A Review of Short-term Event Studies in Operations and Supply Chain Management

Li Ding
School of Management
Zhejiang University
Hangzhou, P. R. China, 310058

The Department of Logistics and Maritime Studies
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

Hugo K.S. Lam
Management School
University of Liverpool
Liverpool, UK

T.C.E. Cheng
The Department of Logistics and Maritime Studies
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

Honggeng Zhou *
School of Management
Zhejiang University
Hangzhou, P. R. China, 310058

* Corresponding author

Abstract

The short-term event study method, grounded in the Efficient Market Hypothesis, is one of the most widely used tools for quantifying the impact of a specific event on a firm's shareholder value. As the short-term event study method has been increasingly employed by researchers to investigate various operations and supply chain management (OSCM) events, it is timely to conduct a systematic review of the method to examine how it has been implemented in the OSCM literature and what could be improved to deploy it for future OSCM research. Analyzing 29 short-term event studies published in renowned OSCM journals between 1995 and 2017, we find that OSCM researchers generally follow the standard procedures in conducting event studies, but pay less attention to some methodological issues ranging from addressing the confounding events to expanding the event windows. Based on our analysis, we provide several recommendations for future event studies in OSCM, such as the opportunity for studying external events in the non-U.S. context, the caution of expanding the event windows, and the need to deal with the self-selection bias.

Keywords: short-term event study; shareholder value; abnormal return; operations management; supply chain management; literature review

1. Introduction

Over the past few decades, there is growing recognition of the strategic importance of operations and supply chain management (OSCM) in creating shareholder value. OSCM plays a vital role in generating shareholder value through the mechanisms of revenue growth, operating cost reduction, and efficient use of fixed and working capital (Martin and Lynette, 1999). Following this theoretical logic, researchers have conducted various empirical studies to analyze the connection between OSCM and shareholder value, among which the event study method represents one of the most popular methodologies adopted in the literature. Grounded in the Efficient Market Hypothesis (Malkiel and Fama, 1970), the short-term event study method relies on the premise that the value of market information will be reflected almost completely in the equity prices in financial markets. By detecting the abnormal equity price changes in response to new market information available in the financial market, the short-term event study method enables researchers to quantify the impact of a specific event on a firm's shareholder value (MacKinlay, 1997).

With its growing popularity in the OSCM literature, the short-term event study method has been employed by researchers to investigate various OSCM topics such as supply chain disruptions (Hendricks and Singhal, 1997; Zhao et al., 2013), environmental management (Jacobs, 2014; Klassen and McLaughlin, 1996), and quality management (Lin and Su, 2013; McGuire and Dilts, 2008). In addition, short-term event studies in OSCM are evolving as a result of advances in asset pricing models and statistical analysis. The method has been modified to address potential statistical issues specific to different research settings (Fama and French, 2015; Kothari and Warner, 2007). In view of its increased popularity and recent methodological improvements, it is timely to conduct a systematic review of the method to examine how it has been implemented in the OSCM literature and what could be improved to deploy it for future OSCM research.

Reviewing 29 short-term event studies published in renowned OSCM journals between 1995 and 2017, we have the following observations: (1) The majority of the short-term event studies in OSCM focus on internal corporate events in the U.S. context. (2) While most studies set standard event windows including at most three days around the event, theoretical justifications are not commonly provided for short-term event studies with longer event windows. (3) Researchers often rely on multiple data sources to identify the events under study, but pay less attention to the issue of confounding events. (4) The market model is the most popular estimation model in the OSCM literature, but some researchers also employ multiple estimation models to increase the robustness of the analysis. (5) Researchers are wary of possible violations of the assumptions for the significance test, so adopting various modifications of the traditional *t*-test according to different research contexts. (6) Researchers often conduct subsequent cross-sectional regression and ANOVA to probe into the operational determinants of variations in abnormal returns.

Based on our analysis, we provide several recommendations for future event studies in OSCM. First, we urge OSCM researchers to take advantage of events external to the firms concerned and occurring outside the U.S. context, advancing our understanding of the financial impacts of these under-studied events. Second, researchers should be careful about expanding the event windows, and provide theoretical explanations to justify the window lengths. Third, removing confounding effect is a critical step in conducting short-term event studies. Fourth, the possible self-selection bias should not be ignored, especially when the events under study are initiated by firms voluntarily. Fifth, employing alternative models to estimate the expected returns could enhance the robustness of the analysis. Sixth, modifications of the traditional *t*-test might become necessary in some research settings such as external events and industry-specific studies. Finally, independence is a vital assumption in testing the significance of cumulative abnormal returns. It thus is important to address the issues arising from time and industry clustering.

Our research is important in several ways. First, it serves as a practical guide for OSCM researchers interested in employing the short-term event study method in their research. We document the detailed steps of conducting a short-term event study and discuss some common issues encountered in each step, thus enabling OSCM researchers to have a better understanding of how a short-term event study should be conducted. Moreover, to the best of our knowledge, this is the first comprehensive review of event studies in the OSCM literature. Given the increased prevalence of event studies in OSCM, it is imperative to provide an overview of the current state of knowledge and best practices adopted in the OSCM literature. Finally, our research identifies several important research design issues that are often ignored by researchers of past short-term event studies in OSCM, as well as some emerging opportunities specific to the OSCM context, so helping advance the adoption of the event study method for OSCM research.

2. Literature review

The first event study reported in the literature was perhaps conducted by James Dolley in 1933. Based on a sample of 95 stock splits from 1921 to 1931, Dolley (1933) investigated the nominal stock price changes at the time of the stock splits. Modern event studies were initiated in the two seminal works of Ball and Brown (1968) and Fama et al. (1969). Modern event studies are developed into different categories in terms of the event window length and performance measurement. Long-term event studies detect abnormal stock returns over a period normally ranging from one to eight years with calendar-time portfolio abnormal return (CTAR) or buy-and-hold abnormal return (BHAR) (Barber and Lyon, 1997; Lyon et al., 1999), while short-term event studies examine abnormal stock returns over a maximum window length of 40 days

(Brown and Warner, 1985; MacKinlay, 1997). A broader definition of event study goes beyond the scope of stock market reaction as it also measures other firm-level outcomes such as operating performance (Barber and Lyon, 1996). In parallel with advances in asset pricing models and statistical analysis, the event study method is still evolving to account for possible deviations from the fundamental assumptions. However, the gist of modern event studies remains the same, which is measuring the significance of sample securities' mean and cumulative abnormal returns around an event period (Kothari and Warner, 2007).

Originally applied in accounting and finance, the event study method has expanded its application to virtually all the business disciplines including management, information systems, marketing, operations and supply chain management (MacKinlay, 1997; McWilliams and Siegel, 1997). For example, in the marketing literature, researchers adopt the event study method to examine the financial impact of such marketing events as new product release, CMO appointment, brand acquisition and disposal, and Internet channel addition (Sorescu et al., 2017), while events attracting information systems researchers' attention include IT outsourcing, IT investment, IT excellence award, software vulnerability, and security breaches (Konchitchki and O'Leary, 2011).

(Insert Table 1 here)

Table 1 summaries previous literature reviews of event studies in different business disciplines. It indicates that the literature reviews in accounting and finance emphasize the econometric and statistical fundamentals and provide guidelines for applications in other fields. For instance, MacKinlay (1997) and Binder (1998) reviewed the use of event studies in finance, outlined the standard procedures for conducting event studies, and discussed the power of analysis and the subsequent regression analysis. Corrado (2011) reviewed variations in the basic short-term event study method to adjust for non-normality, event-induced volatility, and cross-sectional weighting. Kothari and Warner (2007) conducted a comprehensive survey of over 500 studies published in five of the top finance and accounting journals from 1974 to 2005. They found that the properties of the event studies reviewed were different depending on the time period and sample firm characteristics. They also indicated that, compared with short-term event studies, long-term event studies suffer from several important limitations.

As the event study method evolves over time, its statistical properties become well-defined and its applications are widely acknowledged. Literature reviews in other business disciplines place a greater emphasis on the research design issues and economic interpretations of the study results. McWilliams and Siegel (1997) conducted a survey of 29 event studies in three of the top management journals from 1986 to 1995. They discussed several concerns about the validity of the assumptions and research design issues. By replicating three studies in management with alternative research designs, they called for adequate attention towards the aforementioned concerns. They also indicated that the abnormal returns only reflect the effect on the shareholder wealth, rather than the welfare of all the stakeholders. Konchitchki and O'Leary (2011) examined the use of the event study method in over 50 information systems studies. They focused on the research design issues without investigating the actual results and conclusions in specific studies. Sorescu et al. (2017) identified over 40 event studies published in the marketing journals included in the list of *Financial Times*' 50 top business journals. In addition to research designs, their review examines interpretations of event studies as well. They provided economic inferences from the event studies by summarizing the main findings and common determinants of abnormal returns in the marketing literature.

Consistent with other fields, OSCM has witnessed a growth in employing event study as a viable research method. However, to the best of our knowledge, there is no literature review of event studies in OSCM. One related study performed by Min and Wei (2013) reviews the literature linking supply chain management (SCM) and firm-level financial performance. Based on 49 research articles published between 1990 and 2011, they summarized the empirical studies conducted using various research methods, including structural equation modelling, event study, correlation analysis, and multivariate regression. Aiming to provide a better understanding of how SCM affects financial performance, their review is topic-centric and is comprehensive in terms of research methodology without specializing in event studies. Therefore, in order to summarize the current knowledge of short-term event studies in OSCM and to provide guidelines for OSCM researchers interested in applying the methodology, we conduct this literature review and make recommendations on its proper use.

3. The scope of this research

Event studies in OSCM can be classified according to short-term or long-term event windows, along with various performance measurements, such as stock returns (Brandon-Jones et al., 2017), accounting-based operating performance (Lo et al., 2009; Tang et al., 2016), plant productivity (Gopal et al., 2013), safety violations (Lo et al., 2014), and flight delays (Nicolae et al., 2016). Our study focuses on the event studies measuring the short-term stock market reactions for the following reasons. First, among the different types of event studies, the short-term approach is the earliest, as well as the most widely adopted method in OSCM (Hendricks and Singhal, 1996; Hendricks et al., 1995; Klassen and McLaughlin, 1996), providing enough representative samples for us to analyze how the method is implemented in the literature. Second, it is difficult to incorporate both short-term and long-term event studies in a single review paper due to their fundamental differences in theoretical assumptions and methodological execution. Specifically, short-term event studies are based on the Efficient Market Hypothesis (Malkiel and Fama, 1970), assuming that any new information available in the stock market will be reflected almost immediately in security price changes (MacKinlay, 1997). In contrast, long-term event studies are proposed based on the belief that stock prices could partially anticipate and slowly adjust to new available information. In terms of execution, elimination of confounding announcements is a vital step in short-term event studies, whereas this step is unnecessary and impractical in long-term event studies (Sorescu et al., 2017). In addition, short-term event studies are less sensitive to the estimation model of normal returns and assumptions of independence in most cases (Kothari and Warner, 2007). On the contrary, the precision of estimation is important in long-term event studies. Even a small error in risk adjustment of estimation models may ultimately lead to huge differences in cumulative abnormal returns, which are aggregated over a long time period (Kothari and Warner, 2007). Therefore, in consistency with the literature reviews of event studies in other fields (Corrado, 2011; Konchitchki and O'Leary, 2011; MacKinlay, 1997), we focus our review on short-term event studies in OSCM to provide clearer and more specific analysis and discussion.

