Judgment and Decision Making, Vol. 15, No. 5, September, 2020, pp. 861–862

Note on Birnbaum and Wan (2020): True and error model analysis is robust with respect to certain violations of the MARTER model

Michael H. Birnbaum* Bonny Quan†

Abstract

The Markov True and Error (MARTER) model (Birnbaum & Wan, 2020) has three components: a risky decision making model with one or more parameters, a Markov model that describes stochastic variation of parameters over time, and a true and error (TE) model that describes probabilistic relations between true preferences and overt responses. In this study, we simulated data according to 57 generating models that either did or did not satisfy the assumptions of the True and Error fitting model, that either did or did not satisfy the error independence assumptions, that either did or did not satisfy transitivity, and that had various patterns of error rates. A key assumption in the TE fitting model is that a person's true preferences do not change in the short time within a session; that is, preference reversals between two responses by the same person to two presentations of the same choice problem in the same brief session are due to random error. In a set of 48 simulations, data generating models either satisfied this assumption or they implemented a systematic violation, in which true preferences could change within sessions. We used the true and error (TE) fitting model to analyze the simulated data, and we found that it did a good job of distinguishing transitive from intransitive models and in estimating parameters not only when the generating model satisfied the model assumptions, but also when model assumptions were violated in this way. When the generating model violated the assumptions, statistical tests of the TE fitting models correctly detected the violations. Even when the data contained violations of the TE model, the parameter estimates representing probabilities of true preference patterns were surprisingly accurate, except for error rates, which were inflated by model violations. In a second set of simulations, the generating model either had error rates that were or were not independent of true preferences and transitivity either was or was not satisfied. It was found again that the TE analysis was able to detect the violations of the fitting model, and the analysis correctly identified whether the data had been generated by a transitive or intransitive process; however, in this case, estimated incidence of a preference pattern was reduced if that preference pattern had a higher error rate. Overall, the violations could be detected and did not affect the ability of the TE analysis to discriminate between transitive and intransitive processes.

Keywords: simulations, risky decision making, error models, Markov model, sequential effects, robustness

Birnbaum and Wan (2020) proposed an extension of basic true and error (TE) theory: MARkov True and ERror (MARTER) theory, in which parameters of a risky decision making model change gradually over time according to a random walk, generating different true preference patterns at different times. They presented a computer program that can be used to simulate data according to this model and showed that the simulated data can be properly analyzed by computer programs that fit the TE models.

They also showed that TE analysis can correctly identify whether the generating model had satisfied transitivity or if it violated transitivity, and that methods proposed by others to test properties of binary response probabilities, such as testing Weak Stochastic Transitivity and the Triangle Inequality, for example, as in Regenwetter, et al. (2011) were not capable of correctly distinguishing whether a transitive or intransitive generating model had been used to simulate the data. When the data were generated from a MARTER model, the TE analysis correctly detected the generating model and gave accurate estimates of the parameters used to generate the data, even in cases where other methods of analysis failed to correctly identify the process.

It is often said that all models are false but some models can still be useful. A magnetic compass does not point exactly True North, and flat maps are only approximations of the curved earth's surface, and yet these models proved useful in navigation. Signal Detection theory, factor analysis, and analysis of variance are similar quantitative models that are regarded as "robust". That is, they are regarded as useful for answering substantive questions about psychological theory even when their assumptions are only approximations.

In this follow-up to Birnbaum and Wan (2020), we conducted a series of simulations to investigate what happens

We thank Pele Schramm for comments and discussions, including discussion of his own simulations using Bayesian analysis related to this topic (see Schramm, 2020).

*Department of Psychology, California State University, Fullerton. Email: mbirnbaum@fullerton.edu.

†Palos Verdes. Email: quanyi306@gmail.com

if we generate data with violations of the MARTER model and then use a TE model that is only an approximation of the generating process. The computer program of Birnbaum and Wan (2020), *MARTER_sim.htm*, can be used not only to simulate data from MARTER models, but it is also capable of generating data from stochastic processes that violate certain of the assumptions.

We generated data that either satisfied or violated two key TE assumptions. We then asked three related questions: First, would TE analysis correctly detect the violations, so that an investigator would at least be aware of the lack of fit of the model? Second, would violations of the assumptions lead to wrong conclusions regarding transitivity? Third, would violations lead to biased parameter estimates?

We explored violations of these two key assumptions: First, the TE fitting model assumes that people do not change their true preferences in the short time interval of a session; Second, the TE fitting model we used assumed that error rates are independent of true preferences.

We ran 59 simulated sets of data from a hypothetical study of transitivity of preference, each with 10,000 sessions of data with two replications per session, similar to those described in Birnbaum and Wan (2020). Half of the simulated data sets were generated from a transitive process and half from an intransitive process. Half of the data sets were generated from process that satisfied the TE fitting model assumptions, and half violated those assumptions. Further, different stochastic decision models were used and different combinations of error rates were combined with those.

In all of the cases generated by MARTER models (that satisfied the TE assumptions, i.e., "control" conditions), the TE fitting model did an excellent job of identifying the model and recovering the parameters that had been used to generate the data. The TE fitting model always correctly identified whether the generating model was transitive or intransitive. These "control" conditions thus replicate and extend the findings reported by Birnbaum and Wan (2020) to new examples.

In all of the cases generated with violations of the TE fitting model, the tests of fit were large and significant, showing that an investigator could, in principle, realize that there was a systematic violation of the fitting model. That is, the model analysis does not "sweep under the rug" either of these types of violations.[1]

Despite systematic violations in these cases, the TE model correctly identified in every case whether the generating model had been transitive or intransitive. Because the main purpose of the hypothetical experiment was to test transitivity, this apparent robustness of the TE analysis is potentially the most important result of the simulations.

When the assumption that true preferences are stable within a session was violated, the estimates of incidence of preference patterns were surprisingly accurate, but the estimated error rates were inflated by the fact that true changes of preference were absorbed into the error component of the model.

In the cases where error rates were systematically different for different true preferences, it was found that the estimated incidences of preference patterns were decreased when the error rates associated with those patterns were greater, and relatively increased when the error rates were smaller for those patterns.

In summary, the TE fitting model appears to be quite robust with respect to these two sources of violations of the model. Although our conclusions must be restricted to the examples we have studied so far, we think that these conclusions are more generally applicable, but we are not yet able to provide analytic proofs, nor have we yet been able to deduce formulas to describe the limits (if any) of the conclusions that the TE fitting model is quite robust with respect to these two potential sources of violation.

Methods based on analysis of binary response proportions were unable to detect which datasets had been generated from transitive or intransitive models. All of the datasets would be declared by those methods to be compatible with transitive models, despite the fact that half of the datasets had been generated from intransitive models.

We have written a complete paper with the details of our simulations and results, which is available at the journal's Website via the following link: http://journal.sjdm.org/20/200413b/supp.pdf

[1]Of course, if an investigator used small samples, there is always the risk of a Type II error.

References

Birnbaum, M. H., & Wan, L. (2020). MARTER: Markov true and error model of drifting parameters. *Judgment and Decision Making, 15*(1), 47–73.

Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2011). Transitivity of Preferences. *Psychological Review, 118*, 42–56.

Schramm, P. (2020). The individual true and error model: Getting the most out of limited data. *Judgment and Decision Making, 15*(5), 851–860.