Methods for the Identification of Semi-Plausible Response Patterns

Tayfun Terzi - T.Terzi@lse.ac.uk

SpRPs Problem Definition

New challenges concerning bias to measurement error have arisen due to the increasing use of paid participants: semi-plausible response patterns (SpRPs). SpRPs result when participants **only superficially process the information** of (online) experiments or questionnaires and attempt only to respond in a plausible way. This is due to the fact that participants who are paid are generally **motivated by fast cash**, and try to efficiently overcome objective plausibility checks and process other items only superficially, if at all. Thus, those participants produce not only useless but detrimental data, because they attempt to conceal their malpractice from the researcher. The consequences are biased estimations, blurred or even covered true effect sizes, and **contaminating valid models**. Existing and new approaches for the development of measures for the identification of SpRPs are discussed.

Data

Huang et al. (2012) - International Personality Item Pool (IPIP-300)

Experimental 1 Factor - 2 Conditions Design

Warning ($n = 84$)	"Sophisticated statistical control methods are be used to check for validity of responses and that responding without much effort would result in loss of credits."
Semi-plausible ($n = 64$)	"Respond without much effort but pretend that you want your laziness in filling out this survey to remain undetected."

Discussion

Initial findings show that presented measures qualify for the identification of not only model aberrant but semi-plausible responders.

Even simple methods like individual consistency measures significantly identify semi-plausible response patterns more often than participants in the 'Warning' conditions (higher percentage of semi-plausible responders exceed the critical Long String cut value).

Multivariate methods, especially aimed measures for identifying SpRPs in latent models, indicate hypothesis conform more semi-plausible responders (higher mean IC_i score). However, more research about their usefulness is needed.

References

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M. and DeShon, R. P. (2012), 'Detecting and deterring insufficient effort responding to surveys', *Journal of Business and Psychology* **27**(1), 99–114.

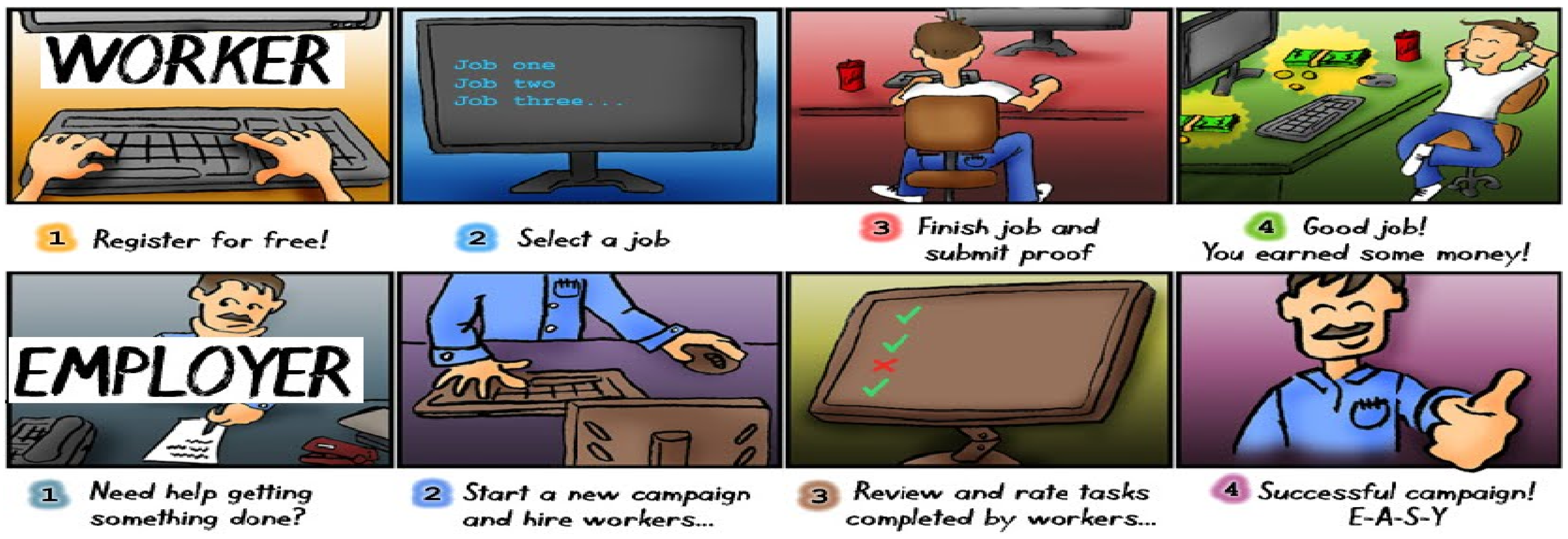
Lange, K., Westlake, J. and Spence, M. A. (1976), 'Extensions to pedigree analysis iii. variance components by the scoring method', *Annals of Human Genetics* **39**(4), 485–491.