4. Data

To identify short-term event study papers in OSCM for this review, we rely on a list of 13 "leading" OSCM journals included in the Korea University Business School (KUBS) Worldwide Business Ranking. The 13 journals are *Computers and Operations Research*, *Decision Sciences*, *European Journal of Operational Research*, *IIE Transactions*, *International Journal of Operations and Production Management*, *International Journal of Production Economics*, *Journal of Operations Management*, *Journal of Supply Chain Management*, *Journal of the Operational Research Society*, *Management Science*, *Manufacturing and Service Operations Management*, *Operations Research*, and *Production and Operations Management*.

We conducted the data collection process in five steps. First, we searched the single keyword "event study" in the aforementioned journals to generate a list of papers fitting our research objective. This single keyword approach could ensure a more comprehensive coverage of event studies about different OSCM topics, which is different from past review studies that are concerned with a specific OSCM topic such as green supply chain management (Srivastava, 2007) and rely on a combination of various keywords. Second, we examined all the papers generated from the preliminary search process and only included those actually adopting the event study method. In particular, we read the methodology section of each paper and excluded those mentioning the event study method but deploying other methods such as content analysis (e.g., Montabon et al., 2007) and regression analysis (e.g., Bayus et al., 2003; Ramdas et al., 2013). Third, as our review focused on short-term event studies based on abnormal stock returns, we excluded other types of event studies such as long-term event studies based on abnormal stock returns (e.g., Hendricks and Singhal, 2001; 2005) or abnormal operating performance (e.g., Corbett et al., 2005; Lo et al., 2012). Fourth, we further filtered the search results to ensure that the event study method is employed to investigate OSCM topics directly. Specifically, after reading the hypotheses and results sections of all the searched papers, we excluded the event study by Fosfuri and Giarratana (2009) that investigated stock market reactions to new product announcements and filed trademarks, which are more related to marketing rather than OSCM. Finally, we cross-checked the references cited in the papers to ensure no qualified articles were missed out from our analysis.

(Insert Table 2 here)

Table 2 lists the final 29 short-term event studies included in this review. The papers were published between 1995 and 2017 in *Journal of Operations Management* (28%), *International Journal of Production Economics* (24%), *Management Science* (21%), *Production and Operations Management* (14%), *International Journal of Operations and Production Management* (7%), *Decision Sciences* (3%), and *European Journal of Operational Research* (3%). In addition, from the publication years, we find that short-term event studies in OSCM are emerging and developing. There were only six papers (20%) published in the first ten years from 1995 to 2004, but 18 papers (62%) were published in the recent eight years from 2010 to 2017.

5. Current practices of short-term event studies in OSCM

Figure 1 summarizes the basic steps for conducting an short-term event study (MacKinlay, 1997), which include: (1) identify an event of interest; (2) define the event window and justify the choice of the window length; (3) collect the sample and eliminate confounding events; (4) predict normal returns with an estimation model; (5) calculate the abnormal returns, aggregate them over the event windows and test their significance; and (6) explain the cross-sectional variations in the abnormal returns. We provide a detailed explanation of each step below and review the current practices of conducting short-term event studies in OSCM.

(Insert Figure 1 here)

5.1 Identify an event of interest

Firms and other third parties often make announcements about significant activities occurring in all the aspects of the firms' internal operations and supply chain management, offering rich opportunities for researchers to identify events of interest for their research. As shown in Table 3, the topics investigated by short-term event studies in OSCM include supply chain disruptions (31%), environmental management (24%), quality management (14%), R&D projects (10%), sourcing strategies (7%), capacity expansion (4%), information technology management (4%), supply chain integration (3%), and purchasing and sales contract (3%).

(Insert Table 3 here)

Although the topics of event studies in OSCM vary, most of them are focused on internal corporate events that are within specific firms or their supply chains, with only one of the 29 papers we reviewed examining an event external to the firms concerned. Specifically, only the recent event study conducted by Jacobs and Singhal (2017) investigates the impact of an external event in terms of the Rana Plaza disaster in Bangladeshi on the shareholder value of global apparel retailers.

The majority of the OSCM literature studies events in the U.S. context, with only five of the 29 studies (17%) being in the non-U.S. context. Specifically, of these five non-U.S. based studies, there is one study about the impact of quality certification on the Spanish stock market (Nicolau and Sellers, 2002), and the other four studies are in the Chinese context. They investigate the reactions of the Chinese stock market to quality management (Lin and Su, 2013), product recall (Zhao et al., 2013), purchase and sales contract (Yang et al., 2014), and environmental initiatives (Lam et al., 2016).

An important consideration when identifying an event of interest is whether an unambiguous definition of the event could be provided. In some cases, defining the event itself or its proxy variable is a straightforward task. For example, product recalls in the U.S. are managed by five specific federal agencies and the announcement of a product recall conveys detailed information about the product being recalled, and the firm recalling it, making the identification of product recalls less subjective (Ni et al., 2014). However, some events have broader meanings in nature, and researchers need to define clear boundaries of the events with a set of keywords. For example, Hendricks and Singhal (2003) relied on a combination of various keywords such as delay, shortfall, shortage, manufacturing, production, shipment, delivery, parts, and components, to identify the announcements of supply chain glitches.

Another important consideration is whether the event is unexpected by the investors before being announced and whether it is visible to investors when being announced. This is because, based on the Efficient Market Hypothesis (Malkiel and Fama, 1970), the underlying assumption of all the short-term event studies, any new information available in the stock market will be reflected immediately in security price changes (MacKinlay, 1997). For example, if there is information leakage of an OSCM event such as a product recall, the firm’s stock price will be affected before the official announcement, and the market reaction captured on the event day may just be a residual adjustment of the real expectations.

5.2 Event window

The event window is the time period over which the effect of an event will be examined. An event window is denoted as (-*x*, +*y*). The announcement date of an event is usually set as day 0. It is also possible that the announcement is made public after the stock market is closed, then day 0 is adjusted as the next trading day after the announcement date. The event window (-*x*, +*y*) includes *x* trading days before day 0 to capture any information leakage, and *y* trading days after day 0 to account for any delay of the market in perceiving the information.

(Insert Table 4 here)

It is customary to expand the event window to several days around the event day. As shown in Table 4, 83% (24 articles) of the short-term event studies in OSCM adopt the standard event windows including day -1, day 0, and day 1, or some combinations of them. However, the event window could also be expanded longer if there are theoretical reasons to justify for the leakage or dissipation of information over a relatively long period (MacKinlay, 1997). In practice, it is a standard procedure to use alternative event windows for the robustness test. For example, Thirumalai and Sinha (2011) used various event windows including (-1, 0), (-1, +1), (-5, +1), (-5, +5), (-10, +1), and (-10, +10) to assess the sensitivity of their results.

The event windows do not typically overlap across different securities. The absence of overlap implies that the abnormal returns are independent across securities, satisfying the assumption for the subsequent significance tests. However, sometimes event window clustering is inevitable. For example, in the case of a single event such as a natural disaster, release of policy or other macroeconomic events, the event days are the same across the firms. A single event day would lead to considerable correlations of the abnormal returns among securities. In order to address the issue of cross-sectional correlation, several modifications of the traditional significance tests need be adopted, which we will discuss in Section 6.

5.3 Collect data

The process of collecting a representative sample of event announcements may not be necessary for external events such as the change of government policies and the occurrence of natural disasters, as these events could affect all firms in specific industries or geographic locations (e.g., Desai et al., 2007). However, for internal events, the process is important and can be further divided into three steps as follows: (1) select suitable data sources, (2) compile a set of keywords and set the time period during which the announcements will be collected, and (3) eliminate the confounding announcements.

Proper data sources have a good coverage of timely press releases and reach the major investors. Table 3 shows that most OSCM event studies collect announcements from two databases, namely Dow Jones Factiva and LexisNexis (e.g., Ba et al., 2013; Hendricks and Singhal, 2003; McGuire and Dilts, 2008; Sabherwal and Sabherwal, 2005; Xia et al., 2016). These two databases aggregate global information from major newswires including Public Relations (PR) Newswire, Business Wire, Dow Jones Newswires, Reuters News, *The New York Times*, *The Wall Street Journal*, and other news sources. While Dow Jones Factiva and LexisNexis are widely used, other databases with specialties are adopted in country-specific studies outside the U.S. and industry-specific studies. For example, in a study of quality management based in the Spanish market, Nicolau and Sellers (2002) collected announcements from the database Baratz, which contains information of news published in important Spanish newspapers. Studies in the Chinese context use databases such as China Infobank (Zhao et al., 2013) and WiseNews (Lam et al., 2016) that cover the major Chinese security newspapers, including *Shanghai Securities News*, *Securities Daily*, and *Secutimes*. In terms of industry-specific research, additional databases gathering industry information are often used as complementary. For example, studying product recalls in the Chinese automobile industry, Zhao et al. (2013) used the Chinese Automobile Recall Website, in addition to China Infobank. Girotra et al. (2007) searched the R&D Insight database developed by Australasian Drug Information Service (ADIS) international to probe into the pharmaceutical industry. In addition to conducting a primary search in multiple databases, a rigorous search process also includes a second search in other databases with wider coverage to address potential information leakage. For example, Modi et al. (2015) double checked Factiva to identify earlier announcements. If multiple announcements regarding the same event are identified, the announcement with the earliest date should be collected.

The selection of keywords and time period used in the searching process can be regarded as a tradeoff and usually requires multiple revisions. On the one hand, the searching process should generate a sufficient sample for statistical analysis. On the other hand, the set of keywords and time period should be conservative to ensure the definition of the event is explicit and consistent over time. In practice, keyword selection is a retrospective process. The primary search usually starts with a small set of keywords. A limited number of announcements well-fitting the boundaries of the event definition are collected. Then researchers read these announcements to identify additional phrases commonly used in the media. Finally, all the keywords identified will be included in searching for the announcements. As seen in Table 3, announcements are collected over time periods ranging from two to 38 years. The lengths of the time periods vary according to different event types. For some events occurring less frequently such as product recalls in the toy industry (Wood et al., 2017), announcements are collected over a longer time period. In spite of the wide range of time periods, most studies set their time periods around ten years. An extremely long time period could be problematic in some cases. For example, information technology adoption and international standards could have different definitions over time. Inconsistent definition of the event could generate biased results. For example, Lo et al. (2009) indicated that ISO 9000 underwent a major revision in 2000 with a change in emphasis, and a time-based investigation of ISO 9000 adoption is necessary.

The last step is to eliminate the confounding announcements. Confounding announcements are made by the same entity on dates around the event date. If not eliminated, other events rather than the event of interest may contaminate the measurement of the abnormal returns and decrease the internal validity, especially in short-term event studies. As in short-term event windows, the distribution of the abnormal returns due to the confounding announcements may not have a mean of zero (Sorescu et al., 2017). Our survey of the literature shows that OSCM researchers do not appear to have been sensitive to this issue. About 45% of the studies do not clearly state that they have eliminated the confounding announcements, as shown in Table 3. Among those studies eliminating the confounding announcements, practices vary across different studies due to a lack of strict guidance as to what type of announcements should be concerned about. For instance, Modi et al. (2015) only considered the announcements of earnings release, merger and acquisition, change of a CEO or CFO, debt restructuring, and an unexpected dividend change. Brandon-Jones et al. (2017) considered a wider range of information including all the announcements within the same event window.

