It’s Your Responsibility: The Impact of Supply Chain CSR Performance on Firm Value

by

Zachary S. Rogers

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2016 by the
Graduate Supervisory Committee:

Craig Carter, Chair
Kevin Dooley
Vinod Singhal

ARIZONA STATE UNIVERSITY
August 2016
Firms are increasingly being held accountable for the unsustainable actions of their suppliers. Stakeholders, regulatory agencies, and customers alike are calling for increased levels of transparency and higher standards of corporate social responsibility (CSR) performance for suppliers. While it is apparent that supplier performance is important, it remains unclear how the stock market weighs the CSR performance of a supplier relative to that of a focal firm. This dissertation focuses on whether these relative differences exist. In addition to capturing the magnitude of the difference in market impact between focal firm and supplier CSR events; I analyze the ways in which these differences have changed over time. To capture this evolution, CSR events ranging over a period from 1994 to 2013 are examined. This research utilizes an event study methodology in which the announcement of over 2,300 CSR events are identified and analyzed to determine the subsequent stock market reaction. I find that while the market evaluated negative supplier CSR events less harshly than events occurring at the buying firm in the early years of the sample, by the turn of the millennium this “supplier discounting” had disappeared. The analysis is broken down by CSR event "type". Findings demonstrate that negative CSR events, particularly those revolving around worker or customer safety, generate the most significant abnormal return. The findings of this dissertation produce valuable managerial insights along with interpretation. Resources are scarce, and understanding where a firm might best allocate their resources to avoid financial penalties will be valuable information for corporate decision makers. These findings present clear evidence that some of these resources should be allocated to supplier CSR performance, not just towards the CSR performance of the focal firm.
DEDICATION

This dissertation is dedicated to two people. First I need to acknowledge my daughter Aurora Beverly Rogers. Aurora was only 5 months old when we uprooted our lives to come down to Arizona. I know that she won’t remember everything that happened in Arizona, but I will certainly remember how happy I have been to have her here with us the whole time. From a young age Aurora has somehow had a sense of when her father was depressed or overwhelmed by all there was to do. More times than I could count, Aurora has come over with a giggle, an explanation of something she’s had on her mind, or a “big squishy hug” and pulled me out of whatever despair I was in. Aurora, it has been a pleasure watching you grow up in Arizona. I only have to look over to you to remember why all of this work has been worth it.

Finally, I also dedicate this to my best friend and the love of my life, Ándrea. I can say with absolute certainty that none of this would have been possible without your love and support. Your kindness, your grace, and your seemingly boundless supply of patience have been a source of comfort and refuge for me as I trudged through what has undoubtedly been the most difficult period of my life. I got through it in no small part because I knew that at the end of every day you would be there when I came home. I could fill the rest of this dissertation with what you have meant to me over the last eight years. But instead I’ll just say this: You the real MVP.

Love,

Zachary
ACKNOWLEDGMENTS

There are many people who contributed to this dissertation and my growth as a scholar over the last four years, but the first person I should thank is Dr. Craig Carter. His wisdom, and more importantly his friendship, has been invaluable over my time at Arizona State. I’m sure I wasn’t the easiest doctoral student, especially when I first arrived, but Dr. Carter’s was unfailingly patient and always believed in me (even when I kind of didn’t). I’m a better scholar and better person for the time I’ve spent with him. He has a seemingly limitless knowledge of the literature and a deep understanding of how to come up with and develop meaningful research ideas. I’ve been in his office at least 100 times, and I’ve never left without learning something (and generally not without something new to read). I can’t imagine having a better mentor over these last four years.

I would be remiss to not also acknowledge the contributions of the other members of my dissertation committee: Dr. Kevin Dooley and Dr. Vinod Singhal. I was very blessed to have scholars of this quality on my committee. As subject matter experts, you can do no better in our field than Dr. Dooley for sustainability and Dr. Singhal for the event study methodology. I am incredibly lucky to count both of them as advisors on this project. This dissertation would not be nearly as strong without the contributions from all three members of the committee.

I should also acknowledge the Supply Chain faculty at Arizona State, who I (perhaps somewhat subjectively) believe is the best on the planet. My abilities as a scholar were undeniably shaped from my time spent in the classrooms of Dr. Choi, Dr. Kull, Dr. Rabinovich, Dr. Webster, Dr. Yin and our Chairman Dr. Fowler. Along with Dr. Carter, Eddie Davila is the guy who taught me how to be a teacher both inside and
outside of the classroom. Though I never took classes from them, I received nothing but positive interactions and good advice from the rest of the faculty and the admin team.

There is one other professor who has had an undeniable influence on me in my time at ASU. Most of the world knows him as Dr. Dale Rogers, but I usually just call him Dad. It took going through the last four years to truly appreciate what he was really up to all this time. Thanks Dad, I hope I can do it as well as you. I also need to acknowledge my mother Deana. She is full of joy and hope, and I have always been able to count on her to cheer me up when everything seems a little overwhelming. I can’t imagine what the last two years would have been like without the sage advice, the unwavering support, and the thousands of dollars’ worth of free babysitting and On The Border. I also need to acknowledge my siblings - Ben, Tre, Sam & Jaesa - each of whom I have come to for advice or a welcome distraction from homework.

At this point my Ph.D. cohort is my family in all but DNA. Eunae Yoo, Sining Song, Sina Golara, Jerry Huff, Min Choi, Sophie Yang, Annibal Sodero, Feng Cheng, Kang Hsu, Piyush Shah, Mehdi Saleh, Scott Cowley and Irfan Kanat, have all been good friends and trusted confidants during my time here. Of course I must also acknowledge the three gentlemen who came into the program with me, my good friends and boon companions: Yousef Abdulsalam, Sangho Chae, and Zhongzhi “The Macho Man” Liu. I know that I could count on these three guys for anything; they have literally helped me move, up and down multi-floor walkups – which is widely acknowledged as the highest level of friendship.

Finally, I must acknowledge God – In whom I can do all things (Holy Bible., 2011).
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
</tbody>
</table>

## CHAPTER

1. INTRODUCTION ................................................................. 1

2. RESEARCH MOTIVATIONS & LITERATURE REVIEW ...................... 10
   - Corporate Social Responsibility ........................................ 10
   - Event Studies in Corporate Social Responsibility ................. 13
   - The Irrational Side of Rationality ..................................... 16

3. DEVELOPMENT OF HYPOTHESES .......................................... 20
   - Hierarchy of CSR Event Types .......................................... 30

4. METHODS ............................................................................ 39

5. RESULTS ............................................................................. 45
   - Internal vs. External Events .............................................. 45
     - Full Sample ..................................................................... 45
     - Negative Sample ............................................................. 47
   - Comparisons between Event Types ...................................... 55
   - Controlling for Endogeneity .............................................. 61

6. DISCUSSION ....................................................................... 63
   - Discussion of Results ...................................................... 63
   - Managerial Implications ................................................... 71
   - Contribution To Academic Literature .................................. 75
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Potential Limitations</td>
<td>77</td>
</tr>
<tr>
<td>Future Research</td>
<td>79</td>
</tr>
<tr>
<td>Future Environmental Event Study</td>
<td>79</td>
</tr>
<tr>
<td>Hierarchy of Environmental Event Types</td>
<td>86</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>95</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
</tr>
<tr>
<td>A SEARCH TERMS</td>
<td>117</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Table Events by Step &amp; Type</td>
<td>41</td>
</tr>
<tr>
<td>2. Selected Event Headlines</td>
<td>41</td>
</tr>
<tr>
<td>3. Full Sample Regression</td>
<td>46</td>
</tr>
<tr>
<td>4. Event Breakdown</td>
<td>48</td>
</tr>
<tr>
<td>5. Negative &amp; Positive Events Descriptive Statistics</td>
<td>49</td>
</tr>
<tr>
<td>6. Negative Event Sample Regression</td>
<td>50</td>
</tr>
<tr>
<td>7. Negative Events Internal vs. External by Period</td>
<td>51</td>
</tr>
<tr>
<td>8. Full Sample Comparisons between Event Types</td>
<td>56</td>
</tr>
<tr>
<td>9. Negative Sample Comparisons between Event Types</td>
<td>58</td>
</tr>
<tr>
<td>10. Mean CAR value by Negative Event Type</td>
<td>60</td>
</tr>
<tr>
<td>11. Hypotheses Results by Sample</td>
<td>65</td>
</tr>
<tr>
<td>12. Positive vs. Negative Waste Management Event Comparison</td>
<td>92</td>
</tr>
<tr>
<td>13. Negative Sample Comparisons: Waste management &amp; CSR Events</td>
<td>93</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>Social Event Types</td>
</tr>
<tr>
<td>2.</td>
<td>Effect of Positive Supplier Development Events</td>
</tr>
<tr>
<td>3.</td>
<td>Average CAR’s For Negative Events</td>
</tr>
<tr>
<td>4.</td>
<td>Environmental Event Types</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Corporate Social Responsibility (CSR) events can have significant impacts on firms, influencing their financial performance and public perception (Story, 2007; Dwyer, 2010). Environmental and social violations committed by firms or their suppliers can result in significant financial losses for the focal firm (Carter & Jennings, 2004). These violations also erode the public perception of a firm’s reputation (Rogers, 2011). Consumers are now more likely to hold buying firms responsible for the unsustainable actions of their suppliers and penalize them for missteps that occur thousands of miles away (Hartmann & Moeller, 2014). Firms have become more serious about improving the CSR performance of their upstream suppliers and realize they will not be considered truly sustainable if those suppliers are not exhibiting sustainable behavior (Zhu & Sarkis, 2004; Ciliberti et al., 2008).

The penalties that firms suffer from CSR violations in their supply chain are both financial and reputational. The 2013 collapse of the Rana Plaza garment factory in Bangladesh is an example of severe financial and reputational penalties. Large US firms such as Wal-Mart, Target, and Macy’s have had to commit over $140 million to correct the issues that led to the collapse with potentially much more to follow (Greenhouse & Clifford, 2013; Dhaka, 2014). Beyond the financial penalties, firms also suffer from a blow to their reputation and brand image (Hartmann & Moeller, 2014). Mattel’s Barbie has still not returned to the prominent position it enjoyed for many years prior to the scandal they faced in 2007. This is partially due to the reputational damage they suffered
as a result of one of their suppliers using lead-paint on children’s toys (Story, 2007; Khouri, 2014).

One of the primary drivers behind increased interest in supplier CSR is the financial consequences associated with failing to achieve acceptable levels of CSR performance. Poor CSR performance can result in “front-page” events, marring a firm and its suppliers. Understanding the costs associated with this is a primary issue in the CSR and sustainability literature. Studies have found strong ties between CSR initiatives and financial performance (Zhu & Sarkis, 2004; Callan & Thomas, 2009). In their meta-analysis of sustainable supply chain research, Golicic and Smith (2013) find a strong positive link between supply chain sustainability and financial performance. Other research has found the opposite result (Hahn, Figge, Pinske, & Preuss, 2010). Still others focus on the potential cost of the headline risk events that come as a result of neglecting sustainability performance until it becomes an issue (Rogers, 2011). However, almost all of these studies rely on long-term financial statistics retrieved from 10K’s or quarterly reports. While these studies are interesting and very useful in the study of the financial impact of more long-term events, they may not sufficiently capture the immediate financial impact of certain types of CSR events within the supply chain. For evaluating this more immediate impact, epitomized by Rogers’ (2011) concept of headline risk events, we turn to the event study methodology. This method has been expansively used in finance and accounting. More recently it has been utilized in the marketing and management literature. It is an attractive methodology since stock prices describe the value of firms, and are not as easily manipulated as other financial statistics such as sales or revenues (McWilliams & Siegel, 1997). Stock prices are updated continuously, with
new information quickly being incorporated such that they are highly reflective of a firm’s perceived value. This continuous assessment holds an advantage over financial statistics that are presented on quarterly reports or 10-K’s as they will more accurately reflect the impact of an event on the value of a firm (McWilliams & Siegel, 1997).

This dissertation makes an original contribution by comparing the differences between events that occur within the four walls of a firm against those that occur within a firm’s supply chain using the event study methodology. Although consumers tend to hold firms responsible for the unsustainable actions of their suppliers (Hartmann & Moeller, 2014), it is not clear to what extent the stock market responds to negative (or positive) CSR events. More precisely, it is not clear if the stock market will penalize firms for the misdeeds (or accomplishments) of their suppliers as well as their own. The theory of psychological distance (Trope & Liberman, 2010) would suggest that they may not.

Extant event studies have been used to measure the impact of CSR events on stock prices. Jayachandran et al. (2013) performed an event study analyzing the impact of both positive and negative “product social performance” and “environmental social performance” events. This study is a logical extension of that work in many ways. Jayachandran and his colleagues used these as blanket terms, lumping the different types of social and environmental events into single categories. This research will differ in that we are taking a more granular approach to our definitions of social responsibility events, as well as extending the study horizon over a period of 20 years. There have been multiple definitions of social responsibility (Carroll, 1994; Garriga & Melé, 2004), with some scholars pointing out that social responsibility can mean different things to different
companies (Votaw, 1972). For instance, a social event involving too few diversity hires may not have the same impact on stock prices as an event involving worker safety issues. Determining if these differences exist, and how they have changed over time, should prove to be a valuable contribution to both industry and the academic literature. In this research we base our division of social events on Carter and Jenning’s (2004) framework, dividing social events into four categories: worker safety, labor conditions, diversity, and community outreach. We will rely on decision making heuristics from the behavioral finance literature to understand the differences in stock price reaction to these granular events.

This type of granular analysis is valuable because it is not always possible to achieve “win-win” scenarios in which improving societal performance is financially beneficial (Voinov & Farley, 2007; Hahn et al., 2010). Seuring & Muller’s (2008) Delphi study reveals that many practitioners see tradeoffs as a necessary evil when developing their firms’ CSR performance. Often, due to resource or oversight limitations, firms must make tradeoffs between the different elements of CSR in their supply chain. Understanding the different ways in which the market penalizes or rewards different types of actions will be useful as they consider the tradeoffs they must make. Given the proliferation of fines and public headline risk events related to sustainability in the last twenty years, it is likely that the market treats these tradeoffs differently today than it did in 1994.
In summary, the overriding research questions of this study are:

*Does the market distinguish between CSR events occurring at the focal firm and events occurring at a supplier? Does the market evaluate alternate, distinct types of CSR events differently from one another? If so, what is the magnitude of these differences? Finally, how have the differences in market impact between all event types studied in this research changed over the last 20 years?*

This study contributes to the academic literature in both CSR event studies and buyer-supplier relationships. It is unlikely that a firm will have the resources or oversight capabilities to ensure high levels of every aspect of CSR performance throughout their supply chain. Inevitably, they will be forced to make tradeoffs. This research helps firms with these decisions in the following ways. First, it provides a time-series comparison of the market’s reaction to CSR events occurring at the focal firm relative to events occurring further up the supply chain. Although both the literature (Carter & Rogers, 2008) and consumers (Hartmann & Mueller, 2014) suggest that the importance of supply chain CSR performance has increased, no studies have taken an in-depth, time-series approach to validate this notion using stock prices. It remains unclear whether or not supply chain CSR performance has the same impact as focal firm performance, as well as how this has changed over time. Secondly, we analyze the relative difference in market costs and/or benefits between different types of CSR performance. The combination of these findings will provide firms with a resource to use as they make decisions on which parts of supply chain CSR performance they will allocate their scarce resources to. This research overcomes a limitation of previous CSR event studies by moving away from the assumption that all social and environmental events should be grouped together. By
conducting a granular analysis on the stock impact of different event types we address a need in the existing CSR event study literature.

The research methodology relies on stock price changes spanning a period from January 1st, 1994 to December 31st, 2013 for publicly traded firms that are listed on U.S. stock exchanges. This range of dates was chosen because at the time this project was started, information from after 2013 was not available using the Eventus program through which we calculate abnormal returns. We wished to observe market differences over at least a 20 year period so 1994 was selected as the starting point. Other event studies have been far shorter (Jayachandran et al., 2013), or substantially longer. Flammer’s (2013) study of environmental events over time ranged from 1980 to 2009. Future studies could incorporate similarly earlier dates and extend the terminus point to a later date. For the purposes of this study, the 20 year range was sufficient to show significant movement in market preferences and reactions. Using historical prices to construct a model, stock prices are predicted for a firm on the day a CSR event occurs. Prices that differ significantly from what the model predicts suggest that the fluctuations may be due to the occurrence of the event. Event studies have been used successfully to measure the impacts of different parts of CSR performance (Jacobs et al., 2010; Jayachandran et al., 2013; Wang & Chen, 2015) as it allows researchers to observe the immediate impact of CSR events while dealing with minimal complications from potential confounding events. This research is in many ways a logical extension of the CSR event studies that have come before. We are however adding layers of complexity in our differentiation between the focal firm and suppliers, and between CSR event types. It is our hope that
this research will not only provide evidence towards answering the questions raised above, but will also illuminate related questions we have yet to consider.