Levine, M. V. and Rubin, D. B. (1979), 'Measuring the appropriateness of multiple-choice test scores', *Journal of Educational and Behavioral Statistics* **4**(4), 269–290.

Mahalanobis, P. C. (1936), 'On the generalized distance in statistics', *Proceedings of the National Institute of Sciences (Calcutta)* **2**, 49–55.

Reise, S. P. and Widaman, K. F. (1999), 'Assessing the fit of measurement models at the individual level: A comparison of item response theory and covariance structure approaches', *Psychological Methods* **4**(1), 3–21.

Get CASH for SURVEYS



Method

The goal is to derive SpRP identication measures in the structural equation (latent model) framework, by drawing on existing measures for the identification of diverse response patterns in the following related fields:

1. Person-Fit Indices in Educational Testing (e.g. Guttman Model, IRT)

In educational testing, item-response theoretical (IRT) models are fitted based on two parameters: (δ_j) Item difficulty & (θ_i) Person skill. P_{ij} indicates a participant's (i) probability of correctly solving an item (j). One way of identifying SpRPs is by estimating the 'individual' log-likelihood for each participant's response pattern under the model (person-fit):

$$l_{0,i} = \sum_{j=1}^{J}\{X_{ij} \ln P_{ij}(\theta_i) + (1 - X_{ij}) \ln[1 - P_{ij}(\theta_i)]\}$$

, where $l_{0,i}$ is small when a response pattern is model conform/consistent and vice versa (Levine and Rubin, 1979).

2. Outlier Analysis and Structural Equation Model Fit (e.g., multivariate χ^2-Test)

The Mahalanobis distance

$$D = (y_i - M)\Sigma^*(y_i - M)$$

, where M is the vector of means, (Mahalanobis, 1936) indicates the distance of a response pattern (y_i) based on the multivariate distribution (e.g., theoretical covariance matrix Σ). Furthermore, analogously to log-likelihood ratio tests for model fit, one can estimate the log-likelihood of a model at the level of an individual response pattern (Lange et al., 1976):

$$R_i = -.5\{p \ln(2\pi) + \ln|\Sigma^*| + D$$

, where p is the amount of observed variables. Hence,

$$IC_i = -2(R_{i,\text{restricted model}} - R_{i,\text{saturated model}})$$

produces a value that is directly interpretable as an individual's contribution to the overall model χ^2, where large positive IC_i values indicate patterns with larger contributions to the overall model misfit (Reise and Widaman, 1999).