5.4 Predict normal returns

In event studies, the effect of a specific event is measured by the stock market reaction, which is computed as the difference between actual and expected stock returns. As only the actual stock returns after the event can be observed, the stock returns in the absence of the event can only be estimated. Table 4 indicates that the most popular estimation model adopted in the literature is the market model (26 articles, 90%). Other statistical models adopted include the mean adjust model, market adjusted model, and Fama-French factor model.

The mean adjust model calculates the average return over the estimation window as the expected return for a specific security. Similarly, the market adjusted model uses the returns of the market portfolio return $R_{m,t}$ over the event period as the estimated normal return. The market model and Fama-French factor model are more sophisticated, which we introduce as follows:

Market model. The market model (Scholes and Williams, 1977) assumes a linear relationship between the return of a specific security and the return of the market portfolio as follows:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t} \text{ with}$$

$$E(\varepsilon_{i,t}) = 0 \text{ and } var(\varepsilon_{i,t}) = \sigma_{\varepsilon_i}{}^2,$$

where $R_{i,t}$ denotes the stock return for security i in period t, $R_{m,t}$ is the period t returns of the market portfolio, $\varepsilon_{i,t}$ is the zero mean disturbance term, and α_i and β_i are estimated for each security over the estimation window.

Fama-French four-factor model. The Fama-French four-factor model is an extension of the three-factor model (Fama and French, 1993) by adding a moment factor (MOM) (Carhart, 1997) as follows:

$$R_{i,t} - R_{f,t} = a_i + \beta_i(R_{m,t} - R_{f,t}) + s_i SMB_t + h_i HML_t + m_i MOM_t + \varepsilon_{i,t} \text{ with}$$

$$E(\varepsilon_{i,t}) = 0 \text{ and } var(\varepsilon_{i,t}) = \sigma_{\varepsilon_i}{}^2,$$

where $R_{i,t}$ denotes the stock return for security i in period t, $R_{m,t}$ is the period t returns of the market portfolio, $R_{f,t}$ is the period t risk-free return rate, SMB_t is the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks, HML_t is the difference between the returns on the diversified portfolios of high and low stocks, MOM_t is the difference between the portfolios of high prior return stocks and low prior returns, lagged one month, and $\varepsilon_{i,t}$ is the zero-mean residual.

The assumptions of these statistical models are that the stock returns are jointly normal, and independently and identically distributed through time. MacKinlay (1997) noted that although the assumptions are strong, they are empirically reasonable and the references using these models are robust to deviations from the assumptions. Therefore, ordinary least square (OLS) regression is often used for estimation.

Once the estimation model is chosen, the parameters in the factor models are estimated over the estimation window. As shown in Table 4, the estimation windows in the literature range from 120 days to 255 days. The estimation windows are usually long in order to address the bias in abnormal returns due to out-of-sample estimation. In addition, the estimation window typically does not overlap with the event window. Table 4 shows that the estimation window ends at least ten days prior to the event day. Avoiding overlap prevents the normal returns used to estimate the model parameters being influenced by the event. After the model parameters are estimated, the expected normal returns $\widehat{R_{i,t}}$ can be calculated over the event window.

5.5 Test abnormal returns

The abnormal return is calculated as a firm's actual *ex post* return minus its expected normal return over the event window. For firm *i* and event day *t*, the abnormal return is

$$AR_{i,t} = R_{i,t} - E(R_{i,t}),$$

where $AR_{i,t}$, $R_{i,t}$, and $E(R_{i,t})$ are the abnormal, actual, and expected returns, respectively. Then the abnormal returns are aggregated through the event window and across securities to capture the overall effect of the event as follows:

$$\overline{CAR}(t_1, t_2) = \frac{1}{N} \sum_{i=1}^{N} CAR_i\,(t_1, t_2),$$

where $\overline{CAR}(t_1, t_2)$ is the average cumulative abnormal returns over the event window (t_1, t_2) for all the securities i , $i = 1, \ldots, N$.

An important assumption for aggregation is that there is no clustering of the event windows among the securities, so $CAR_i(t_1, t_2)$ is assumed to be independent across the securities. The assumption of independence simplifies the calculation of the variance of $\overline{CAR}(t_1, t_2)$, as the covariance across the securities will be zero. In addition, the abnormal return is actually the disturbance term of the estimation model calculated on an out-of-sample bias. The additional variance due to the sampling error approaches zero after divided by the long estimation window. So the conditional variance of abnormal returns can be estimated as the disturbance variance $\sigma_{\varepsilon_i}{}^2$ in the estimation period.

Under the null hypothesis that the event has no impact on the stock returns, the cumulative abnormal return follows the distribution as follows:

$$\overline{CAR}(t_1, t_2) \sim N[0, var(\overline{CAR}(t_1, t_2))],$$

where

$$var(\overline{CAR}(t_1, t_2)) = \frac{1}{N^2} \sum_{i=1}^{N} (t_2 - t_1 + 1) \sigma_{\varepsilon_i}{}^2.$$

The null hypothesis that the cumulative abnormal return is zero can be tested using

$$\theta = \frac{\overline{CAR}(t_1, t_2)}{var(\overline{CAR}(t_1, t_2))^{1/2}} \sim N(0, 1).$$

The parametric *t*-test above is the traditional approach to assess the significance of the cumulative abnormal returns and has been used in many of the OSCM event studies (55%) (e.g., Hendricks and Singhal, 1997; Jacobs, 2014; Lin and Su, 2013; McGuire and Dilts, 2008). This approach, though simple, relies on relatively strong assumptions of independence and homoscedasticity among the abnormal returns. However, in practice, the assumptions sometimes can be violated in circumstances of clustering of the event days and event-induced volatility. Table 4 presents the traditional approach and the modifications adopted by OSCM researchers. The most commonly adopted modifications are the crude dependence adjustment test (Brown and Warner, 1985), standardized residual test (Patell, 1976), and standardized cross-sectional test (Boehmer et al., 1991). In addition to parametric tests, researchers also conduct non-parametric tests such as the Wilcoxon signed-rank test and binomial sign test to address the concern of skewness in the distribution of the abnormal returns (Hendricks and Singhal, 1996; Lam et al., 2016).

5.6 Cross-sectional analysis

Event study is powerful as it links the new information about an event of interest and stock prices by isolating the component of price changes due to the firm-specific event from other factors such as market-wide movements. Generally, significant positive abnormal returns indicate increased future performance expected by investors due to a specific event, and vice versa. As indicated in our survey, the market reaction to the same type of event varies in different contexts. For instance, while some studies show that product recalls have a negative impact on the financial performance of both manufacturers and retailers (Ni et al., 2014; Wood et al., 2017; Zhao et al., 2013), Thirumalai and Sinha (2011) found that firms in the medical device industry are not significantly affected by product recalls. Mixed results in the literature indicate that it would be informative to further investigate the patterns or determinants of variations in abnormal returns. However, the event study is limited in explaining the mechanisms of how the effect will vary across firms. Therefore, researchers of OSCM event studies often conduct cross-sectional regression and ANOVA to provide further insights (23 articles, 80%).

Cross-sectional regression is conducted to identify the determinants of variations in abnormal returns. The dependent variable is the cumulative abnormal return for each security over the event window, and the independent variables usually include the moderating variables specific to each research context. For instance, Kalaignanam et al. (2013) found that, in Customer Relationship Management (CRM) outsourcing, capabilities of the outsourcing firms, distance between the outsourcing firm and the vendor, and the type of CRM process being outsourced moderate the shareholder value of CRM outsourcing. Jacobs (2014) showed that the market reaction to voluntary emission reduction is associated with the time, emissions type, and whether the reduction is announced *ex ante* or *ex post*.

In addition to the moderating variables unique to each research context, it is also important to include firm-level, industry-level, and macro-level control variables to account for the influences of other factors on the stock returns. In line with the finance literature, most OSCM event studies adopt firm-level variables such as firm size, financial leverage, and book-to-market ratio; industry-level variables such as industry dummy variables and industry competition; and macro-level variables including recession dummy variables and time trend.

ANOVA is adopted to separate the mixed effects among different subgroups from the overall effect (Paulraj and Jong, 2011; Zhao et al., 2013). In essence, ANOVA is equivalent to linear regression in terms of the estimation model, whereas they have different concentrations. Linear regression is mostly concerned about identifying variables that either mitigate or magnify the abnormal returns, while ANOVA concentrates on discerning the mixed effects between subgroups with different characteristics.

6. Recommendations for future short-term event studies in OSCM

The systematic review of the practices in conducting short-term event studies in OSCM allows us to uncover several methodological issues that need further attention. We identify several research design issues regarding event identification, event window selection, confounding effect, self-selection bias, estimation model, significance test, and time and industry clustering, and suggest ways to address them, thus providing OSCM researchers with practical recommendations for conducting future short-term event studies.

6.1 External events and non-U.S. context

Our analysis of short-term event studies in OSCM indicates that most researchers focus on internal corporate events in the U.S. context, while less is known about the effects of external events and in the non-U.S. context. While it seems to be the same case as in other areas such as marketing (Sorescu et al., 2017), we believe OSCM researchers should pay special attention to such research opportunities due to the emergence of the global supply chain. In particular, firms are more closely related than ever and can hardly be isolated from the risks originating from external supply chain partners or catastrophic disasters across national borders. In addition, non-U.S. countries, especially developing countries, have been playing the prominent role of being sourcing destinations in global supply chains. Validating findings from previous studies across different countries is important in advancing our understanding of the global value of OSCM events.

First, while it is intuitively compelling that supply chain disruptions have negative impacts within a specific company, it remains unsettled as to the transmission effects on external parties. Negative or positive transmission effects have been documented for firms having cooperative or competitive supply chain relationships with initially-disrupted firms (Barrot and Sauvagnat, 2016; Erwin and Miller, 1998; Ferstl et al., 2012). In our survey of short-term event studies in OSCM, only one study conducted by Jacobs and Singhal (2017) documents the shareholder value effect of external events in terms of the Rana Plaza disaster in Bangladesh.

Second, despite the important role of developing countries in global supply chains, event studies in developing countries are far from adequate. Event studies in developing countries complement our existing knowledge in developed countries. The same type of events could have different or even opposite effects in the context of developed and developing countries having different cultural, political, and institutional environments. For example, Lam et al. (2016) found that in contrast to the Western context, Chinese investors react negatively to corporate environmental initiatives in China. They believe that the difference could be explained by Chinese investors' risk-taking investment strategy and China's fluctuating environmental policies and regulations.

One challenge of conducting event studies regarding external events is the concern about cross-sectional correlation in the significance test for abnormal returns. As previously argued, an important assumption for the traditional significance test of cumulative abnormal returns is independence among the securities. This assumption requires that the event days do not overlap and the correlation among the securities is assumed to be zero. Otherwise, in the case of total clustering, meaning the event days for all the securities are the same, the under-estimated covariance between abnormal returns will lead to a substantial over-rejection problem (MacKinlay, 1997; Kolari and Pynnönen, 2010). In event studies of internal activities, the event announcements are checked before analysis to ensure that there is no overlapping of the event windows. However, in event studies of external events, especially in the cases of policy change, industrial regulations, catastrophic disasters, and wars, the event days are the same. We suggest that researchers studying external events modify the traditional significance test to correct the problem of cross-sectional correlation. Two common modifications are the test using time-series mean abnormal returns (Brown and Warner, 1985) and the test using calendar-time abnormal returns (Jaffe, 1974). Jacobs and Singhal (2017) tested the time-series mean abnormal returns in their study of Bangladesh collapse to address the problem of correlation resulting from the same event day.

