A TWO-COMPONENT MARKOV SWITCHING REGRESSION MODEL*

Roberto Colombi [1] and Sabrina Giordano[2]

[1] Department of Management, Information and Production Engineering, University of Bergamo, (e-mail: roberto.colombi@unibg.it)

[2] Department of Economics, Statistics and Finance "Giovanni Anania", University of Calabria, (e-mail: sabrina.giordano@unical.it)

ABSTRACT: The proposed approach addresses the issue of response styles in longitudinal ordered categorical data, where respondents tend to endorse certain options on a Likert scale regardless of the item's content. These response styles, including middle, extremes, acquiescence, and disacquiescence, can introduce bias in the results. To tackle this, the approach uses a Markov switching logit model with two latent components. One component captures serial dependence and respondent's unobserved heterogeneity, while the other accommodates the responding attitude (RS or no-RS). The responses' dependence on covariates is modeled using a flexible stereotype logit model with parameters varying based on the two latent components.

KEYWORDS: Latent variables; Response styles; Stereotype logit models.

Motivation and Models

Longitudinal ordered categorical data is susceptible to response styles, where respondents, when asked to assess items using Likert scales at various time points, tend to consistently select only a few specific options on the rating scale, regardless of the item's actual content.

Numerous studies in psychometrics and statistics have explored different types of response styles (RS) and their consequences (Van Vaerenbergh & Thomas, 2013). The commonly recognized response styles include acquiescence, disacquiescence, extreme and middle RS as described in Baumgartner & Steenkamp, 2001, among many others.

*The author Sabrina Giordano received partial financial support by MUR, grant number 2022XRHT8R - The SMILE project and by COST Action CA19130 Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry (FinAI), funded by COST (European Cooperation in Science and Technology).

Ignoring the RS mechanism can introduce response heterogeneity, biases the estimated model parameters, and consequently leads to inaccurate results (e.g., Colombi *et al.*, 2021).

In this study, our goal is to account for the temporal evolution of the RS behavior, which contrasts with previous approaches that either ignore it or model it as a time-invariant latent trait or random effect (Billiet & Davidov, 2008; Schauberger & Tutz, 2022).

Recognizing the significance of RS and its temporal dynamics, we propose a Markov switching model (Fruhwirth Schnatter, 2001) driven by a bivariate latent Markov chain. One component of this chain has k states or regimes, known as the k-regime switching indicator, which captures serial dependence and respondent heterogeneity due to unobserved covariates. The other binary component, called the response style regime switching indicator, dictates whether respondents answer according to an RS or use the rating scale appropriately to accurately represent their feelings.

Given the k-regime switching indicator, the observed categorical responses' dependence on time-varying and subject-specific covariates under the no-RS regime is modeled using a stereotype logit model (Anderson, 1984), while under the RS regime, it is modeled using a parallel local logit model with restricted intercepts, accommodating the tendency of respondents to select categories due to RS.

This approach adds a contribution to the literature on multivariate Markov chains in the context of Markov switching models (e.g., Pohle *et al.*, 2021, among others). The novelty of our approach, which extends existing models for longitudinal categorical data (e.g., Bartolucci *et al.*, 2012), lies in providing a Markov switching regression model for ordered responses that simultaneously considers attitude towards response styles, unobserved heterogeneity, serial dependence, and the impact of time-varying covariates.

The fundamental assumption in our current approach is that transition probabilities remain identical across subjects, as unit-specific covariate effects are accounted for at the observation level. However, an alternative scenario has been explored by Colombi *et al.*, 2023, where a non-homogeneous latent process is considered. In this case, the initial probabilities can be influenced by time-invariant regressors, while the transition probabilities by time-specific covariates. This approach considers the restriction of subject and time-invariant observation probability functions. In the mentioned paper, the observed variables are treated as indicators of a latent construct of interest, allowing covariates to naturally affect only the latent component of the model. The primary focus in the proposed work is centered on logit regression models fea-

turing time-varying parameters for the observable variables, with the k-regime switching indicator serving as a tool to model both unit heterogeneity and time dependence arising from unobserved covariates. For a deeper comprehension of the proposed model and real-world applications, refer to Colombi & Giordano, 2023's work, which provides extensive details on both aspects.

Our approach has potential applications in various longitudinal surveys that collect opinions on health status, risk of illness, economic difficulties, the impact of climatic events, discriminatory and racist beliefs, and political attitudes. These surveys may reveal biased perceptions due to response styles that vary over time, reflecting the ever-changing nature of human behavior.

References

ANDERSON, J. A. 1984. Regression and ordered categorical variables. *Journal of the Royal Statistical Society: Series B (Methodological)*, **46**(1), 1–22.

BARTOLUCCI, F., FARCOMENI, A., & PENNONI, F. 2012. *Latent Markov Models for Longitudinal Data*. CRC Press.

BAUMGARTNER, H., & STEENKAMP, J. B. E. M. 2001. Response styles in marketing research: a cross-national investigation. *Journal of Marketing Research*, **38**, 143–156.

BILLIET, J. B., & DAVIDOV, E. 2008. Testing the stability of an acquiescence style factor behind two interrelated substantive variables in a panel design. *Sociological Methods & Research*, **36**(4), 542–562.

COLOMBI, R., & GIORDANO, S. 2023. Markov Switching Stereotype Logit Models for Longitudinal Ordinal Data Affected by Unobserved Heterogeneity in Responding Behavior. *Submitted (second revision)*.

COLOMBI, R., GIORDANO, S., & TUTZ, G. 2021. A Rating Scale Mixture Model to Account for the Tendency to Middle and Extreme Categories. *Journal of Educational and Behavioral Statistics*, **46**(6), 682–716.

COLOMBI, R., GIORDANO, S., & KATERI, M. 2023. Hidden Markov Models for Longitudinal Rating Data with Dynamic Response Styles. *Statistical Methods and Applications (in press)*.

FRUHWIRTH SCHNATTER, S. 2001. *Finite Mixture and Markov Switching Models*. Springer.

POHLE, J., LANGROCK, R., SCHAAR, M. VAN DER, KING, R., & JENSEN, F. H. 2021. A primer on coupled state-switching models for multiple interacting time series. *Statistical Modelling*, **21**(3), 264–285.

SCHAUBERGER, G., & TUTZ, G. 2022. Multivariate ordinal random effects models including subject and group specific response style effects. *Statistical Modelling*, **22**(5), 409–429.

VAN VAERENBERGH, Y., & THOMAS, T. D. 2013. Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research*, **25**(2), 195–217.