3. Individual Consistency in Personality Measurement (e.g., Long String)

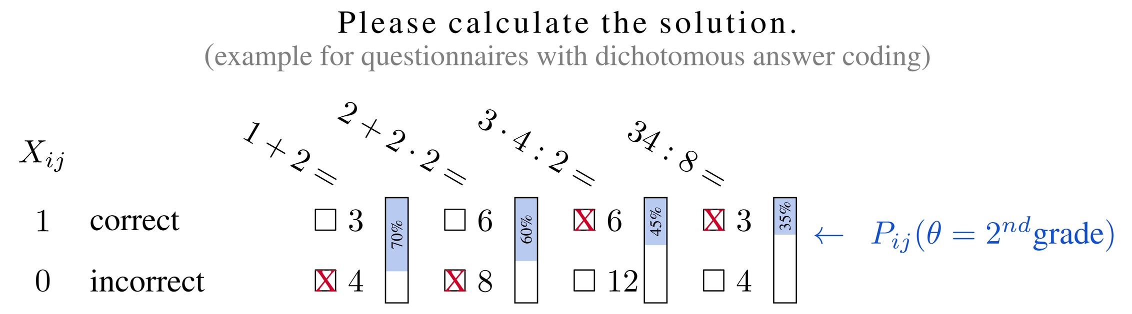
Long String indices indicate the longest consecutive occurrence of the same answer category for each individual. High LS_i values on any of the available answer categories indicate less plausible response pattern.

4. Para-Data (e.g., Response Time)

Assessing an individual's response time per item can also help to identify SpRPs (e.g. too long or too short response time).

Intuitive Identification

The following examples illustrate three of the discussed identification measures.

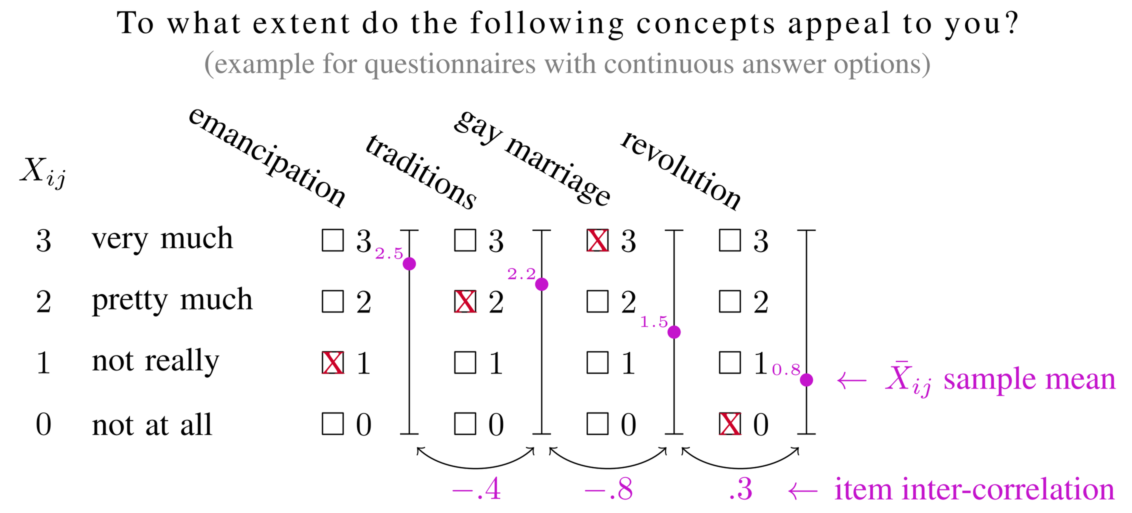


Above is the least plausible response pattern shown:

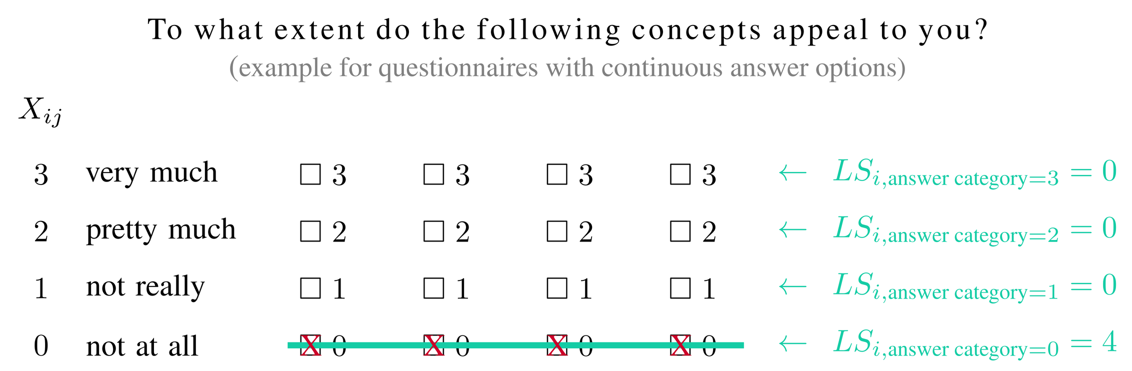
$$l_{0,i} = \{0 \cdot \ln(.7) + (1-0)\ln[1-.7]\} + \{0 \cdot \ln(.6) + (1-0)\ln[1-.6]\}$$
$$\{1 \cdot \ln(.45) + (1-1)\ln[1-.45]\} + \{1 \cdot \ln(.35) + (1-1)\ln[1-.35]\} = -3.96877$$

