Journal of Operations Management 36 (2015) 178–186

Contents lists available at ScienceDirect

Journal of Operations Management

journal homepage: www.elsevier.com/locate/jom

Technical note

Supply chain management research: Key elements of study design and statistical testing

Catherine A. Helmuth[a], Christopher W. Craighead[b], Brian L. Connelly[a,*], Donovan Y. Collier[a], Joe B. Hanna[a]

[a] Raymond J. Harbert College of Business, Auburn University, 415 West Magnolia Avenue, Auburn, AL 36849, United States
[b] Smeal College of Business, The Penn State University, Business Building, Suite 483, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 19 July 2013
Received in revised form 3 November 2014
Accepted 14 December 2014
Available online 22 December 2014
Accepted by Thomas Younghoon Choi

Keywords:
Supply chain
Effect size
Statistical power
Reliability
Empirical research

ABSTRACT

Over the past three decades, supply chain management (SCM) has evolved from its origins as a nascent field of study to encompass construct definition, identification of the field's central issues, and establishment of its conceptual boundaries. At this point, a sufficient body of empirical SCM research has been put forward to allow for quantitative assessment of the field. Therefore, we examine three key elements of study design to assess what has happened, what is currently happening, and where we should be heading as a field. To do so, following a pattern of reviews in similar disciplines, we begin with an examination of effect sizes of the relationships under investigation. Results show that effect sizes in SCM research have marginally increased over time and that sub-domains within SCM that receive the most scholarly attention also have higher effect sizes. We also conduct a post hoc analysis of statistical power and empirically examine a range of factors and study contexts that could influence power. Findings suggest that average statistical power in SCM research exceeds the statistical power of most related disciplines and is particularly high in several unique contexts. Lastly, we find that measurement reliability and the use of control variables have increased over time, possibly suggesting the field has matured, instilling a degree of confidence in its research. Overall, our results show that SCM research is becoming more empirically rigorous, but we also uncover key areas that warrant improvement. We describe implications of our review for the design of future SCM empirical studies.

1. Introduction

A "supply chain" refers to the activities, functions and entities that are connected via product and information flow from source to consumer (Craighead et al., 2007). Research on supply chain management (SCM) has evolved from its origins as a nascent field of study to encompass construct definition, identification of the field's central issues, and establishment of its conceptual boundaries. We believe a sufficient body of empirical work has emerged (Craighead and Meredith, 2008), particularly over the past ten years, to warrant more quantitative assessment of study designs. Reviews of specific topic areas and theories are useful for summarizing content (Short, 2009), but it is also important to examine and assess a disciplines' methodological rigor. Such assessments have been instrumental in the social sciences and important to assessing study design in fields such as entrepreneurship (Connelly et al., 2010), management (Cashen and Geiger, 2004), and industrial and organizational psychology (Mone et al., 1996). Therefore, in this technical note, we examine three core aspects of SCM study design and consider what has happened, what is currently happening, and where we should be heading as a field of study.

To evaluate study design of empirical SCM research, the most basic question is: what are researchers trying to measure? This speaks to the issue of effect size, which describes the strength of association between two variables, a predictor and criterion (Cohen et al., 2003) (other reviews answer this question from a theoretical standpoint, examining the range of research questions and theories that SCM researchers employ). Effect size is important because empirical SCM research is largely built around statistical inference testing, so we should begin by examining what it is researchers are testing. That is, whether it is something that is actually occurring or not. Effect size captures the extent to which the theoretical phenomenon that the researcher has chosen to examine actually exists in the population (Cohen, 1988).

Assuming the presence of an effect for the relationships under investigation, we also consider whether researchers are examining a sufficient pool in order to draw conclusions about that

* Corresponding author. Tel.: +1 334 703 7070.
E-mail address: bconnelly@auburn.edu (B.L. Connelly).

http://dx.doi.org/10.1016/j.jom.2014.12.001
0272-6963/

relationship. This speaks to the issue of statistical power, which is foundational to a discipline's research methodology (Faul et al., 2007). As a result, reviews of study design often survey, analyze, and report on statistical power to determine the extent to which researchers are attending to, or neglecting, the subject of statistical power (Connelly et al., 2010; Mone et al., 1996; Verma and Goodale, 1995). Pertinent examples include reviews by Brock (2003) for international business, Ferguson and Ketchen (1999) for strategic management, and Borkowski et al. (2001) for behavioral accounting. To date, however, scholars have not yet assessed this issue in supply chain management (SCM) research.

Our examination of SCM study designs, therefore, considers what researchers are measuring (effect size), on whom they measure it (statistical power), and also examines how well they are measuring those relationships. To do so, we first consider the extent to which researchers are measuring what they think they are measuring. While there are several ways of considering this issue, not all lend themselves well to post hoc analysis (e.g., ensuring proper randomization in sample selection). However, we can examine the reliability of researcher's measures, which helps determine the extent to which their measures are psychometrically sound. Moreover, even if a relationship is occurring in the population (i.e., effect size), the researcher has a sufficiently large sample (i.e., statistical power), and sound measures (i.e., reliability), the study may not be capturing the intended effect if the results are subject to an extensive range of alternative explanations. Therefore, as a final component, we also consider the use of control variables in SCM study design, because these variables are designed to account for explanations of the relationship under investigation beyond the effects of the researcher's chosen predictors.

