Does It Pay to Recall Your Product Early? An Empirical Investigation in the Automobile Industry

Meike Eilert (meike.eilert@unl.edu)
Assistant Professor
College of Business Administration
University of Nebraska-Lincoln
Lincoln, NE

Satish Jayachandran*(satish@moore.sc.edu)
James F. Kane Professor of Business
Darla Moore School of Business
University of South Carolina
Columbia, SC

Kartik Kalaignanam (kartik.kalaignanam@moore.sc.edu)
Associate Professor
Darla Moore School of Business
University of South Carolina
Columbia, SC

Tracey A. Swartz (Tracey.Swartz@grad.moore.sc.edu)
Doctoral Candidate
Darla Moore School of Business
University of South Carolina
Columbia, SC

Forthcoming, Journal of Marketing

*Contact Author. The first three authors contributed equally.

Does It Pay to Recall Your Product Early? An Empirical Investigation in the Automobile Industry

Abstract

Defective products are often recalled to limit harm to consumers and damage to the firm. However, little is known about why the timing of product recalls varies after an investigation is opened. Likewise, there is little evidence on whether stock markets care about recall timing. This study tests the effect of problem severity on time to recall and the role of brand characteristics in moderating this relationship, and the stock market impact of time to recall. The hypotheses are tested on a sample of 381 recall investigations in the automobile industry between 1999 and 2012. The results show that while problem severity increases time to recall, this relationship is weaker when the brand under investigation has a) a strong reputation for reliability and b) experienced severe recalls in the recent past. However, the relationship between problem severity and time to recall is stronger when the brand is diverse. Importantly, the results reveal that recall delays are punished by stock markets. The study suggests that time to recall has significant implications for managers and policy makers.

Keywords: Product recalls, time to recall, brand reliability, brand diversification, stock market performance.

Defective products affect the physical safety of consumers and expose manufacturers to liability claims, fines, and loss of reputation. Consequently, defective products are often recalled to limit damage to consumers and firms. In the United States, the Consumer Products Safety Commission (CPSC) reported a total of 390 recalls in 2014, ranging from dishwashers to toys and cribs.[1] In the U.S. automobile industry, the National Highway Transportation and Safety Agency (NHTSA) has overseen recalls involving hundreds of millions of vehicles (Rupp and Taylor 2002).

Typically, when a product is suspected of defects, a government agency can initiate an investigation. While many investigations end with the product being cleared of suspected defects, a significant number of investigations also culminate in a recall. The decision of whether and when to recall is not a simple one. Recalls are costly; announcing and implementing one is associated with both direct costs in repair, restitution, or liability and indirect costs such as losses in reputation and market value (Chen, Ganesan, and Liu 2009; Hora, Bapuji, and Roth 2011). Consequently, recalls could have a devastating impact on a firm's performance, sometimes even threatening its survival. Thus, a firm has reasons to avoid a quick recall and instead wait for the investigation to conclude.

However, delaying a product recall may lead to higher direct and indirect costs through fines, liability damages, and most importantly, diminished reputation (Tang 2008). In 2012, Toyota was fined $17.35 million dollars for delaying a floor mat recall (Koyitty 2012). The U.S. Department of Justice fined General Motors $900 million for willfully delaying the recall for a faulty ignition switch and, thus, defrauding customers (Isidore and Marsh 2014). Hence, though recalls are adverse events in general, a quick response may attenuate the damage (Dani and Deep 2010), and enhance consumer welfare through improved product safety and performance. In

[1] https://www.cpsc.gov/Global/About-CPSC/Budget-and-Performance/CPSCFY2014AFR.pdf

effect, the recall behavior of a firm, especially the timing of recall, is important because it has legal and financial consequences for the firm and economic and safety-related consequences for consumers (Marucheck et al. 2011). The considerable variation in time taken by firms to recall after an investigation is initiated (Wieder 2011), raises important questions for managers and policy makers.

Why do some firms recall earlier than others after an investigation is opened? Do stock markets respond to the timing of recalls? The objective of this research is to investigate the variation in and the performance consequences of time to recall. We define time to recall as the time elapsed between the opening of an external, formal defect investigation and the announcement of a recall by the firm. The manuscript makes two contributions to marketing theory and practice. *First*, to date, no study to our knowledge has examined the factors that influence the timing of product recalls. The product recall literature in marketing, as summarized in Table 1, has predominantly focused on the consequences of recalls for the firm's bottom line (e.g., Borah and Tellis 2016; Chen, Ganesan, and Liu 2009; Cleeren, van Heerde, and Dekimpe 2013; Liu and Shankar 2015; van Heerde, Helsen, and Dekimpe 2007) and their ability to learn and prevent future recalls (Haunschild and Rhee 2004; Thirumalai and Sinha 2011), while little attention has been paid to the actual recall behavior of the firm. By examining the time to recall, our study offers insights on when firms are likely to respond swiftly to defects that are under external investigation.

--- Insert Table 1 about here ---

Drawing on insights from the behavioral theory of the firm, we posit that time to recall is influenced by the firm's *ability to fully investigate the defect* and *its motivation to wait for the outcome of the investigation*. We argue that a key factor that delays recall post an investigation is the severity of the problem. Severe problems are defined as those with serious consequences. We test our hypothesis using defect investigations involving the major automakers from 1999 to 2012 in the United States. We analyze time to recall using a Weibull accelerated failure time hazard model. Indeed, this study finds that, as expected, problem severity increases the time to recall.

Interestingly, we find that there is variation in the magnitude and direction of the relationship between problem severity and time to recall. For a given level of problem severity, increase in brand reputation for reliability reduces the time to recall while increase in brand diversification (i.e., breadth of offerings of the brand) increases the time to recall. Furthermore, our study finds that increase in past recall intensity of the brand helps lower the time to recall. This finding implies that product recalls do have a positive effect in regulating the behavior of firms and promoting consumer safety. These insights help clarify why recall timing varies across brands. Collectively, our study shows that the recall timing behavior of firms is complex and nuanced.

Second, we test the performance impact of product recall timing decisions. We assess performance using a short-term event study methodology. As noted earlier, prior research has not examined whether stock markets care about the timing of product recalls. Our study finds that stock markets react more negatively when time to recall increases. This finding serves a cautionary note to firms that seek to respond slowly to potential problems in the hopes of avoiding a recall altogether.

Background: Product Recalls

Products are recalled when suspected defects undermine their performance and/or safety. Recalls are offered to all consumers of a product including those who may have not yet experienced any problem associated with the defect. Product defects are investigated and recalls are often supervised by government agencies that, among other activities, inform the public about a recall and monitor its completion. In the United States, these include agencies such as the *Consumer Product Safety Commission*, *Food and Drug Administration*, and *National Highway Traffic Safety Administration*. To complete a recall, firms either repair the product or allow customers to return it for a refund.

Product failures occur for many reasons. They can arise from flaws in design or production (Mackelprang, Habermann, and Swink 2015; Ramdas and Randall 2008), problems with materials and suppliers (Chao, Iravani, and Savaskan 2009), and unanticipated use (and misuse) of the product by consumers (Berman 1999). Sources internal or external to the firm may be responsible for the defect (Folkes and Kotsos 1986). As a Ford spokesperson noted during an investigation into a defect causing a fire hazard involving the Ford F-150 series: "Fires happen for a variety of reasons from faulty repair, improper modification to the vehicle with aftermarket parts and wiring, prior accident damage, and even arson. This is why each complaint or allegation must be reviewed on a case-by-case basis" (Thomas 2005). Thus, thorough investigations are often required before the source of the problem can be identified and a remedy provided. The investigation could show that the product is safe and need not be recalled. For these reasons, a hasty recall might lend credibility to an unsubstantiated claim (Smith, Thomas, and Quelch 1996).

Product recalls are costly propositions. Recall costs can be direct or indirect. Direct costs include all expenditures related to the recall process - expenses for repair, refund, or replacement, including costs associated with retrieving the defective product (Jayaraman, Patterson, and Rolland 2003). The magnitude of these costs depends on the nature of the problem, the size of the product population to be recalled, and consumer response to the product recall. Thus, firms could be motivated to delay recalls and push the direct, tangible costs of recall into future time periods. For example, in internal company documents, Toyota claimed that it saved $100 million by delaying a full recall (McCurry 2010). Furthermore, the potential for short-term loss of sales might also lead to a delay in recall to meet sales and profit targets even as the firm considers the effect of the recall on brand reputation.

Recalls also lead to indirect costs associated with declines in reputation and market performance. Brand reputation depends on stakeholder perceptions about the brand's safety and reliability (Keller 1993; Stahl et al. 2012). Recalls constitute negative information about a brand's performance and could thus damage its reputation if stakeholders update their beliefs about the brand. Recalls can also lead to a downturn in the firm's market performance (Rhee and Haunschild 2006). This decline can occur because firms may withdraw a recalled product from the market or consumers may switch to competitors products.

These indirect recall costs, however, are often contingent on how the firm responds to a safety problem. If a recall becomes inevitable, quick responses lead to lower losses in brand reputation than a stonewalling or defensive response (Dawar and Pillutla 2000; Siomkos and Kurzbard 1994). Delayed recalls, in contrast, can damage the brand's reputation and also increase litigation risk. If indirect costs of a recall could exceed its direct costs (Rupp 2004), a quick recall might be warranted. Thus, although firms may have compelling reasons to delay recalls, a quick response may help in containing the indirect costs (Dani and Deep 2010). Overall, the recall timing decision is far from straightforward because firms face competing pressures and constraints.