vs. a perfect model consistent response pattern:

$$l_{0,i} = \{1 \cdot \ln(.7) + (1-1)\ln[1-.7]\} + \{1 \cdot \ln(.6) + (1-1)\ln[1-.6]\}$$
$$\{0 \cdot \ln(.45) + (1-0)\ln[1-.45]\} + \{0 \cdot \ln(.35) + (1-0)\ln[1-.35]\} = -1.89612$$



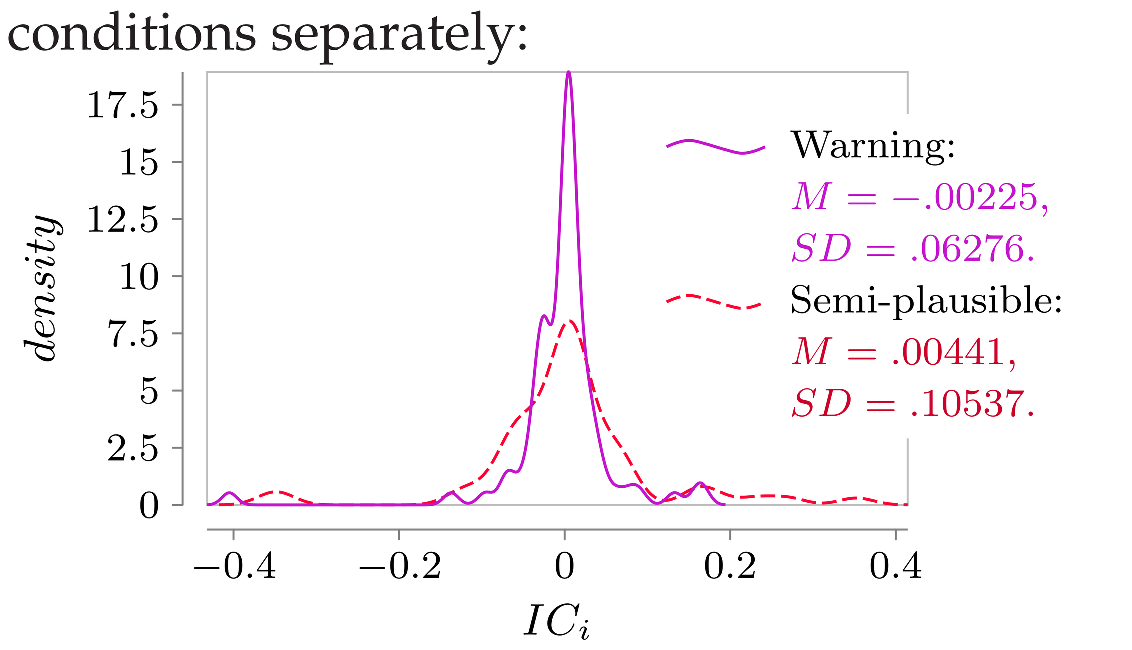
Above is a less plausible response pattern shown.
This response pattern does not follow (even oppose) the underlying model, which is (partly) specified by the correlations between adjacent items.



Above is a less plausible response pattern shown.
The longest consecutive string for answer category 'not at all' takes on the maximum value 4 which might exceed the critical value in relation to questionnaire length.

Results

Density distribution of IC_i values for both conditions separately:



Cumulative distribution of Long String values in answer category $u = 3$: