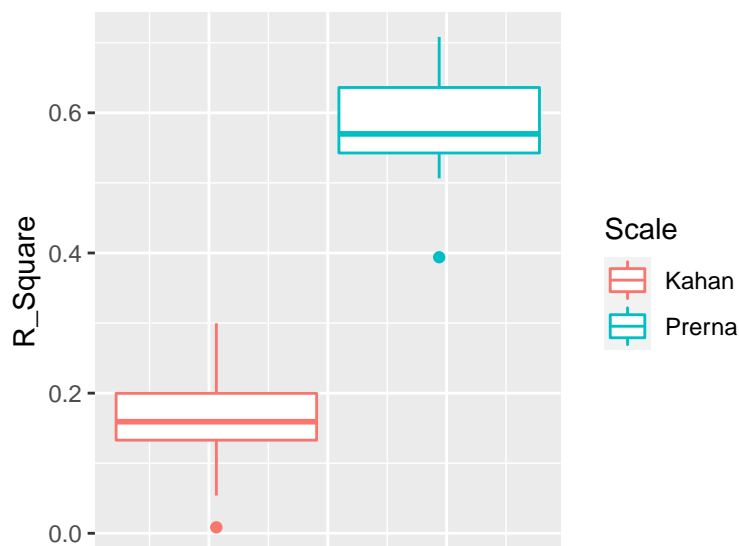


Figure 2: R Square obtained in Linear Regression

```
ggplot(data = AICResult , aes(y = AIC, color =Scale)) + geom_boxplot() + theme(axis.ticks.x = element_b
```

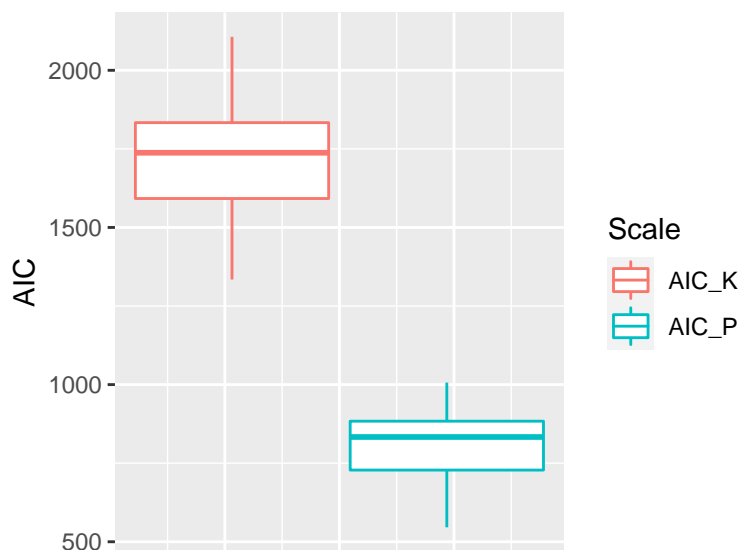


Figure 3: AIC achieved in Logistic Regression

References

- Dan M Kahan, Donald Braman, John Gastil, Paul Slovic, and CK Mertz. Culture and identity-protective cognition: Explaining the white-male effect in risk perception. *Journal of Empirical Legal Studies*, 4(3): 465–505, 2007.
- Roslina Kamaruddin and Shamzaeffa Samsudin. The sustainable livelihoods index: A tool to assess the ability and preparedness of the rural poor in receiving entrepreneurial project. *Journal of Social Economics Research*, 1(6):108–117, 2014.