Theory and Hypotheses

The behavioral theory of the firm characterizes firms as systems of structurally distributed attention- the noticing, encoding, interpreting, and focusing of time and efforts by firms' decision makers. This theory implies that firms selectively attend to market information, conduct limited search, and find satisficing solutions to problems. A satisficing solution involves striving to meet multiple goals such as market share and brand reputation while operating within a profit constraint (Cyert and March 1963; 1992). We approach recall timing decisions from the vantage point of the behavioral theory of the firm. From this perspective, the ability and motivation of firms to respond to a defect investigation in light of these multiple and often conflicting goals are key determinants of their actions.

Firms divide their managerial attention and resources amongst numerous events and prioritize the ones that need a quick response. Attention is likely to be focused on external events likely to affect the firm's long-term ability to generate cash flows (Argote and Greve 2007; Cyert and March 1992). Prior research has examined firm responses to such events, including competitive actions such as new product introductions or market entry (Chen and Hambrick 1995; Gatignon, Anderson, and Helsen 1989; Jayachandran and Varadarajan 2006), or responses to technological advances (Lee and Grewal 2004). The opening of a safety investigation constitutes such an event given that the investigation could result in a recall, which studies have shown to be a value-relevant event (e.g., Chen, Ganesan, and Liu 2009). Firm responses can be categorized along several dimensions but in the context of safety investigation, the most relevant response characteristics are the likelihood of a recall and the associated timing. Prior research has shown that these response dimensions are a function of the motivation and ability of the firm to respond to the external event (Chen and Hambrick 1995; Smith et al. 1991).

We argue that the more severe the defect, the greater the attention to the investigation since a recall involving a severe defect is costlier than a recall involving less severe problems. However, problem severity, while triggering search for a solution, does not necessarily lead to a quick response because it reduces the ability and motivation of firms to respond. We contend that the primary factor that drives recall timing is problem severity.

However, we propose that the effect of problem severity on time to recall will be contingent on the characteristics of the brand. Specifically, we argue that the impact of problem severity on time to recall will be moderated by a) brand reliability, b) brand diversification and c) past recall intensity of the brand. Drawing on insights from the behavioral theory of the firm, we explain how these brand characteristics moderate the relationship between problem severity and time to recall by altering the ability and motivation of the firm to respond to a potential defect of high severity that is under investigation.

Problem Severity and Time to Recall

Problem severity refers to the seriousness of the consequences of product defect from the standpoint of consumer safety. We expect problem severity to be significantly related to recall timing because of its impact on the ability and motivation of firms to respond. The ability of the firm to provide a quick response is closely linked to whether the firm can identify a potential solution to fix the defect. In this regard, severe problems will trigger 'problemistic' or problem-oriented investigations in firms. However, problemistic search behavior is myopic in that the investigation will rely on traditional routines and, thus, may not quickly arrive at a solution (Argote and Greve 2007). Thus, although all safety investigations might trigger the search for a solution, the *ability* of the firm to provide a quick response to severe problems is especially limited.

When a product defect is suspected, apart from searching for a solution, the firm also strives to determine who is to be held accountable for the failure (Sitkin 1992). Assessing accountability becomes more consequential for severe problems, especially because of the desire to prevent such problems in the future. However, fixing internal responsibility may put employees who are entrusted with finding a solution to the problem at risk for loss of reputation, demotion, or even job loss. Hence, the desire to avoid responsibility might lower motivation to share information and pursue a solution amongst those likely to be held accountable (Madsen and Desai 2010). The ensuing response would be similar to that predicted by the "threat-rigidity" hypothesis (Staw, Sandelands, and Dutton 1981), where the key concern may not be solving the problem but protecting the interest of the dominant coalition by controlling decisions.

Research has also shown that severe recalls are more likely to be punished by stakeholders than less severe recalls (Cheah, Chan, and Chieng 2007; Chen, Ganesan, and Liu 2009; Liu and Shankar 2015). Furthermore, lawsuits are more likely in cases of severe defects. Therefore, stakes are higher for a firm in case of severe defects. Thus, as problem severity increases, firms will also be motivated to avoid external accountability and delay the recall. The benefit of delaying is that the firm could avoid a recall altogether if the investigation finds that the problem does not constitute a safety defect that warrants a recall. Therefore, firms may be motivated to delay a response during the investigation even when there are numerous complaints. For example, the New York Times recently reported that despite 150 complaints of injuries or deaths attributed to a steering problem with the Ford Focus, Ford had not initiated a recall (Jensen 2015). Overall, when the firm is investigated for severe defects, the process may take longer and delay the recall decision. Hence, we offer the following baseline hypothesis:

H1: Problem severity increases time to recall.

Moderating Effect of Brand Reliability

How does the relationship between problem severity and time to recall vary as a function of brand reliability? *Brand reliability* is a brand's reputation as a provider of dependable products. Reputation - whether for a brand or for the firm itself - is a critical asset that a firm strives to protect. For instance, Warren Buffet, CEO of Berkshire Hathaway, in a July 2010 letter, exhorted his managers to zealously guard Berkshire's reputation: "We can afford to lose money – even a lot of money, but we can't afford to lose reputation – even a shred of reputation" (Protess, Rusli, and Craig 2011). A strong brand reputation confers several advantages. It attracts and retains customers, reduces their price sensitivity, and enhances revenues (Keller 1993; Stahl et al. 2012). There are two reasons why brand reliability should moderate the problem severity–time to recall relationship.

The damage to brand reputation from a crisis is a function of prior expectations and whether the recall strategy is consistent with those expectations (Cleeren, Dekimpe, and Helsen 2008; Dawar and Pillutla 2000). Consumers are particularly sensitive to information that violates expectations of high reliability (Heath and Chatterjee 1995). Defects of high problem severity already violate those expectations. Consequently, a quick recall would be more consistent with the expected response from brands of high reliability reputation. A delayed recall for severe problems by higher reliability brands would further violate expectations. That is, although the brand's reputation for reliability will be harmed by the negative information conveyed by a recall, a quick response could potentially limit the damage (Dawar and Pillutla 2000; Siomkos and Kurzbard 1994). Thus, higher reliability brands will be motivated to recall faster than their lower reliability counterparts do as they face investigations for serious problems.

Second, whether the affected brand has a reputation for being reliable is indicative of the firm's ability to investigate a defect and determine the root cause faster. Firms that position their brands on reliability are likely to do so by building strong research and engineering expertise. This expertise could help the firm identify the type and source of the flaw (e.g., materials/sourcing, product) more quickly. Thus, reliable brands will be quicker to offer remedies/solutions for severe problems than brands that are less reliable. Overall, as brand reliability increases, the motivation and ability to hasten recalls for severe problems will increase. Thus, we expect:

H2: The higher a brand's reliability, weaker the relationship between problem severity and time to recall.

Moderating Effect of Brand Diversification

Brand diversification refers to the number and variety of products marketed under a brand. How is the relationship between problem severity and recall timing affected when brand diversification increases? Brand diversification influences the need that a firm has to investigate the problem more thoroughly, thus deepening the search process. While diverse brands market different products, firms often share components and systems across products to reduce costs. This practice introduces uncertainty around safety investigations. When a safety investigation is opened, firms need to perform checks on multiple models to accurately determine the source of the defect. Further, the investigation could get complex as shared components or systems could perform differently in sub-brands because of interactions with other systems (Ramdas and Randall 2008). Therefore, failure in a brand does not always imply that all sub-brands would be necessarily at risk. This uncertainty places greater pressure on diverse brands to accurately investigate safety issues (Ramdas and Randall 2008).

To exacerbate the problem, prior research has also shown that firms that address multiple segments through diversification are less capable of integrating complex knowledge and learning from their own experience, compared to their more focused rivals (Haunschild and Sullivan 2002; Ingram and Baum 1997). Hence, diversified brands are likely to be less adept at investigating severe problems, resulting in delayed recalls. Also, the fear of negative spillovers to related products in the portfolio that expands the scope of the recall may inhibit the motivation for a quick recall in the hope of avoiding one after the investigation.

Overall, increase in brand diversification is likely to reduce both the ability and motivation of firms to recall quickly when faced with problems of high severity. Thus, we propose:

H3: The higher a brand's diversification, stronger the relationship between problem severity and time to recall.

Moderating Effect of Past Recall Intensity

Past recall intensity refers to the extent of product recalls experienced by the firm in the recent past. The arguments for how past recall intensity could influence the relationship between defect severity and time to recall are equivocal. Prior research has shown that firms can learn from failures such as accidents or product recalls (Madsen and Desai 2010; Thirumalai and Sinha 2011). Recent recalls should help firms develop the managerial and administrative competence to deal with information about product failures more effectively, making their own investigation process more efficient. Thus, a firm that has experienced recent recalls is more likely to have relevant knowledge stock that it can tap into in case of a new investigation. Consequently, recent recalls sensitize the firm to the trajectory of investigations and expedite recalls. This learning will be especially valuable when the products under investigation are potentially facing severe problems.

However, past recalls could also limit a firm's ability to recall early. The recall process is tedious. Details about the recall should be communicated to channel members and consumers, a remedy should be offered, and the supply chain reconfigured for the reverse flow of goods (Jayaraman, Patterson, and Rolland 2003). The remedial process can take a long time as consumer return rates are often low. Even after several months, firms have to devote resources to motivate consumers to comply with the recall. A recent US Government Accountability Office report shows that the remediation rates average around 70% for automobiles 18 months after the recall. For other products, where firms have a lower ability to track consumers, these rates are lower. Thus, firms with high past recall intensity are likely to be resource-constrained and have lower ability to issue a quick recall. Given the equivocal arguments for the moderating role of past recall intensity, we do not offer a formal hypothesis but treat it as an empirical issue. Table 2 summarizes the rationale for these hypotheses.

--- Insert Table 2 about here---

Time to Recall and Shareholder Value

We expect time to recall to adversely influence stock market performance. As previously discussed, recalls are costly events that not only require a resource commitment in the short- and medium-run, but also can damage marketing assets such as brands for many years. In the case of a recall, time to recall reveals information to the firm's stockholders about its responsiveness to safety concerns. Stock markets are likely to respond to recall timing decisions because they significantly shape consumer sentiment towards the brand. That is, time to recall has value relevance for stock markets beyond information about the recall's direct and indirect costs.[2] As we noted earlier, firms turn inward and investigate a defect when they suspect a severe problem, slowing down the recall process. Longer time to recall signals that the firm does not have either the ability or the motivation to respond to safety problems in a timely manner. That is, it does not have the knowledge and processes in place to make a decision quickly and it does not have the motivation to recall quickly to safeguard consumer safety. In contrast, a quick recall signals to stock markets that the firm is responsive to safety concerns and can make recall decisions quickly. Hence, stock markets are likely to view time to recall as a proxy for the firm's commitment to safety concerns. Therefore:

[2] Prior research has shown that the stock market reacts negatively to direct and indirect costs of a recall (Chen, Ganesan and Liu 2009; Jarrell and Peltzman 1985; Thirumalai and Sinha 2011). Indicators of recall-related costs are, for example, recall size and problem severity.

H4: Time to recall is negatively associated with stock market performance of the firm.

Research Methodology

To test the hypotheses, we collected data on safety investigations in the automotive industry. We focus on the automotive industry for several reasons. First, examining recall timing in one industry eliminates confounding from extraneous industry-specific effects and enhances internal validity. In addition, the automotive industry is a salient one because of the comprehensiveness of the data available - the opening of an investigation and the announcement of a recall are well documented in this industry. For these reasons, the automotive industry has been the context of other studies in the recall area (e.g., Borah and Tellis 2016; Haunschild and Rhee 2004; Jarrell and Peltzman 1985; Kalaignanam, Kushwaha, and Eilert 2013; Liu and Shankar 2015).

Investigations in the automotive industry are initiated by the NHTSA. The NHTSA was instituted in response to the National Traffic and Motor Vehicle Safety Act of 1966. The responsibilities of NHTSA include establishing minimum performance standards for automobiles, verifying whether these standards are met, investigating noncompliance, and directing recall campaigns (Rupp and Taylor 2002). The NHTSA has overseen thousands of recalls involving hundreds of millions of vehicles since its inception in 1966. Most of the recalls are voluntary yet supervised by the NHTSA.

If the NHTSA suspects a product defect, it can open an investigation into this particular issue. Investigations can be triggered by consumer complaints, queries into ongoing or past recalls, service bulletins, product testing, or by petitions. At the end of the investigation, NHTSA will either require the manufacturer to recall the product or close the investigation into the issue for the time being. The manufacturer, however, can issue a recall at any point during the investigation. NHTSA assesses the proposed recall based on whether it will remedy the safety problem.

Our dataset includes all investigations on safety issues related to passenger vehicles between 1999 and 2012. The year 1999 was the first for which data were available for all investigations, and hence serves as the starting year for the sample. The final dataset includes 381 investigations; 201 of these eventually ended in a recall.

Data Sources and Measures

We assembled the data for the study from numerous sources. Investigation-related information such as opening, closing, and recall dates, as well as information about problem severity was collected from NHTSA. Other data sources, particularly for the brand-related variables, include *Consumer Reports, Ward's Automotive Yearbook*, *Automotive News Market Data Book*, and firm's annual reports. In our empirical setting, the term 'brand' refers to the auto-make and the term 'firm' refers to the manufacturer. For example, Acura is the brand in our context whereas Honda is the manufacturer. Table 3 lists the specific measures and data sources.

--- Insert Table 3 about here ---

Dependent variables. Time to recall (*TIMETORECALL*) refers to the time elapsed between the opening of an investigation and the time of recall announcement by the manufacturer. We collected information about investigation opening and recall announcement dates from NHTSA investigation and recall reports. We treat investigations not ending in a recall as right censored because of the possibility that a closed investigation could be reopened if additional problems are observed (see 'Additional Analyses' for an alternative model that relaxes this assumption). The measure of time to recall represents the difference between the investigation opening date and recall announcement date (or investigation closing date for censored observations), measured in days.

Our performance measure is the short-term abnormal returns (AR) accruing from the recall announcement to the focal firm, using the event study methodology (e.g., Hendricks and

Singhal 1996). Event studies typically allow 1) testing for the existence of information effects (i.e., the impact of an announcement on shareholder value) and 2) identifying factors that explain changes in shareholder value (Kalaignanam et al. 2013, p. 754). The information effects of an announcement are assessed by computing the difference between the observed return, R_{it}, on the event date and the expected returns, $E(R_{it})$, estimated on a benchmark model. The premise of an event study methodology is that an announcement reveals new information to stock markets causing it to adjust the valuation of the firm based on the expected impact of the new information. A product recall announcement reveals to stock markets at least two pieces of new information. *First*, the stock market learns that a safety investigation underway was serious enough to lead to a product recall. This is new information because there is uncertainty about the outcome of a safety investigation. *Second*, the stock market also learns about the time taken by the firm to investigate and issue a product recall. This constitutes new information to stock markets because the time taken in a safety investigation would only be relevant after a product recall is announced. We estimate the expected returns, $E(R_{it})$, using the Fama-French four-factor model (Carhart 1997; Fama and French 1993):

$$E(R_{it}) = \hat{\alpha}_i + \hat{\beta}_i R_{mt} + \hat{\gamma}_i SMB_t + \hat{\delta}_i HML_t + \hat{\sigma}_i UMD_t$$

where, R_{mt} is the stock returns of the benchmark market portfolio, SMB_t is the difference between rate of returns of small and big stock firms (i.e., small – big), HML_t is the difference in returns between high and low book-to-market ratio stocks (i.e., high – low) and UMD_t is the momentum factor defined as the difference in returns between firms with high and low past stock performance (i.e., winners – losers). α, β, γ, δ, and σ are the parameter estimates obtained by regressing R_{it} on the four factors.

We estimate the daily stock returns for each firm between 260 and 30 days prior to the event day using the Fama-French Four-Factor model. Abnormal returns are estimated as the difference between the observed returns and the expected returns:

$$AR_{it} = R_{it} - E(R_{it}) = R_{it} - (\widehat{\alpha_i} + \widehat{\beta_i}R_{mt} + \widehat{\gamma_i}SMB_t + \widehat{\delta_i}HML_t + \widehat{\sigma_i}UMD_t)$$

The abnormal returns are aggregated for a firm over an event period [$-t_1$, t_2] and are given by

$$CAR_{(-t_1,t_2)} = \sum_{t=-t_1}^{t_1} AR_i.$$

When information leakage (for t_1 days before the event) and/or dissemination over time (for t_2 days after the event) occur, the abnormal returns for a firm are aggregated over the 'event period' [t_1, t_2] into a cumulative abnormal return (*CAR*).

Independent variables. We operationalize problem severity (*PROBSEV*) using complaints/accidents data that is available at the beginning of the investigation. NHTSA collects information about the number of a) consumer complaints, b) crashes or fires, c) injuries, and d) fatalities to analyze the safety hazards posed by automobiles in the market. We performed principal component analysis on these four items and generated a univariate score to operationalize problem severity. Higher scores imply greater problem severity.

The data for brand reliability (*RELIABILITY*) was obtained using *Consumer Reports'* assessment of vehicle reliability. The reliability ratings by *Consumer Reports* influence perceptions of brand reliability (Rhee and Haunschild 2006). In line with prior research, we aggregate the information to the make, or brand, level. As noted before, the term 'brand' in our setting refers to the auto-make and 'firm' refers to the manufacturer. For example, Cadillac is the brand and GM is the firm in our setting. In the automotive industry, metrics such as brand equity (e.g., Harris Equitrend) or customer satisfaction (e.g., ACSI) is commonly assessed at this level. *Consumer Reports* surveys consumers regarding problems with a particular model and aggregates the information into problem rates. From this data, brand reliability is measured using a five-point scale of problem rates, with higher scores reflecting higher reliability (and thus fewer problems; Rhee 2009). We collected problem scores for the previous three years for each of the models associated with a particular make (our level of aggregation) and averaged these model scores across the make to generate the reliability score for the make (Rhee and Haunschild 2006).

Brand diversification (*DIVERSE*) represents the number and variety of products marketed under a brand. Consistent with past research (Dobrev 2000; Rhee and Haunschild 2006), we operationalize brand diversification in terms of the number of product lines and range of engine capacities of the models produced by the brand. We performed principal component analysis on these two indicants to generate a univariate score to represent brand diversification (Rhee and Haunschild 2006). Higher scores on this component reflect greater brand diversification. We collected information about the brand's past recall intensity (*PASTRECINT*) from recall reports published by NHTSA. We operationalize past recall intensity as the count of the total number of recalled vehicles of the brand in the previous year.