It should be noted that this dissertation focuses specifically on social performance, eschewing any analysis of financial or environmental events. This runs somewhat contrary to the triple bottom line (TBL) conceptualization of sustainability in recent supply chain research (Elkington, 1998). This research has taken many different forms. Carter and Rogers (2008) expanded on Elkington’s (1998) triple bottom line to establish a sustainable supply chain framework, positing that supply chains cannot be truly sustainable unless they are meeting satisfactory levels of financial, social, and environmental performance. Seuring and Müller (2008) conducted a Delphi study to determine what priorities exist for businesses looking to establish a sustainable supply chain, finding that both environmental and social performance were considered to be important barometers of success. Walker and Jones utilized a similar strategy in their exploration of sustainability practices in the U.K. (2012). Pagell and Wu (2009) utilized a series of case studies, eventually echoing the triple bottom line notion when they conclude that a sustainable supply chain would “... do no net harm to natural or social systems while still producing a profit over an extended period of time,”.

While we strongly agree with the assessment that a firm’s supply chain is only truly sustainable when it has achieved sufficient levels of social, environmental and financial performance, neither environmental or financial events are considered in this research. There are multiple reasons for their exclusion. The primary reason to forgo an analysis of financial events is that the impact of financial events has already been extensively studied using the event study methodology. The first event study was
conducted to analyze the impact of stock splits on market prices (Fama et al., 1969).

Since then event studies have been used extensively to study the impact of many types of financial events. Kothari & Warner’s survey of the five largest finance journals reveals that between 1974 and 2000, 565 articles utilizing event studies to analyze financial events were published (Kothari & Warner, 2004). No matter how well done our analysis is, it is unlikely this dissertation would advance the literature by including financial events. While event studies have been used to measure the impact of social events, it is not nearly to the extent to which financial events have been studied, and the studies that have been conducted have not been in the context we will use here (which is covered in greater detail in Chapter 2). Additionally, the breadth of financial events is such that it would be difficult, if not impossible, to include them all in this research. Any conclusions drawn from such a study would include major caveats on the limitations of the sub-sample. Firms are clearly interested in the potential financial impact of their supply chain’s environmental performance (Zhu & Sarkis, 2004; Seuring and Müller, 2008) and studying it will be worthwhile. Due to the volume of events to be analyzed, it was decided that this dissertation needed to focus on either social or environmental events, as to do both would be untenable. The decision was made to focus solely on the impact of social events in the supply chain. While they have not focused on the contrast between focal firm and supplier events, extended time-series studies have recently been carried out on the market impact of environmental events (Flammer, 2013); this is not true of social events. Due to these factors, an analysis of the impact of social events on firm value was prioritized. However, to my knowledge a study comparing the financial impact of environmental events occurring at the focal firm relative to those occurring in
the supply chain does not exist. Therefore a study of these events is still needed, and a follow-up study of the impact of environmental events is currently underway. Preliminary results stemming from this follow-up study will be presented in the Future Research section in Chapter 6.

The remainder of this dissertation proceeds as follows: Chapter 2 presents the research motivations and in-depth review of the literature surrounding our topic. Chapter 3 develops the conceptual framework and lays out the hypotheses to be tested. Chapter 4 details the methodology used in this study. The data analysis and results are presented in Chapter 5. Finally, Chapter 6 presents the discussion, including the implications of our findings along with limitations and potential directions for future research.
CHAPTER 2

RESEARCH MOTIVATIONS & LITERATURE REVIEW

Corporate Social Responsibility.

The importance of a firm’s social responsibility has been brought to the forefront of corporate consciousness in the last 20 to 25 years. Porter and van der Linde (1995) give multiple examples of firms that, in some cases by accident, experienced significant financial savings due to social or environmental initiatives. This led to a cascade of research exploring the financial benefits of social and environmental performance (Price, 1995; Shrivastava, 1995; Milne, 1996). Margolis and Walsh (2003) review 127 papers written between 1972 and 2002. Of these, they find that 109 papers suggest a causal relationship between CSR (or more commonly, CSP) and financial performance. Of these, over half (57) posit a positive relationship, along with 7 reporting a negative relationship, 20 indicating mixed relationships, and 28 finding no relationship at all.

As far back as 1953 Bowen and Johnson put forth a definition of corporate social responsibility, stating that it was the responsibility of business to “follow those lines of action which are desirable in terms of the objectives and values of society, (pp. 28)”. However it wasn’t until 1972, with Bragdon and Marlin's examination of the correlation between pollution control and profitability and Moskowitz's comparison of stock prices for socially responsible firms, that the relationship between CSR and financial performance was first explored.

As mentioned above there were also a number of studies that had weak or neutral findings. Cochran and Wood find a weak correlation between CSR and financial
performance when controlling for average age of corporate assets (1984). McGuire et al. (1988) show that a firm’s previous financial performance is more closely related to CSR performance levels than their subsequent performance. Johnson finds that being socially irresponsible can have a negative impact on financial performance, but that going above and beyond legal mandates offers no financial bonuses. However, he was able to show that certain parts of CSR, such as employee treatment, did have some financial benefits (2003). Nelling and Webb (2009) were unable to find a strong causal correlation between financial and CSR performance. However they do find that there is a link between CSR and recent stock performance - a result that is indicative of the findings for this research.

However as mentioned earlier, many more findings have indicated a positive correlation between CSR and financial performance (Preston, 1978; Wokutch & Spencer, 1987; Simerly, 1995; Pava & Krausz, 1996;) Wokutch and Spencer (1987) were able to find a link between corporate giving and financial performance – with financial performance being somewhat negated by “CSR crimes”. Pava and Krausz’s (1996) investigation of the Social Investment Forum found over 500 firms that had begun to allocate investment funds based on social performance criteria. Lin et al. (2009) find that CSR can have a positive association with financial performance when R&D expenditures are properly controlled for. Zhara et al. (1993) and his coauthors display evidence of the effect of ownership and board structure on the association between CSR and financial performance.

Beyond academic research, there has also been a significant amount of anecdotal evidence that firms risk financial penalties when their (or their suppliers’) social and environmental performance is insufficient. Notable examples of this are the suicides at
Foxconn and the Rana Plaza Factory Collapse. Beyond headline risk events (Rogers, 2011) such as these, firms may also face financial pressures from customers. Large retailers such as Wal-Mart and Target have both implemented initiatives requiring suppliers to maintain certain levels of CSR performance (Buss, 2013; Gunther, 2013). These requirements are tightly enforced; suppliers can lose their contracts for failing to maintain buyer standards, even in cases where the suppliers are maintaining the regulatory standards of their country of operation (Hower, 2013). Regulation also plays a role in the adoption of CSR values. The Dodd-Frank act requires firms to fully disclose their use of conflict minerals, such as tin or gold, in their products (Anand, 2011). In response to this, firms like Intel, HP, and Apple have pressured suppliers to ensure that the minerals they supply are conflict-free (Schwartz, 2014). If suppliers cannot prove their inventory is conflict-free they are essentially closed off from western markets, and are only able to sell to less regulated regions at a discount of 30-60 percent (Baflemba, Mueller, & Lezhnev, 2014).

In their analysis of past attempts to link CSR and financial performance, McWilliams and Siegel (2000) state that there are two types of analyses: those that observe long-term effects by relying on annual and quarterly financial measures, and those that observe short term effects by studying shifts in the stock market. Many of the former have been discussed above. As mentioned earlier there have been hundreds of CSR studies in the past 60 years, making it intractable to give them all their proper due in this space. This dissertation falls under McWilliams and Siegel’s latter category, as it will utilize an event study methodology. The reasons for this are once more discussed in-depth in the section on event studies in corporate social responsibility below.
Event Studies in Corporate Social Responsibility.

This research will primarily rely on an event study methodology. Event studies were pioneered by Fama et al. (1969) in order to determine whether or not splitting securities changed the behavior of stock prices in an “unusual” amount. By establishing a model predicting the way in which stock prices move over time, the authors were able to determine the effect of splits by observing the amount in which actual stock prices differed from the predicted price on the day the split was announced (Fama et al., 1969). This methodology has been very influential, with the original paper having over 4,500 citations as of May 2016. This method operates on the assumption that, given a reasonable level of rationality exists in the stock market, the unexpected change in security prices may be attributable to the event in question (MacKinlay, 1997). This methodology has been widely used in finance (MacKinlay, 1997; Cowan & Sergeant, 1996), accounting (Duso, Gugler, & Yurtoglu, 2010), and marketing (Agrawal & Kamakura, 1995; Geyskens, Gielens, & Dekimpe, 2002; Wiles & Danielova, 2009; Chen, Ganesan, & Liu, 2009). Due to its predictive ability, it has become the primary method of measuring the effect of any type of event on security prices (Binder, 1998).

Event studies are an effective tool for measuring the impact of supplier social and environmental performance for a number of reasons. Large “headline risk” events are often pointed to as the occurrences that managers and shareholders are most trying to avoid when they push for CSR initiatives or improved supply chain sustainability (Anderson, 2006; Rogers, 2011). The event study methodology lends itself to
examinations of headline-type events as it allows for analysis of the stock market in the
days surrounding the headline-generating event. In terms of understanding the impact of
social and environmental responsibility on financial performance, this type of
examination is superior to year-over-year or quarterly analysis. Yearly and quarterly
reports can be influenced by many elements, and sorting out the confounding factors
would likely be untenable.

One of the first event studies with a focus on sustainability was Klassen and
McLaughlin’s (1996) work on environmental announcements, in which they found that
both environmental crises and performance awards have a significant impact on stock
price. Interestingly, 14 years later Jacobs et al. (2010) find opposing results, suggesting
that the announcement of environmental initiatives did not significantly impact market
valuation (although certain types of initiatives did have an effect in certain industries). A
similar phenomena was explored in Wang and Chen’s (2015) comparison of the impact
of CSR policy implementations in the U.S. and Taiwan. Bose and Pal (2012) find that the
impact of a green supply chain (GSM) announcement depends greatly on the industry and
R&D intensity of a firm. They do not specify the type of initiative, rather, the type of
firm. Jayachandran et al. (2013) compare the impact of a product’s “social” and
“environmental” performance, concluding that the social performance of a product is
much more closely associated with positive returns (with environmental performance
actually associated with negative returns).

Event studies have also been applied to the relationship between buyers and
suppliers, although in those cases CSR performance has not always been the context.
Homburg et al. (2014) used the methodology to understand the effects of channel
expansions in different industry environments. Eshleman and Guo (2014) find that supplier earnings announcements can affect their customers’ stock prices, potentially due to analysts being able to extrapolate their customers’ costs of goods sold. Event studies have also been used to analyze the impact of signals moving from the customer outwards to their suppliers. Son et al. (2015) show that the announcement of a new model of iPhone has a significant impact on Apple’s suppliers’ stock price. The market impact of outsourcing to different locations, primarily overseas versus local or nearshoring, for Fortune 500 firms is shown to be beneficial by Chakravarty et al. (2014). Wang et al. (2014) examine the impact that sustainability initiatives emanating from a buying firm have on the market evaluations of their suppliers. While this is an important step towards the use of event studies in supply chain sustainability, this piece concentrates on only the upstream effect of a single positive event. Deng et al. (2013) find that acquiring firms with higher CSR reputations achieve higher immediate and long-term returns when executing a merger, the potential ramification being that a firm’s social responsibility is a strong indicator of future success. Singhal and Jacobs (working paper) conduct a similar study in their examination of the effects of the Rana Plaza factory collapse on their customers. While this research differed in that it focused on the impact of a supplier’s performance on their customers, it was similar to Wang et al., (2014) in that it focused on a single event. The natural extension of these supply chain event studies is to expand the scope of events under consideration. Flammer’s (2013) study of environmental events over time shines a light on the path this scope expansion may follow. Flammer considers the impact of environmental events occurring from 1980 to 2009. She examines the way in which the market impact of negative and positive event types have changed over time.
As discussed above, the attention paid to supplier CSR performance by many parties has increased over time. Using an extended time horizon similar to the one used by Flammer (2013) will allow us to generate empirical data examining just how these attitudes have changed.

A gap in the literature exists in that event studies have not been used to measure granular social responsibility events within the focal firm relative to events occurring at suppliers over an extended period of time. By contrasting market reactions to these different types of events, we paint a picture of the financial tradeoffs firms must make. If a firm suffers financial consequences due to the implementation of a program to improve working conditions, how does that effect compare to the potential costs of never implementing such a program? The costs of headline risk events (Rogers, 2011) are often portrayed as binary events in which costs either happen or do not happen. In reality there are costs associated with every business decision. By studying the market impact of different types of CSR performance types we will capture some of these costs. Additionally, by relying on time-series data we will observe how the magnitudes of differences between four distinct elements of CSR performance events have changed over time.

**The Irrational Side of Rational Evaluation.**

One of the primary advantages of the event study methodology is the supposed objectivity of the stock market (Fama, 1969; Corrado, 2011). Efficient market theory espouses that the security prices will fully reflect all available information (Fama, 1969). While the stock market is likely a more objective evaluator of a company’s future
performance than other entities, even Fama (1991) concedes the possibility that other non-rational factors might be at play in the market’s evaluation of a firm. A growing body of research suggests that psychological heuristics play a significant role in firm evaluation. The behavioral finance field is in part a reaction to the notion that markets and analysts are perfectly rational. Shiller (1999) and Bazerman and Moore (2009) provide very comprehensive reviews of the different types of decision heuristics that can affect stock value evaluation.

For instance, Andreassen (1987) finds evidence that the way in which a story is reported by the media can have a significant effect on the way stock prices are evaluated by analysts. Yates (1990) introduced the concept of overconfidence to stock evaluations, showing that investors are often overconfident in their abilities to predict future stock prices, leading to an excess of trades and increased volatility. An overconfident market can fluctuate wildly, with significant swings in stock value. The findings of Chen et al. (2007) supported the effects of overconfidence in Chinese investors. Chen and his coauthors also found investors suffered from a representative bias, and that they hold on to depreciating stocks longer than they should because they do not want to perceive previous investments as sunk costs. The effects of representative bias on stock evaluation has been a common finding in this stream of literature (Bazerman & Moore, 2009). Other biases that have been found to affect stock market evaluation are inflated senses of self, mistaking randomness for steady patterns (Jordan & Kaas, 2002; Moore et al, 1999), ambiguity (Ellsberg, 1961), and as discussed in Chapter 3, the anchoring effect (Tversky & Kahneman, 1974).
Efficient market theory would lead one to assume that the stock price impacts of different types of CSR events would be entirely based on the actual impact these events will have on the firm in question. However, with so many decision heuristics potentially at play, a truly objective evaluation of the impact of these different events is very unlikely (if not impossible). Therefore we must consider the impact of different heuristics that analysts might rely on as they assign a market penalty or reward based on the focal firm’s (or their supplier’s) performance. Different psychological heuristics will tie into the construction of our hypotheses in Chapter 3, but the lenses we will most rely on are Trope and Liberman’s (2010) concept of Psychological Distance, emotional subjectivity (Loewenstein et al., 2001) and unit saliency (Whittlesea, 1990; Shiller, 2001). These lenses are used to construct our hypothesized hierarchy of market impact resulting from different event types. They also come into play as we evaluate the reasons investors may place different values on similar events occurring at the focal firm versus those occurring in their supply chain. They are discussed in-depth in Chapter 3 during the hypothesis development.

Opportunities for arbitrage can last for seconds or even, as we’ve seen recently with the emergence of high-frequency training, micro seconds (Lewis, 2014). Analysts need to move quickly if they hope to take advantage of these opportunities. Due to the time pressures associated with market evaluation, analysts rarely have time to gather and evaluate all of the relevant information regarding an event. It is expected that at some point analysts must fall back on heuristics to fill in the gaps in their information. Due to this, the market impact that different CSR events have is on some level a function of different psychological heuristics, implying that the rational evaluation of the market may
on occasion be quite irrational. Understanding how this irrationality plays a role will be crucial in our understanding of why certain events, occurring in certain places, leave more significant impacts than others.

To our knowledge this is the first event study to compare the market reactions to CSR events occurring at the focal firm to those occurring at a supplier. The prominence of decision heuristics in market evaluation may have contributed to the discernable differences we find in how these events were evaluated. Understanding the effect of these heuristics helps us to understand our results as we move forward with this research.

The argument could be made that the use of heuristics in market evaluations devalues the usefulness of the event study methodology as an indicator of the true financial impact CSR events have on firms. One might imagine that by undercutting efficient market theory we undercut our own analysis. In his defense of event studies, Malkiel (2003) makes the argument that even if analysts act irrationally, markets maintain a level of collective rationality superior to other forms of financial evaluation. We agree with this assessment and whole-heartedly believe that event studies are the appropriate methodology for this research. An understanding of the heuristics at play does not discount the representative power of stock prices; it adds to our understanding of them. Our hypothesized hierarchy of market reactions to different types of CSR events is based largely on the effects of heuristics on the mental accounting performed by market analysts. Without an understanding of the heuristics underlying stock price evaluations we would have a difficult time hypothesizing the differences that exist between certain types of CSR events. The existence of heuristics in stock evaluation does not take away from returns as an indicator of value; it merely adds nuance.
CHAPTER 3
DEVELOPMENT OF HYPOTHESES

To answer its research questions, this dissertation builds on previous event studies in the realm of corporate social responsibility. In addition, it builds on literature in which agency theory had been used to examine CSR oversight in outsourcing relationships, contrasting this with the theory of psychological distance as it attempts to ascertain how perceptions of supplier and focal firm CSR events differ. In addition, it draws from behavioral finance, specifically decision-making heuristics, to hypothesize a hierarchical relationship between the levels of market impact preceded by different types of CSR events.

