

# RM-SIG NEWS

No 3

June 2017

## Meet this year's Coordinators of the AIB Research Method Clinics

### Aggie Chidlow

I am interested in discrete choice models as well as data collection techniques and issues relating to survey data. My goal for the RM SIG is to enhance the rigour and diversity of research methodologies applied in IB research.



### Stewart Miller

I tend to use Heckman models and event study methodology. My goal for the RM-SIG is to provide a learning environment that helps scholars to embrace the methodological challenges and opportunities that arise with IB research.



### In this Newsletter:

- Research Methodology Offerings at AIB Annual Meeting 2017
- Introducing Necessary Condition Analysis, PLS-SEM
- Methodological guidance from key journals
- Methodological Myth: Interviews are a better data source than documents
- RM-SIG Soapbox: Herman Aguinis
- SIG Members' Activities, workshops, Call for Papers, Best Paper award finalists

# Ouch, That Hurts! On Torturing the Data Until They Confess



In a forthcoming article in *Journal of International Business Studies*, Herman Aguinis, Wayne F. Cascio, and Ravi S. Ramani argue that international business is not immune to science's reproducibility and replicability crisis. They provide evidence that this crisis is not entirely surprising given methodological practices that enhance systematic capitalization on chance. This occurs when researchers search for a maximally predictive statistical model based on a particular dataset and engage in several trial-and-error steps that are rarely disclosed in published articles. In other words, many researchers "torture the data until they confess" that effects are statistically significant, large, and supportive of favored hypotheses and models. Each of these outcomes—which together are more likely to result in the desired result of a successful publication—can be reached more easily by systematically capitalizing on chance.

Their JIBS article describes five of several "data torturing" practices that enhance systematic capitalization on chance: (1) selection of variables to include in a model, (2) use of control variables, (3) handling of outliers, (4) reporting of  $p$ -values, and (5) hypothesizing after results are known (HARKing). They also illustrate the pervasiveness of each of these practices using articles published in JIBS and then offer best-practice recommendations on how to minimize capitalization on chance in future international business research. These recommendations serve as resources for researchers, including doctoral students and their training, as well as for journal editors and reviewers evaluating manuscript submissions.

A key issue regarding "data torturing" practices is lack of transparency. Aguinis, Cascio, and Ramani clarify that "epistemological approaches other than the pervasive positivistic model, which has become dominant in management and related fields since before World War II, are indeed useful and even necessary. For example, inductive and abductive approaches can lead to important theory advancements and discoveries." In short, they are not advocating a rigid adherence to a positivistic approach but, rather, methodological plurality that is *fully transparent* so that results can be reproduced and replicated.

## RM-SIG Soapbox

This issue's Soapbox is contributed by Herman Aguinis, Avram Tucker Distinguished Scholar and Professor of Management at the George Washington University School of Business

Please contact us if you have an idea for the next Soapbox

As an example of one of these five “data torturing” practices, consider the issue of outliers, which are data points that deviate markedly from others. Outliers are a challenge because they can substantially affect results obtained when testing hypotheses. Because of their outsized influence, the management of outliers presents an opportunity for researchers to systematically capitalize on chance when analyzing data, often in the direction of supporting their hypothesis. However, many researchers routinely fail to disclose whether they tested for outliers within their datasets, whether any outliers were identified, the type of outliers found, and the rationale behind choosing to include or exclude outliers from analyses.

As noted in their JIBS article, “Recently published articles in JIBS suggest the presence of systematic capitalization on chance regarding the management of outliers. For example, reported practices include winsorizing firm-level variables at the 5% level to account for outliers, trimming the sample by excluding observations at the top and bottom one percentile of variables, and removing an outlier based on other techniques and criteria. In none of these cases did the authors define the type of outlier they were addressing. Specifically, error outliers (i.e., data points that lie at a distance from other data points), interesting outliers (i.e., non-error data points that lie at a distance from other data points and may contain valuable or unexpected knowledge), or influential outliers (i.e., non-error data points that lie at a distance from other data points, are not error or interesting outliers, and also affect substantive conclusions). In addition, in none of these published articles did the authors take appropriate steps such as correcting the data for error outliers and reporting the results with and without outliers. Therefore, by not providing clear and detailed reporting of the manner in which they addressed the issue of outliers, it is virtually impossible to reproduce and replicate substantive conclusions.”

Here’s their suggestion pertaining specifically to handling outliers: “Researchers should provide evidence showing that they tested for outliers in their datasets. They should specify the rules used to identify and classify outliers as error, interesting, or influential, and disclose whether influential outliers affect model fit or prediction. Finally, they should test their models using robust approaches (e.g., absolute deviation) and report results with and without outliers.” Similarly, their article offers best-practice recommendations regarding how to minimize the negative effects of systematic capitalization on chance that results from other “data torturing” practices such as selection of variables to include in a model, use of control variables, reporting of *p*-values, and hypothesizing after results are known.

Their article is available on “Online First Articles” at <https://link.springer.com/journal/41267/onlineFirst>

**Reference:**

**Aguinis, H., Cascio, W. F., & Ramani, R. S. In press. Science’s reproducibility and replicability crisis: International business is not immune. *Journal of International Business Studies*. doi: 10.1057/s41267-017-0081-0**