We also control for firm profitability (*ROA*), recall size (*RECSIZE*) and past publicity (*PASTPUBLICITY*). The relationship between firm profitability and time to recall is not altogether clear, as strong financial performance might reduce the firm's motivation to respond but enhance its ability to respond because of the availability of slack resources (Jayachandran and Varadarajan 2006). We expect recall size to enhance recall delays and past publicity to shorten the time to recall. Further, we collected data on the type of the investigation (*INVTYPE*) at the time of closing. Investigations by NHTSA could be categorized as a) preliminary evaluation (PE) b) service query (SQ), c) recall query (RQ) and d) engineering analysis (EA). PE is triggered by consumer complaints, petitions, or outcomes of NHTSA product testing. SQ is initiated after NHTSA reviews technical service bulletins issued by the manufacturer. RQ is initiated when NHTSA assesses the adequacy of the scope of prior recalls. Finally, EA is initiated when detailed analyses is needed to evaluate the product defect.

Model Estimation

Our principal interest is in testing factors that influence the time to recall and stock market reactions to time to recall. A potential econometric concern for assessing stock market reaction the second stage is that time to recall is endogenous. That is, firms are likely to be aware of the consequences of time of recall and would therefore (rationally) make decisions about when to recall based on strategic considerations, some of which may be unobserved. Accordingly, the model needs to account for the endogeneity of time to recall to yield unbiased coefficients. We account for this endogeneity using the control function approach (Petrin and Train 2010). The logic of this approach is that if there are unobservables that drive the time to recall decision, including a control variable in the structural regression model (second stage) reduces the correlation between the endogenous regressor and the error term.

The model estimation proceeds in two stages. In the first stage, we model the likelihood of recall and time-to-recall using a hazard model. Not much theory or empirical evidence is available to predict the shape of the time to recall process. Therefore, to overcome misspecification bias, we compare several distributions. These distributions are continuous because the recall could occur at any time. In particular, we estimate accelerated failure time (AFT) models incorporating exponential, Weibull, log-logistic, and Gamma time to recall distributions. Table 4 reports the model fit for these distributions. As seen in Table 4, the Weibull distribution offers the best fit based on Bayesian Information Criteria (BIC) statistic. Accordingly, we use a Weibull distribution for the time to recall process (t) with a scale parameter λ and shape parameter *p*. The probability density function for the Weibull time-to-recall process can be expressed as:

$$f(t) = \underbrace{\lambda p t^{p-1}}_{h(t)} \underbrace{\exp(-\lambda t^{p})}_{S(t)}$$

where h(t) is the hazard function and S(t) is the survivor function. To facilitate direct interpretation in terms of time, we use an accelerated failure time (AFT) metric. The survivor function can be written as:

$S(t)= \exp(-\lambda t^{p})$

Solving for t yields:

$t= (-\log(S(t)))^{1/p} * 1/\lambda^{1/p}$

Reparametrizing $1/\lambda^{1/p}$ in terms of exp(Xβ), the Weibull time to recall process in an AFT metric can be expressed as in equation (1).

$$TIMETORECALL_T = [-log S(t)]^{1/p} * \exp(\propto_1 PROBSEV_i + \propto_2 RELIABILITY_j + \propto_3 DIVERSE_j + \propto_4 PASTRECINT_j + \propto_5 (PROBSEV_i * RELIABILITY_j) + \propto_6 (PROBSEV_i * DIVERSE_j) + \propto_7 (PROBSEV_i * PASTRECINT_j) + \propto_8 ROA_k + \propto_9 RECSIZE_i + \propto_{10} PASTPUBLICITY_j + \propto_{11-13} INVTYPE_i + \sigma\varepsilon_I \quad \textbf{(1)}$$

where TIMETORECALL= time elapsed between opening of investigation and recall for investigation ‘i’, *S(t)* is the survivor function, PROBSEV= severity of problem in investigation ‘i’, RELIABILITY= reliability of brand ‘j’, DIVERSE= diversity of brand ‘j’, PASTRECINT= past recall intensity of brand ‘j’, ROA= return on assets of firm/manufacturer ‘k’, RECSIZE = number of vehicles included in the investigation ‘i’, PASTPUBLICITY= past publicity of recalls for make ‘j’, INVTYPE= investigation type, α_1- α_{13} are the parameter estimates and ε_i is the error term. Equation (1) could suffer from specification issues if unobservable firm characteristics that influence time to recall are ignored. To account for this possibility, we use clustered robust standard errors (Liang and Zeger 1986). We cluster the standard errors on the firm and account for within cluster correlations or unobserved firm heterogeneity in the model estimation. Note that clustering standard errors by firms is equivalent to modeling firm-specific random effects for the intercept (Moulton 1986) and are larger than those obtained from conventional estimation, thereby making the hypotheses tests more conservative. The clustered robust standard errors also control for firm-specific heteroskedasticity (Bertrand, Duflo, and Mullainathan 2004). As a robustness check, we also estimate the model with fixed firm effects.

---Insert Table 4 about here---

In this model, past publicity is the exogenous variable that identifies the system. From a theoretical perspective, past publicity is exogenous because while it is likely to influence time to recall, it should not be related to cumulative abnormal returns in the second stage. The exogeneity of past publicity stems from the efficient market hypothesis, which argues that the effects of past recalls that were publicized should already be incorporated in the stock prices. Hence, past publicity is not 'news' per se and should not elicit any reaction from stock markets. To understand the relevance of past publicity for time to recall, we estimated an alternative specification that excludes past publicity from the model. The BIC for this alternative model is 721.37 which is greater than the model which includes past publicity (BIC=706.89). Thus, past publicity improves the model fit and does a good job of predicting time to recall.

In the second stage, we test the effect of time to recall on stock market reactions. As noted before, we account for the endogeneity of time to recall using a control function. The control variable we include in the second stage are the residuals from the first stage (actual time to recall- predicted time to recall). By including residuals from the first stage, OLS can be used to generate unbiased coefficients of stock market reactions to time to recall decisions. Also, note that the second stage only features investigations that ended in a recall. The cumulative abnormal return is modeled as a linear combination of time to recall, recall size, problem severity and brand sales. We include recall size and problem severity to account for the direct and indirect costs of the recall. To account for the possibility that stock markets respond more adversely to product recalls of significant brands, we include brand sales (*BRANDSALES*) in the model.

$$CAR_{i(t1,t2)} = \eta_0 + \eta_1 TIMETORECALL_i + \eta_2 RECSIZE_i + \eta_3 BRANDSALES_j + \eta_4 PROBSEV_i + \eta_5 RESIDUALFS_i + \upsilon_i \quad \textbf{(2)}$$

Where TIMETORECALL = time to recall, CAR= cumulative abnormal returns for the recall announcement in the event window [t_1, t_2], RECSIZE= number of vehicles included in the investigation ‘i’, BRANDSALES= sales of the brand involved in the recall, RESIDUALFS = residuals from the first stage (i.e., actual time to recall - predicted time to recall), υ_i is random error. As before, we use clustered robust standard errors for inferences from equation (2). We discuss the results next.

Results

Overall Descriptive Findings

Table 5a presents the summary statistics and correlations between key variables in the study. The correlations between the independent variables are within prescribed limits (vifs <10) and do not pose a threat to the validity of the findings. The mean time to recall is 295 days and the standard deviation of time to recall across investigations is 248 days. Table 5b reports the average time to recall across manufacturers in our sample. As seen in Table 5b, the mean time to recall for manufacturers such as General Motors and Ford is higher than other manufacturers. Toyota has the lowest mean recall time in our sample.

We also find considerable variation in the severity of problems across investigations. On average, the mean number of complaints at the time of investigation is 27 with a standard deviation of 96, mean number of crashes is 1.72 with a standard deviation of 7.5, mean number of injuries is .71 with a standard deviation of 2.24 and mean number of fatalities is .04 with a standard deviation of .32. We also find reasonable within and between variance in the brand characteristics. For instance, the between and within variances for a) brand reliability are 69% and 31% respectively, b) brand diversification are 80% and 20% respectively and c) past recall intensity are 56% and 44% respectively. The within variance arises because the data spans investigations of brands over a thirteen-year period between 1999 and 2012. The mean cumulative abnormal returns to recall announcements is -0.6%, which translates into a mean shareholder loss of $168 million (average market capitalization = $28 billion).

--- Insert Tables 5a & 5b about here ---

Impact of Problem Severity on Time to Recall: Results

The results for the time to recall model are reported in Column I of Table 6. The standard errors are clustered robust standard errors that controls for within firm correlation (i.e., equivalent to random effects) and for heteroskedasticity. As shown in Table 6, the shape parameter is greater than one (1.76, p<.01), suggesting that the hazard of recall increases over the time of the investigation. H1 states that problem severity enhances time to recall. The coefficient for the impact of problem severity on time to recall is positive and significant (.84, p<.01).