The other challenge arising from the non-U.S. context is the concern of market efficiency in emerging markets. The fundamental assumption of conducting short-term event studies is Efficient Market Hypothesis, a violation of which may lead to unconvincing conclusions. Some event studies in finance also cast doubt on the efficiency of emerging markets with empirical evidence. For instance, based on a study of Mexican Stock Exchange, Bhattacharya et al. (2000) found that firms' stock prices are not sensitive to a variety of corporate news announcements, as the unrestricted insider trading causes the stock prices to fully incorporate the superior information before public announcements. Moreover, Bekaert and Harvey (2002) pointed out that emerging markets are typically characterized as thin markets, where infrequent trading and slow adjustment to information may result in high serial correlation in daily returns. In addition, Chinese stock market was not completely open until the non-tradable shares (NTS) reform initiated in 2005 (Liu and Tian, 2012). Before the NTS reform, holders of non-tradable shares had almost the same rights as holders of tradable shares, except for public trading. Therefore, OSCM researchers who are interested in conducting short-term event studies in emerging markets should pay close attention to the issue of market efficiency and perform additional tests (e.g., alternative event windows, adjusted significance tests) to verify the robustness of their findings. For instance, in addition to the three-day event window, Lam et al. (2016) recalculated the abnormal stock returns over longer event windows ranging from 5 to 21 days to verify their findings regarding Chinese investors' reactions to corporate environmental initiatives. On the other hand, in order to address the concern of serial correlation resulted from non-synchronous trading, Chen et al. (2009) adopted the cross-sectional test and standardized cross-sectional test (Boehmer et al., 1991) to address the concern of serial correlation in the Chinese stock market. Moreover, in an investigation of environmental incidents in the Chinese context, Lo et al. (2017) excluded the announcements made in or before 2005 in consideration of potential violation of the Efficient Market Hypothesis due to non-tradable shares.

6.2 Justify the event window

Although there is no universal rule on the lengths of the event windows, our survey of short-term event studies shows that the event windows are usually short. About 83% (24 articles) of the studies set the event window as combinations of -1, 0, +1 days. Short event windows are recommended not only based on the Efficient Market Hypothesis, but also due to the costs of expanding them. According to the Efficient Market Hypothesis, the stock market reacts almost immediately to any new information available. Therefore, without theoretical justifications for information leakage or slow dissipation, including one pre-event day and one post-event day should be sufficient to account for possible information leakage, as well as the market reaction after the stock market is closed. Moreover, expanding the event windows leads to decreased sample size and reduced power of analysis (Brown and Warner, 1985). As discussed previously, preliminary sample announcements need to be checked to remove confounding events and overlapping event windows. Longer event windows are more likely to be affected by confounding events, as well as overlapping with the event windows of other firms. Decreasing the sample size can be costly, especially when the preliminary sample size is already small. In addition, the power of analysis will be substantially decreased. Brown and Warner (1985) compared the power of analysis when the abnormal returns are measured over the event windows of 0 and (-5, +5). They found that with an actual level of 1% abnormal performance, the rejection frequency for market adjusted returns is only 13.2% in the 11-day event window, compared with 79.6% in the one-day event window.

However, with theoretical justifications, event windows can be expanded according to the nature of the event. One example is the event window of (0, +11) in a study of a catastrophic disaster (Jacobs and Singhal, 2017). The authors argued that a disaster such as the collapse of a garment factory is unexpected and unintended, so there is no evidence of information leakage.

Besides, the information about the severity of the disaster may be gradually revealed, so it is reasonable to include longer post-event days. Unfortunately, our survey shows that two of the five event studies with longer event windows do not provide clear justifications (i.e., Lin and Su, 2013; Nicolau and Sellers, 2002).

6.3 Confounding announcements

The isolation of the confounding effect of other financially related events is perhaps one of the most critical assumptions of the short-term event study method (McWilliams and Siegel, 1997). McWilliams and Siegel (1997) demonstrated the importance of controlling confounding announcements by replicating three event studies of corporate social responsibility published in *Academy of Management Journal.* They found that after controlling the confounding effect, the significant abnormal returns reported in the three event studies all became insignificant.

However, our survey shows that efforts should be made to strengthen the awareness of controlling confounding announcements among OSCM researchers. In particular, in addition to emphasizing the necessity of controlling confounding effect, more discussion is needed about the execution of identifying confounding announcements, as there is no strict guidance in the literature as to what announcements should be controlled. Table 3 shows that some researchers examined the sample announcements and excluded those containing both the event of interest and other material information (Hendricks and Singhal, 2003; Hendricks et al., 2009; Jacobs, 2014). Some other researchers considered the announcements which have been shown to significantly affect stock returns including earnings or dividends announcements, key executive appointments, merger and acquisitions, restructuring or divestiture (Klassen and McLaughlin, 1996; Lam et al., 2016; Modi et al., 2015; Nicolau and Sellers, 2002; Paulraj and Jong, 2011; Sabherwal and Sabherwal, 2005). Other researchers set a wider range and argued that any other announcements released by the sample firm around the event date may cause potential contamination (Brandon-Jones et al., 2017; Hendricks and Singhal, 1996; Jacobs and Singhal, 2017; Jacobs et al., 2010). It is noteworthy that eliminating confounding announcements with a broader definition or over a longer time period may reduce the possibility of contamination, but it could also reduce the sample size significantly. To strike a balance, we recommend researchers to at least control those common confounding announcements identified by McWilliams and Siegel (1997), such as dividend declarations, earnings announcements, key executive appointments, restructuring or divestiture, merger and acquisition, joint ventures, major litigation or labor unrest, forecasted changes in sales or earnings, and major contracts over the event window.

6.4 Self-selection bias

The majority of the event studies we reviewed are based on self-announced events adopted voluntarily by firms. Firms proactively initiate events such as environmental management, quality management, R&D projects, sourcing strategies rather than being passively prompted to pursue them. For instance, Ni et al. (2014) are interested to assess how product recalls may affect the U.S. public-listed retailers' stock returns. In the cumulative abnormal return (CAR) analysis, the effect of product recalls is quantified as the actual ex-post return minus the estimated normal return of the firms making recall announcements. However, as suggested by the authors, retailers who choose to initiate product recalls may differ from those who choose not to. Specifically, firms with better reputation are more likely to initiate product recalls. Due to the self-selection, a significant difference in mean abnormal returns could be observed between the two populations independent of the impact of product recalls. For example, firm reputation has been shown to affect consumers' reactions to product harm crisis (Siomkos and Kurzbard, 1994). Consumers felt that the products failures are less severe when sold by firms with better reputation. Therefore, the average treatment effect calculated with only the treated group (i.e., CAR for the U.S. public-listed retailers making announcements) may underestimate the average treatment effect on the population (i.e., the "true" effect on all U.S. public-listed retailers) (Austin, 2011; Heckman, 1979).

In the cross-sectional analysis, the CAR of a particular firm is usually regressed on its observable characteristics to explain the variations in the CAR. However, as CAR is only observed for a subsample of the population (i.e., the firms making announcements), there could be a problem of endogeneity if the self-selection process is omitted from the cross-sectional model. In the example we mentioned above, an unobserved factor (i.e., firm reputation) may affect a firm's decision to initiate a product recall as well as its abnormal stock return (Ni et al., 2014). In this case, the unobserved factor manifests in the residual of the cross-sectional model, making the residual correlated with the explanatory variables (i.e., observable characteristics such as recall size and remedy strategies) and the dependent variable (i.e., CAR). Consequently, omitting the self-selection process in the cross-sectional model potentially violates OLS' assumption of exogeneity, leading to the bias in the estimation of coefficients (Clougherty et al., 2016).

Researchers should address the potential sample selection bias resulting from the systematic differences between the sample and non-sample firms. Our survey shows that only seven out of the 29 studies address the potential sample selection bias issue (i.e., Paulraj and Jong, 2011; Dam and Petkova, 2014; Hendricks et al., 2009; Jacobs, 2014; Kalaignanam et al., 2013; Modi et al., 2015; Ni et al., 2014).

To correct the biased estimation of treatment effect in the CAR analysis, a common practice is to mimic the random selection process. Researchers construct a benchmark group and directly compare the abnormal stock returns between the sample firms and the benchmark firms. The benchmark firms are selected from the pool of firms not involved in the events based on certain criteria. Conditional on the specific matching criteria, the distribution of observed baseline characteristics is similar between the sample firms and benchmark firms. Then the differences in abnormal stock returns during the event window are calculated and tested for significance. While the rationale to control for self-selection bias is the same, approaches to generate the benchmark group vary across different studies.

Traditionally, researchers use the one-to-portfolio or one-to-one matching approach to develop the benchmark group (e.g., Paulraj and Jong, 2011; Hendricks et al., 2009). Specifically, all the listed firms are assigned to portfolios based on various characteristics that are believed to influence stock returns. The characteristics frequently included in the OSCM event studies are industry, firm size, and prior firm performance. Then a group of firms or a single closest firm in the same portfolio to the sample firm is selected as the benchmark. Admittedly, it is difficult to get benchmarks that are all well matched on all the criteria and there are tradeoffs among criteria. There are also some limitations when high-dimension criteria are used because it is difficult to determine along which dimensions to match and which weighting scheme to adopt (Dehejia and Wahba, 2002).

Propensity score matching (PSM) is another approach used in the OSCM literature to construct the benchmark group (Modi et al., 2015). Different from the portfolio matching method, PSM reduces the dimensionality by generating a propensity score. The propensity score is the probability of treatment assignment conditional on observed baseline characteristics. It can be estimated with a probit or logit model from the observational data on treatment assignment and baseline characteristics. Based on the estimated propensity scores of all the firms, the firms in the comparison group that have the closest scores to the sample firms are identified as the benchmark.

To address the omitted variable bias in the cross-sectional analysis, an approach commonly adopted is Heckman’s two-stage selection model (Dam and Petkova, 2014; Kalaignanam et al., 2013; Ni et al., 2014). Different from the two aforementioned matching methods that mimic the random selection process in the context of observational studies, this model corrects the sample selection bias by first estimating the values of the omitted variables, and then using the values as regressors in estimating the effect of the event on the stock returns (Heckman, 1979). Accordingly, Heckman's model includes two equations. In the first equation, the probability of a firm undertaking a specific event is modelled with probit analysis for the full sample. The inverse Mills ratio is generated from the first equation and represents the probability that an observation is selected to include in the sample. In the second equation, the effect of the event on abnormal returns is estimated with the OLS function. The inverse Mills ratio is added as an additional explanatory variable in the OLS function and indicates whether selection bias is an issue. One of the concerns in implementing this method is the selection of variables that may account for the selection bias.

A key challenge to implementing both PSM and Heckman's two-stage model is to determine the explanatory variables to be included in estimating the selection model. The possible sets of variables recommended in the literature include baseline variables that influence the outcome (i.e., stock returns in event studies) and baseline variables that influence the treatment assignment (i.e., the probability of occurrence of the event) (Austin, 2011; Heckman and Navarro-Lozano, 2004). In practice, the baseline variables are usually selected specific to each research context, based on theoretical justifications, and tested with difference analysis. For example, Dam and Petkova (2014) assumed that consumer pressure that differs across industries explains firms' participation in supply chain sustainability programmes. They further tested whether there are differences in firm-level characteristics that could serve as potential baseline variables. Based on the information from the two steps, they included industry dummy as the explanatory variable in the probit model. Modi et al. (2015) included the variables of productivity, leverage, capital resource slack, market-to-book ratio, and firm size that affect abnormal returns as the baseline variables.