The question of whether financial analysts discount CSR events occurring at supplier facilities relative to those occurring at focal firms can be considered as a test between the strength of two theoretical lenses: agency theory and psychological distance. While agency theory has been employed by researchers across multiple business disciplines, psychological distance has primarily been used as a theoretical lens in the psychology literature. To the best of our knowledge the use of this theory in a supply chain study is somewhat novel.

Stakeholders and the stock market may perceive that suppliers operate as agents of the firms purchasing their goods. Classic agency theory arose partially from situations in which it was resource-prohibitive for the principal (in this case the buying firm) to comprehensively monitor the actions of their agent (the supplier) (Eiseenhardt, 1989). Lack of transparency into the operations of third-party logistics suppliers is one of the leading reasons for relational failure with 3PL’s and their clients (Boyson et al., 1999).
The agent works on behalf of the principal, and is in many cases being paid by the principal to perform a service or produce a good. Agency theory would suggest that because a supplier (in this case the agent), is manufacturing or assembling products on behalf of a firm (the principal), their actions in doing so are reflective of the firm on whose behalf they are operating. This could explain Hartmann and Mueller’s (2014) findings that consumers are becoming increasingly likely to hold firms accountable for the sustainability performance of their suppliers.

Agency theory has been used extensively in the past to study the oversight and implementation of CSR in outsourcing relationships. Agency issues can occur when the risk resulting from a certain course of action is not the same for the agent and principal. When the risk or blame will likely be passed to the principal, the agent may engage in riskier behavior (Wiseman & Gomez-Mejia, 1998). Zsidisin and Ellram (2003) use an agency framework to explain how firms need to coordinate with their suppliers to prevent supply disruptions. Weak governance by the buying firm can lead to damaging labor issues with suppliers in emerging economies (Dharwadkar, George, & Brandes, 2000). McWilliams and Siegel (2001) discuss a hypothetical “optimum” level of supply-chain wide CSR, and how that might be attained. Pedersen and Andersen (2006) show IKEA acting to align their supplier with certain codes of conduct as they know that any supplier misbehavior will have a direct effect on IKEA. Wiseman et al. (2012) apply agency theory to understand outsourcing to diverse cultures.

The underlying context of the agency theory research discussed above is that the outsourcing firm will be held responsible for the irresponsible actions of their suppliers. Preventing supplier irresponsibility in order to save one’s own reputation was the primary
topic in much of that research. The proliferation of research on this topic exists because the consequences of being held responsible for a supplier’s poor CSR performance is very real. Going by a strictly agency-theory based perspective, we would expect that the outsourcing firm is held responsible for the actions of its suppliers. Therefore, we would not expect to find a significant difference in stock price impact between events occurring at the focal firm and events occurring at a supplier.

However, it may not be the case that the market weighs supply chain CSR events as heavily as they do those that occur at facilities that are actually owned by the firm whose stock is being evaluated. If differences do exist, an explanation may be found in the theory of psychological distance. The psychological distance of an event is measured by an individual’s perception of the spatial and temporal proximity of the event relative to themselves (Trope & Liberman, 2010). Psychological distance increases as the perceived spatiotemporal distance between an actor and an event increases. Heightened levels of psychological distance will decrease an actor’s perceived likelihood that the event will occur at all (Trope & Liberman, 2010; Stephan, et al., 2010). Thus individuals will discount the probability that events with high spatiotemporal distance will occur, potentially leading them to take irrational actions when assessing the risk of those events. When events happen further away from a firm, for example at a supplier site, market analysts may subconsciously degrade the impact on the focal firm due to the physical distance of the event. Estimations of distance are not subjective, but rather based on the perspective of the one doing the measuring. Artificial distances, such as the differences in stages or supply tiers can have an effect similar to actual distance (Fiske et al., 2010).
The effect of psychological distance might be best understood through the consideration of a murder. If a murder happens in another country one may think about it for a few minutes (if at all) and then move on. If this murder happens in your town, or neighborhood, you might devote more attention to it. Now imagine the murder happens in your house; it may be the only thing you think of for the rest of the day. In the way the hypothetical murder increased in importance and gained a greater “weight” as it got closer to you, CSR events may increase in importance to analysts as they grow closer to the firm whose stock is being evaluated. The greater the perceived psychological distance of an event, the greater the level of mental discounting that takes place in the estimation of its value.

Even in cases where the distance being perceived by the financial analyst is artificial rather than actual, some mental discounting may still occur. A factory owned by the focal firm could be sitting adjacent to a supplier’s facility and still, because of the perception of distance given by the perceived divide that exists between two firms, events occurring at the latter may be perceived as less impactful than those at the former. We side with psychological distance, believing that market analysts will assign greater impact to events occurring inside the firm. Therefore:

_Hypothesis 1: Internal events will have a larger impact on stock price than external events._

Though internal CSR events occurring at the focal firm may have a larger stock market impact than those occurring further up the supply chain, the disparity between the
two may be decreasing over time. Who and what a firm has responsibilities to have long been debated in the strategy literature. Milton Friedman famously stated that the only social responsibility of a firm is to turn a profit for its owners (Friedman, 1962). This somewhat narrow view stands in contrast to more recent thoughts on the responsibility of firms. While trade-offs must sometimes be made between different CSR elements, socially responsible behavior should not be seen as being mutually exclusive from “good business decisions”. Wicks (1996) defines the notion that they must be distinct from one another as the “separation fallacy”. Freidman justified this separation by claiming that a business should use its resources only for the purpose of “increase(ing) its profits so long as it stays within the rules of the game” (Friedman, 1962). However, the rules have changed since Friedman’s time. Social responsibility has become more important to academics (Carter & Easton, 2011), employees (Ditlev-Simonsen, 2015), investors (Anderson, 2006), and customers (Hartmann & Moeller, 2014); the latter three of which would be considered “stakeholders” for most firms. Stakeholder theory has been widely used in the management literature to explore the ways in which firms choose to engage in socially responsible behavior; and it is the theoretical lens we rely on here as we consider how the market impact of CSR events within the focal firm and supply chain have changed over time (Freeman, 2010).

In his seminal definition of stakeholder theory, Freeman tells us that a firm’s stakeholders are “those groups without whose support the organization would cease to exist” (2010). These “groups” (firms) are responsible to have unquestionably changed since the inception of stakeholder theory 50 years ago (Ansoff, 1965). Mitchell et al.’s (1997) typology defines stakeholders as those entities viewed as having legitimate
sources of influence that might influence a firm with some modicum of urgency.

Stakeholder theory has been frequently applied as a theoretical lens to event studies in CSR (McWilliams & Siegel, 1997; Flammer, 2003; Godfrey et al., 2009). CSR initiatives are often presented as being implemented in order to appease stakeholder wishes or ease their concerns about headline risk events. Event studies also help to understand the way that the implementation of these initiatives impact shareholders (who some – particularly Milton Friedman – would argue are the most important stakeholders). Stakeholder theory lends itself to analysis of shareholder-backed initiatives. It is relied on in this dissertation to understand how pressures for higher levels of CSR performance and the associated market reactions have changed over time.

The actions of firms and their suppliers have come under increasing levels of scrutiny from a number of different entities. Fifty years ago, firms did not have to contend with the specter of non-governmental organizations (NGOs) the way they do today. The ability of NGOs to uncover labor or abuse scandals, which can lead to subsequent fines and reputational damage, has increased exponentially (Locke & Romis, 2012). Many of the events in our sample are illegal practices first reported by NGO’s (Greenpeace, 2012; Institute for Global Labour And Human Rights, 2007), practices that could be quite costly to the responsible firms. Media has also showcased the ability to hold firms accountable for unsustainable behavior in their supply chain (Schanberg, 1996; Kenyon, 2000; Dwyer, 2010) and should be considered an influential stakeholder for many firms. The efforts of NGOs and media organizations matter because of the effect they have on the buying habits of a firm’s potential customers. These potential customers have become more sensitive to reports of a firm’s (and their supply chain’s)
social performance (Miles & Covin, 2000; Albinger & Freeman, 2000; Gershoff & Frels, 2015). Consumers’ increasing awareness and sensitivity to CSR issues combined with the increasing number of powerful, legitimate entities investigating their performance has created an unprecedented business environment where firms are arguably being held accountable to more stakeholders than ever before.

This change in stakeholder attitudes can be partially attributed to the anchoring effect and the shifting of reference points. The idea behind the anchoring effect is that people set a reference point from which they decide whether or not an event can be considered positive or negative, and to what extent (Tversky & Kahneman, 1974). The same event can be perceived very differently depending on the positioning of the reference point against which it is judged. Consider for instance a child who can write their full name with only one misspelling. If they do this when they are three years old, it is quite impressive; however if the child is 14 it might be somewhat concerning. The change in opinion for two instances of the same event is a function of a shifted reference point. Expectations for preschoolers are not the same as they are for high school students. In the same way expectations on writing ability shift between the ages of three and 14, expectations for supply chain CSR performance seem to have shifted over the last 20 years. Flammer’s (2013) recent event study lends some empirical evidence to the impact of shifting reference points. In her analysis of environmental events between 1980 and 2009, she finds evidence that the market’s reaction to negative environmental events has increased, while the reaction to positive events has lessened significantly. Anecdotal evidence suggests that a similar shift has occurred in stakeholder perception of social responsibility events as well.
We see evidence of this shifted reference point in customers (Hartmann & Mueller, 2014) and in academia (Carter & Easton, 2011), but perhaps the most telling evidence comes from industry. This shift in expectations is epitomized in the evolution Nike’s stance on supplier responsibility. When Nike first fell under fire for labor violations at their suppliers in the early 1990’s, they took a defensive stance, stating they had no responsibility for the treatment of employees at other firms, even if those firms were their direct suppliers (Locke & Romis, 2012). This is a stark contrast to their current stance. Recently they have shown a willingness to not only drop suppliers who are unable to meet their desired levels of CSR performance, but have recalled over $100 million of inventory revealed to be produced using child labor (Banjo, 2014). The attitude Nike displayed in their assertion that they were not responsible for the actions of their suppliers would be unthinkable today. However at the time, executives at Nike must have assumed this was par for the course. Nike’s change in attitude on supplier oversight is a micro-level indicator of the reference point shift that has occurred on a macro-level.

It is no longer impressive to shareholders for a firm to announce that they will be holding suppliers to high standards of worker treatment, it is expected that they do these things (Slater & Gilbert, 2004). The CSR actions that were once considered to be proactive are now seen as the cost of doing business. Recall how the quality initiatives that were revolutionary during the late 1970’s have become common (Yong & Wilkinson, 2002; Schroeder, Linderman, & Zhang, 2009). Research, popular press, and anecdotal evidence would suggest that a similar move has occurred (or is occurring) with CSR. The flip side of this shifted reference point is that negative events may now have a
larger impact as firms are being held to higher expectations. The increasing amount of stakeholders with increased visibility into the full supply chain combined with the shifting of reference points is the strongest argument to support the idea that the difference between how markets perceive the importance of supplier versus focal firm CSR performance in closing. We expect our analysis to reveal that the stock-market impact of CSR events occurring at supplier facilities have increased relative to the stock-market impact of CSR events occurring at the focal firm over the 20 year period of the analysis.

Hypothesis 2: The gap between the stock price impact of internal CSR events and external CSR events has decreased over time.

An issue with previous event study research is the tendency to group somewhat disparate events into overarching categories. There is room in the literature for a more granular approach to corporate social responsibility events. There have been multiple definitions of social responsibility (Carroll, 1994; Garriga & Melé, 2004), with some scholars pointing out that social responsibility can mean different things to different companies (Votaw, 1972). For instance, a social event involving too few diversity hires may not have the same impact on stock prices as an event involving worker safety issues. Determining if these differences exist, and if so what the magnitude of the differences between events types are, should prove to be a valuable contribution to both industry and academic literature.
Although many frameworks classifying social responsibility as it relates to business have been developed in academic literature (Carroll, 1979; Carroll, 1991; Clarkson, 1995), we adopt the frameworks established by Carter and Jennings (2002; 2004). This framework is preferable for our purposes as it was developed through interviews with purchasing and logistics managers. These frameworks were established in part to help purchasing and logistics practitioners understand the cost and revenue drivers of CSR. Through improved understanding they might better tailor their CSR strategy in a way that optimizes both efficiency and revenue. In-depth interviews with managers of large companies led to the division of logistics social responsibility into four categories: “diversity, working conditions and human rights, safety, and philanthropy/community involvement” (Carter & Jennings, 2002). We rely on this same division to define the categories of social events that will be compared in this analysis. The full definitions for these categories are fleshed out in section 3.1. These categories, along with some key literature are presented in Figure 1 below.

Figure 1: Social Event Types
We believe that there will be different levels of market reactions to different types of events. Although they are both “social performance events” it is unlikely that the stock market’s reaction of a factory fire will be comparable to their evaluation of a new community outreach program; similarly an oil spill might engender a large reaction than the receipt of an environmental citizenship award.

*Hypothesis 3: Significant differences in stock price impact will exist between different social event types.*

**Hierarchy of CSR Event Types.**

It is likely that these different types of social responsibility events will have differing levels of impact on stock evaluation. It is not enough to merely hypothesize that these differences will exist. It will be more valuable for both firms and academics to understand *why* these differences exist. Therefore in this section we develop and propose a hierarchy of CSR events sorted by the absolute value of market impact (as measured by average cumulative return abnormality – the meaning of which will be covered in greater detail in Chapter 4). However before developing this hierarchy of event types, we will explicitly define each one. We start with the four social event types.

We define diversity as any event having to do with the outreach, award reception, or lawsuits (Greehouse, 2001) having to do with corporate diversity at the employee (Cantor, 2008), supplier (Shah & Ram, 2006; Adobor & McMullen, 2007), or board level (Wright et al., 1995; Rose, 2007). Diversity has been studied extensively in supply chain research, although often in the context of financial performance (Richard, 2000; Kochan et al., 2003; Nelling & Webb, 2009), new product development (Shibyama, 2008;
Hewlett et al., 2013), or cultural issues (Pelled et al., 1999). Diversity has been included as a part of social performance in a number of event studies (McWilliams & Siegel, 2000; Hillman & Keim, 2001; Jayachandran et al., 2013); however, we are only aware of one instance where it was the single focus of an event-based study. Wright et al. (1995) provide evidence that both positive and negative diversity events have an impact on stock prices. However, this study was conducted 20 years ago. As discussed above, shifting reference points may have altered market reactions to different types of events. We see evidence of this with environmental initiatives in the competing findings of Klassen and McLaughlin (1996) and Jacobs et al. (2010). It will be interesting to see if any differences exist between Wright et al.’s (1995) findings and the findings of this research.

In the context of this dissertation, the term working conditions refers to the treatment of employees at supplier and focal firm facilities. This could include matters like criminally low pay (Deavers, 1997; Ross, 2004), forced overtime (Reilly, 2013), the use of child or slave labor (O’Rourke, 2003; Abagail McWilliams et al., 2006), labor disputes (Knoepfel, 2001; Brown, Deardorff, & Stern, 2004), or any other sweatshop-like conditions (Emmelhainz & Adams, 1999; Nisen, 2013).

Working conditions and employee treatment have been included in definitions of corporate social responsibility (CSR) as far back as the early 1950’s (Bowen & Johnson, 1953). Roberts (2003) found a correlation between poor working conditions at supplier facilities and a worsening of the buying firm’s reputation among their customers. Locke et al., (2007) explores the ways in which Nike has evolved in regards to the levels of worker treatment they hold their suppliers to, and how it has improved their overall financial performance. Locke and his coauthors also suggest that treatment of supplier
employees has grown increasingly important for firms across multiple industries (2007). Andersen and Skjoett-Larsen's (2009) use a case-study methodology to explore IKEA’s method of changing supplier culture to reflect their own beliefs on CSR. Their analysis finds that embedding CSR and sufficient worker treatment throughout the entire supply chain has been a major part of IKEA’s success.

Firms are very careful to avoid these poor sweatshop-like conditions at their suppliers’ manufacturing plants. In an in-depth examination of Asian sweatshops, Schrage (2004) found that many U.S. buying firms are implementing private labor regulations that are more stringent than the local standards. Even though they may be more costly in the short-run, firms are implementing these private standards voluntarily in order to avoid the harsh penalties they may face as a result of a labor scandal (Schrage, 2004). Firms are concerned that labor rights NGOs will uncover sub-standard labor conditions and put pressure on firms to change their practices, potentially through damaging firm reputations in the eyes of their customers (Locke & Romis, 2012).

Although safety issues may come as a direct result of poor or dangerous working conditions, a distinction is made between the two categories in this research. Safety is specifically being interpreted only as events that explicitly involve the death or injury of workers or employees. Safety has been studied extensively in supply chain and outsourcing literature (Mayhew et al., 1997; James et al., 2007; Gavious et al., 2009; Leigh, 2012; Harvey, 2014). Carter and Easton (2011) find that “Safety” was the second most frequent subject of supply chain literature focused on sustainability from 1991-2011, trailing only “Environment” in prominence. Research from the National Association of Environmental Management (NAEM, 2011) backs this finding, showing
that metrics tracking “injuries and fatalities” are reported to 75 percent of C-level executives surveyed. Of the firm’s surveyed in Cohen et al.’s (2014) examination of annual reports, “Safety” is the most measured social metric.

Like diversity, working conditions and safety have primarily been operationalized as a portion of social performance in previous event studies (Hillman & Kiem, 2001; Wai Kong Cheung, 2011; Jayachandran et al., 2013). Also like diversity, the return abnormalities generated as a result of working condition events have not been compared to return abnormalities from the other social event categories.