Substantively, this result implies that a unit increase in the severity of the problem increases time to recall by 52%.[3] While our model does account for unobserved firm effects (through clustered standard errors), we tested the robustness of the results by including fixed firm effects. As seen in Column II in Table 6, the main effect of problem severity is positive and significant. We also performed additional analyses to check if this result is robust to an alternative measure of problem severity that includes crashes, injuries and fatalities but excludes a number of complaints. The results from this analysis are reported in Column III of Table 6. As seen in Table 6, the finding remains positive and significant.

[3] Computed as 100*[exp(α)-1], where α= .74 * 1/1.76 (1.76 is the shape parameter coefficient from Table 6).

--- Insert Table 6 about here ---

Moderating Influence of Brand Reliability, Brand Diversification, and Past Recall Intensity

H2 hypothesizes that the relationship between problem severity and time to recall would be weaker as brand reliability increases. Consistent with H2, the coefficient for the interaction between problem severity and brand reliability is negative and significant (-.18, $p<.05$). Given that our model is non-linear, we compute average marginal effects to gain a better understanding of this interaction effect. Specifically, we compute the average marginal effect of problem severity on time to recall at low and high levels of brand reliability. We set low and high levels of brand reliability at 10% and 90% respectively. We find that a unit increase in problem severity increases time to recall by 357 days at low levels of brand reliability and by 259 days at high levels of brand reliability. Therefore, the time to recall for brands of higher reliability is shortened by 98 days (i.e., 357days - 259 days) for a unit increase in problem severity. The direct effect of brand reliability on time to recall is also negative and significant (-.29, $p<.05$).

H3 states that the relationship between problem severity and time to recall would be stronger as brand diversification increases. Consistent with this prediction, the results in Table 6 show that the interaction between problem severity and brand diversification is positive and significant (.08, $p<.05$). As before, we compute average marginal effects to gain substantive insights into this interaction. The average marginal effect of a unit increase in problem severity on time to recall at low levels of brand diversification is 279 days and at high levels of brand diversification is 344 days. Substantively, this finding means that the time to recall for problems of higher severity increases by 65 days at high levels of brand diversification. Therefore, when brands target multiple segments of customers through their offerings, their response to product safety investigations slows down significantly. This result is similar in spirit to results from previous research on structural-inertia, which finds that a firm in a “wide niche” (addressing multiple segments) would have lower ability to act as a reliable and accountable collective entity (Carroll and Swaminathan 2000; Dobrev 2000), and hence respond more slowly to crises. We do not find the direct impact of brand diversification on time to recall to be significant (p>.10).

We did not propose a hypothesis for the moderating influence of past recall intensity on the relationship between problem severity and time to recall. The results in Table 6 show that though the impact of past recall intensity on time to recall is insignificant (p>.10), the interaction of problem severity and past recall intensity is negative and significant (-.0074, p<.01). A unit increase in problem severity increases time to recall by 368 days at low levels of past recall intensity and by 172 days at high levels of past recall intensity. Thus, problem severity shortens time to recall by 196 days for brands with high levels of past recall intensity. The evidence strongly supports the moderating influence of past recall intensity on the relationship between problem severity and time to recall and is in favor of the learning argument and not the resource constraint argument.

As regards control variables, we find that past publicity is negatively related to time to recall (-.006, p<.01). Therefore, in addition to past recall intensity, publicity of past recalls appears to sensitize firms to new investigations and recall earlier. We also find that the engineering analysis and recall query investigation types have higher times to recall relative to the service query investigation type. This finding suggests that certain investigations offer clearer assessments of the safety risks resulting in firms recalling earlier.

Results for the Impact of Time to Recall on Stock Market Performance

H4 states that time to recall is negatively associated with stock market performance. We examined the daily average abnormal returns for the 73 recall announcements for which stock returns were available. We do not find a statistically significant abnormal return on the announcement day -.12% (p>.10). We examined abnormal returns for a period of seven days around the event. The only time window with a significant *CAR* is [-2, 2]: -.6% ($p<.05$). Although the *CAR* appears to be modest in size, the associated shareholder losses are not trivial given the large market capitalization of automobile firms in the United States. To test H4, we estimated Equation (2) using *CAR* [-2,2] as the measure of stock market performance. We wish to note that examining cross-sectional abnormal returns in the significant window is consistent with previous research in marketing (Geyskens, Gielens, and Dekimpe 2002; Raassens, Wuyts, and Geyskens 2014). Table 7, Column I reports the results of this analysis. The reported standard errors in Table 7 are clustered robust standard errors. Consistent with H4, the coefficient for the impact of time to recall on cumulative abnormal returns is negative and significant (-4.11×10^{-5}, $p<.05$) (see Table 7 Column IV for results with bootstrapped standard errors). Therefore, firms that delay recalls tend to be punished more by stock markets. Regarding control variables, recall size is negatively associated with cumulative abnormal returns (-2.6×10^{-5}, $p<.05$). We do not find problem severity to be significantly related to cumulative abnormal returns ($p>.10$). We also cumulative abnormal tested for the impact of time to recall on an alternative event window [-1,1] despite the fact that returns were statistically insignificant for this window. The results of these analyses are reported in column II of Table 7. We find that time to recall is again negatively related to cumulative abnormal returns (-1.34×10^{-5}, $p<.05$).

As noted before, we tested the stock market reactions for product recall timing decisions after accounting for its endogeneity. We also tested the impact of time to recall on cumulative abnormal returns without endogeneity controls. To facilitate comparisons, we used the full set of variables from the first stage. These results are reported in columns III of Table 7. The results again suggest that time to recall is negatively related to cumulative abnormal returns (-4.8×10^{-5}, $p<.05$). Collectively, our results suggest that even after accounting for size of the recall (i.e., proxy for direct costs for recalls) and problem severity, time to recall has a negative and significant impact on stock market reactions. The implication is that the time to recall is an important variable with value relevance for stock markets.

--- Insert Table 7 about here---

Additional Analyses

We performed numerous additional analyses to examine the robustness of the results to alternative explanations, models and specifications.

Does the timing of product recall differ for U.S., Europe and Asian nameplates? Although our empirical analyses accounted for manufacturer or firm specific heterogeneity by using clustered standard errors, one could wonder if there are significant differences between U.S., Europe and Asia based manufacturers with respect to how they respond to safety investigations. To test this possibility, we included dummy variables for the manufacturer headquarters and re-estimated the recall timing model. These results appear in column I of Table 8. As seen in Table 8, the U.S. and Asian headquarters dummies are not significantly different (p>.10) compared to the base manufacturer headquarter dummy (i.e., Europe). Also, the pattern of results for our hypothesized variables is remarkably similar after the inclusion of these dummies. Thus, it is reasonable to conclude that there are no significant differences in the timing of product recalls based on where the manufacturer is based.

--- Insert Table 8 about here ---

Does the timing of product recall vary depending on the price tier of the brand? Recall that the unit of analysis for this study is the 'make' or 'brand'. That is, Chevrolet and Cadillac are makes or brands for the parent firm (or manufacturer) General Motors. A potential difference between various brands from firms is that they belong to different price tiers. Does the response of firms to safety investigations differ based on the price tier of the make involved in the investigation? A priori expectations would be that luxury brands are likely to recall earlier compared to other brands. We performed additional analysis to evaluate this possibility. Based on the average base prices of models of the make, we created a 'Luxury Brand' dummy variable with luxury makes (or brands) coded as '1' and others as '0'. The results of the recall timing model after including the 'Luxury Brand' variable appear in column II of Table 8. The luxury brand dummy is not statistically significant although it is directionally consistent with expectations. Thus, we rule out the possibility of the brand's price tier driving the time to recall results.

Alternative model to allow for the possibility that the right censored investigations never experience product recalls in the future? The Weibull AFT regression model used to test the hypotheses assumes that the investigations closed without a product recall are right-censored. That is, these investigations would also lead to a recall in the future although we do not observe a product recall at the end of 2012. One could wonder if all investigations would indeed culminate in a product recall. To relax this assumption, we turn to an alternative model, the split hazard model, which incorporates both incidence and the time to recall in the same framework. The logic of this model is that time to recall is conditioned on the likelihood of a recall occurring. We used a logit specification to model product recall incidence and Weibull specification to model time to recall (as before). The results from this new model appear in columns IIIa and IIIb in Table 8. As seen in Table 8, the conclusions from this model are identical to those from the standard duration model.

Are there long-term abnormal returns for product recall announcements? In this subsection, we examine whether the market quickly and accurately incorporates the performance implications of product recall announcements in the share price. Alternatively, it may take the market a long time to figure out the performance consequences of the product recall announcement[4]. Support for the Efficient Market Hypothesis can be found if the long-term abnormal returns for the announcing firm in question are not significantly different from the long-term abnormal returns for a comparison benchmark. Consistent with previous research, we use the calendar-time portfolio methodology to test the long-term abnormal returns for product recall announcements (Sorescu, Shankar, and Kushwaha 2007).

The calendar-time portfolio methodology involves creating portfolios of stocks of firms who have announced a product recall. Firms are added to the portfolio on the date of announcement and held in the portfolio for the period of time for which we wish to calculate returns (see Fama 1998; Mitchell and Stafford 2000). We use the Fama-French three-factor model to calculate the one-year and two-year abnormal returns. The intercept term of these models is the measure of the average monthly abnormal returns of the portfolio. The results appear in Table 9. The results suggest that the intercept term is not statistically significant for either the one-year or two-year portfolio returns (p>.10). Thus, the long-term abnormal returns analyses indicate that the Efficient Market Hypotheses cannot be rejected and that the returns to product recall timing decisions accrue in the short-term.