6.5 Estimation model

The statistical asset pricing models adopted in short-term event studies in OSCM are two simple models including the mean adjusted model and market adjusted model, and two factor models including the market model and Fama-French factor model. Among the four models, the factor models are commonly adopted for major data analysis, while the other two simple models are often used in the sensitivity test. The factor models are believed to be superior to the simple models in that they account for the movement in market returns in estimating the normal returns (MacKinlay, 1997). Consequently, they will reduce the variance in the estimated returns and enhance the ability to detect abnormal returns. In recent years, a number of sophisticated statistical asset pricing models have been proposed. For example, the Fama-French three-factor model (Fama and French, 1993) extends the capital asset pricing model (CAPM) by adding the size and value factors to the market risk factor. The model is further extended by adding a momentum factor by Carhart (1997), and the profitability and investment factors by Fama and French (2016).

Our survey reveals a surprising fact that despite the increased sophistication, the market model has been consistently used by most researchers for stock return estimation from the earliest study we identified (Hendricks et al., 1995) to the latest research (Brandon-Jones et al., 2017; Jacobs and Singhal, 2017; Wood et al., 2017). This is because the improvement is very conservative with the increase in model sophistication in short-term event studies, and more sophisticated models usually yield similar results with the market model (Brown and Warner, 1985). As the daily expected normal returns usually approach zero, the reduced variance in the expected returns is too limited compared with the much larger abnormal returns. The lack of sensitivity to the models explains the prevalence of the market model across different studies in all the time periods. Therefore, we suggest that researchers choose the factor models according to the availability of data with little preference for the more sophisticated models.

However, in some cases, employing the multi-factor model could bring substantial improvement. MacKinlay (1997) suggested that if firms share common characteristics such as coming from the same industry or concentrating in the same capitalization group, researchers should consider a more sophisticated model. Since there are no specific guidelines as to under which circumstances the more sophisticated models are necessary, we suggest that researchers, whenever possible, should estimate the expected returns using alternative models to enhance the robustness of the analysis.

6.6 Significance tests

The most widely adopted parametric test (16 articles, 55%) in the studies we reviewed is the classical *t*-test. As previously introduced, the test assumes that the stock returns are jointly multivariate normal, and independently and identically distributed across time and among individuals (MacKinlay, 1997). Yet, in some research settings, these statistical assumptions are likely to be violated and the inferences from the classical *t*-test tend to be problematic. Researchers have modified the test to correct for prediction errors. OSCM researchers seem to be sensitive to the issue of significance tests and the most widely adopted modifications are those developed by Patell (1976), Brown and Warner (1985), and Boehmer et al. (1991). Table 5 presents a summary of the parametric tests commonly adopted in OSCM studies with key references, strengths, weaknesses, and representative OSCM studies identified for each test.

(Insert Table 5 here)

Since there is no universal best significance test that is well-specified in all the circumstances, the choice of test statistic should be based on the specific research setting and statistical features of the dataset under investigation. For example, Brown and Warner (1985) suggested that the adjustment of cross-sectional dependence is only necessary in special cases of extreme cross-sectional correlation such as those when firms come from the same industry or share the same event day. We suggest that researchers of industry-specific studies or studies allowing for clustering of event days should be sensitive to the problem of cross-sectional correlation. Examples of such modifications are studies conducted by Hendricks and Singhal (1996), and Jacobs and Singhal (2017).

6.7 Time and industry clustering

Time and industry clustering are two critical issues which potentially cause misspecification in significance tests, but they are sometimes ignored by OSCM researchers. Time clustering could be an issue when the events occur at or near the same calendar date (Henderson, 1990). It is often observed in the event studies with a focus of external events such as regulations, legislations, policies, and disasters, where firms share common event days (Kolari and Pynnönen, 2010). For example, in an investigation of the impact of Bangladeshi garment factory collapse on apparel retailers, the event day is set as the date of the Rana Plaza disaster on April 24, 2013 (Jacobs and Singhal, 2017). When the event windows overlap or are the same, the abnormal returns of sample firms are potentially correlated, which may result in non-zero covariance among abnormal returns (MacKinlay, 1997). On the other hand, industry clustering refers to the situation when the events are concentrated in the same or a small number of industries (Henderson, 1990). For instance, Girotra et al. (2007) investigated the influence of phase III clinical trial failures on pharmaceutical companies. Wood et al. (2017) examined the effect of product recalls on toy manufacturers and retailers. In the case of industry clustering, abnormal returns of industry peers tend to contemporaneously move together as they usually share common fundamentals such as supply and demand shocks. Dyckman et al. (1984) found that the variance of the return residuals across securities in the same industry is significantly higher, even if their returns are sufficiently diversified over time.

Time and industry clustering may cause problems in the significance test, as the vital assumption of cross-sectional independence is likely to be violated. The first step in the significance test is to aggregate abnormal returns across securities. For the aggregation, it is assumed that there is no clustering across securities so that the covariance term can be regarded as zero (MacKinlay, 1997). However, in the case of time and industry clustering, the abnormal returns across securities are potentially correlated. Ignoring the cross-sectional correlation may cause a downward bias in the estimation of the standard deviation of abnormal returns. As a result, the null hypothesis of zero abnormal returns will be rejected too frequently. Moreover, the significance test could be further misspecified in the case of both time and industry clustering, as both problems reinforce one another (Dyckman et al., 1984).

To address the concern of cross-sectional correlation, various approaches have been proposed in the literature. One of the most popular approaches is the portfolio approach (Brown and Warner, 1985; Jaffe, 1974). In this approach, the significance test is performed at the portfolio level so that the cross-sectional correlation across securities in the portfolio is allowed. Specifically, the securities in a specified time period are first included into one or several portfolios. Next, the average abnormal return for the portfolio is calculated as the abnormal returns aggregated over securities in the portfolio divided by the number of the securities. With the assumption that the portfolio abnormal returns are independently, identically and normally distributed over time, Student *t*-test can be employed to test the time-series of portfolio abnormal returns. The other approach is to correct the underestimated standard deviation by taking into account a correlation factor (Kolari and Pynnönen, 2010). For example, based on the BMP test (Boehmer et al., 1991), Kolari and Pynnönen (2010) proposed an ADJ-BMP test which adjusts the cross-sectional correlation. In the BMP test, the abnormal returns during the event period are standardized by the estimation-period standard deviation, and then the standardized abnormal returns are divided by its contemporaneous cross-sectional standard deviation. BMP test allows serial correlation, heteroscedasticity among abnormal returns and event-induced volatility, but it is prone to cross-sectional correlation. The ADJ-BMP test modifies the cross-sectional standard deviation by adding the average of the cross-correlation of the estimation-period residuals, which accounts for the cross-sectional correlation among abnormal returns in the event period.

7. Conclusions and limitations

Reviewing 29 short-term event studies in OSCM published between 1995 and 2017, we observe that the short-term event studies in OSCM are on the increase and about 62% of the papers were published in the recent eight years from 2010 to 2017. As the basic steps of short-term event studies remain essentially the same, our study first outlines the basic steps as suggested by MacKinlay (1997). For each step, we then analyze the practices adopted in these OSCM papers in detail. First, we find that 28 articles (97%) focus on internal corporate events, with only one article (3%) examining an external event in terms of a catastrophic disaster. Most event studies are in the U.S. context, and only five studies (17%) are in the non-U.S. context. Second, the study demonstrates that the standard event windows (i.e., including day -1, day 0, and day 1) are widely adopted in short-term event studies. However, theoretical justifications are not provided in some event studies with longer event windows. Third, multiple data sources are often used to enhance the rigour of data collection, but elimination of confounding announcements is not implemented well. About 45% of the studies do not clearly state that they have eliminated the confounding announcements, and practices vary across different studies with confounding eliminations. Fourth, our study shows that researchers are not sensitive to the estimation model of normal returns. The market model is the most popular estimation model, which is adopted in 26 articles (90%) from 1995 to 2017. Fifth, OSCM researchers are wary of possible violations of the assumptions for the significance tests. Various modifications of the classical *t*-test are adopted according to different research contexts. Sixth, subsequent cross-sectional regression and ANOVA are usually conducted to probe into the operational determinants of variations in abnormal returns (23 articles, 80%).

Based on the above analysis, we propose several recommendations for future short-term event studies in OSCM. First, we suggest that OSCM researchers pay special attention to external events that may create transmission effects along global supply chains. In addition, researchers should be careful about expanding the event windows, and provide theoretical explanations to justify the window lengths. Third, as removing confounding effect is a critical step in conducting short-term event studies, researchers should at least control those commonly identified newsworthy confounding announcements over the event window. Fourth, self-selection bias should be tested and well controlled, especially in short-term event studies with voluntary announcements. Fifth, employing the multi-factor model could bring substantial improvement. We recommend that researchers estimate the normal returns using alternative models to enhance the robustness of the analysis. Sixth, it is necessary to modify the significance tests according to research settings in the case of external events and industry-specific studies. Finally, we urge researchers to address the concern of cross-sectional correlation in the cases of time and industry clustering.

We acknowledge that our study is limited in terms of the scope. Not all types of event studies have been taken into account. However, considering the fact that short-term event studies are the most widely adopted in OSCM research, the summary and recommendations are valuable to shed light on this topic. Also, as this study primarily deals with the methodological issues in short-term event studies, we do not focus on the results and conclusions in specific studies. To further enhance our knowledge about event studies in

OSCM, this study can be extended in two ways. First, our study provides a comprehensive but not exhaustive review of the event studies in OSCM. It is possible to review the research undertaken with other types of event study methodologies such as long-term event studies and event studies with operating performance measures. Second, it would also be informative to investigate the consequences of various OSCM events and operational variables that account for variations in abnormal returns from the theoretical perspective. Different from traditional OSCM research that only focuses on one key outcome such as speed or quality, event studies in OSCM are based on the notion of strategic OSCM aimed at yielding competitive advantage and creating superior financial performance. Event studies in OSCM usually conduct ANOVA and cross-sectional regression to explain variations in abnormal returns, which rely on various theoretical lens and frameworks. Therefore, a future review of the diverse theoretical perspectives adopted in OSCM event studies will deepen our understanding of the financial impact of OSCM practices.

References

Austin, P.C. (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46(3), 399-424.

Ba, S., Lisic, L.L., Liu, Q., & Stallaert, J. (2013) Stock market reaction to green vehicle innovation. *Production and Operations Management*, 22(4), 976-990.

Ball, R., & Brown, P. (1968) An empirical evaluation of accounting income numbers. *Journal of Accounting Research*, 6(2), 159-178.

Barber, B.M., & Lyon, J.D. (1996) Detecting abnormal operating performance: The empirical power and specification of test statistics. *Journal of Financial Economics*, 41(3), 359-399.

Barber, B.M., & Lyon, J.D. (1997) Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics*, 43(3), 341-372.

Barrot, J.-N., & Sauvagnat, J. (2016) Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, 131(3), 1543-1592.

Bayus, B.L., Erickson, G., & Jacobson, R. (2003) The financial rewards of new product introductions in the personal computer industry. *Management Science*, 49(2), 197-210.

Bekaert, G., & Harvey, C. R. (2002) Research in emerging markets finance: looking to the future. *Emerging Markets Review,* 3(4), 429-448.

Bhattacharya, U., Daouk, H., Jorgenson, B., & Kehr, C.-H. (2000) When an event is not an event: the curious case of an emerging market. *Journal of Financial Economics*, 55(1), 69-101.

Binder, J. (1998) The event study methodology since 1969. *Review of Quantitative Finance and Accounting*, 11(2), 111-137.