Research has been done on the role of philanthropy and community outreach in firm reputation (Alsop, 2004), in achieving higher levels of organizational identification (Bartel, 2001). Dean (2002) found a link between the number of local events sponsored by a firm and their perception in the community. There has also been research on the consequences of not engaging in corporate community relations. Idemudia & Ite (2006) argue that the failure of oil companies operating in Nigeria to establish significant bonds with the local community is a contributing factor to the escalating violence that has consistently derailed drilling. Frynas (2005) makes a similar claim in regards to the oil companies operating in the Gulf of Guinea. Higher levels of community outreach have also been correlated with the desire of potential employees (Albinger & Freeman, 2000). Philanthropy and community outreach have previously been incorporated into event studies (Wokutch & Spencer, 1987; Porter & Kramer, 2002). However those have mainly focused on the impacts of “positive” events. The lack of research on the short-term effects of negative philanthropy events will be discussed below.
Although I have provided references to research that has studied all four of our social event categories, we are not merely updating preexisting studies. Our work differs in that we are studying market reactions to all four categories and then making explicit comparisons between them. None of the studies discussed above have performed this type of analysis before. It would seem that making these explicit comparisons is a logical extension of the work that has preceded this dissertation.

In order to further develop our hypotheses we draw once more on behavioral finance and decision making heuristics. In section 2.3 we discussed the effect of heuristics in stock evaluations. When confronted with a lot of information at once, evaluators have a tendency to focus only on the most salient information (Shafir et al., 1993; Shiller, 2002). Whittlesea et al.’s (1990) theory of Unit Saliency states that when individuals have a better grasp of the value of an object’s unit of measurement, the object will seem more likely to occur or exist. People have an easier time processing and assigning value to concepts that are more familiar or that they have direct experience with (Whittlesea, 1987). This is because humans are better able better to comprehend situations in which they may have some experiential, rather than theoretical or “declarative”, knowledge (Musch & Klauer, 2003) and assign them proper value. The more familiar an analyst is with a concept, the more accurately he/she is when comprehending its value.

We would imagine that injuries and fatalities are more familiar to evaluators than issues regarding working conditions. Furthermore, it is more likely analysts would have some experience with injuries than any of the other event types. While we would lean towards safety events engendering a more significant move in stock prices than working
condition events, it is difficult to know for certain. A study by (Rogers, Carter, & Kwan, Working Paper) suggests that working conditions and safety are equally important to managers when deciding between supplier improvement initiatives.

Emotion and “vividness” can also play a role in the evaluation of an ambiguous event (Loewenstein et al., 2001). Specific event types within the working conditions category may not engender a greater emotional response from evaluators than every event type within the safety category. For instance evidence of large-scale child slavery may illicit a more vivid, emotional response than a report of a single injured adult. This increased emotional response may lead to a larger market reaction. However the hierarchy hypothesized in this chapter only draws comparisons between the higher-level categories laid out in the frameworks in Figures 1 and 2. On balance we believe that safety events will engender higher levels or vividness and emotion from evaluators. Therefore:

**Hypothesis 3a: Safety events will have a larger impact on stock prices than working condition events.**

Both safety and working conditions may be more vivid to evaluators than either diversity or philanthropy events. Injuries and working conditions may be easier for those evaluating the significance of events to conceptualize than more abstract concepts like diversity or philanthropy. Safety and working conditions events are generally very salient; not only in terms of units, but also in terms of immediate business impact. On-site deaths or injuries may lead to a costly supply chain disruption (Hendricks & Singhal,
Diversity and philanthropy events would appear to be less likely to precede any type of major disruption. Furthermore, diversity and philanthropy initiatives that go above what is required by law may be seen by analysts as discretionary (Arya & Zhang, 2009) and actually have a negative impact on stock price (Jayachandran et al., 2013). We would argue that it is easier to imagine the business implications of a factory collapse or report on poor labor conditions will be more tangible than the effect of diversity or community outreach events. Therefore:

**Hypothesis 3b: Safety and working condition events will have a larger impact on stock price than diversity or philanthropy and outreach events.**

In the view of market evaluators, community outreach and philanthropy may fall under the umbrella of “discretionary” actions (Johnson, 2003). Negative diversity events, while perhaps more difficult to conceptualize than worker injuries or labor conditions, are still widely acknowledged as important in supply chain literature (Robinson & Dechant, 1997; Shah & Ram, 2006; Adobor & McMullen, 2007). While there have also been findings associating positive stock prices with philanthropy announcements (Jacobs et al., 2010), the impact of “negative” philanthropy events is unclear. Popular press has chronicled corporate discrimination issues and the associated penalties very thoroughly, recently in cases with Macy’s (Feuer, 2014) and Saks Fifth Avenue (Zillman, 2015). The penalties for failure in philanthropy and community outreach are not as clear as those for the three other social event categories, as negative events often have larger market impacts than negative events (Flammer, 2013). Flammer’s finding is likely due to
negativity bias, or the propensity to give greater weights to negative entities and/or situations than to their positive counterparts. The idea of negativity bias was established by Kahneman & Tversky in their exploration of prospect theory (Kahneman & Tversky, 1979). Rozin & Royzman (2001) develop a framework to better understand the underlying causes of negativity bias, positing that negative events can seem to come more quickly and “spread” more swiftly than positive events. Shifting back to event studies, Akhtar and his coauthors (2011) reaffirm this bias as bad news has a much larger impact than good in the Australian market. Although firm reputation was a moderating variable on overall effect, Jayachandran et al (2013) also find that negative events make a larger impact on stock prices.

Although short-term “negative philanthropy events” may exist it is unclear what exactly one would be and how the market might react to one. On the other hand, we have seen many examples of negative diversity events, there are a number of discrimination lawsuits to be analyzed in our data sample (Greenhouse, 2003; Associated Press, 2009; Feuer, 2014). A database search of the phrases “philanthropy lawsuit” and “charity lawsuit” returned only one article which was in reference to a fake charity being sued for running a scam (not applicable to our study). It seems unlikely that we will find many negative short-term philanthropy events, as such it seems unlikely the cumulative return abnormality for these types of events will be as significant as those for our three other social event types. That being said, we would be remiss to not mention that philanthropy has been repeatedly found to have a positive correlation with financial performance (Wokutch & Spencer, 1987; Porter & Kramer, 2002). However, many of these studies rely on long-term financial measures, and there is a question of causality (are firms more
successful because of their charitable works, or are firms more willing to engage in charitable works due to their success?). Similarly, the few studies that have linked a lack of philanthropic activity to negative consequences (Frynas, 2005; Idemudia & Ite, 2006); the negative consequences in question have either been indirect or were long-run outcomes. It seems unlikely that this negative relationship will be reflected in our event study. Low levels of saliency on the negative impact of philanthropy and community outreach may lead to this performance type having a smaller market impact than the other types of social performance tracked in this study. Therefore:

*Hypothesis 3c: Diversity events will have a larger market impact than philanthropy and outreach events.*
CHAPTER 4

METHODS

To generate our sample we first collected a wide pool of 10,189 potential events by conducting a number of key-word searches via Factiva, considering publications from the Wall Street Journal (WSJ) and Business Wire News Service (BW). These key words are presented in Appendix A. Selection of search term keywords was based on preliminary searches and surveys of relevant literature. Using ProQuest, we searched for any article published in either Business Wire or the WSJ from January 1st, 1994 to December 31st, 2013 that contained any of these key-words or phrases in either the title or full body text of the article. Following the example of Jacobs et al. (2010), if any of the announcements contained evidence that the information about the specific event may have been announced earlier, we checked alternate sources from preceding dates to check for early announcements. If earlier announcements were found, the earlier date was the one incorporated into our analysis. This was done in order to prevent any potential confound brought about by the market having previously reacted to a particular event. This happened infrequently and we have no reason to believe the dates of our event announcements are not accurate. That being said, if any of the announcement dates were slightly off, running analysis in our four distinct time windows should offset any potential issues.

Once the sample of potential events was established, a coding system was utilized to cull the sample down to include only relevant articles. This process included reading every article and evaluating whether or not the event mentioned was relevant to our research. The event collection and coding process was carried out by the first author and
validated by another researcher through random checks in order to avoid reviewer reliability issues (Slim et al., 2003). Analysis of this validation produced a Cohen’s Kappa of \( k = 0.891 \), indicating a high level of agreement between raters. This provides a high level of confidence in the coding process utilized in this research (Cohen & Cohen, 2003). Through this coding process we reduced our sample down to 2,886 relevant events. As is common practice in event studies, the relevant events were then pared down once more through an analysis of potential confounding events. This was carried out by using ProQuest to search for any newspaper or wire service article that mentioned the firm in question’s name in the title within a three day window on either side of (before or after) the event. Unlike the initial search where only the Wall Street Journal or Business Wire were considered, every newspaper or wire service on ProQuest was eligible for the confound check. If any event that could potentially effect the firm’s stock price was announced during this time window the event was discarded. For example, if a firm released a quarterly report or announced the appointment of a new board member within a few days of an event, the stock price may be reflective of those confounding events. We would be unable to confidently correlate any abnormal stock price movement with the event in question. This is similar to the practice recommended by Corrado (2011) in his event study methodology review; and consistent with the precautions taken against confounding in multiple event study papers (Hendricks & Singhal, 2003; Jacobs et al., 2010). Through the subsequent review of potential confounding events our final sample was reduced from 2,886 to 2,301 events. A breakdown of these events by step and event type is displayed in Table 1.
Table 1: Events by Step & Type

<table>
<thead>
<tr>
<th>Event Type</th>
<th>All</th>
<th>Relevant (Pre-Confound)</th>
<th>Final (Post-Confound)</th>
<th>Percent Kept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safety</td>
<td>3,234</td>
<td>689</td>
<td>531</td>
<td>16.42%</td>
</tr>
<tr>
<td>Conditions</td>
<td>2,943</td>
<td>686</td>
<td>532</td>
<td>18.08%</td>
</tr>
<tr>
<td>Diversity</td>
<td>2,073</td>
<td>837</td>
<td>689</td>
<td>33.24%</td>
</tr>
<tr>
<td>Philanthropy</td>
<td>1,939</td>
<td>674</td>
<td>549</td>
<td>28.31%</td>
</tr>
<tr>
<td>Total</td>
<td>10,189</td>
<td>2,886</td>
<td>2,301</td>
<td>22.58%</td>
</tr>
</tbody>
</table>

The final sample of 2,301 social events spans 694 unique firms. Included in the 694 are iterations of firms before and after mergers. For instance both Conoco and Phillips are present in our sample as individual firms prior to their merger in 2002, after which they are counted as a new entity, Conoco-Phillips. The sample encompasses a broad range of industries, including firms with 67 unique two-digit SIC codes. Example headlines representing all four CSR event types, as well as internal and external events are presented in Table 2.

Table 2: Selected Event Headlines

<table>
<thead>
<tr>
<th>Headlines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate News: Amazon Acts on German Controversy --- Online Retailer Cuts Ties With Security Firm</td>
</tr>
<tr>
<td>Corporate News: Hon Hai to Raise Workers' Pay --- Manufacturer Says Move Isn't Related to Recent Suicides</td>
</tr>
<tr>
<td>American Airlines flight 1572 severely damaged</td>
</tr>
<tr>
<td>Dell No 1 Corporation in America for Promoting Multicultural Businesses</td>
</tr>
<tr>
<td>HK Grp Protests Conditions In Disney's China, Macau Plants</td>
</tr>
<tr>
<td>General Mills Donates $1 Million for Tsunami Relief Effort in Southern Asia</td>
</tr>
</tbody>
</table>
Once the sample of events was finalized, unique models predicting the expected price on the day of the event for the firm in question were constructed. This was accomplished by analyzing the previous 255 days (Hendricks, Singhal, & Stratman, 2007) of prices for the firm. The predicted prices were determined using Carhart’s (1997) 4-factor model. The four factor model is superior to the CAPM or 3-factor (Fama & French, 1993) as it includes Jegadeesh and Titman’s, 1993 momentum anomaly factor. The incorporation of this factor allows us to account for the “momentum” of the index in which the security in question resides. This model is presented in Equation 1 below:

\[
R_{it} = \alpha_i + b_{it}RMRF_t + s_{it}SMB_t + h_{it}HML_t + p_{it}PR1YR_t + \varepsilon_{it} \quad (1)
\]

Where:
\( R_{it} \) = R.O.R. on share price of firm \( i \) on day \( t \),
\( \alpha \) = intercept
\( \beta \) = systematic risk of stock \( i \)
\( RMRF_t \) = The excess return on value-weighted aggregate market proxy on day \( t \) (determined from equally weighted index of all securities traded on American, Nasdaq, and New York exchanges),
\( SMB \) = monthly premium of size factor
\( HML \) = monthly premium of book-to-market factor
\( PR1YR \) = One year momentum factor of NYSE, AMEX, and Nasdaq stock portfolios

Once this was determined, the expected price was subtracted from the actual price on the day of the event and the day immediately following and preceding it in order to determine the abnormality of the return. Two-day event periods are common in many studies of this kind (Hendricks & Singhal, 2003; Hendricks et al., 2007) as they negate any issues associated with the market capturing the effects of events that happened late in the trading day. The determination of abnormal returns is displayed in Equation 2 below:
\[ AR_{it} = R_{it} - (\alpha_i + \beta_i R_{mt}) \]  

\( AR_{it} \) = Abnormal Return  
The abnormal returns will then be averaged across all firms in the sample on day \( t \). This average is expressed in Equation 3:

\[ A_t^- = \frac{\sum_{i=1}^{N} AR_{it}}{N} \]  

Where:
\( N \) = the number of sample observations on day \( t \),

Finally, the cumulative abnormal returns will be calculated by aggregating the average abnormal returns over the time period in question, as expressed in Equation 4:

\[ CAR (t_1, t_2) = \sum_{t=t_1}^{t_2} A_t^- \]  

The CARS were determined and the models were built and using the Eventus software (Hendricks et al., 2007) embedded in the Wharton Research Data Service (WRDS).

After the final set of events was identified, abnormal returns were calculated using the WRDS database. Four potential windows were checked for our analysis. These windows consisted of market reactions on: only the day of the event announcement (0, 0); the day before and day of the announcement (-1, 0); the day of and the day after the announcement (0, 1); and finally the day before and after the event announcement (-1, 1).

If the announcement was made after markets closed at 4:00 pm Eastern Standard Time or on a non-trading day, the subsequent trading day was then treated as Day 0.

There is a strong precedent for the use of these two and three day event windows (Khotari & Warner’s chapter on the history of event studies 2006). The authors note that in many cases the announcement or occurrence of events can be anticipated by savvy
investors and a portion of abnormal return-related behavior may surface in the pre-event period. Beyond this, certain less-efficient portions of the market may react more slowly to the announcement of an event. This could be due to waiting to observe the reactions of other investors, or taking a few extra days to observe the announcement of other details surrounding the event in question (Khotari & Warner, 2006).

After these calculations the abnormal returns from our four time windows of interest were regressed against a number of binary variables using both ordinary least squares via the PROC REG command and with generalized linear models via the PROC GLM command in SAS. Generalized linear models were used as they allow for one-to-one comparisons between event types. This was useful in our granular analysis of market impact by event type. In order to test fluctuation in market preferences and priorities over time, events were divided into three distinct time periods. Period 1 includes all announcements from 1994 to 2000, Period 2 includes announcements from 2001 to 2007, and Period 3 is made up of the announcements from 2008 to 2013. T-tests were employed in order to determine the differences that exist between internal and external events in each of our three time-periods. All predictive factors in these analyses were nominal, dichotomous variables. All of these measures were repeated for all four time windows and then repeated between time periods. Due to factors that will be discussed in Chapter 5, the tests were repeated with sub-samples of both positive and negative events.
CHAPTER 5

RESULTS

Internal vs. External Events.

Full Sample.

Both Ordinary Least Square (OLS) regression and Generalized Linear Models (GLM) were used to analyze the full sample of events in order to discern the factors that may have an impact on the stock price abnormality. The six binary independent variables in this model were the four social event types (Safety, Working Conditions, Diversity and Philanthropy), whether or not the event is internal or external, and if an event is positive or negative. The dependent variables in these models were the cumulative abnormal returns in our four time windows of interest. The calculations used in these models are displayed in Equation 5.

\[
 CAR_{ij} = B_0 + B_1 Supplier + B_2 Negative + B_3 Safety + B_4 Conditions + B_5 Diversity + B_6 Philanthropy + e
\]  

Regression analysis suggests that the independent variables are very predictive of abnormal stock price variation on the day the event is announced \((p = 0.0003)\) and in the window \((-1,0)\) where \(p = 0.0286\), predictive at a level of \(p < 0.10\) in window \((0,1)\) where \(p = 0.0898\), and not at all predictive in the three day window ranging from day -1 to day 1 \((p =0.455)\). However, the majority of factors were not significant predictors of stock price variation, with only the predictor indicating whether or not an event was negative, generating significance in more than one time window. Whether or not an event is
negative has a negative impact on stock price variation in windows (0,0), (-1,0) and (0,1). Whether or not an event was internal or external was only significant (p = 0.048) in the time window (0,1). The standardized results of the OLS regression analysis are displayed in Table 3 below. The GLM models were used to take advantage of the “contrast” statements, allowing us to make one-to-one comparisons between social event types.