--- Insert Table 9 about here ---

Do the results hold across individual components of problem severity? Recall that problem severity is derived using principal component analyses on the number of complaints, number of crashes/fires, number of injuries and number of deaths. One could wonder whether the results hold across these individual components because of the possibility that they are qualitatively different. To assess this possibility, we re-estimated the results using these

[4] The difference between short-term abnormal returns and long-term abnormal returns is related to the issue of market efficiency and security mispricing. Evidence of significant long-term abnormal returns implies that there are biases in information processing that result in over-reaction or under-reaction to announcement information (Kalaignanam et al. 2013). In other words, short-term abnormal returns and long-term abnormal returns only suggest the time horizons over which markets respond to information. As such, they do not connote short-term and long-term firm performance.

individual components as proxies for problem severity. The results of this analyses are reported in Table WA2 in the '*Web Appendix'.* As seen in Table WA2, the pattern of results is fairly consistent across crashes/fires, injuries and deaths. The only difference is that some of the effects are not statistically significant ($p>.10$) for the number of complaints (although they are directionally consistent with our theory). This analyses provides additional evidence supporting our theorized effects.

Discussion

Research has shown that product recalls have significant consequences for firms as well as end customers. However, little evidence is available regarding why some firms recall earlier than others do. We develop and test hypotheses that examine the impact of problem severity on time to recall. Then, we examine how brand reliability, brand diversification, and brand's past recall intensity pose boundary conditions, affecting the relationship between problem severity and time to recall. Thereafter, we demonstrate the effect of time to recall on firm performance. Next, we summarize our contributions in relation to the research questions identified at the beginning of the article.

Research Contributions

Our study makes two important contributions to the product recalls and brand management literature. *First*, we contribute to the product recall literature by examining time to recall when firms face safety investigations by external entities. Extant research has thus far focused on how product recalls impact stock market performance (Chen, Ganesan, and Liu 2009), effectiveness of marketing instruments (Cleeren, van Heerde, and Dekimpe 2013; Liu and Shankar 2015; van Heerde, Helsen, and Dekimpe 2007) and on how firms learn and recover following product-harm crises (Cleeren, Dekimpe, and Helsen 2008; Gao et al. 2015; Rubel, Naik, and Srinivasan 2011). By examining the timing and incidence of product recalls, we generate valuable insights on the pre-recall process – an aspect of recall decisions that has not yet been investigated in the literature. Our findings reveal that firms experience competing pressures with respect to the timing of recall decisions. The satisficing responses they pursue are contingent on the brand's position in the marketplace.

Our main effect finding is that the time to recall for firms is higher when the problems encountered are severe in nature. When firms receive information about a potentially severe problem, they often turn inward which adversely influences their ability to find a quick solution to the problem. In addition, following severe problems, determining accountability is an important issue. Under such conditions, firms may be less likely to share information about the failure, preferring instead to protect themselves from the political fallout surrounding failure investigation (Madsen and Desai 2010). Thus, the threat of being held accountable may cause firms to be somewhat rigid in their response to investigations (Staw, Sandelands, and Dutton 1981).

However, we find significant variation in time to recall decisions for severe problems. Brand characteristics play a critical role in mitigating or exacerbating response times. When a higher reliability brand is faced with a severe problem, we find that firms are likely to make a recall decision faster than when the brand is perceived as less reliable. This finding is consistent with previous research findings that brands with higher reliability are penalized more by consumers for violating expectations (Liu and Shankar 2015; Rhee and Haunschild 2006). We argue that because of this increased market expectation, reliable brands tend to recall early and soften the damage to their reputational assets. Further, consistent with past research (Kalaignanam et al. 2013), this finding suggests that reliable brands should be better equipped to understand root causes for potential problems and provide remedial action in a timely manner.

Our study also finds that brand diversification can be an organizational liability in certain situations. Although diversified brands are able to cater to varying needs of the market, this characteristic impedes firm response when faced with a crisis. Specifically, the threat of demand side negative spillovers is significant for a more diverse brand. When a brand faces a safety investigation involving a severe problem, the specter of this incident spilling over to sub-brands in the portfolio looms large and adversely impacts the recall timing decision. This finding is consistent with research that finds negative spillovers after a brand scandal to be driven by the relatedness of the brands in the portfolio (Lei, Dawar, and Lemmink 2008). In some industries, firms are able to develop a diverse brand portfolio by using common components and systems (e.g., Fisher, Ramdas, and Ulrich 1999; Ramdas and Randall 2008). This supply side spillover exposes the sub-brands in the portfolio to greater risk when faced with a crisis and makes them more vulnerable. The pattern of results in our study should encourage researchers to further explore the growth versus vulnerability tradeoffs (e.g., higher time to recall) from brand diversification and the implications thereof for performance.

Perhaps the most interesting set of results involves the moderating impact of past recall intensity on the relationship between problem severity and time to recall. The results suggest that problem severity is less likely to lead to a recall delay when the firm has faced recalls in its recent past. While it may still be the case that managing these recalls puts constraints on the firm, the significant negative interaction effect shows that firms are more able or more willing to recall earlier, giving support to the learning argument. Past research suggests that negative publicity from recalls might actually heighten awareness toward the brand and the category (Cleeren, van Heerde, and Dekimpe 2013). Along similar lines, we find that increased publicity around past recalls has a significant impact in reducing the time to recall for subsequent investigations. Thus, negative publicity around previous product recalls acts as a catalyst in focusing firm attention.

We note that only a small proportion of recalls of a firm typically receive media attention. Therefore, minor product recalls that are not picked up by major media outlets may not be as effective in altering firm behavior. Collectively, the moderating influence of past recall intensity and direct influence of past publicity is encouraging from a public policy standpoint because large and publicized product recalls appear to have a positive effect in regulating firm responses to subsequent safety investigations by increasing their ability and motivation to respond to an investigation.

The *second* contribution is that we offer empirical evidence on the performance consequences of product recall timing decisions. The findings reveal that firms that take longer to recall are penalized more by stock markets than firms that recall earlier. Previous research has predominantly focused on testing the stock market reactions to product recall announcements and shown when the stock market reacts negatively to these announcements (Chen, Ganesan, and Liu 2009; Thirumalai and Sinha 2011). Our study diverges from this in showing that beyond the announcement of the recall, its timing also has value relevance for stock markets. Market analysts should care about the timing of product recalls, especially those for serious defects, because of its direct implications for cash flows and potential fines. Our study should encourage future researchers to include time to recall as a variable in models that examine the performance consequences of product recalls.

Managerial and Policy Implications

The findings from this manuscript have implications for both managers and policy makers. Whether and when firms should announce and implement a recall has been mostly looked at from a product safety perspective. Clearly, it is imperative that firms respond to significant product safety concerns. Recalls are initiated to remedy serious product defects to reduce product-associated injuries and accidents. Our study shows that not all firms are willing and able to respond to information about a severe defect quickly. Yet, the longer recall is delayed, the more likely it is that the product increases costs to the firm and society.

One key implication for managers is that the market reacts to the information embedded in the time that the firm takes to announce a recall during an investigation. We performed additional analyses to better understand the economic implications of the findings. We computed the direct impact of a unit change (i.e., marginal effects) in problem severity on time to recall and the indirect impact on shareholder wealth. Table 10 presents the results of this post-hoc analysis. The results suggest that a unit increase in problem severity increases time to recall by 309 days (at average levels of brand reliability), 304 days (at average levels of brand diversification) and 256 days (at average levels of past recall intensity).[5] In terms of economic significance, Table 10 suggests that the corresponding losses in shareholder wealth (at average levels of brand characteristics) because of delayed recalls in time to recall are $112 million, $109 million and $75 million, respectively. Thus, our finding implies that while recalls are adverse events for firms, delaying them when faced with severe problems imposes significant costs on firms in addition to the possibility of lawsuits and fines. The recent recall by General Motors in 2015 for faulty ignition switches was the outcome of a prolonged safety investigation that was initiated because of significant number of injuries and deaths in the market (Valdes-Dapena and Yellin 2015). GM was fined $900 million[6] and faces numerous lawsuits for the delay.[7]

--- Insert Table 10 about here ---

[5] The marginal effects at average levels of the moderators are different from each other because these are average marginal effects (and not marginal effects at the means). That is, the average marginal effect of a given moderator is computed over the entire range of observed values of other moderators rather than at their mean levels. For example, to compute average marginal effects at average levels of brand reliability, we set brand reliability at 4.1 (the mean) and estimate the average marginal effect over all the observed values of the other moderators (brand diversification and past recall intensity).

[6] http://www.freep.com/story/money/cars/general-motors/2015/09/16/gm-feds-near-settlement-ignition-switch-probe/32522599/ (Accessed on 2/22/16)

[7] http://michiganradio.org/post/one-gm-ignition-switch-lawsuit-down-five-go#stream/0 (Accessed on 2/22/16)

Moreover, managers need to recognize that multiple factors either impede or enable a fast recall decision. Our study finds evidence that brands with greater reputation for reliability are able to speed up recall decisions for severe problems. Therefore, brands with lower reliability face a significant challenge when responding to safety investigations. Lower reliability damages the competitive position of brands by preventing a quick response to safety investigations. In economic terms, we find that the difference in shareholder losses for lower and higher reliability brands because of recall timing is $84 million (see Table 10). We note that this economic loss is for a single safety investigation. Thus, if a lower reliability brand is faced with multiple safety investigations of severe problems, the economic losses because of delayed recalls would be even higher. The implication for lower reliability brands is that they need to invest at least some resources in building routines and procedures that enable them to respond to safety investigations in a timely manner. We caution that building a reputation for reliability is a long- run process that cannot occur instantaneously. Yet, often investments in reliability are not seen as priority because of the need to pursue growth through new products. In this context, the magnitude of the economic losses we document should provide a compelling rationale for brands to also invest in reliability improvements.