Boehmer, E., Masumeci, J., & Poulsen, A.B. (1991) Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics*, 30(2), 253-272.

Brandon-Jones, E., Dutordoir, M., Frota Neto, J.Q., & Squire, B. (2017) The impact of reshoring decisions on shareholder wealth. *Journal of Operations Management*, 49, 31-36.

Brown, S.J., & Warner, J.B. (1985) Using daily stock returns: The case of event studies. *Journal of Financial Economics*, 14(1), 3-31.

Carhart, M.M. (1997) On persistence in mutual fund performance. *The Journal of Finance*, 52(1), 57-82.

Chen, D. H., Liu, H. H., & Huang, C. T. (2009) The announcement effect of cash dividend changes on share prices: An empirical analysis of China. *Chinese Economy*, *42*(1), 62-85.

Clougherty, J. A., Duso, T., & Muck, J. (2016) Correcting for self-selection based endogeneity in management research: Review, recommendations and simulations. *Organizational Research Methods*, 19(2), 286-347.

Corbett, C.J., Montes-Sancho, M.J., & Kirsch, D.A. (2005) The financial impact of ISO 9000 certification in the United States: An empirical analysis. *Management Science*, 51(7), 1046-1059.

Corrado, C.J. (2011) Event studies: A methodology review. *Accounting & Finance*, 51(1), 207-234.

Dam, L., & Petkova, B.N. (2014) The impact of environmental supply chain sustainability programs on shareholder wealth. *International Journal of Operations & Production Management*, 34(5), 586-609.

Dehejia, R.H., & Wahba, S. (2002) Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84(1), 151-161.

Desai, M. A., Dyck, A., & Zingales, L. (2007) Theft and taxes. *Journal of Financial Economics*, *84*(3), 591-623.

Dolley, J.C. (1933) Characteristics and procedure of common stock split-ups. *Harvard Business Review*, 11(3), 316-326.

Dyckman, T., Philbrick, D., & Stephan, J. (1984). A comparison of event study methodologies using daily stock returns: A simulation approach. *Journal of Accounting Research*, 22, 1-30.

Erwin, G.R., & Miller, J.M. (1998) The intra-industry effects of open market share repurchases: Contagion or competitive? *Journal of Financial Research*, 21(4), 389-406.

Fama, E.F., Fisher, L., Jensen, M.C., & Roll, R. (1969) The adjustment of stock prices to new information. *International Economic Review*, 10(1), 1-21.

Fama, E.F., & French, K.R. (1993) Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.

Fama, E.F., & French, K.R. (2015) A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22.

Fama, E.F., & French, K. R. (2016) Dissecting anomalies with a five-factor model. *The Review of Financial Studies*, 29(1), 69-103.

Ferstl, R., Utz, S., & Wimmer, M. (2012) The effect of the Japan 2011 disaster on nuclear and alternative energy stocks worldwide: An event study. *Business Research*, 5(1), 25-41.

Fosfuri, A., & Giarratana, M.S. (2009) Masters of war: Rivals' product innovation and new advertising in mature product markets. *Management Science*, 55(2), 181-191.

Girotra, K., Terwiesch, C., & Ulrich, K.T. (2007) Valuing R&D projects in a portfolio: Evidence from the pharmaceutical industry. *Management Science*, 53(9), 1452-1466.

Gopal, A., Goyal, M., Netessine, S., & Reindorp, M. (2013) The impact of new product introduction on plant productivity in the North American automotive industry. *Management Science*, 59(10), 2217-2236.

Heckman, J., & Navarro-Lozano, S. (2004) Using matching, instrumental variables, and control functions to estimate economic choice models. *The Review of Economics and Statistics*, 86(1), 30-57.

Heckman, J.J. (1979) Sample selection bias as a specification error. *Econometrica*, 47(1), 153-161.

Henderson, G.V. (1990) Problems and solutions in conducting event studies. *Journal of Risk and Insurance*, 57(2), 282-306.

Hendricks, K.B., & Singhal, V.R. (1996) Quality awards and the market value of the firm: An empirical investigation. *Management Science*, 42(3), 415-436.

Hendricks, K.B., & Singhal, V.R. (1997) Delays in new product introductions and the market value of the firm: The consequences of being late to the market. *Management Science*, 43(4), 422-436.

Hendricks, K.B., & Singhal, V.R. (2001) The long-run stock price performance of firms with effective TQM programs. *Management Science*, 47(3), 359-368.

Hendricks, K.B., & Singhal, V.R. (2003) The effect of supply chain glitches on shareholder wealth. *Journal of Operations Management*, 21(5), 501-522.

Hendricks, K.B., & Singhal, V.R. (2005) An empirical analysis of the effect of supply chain disruptions on long-run stock price performance and equity risk of the firm. *Production and Operations Management*, 14(1), 35-52.

Hendricks, K.B., Singhal, V.R., & Wiedman, C.I. (1995) The impact of capacity expansion on the market value of the firm. *Journal of Operations Management*, 12(3), 259-272.

Hendricks, K.B., Singhal, V.R., & Zhang, R. (2009) The effect of operational slack, diversification, and vertical relatedness on the stock market reaction to supply chain disruptions. *Journal of Operations Management*, 27(3), 233-246.

Jacobs, B.W. (2014) Shareholder value effects of voluntary emissions reduction. *Production and Operations Management*, 23(11), 1859-1874.

Jacobs, B. W., & Singhal, V. R. (2014) The effect of product development restructuring on shareholder value. *Production and Operations Management,* 23(5), 728-743.

Jacobs, B.W., & Singhal, V.R. (2017) The effect of the Rana Plaza disaster on shareholder wealth of retailers: Implications for sourcing strategies and supply chain governance. *Journal of Operations Management*, 49, 52-66.

Jacobs, B.W., Singhal, V.R., & Subramanian, R. (2010) An empirical investigation of environmental performance and the market value of the firm. *Journal of Operations Management*, 28(5), 430-441.

Jaffe, J.F. (1974) Special information and insider trading. *The Journal of Business,* 47(3), 410-428.

Kalaignanam, K., Kushwaha, T., Steenkamp, J.-B.E., & Tuli, K.R. (2013) The effect of CRM outsourcing on shareholder value: A contingency perspective. *Management Science*, 59(3), 748-769.

Klassen, R.D., & McLaughlin, C.P. (1996) The impact of environmental management on firm performance. *Management Science,* 42(8), 1199-1214.

Konchitchki, Y., & O'Leary, D.E. (2011) Event study methodologies in information systems research. *International Journal of Accounting Information Systems,* 12(2), 99-115.

Kothari, S., & Warner, J. (2007) Econometrics of event studies, *Handbook of Corporate Finance: Empirical Corporate Finance,* B. Espen Eckbo, ed: Elsevier/North-Holland.

Kolari, J., & Pynnönen, S. (2010) Event study testing with cross-sectional correlation of abnormal returns. *The Review of Financial Studies*, 23(11), 3996-4025.

Lam, H.K.S., Yeung, A.C.L., Cheng, T.C.E., & Humphreys, P.K. (2016) Corporate environmental initiatives in the Chinese context: Performance implications and contextual factors. *International Journal of Production Economics*, 180, 48-56.

Lin, C.-S., & Su, C.-T. (2013) The Taiwan national quality award and market value of the firms: An empirical study. *International Journal of Production Economics,* 144(1), 57-67.

Liu, Q., & Tian, G. (2012) Controlling shareholder, expropriations and firm's leverage decision: Evidence from Chinese non-tradable share reform. *Journal of Corporate Finance*, 18(4), 782-803.

Lo, C.K.Y., Pagell, M., Fan, D., Wiengarten, F., & Yeung, A.C.L. (2014) OHSAS 18001 certification and operating performance: The role of complexity and coupling. *Journal of Operations Management*, 32(5), 268-280.

Lo, C.K.Y., Yeung, A.C.L., & Cheng, T.C.E. (2009) ISO 9000 and supply chain efficiency: Empirical evidence on inventory and account receivable days. International Journal of *Production Economics*, 118(2), 367-374.

Lo, C.K.Y., Yeung, A.C.L., & Cheng, T.C.E. (2012) The impact of environmental management systems on financial performance in fashion and textiles industries. *International Journal of Production Economics*, 135(2), 561-567.

Lo, C.K.Y., Tang, C.S., Zhou, Y., Yeung, A.C.L., & Fan, D. (2017) Environmental incidents and the market value of firms: An empirical investigation in the Chinese context. Working Paper.

Lyon, J. D., Barber, B. M., & Tsai, C. L. (1999) Improved methods for tests of long-run abnormal stock returns. *The Journal of Finance*, 54(1), 165-201.

MacKinlay, A.C. (1997) Event studies in economics and finance. *Journal of Economic Literature*, 35(1), 13-39.

Malkiel, B.G., & Fama, E.F. (1970) Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383-417.

Martin, C., & Lynette, R. (1999) Supply chain strategy: Its impact on shareholder value. *The International Journal of Logistics Management*, 10(1), 1-10.

McGuire, S.J., & Dilts, D.M. (2008) The financial impact of standard stringency: An event study of successive generations of the ISO 9000 standard. *International Journal of Production Economics*, 113(1), 3-22.

McWilliams, A., & Siegel, D. (1997) Event studies in management research: Theoretical and empirical issues. *Academy of Management Journal*, 40(3), 626-657.

Min, S., & Wei, Y. (2013) Supply chain management and financial performance: Literature review and future directions. *International Journal of Operations & Production Management*, 33(10), 1283-1317.

Mitra, S., & Singhal, V. (2008) Supply chain integration and shareholder value: Evidence from consortium based industry exchanges. *Journal of Operations Management,* 26(1), 96-114.

Modi, S.B., Wiles, M.A., & Mishra, S. (2015) Shareholder value implications of service failures in triads: The case of customer information security breaches. *Journal of Operations Management*, 35, 21-39.

Montabon, F., Sroufe, R., & Narasimhan, R. (2007) An examination of corporate reporting, environmental management practices and firm performance. *Journal of Operations Management*, 25(5), 998-1014.

Ni, J.Z., Flynn, B.B., & Jacobs, F.R. (2014) Impact of product recall announcements on retailers′ financial value. *International Journal of Production Economics*, 153, 309-322.

Nicolae, M., Arıkan, M., Deshpande, V., & Ferguson, M. (2016) Do bags fly free? An empirical analysis of the operational implications of airline baggage fees. *Management Science*, 63(10), 3187-3206.

Nicolau, J.L., & Sellers, R. (2002) The stock market's reaction to quality certification: Empirical evidence from Spain. *European Journal of Operational Research*, 142(3), 632-641.

Patell, J.M. (1976) Corporate forecasts of earnings per share and stock price behavior: Empirical test. *Journal of Accounting Research*, 14(2), 246-276.

Paulraj, A., & Jong, P. (2011) The effect of ISO 14001 certification announcements on stock performance. *International Journal of Operations and Production Management*, 31(7), 765-788.

Penman, S.H. (1982) Insider trading and the dissemination of firms' forecast information. *Journal of Business*, 55(4), 479-503.

Ramdas, K., Williams, J., & Lipson, M. (2013) Can financial markets inform operational improvement efforts? Evidence from the airline industry. *Manufacturing & Service Operations Management*, 15(3), 405-422.

Rosenbaum, P.R., & Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

Sabherwal, R., & Sabherwal, S. (2005) Knowledge management using information technology: Determinants of short-term impact on firm value. *Decision Sciences*, 36(4), 531-567.

