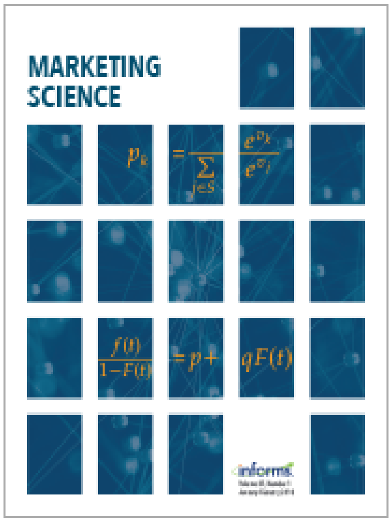

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Commentary—Discussion of "Alleviating the Constant Stochastic Variance Assumption in Decision Research: Theory, Measurement, and Experimental Test"

Jordan Louviere, Joffre Swait,

MARKETING SCIENCE
Vol. 29, No. 1, January–February 2010, pp. 18–22
ISSN 0732-2399 | EISSN 1526-548X | 10 | 2901 | 0018

DOI 10.1287/mksc.1080.0474

Commentary

Discussion of "Alleviating the Constant Stochastic Variance Assumption in Decision Research: Theory, Measurement, and Experimental Test"

Jordan Louviere
Centre for the Study of Choice (CenSoC), Faculty of Business, University of Technology, Sydney,
Broadway, New South Wales 2007, Australia, jordan.louviere@uts.edu.au

Joffre Swait
Faculty of Business, University of Alberta, Edmonton, Alberta T6G 2R6, Canada, swaitjd@swaitworld.org

We discuss the Salisbury and Feinberg paper [Salisbury, L. C., F. M. Feinberg. 2010. Alleviating the constant stochastic variance assumption in decision research: Theory, measurement, and experimental test. *Marketing Sci.* **29**(1) 1–17], setting their contribution in the historical context of the wider literature on the role of error variability in discrete choice models. We discuss the seminal nature of their contribution and suggest that the paper should be required reading for current and future Ph.D. students.

Key words: choice models; stochastic utility; variance
History: Received: October 23, 2008; accepted: November 11, 2008. Published online in *Articles in Advance* November 3, 2009.

It is an honor to be invited to discuss this paper by Salisbury and Feinberg (2010; hereafter S&F), which we hope will be seen as a watershed in the application of random utility theory (RUT) to behavioral research problems in marketing and other fields.

S&F emphasize the importance of understanding and analyzing the entire utility (preference, decision) construct, including both systematic and stochastic (random) components. Their paper focuses on potential confounds that arise because of violations of constant error variance assumptions. We see their paper as a seminal, groundbreaking contribution to the analysis of choice response data in decisionmaking and choice behavior, which we hope will be required reading for current and future Ph.D. students in marketing and other fields. We also hope their paper will lead to more communication and interaction among researchers in consumer behavior (CB) and behavioral decision theory (BDT) with researchers in other areas in marketing and related fields.

Of particular note are the following takeaways germane to all researchers interested in choices: (a) choices are discrete dependent variables; (b) design theory exists to study these types of dependent variables under controlled conditions (e.g., Street and Burgess 2007); (c) there is a large literature that addresses the analysis and modeling of discrete choices and other related phenomena in mathematical psychology, discrete multivariate statistics, and econometrics; and (d) failure to fully understand and apply the learning from this body of literature significantly diminishes the impacts of scientific contributions and perpetuate artificial barriers to collaboration and communication. This also serves to remind us of our intellectual and scientific obligations to understand and acknowledge relevant literature, while at the same time reminding us that like any product or service, managers of scientific or scholarly products must constantly be alert to new developments and insights from little-known or unexpected areas lest the value of their products be undermined or lost to more innovative or superior advances. S&F provide us with a clear example of a superior innovation likely to take many by surprise.

To put the above opinions and assertions in context, it is useful to review a bit of the history and evolution of several research streams that culminated in the paper by S&F. To this end we will mention some key events and highlights in that history and evolution, bring together some prior work likely viewed by many researchers as diverse and nonrelevant strands of literature, and place the contribution of S&F in a historical and contemporary context. Naturally, this review is from our perspective, and others may view this history and evolution differently.