Table 3: Full Sample Regression

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(-1,0)</th>
<th>(0,1)</th>
<th>(-1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model p-Value</td>
<td>0.0003***</td>
<td>0.0286**</td>
<td>0.0898*</td>
<td>0.4552</td>
</tr>
<tr>
<td>Intercept</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>-0.0729**</td>
<td>-0.04434**</td>
<td>-0.0537**</td>
<td>-0.03906</td>
</tr>
<tr>
<td>Safety</td>
<td>-0.0378</td>
<td>-0.00522</td>
<td>-0.03174</td>
<td>-0.00943</td>
</tr>
<tr>
<td>Supplier</td>
<td>0.02292</td>
<td>0.00266</td>
<td>0.04648*</td>
<td>0.02934</td>
</tr>
<tr>
<td>Working Conditions</td>
<td>0.03519</td>
<td>0.06178*</td>
<td>-0.00961</td>
<td>0.01984</td>
</tr>
<tr>
<td>Philanthropy</td>
<td>0.04609</td>
<td>0.03891</td>
<td>0.00862</td>
<td>0.01105</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.02775</td>
<td>0.05723</td>
<td>0.00052381</td>
<td>0.02868</td>
</tr>
</tbody>
</table>

Notes: N = 2,301 observations, ***p<0.01, **p<0.05, *p<0.1

As mentioned in Chapter 4, the sample contains firms from 67 unique two-digit SIC codes. In order to control for any potential influence due to the industry in which the events took place, multi-level models were run using the PROC MIXED procedure in SAS, with SIC codes as the level-1 variable in which the rest of the variables were nested. This analysis suggests that industry by two-digit SIC code did not have a significant effect on the shift in stock-price caused by our independent variables. Due to the proliferation of SIC codes in our sample, many of the 2-digit SIC codes only contained 1-2 events nested within them. To compensate for this, we followed the example of Flammer (2013) and “went up a level”, running additional multi-level models
using only the first digit of the SIC code as the level-1 variable. This limited us to only eight dummy variables as opposed to 67, all except one of which had a sample size of at least 30 (with seven having sample sizes greater than 100). This smaller range of SIC values allowed us to utilize the “contrast” statement to compare the effects of the different groups to one another. These multi-level models were run with both two and one-digit SIC codes as the level-1 variable across all potential time windows. None of these tests displayed anything to suggest that industry played a significant factor in our results.

**Negative Event Sample**

Hypothesis 1 stated that “internal events will have a larger impact on stock price than external events”. Analysis shows no evidence to support this hypothesis so we must reject it. However there is an issue to be considered with this finding. When observing the sample as a whole, the majority of our external events are also negative. Negative events make up a minority of the total sample, yet were still the most significant factor in predicting stock-price variation. The authors reasoned that perhaps the impact of external events would be more apparent if the sample were whittled down to only the negative events. As displayed in Table 4, external events make up only 10.26% of the total event sample. This stands in stark contrast to the 36.33% of negative events made up by negative external events.
Noting this, the sample was drilled down to the 501 negative social events for further analysis. It is interesting that of these events, the majority of CAR’s on the day of the event are negative, with 214 positive CAR’s and 276 negative CAR’s. Furthermore, the “maximum” CAR reported on day zero is 10% above the expected return; the “minimum” CAR reported on day zero is 26% below the expected return. This is particularly striking when contrasted with the 1,800 positive events. The majority of positive events are actually negative, with CAR’s ranging from 48.61% to 49.06% actually registering as positive (although the average abnormal return across all positive events is positive – suggesting that although there are fewer positive CAR’s, they are more strongly positive than the negative CAR’s are negative). Across all four time windows, the average CAR for positive events is much closer to zero than the average CAR for negative events. This, along with the results of regression analysis, suggests that the positive events in this sample have a negligible effect on abnormal stock price. Further descriptive statistics for both negative and positive events in all four time windows are displayed in Table 5.

Table 4: Event Breakdown

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Supplier Events as %</th>
<th>Supplier Dev. Events as %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Events</strong></td>
<td>2301</td>
<td>10.26%</td>
<td>20.03%</td>
</tr>
<tr>
<td><strong>Negative Events</strong></td>
<td>501</td>
<td>36.33%</td>
<td>5.39%</td>
</tr>
<tr>
<td><strong>Positive Events</strong></td>
<td>1800</td>
<td>3.00%</td>
<td>24.11%</td>
</tr>
<tr>
<td><strong>Supplier Events</strong></td>
<td>236</td>
<td>-</td>
<td>24.11%</td>
</tr>
<tr>
<td>**Neg. Supplier Events</td>
<td>182</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>**Supplier Dev. Events</td>
<td>461</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>**Neg. Supplier Dev</td>
<td>27</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Events</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 5: Negative & Positive Events Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(-1,0)</th>
<th>(0,1)</th>
<th>(-1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Negative Events</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative CARs</td>
<td>284</td>
<td>262</td>
<td>261</td>
<td>261</td>
</tr>
<tr>
<td>Positive CARs</td>
<td>217</td>
<td>239</td>
<td>240</td>
<td>240</td>
</tr>
<tr>
<td>N</td>
<td>501</td>
<td>501</td>
<td>501</td>
<td>501</td>
</tr>
<tr>
<td>% neg</td>
<td>56.69%</td>
<td>52.30%</td>
<td>52.10%</td>
<td>52.10%</td>
</tr>
<tr>
<td>Average CAR</td>
<td>-0.41%</td>
<td>-0.34%</td>
<td>-0.33%</td>
<td>-0.26%</td>
</tr>
<tr>
<td>Max CAR</td>
<td>10.45%</td>
<td>12.03%</td>
<td>16.45%</td>
<td>16.77%</td>
</tr>
<tr>
<td>Min CAR</td>
<td>-26.09%</td>
<td>-29.90%</td>
<td>-26.79%</td>
<td>-27.50%</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>2.81%</td>
<td>3.45%</td>
<td>3.53%</td>
<td>4.01%</td>
</tr>
<tr>
<td><strong>Positive Events</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative CARS</td>
<td>917</td>
<td>915</td>
<td>931</td>
<td>925</td>
</tr>
<tr>
<td>Positive CARS</td>
<td>883</td>
<td>885</td>
<td>869</td>
<td>875</td>
</tr>
<tr>
<td>N</td>
<td>1800</td>
<td>1800</td>
<td>1800</td>
<td>1800</td>
</tr>
<tr>
<td>% Positive</td>
<td>49.06%</td>
<td>49.17%</td>
<td>48.28%</td>
<td>48.61%</td>
</tr>
<tr>
<td>Average CAR</td>
<td>0.022%</td>
<td>0.021%</td>
<td>0.021%</td>
<td>0.020%</td>
</tr>
<tr>
<td>Max CAR</td>
<td>20.32%</td>
<td>20.12%</td>
<td>30.89%</td>
<td>30.69%</td>
</tr>
<tr>
<td>Min CAR</td>
<td>-21.93%</td>
<td>-34.33%</td>
<td>-46.80%</td>
<td>-29.13%</td>
</tr>
<tr>
<td>Std. Dev</td>
<td>1.97%</td>
<td>2.75%</td>
<td>3.11%</td>
<td>3.53%</td>
</tr>
</tbody>
</table>

Similar to the analysis of the full sample of events, OLS regression and GLM analysis were utilized in the analysis of the negative event sample. This was carried out using all of the same independent variables, except for the “Negative” variable as all events in this sample are negative. The variables used in this equation are displayed in Equation 6.

\[
CAR_{ij} = B_0 + B_1Supplier + B_2Safety + B_3Conditions + B_4Diversity + B_5Philanthropy + e
\]

Analysis suggests that the explanatory power of our variables is highly significant in all four windows of interest. The model fit statistics and regression weights for all four
windows of interest as displayed in Table 6. Analysis suggests that whether or not an event has a safety component is strongly significant in two of our time windows (most strongly on the day of the event announcement). However, whether or not an event is internal or external is most pronounced in the two day window (0, 1). This (0,1) window is interesting as additional information regarding the event can leak out on the day after the event. Our analysis suggests that firms suffer a significantly reduced drop the day after the announcement of the event if the event occurred at a supplier’s facility. On the day of the event, whether or not an event is internal or external is of no identifiable consequence. The primary significant factor is whether or not the announcement included an element of safety. Perhaps the discounting that occurs due to the externality of an event comes after the initial shock of the event? Or perhaps the ability of the buying firm to disavow/rebuke the actions of their supplier allows them to avoid any subsequent penalties? Interestingly, Diversity events have a significant positive effect in two of our windows of interest. This suggests that the market does not consider negative diversity events to be particularly impactful to a firm’s future value.

Table 6: Negative Event Sample Regression

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(-1,0)</th>
<th>(0,1)</th>
<th>(-1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Fit</td>
<td>p = 0.0106***</td>
<td>p = 0.0121**</td>
<td>p = 0.0026***</td>
<td>p = 0.0061***</td>
</tr>
<tr>
<td>Supplier</td>
<td>0.064</td>
<td>0.013</td>
<td>0.114**</td>
<td>0.066</td>
</tr>
<tr>
<td>Safety</td>
<td>-0.1303**</td>
<td>-0.088</td>
<td>-0.112*</td>
<td>-0.084</td>
</tr>
<tr>
<td>Working Conditions</td>
<td>0.008</td>
<td>0.049</td>
<td>0.024</td>
<td>0.058</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.067</td>
<td>0.109**</td>
<td>0.085</td>
<td>0.114**</td>
</tr>
<tr>
<td>Philanthropy</td>
<td>0.021</td>
<td>0.040</td>
<td>-0.013</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Notes: N = 501 observations, ***p<0.01, **p<0.05, *p<0.1
This analysis would suggest that the only robustly significant predictive variable in our analysis is whether or not the event featured some sort of safety component. This also suggests that whether or not an event occurred internally or externally is not relevant (except in a single window of time) and the market treats events occurring at either the buyer or the supplier equally harshly. However it is important to note the majority of our sample of negative events occurred in period 3. It is possible that a difference in market reaction to external relative to internal events existed in earlier periods. In order to test this we next segmented events by period and used t-tests to compare the CARs in external and internal events by period.

Table 7: Negative Events Internal vs. External Events by Period

<table>
<thead>
<tr>
<th>Period</th>
<th>Internal Events Mean (n = 95)</th>
<th>External Events Mean (n = 49)</th>
<th>Mean Difference</th>
<th>Period</th>
<th>Internal Events Mean (n = 111)</th>
<th>External Events Mean (n = 54)</th>
<th>Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Period 1 (1994 - 2000)</strong></td>
<td>(0,0)</td>
<td>(-1,0)</td>
<td>(0,1)</td>
<td>(-1,1)</td>
<td><strong>Period 2 (2001 - 2007)</strong></td>
<td>(0,0)</td>
<td>(-1,0)</td>
</tr>
<tr>
<td>Internal Events</td>
<td>-0.935%</td>
<td>-0.740%</td>
<td>-0.850%</td>
<td>-0.657%</td>
<td>Internal Events</td>
<td>-0.392%</td>
<td>-0.332%</td>
</tr>
<tr>
<td>Mean Difference</td>
<td>0.0301**</td>
<td>0.0587*</td>
<td>0.0228**</td>
<td>0.0432**</td>
<td>Mean Difference</td>
<td>0.6705</td>
<td>0.1712</td>
</tr>
</tbody>
</table>

***p<0.01, **p<0.05, *p<0.1

Table 7 displays the results of a series of t-tests that were conducted in order to better understand how the difference in market perception of negative supplier events has changed over time. Prior to running the t-tests, F-tests were conducted in order to determine whether the variance between internal and external samples were equal or unequal. F-test analyses revealed that both types of variance existed among our pairs. In
cases with unequal variance, the Satterthwaite significance test was used to correct for this. When variance was relatively equal, pooled p-values were employed. In the period from 1994 to 2000, we find significant differences in the way the market reacted to negative internal versus negative external events – reacting significantly less negatively to external events in all four event windows. The difference between internal and external events was significant at a $p < 0.10$ level in the time window (0,1) and significant at a $p < 0.05$ level in the other three windows of interest. It is interesting to note that the abnormal return means of the external events are negative across all time periods and event windows. External event means are actually positive in period 1. They became negative in three of the four windows in period 2 and were negative in all four event windows in period 3. The shifts in return abnormality for internal and external events over our years of interest are displayed in Figure 2.

Figure 2: Internal & External Average Negative Event CAR’s Over Time
These findings show a clear discounting applied to the market penalty associated with negative external events. The market seems to have reflected the attitude of the Nike executive quoted earlier (Locke & Romis, 2012) in which the socially irresponsible actions of a supplying firm were not considered to be the responsibility of their customers. Interestingly, this discounting does not hold in later periods. In the period from 2001 to 2007 the reaction only differs in the (0, 1) event window. In the final period of analysis, from 2008 to 2013 there are no significant differences in the penalties levied by the market in any of the event windows. There is a clear implication that whatever “discounting” existed between negative internal and external events has not only decreased over the last 20 years, but has virtually disappeared. Although differences in the market reaction to internal external events may have once existed, there is no evidence to suggest that it exists today. These results lead us to reject Hypothesis 1.

It is interesting to note that the magnitude of an average abnormal return for a negative event has decreased over the course of our sample. This differs from Flammer’s findings that CAR’s associated with “eco-negative’ events actually got more negative over time (2013). Two potential reasons may explain this phenomenon. First, as events are “repeated” over time, the “shock” associated with them may wane, leading to increasingly indifferent market reactions (although it could be argued that the opposite is true). A more likely factor is that the volumes of events that are reported have increased steadily over time. Major events like airline crashes or factory fires have been reported since the beginning of this sample, but smaller-scale events that the market places less importance on may have been underreported in earlier periods, or are simply more widely reported now. If an increased number of what the market considers to be “low
importance” events have flooded into more recent time periods, these events could
decrease the magnitude of observed abnormal returns. However, the fact that the negative
events in our sample generate such strongly negative CAR’s despite the potential
existence of these factors may actually strengthen the argument that the market’s attitude
on the impact of negative internal relative to external events has changed.

Similar tests were carried out with positive events but no robustly significant
differences were found between internal and external events. The distinct lack of positive
external events mentioned earlier in the chapter is likely a factor here. Finally, measures
were taken to ensure that the non-significant difference found between negative internal
and external events is not due to the apparent decrease in impact of negative internal
events in Period 3. T-tests confirmed that no significant differences exist between the
CAR’s associated internal events in period 1 and the external events in period 3. This
provides us with confidence that the disappearance of the disparity in the market’s
reaction to negative internal relative to negative external events is due to the increased
emphasis put on external events, not any potential decreased emphasis on internal events.

Hypothesis 2 stated that “The gap between the stock price impact of internal CSR
events and external CSR events has decreased over time”. There is no evidence for this
hypothesis when observing only positive events, and limited evidence when all events are
considered. However, when solely observing negative events, we find strong evidence
that an “impact gap” existed in the period from 1994 – 2000 and not in the subsequent
periods from 2001 – 2013. The implications suggested by this analysis provide support
for Hypothesis 2.
The primary objective of this research is to determine if financial drivers exist behind what is perceived as an increasing importance and focus on supplier responsibility. The rejection of Hypothesis 1 and failure to reject Hypothesis 2 suggest that the financial realities of supplier CSR performance have changed. Where the market may have once given firms a free ride or reduced penalty in the past, it now seemingly treats the actions of direct suppliers as the responsibility of their customer.