Scholes, M., & Williams, J. (1977) Estimating betas from nonsynchronous data. *Journal of Financial Economics*, 5(3), 309-327.

Shafer, S.M., & Moeller, S.B. (2012) The effects of Six Sigma on corporate performance: An empirical investigation. *Journal of Operations Management*, 30(7), 521-532.

Siomkos, G., Kurzbard, G. (1994) The Hidden Crisis in Product-harm Crisis Management. *European Journal of Marketing*, 28(2), 30-41,

Sorescu, A., Warren, N.L., & Ertekin, L. (2017) Event study methodology in the marketing literature: An overview. *Journal of the Academy of Marketing Science*, 45(2), 186-207.

Srivastava, S.K. (2007) Green supply chain management: A state-of-the-art literature review. *International Journal of Management Reviews*, 9(1), 53-80.

Tang, A.K.Y., Lai, K.-h., & Cheng, T.C.E. (2016) A multi-research-method approach to studying environmental sustainability in retail operations. *International Journal of Production Economics*, 171, 394-404.

Thirumalai, S., & Sinha, K.K. (2011) Product recalls in the medical device industry: An empirical exploration of the sources and financial consequences. *Management Science*, 57(2), 376-392.

Travlos, N. G., Trigeorgis, L., & Vafeas, N. (2001). Shareholder wealth effects of dividend policy changes in an emerging stock market: The case of Cyprus. *Multinational Finance Journal*, 5, 87-112.

Wood, L.C., Wang, J.X., Olesen, K., & Reiners, T. (2017) The effect of slack, diversification, and time to recall on stock market reaction to toy recalls. *International Journal of Production Economics*, 193, 244-258.

Xia, Y., Singhal, V.R., & Zhang, P.G. (2016) Product design awards and the market value of the firm. *Production and Operations Management*, 25(6), 1038-1055.

Yang, J., Lu, W., & Zhou, C. (2014) The immediate impact of purchasing/sales contract announcements on the market value of firms: An empirical study in China. *International Journal of Production Economics*, 156, 169-179.

Zhao, X., Li, Y., & Flynn, B.B. (2013) The financial impact of product recall announcements in China. *International Journal of Production Economics*, 142(1), 115-123.

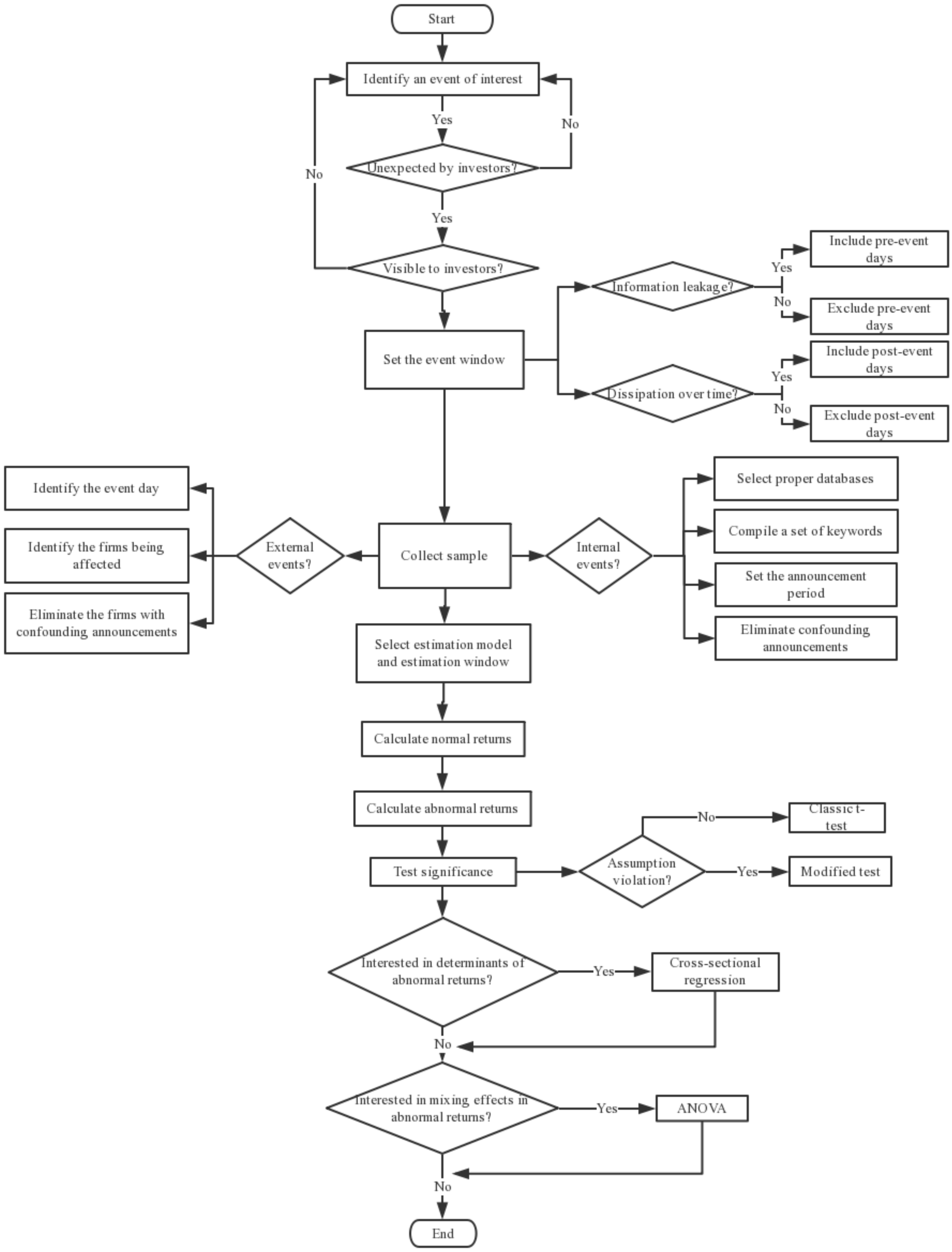

Figure 1 Steps of conducting a short-term event study

Table 1 Previous literature reviews of the event study method

Discipline	Literature review	Articles	Time range	Source	Content description
	MacKinlay (1997)	N.A.	N.A.	N.A.	1. Procedure for conducting an event study; 2. Measuring the expected returns; 3. Making statistical inferences; 4. Analysis of the power of an event study; 5. Nonparametric approaches; 6. Cross-sectional regression approach; 7. Further issues relating to event study design.
	Binder (1998)	N.A.	N.A.	N.A.	1. Hypothesis testing; 2. Different benchmarks for the normal rate of return; 3. The power of the methodology in different applications; 4. The modeling of abnormal returns as coefficients in a regression framework.
Accounting and finance	Corrado (2011)	N.A.	N.A.	N.A.	1. Outlines the econometric skeleton of an event study; 2. A survey of results obtained from studies of event study methodology; 3. Problem of event-induced variance and attempts to cope with the problem.
	Kothari and Warner (2007)	565	1974-2005	*Journal of Business*, *Journal of Finance*, *Journal of Financial Economics*, *Journal of Financial and Quantitative Analysis*, *Review of Financial Studies*	1. Describe the changes in event study methodology over time; 2. Procedures for conducting an event study, properties of the event study test; 3. Critical issues of conducting long-horizon event studies.
Management	McWilliams and Siegel (1997)	29	1986-1995	*Academy of Management Journal*, *Strategic Management Journal*, *Journal of Management*	1. Assumptions and research design issues of event studies in Management literature; 2. Replications of previous event studies.
Information systems	Konchitchki and O'Leary (2011)	over 50	N.A.	N.A.	1. A survey of research that uses event study methodologies; 2. Key parameters and concerns associated with implementation of event studies; 3. Remarks on key event study modeling issues and recommendations to researchers.
Marketing	Sorescu et al. (2017)	over 40	2000-2015	*Journal of Marketing Research*, *Journal of Marketing*, *Journal of the Academy of Marketing Science*, *Marketing Science*	1. Theoretical foundations and research design of event studies used in the marketing literature; 2. Interpretation of event studies; 3. Event study implementations and alternative methods; 4. Guidelines for future research.

Table 2 Publication journals and years of short-term event studies in OSCM

Classification	No. of papers	Event studies
Panel A: Publication Journal		
JOM	8	Brandon-Jones et al. (2017), Hendricks and Singhal (2003), Hendricks et al. (1995), Hendricks et al. (2009), Jacobs and Singhal, (2017), Jacobs et al. (2010), Mitra and Singhal (2008), Modi et al. (2015)
IJPE	7	Lam et al. (2016), Lin and Su (2013), McGuire and Dilts (2008), Ni et al. (2014), Wood et al. (2017), Yang et al. (2014), Zhao et al. (2013)
MS	6	Girotra et al. (2007), Hendricks and Singhal (1996), Hendricks and Singhal (1997), Kalaignanam et al. (2013), Klassen and McLaughlin (1996), Thirumalai and Sinha (2011)
POM	4	Ba et al. (2013), Jacobs and Singhal (2014), Jacobs (2014), Xia et al. (2016)
IJOPM	2	Dam and Petkova (2014), Paulraj and Jong (2011)
DS	1	Sabherwal and Sabherwal (2005)
EJOR	1	Nicolau and Sellers (2002)
Panel B: Publication year		
1995-1999	4	
2000-2004	2	
2005-2009	5	
2010-2014	12	
2015-2017	6	
Total	29	

DS = Decision Sciences, *EJOR = European Journal of Operational Research*, *IJOPM = International Journal of Operations and Production Management*, IJPE = *International Journal of Production Economics*, *JOM = Journal of Operations Management*, *MS = Management Science*, *POM = Production and Operations Management.*

Table 3 Summary of short-term event studies in OSCM

Event Study	Journal	Topic	Event Type	Event	Event Period	Data Source	Confounding Announcements
Ba et al. (2013)	*POM*	Environmental management	Internal	Environment initiatives and innovation (Green Vehicle Innovation)	1996-2009	Factiva	Adjacent announcements in (-2, +2)
Brandon-Jones et al. (2017)	*JOM*	Sourcing strategy	Internal	Reshoring	2006-2015	Factiva, Google News, the website of the Reshoring Initiative	Any announcements released on the prior trading day after stock market closure or on the announcement date itself
Dam and Petkova (2014)	*IJOPM*	Environmental management	Internal	Environmental supply chain sustainability program	2005-2011	BW, Google	Not reported
Girotra et al. (2007)	*MS*	R&D projects	Internal	R&D projects	1994-2004	R&D Insight database developed by ADIS international (the pharmaceutical industry)	Not reported
Hendricks and Singhal (2003)	*JOM*	Supply chain disruptions	Internal	Supply chain glitches	1989-2000	*WSJ*, DJNS	Earnings pre-announcements where supply chain glitches were mentioned as one of the many factors affecting earnings expectations
Hendricks and Singhal (1996)	*MS*	quality management	Internal	Quality award	1985-1991	TRND, DJNS	Any other announcements in (-2, +2)
Hendricks and Singhal (1997)	*MS*	Supply chain disruptions	Internal	Product introduction delay	1984-1991	TRND, DJNS	Not reported
Hendricks et al. (1995)	*JOM*	Capacity expansion	Internal	Capacity expansion	1979-1990	TRND, WSJ, PR Newswire	Earnings or any other types of announcements (dividends, change in CEO, product recalls, product delays, lawsuits, new product introductions, etc.) made in (-2, +2)