We begin our history at the first Invitational Choice Symposium at Banff, Alberta, Canada in 1990, where Taka Morikawa discussed his Ph.D. work (Morikawa 1989) comparing real market (revealed) and experimental (stated) preferences/choices for transport mode choice in The Netherlands. Morikawa noted that RUT made a strong prediction about the relationship between preferences represented by the two types of data and showed how to compare and test whether preferences underlying both data sources differ. He showed that preferences were the same in his data set, which many at the symposium declared "a fluke." Of course now we know that it was not a fluke, as many dozens of papers in marketing, transport, and applied economics show that both data types often give similar preference estimates (e.g., see Louviere et al. 2000, Chapter 13; Ohler et al. 2000; Swait and Andrews 2003; Louviere 2006).

Morikawa's key insight from RUT was that if the underlying preferences are the same in different data sources, parameter estimates from RUT-based preference (choice) models should be proportional to one another. Specifically, RUT predicts that model estimates should be the same "up to scale"; that is, they should differ only by a constant of proportionality that equals the ratio of the error variances in the data sources being compared. To introduce some notation into our discussion, let $U^k = V^k + \varepsilon^k$, where U is total utility, V is its systematic component, and ε is the corresponding stochastic component for the kth data source. Let the scale μ^k of ε be inversely proportional to its variance σ^k. Then, if two data sources, k and k', both reflect the same underlying preferences, it should be the case that (see, e.g., Swait and Louviere 1993)

$$V^k = (\mu^{k'}/\mu^k)V^{k'}, \tag{1}$$

because only the ratio of the scales is identifiable. Part of Morikawa's insight was not new, as it was known that comparing different sources of revealed preference data required knowledge of scale differences. For example, Ben-Akiva and Lerman (1985) discuss and compare MNL and independent probit preference estimates; they show that scale differences are associated with slightly different identification requirements in each model. What was new was his insight that stated intentions data from surveys or choices in discrete-choice experiments (DCEs) should be related to choices in real markets in a specific way to be consistent with RUT. This insight is widely used in transportation and environmental economics to combine or fuse stated and revealed preference data sources. In transportation, it is often used as a way to enhance statistical efficiency and/or to fill in new products and extend attribute ranges in real market data.

At the time, we both were excited by Morikawa's result. We had several sources of real market and stated preference data, so after the symposium we began to see if his result held in our data sources. To say that it did would be an understatement! In fact, we made many comparisons that eventually became the basis for Chapter 13 in Louviere et al. (2000), where we laid out what we thought was a compelling case for a widespread empirical regularity in different types of preference elicitation data. However, the more we thought about this phenomenon, the more we came to see that it was much more profound than simply comparing data sources and/or filling in attribute gaps. That is, we came to see it as fundamental to all data source comparisons based on limited dependent variable statistical models, as we noted in Swait and Louviere (1993). Few academic and practical marketing researchers seem aware of the now-large literature on this topic, nor do they seem aware that Morikawa's insight also provides basic theory to understanding and quantifying relationships between stated intentions/choices and real market behavior.

As earlier noted, Morikawa's result suggested that data sources can differ in error variances. We felt that a logical next question was whether error variances are constant within the *same* data source. The more we looked, the more we found that they rarely were constant. From the mid-1990s we began reporting these results and their implications at conferences and discussed them with colleagues in various fields. We tried to translate and communicate the implications of our results for our colleagues in marketing and other fields to help them see that statistical analyses performed on limited dependent variables are not "clean," but instead are confounded with error variance differences. This was a much more difficult and frustrating task than we first imagined!

Met largely with indifference and/or disinterest, we turned our attention to discussing these issues with colleagues more directly concerned with discrete-choice models in fields like transport and applied economics. Again, we largely experienced deep skepticism or utter dismissal, such as the following typical comment: "Even if this is a problem, which it isn't, it can easily be corrected by simply allowing for more flexible error variance-covariance matrices and/or more flexible preference heterogeneity specifications." Kuhn (1962) was an astute observer of scientists, and his observations about paradigm changes have held true for many paradigm shifts that we have seen since the early 1970s. They once again held here: No one wanted to know about differences in error variances, the RUT paradigm was fine, and only small changes/tweaks needed to be made to existing models to deal with the problem even if error variances weren't constant. That is, we faced a classic "one can easily save the paradigm by adding this or that parameter or effect" type of reaction. Unfortunately, such statements typically blind paradigm-savers to other ways of viewing how the world works or potentially new and interesting explanations of empirical processes.