In sum, we examine three core aspects of study design: (1) effect size, (2) statistical power, and (3) reliability and controls. There are, of course, more nuanced aspects of study design that SCM scholars could also consider, such as the empirical challenges associated with common method variance (Craighead et al., 2011) or action-research perspectives (Näslund et al., 2010). However, taken together, we believe these three basic issues form a foundation upon which scholars can build, that they constitute a benchmark for evaluating the state of SCM research, and that they provide a useful tool for future study designs.

Toward this end, we conduct an in-depth review of SCM research over the years 2002–2012, examining 4235 statistical tests that appear in 217 unique studies. We begin by describing average reported effect size in the different types of studies under investigation, commonly referred to as the meta-analytic effect size (Henson, 2006; Thompson, 2005). We also conduct a post hoc power analysis to determine if there is sufficient statistical power to detect large, medium, and small effect sizes in SCM research and examine power levels for a range of different study types. Lastly, we examine measurement reliability (i.e., Cronbach's alpha) and the use of control variables over time.

This technical note holds the potential of contributing to SCM research in several important ways. First, it adds to the discipline's methodological knowledge by assessing the likelihood that researchers could have correctly detected a treatment effect if one was actually present. When statistical power levels are inadequate, researchers may determine there is no effect when actually their research design prohibited them from detecting a significant relationship (Nickerson, 2000). These types of erroneous conclusions restrict the field's theoretical development and advancement. Second, our overall evaluation of empirical SCM research can be compared to those of similar disciplines in order to help establish the credibility of the academic knowledge emerging from SCM scholars. As a field of study establishes itself, it is important to ensure that scholars working in that area maintain the requisite level of methodological rigor, particularly as compared to other areas. Third, and more broadly, our study examines key components that could serve as indicators of methodological maturity. Given the relative newness of the SCM field, this kind of introspection could be important to understanding where we are at, and guiding where we should go, as a community of scholars. Finally, our findings constitute the basis of forming standards for meta-analytic thinking that scholars may employ when designing their studies (Henson, 2006; Thompson, 2005). Our review reveals that SCM research has come a long way but, despite the progress the field has exhibited, we also uncover some caveats that scholars should consider in future research endeavors.

2. Method and sample

We reviewed empirical articles from ten top-tier journals: *Journal of Operations Management*, *Production and Operations Management*, *Decision Sciences*, *Management Science*, *Manufacturing and Service Operations Management*, *Journal of Business Logistics*, *International Journal of Logistics Management*, *International Journal of Physical Distribution and Logistics Management*, *Journal of Supply Chain Management*, *Journal of Purchasing and Supply Management*. These journals have been used in SCM research (Näslund et al., 2010; Schoenherr, 2009) and are all highly ranked (Zsidisin et al., 2007; Chapman and Ellinger, 2009).

We evaluated all empirical articles in the above journals between the years 2002–2012. To identify SCM research studies during this time frame, we collected and individually reviewed all articles that contain the terms *supply chain* or *supply chains* in the title or abstract (scholars may define supply chain differently, but searching on these terms allowed us to cast a wide net). While formal meta-analyses sometimes include developmental work (e.g., doctoral dissertations and conference presentations) to avoid the "file drawer" problem, our stated goal is to review the current state of published research. Including studies under development could bias our results, so we limited our attention to the top peer-reviewed journals (cf., Aguinis et al., 2005; Cashen and Geiger, 2004).

Following prior reviews from other disciplines, we excluded studies that used mathematical modeling, optimization techniques, simulations, sensitivity analyses, generated data, or research designs that did not align with statistical tests for which power analyses have been established (Cohen, 1992; Connelly et al., 2010; Faul et al., 2007). We omitted articles that failed to report the specific data analytic technique and test statistic used as well as articles that investigated journal rank or citation counts. Non-parametric tests were excluded with the exception of the Wilcoxon Signed Rank Test and the Wilcoxon Mann Whitney Test. Finally, as a result of the multiple approaches to, challenges with, and conflicting recommendations for conducting statistical power analysis in structural equation modeling (MacCallum et al., 2006; Saris et al., 2009; Li and Bentler, 2011), we omitted studies that used SEM from our sample. We individually coded each statistical test for the remaining 217 empirical SCM research studies under investigation (see Table 1 for a breakdown by journal), which included 4235 statistical tests (excluding post hoc, manipulation, or reliability tests).

3. Review of SCM empirical research: assessment of key elements

3.1. Effect size

Perhaps the most foundational element of study design is the effect size under investigation, which describes the strength of the association between the variables at hand (Cohen et al., 2003). Effect size captures the extent to which the theoretical