**Comparison Between Event Types.**

In order to test for differences among the varying social event types, Generalized Linear Models (GLM) were used. These models were utilized in order to take advantage of their “contrast” statements, allowing for one-to-one comparisons between social event types. This analysis was expanded to include any event type interactions that appeared more than 30 times in our sample. These interactions were events that contained elements of more than one social event type. The interactions included here were events containing elements of both safety and working conditions, and events containing elements of both diversity and philanthropy. Models were run for both the full sample of all 2,301 events, and for the 501 event sample of only negative events.
Table 8: Full Sample Comparisons between Event Types

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>p = 0.014</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>0.0279**</td>
<td>0.0574*</td>
<td>0.246</td>
<td>0.2201</td>
</tr>
<tr>
<td>Conditions</td>
<td>0.0071**</td>
<td>0.0265**</td>
<td>0.1356</td>
<td>0.1734</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.4455</td>
<td>0.2631</td>
<td>0.4554</td>
<td>0.2631</td>
</tr>
<tr>
<td>Phil.</td>
<td>0.0009***</td>
<td>0.262</td>
<td>0.3163</td>
<td>0.3163</td>
</tr>
<tr>
<td>WCSafe</td>
<td>0.0343*</td>
<td>0.1111</td>
<td>0.0409**</td>
<td>0.1267</td>
</tr>
<tr>
<td>DivPhil</td>
<td>0.4494</td>
<td>0.5785</td>
<td>0.4528</td>
<td>0.8035</td>
</tr>
<tr>
<td><strong>p = 0.265</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>0.0574*</td>
<td>0.246</td>
<td>0.4554</td>
<td>0.2631</td>
</tr>
<tr>
<td>Conditions</td>
<td>0.0265**</td>
<td>0.1356</td>
<td>0.3163</td>
<td>0.3163</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.0009***</td>
<td>0.0409**</td>
<td>0.0409**</td>
<td>0.0409**</td>
</tr>
<tr>
<td>Phil.</td>
<td>0.1111</td>
<td>0.5785</td>
<td>0.5951</td>
<td>0.5951</td>
</tr>
<tr>
<td>WCSafe</td>
<td>0.262</td>
<td>0.0674</td>
<td>0.5951</td>
<td>0.5951</td>
</tr>
<tr>
<td>DivPhil</td>
<td>0.4455</td>
<td>0.5555</td>
<td>0.3785</td>
<td>0.3785</td>
</tr>
<tr>
<td><strong>p = 0.243</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>0.246</td>
<td>0.4554</td>
<td>0.3163</td>
<td>0.2631</td>
</tr>
<tr>
<td>Conditions</td>
<td>0.2201</td>
<td>0.8035</td>
<td>0.4528</td>
<td>0.8035</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.411</td>
<td>0.4554</td>
<td>0.0781*</td>
<td>0.3032</td>
</tr>
<tr>
<td>Phil.</td>
<td>0.9251</td>
<td>0.0409**</td>
<td>0.7612</td>
<td>0.3032</td>
</tr>
<tr>
<td>WCSafe</td>
<td>0.2275</td>
<td>0.262</td>
<td>0.0781*</td>
<td>0.262</td>
</tr>
<tr>
<td>DivPhil</td>
<td>0.5859</td>
<td>0.5555</td>
<td>0.2374</td>
<td>0.1457</td>
</tr>
<tr>
<td><strong>p = 0.477</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety</td>
<td>0.2201</td>
<td>0.411</td>
<td>0.9251</td>
<td>0.5859</td>
</tr>
<tr>
<td>Conditions</td>
<td>0.2275</td>
<td>0.411</td>
<td>0.2236</td>
<td>0.5762</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.2201</td>
<td>0.411</td>
<td>0.4307</td>
<td>0.4307</td>
</tr>
<tr>
<td>Phil.</td>
<td>0.2275</td>
<td>0.411</td>
<td>0.4307</td>
<td>0.4307</td>
</tr>
<tr>
<td>WCSafe</td>
<td>0.5859</td>
<td>0.5762</td>
<td>0.2374</td>
<td>0.1457</td>
</tr>
<tr>
<td>DivPhil</td>
<td>0.5859</td>
<td>0.5762</td>
<td>0.2374</td>
<td>0.1457</td>
</tr>
</tbody>
</table>

***p<0.01, **p<0.05, *p<0.1

Table 8 displays the p-values associated with the one-to-one differences between factors in the overall sample. The stock price fluctuations generated by safety events are shown to be significantly different from working condition, philanthropy, and diversity events in window (0,0) and different than the former two in window (-1,0). Events with
elements of both safety and working conditions also display differences from many of the
other factors in windows (0,0) and (0,1).

The negative sample of events was analyzed using the nearly exact same GLM
contrast method. The major difference between the two analyses was the dis-inclusion of
two event types. As there are so few negative philanthropy events in the sample, they
were not included among the other three event types in the analysis of negative events.
For the same reason, the interaction between diversity and philanthropy events was not
included in this analysis.

As the variables in the negative sample had greater predictive power than those in the
overall sample, we expected to see a greater distinction between event types in the
analysis of the negative sample. The p-values signifying the one-to-one differences
between negative event types are displayed in Table 9.
In Hypothesis 3 it was posited that *significant differences in stock price impact will exist between different social event types*. The results of the GLM models run in both the overall and negative samples of events suggest that there are in fact differences in stock price impact between social event types. In time window (0,0) safety events were significantly more impactful than working condition, diversity, philanthropy or events
that contained both diversity and philanthropy elements. It should also be noted that events containing elements of both safety and working conditions had a significantly larger impact on stock price fluctuations than events that contain both diversity and philanthropy elements. In time window (-1,0) significant differences exist between safety and working conditions as well as between safety and diversity events. In time window (0,1) differences between events containing both safety and working condition elements relative to diversity and philanthropy, as well as events containing both philanthropy and diversity events. No indication of significant differences among any of the event types in time window (-1,1).

GLM models were used to analyze the sub-sample of positive events – none of the event types were shown to have significantly different market impacts. This finding is in line with the lack of significant findings of any kind in the positive sample of events. However, analysis of negative events suggests clearer distinctions between event types than the full-sample analysis discussed above. GLM contrasts consistently demonstrate that safety events have a significantly larger impact on stock price fluctuations than any of our other event types across all time windows. Both Hypothesis 3 and 3a are thus supported by our analysis of negative events by event type.

What is surprising about our findings for hypothesis 3b is that working condition events did not seem to have a significantly greater impact on stock-price fluctuations than diversity or philanthropy events. However, it should also be noted that events that were classified as having both safety and working condition elements (heretofore referred to as WCSafe events) had a significantly greater impact on stock price fluctuations than diversity events. Additionally, the mean impact of WCSafe events is actually more
negative than the impact of safety-only events in time windows (0,0) and (0,1). The implications of this will be discussed further in Chapter 6. The mean CAR generated from each negative social event by type is presented across all four windows of interest in Table 10.

Table 10: Mean CAR Value by Negative Event Type

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(0,1)</th>
<th>(-1,0)</th>
<th>(-1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCSafe</td>
<td>-1.27%</td>
<td>-1.14%</td>
<td>-0.89%</td>
<td>-0.76%</td>
</tr>
<tr>
<td>Safety</td>
<td>-0.81%</td>
<td>-0.87%</td>
<td>-0.92%</td>
<td>-0.98%</td>
</tr>
<tr>
<td>Conditions</td>
<td>-0.06%</td>
<td>0.13%</td>
<td>-0.02%</td>
<td>0.17%</td>
</tr>
<tr>
<td>Diversity</td>
<td>0.28%</td>
<td>0.63%</td>
<td>0.58%</td>
<td>0.93%</td>
</tr>
</tbody>
</table>

Hypothesis 3b stated that: *Safety and working condition events will have a larger impact on stock price than diversity or philanthropy and outreach events.* The analysis does not demonstrate any significant difference between working condition events and diversity and philanthropy events. A potential explanation is that our definition of working condition events was too broad, and we included events that would be less of a “shock” to markets. In future research working condition events, along with all of the other event type categories, could be broken down further to even more granular levels to verify if this is the case. Another possibility is that many of the most impactful working conditions events also contained some element of safety. Working conditions and safety often go hand-in-hand. Analysis of negative events *did* suggest that events containing both working condition and safety elements had a larger impact than diversity events. However because these were the only working conditions events significantly more impactful than diversity events, it is difficult to attribute this difference to the presence of
working conditions elements. Based on these findings we do not find support for Hypothesis 3b.

In Hypothesis 3c it was posited that: Diversity events will have a larger market impact than philanthropy and outreach events. This could not be tested in the negative event sample as there were too few negative philanthropy events to make any meaningful type of comparison. Contrasts between the two event types carried out with the positive and overall samples also failed to reveal any significant differences between philanthropy and diversity events. Based on these findings, we do not find support for Hypothesis 3c.

Controlling For Endogeneity.

The potential for endogeneity is a major issue in academic research (Shugan, 2004). Much of this is due to issues with omitted variable bias (Bascle, 2008). Empirically speaking, endogeneity was addressed through 2SLS models and the introduction of an instrumental variable, as well as through the creation of our independent variables using the Carhart Four-Factor Model. By its nature the Carhart model controls for many potential instrumental variables. As mentioned in Chapter 4, this estimation method controls for the, “momentum” of the exchange in question, the size and book-to-market value of the firm in question, and any excess returns via a weighted index of returns across several security exchanges.

Leaving alone the potential risk of reverse causality, it is possible that an omitted predictor is influencing stock price abnormality and has not been accounted for in our analysis. A dichotomous measure indicating whether or not the event in question was not, or was, customer-facing was used as an instrumental variable. This variable (called
“process”) was chosen because it is highly correlated with the focal independent variable (whether or not the event was external), and lowly correlated with the dependent variables (CAR’s in all four time windows). 2SLS models were run for both the overall event sample and the negative event sample. All of the other variables in the 2SLS models are the same as those specified above in equations 5.1 and 5.2 respectively. The results of these models were highly consistent with the non-instrumental regression models. The consistency of these models was formally tested using the Durbin-Wu-Hausman test. These tests suggested no significant difference between the instrumented and non-instrumented models for either the overall sample model ($\chi^2 = 0.81, p = 0.662$) or the negative event sample model ($\chi^2 = 1.98, p = 0.3708$). Thus we fail to reject the null hypotheses that our models and instrumented models vary from each other in any significant fashion. The additional endogeneity analysis increases our confidence that potential endogeneity between whether or not an event is internal or external and stock price abnormality is not a concern.
CHAPTER 6

DISCUSSION

**Significance to the Field.**

The research questions posed at the beginning of this paper were as follows: 

*Does the market distinguish between CSR events occurring at the focal firm and events occurring at a supplier? Does the market evaluate alternate, distinct types of CSR events differently from one another? If so, what is the magnitude of these differences? Finally, how have the differences in market impact between all event types studied in this research changed over the last 20 years?*

All of these questions were answered in our analysis. More importantly, through our findings we may posit an answer to the questions that underlies them all: *What are the financial reasons for the (seemingly) increasing amount of attention being paid to supplier CSR?*

Since the emergence of sustainable supply chain literature to the present, there have been calls for research that answers these questions empirically (Carter & Rogers, 2008; Seuring & Müller, 2008; Markman & Krause, 2016; Busse, 2016). These calls are undoubtedly influenced by what appears to be a sea-change in terms of the way firms think about, and ultimately treat, supplier CSR performance. Where firms were once able to project plausible deniability when it comes to the actions of their suppliers, this no longer seems to be the case. There are several prominent examples of firms investing in the CSR performance of their suppliers. Perhaps the most well-known example of a firm committing significant resources to supplier responsibility is Wal-Mart’s supplier
sustainability program. Wal-Mart’s large-scale supplier sustainability programs were hailed as revolutionary when they were first introduced (Rosenbloom, 2009). However some skeptics questioned whether a firm like Wal-Mart, which relies on fairly slim margins for many of their products, could afford a potentially costly supplier sustainability program (Lurie, 2010). When commenting on whether or not Wal-Mart would be able to afford their supplier program, chief executive officer Mike Duke stated: “We can’t afford not to (be so aggressive in sustainability)” (Wal-Mart, 2010). Perhaps what Mr. Duke is alluding to is the necessity to avoid potentially costly situations in which a supplier’s actions can hurt their customer’s bottom line.

Chipotle’s recent issue with food-safety provides a well-publicized example of such a situation. Recently, Chipotle went through a crisis in which their suppliers provided them with produce that resulted in multiple cases of salmonella for Chipotle customers, and was the primary factor in Chipotle posting their first ever losing quarter, with losses of over $26 million and same-store sales falling nearly 30 percent. Beyond the financial consequences, Chipotle also saw their reputation suffer as their brand admiration slipped from 70 to 50 percent (Jargon & Minaya, 2016). There are other, numerous examples of supplier actions causing similar issues that have caused serious repercussions for firms (i.e., the aforementioned cases of Mattel and Nike). However they tend to be anecdotal in nature and it is difficult to generalize the financial effects of supplier misdeeds. In this dissertation we attempted to shed light on these effects by empirically investigating the subnormal stock price shifts associated with external relative to internal CSR events, thereby uncovering any shifts in financial incentives that may explain the phenomenon.
A summary of the hypotheses that were supported or rejected by analysis of both the overall and negative event samples are displayed in Table 11.

Table 11: Hypotheses Results by Sample

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Supported in Overall Event Sample?</th>
<th>Supported in Negative Event Sample?</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H2</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>H3</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>H3a</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>H3b</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H3c</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

A potential financial explanation for why firms seem to be increasing their focus on supplier responsibility is given by our findings for Hypotheses 1 and 2. In Hypothesis 1 it was postulated that some type of “discounting” may be applied to external events. The theory of psychological distance (Trope & Liberman, 2013) purports that the further removed an event is from the focal point of evaluation, the greater the level of discounting that is applied to the significance of the event. This was contrasted with agency theory, in which principals (buying firms) are held responsible for the actions of the agents (suppliers) working on their behalf (Eisenhardt, 1989). Hypothesis 1 was essentially a test of the tension between these two theoretical lenses. Our analysis suggested that there is currently no difference in how the stock market treats external CSR events relative to internal CSR events. This finding suggests the primacy of agency theory over psychological distance in this particular context. However, our findings for Hypothesis 2 indicate that this may have not always been the case.

Although we find no evidence that a “supplier discount” currently exists when it comes to CSR events, our analysis suggests this discount did exist in the past, particularly
in the period from 1994 to 2000. Evidence of this discount disappears from our sample by the turn of the millennium, with no evidence of its existence from 2001 to 2013. This change may be evidence of a shifting reference point (Tversky & Kahneman, 1974) in terms of what level of supplier social responsibility is acceptable to market evaluators. These findings suggest that from 1994 to 2000 firms were not expected to monitor (or perhaps even to be aware of) the actions of their suppliers. Even when faced with negative external events, which carry the most weight with market evaluators, firms actually saw their abnormal stock valuations tick upwards. When it came to negative supplier CSR events in this period, market evaluators were either unconcerned or unaware of a firm’s connection to the offense-committing supplier.

However since 2001, it would appear that these expectations have shifted. Plausible deniability is no longer a viable defense as firms are being held equally accountable for the missteps of their suppliers as they are for their own. Perhaps this is due to the preponderance of events, such as those detailed earlier in the chapter, in which firms suffered clear financial consequences as the results of the actions of their suppliers. We might also attribute this to the increased level of transparency that market evaluators currently have into supply chains. A combination of reporting legislation such as Dodd-Frank and developments in communication technology has increased the visibility into, and scrutiny of, supplier activities. These factors were not in play to provide supplier clarity in the beginning of our sample, particularly for international suppliers. Whatever the reason is for this paradigm shift, the key for firms going forward is that it has taken place.
The combined findings of Hypotheses 1 and 2 provide significant managerial implications. The first and most obvious is that there are clear financial incentives behind the increased attention being paid to supplier CSR performance. In a recent conceptual piece, Busse (2016) posits a strong positive relationship between supplier sustainability programs and the financial performance of their customers. What Busse establishes conceptually, our findings establish empirically. The results of this study demonstrate that firms do not just have a moral imperative to monitor and ensure supplier responsibility, but a financial one as well. Negative CSR events have a clear and significant negative impact on stock prices, even when occurring at a supplier’s facility. In order to avoid the financial consequences associated with negative social events, firms need to mitigate their exposure to risk in any way that they can.

This exposure to risk is most apparent in events involving safety issues. Analysis consistently shows that the presence of safety elements was the key driver in the impact of negative social events on stock price. In one-to-one comparison tests in both the negative and overall sample of events, safety events were significantly more impactful than the other three social event types – particularly on the day the event was reported. This finding provides a clear indication that negative safety incidents are the supplier actions that put buying firms at the highest level of financial risk.

There is no clear evidence that any of the other three social event types examined in this study led to any significant financial penalties. It should be noted that this lack of effect was consistent for both external and internal events. This finding was somewhat surprising, particularly for working conditions events.
It had been expected that while working condition events would not have the same impact as safety events, that their impact would easily outstrip the impact of either diversity or philanthropy events. It is possible that our definition of a working condition event is too broad, and that events were included within this group that counteracted the effects that came from events such as discoveries of child labor or sweat-shop like conditions. For instance, many of the working condition events included in the sample revolved around strikes or similar labor disputes. Not only might these events not carry the same weight as a child or slave labor incident, but they might also be seen coming well ahead of time, with the market having already adjusted to account for them.

However, when allowing for interactions between event types, events containing elements of both safety and working conditions violations (WCSafe events) actually had a greater negative impact than negative safety-only events. This suggests that firms are more harshly penalized for accidents resulting in injury or death when those accidents may be directly attributable to poor, and potentially unsafe, working conditions. This finding is interesting as it may indicate that firms face harsher penalties when negative safety occurrences appear to have been preventable – potentially stemming from neglect on the part of the focal firm or their supplier (or both). The market may perceive an element of laziness or neglect from the firm and/or their suppliers in these types of events – perhaps signifying some sort of mismanagement. This suggests that the market is “smart” in the penalties it assigns to firms, taking more into consideration than simply the type of event. It is possible that one of the reasons the market penalizes firms more harshly for the announcement of WCSafe events is that the negligence (or possibly ignorance) required to allow such an event to happen may be indicative of a susceptibility
to future CSR issues. The market may be more willing to move past what they believe to be a tragic, one-time event, than they are to move past an event that suggests systemic managerial failings. The notion that the market is “smart” and considers so many factors in their assessment of CSR announcements highlights the need for a more granular study of CSR event types. This could potentially be done through some form of content analysis (Berelson, 1952; Krippendorff, 2013) in which we identify different, distinct subsets of our social event types.

Clearly the sample of negative events loaned itself more readily to this research. Our key finding, that the difference in market reaction to external events relative to internal events had changed over time, was only evident in analysis of the negative sample. This is consistent with the findings of recent event studies (Jayachandran et al., 2013; Flammer, 2013) in which negative events had the most pronounced impact on abnormal stock price returns. The implication in this finding is that while the “reward” for positive CSR events is negligible, the “penalty” incurred for negative CSR events is very real. This is also aligned with anecdotal evidence that firms invest in supplier responsibility in order to avoid potential negative events (Wal-Mart, 2009; Hajmohammad & Vachon, 2016). This provides us with confidence in interpreting the findings garnered from analysis of the negative sample.