Over our careers we have found that persistence is a key to eventual success in getting people in well-established paradigms to pay attention to potentially paradigm-shifting issues. We slowly managed to get a few things published on the topic, and gradually others began to get it, most notably, researchers in environmental and resource economics (e.g., Adamowicz et al. 1994, 1997; Swait and Adamowicz 2001; DeShazo and Fermo 2002; Swait et al. 2004). We proposed a workshop on the general topic of stochastic utility for the Berkeley Invitational Choice Symposium (Louviere et al. 2002), and we tried for more traction in a workshop on endogeneity in choice models at the Colorado Symposium (Louviere et al. 2005), but our message was greatly diluted. We tried again at the recent Wharton Choice Symposium (Adamowicz et al. 2008), weaving it into a general theme of putting more behavioral theory into choice models. We got traction in the latter workshop, but only because most participants already understood and agreed with the issues.

Following the Wharton Symposium, Jordan Louviere, Michael Keane, their colleagues, and students at the University of Technology in Sydney, began studying the topic of differences in variances in earnest, as we had gone on record in the Wharton Symposium workshop report as saying that what was needed was a theory of scale (the inverse of the error variance). We made such a bold statement because of the many empirical results that not only showed nonconstant error variances, but also showed that they were systematically related to predictable sources. These sources included such things as how analysts design and implement experiments; how survey researchers design and implement surveys; differences in contexts, geography, and time; and differences in humans studied. We then developed new ways to separate scale and preference differences (Fiebig et al. 2009, Islam et al. 2007, Louviere et al. 2008). Empirical results associated with these references clearly show that error variances differ within and between people, and these differences play significant roles in driving differences in choices. Additional results are in Louviere and Eagle (2006) and Louviere and Meyer (2007).

Earlier, Louviere (2001) did a thought piece on the role of the error variance at the invitation of *Journal of Consumer Research* Editor David Mick. It was so persuasive that it had virtually no citations until about 12 months ago when it began to be cited (as of this writing it has 38 cites on Scholar Google). We began to think that this was a losing battle. Then came S&F, which gave us renewed hope that maybe views can change, and maybe others will begin to see the many potential research opportunities provided by viewing the world through a nonconstant variance lens.

S&F are the focus of the remainder of our discussion, and we want to say immediately that their paper is one of those rare, special papers that markedly impact one's academic life and career. They are braver than we were; we made a few attempts to convince our behavioral colleagues about the need to be aware of the influences of stochastic error variance on systematic and stochastic utility, and how inferences and conclusions could be seriously biased and incorrect unless one did so. Our lack of success eventually discouraged us, and we returned to our own research patch, frustrated by indifference and even outright hostility. We again can turn to Kuhn (1962) to anticipate that there will be a strong reaction to the paper by S&F because it has many implications for work in many different areas in marketing, psychology, economics, and other fields. That is, S&F have made a statement that more or less says that "the emperor has no clothes" (technically, their result applies only to the one research stream studied), but they made very careful observations from many different angles before going public. Those of us who have labored in the error variance wilderness for many years now have a sign that points the way to normal science.

We hope that the stream of work studied by S&F will be only the first of many subjected to careful and rigorous tests like those in their paper. Indeed, we would not be surprised if many other phenomena now come under close and careful scrutiny following their lead. S&F make this more likely by being brave enough to give intellectual permission to others who want to follow their lead; they do this in an exemplary way: the clarity of their presentation is superb, as is the thoroughness and rigor of their entire approach. We also hope that many researchers will now take up the challenge of extending S&F's ideas to other streams of work in the social and behavioral sciences. In fact, it may well be the case that their work leads to changes in the way researchers view many long-held beliefs, which in turn will lead to novel ideas about processes underlying decision-making and choice behavior.

The first author of this discussion piece was originally trained as a behaviorist and has since worked extensively with economists. Indeed, both authors have worked extensively with economists and are hopeful that S&F's work will open new doors to collaboration and cooperation between behavioral researchers and economists. S&F make this more likely by clearly laying out how RUT can be further instrumented to provide robust tests not confounded by untenable assumptions about stochastic utility. Their paper is a clear example of how careful, rigorously designed experiments and sophisticated statistical analysis work together to separate sources of influence on choice behavior. Such an extended view of RUT provides common ground and new ways for marketers and other fields to communicate and collaborate with economists, and with one another.