Hendricks et al. (2009)	*JOM*	Supply chain disruptions	Internal	Supply chain disruptions	1989-1998	*WSJ*, DJNS	Announcements that mention the supply chain disruption as one of many issues
Jacobs and Singhal (2014)	*POM*	R&D projects	Internal	Product development restructuring	2002-2011	DJNS, *WSJ*	Not reported
Jacobs and Singhal (2017)	*JOM*	Supply chain disruptions	External	Catastrophic disaster	N.A.	N.A.	Any announcements over the event window
Jacobs et al. (2010)	*JOM*	Environmental management	Internal	Environmental initiatives and environmental awards	2004-2006	BW, *Chicago Tribune, Denver Post*, *Dow Jones Business News*, *Financial Times*, *Houston Chronicle*, *Los Angeles Times*, *New York Daily News*, *Philadelphia Inquirer*, PR Newswire (US) *The New York Times*, *WSJ*, *USA Today*, *Washington Post*	Any other announcements in (-2, +2)
Jacobs (2014)	*POM*	Environmental management	Internal	Voluntary emissions reduction	1990-2009	*WSJ*, PR Newswire, BW, DJNS	Multiple VER announcements for the same firm within 20 trading days; VER announcements that also contain earnings or other material information
Kalaignanam et al. (2013)	*MS*	Sourcing strategy	Internal	CRM outsourcing	1996-2006	LexisNexis, Factiva, ACSI	Not reported
Klassen and McLaughlin (1996)	*MS*	Environmental management	Internal	Environmental management	1985-1991	Nexis	Financial and management announcements identified from the NEXIS financial database in (-1, +1)
Lam et al. (2016)	*IJPE*	Environmental management	Internal	Environmental initiatives	2005-2014	WiseNews (Shanghai Securities News, China Securities Journal, and Secutimes)	Announcements such as key executive appointments and annual earnings announcements
Lin and Su (2013)	*IJPE*	Quality management	Internal	Quality award	1991-2009	N.A.	Not reported
McGuire and Dilts (2008)	*IJPE*	Quality management	Internal	ISO9000	1999-2002	BW, Dow Jones Business News, DJNS, PR News, PR Newswire, Reuters News, *WSJ*	Announcements with more than one article in the Wall Street Journal in (-2, +2)

Mitra and Singhal (2008)	*JOM*	Supply chain integration	Internal	Supply chain integration	2000-2001	*WSJ*, Dow Jones Newswire, BW, PR NewsWire	Not reported
Modi et al. (2015)	*JOM*	Supply chain disruptions	Internal	Service failure	2005-2010	Identity Theft Resource Center (ITRC) (report information security breaches), Factiva	A quarterly earnings release, a merger/acquisition, a change of a CEO or CFO, a debt restructuring, or an unexpected dividend change within two trading days of the event date
Ni et al. (2014)	*IJPE*	Supply chain disruptions	Internal	Product recall	2000-2009	Consumer Product Safety Commission (CPSC)	Not reported
Nicolau and Sellers (2002)	*EJOR*	Quality management	Internal	ISO9000	1993-1999	Baratz (cover Spanish newspapers)	News items within whose windows a public offer of stock acquisition, a take-over or any large purchases of shares were announced
Paulraj and Jong (2011)	*IJOPM*	Environmental management	Internal	ISO14001	1996-2008	BW, PR NewsWire	Potentially newsworthy announcements, such as dividend declarations and earnings announcements in (-5, +5)
Sabherwal and Sabherwal (2005)	*DS*	IT	Internal	IT governance (IT-based knowledge management efforts)	1995-2002	LexisNexis (BW, PR Newswire, *The New York Times*, *The San Francisco Chronicle*, *USA Today*)	Earnings, dividends, merger, acquisition, divestiture, or change in top management announcements in (-2, +2)
Thirumalai and Sinha (2011)	*MS*	Supply chain disruptions	Internal	Product recall	2002-2005	FDA, Lexis-Nexis, Google News archives	Not reported
Wood et al. (2017)	*IJPE*	Supply chain disruptions	Internal	Product recall	1979-2016	Consumer Product Safety Commission (CPSC) database, Factiva	Not reported
Xia et al. (2016)	*POM*	R&D projects	Internal	Product design awards	1998-2011	Factiva, LexisNexis	Not reported
Yang et al. (2014)	*IJPE*	Purchasing/sales contract	Internal	Purchasing/sales contract	2001-2012	Shanghai SE website, Shenzhen SE website	Not reported
Zhao et al. (2013)	*IJPE*	Supply chain disruptions	Internal	Product recall	2002-2011	China Infobank database, Chinese automobile recall website	Not reported

WSJ = *The Wall Street Journal*, DJNS = Dow Jones News Service, TRND = Trade and Industry Index, BW = Business Wire

Table 4 Summary of tests for significance of abnormal returns

Parametric test	Studies	Sample size	Estimation windows (day)	Event windows (day)	Model for estimation	Nonparametric test
Panel A: Traditional t-test						
	Ba et al. (2013)	261	(-259, -10)	(-1, +1)	Market Model	Wilcoxon signed-rank test, binomial sign test
	Jacobs et al. (2010)	780	(-210, -11)	(-1, 0)	Market model	Wilcoxon signed-rank test, binomial sign test
	Jacobs (2014)	450	(-210, -11)	(-1, 0)	Market Model	Wilcoxon signed-rank test, binomial sign test
	Hendricks and Singhal (2003)	519	(-210, -11)	(-1, 0)	Market model, market adjusted model, mean adjusted model	Wilcoxon signed-rank test, binomial sign test
	Hendricks and Singhal (1997)	101	(-210, -11)	(-1, 0)	Market model	N.A.
	Hendricks et al. (1995)	128	(-214, -15)	(-1, +1)	Market model	N.A.
	Hendricks et al. (2009)	307	200-day	(-1, 0)	Market model	N.A.
t test	Jacobs and Singhal (2014)	165	(-210, -11)	(-1, 0)	Market Model	Wilcoxon signed-rank test, binomial sign test
	Klassen and McLaughlin (1996)	162	(-209, -10)	(-1, +1)	Market model	Wilcoxon signed-rank test, binomial sign test
	Lam et al. (2016)	556	200-day	(-1, +1)	Market model	Wilcoxon signed-rank test
	Lin and Su (2013)	20	(-210, -11)	(-1, +10)	Market model, market adjusted model, mean adjusted model	Wilcoxon signed-rank test, binomial sign test
	McGuire and Dilts (2008)	204	(-210, -11)	(-1, +1)	Market model, market adjusted model, mean adjusted model	Wilcoxon signed-rank test, binomial sign test
	Paulraj and Jong (2011)	140	(-261, -10)	(-1, +1)	Market model, mean adjusted model, market adjusted model	Wilcoxon signed test, generalized sign test, rank test
	Xia et al. (2016)	264	(-220,-21)	(-1, 0)	Market model	Wilcoxon signed-rank test, binomial sign test
	Yang et al. (2014)	318	N.A.	2-day	N.A.	N.A.
z test	Dam and Petkova (2014)	66	(-110, -11)	0	Market model	N.A.

Parametric test	Studies	Sample size	Estimation windows (day)	Event windows (day)	Model for estimation	Nonparametric test
Panel B: Modifications to the traditional t-test						
Brown and Warner (1985) t-test	Hendricks and Singhal (1996)	91	(-210, -11)	0	Market model, market adjusted model, mean adjusted model	Wilcoxon signed-rank test, binomial sign test
	Mitra and Singhal (2008)	144	200-day	(-1, 0)	Market model, mean adjusted model	Wilcoxon signed-rank test, binomial sign test
	Jacobs and Singhal (2017)	39	200-day	(0, +10)	Market model	Wilcoxon signed-rank test, binomial sign test
Time-series standard deviation test, portfolio *t*-test	Modi et al. (2015)	146	255-day	(-1, +1), (-2, +2)	Fama-French four-factor model	Generalized sign test, Wilcoxon signed rank test
Jaffe test	Nicolau and Sellers (2002)	27	147-day	(-3, +3)	Market model	Corrado rank test
Patell *Z* test	Zhao et al. (2013)	42	(-130, -11)	(0, +1), (-5, +1)	Market model	Wilcoxon signed-rank test, binomial sign test
t-test, Patell *Z*-test, standardized cross-sectional *t*-test	Ni et al. (2014)	164	(-270, -21)	(-1, 0)	Market model, market adjusted model, mean adjusted model, size-and-industry adjusted model	Wilcoxon-Mann-Whitney test
Cross-sectional standard deviation test, standardized Patell *Z* test, crude dependence adjustment test	Girotra et al. (2007)	132	(-255, -10)	(-2, +4), (-3, +3), (-4, +4)	Comparison period model, market model, Fama-French three-factor model	Generalized sign-z test, Wilcoxon signed rank test
Cross-sectional *t*-test, Patell *Z* test, BMP *t*-test	Wood et al. (2017)	135	(-131, -11)	(0, 1)	Market model	Wilcoxon signed ranks test, sign test
Cross-sectional variance-adjusted Patell test	Kalaignanam et al. (2013)	158	(-260, -30)	(0, +1), (0, +2), (-2, 0), (-1, +2), (-2, +1), (-2, +2)	Fama-French four-factor model	N.A.
Heteroscedasticity consistent standard errors *t*-test	Sabherwal and Sabherwal (2005)	89	(-300, -46)	(-2, +2), (-3, +3)	Market model	N.A.
Patell test	Thirumalai and Sinha (2011)	223	120-day	(0, +1), (-1, 0), (-1, +1), (-5, +5), (-10, +1), (-10, +10)	Market model	Binomial sign test
Patell test, standardized cross-sectional test	Brandon-Jones et al. (2017)	37	(-210, -11)	0	Market model, market adjusted model, mean adjusted model	Rank test, generalized sign test

Table 5 Comparison of parametric tests for significance of abnormal returns

Significance Test	Key Reference	Key Assumptions	Strength	Weakness	Representative Studies
Traditional t-test	MacKinlay (1997)	Cross-sectional independence of abnormal returns; Event-induced variance is insignificant; homoscedasticity of abnormal returns	Simplicity	Prone to cross-sectional correlation; Prone to event-induced volatility; Prone to heteroskedasticity among observations	Hendricks and Singhal (2003), Jacobs et al. (2010), Ba et al. (2013), Xia et al. (2016)
Crude dependence adjustment test	Brown and Warner (1985)	Homoscedasticity of abnormal returns	Allow for cross-sectional correlation	Prone to heteroskedasticity among observations, less powerful	Hendricks and Singhal. (1996), Girotra et al. (2007), Mitra and Singhal (2008), Jacobs and Singhal (2017)
Cross-sectional test	Penman (1982)	Cross-sectional independence of abnormal returns	Allow for event-induced volatility; Allow for serial correlation	Prone to cross-sectional correlation	Wood et al. (2017)
Standardized residual test	Patell (1976)	Cross-sectional independence of abnormal returns; Event-induced variance is insignificant	Allow for the heteroskedasticity among abnormal returns over the event period	Prone to cross-sectional correlation and event-induced volatility	Girotra et al. (2007), Thirumalai and Sinha (2011), Kalaignanam et al. (2013), Zhao et al. (2013), Ni et al. (2014), Wood et al. (2017), Brandon-Jones et al. (2017)
Standardized cross-sectional test	Boehmer et al. (1991)	Cross-sectional independence of abnormal returns	Allow for the heteroskedasticity among abnormal returns over the event period; Allow for event-induced volatility; Allow for serial correlation	Prone to cross-sectional correlation	Kalaignanam et al. (2013), Wood et al. (2017), Brandon-Jones et al. (2017)