Task-Level Error Scale Modeling Using Eye and Emotion Tracking Data

Roger Bailey and Jeff Dotson

November 15, 2016

1 Introduction

For discrete choice experiments to be informative, respondents must actively weigh the tradeoffs between the alternatives in a choice task and make their decisions accordingly. Respondents that are more engaged in these tasks make choices that a more consistent with their underlying preferences, leading to estimated coefficients that yield better predictive results. Though there are modeling techniques that allow for individual respondents to differ in terms of their level of engagement, the reality is that an individual respondent's level of engagement may change significantly between choices.

The purpose of this project is to provide a practical solution to modeling respondent engagement in each task. This is accomplished through the use of task-level tracking data, including response time, mouse-tracking, eye-tracking, and emotion measures. By allowing for different levels of engagement, choice tasks wherein the respondent is more engaged can be weighted more heavily in the estimation, allowing for better model fit and improved predictive power. Moreover, the simplicity of the proposed model lends itself easily to application in any discrete choice experiment where tracking data can be collected.

2 Previous Work

When respondents engage in discrete choice experiments, they are typically assumed to evaluate the alternatives within each choice task with equal care. This assumption manifests itself in the standard multinomial logit model via the assumption of a constant error scale. More recent modeling contributions, such as the generalized multinomial logit model of [Fiebig et al., 2010] and [Feit, 2009], allow for heterogeneity in the error scale across respondents, but still assume a fixed error scale across tasks. Though these models may not handle task-level changes in respondent engagement, the results demonstrate that error scale can be effectively modeled using covariates in the upper level of a hierarchical model. This bodes well for the underlying hypothesis of this project: that task-level data for each respondent can be informative of task-level error scale.

Recent research by [Meißner et al., 2016] demonstrates that eye-tracking data can be informative of how respondents make choices. Given the potential uses of tracking data as a proxy for respondent engagement at the task-level, we seek a model that can include additional measures. The literature does include techniques that could be used at the task-level to include these covariates, but these techniques are complex. For example, [Otter et al., 2008] construct a model that includes response time as a way to improve model fit in a Poisson race model of choice. While the inclusion of response time is shown to improve fit, the nature of response time lends itself easily to the Poisson race model, and other task-level covariates cannot be integrated as easily. As a result, the complexity of the poisson race model make it ill-suited to our goal of general application in discrete choice, whereas the parsimonious model we propose below can include any number of tracking measures in a way that is easy to interpret.

3 Modeling Respondent-Task Level Error

A parsimonious model of task-level error can be estimated using task-level tracking variables as proxy variables for respondent engagement. These tracking measures are included as covariates in the upper level of a hierarchical model, with error scale assumed to be a function of these variables at the respondent and respondent-task level. This structure is an extension of the generalized multinomial logit model mentioned above.

3.1 Choice Model

The choice by respondent h in choice task t, denoted y_{ht}, is given by

$$y_{ht} = \max\{X_{jt}\left(\frac{\beta_h}{\lambda_{ht}}\right) + \epsilon_{ht}\}$$

where X_{jt} is a vector of the attributes of alternative j in choice task t, β_h is the vector of part-worth utilities for respondent h, ϵ_{ht} is the error for respondent h in choice task t, and λ_{ht} is the error scale for task t.

The error scale λ_{ht} is modeled as:

$$\lambda_{ht} = \exp((Tr1)_h\delta + (Tr2)_{ht}\omega + \xi_h)$$

where $(Tr1)_h$ is the set of means of respondent h's tracking variables, mean-centered using the grand means of these variables across all respondents. $(Tr2)_{ht}$ is the set of respondent h's task-level tracking variables, mean-centered by the respondent-level means of these variables across all tasks.

As in [Feit, 2009] and other work, mean-centering these covariates is required for identification. The standard multinomial logit model achieves identification by assuming an error scale of one. For identification in this model, the expectation of λ_{ht} is one. By mean-centering $(Tr1)_h$ and $(Tr2)_{ht}$, it must be that $(Tr1)_h\delta$ and $(Tr2)_{hi}\omega$ have an expectation of zero across the respective variation of the respondent-level and aggregate-level tracking data. Thus, the expected respondent-level component and task-level component of λ_{ht} have an expectation of zero, resulting in the required identification restriction.

3.2 Upper Level

As in the standard heirarchical model for discrete choice experiments, the heterogeneous part-worth utilities are given by

$$\beta_h \sim N(\Delta Z_h, V_\beta)$$

where Z_h is a vector of covariates for respondent h, and Δ converts these covariates into part worth utility. Note that since $(Tr1)_h$ includes respondent-level covariates, these covariates could be included in Z_h. Moreover, other non-tracking covariates could potentially be included in $(Tr1)_h$ to inform the distribution of heterogeneity for λ.

3.3 Estimation

Estimation of this model is achieved using Bayesian techniques with standard normality assumptions for the hyper-parameters. The error term ϵ is assumed to be distributed extreme-value, resulting in an extension of multinomial logit model. The respondent-level error term in the error scale, ξ is assumed to be normally distributed. Estimation proceeds using a hybrid MCMC sampler.

This sampler has been constructed and sun on multiple simulated data sets, and the true values used to simulate the data are all at or near their 95% credible intervals from the posterior sample.

4 Conclusion

The goal of this project is to leverage advances in eye and emotion tracking technology to improve the results from discrete choice experiments. This will be accomplished by using these task-level tracking variables as a proxy measure for respondent engagement, and weighting choices made under high engagement more heavily in the estimation of the model. This is achieved by allowing for task-level deviations in the respondent's error scale. The end result should provide a modeling technique that consistently improves model fit and predictive results.

References

[Feit, 2009] Feit, E. M. (2009). Extending the generalized multinomial logit model: Error scale and decision-maker characteristics. *Available at SSRN 1566068*.

[Fiebig et al., 2010] Fiebig, D. G., Keane, M. P., Louviere, J., and Wasi, N. (2010). The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science*, 29(3):393–421.

[Meißner et al., 2016] Meißner, M., Musalem, A., and Huber, J. (2016). Eye tracking reveals processes that enable conjoint choices to become increasingly efficient with practice. *Journal of Marketing Research*, 53(1):1–17.

[Otter et al., 2008] Otter, T., Allenby, G. M., and Van Zandt, T. (2008). An integrated model of discrete choice and response time. *Journal of Marketing Research*, 45(5):593–607.