The simulation work in S&F is especially noteworthy as it addresses a likely objection by others that their choice model specification might lead to misattribution of effects to incorrect model components in the cases they studied. Their simulations lay that possibility to rest, clearly showing that one cannot simply make error variance results go away by including certain interactions in systematic utility functions, or by richer specification of error correlations. In turn, their results should be heard as a call to action by researchers in different fields: One must exercise great caution on explicit and implicit assumptions about stochastic utility components (random error terms) when analyzing limited dependent variables such as multiple-choice responses.

It is also noteworthy that their experimental and simulation work may imply a generalizable assertion that can be tested in future work. Their results clearly suggest that sophisticated methods for modeling preference heterogeneity and representing complex error correlations in RUT-based statistical models should receive less attention than modeling stochastic utility variances (or their inverse, scales). Given S&F's empirical evidence, it is likely to be beneficial if the process adopted in RUT modeling for making inferences about the jointly distributed random variables of interest (i.e., the utilities, U) becomes more disciplined. We conjecture that analysis should proceed in the following way:

1. Focus on the specification of the mean or systematic utility V (by which we specifically *do not* mean complex stochastic representations of taste heterogeneity, but *do* mean modeling of systematic taste heterogeneity based on exogenous covariates).
2. Shift attention to specifying the stochastic utility variance (i.e., the diagonal of the error covariance matrix, Σ_ε).
3. Focus on the off-diagonals of the covariance matrix (i.e., covariances between the stochastic utilities of alternatives).
4. Loop back to complete a specification with considerations like preference heterogeneity (essentially, introducing random taste variations into V).

Those acquainted with the history of choice modeling will recall that interest first focused on the systematic utility, V, then shifted to specification of the off-diagonals of the covariance matrix, Σ_ε (because of concerns with potential violation of IIA) then shifted back strongly to modeling preference heterogeneity in V. Variances in Σ_ε received little attention until the mid-1990s. Now, however, S&F's work shows that if underlying data generation processes are heteroscedastic, but misspecified to be homoscedastic, preference heterogeneity inferences will be incorrect and biased. That is, means and variances of stochastic preference distributions will be directly confounded with scale differences within and between choice options and people (e.g., Swait 2006). We suspect that simply including off-diagonal elements in the stochastic utility covariance matrix will partially capture true variance differences.

Both of the latter confounds are undesirable. This conjecture should serve to remind researchers in various choice modeling paradigms that when one works with latent utilities and highly nonlinear links to choice, there can be unpleasant consequences to misspecification not found with general linear models. If one treats heteroscedasticity as homoscedasticity in choice or other limited dependent variable models, it will lead to biases, not efficiency losses (e.g., see Yatchew and Griliches 1985 for the case of the probit model). The bottom line for those working in the RUT framework is that scale differences cannot be ignored or dealt with in a cavalier manner. The bottom line for researchers who think that they do not work in the RUT paradigm is that the issues noted in S&F, and in our discussion, apply to *all* limited dependent variable statistical models, which include all logit and probit models, tobit models, ordinal probit and logit models, and structural equation models.

Many will see the paper by S&F as a problem. Perhaps they cannot see the new opportunities it affords researchers in many fields. We think it is obvious that many new and exciting opportunities await those who do not fear to tread where others have not previously walked. To cite only one instance, Louviere and colleagues have been studying differences in preferences over choice occasions (choice sets) as a function of the order in which participants see them. They impose latin square designs on the choice sets to control for order and ensure that preferences and error variances can be separated at each order. In this experimental setup they find little evidence of preference evolution, but instead find evidence of error variance evolution (e.g., see Louviere and Eagle 2006, Louviere and Meyer 2007). Of course, if one does not control for error variance differences across orders and choice sets, one is likely to conclude that preferences change over choice sets. Moreover, despite a widespread practice of randomizing choice set order appearance, this may not be sufficient to control for, nor allow one to capture, evolution of variances or preferences over choice sets as it introduces correlations and confounds that researchers typically do not look for and cannot see in their data.

How rich and complex the world now is! How much more interesting and complex human decisions and choices now are! We congratulate and thank Salisbury and Feinberg on a landmark paper.

References

Adamowicz, W., J. Louviere, M. Williams. 1994. Combining revealed and stated preference methods for valuing environmental amenities. *J. Environ. Econom. Management* **26** 271–292.

Adamowicz, W., J. Swait, P. Boxall, J. Louviere, M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. *J. Environ. Econom. Management* **32** 65–84.

Adamowicz, W., D. Bunch, T. A. Cameron, B. G. C. Dellaert, M. Hanneman, M. Keane, J. Louviere, R. Meyer, T. Steenburgh, J. Swait. 2008. Behavioral frontiers in choice modeling. *Marketing Lett.* **19**(3/4) 215–228.