Our findings also suggest that brand diversification might be a double-edged sword for firms. In general, brand diversification enables firms to target multiple customer segments and pursue rapid growth. However, in many industries (e.g., personal computers, consumer electronics, automobiles), firms pursue brand extensions by using shared components or platforms across sub-brands in order to lower the cost of offering product variety (Ramdas and Randall 2008). Our study suggests that such sharing places firms at a disadvantage when investigating potential product recalls. That is, when such brands face an investigation of a severe problem, there is additional burden on diverse brands in assessing the scope of the problem as well as preparing the supply chain to remedy the defect. It is not surprising that when faced with an investigation for a power steering system for Chevrolet (i.e., a diverse brand in our sample), a senior spokesperson in the company noted "We are redoubling efforts on pending product reviews to bring them forward and to resolve them quickly. We will not sacrifice accuracy for speed…" (Healey and Meier 2014). Furthermore, the fear of spillover to other related products will also limit the motivation to undertake a quick recall for serious problems. The post-hoc analyses reveals that the difference between low and high diversification brands in time-to recall is 65 days and the subsequent shareholder losses because of recall delays is $103 million. From a policy standpoint, the implication is that regulators may need to be proactive in nudging diverse brands to respond to safety investigations in a timely manner.

Our findings should interest policy makers because product recalls appear to have a silver lining in the long run. In line with prior research, we find that firms are able to learn from prior recalls (e.g., Haunschild and Rhee 2004; Thirumalai and Sinha 2011) and make recall decisions more quickly for severe defects. Initially, we did not advance a hypothesis regarding the moderating effect of the brand's past recall intensity as managing several recalls in a short period could also pose a constraint on a firm's resources and attention. However, our results suggest that firms are more sensitive to severe problems and recall earlier when they have experienced recalls in recent time periods. In terms of economic significance, the difference in shareholder loss for low and high levels of past recall intensity is approximately $142 million (see Table 10). These differences are significant because firms often lament that product safety is over-regulated in North America and that many defects do not pose significant threats to consumers. Consistent with previous research (Kalaignanam et al. 2013), we find that product recalls, although expensive, have a positive role in altering firm behavior. Therefore, in assessing the shareholder losses because of product recalls, managers are well advised to adopt a broader perspective by taking into account how product recalls (involving severe problems) help them in responding early to new investigations and recouping some of the incurred losses.

Limitations and Future Directions

While the manuscript provides interesting insights into product recalls, the fact that the study is limited to the automobile industry implies that caution is warranted in generalizing our findings to other settings. The benefit of using data from a single industry is that we are able to improve the internal validity of the findings. In addition, the recall process varies across industries. The legal and other aspects that drive product investigations have substantial industry-specific idiosyncrasy that clear empirical assessments are potentially possible only using within-industry samples. In addition, since we are focusing on the delay in recall from the opening of a government investigation, the auto industry, unlike other industries, gives us a clear measure of recall delay. It is therefore not surprising that many studies on product recalls have also focused on a single industry setting rather than multiple industry settings (e.g., Borah and Tellis 2016; Hora, Bapuji, and Roth 2011; Rhee and Haunschild 2006; Rubel, Naik, and Srinivasan 2011; Thirumalai and Sinha 2011; van Heerde, Helsen, and Dekimpe 2007). Furthermore, it is important to note that the automotive industry is a highly relevant industry from an economic perspective, representing 3-3.5% of the GDP in the U.S. Nonetheless, an avenue for future research is to investigate firm's recall timing decisions in other industries to help in generating empirical generalizations.

Second, although our empirical analysis accounts for the direct and indirect costs of product recalls, the adverse reaction of stock markets does not necessarily imply that investors construe recall delays as stonewalling by firms. It is plausible that firms acknowledge the defect but simply lack the ability to issue a recall in a timely manner. More research is needed to disentangle and distinguish willful ignorance from the inability to respond faster to safety investigations. Likewise, although our study relied on theoretical arguments related to learning and problemistic search to develop our predictions, these effects are not estimated in the study. The challenge in explicitly measuring learning and search is that there are no reasonable proxies for these constructs. Therefore, it may not be appropriate to directly attribute the study's findings to these mechanisms. We hope our study spurs more research on the timing decisions of product recalls.

References

Argote, Linda and Henrich R. Greve (2007), “A Behavioral Theory of the Firm- 40 Years and Counting: Introduction and Impact,” *Organization Science,* 18 (3), 337-49.

Berman, Barry (1999), “Planning for the Inevitable Product Recall,” *Business Horizons,* 42 (2) 69-78.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004), “How Much Should We Trust Differences-in-Differences Estimates?,” *Quarterly Journal of Economics*, 119, 249–275.

Borah, Abhishek and Gerard J. Tellis (2016), “Halo (Spillover) Effects in Social Media: Do Product Recalls of One Brand Hurt or Help Rival Brands?,” *Journal of Marketing Research*, 53 (2), 143-160.

Carhart, Mark M. (1997), "On Persistence in Mutual Fund Performance," *The Journal of Finance,* 52 (1), 57-82.

Carroll, Glenn R. and Anand Swaminathan (2000), "Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the US Brewing Industry," *American Journal of Sociology,* 106 (3), 715-762.

Chao, Gary H., Seyed M.R. Iravani, and R. Canan Savaskan (2009), “Quality Improvement Incentives and Product Recall Cost Sharing Contracts," *Management Science,* 55 (7), 1122-38.

Cheah, Eng Tuck, Wen Li Chan, and Corinne Lin Lin Chieng (2007), “The Corporate Social Responsibility of Pharmaceutical Product Recalls: An Empirical Examination of U.S. and U.K. Markets,” *Journal of Business Ethics*, 76 (4), 427-449.

Chen, Yubo, Shankar Ganesan, and Yong Liu (2009), “Does a Firm’s Product-Recall Strategy Affect Its Financial Value? An Examination of Strategic Alternatives During Product-Harm Crises,” *Journal of Marketing*, 73 (6), 214-26.

Chen, Ming-Jer and Donald C. Hambrick (1995), “Speed, Stealth, and Selective Attack: How Small Firms Differ from Large Firms in Competitive Behavior,” *Academy of Management Journal*, 38 (2), 453-482.

Cleeren, Kathleen, Marnik G. Dekimpe, and Kristiaan Helsen (2008), "Weathering Product-Harm Crises," *Journal of the Academy of Marketing Science,* 36 (2), 262-270.

----, Harald J. van Heerde, and Marnik G. Dekimpe (2013), “Rising from the Ashes: How Brands and Categories Can Overcome Product-harm Crises,” *Journal of Marketing,* 77 (2), 58-77.

Cyert, Richard M. and James G. March (1963, 1992), "A Behavioral Theory of the Firm," *Englewood Cliffs, NJ,* 2.

Dani, Samir and Aman Deep (2010), “Fragile Food Supply Chains: Reacting to Risks,” *International Journal of Logistics: Research & Applications*, 13 (5), 395-410.

Dawar, Niraj and Madan M. Pillutla (2000), “Impact of Product-Harm Crises on Brand Equity: The Moderating Role of Consumer Expectations,” *Journal of Marketing Research*, 37 (2), 215-26.

Dobrev, Stanislav D. (2000), “Decreasing Concentration and Reversibility of the Resource Partitioning Process: Supply and Deregulation in the Bulgarian Newspaper Industry,” *Organization Studies,* 21 (2), 383–404.

Fama, Eugene F. (1998), “Market Efficiency, Long-Term Returns, and Behavioral Finance.” *Journal of Financial Economics*, 49 (3), 283-306.

---- and Kenneth R. French (1993), "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics,* 33 (1), 3-56.

Fisher, Marshall, Kamalini Ramdas, and Karl Ulrich (1999), "Component Sharing in the Management of Product Variety: A Study of Automotive Braking Systems," *Management Science,* 45 (3), 297-315.

Folkes, Valerie S. and Barbara Kotsos (1986), "Buyers' and Sellers' Explanations for Product Failure: Who Done It?," *Journal of Marketing,* 50 (2), 74-80.

Gao, Haibing, Jinhong Xie, Qi Wang, and Kenneth C. Wilbur (2015), "Should Ad Spending Increase or Decrease Before a Recall Announcement? The Marketing-Finance Interface in Product-Harm Crisis Management," *Journal of Marketing,* 79 (5), 80-99.

Gatignon, Hubert, Erin Anderson, and Kristiaan Helsen (1989), "Competitive Reactions to Market Entry: Explaining Interfirm Differences," *Journal of Marketing Research*, 26 (1), 44-55.

Geyskens, Inge, Katrijn Gielens, and Marnik G. Dekimpe (2002), "The Market Valuation of Internet Channel Additions," *Journal of Marketing*, 66 (2), 102-119.