Although we see limited stock-price benefits occurring as a result of positive CSR events, it is possible firms may see “positive” CSR initiatives as preventative measures – a necessary long-term investment designed to help a firm to avoid potential negative events. Varadarajan, (1985) champions the benefits of “failure preventing” strategies as or nearly as important as “success producing” variables. Disney recently avoided what
could have been a particularly costly headline risk event when they pre-emptively chose to remove all textile manufacturing from Bangladesh due to perceived levels of safety and oversight in the Bangladeshi textile manufacturing industry (Palmeri & Rupp, 2013). Although the production costs were lower in Bangladesh, Disney believed they would be exposing themselves to too much risk to continue to viably source from the country. Mere months after this decision was announced, the tragic Rana Plaza factory collapse occurred, and a slew of firms that had continued to produce clothing in Bangladesh were on the hook for millions of dollars in lawsuits and factory audits (Greenhouse & Clifford, 2013). However, the benefits of these failure-preventing initiatives are not always so apparent.

One of the issues with our analysis of positive events is that there were very few positive external events. Of the 1,800 positive social events, a mere 54 of them were classified as external. However, this does not indicate that there were not any positive supplier-related events in our sample. There were 434 “supplier development” events in the positive event sample. These are events in which it was reported that the buying firm had undertaken some sort of initiative or development program aimed at improving supplier CSR performance and/or capabilities. It is important to remember that many of the events in the sample are reported by the firms themselves rather than by a third party, particularly those reported via Business Wire. It is possible that firms are more likely to report the efforts that they themselves make to improve supplier performance than to report on what suppliers are doing autonomously on their own. While it is outside the scope of this research, an empirical examination of the market impact of supplier development programs and initiatives would provide unique and valuable managerial
implications. This possibility, along with some early analysis on the topic, is discussed in greater levels of detail in the Future Research section.

**Managerial Implications.**

The findings of this research carry clear managerial implications for any firms that are engaged in some form of outsourcing. The quote from the senior Nike executive, in which it was stated that supplier actions “were not their responsibility” (Locke & Ramis, 2012) seems to have been very much a product of its time. When the statement was made in the early 1990’s, the stock market was still applying a “supplier discount” in their reactions to supplier CSR violations. The findings of this study suggest that while this discount existed in the 1990’s (and most likely before then as well), by the turn of the millennium it had essentially vanished.

Clearly the types of supplier violations that firms are most penalized for are those that have to do with safety. Firms should take all necessary measures to avoid these incidents. Although many firms have taken measures to attempt to curtail these incidents, they still occur all over the world. Honda and Toyota recently recalled over 60 million airbags around the world due to a design issue attributed to their supplier Takata (Abrams, 2016). Airbags provided by Takata were inflating prematurely and caused at least 11 deaths and over 100 injuries from 2008 to 2016. This led to the recent recall which caused Honda to miss their expected earnings by over one billion dollars (realizing a Q1 loss of $860 million) in the first quarter of 2016 (Yamakazi, 2016). By failing to completely assure the safety of these airbags, automakers exposed their customers to potential injury and cost themselves a significant amount of money.
One method many firms have turned to is on-site supplier auditing. In the wake of an expired meat scandal McDonald’s changed their supplier management policy to include a number of unannounced on-site audits. In addition to this, McDonald’s also installed a hotline number that whistleblowers could use to report any future product safety violations (Jourdan, 2014). Apple promised to perform similar audits following the rash of suicides at their supplier Foxconn in 2011 (Schroeder, 2013). These findings also give credence to Varadarajan’s (1985) assertion that firms are better off allocating their resources towards failure or risk prevention as opposed to the achievement of gains. As firms are generally risk-averse (Tversky & Kahneman, 1991; Rabin, 2000), this should actually be a fairly intuitive concept. These audits provide an additional benefit in that they allow firms to ensure that exceedingly poor (and potentially dangerous) conditions are not present at supplier facilities. This is particularly important given the finding that safety incidents that appear to be a product of abhorrent or dangerous working conditions are penalized more harshly than any other category of CSR event.

In Williamson’s classical work on outsourcing and transaction cost economics (TCE), the primary question a firm had to answer was whether or not the savings realized through outsourcing would outweigh the necessary governance costs (Williamson, 1971). Firms must answer this same question when considering whether or not the cost added by supplier audits is worth the savings provided through outsourcing. The difference between the situations examined in this research and Williamson’s example is that the primary governance risk is not opportunism, but rather exposure to CSR risk. When faced with this decision, it is highly likely that firms will continue to outsource. In light of that decision they would have a clear economic responsibility to shareholders to ensure
they are taking the necessary governance steps. Before buying firms were held responsible for the actions of their suppliers, firms may have thought it possible to take advantage of the cost savings while shirking their responsibilities in terms of governance costs. In the new era of supply chain accountability indicated by the findings of this research, this type of “free ride”, in which firms could take advantage of lower production costs while avoiding the costs necessary to ensure supplier responsibility, is effectively over. While there are no clear rewards for positive events, it is still economically advisable for companies to pursue them as long as they are preventative measures aimed at avoiding costly negative safety events.

The importance of supplier CSR performance has been shown to have grown in the minds of consumers (Hartmann & Moeller, 2014) and stakeholders (Reuter et al., 2010), and with the findings of this dissertation it is now clear what its importance is relative to CSR performance at the focal firm. Firms may be hesitant to allocate resources to supplier development programs if they are not certain that it will maximize the “bang for their buck”. If it is believed that issues at the focal firm weigh more heavily on the minds of investors than issues at suppliers, firms may be hesitant to use precious resources to improve CSR performance within the supply chain. However, our findings suggest that issues at the focal firm do not outweigh supplier issues. Instead this dissertation provides firms with a clear financial rationale to devote scarce resources to improving the CSR performance throughout their supply chain – particularly towards preventing negative safety issues.

As mentioned in Chapter 4, the sample contains 694 firms across 67 2-digit SIC codes. Multi-level hierarchical analysis failed to reveal any significant industry effects
on stock price evaluations. This indicates that our findings are not industry-specific and are somewhat generalizable across multiple industries. Though they were not explored in this research, future studies might consider analyzing other firm-specific variables in order to check for other firm-level factors that may partially determine any differences in the penalties associated with negative external CSR events. For instance, Bose and Pal (2012) find evidence suggesting that the market’s reaction to green supply chain management initiatives is at least somewhat governed by firm-level R&D expenditures. If a theoretical reason can be found, similar firm-level variables should be collected and analyzed in a future study using this sample.

The managerial implications of this study are clear. In 2016 firms can no longer make the case that allocating resources to improve supplier CSR performance is an unnecessary use of funds. This study demonstrates the very real financial consequences that can befall firms that neglect the CSR policies and capabilities of their suppliers. The time of credibly distancing oneself and avoiding any financial consequences would appear to be over. Where devoting significant resources to improving supplier CSR performance may once have been considered to be discretionary, at this point it can only credibly be considered to be necessary.

**Contribution to Academic Literature**

Over the last decade of sustainable supply chain research, authors have called for an empirical study detailing the effects of supplier sustainability and CSR performance on the financial performance of their customers (Carter & Rogers, 2008; Carter & Easton,
2011; Busse, 2016). This research, at least partially, answers that call. The vast majority of supply chain sustainability and CSR research has been either theoretical (Carter & Rogers, 2008; Sarkis et al., 2010; Rogers, 2011; Golicic & Smith, 2013; Markman & Krause, 2016), or while empirical, anecdotal in scope (Price, 1995; Zhu & Sarkis, 2004; Idemudia & Ite, 2006). The generalizable, empirical supply chain CSR research that has been carried out did not address the dichotomous buyer-supplier relationship in the way this dissertation does (Jacobs et al., 2010; Bose & Pal, 2012; Wang et al., 2014; Wang & Chen, 2015). This dissertation is the next logical step in this stream of research. With this study we extend empirical CSR supply chain research to a broader, more generalizable level. While this dissertation may be an important step, there are still many more steps that must be taken for the role of CSR in supply chains to be truly understood. The event study methodology is a powerful tool through which future empirical research in this area might be conducted.

In one aspect this dissertation acts as a companion piece to Hartmann and Mueller (2014). Where they observed the changing attitudes of customers towards supply chain sustainability, we observe those of the stock market. Our findings demonstrate an interesting symmetry between the preferences of the market and the preferences of customers. These findings confirm the notion put forth by Pagell and Shevchenko (2014) that all supply chains must now be sustainable supply chains – setting out clear empirically-obtained parameters for why that paradigm shift has occurred.

This research also complements Flammer’s (2013) examination of environmental events over time by using similar methods to examine social and environmental events more granularly over a similar period. We shift the focus to social events and also add
the nuance of comparing the impact of events occurring at a supplier relative to the focal firm. Interestingly, our findings differ somewhat from Flammer’s, suggesting an underlying difference in the way the market reacts to social relative to environmental events.

In Flammer’s (2013) study of market reactions to environmental events, she finds that reactions to negative events get “more negative” while reactions to positive events get “less positive” over time. This is not the case in this research. While the market’s reaction to negative external events became more negative over time, we find no evidence of this with negative internal events. Similarly, we find no evidence that market reactions to positive events became “less positive” over time. In our analysis positive events never garnered any type of significant reaction in either direction. A potential explanation for this disparity is that the expectations for what constitutes acceptable environmental performance have shifted more than the expectations for social performance over the duration of this study. It is possible that the shift detected by Flammer indicates that the “reference point” for what is considered to be the baseline level of performance has moved further for corporate environmental behavior than for social behavior (Tversky & Kahneman, 1974). One explanation could be that Flammer’s period of study begins 14 years earlier than ours. Alternatively, it is reasonable that markets were concerned with areas such as safety before they were concerned with emissions. In order to truly understand these differences further research must be carried out to contrast the market’s reaction to environmental events over time to the reactions to social events reported here. 11,134 potential environmental events have already been
identified and are currently being coded by the authors for the purpose of making such a contrast in a subsequent study.

Finally, we make a novel contribution by incorporating psychological distance theory into supply chain management research. Psychological distance is a fairly pliable theoretical lens with many potential applications in evaluating buyer-supplier relationships.

**Potential Limitations.**

There has been a movement away from the traditional view of the supply chain as a series of dyadic buyer-supplier relationships. Some scholars have argued that the supply chain should be discussed in units no smaller than the triad (Wu & Choi, 2005; Choi & Wu, 2009). Other research has framed the supply chain as a network in which relationships form and decay (Choi et al., 2001; Lazzarini et al., 2001; Li & Choi, 2009). Carter, Rogers, and Choi (2015) recently presented a theory of the supply chain, in which they paint the supply chain as a network of links and nodes. While we agree with many of the stances taken in these works, and agree with the assertion that the supply chain is more than the classic dyadic relationship, we only used the dyad as our unit of analysis when considering the effect of external supply chain events on the stock price of a firm. We limit ourselves to dyadic relationships for a number of reasons. First, while it has been shown that firms are affected by the sustainability performance of their immediate suppliers (Hartmann & Moeller, 2014), there has been little empirical evidence that they are affected by suppliers further up the chain. Additionally, one of the weaknesses of the event study methodology is that it is difficult to parse out the factors beyond the event in
question that may be affecting stock prices (McWilliams & Siegel, 1997). Dealing with the additional noise inherent to a multiple intermediaries between firms may prove difficult. Thus when exploring the difference in the financial impact between events that happen inside or outside the firm, we limit our scope to dyadic relationships with first-tier buyers or suppliers.

As always in event studies, there is a possibility of not capturing every possible event in our sample (Binder, 1998). However while it is possible (or even perhaps likely) that we did not capture every potential event, we are confident that the identified sample is sufficient for the purposes of this research. The search terms (reported in Appendix A) were broad and informed by the literature. At over 10,000 potential articles and 2,301 used, the size of our sample appears to have been fairly robust, comparing favorably to recently published event studies (Hendricks, Singhal, & Zhang, 2009; Hendricks & Singhal, 2013; Flammer, 2013). Furthermore, any keywords or articles that were incorrectly omitted would actually reduce the power of our tests. If anything, we have underestimated the significance of our results.

Abnormal stock price shifts are not likely to capture the full economic impact of their associated CSR events. It is possible, if not likely, that some of the abnormal returns in our study either over or underestimate the actual financial impact of the events in our sample. Despite these potential limitations, abnormal returns *do* offer value as they can act as a proxy variable by which the myriad event types present in this study might be compared. The research goal of this work was not to identify the precise financial impact of a given CSR event type. Instead we wished to draw contrasts between the event types so that we might observe the comparative financial impact of the different types of CSR
events in our sample. Abnormal returns are effective for the purposes of this research in that they allow for apples-to-apples comparisons to be drawn between multiple event types. They likely do not, however, capture the full economic or reputational impact of the events in question.

**Future Research.**

Hajmohammad and Vachon, (2016) champion the benefits of a buyer and supplier collaborating on a sustainability risk-mitigation strategy. The authors position this as a viable strategy in cases in which the supplier possesses a serious potential for sustainability risk. Our findings suggest that supplier development programs and initiatives may have been seen as an unnecessary waste of resources between 1994 and 2007 (our first two periods were combined due to lack of supplier development events in our first period), garnering a significant negative abnormal stock fluctuation. A potential explanation for this is that in the first two periods of the sample supplier development was considered a discretionary expenditure. Analysts may have considered the potential opportunity cost of devoting resources to supplier CSR development and concluded that it was a sub-optimal use of resources that may not materially effect firm value. This does not appear to be the case in the period from 2008-2013, in which the market’s reactions to supplier development initiatives are not different than its reaction to internal positive events. The shift in market reactions to positive supplier CSR development initiatives over time is captured in Figure 3.
These findings warrant further analysis. Based on how few positive internal events existed in our sample relative to positive supplier development events, it may be that this is the way to compare the differing financial effects associated with positive internal and external CSR events. The key difference here is that whereas with negative events we focused on the place the event originated, with positive events we would instead focus on where the effects of the event were directed.

As mentioned above, there was some surprise due to working conditions events not distinguishing themselves from diversity or philanthropy events. The potential reasons of why this may have occurred were detailed in section 6.1 above. Although no significant difference was found between our overall sample of working condition events and the diversity and philanthropy events, there were significant differences with the events containing both elements of working conditions and safety. This indicates that
there may be a certain class of working condition events that *does* significantly impact share prices. What these specific classes are might be identified by breaking event “types” down into even more granular levels in an opportunity for future research.

The subsets of social events presented in Figure 1 are not meant to be the final level of granularity to which social events should be explored. Rather, they are a starting point to be used by future researchers who wish to explore the financial effects of different types of social events on an increasingly granular level. For instance, safety events might be separated into events that lead to injuries relative to events that ended in death. Diversity events could be drilled down to account for initiatives aimed at gender, race, or sexual preference. This would allow researchers to move from the cross-type comparisons of this research, down to within-type comparisons in future work.

Although our sample includes 2,301 data points, it does not include 2,301 unique “events”. This is because there are some events in which multiple firms are implicated on the same trading day. For instance, over 30 firms were involved in the Rana Plaza factory collapse of 2013. Interestingly, Jacobs and Singhal, (Forthcoming) find that the firms implicated in the disaster, in which 1,133 fatalities occurred, suffered no significant abnormal stock returns. A potential reason for this may be that when compared to the actions of the other firms implicated in the scandal, no one firm seemed to have done anything particularly egregious. As discussed earlier in this dissertation, evaluators tend to weigh events against some sort of reference point or absolute zero (Tversky & Kahneman, 1974). When a single firm is implicated in an event, they are likely compared against all of the firms that were *not* implicated in said event. In this case the offending firm would appear to be an outlier that committed an offense consistent with a level of
performance that is notably lower (or higher) than any other firm. However, in an event such as the Rana Plaza collapse, when so many major apparel firms were implicated in the incident, what were market analysts to compare this behavior to? It is possible that the abundance of firms that were involved with this scandal “shared” the blame among themselves, protecting any one of them from taking a significantly negative hit on the day of and the days immediately following the event. Many of the events in our sample involve two or more firms. It is possible this “blame sharing” exists in our sample. A hierarchical regression, in which firms are clustered around the event in question, may shed some light on the counterintuitive findings of Jacobs and Singhal.

Jayachandran et al. (2013) incorporate the Kinder, Lydenberg, and Domini (KLD) database as a moderating variable in their study of firm social performance. The KLD index is a capitalization weighted index measuring the exposure risk to hundreds of firms with regard to environmental, social and governmental risk. KLD essentially functions as a CSR reputation index. A firm’s reputation for social and environmental practices could be reasonably expected to have an impact on the stock market’s reaction to the CSR events captured in our data. Ratings are given to all firms included in the database that essentially act as measures of corporate reputation. While we found no evidence of industry-level variance in our analysis, it would be fascinating to explore the potential of firm-level variance, particularly in regards to their reputation.

While we could have run analysis looking solely at the firm in question as a grouping variable, this would not provide the same explanatory power as that same firm’s reputation. Take for example Mattel. Prior to the 2007 ’s lead paint scandal, Mattel was considered to be a standard bearer for supply chain responsibility in the toy industry.
(Teagarden, 2009). After the scandal this was no longer the case, with Mattel often being held up as a cautionary tale rather than as an aspirational model (Quelch, 2007). The effect of an event being a “Mattel” event would be inconsistent over the full range of our sample; this would be corrected via the use of the KLD data. Beyond the potential contribution of explanatory power offered by this index, incorporating KLD will allow us to contrast the impact of both long and short-term CSR performance on market reaction. Will a firm with a strong reputation for social responsibility suffer a more significant market penalty for a negative event because it is so unexpected? Or might the inverse be true? Will the market “take it easy” because they view the event as an anomaly? Incorporating this CSR reputation variable will add a level of richness to this project that will provide valuable insight for both practitioners and academics.