Ben-Akiva, M., S. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Predict Travel Demand*. MIT Press, Cambridge, MA.

DeShazo, J. R., G. Fermo. 2002. Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *J. Environ. Econom. Management* **44**(1) 123–143.

Fiebig, D. G., M. P. Keane, J. J. Louviere, N. Wasi. 2009. The generalized multinomial logit model: Accounting for scale and coefficient heterogeneity. *Marketing Sci.*, ePub ahead of print July 23, http://mktsci.journal.informs.org/cgi/content/abstract/mksc.1090.0508v1.

Islam, T., J. J. Louviere, P. F. Burke. 2007. Modeling the effects of including/excluding attributes in choice experiments on systematic and random components. *Internat. J. Res. Marketing* **24**(4) 289–300.

Kuhn, T. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago.

Louviere, J. J. 2001. What if consumer experiments impact variances as well as means? Response variability as a behavioral phenomenon. *J. Consumer Res.* **28**(3) 506–511.

Louviere, J. J. 2006. What you don't know might hurt you: Some unresolved issues in the design and analysis of discrete choice experiments. *Environ. Resource Econom.* **34**(1) 173–188.

Louviere, J. J., T. Eagle. 2006. Confound it! That pesky little scale constant messes up our convenient assumptions! *Proc. 2006 Sawtooth Software Conf., Delray Beach, FL*, Sawtooth Software Inc., Sequim, WA, 211–228.

Louviere, J. J., R. J. Meyer. 2007. Formal choice models of informal choices: What choice modeling research can (and can't) learn from behavioral theory. N. K. Malhotra, ed. *Review of Marketing Research*, Vol. 4. M.E. Sharpe, Armonk, NY, 3–32.

Louviere, J. J., D. A. Hensher, J. Swait. 2000. *Stated Choice Methods*. Cambridge University Press, Cambridge, UK.

Louviere, J. J., T. Islam, N. Wasi, D. Street, L. Burgess. 2008. Designing discrete choice experiments: Do optimal designs come at a price? *J. Consumer Res.* **35**(2) 360–375.

Louviere, J. J., D. Street, L. Burgess, N. Wasi, I. Islam, A. A. J. Marley. 2008. Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *J. Choice Model.* **1**(1) 128–163.

Louviere, J., D. Street, R. Carson, A. Ainslie, J. R. DeShazo, T. Cameron, D. Hensher, R. Kohn, T. Marley. 2002. Dissecting the random component of utility. *Marketing Lett.* **13**(3) 177–193.

Louviere, J., K. Train, M. Ben-Akiva, C. Bhat, D. Brownstone, T. A. Cameron, R. Carson et al. 2005. Recent progress on endogeneity in choice modeling. *Marketing Lett.* **16**(3/4) 255–265.

Morikawa, T. 1989. Incorporating stated preference data in travel demand analysis. Ph.D. dissertation, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge.

Ohler, T., A. Le, J. J. Louviere, J. Swait. 2000. Attribute range effects in binary response tasks. *Marketing Lett.* **11**(3) 249–260.

Salisbury, L., F. Feinberg. 2010. Alleviating the constant stochastic variance assumption in decision research: Theory, measurement, and experimental test. *Marketing Sci.* **29**(1) 1–17.

Street, D. J., L. Burgess. 2007. *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. John Wiley & Sons, New York.

Swait, J. 2006. Advanced choice models. B. J. Kanninen, ed. *Valuing Environmental Amenities Using Stated Choice Methods: A Common Sense Approach to Theory and Practice.* Springer, Dordrecht, The Netherlands, 229–293.

Swait, J., W. Adamowicz. 2001. Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organ. Behav. Human Decision Processes* **86**(2) 141–167.

Swait, J., R. Andrews. 2003. Enriching scanner panel models with choice experiments. *Marketing Sci.* **22**(4) 442–460.

Swait, J. D., J. J. Louviere. 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *J. Marketing Res.* **30**(3) 305–314.

Swait, J., W. Adamowicz, M. van Bueren. 2004. Choice and temporal welfare impacts: Incorporating history into discrete choice models. *J. Environ. Management Econom.* **47** 94–116.

Yatchew, A., Z. Griliches. 1985. Specification error in probit models. *Rev. Econom. Statist.* **67**(1) 134–139.