In this dissertation we have only considered supplier CSR performance, not supplier sustainability. We define sustainability using Elkington’s (1998) classic triple bottom line model, in which satisfactory levels of environmental stewardship, social responsibility, and economic success must be achieved for a firm to be considered sustainable. This research has demonstrated how the priorities of investors have changed regarding supplier social performance. It would be fascinating to understand how the ways in which investors prioritize the three elements of the triple bottom line relative to one another has changed over the same period. Rogers, Carter and Kwan (Working Paper) find that managers prioritize supplier development initiatives that emphasize improvements in either supplier financial savings or worker safety over those emphasizing environmental performance improvements. An event study comparing all three elements of the triple bottom line would allow researchers to make comparisons...
between the priorities of the managers surveyed in that study and the priorities of financial markets. Additionally, such a study would allow for analysis of how these priorities may have shifted over time, shedding further light on the ongoing evolution of supply chain sustainability.

Porter and van der Linde (1995) discussed the impact of green performance on the bottom line over 20 years ago. Subsequently, environmental performance has often included in discussions of supply chain CSR or sustainability performance as well as in the triple-bottom line definition of sustainability (Elkington, 2008; Carter & Rogers, 2008). The link between the financial performance of a firm and the environmental performance of their suppliers has been widely studied in the supply chain literature (Beamon, 1999; Bai & Sarkis, 2010; Bose & Pal, 2010). A logical follow-up to this dissertation is to collect environmental events over the same period studied here and perform an additional round of analysis. As mentioned in Chapter 1, this follow-up study is already underway. Below, I will present the initial hypotheses, as well as preliminary analysis of one of the four environmental sub-categories we will be exploring in this research.

**Future Environmental Event Study**

As mentioned in Chapter 1, a study focused on the difference in market impact of internal relative to external environmental events is currently underway. As this study acts as a follow-up to, and quite literally the future research of, this dissertation it will be briefly described here. Similarly to event studies carried out on social performance, those conducted on environmental performance have also taken a broader, more generalized view of what constitutes an environmental event (Klassen & McLaughlin, 1996; Jacobs
et al., 2010; Bose & Pal, 2012; Wang et al., 2014). Many of the significant categories utilized in past environmental event study research revolve around firm size, industry, or R&D expenses. While some of these studies have divided events into categories such as types of environmental awards or initiatives (Jacobs et al., 2010), they have not used the specific event types we will use to divide our events in this research. In order to divide environmental events into more granular categories we again turn to a framework developed with suppliers in mind. We adopt Handfield et al.’s 2002 framework of supply chain environmental performance. Handfield and his colleagues divide environmental performance into six distinct metrics that have been widely-studied in CSR research. The performance elements we will include from this framework are waste management (Hart, 1995; Klassen & McLaughlin, 1996); resource management (Guang Shi et al., 2012); packaging and reverse logistics (Tibben-Lembke & Rogers, 2002; Vachon & Klassen, 2006; Lee & Lam, 2012); labeling and certification (Rondinelli & Vastag, 2000; Klooster, 2005); and environmental programs (Sarkis, 2003; Humphreys et al., 2003; Vachon & Klassen, 2008). The sixth pillar of Handfield et al.’s framework is regulation (Walker et al., 2008; Zhu & Sarkis, 2007). However because the scope of this research concerns the types of performance that firms are able to control or influence, regulation will be dropped as a category in our analysis. As regulation occurs at a macro industry or country-level, it is unlikely that individual firms are able to influence it significantly. The categorization we will be employing for environmental events is presented below in Figure 4.
Hierarchy of Environmental Event Types.

Although the environmental framework used in this dissertation is largely inspired by Handfield et al. (2002), there are a few key differences. Handfield et al.’s framework included product attributes and compliance to government regulation (2002). These were both dropped from our analysis. Product attributes specifically apply to internal design and manufacturing processes. Many of the reports that are made available, such as Chipotle moving to non-GMO pork (Alesci & Gillespie, 2015), are disclosed by the firm in question. As such they are likely to skew positive. More importantly, it seems they will be few and far between. It is unlikely that changes to proprietary manufacturing practices will be widely reported in popular press. The “compliance with government regulation” category was also dropped. The managerial motivations behind this research
are to inform firms on where they might best spend their scarce resources. Firms cannot enact government regulation, they can merely respond to it. Comparing the impact of regulation to other environmental events would not add any insight into how firms should allocate their resources. Beyond this, the way regulation is reported is not generally firm-specific, and guessing which firms are effected by new regulation is fraught with potential methodological issues.

In their place we include resource management events. We define resource management events to include an unsustainable use of resources. This may include resource-depletion events such as deforestation (Scherer & Palazzo, 2011; Jorgensen et al., 2012) or overfishing (Lai & Yu, 1995; Punt & Methot, 2005). We retain Handfield et al.’s constructs of waste management, environmental certification, environmental initiatives, and packaging and reverse logistics.

We classify waste management as any situation dealing with the disposal of run-offs such as water waste (Lambooy, 2011), carbon emissions (Sundarakani et al., 2010; Lee, 2011), nuclear waste (Beamon, 1999), spills, or other pollutants. Event studies have been used in the past to study waste management, Hamilton (1995) shows evidence of negative stock price reactions to pollution reports. Generally, if waste management is considered in other event studies it is generally as part of a broad group of environmental events, often revolving around awards or initiatives (Klassen & McLaughlin, 1996; Jacobs et al., 2010).

Environmental initiatives include the implementation of any program designed to improve the environmental performance of a firm or its suppliers. To the best of our knowledge this is the CSR category that has been studied most extensively using event
studies. Event studies in this area include those of Klassen & McLaughlin (1996); Jacobs et al. (2010), Bose and Pal (2012), and Jayachandran et al. (2013). We define certification as the awarding of either industry or governmental awards, including ISO 14000 or LEED certification. Previous event studies analyzing CSR certification suggest that for the right type of company, industry awards and certifications can have a significant stock price impact (Lo et al., 2010; Wang & Chen, 2015). The final environmental event category is packaging and reverse logistics. We define this as any mention of sustainable packaging or product disposition (Tibben-Lembke, 2002).

Generally, recalls are a function of reverse logistics, not a product of them. Recalls stemming from a product defect or safety problem aren’t logistics issues, they’re product issues. For this reason counting recalls as reverse logistics events would be a mis-categorization.

Behavioral finance heuristics are also applied to construct the hypothesized hierarchy of environmental event stock price impacts. Unit saliency (Whittlesea, 1990; Shiller, 2001) is the primary theoretical lens used in this endeavor. The comparison of waste and resource management is difficult because both event types involve the destruction or depletion of natural resources. This is hard to obscure from the eyes of stakeholders as generally the areas being polluted or depleted are out in the open. Different resource and waste management events may have more significant market impacts than one another, but for the most part we expect events from these two performance categories to be fairly similar. We also expect that the impact of these event types will have the most significant impact of all social event types.
Evaluators may find things like carbon emissions, felled trees, or spilled oil more tangible than certifications or the impact of green labels. Anecdotally, in the aftermath of the Deepwater Horizon oil spill in 2010 BP’s stock price dropped from $59.50 to $28.90 (Chamberlain, 2014). While this was partially a function of increased awareness due to the wall-to-wall media coverage of the event, it is also likely that the ability to visualize millions of gallons of oil pouring into the Gulf of Mexico enabled evaluators to bypass any mental discounting associated with less tangible events. Consideration of environmental performance is often tied to the natural environment in the public eye (Bickerstaff & Walker, 2001; Kelly & Fussell, 2015); resource and waste management events are the event types most tied with traditional ideas of pollution and emissions. Furthermore, as these types of environmental issues have been in the public consciousness for many years (Murch, 1971), evaluators are more familiar with them. Tversky and Kahneman’s (1973) availability characteristic demonstrates that when evaluators are able to more easily recall more events of a certain type they assign a higher value to them as well as a higher probability of occurring again. If market evaluators are more familiar with the waste and resource management aspects of environmental performance (as we expect them to be), than they will assign a larger impact to them relative to other types of environmental events types.

Hypothesis: Resource management and waste management events will have a larger market impact than introduction of environmental programs, certification, or packaging and reverse logistics events.
Studies have revealed the positive association of ISO certification and financial performance (Lo et al., 2012), as well as positive market impact upon certification (Jacobs et al., 2010). However, Jacobs et al. (2010) find the impact of other firm-initiated environmental initiatives is negligible, with some even leading to a negative market reaction. This is likely in part due to initiatives being seen as “discretionary” or going above and beyond the actions that add value for the shareholders (Friedman, 1962). The processes behind certifications like ISO14000 are transparent, giving them a cache of legitimacy beyond that of firm-initiated programs or awards.

Hypothesis: Certification Events will have a larger market impact than environmental programs or packaging and reverse logistics.

As mentioned above, voluntary environmental programs may be seen as discretionary actions on the part of firms. Bose and Pal (2012) find that the market impact stemming from the announcement of green supply chain initiatives can vary depending on the attributes of the firm making the announcement. One of the primary findings of Flammer’s (2013) study was that the impact of positive environmental event announcements are “expected” and seen more as fulfilling an obligation than as something that will provide any type of competitive advantage. It is likely that shifting reference points are again at play here (Epley & Gilovich, 2006). Finally, it may be easier to conceptualize the business impact of recalls, returns and product life-cycle programs or events than the impact of environmental programs. This is partially due to unit saliency (Whittlesea, 1990) and partially due to future discounting (Camerer et al., 2004), which
states that evaluators put more weight on events that will occur soon or have occurred, particularly in comparison to events that may happen in the future. Environmental initiatives promise improvements or issue avoidance in the future, causing evaluators to discount whatever their true impact might be.

*Hypothesis: Packaging and reverse logistics events will have a larger market impact than environmental program events.*

The methods being used to analyze environmental events precisely mirror those used to analyze social events in the main body of the dissertation. Key search terms were identified (and are presented in Appendix A below) and used to search Business Wire and Wall Street Journal articles from January 1, 1994 to December 31, 2013. This search yielded over 11,000 potential environmental performance events. As of this writing, only the 4,164 potential waste management events have been fully analyzed and contrasted with the social events that are the focus of this dissertation. Of these 4,164 potential waste management events, 774 were identified as being relevant. Confound checks revealed that 569 of these potential events (of which 92 were negative) were acceptable for use in this analysis.

T-tests indicate a significant difference in the impact of negative and positive waste management event announcements in all four windows of interest. Positive and negative means, as well as the t-value for each comparison, are displayed in Table 12.
Hypotheses contrasting waste management events to the other environmental event types will be tested when the remaining 7,000 environmental events are analyzed. However, analyses were carried out juxtaposing waste management events with the social event types discussed in the body of the dissertation. The results are similar to what was observed in the analysis of this dissertation, with safety events and events combining elements of both safety and working conditions having a greater level of impact than any of the other event types – including waste management events and events that combined elements of both waste management and safety. These comparisons are presented in Table 13.
Interestingly, the waste management events skewed further towards the most recent time period than did any of the samples of CSR events presented in the dissertation. While 511 waste management events were announced in period 3, only 33 events occurred in period 1, and only 26 in period 2. The small sample sizes of waste management events in periods 1 and 2 limit our ability to test our temporal hypotheses until the remaining environmental events can be included in our analysis. A similar sample-size issue also exists in the split between internal and external announcements, with only 17 announcements falling into the latter category.
When completed, analysis of environmental events will also be combined with the research presented in the body of this dissertation. The combination of these two projects will generate a more comprehensive picture of the effects of both social and environmental events in the supply chain and over time. Of particular interest will be the comparison across social and environmental events. It will be value-adding for firms to know if they would be better off, from a firm value perspective, allocating resources to issues like factory-improvement initiatives or reduced/green packaging rollout. The initial, preliminary analysis presented here indicates that the market considers CSR events to be more impactful than environmental events. The collection and analysis of additional environmental data, representing the other three event types laid out above, will be required to validate this finding.
REFERENCES


http://doi.org/10.1509/jm.12.0179


Leigh, J. (2012). *Numbers and Costs of Occupational Injury and Illness in Low-Wage Occupations* (pp. 1–18). Davis, California: Center for Poverty Research, and Center for Health Care Policy and Research, University of California Davis.


Grand Rapids MI: Zondervan Publishing.


http://doi.org/10.1002/pa.232


<table>
<thead>
<tr>
<th>Category</th>
<th>Search Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversity</td>
<td>(&quot;corporate diversity&quot; OR &quot;labor diversity&quot; OR &quot;corporate discrimination&quot; OR</td>
</tr>
<tr>
<td></td>
<td>&quot;corporate discrimination suit&quot; OR &quot;corporate diversity policy&quot; OR &quot;employee</td>
</tr>
<tr>
<td></td>
<td>discrimination&quot; OR &quot;minority supplier&quot; OR &quot;supplier diversity&quot; OR &quot;diversity</td>
</tr>
<tr>
<td></td>
<td>program&quot; OR &quot;diversity initiative&quot; OR MBE OR MWBE)</td>
</tr>
<tr>
<td>Working Conditions</td>
<td>&quot;slave labor&quot; OR &quot;forced labor&quot; OR &quot;labor conditions&quot; OR &quot;child labor&quot; OR</td>
</tr>
<tr>
<td></td>
<td>&quot;working conditions&quot; OR &quot;worker's rights&quot; OR &quot;employee rights&quot; OR &quot;employee</td>
</tr>
<tr>
<td></td>
<td>conditions&quot; OR sweatshop</td>
</tr>
<tr>
<td>Safety</td>
<td>employee safety&quot; OR &quot;worker safety&quot; OR &quot;employee injuries&quot; OR &quot;worker</td>
</tr>
<tr>
<td></td>
<td>injuries&quot; OR &quot;employee injury&quot; OR &quot;worker injury&quot; OR &quot;employee death&quot; OR &quot;</td>
</tr>
<tr>
<td></td>
<td>employee fatality&quot; OR &quot;worker death&quot; OR &quot;worker fatality&quot; OR &quot;safety program&quot;</td>
</tr>
<tr>
<td></td>
<td>OR &quot;safety initiative&quot; OR &quot;customer safety&quot; OR &quot;customer death&quot; OR &quot;</td>
</tr>
<tr>
<td></td>
<td>customer fatality&quot; OR &quot;customer injury&quot; OR &quot;customer injuries&quot; OR &quot;factory</td>
</tr>
<tr>
<td></td>
<td>&quot;factory collapse&quot; OR &quot;store fire&quot; OR &quot;store collapse&quot; OR &quot;employee suicide&quot;</td>
</tr>
<tr>
<td></td>
<td>OR &quot;worker suicide&quot; OR &quot;factory explosion&quot; OR &quot;facility explosion&quot; OR &quot;</td>
</tr>
<tr>
<td></td>
<td>factory collapse&quot; OR &quot;factory explosion&quot; OR &quot;facility collapse&quot;</td>
</tr>
<tr>
<td>Philanthropy/Community</td>
<td>&quot;corporate philanthropy&quot; OR &quot;corporate gifts&quot; OR &quot;Community Involvement</td>
</tr>
<tr>
<td>Outreach</td>
<td>initiative&quot; OR &quot;community outreach program&quot; OR &quot;supplier community outreach</td>
</tr>
<tr>
<td></td>
<td>program&quot; OR &quot;corporate outreach initiative&quot; OR &quot;corporate giving&quot; OR &quot;</td>
</tr>
<tr>
<td></td>
<td>corporate charity&quot; OR &quot;supplier outreach initiative&quot; OR &quot;supplier giving&quot; OR</td>
</tr>
<tr>
<td></td>
<td>&quot;supplier charity&quot; NOT &quot;annual report&quot; NOT &quot;quarterly report&quot;</td>
</tr>
<tr>
<td>Waste Management</td>
<td>hazardous waste OR radioactive waste OR toxic waste OR water runoff OR</td>
</tr>
<tr>
<td></td>
<td>nuclear waste OR manufacturing waste OR hazmat pollution OR contamination OR</td>
</tr>
<tr>
<td></td>
<td>spill OR carbon emissions OR pollutant OR deforestation OR water contamination</td>
</tr>
<tr>
<td>Resource Management</td>
<td>corporate deforestation OR overfishing OR hazardous materials</td>
</tr>
<tr>
<td>Packaging/Reverse</td>
<td>packaging reduction OR green packaging OR sustainable packaging OR harmful</td>
</tr>
<tr>
<td>Logistics</td>
<td>packaging OR toxic packaging OR recycling program OR supplier remanufacturing</td>
</tr>
<tr>
<td></td>
<td>OR remanufacturing program OR returns program OR Green AND product disposition</td>
</tr>
<tr>
<td>Labeling/certification</td>
<td>supplier certification OR LEED certification OR ISO 14000 certification OR green certification OR environmental certification OR green label OR environmental label OR supplier award OR preferred supplier OR industry award OR supplier award OR environmental award</td>
</tr>
<tr>
<td>------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Environmental Programs</td>
<td>Supplier development AND environment OR supplier mission statement AND environment OR environmental initiative OR green initiative</td>
</tr>
</tbody>
</table>