Haunschild, Pamela R. and Mooweon Rhee (2004), "The Role of Volition in Organizational Learning: The Case of Automotive Product Recalls," *Management Science*, 50 (11), 1545-60.

---- and Bilian N. Sullivan (2002), "Learning from Complexity: Effects of Prior Accidents and Incidents on Airlines' Learning," *Administrative Science Quarterly,* 47 (4), 609-43.

Healey, James and Fred Meier (2014), "Other GM Cars Have Recalled Cobalt Steering," *USA Today,* March 21.

Heath, Timothy B. and Subimal Chatterjee (1995), "Asymmetric Decoy Effects on Lower-Reliability versus Higher-Reliability Brands: Meta-analytic and Experimental Evidence," *Journal of Consumer Research*, 22 (3), 268-84.

Hendricks, Kevin B. and Vinod R. Singhal (1996), "Quality Awards and the Market Value of the Firm: An Empirical Investigation," *Management Science,* 42 (3), 415-436.

Hora, Manpreet, Hari Bapuji, and Aleda V. Roth (2011), "Safety Hazard and Time to Recall: The Role of Recall Strategy, Product Defect Type, and Supply Chain Player in the U.S. Toy Industry," *Journal of Operations Management*, 29 (7/8), 766-77.

Ingram, Paul and Joel A.C. Baum (1997), "Opportunity and Constraint: Organizations' Learning from the Operating and Competitive Experience of Industries," *Strategic Management Journal,* 18 (1), 75-98.

Isidore, Chris and Rene Marsh (2014), "GM to Pay $35 Million over Delayed Recall," (accessed January 13, 2016), [available at http://money.cnn.com/2014/05/16/news/companies/gm-nhtsa/].

Jarrell, Gregg and Sam Peltzman (1985), "The Impact of Product Recalls on the Wealth of Sellers," *Journal of Political Economy*, 93 (3), 512-536.

Jayachandran, Satish and Rajan Varadarajan (2006), "Does Success Diminish Competitive Responsiveness? Reconciling Conflicting Perspectives," *Journal of the Academy of Marketing Science*, 34 (4), 284-94.

Jayaraman, Vaidyanathan, Raymond A. Patterson, and Erik Rolland (2003), "The Design of Reverse Distribution Networks: Models and Solution Procedures," *European Journal of Operational Research,* 150 (1), 128-49.

Jensen, Christopher (2015), "Ford Focus Steering Problem Prompts Many Complaints, but No Inquiry," (accessed January 27, 2016), [available at http://www.nytimes.com/2015/12/25/automobiles/ford-focus-steering-problem-prompts-many-complaints-but-no-inquiry.html].

Kalaignanam, Kartik, Tarun Kushwaha, and Meike Eilert (2013), "The Impact of Product Recalls on Future Product Reliability and Future Accidents: Evidence from the Automobile Industry," *Journal of Marketing,* 77 (2), 41-57.

----, ----, Jan-Benedict EM Steenkamp, and Kapil R. Tuli (2013), "The Effect of CRM Outsourcing on Shareholder Value: A Contingency Perspective," *Management Science,* 59 (3), 748-769.

Keller, Kevin Lane (1993), "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (1), 1-22.

Koyitty, Bijoy Anadoth (2012), "Toyota Fined $17.35 Million for U.S. Floormat Recall Delay,"(accessed January 13, 2016), [available at http://www.reuters.com/article/us-toyota-penalty-idUSBRE8BH0J820121218].

Lee, Ruby P., and Rajdeep Grewal (2004), "Strategic Responses to New Technologies and Their Impact on Firm Performance," *Journal of Marketing*, 68 (4), 157-71.

Lei, Jing, Niraj Dawar, and Jos Lemmink (2008), "Negative Spillover in Brand Portfolios: Exploring the Antecedents of Asymmetric Effects," *Journal of Marketing,* 72 (3), 111-123.

Liang, Kung-Yee and Scott L. Zeger (1986), "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika,* 73 (1), 13-22.

Liu, Yan and Venkatesh Shankar (2015), "The Dynamic Impact of Product-Harm Crises on Brand Equity and Advertising Effectiveness: An Empirical Analysis of the Automobile Industry," *Management Science,* 61 (10), 2514-2535.

Mackelprang, Alan W., Marco Habermann, and Morgan Swink (2015), "How Firm Innovativeness and Unexpected Product Reliability Failures Affect Profitability," *Journal of Operations Management,* 38, 71-86.

Madsen, Peter M. and Vinit Desai (2010), "Failing to Learn? The Effects of Failure and Success on Organizational Learning in the Global Orbital Launch Vehicle Industry," *Academy of Management Journal,* 53 (3), 451-76.

Marucheck, Ann, Noel Gries, Carlos Mena, and Linning Cai (2011), "Product Safety and Security in the Global Supply Chain: Issues, Challenges and Research Opportunities," *Journal of Operations Management*, 29 (7/8), 707-20.

McCurry, Justin (2010), "Toyota Memo Shows $100m Savings Over Floor Mat Recall," (accessed January 27, 2016), [available at http://www.theguardian.com/business/2010/feb/22/toyota-memo-safety].

Mitchell, Mark L. and Erik Stafford (2000), "Managerial Decisions and Long-Term Stock Price Performance," *The Journal of Business,* 73 (3), 287-329.

Moulton, Brent R. (1986), "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics,* 32 (3), 385-397.

Petrin, Amil and Kenneth Train (2010), "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47 (1), 3-13.

Protess, Ben, Evelyn M. Rusli, and Susanne Craig (2011), "Buffet's Handling of Deputy Baffles Some Experts," (accessed June 10, 2011), [available at http://dealbook.nytimes.com/2011/03/31/buffetts-handling-of-deputy-baffles-some-experts/?ref=global].

Raassens, Néomie, Stefan Wuyts, and Inge Geyskens (2014), "The Performance Implications of Outsourcing Customer Support to Service Providers in Emerging versus Established Economies," *International Journal of Research in Marketing*, 31 (3), 280-292.

Ramdas, Kamalini and Taylor Randall (2008), "Does Component Sharing Help or Hurt Reliability? An Empirical Study in the Automotive Industry," *Management Science,* 54 (5), 922-38.

Rhee, Mooweon (2009), "Does Reputation Contribute to Reducing Organizational Errors? A Learning Approach," *Journal of Management Studies*, 46 (4), 676-703.

---- and Pamela R. Haunschild (2006), "The Liability of a Good Reputation: A Study of Product Recalls in the U.S. Automotive Industry," *Organization Science*, 17 (1), 101-17.

Rubel, Olivier, Prasad A. Naik, and Shuba Srinivasan (2011), "Optimal Advertising When Envisioning a Product-Harm Crisis," *Marketing Science,* 30 (6), 1048-65.

Rupp, Nicholas G. (2004), "The Attributes of a Costly Recall: Evidence from the Automotive Industry," *Review of Industrial Organization*, 25 (1), 21-44.

---- and Curtis R. Taylor (2002), "Who Initiates Recalls and Who Cares? Evidence from the Automobile Industry," *Journal of Industrial Economics*, 50 (2), 123-49.

Siomkos, George J. and Gary Kurzbard (1994), "The Hidden Crisis in Product-Harm Crisis Management," *European Journal of Marketing*, 28 (2), 30-41.

Sitkin, Sim B. (1992), "Learning Through Failure - The Strategy of Small Losses," *Research in Organizational Behavior,* 14, 231-66.

Smith, Ken G., Curtis M. Grimm, Martin J. Gannon, and Ming-Jer Chen (1991), "Organizational Information Processing, Competitive Responses, and Performance in the U.S. Domestic Airline Industry," *Academy of Management Journal*, 34 (1), 60-85.

Smith, Craig N., Robert J. Thomas, and John A. Quelch (1996), "A Strategic Approach to Managing Product Recalls," *Harvard Business Review*, 74 (5), 102-12.

Sorescu, Alina, Venkatesh Shankar, and Tarun Kushwaha (2007), "New Product Preannouncements and Shareholder Value: Don't Make Promises You Can't Keep." *Journal of Marketing Research*, 44 (3), 468-489.

Stahl, Florian, Mark Heitmann, Donald R. Lehmann, and Scott A. Neslin (2012), "The Impact of Brand Equity on Customer Acquisition, Retention, and Profit Margin," *Journal of Marketing,* 76 (4), 44-63.

Staw, Barry M., Lance E. Sandelands, and Jane E. Dutton (1981), "Threat Rigidity Effects in Organizational Behavior: A Multilevel Analysis," *Administrative Science Quarterly,* 26 (4), 501-24.

Tang, Christopher S. (2008), "Making Products Safe: Process and Challenges," *International Commerce Review*, 8 (1), 48-55.

Thirumalai, Sriram and Kingshuk K. Sinha (2011), "Product Recalls in the Medical Device Industry: An Empirical Exploration of the Sources and Financial Consequences," *Management Science*, 57 (2), 376-92.

Thomas, Ken (2005), "Government to Investigate 3.7 Million Ford Motor Co. Pickups and SUVs," *The Associated Press*, March 23, 2005.

Valdes-Dapena, Peter and Tal Yellin (2015), "GM: Steps to a Recall Nightmare." (accessed February 13, 2016), [available at http://money.cnn.com/infographic/pf/autos/gm-recall-timeline/].

van Heerde, Harald, Kristiaan Helsen, and Marnik G. Dekimpe (2007), "The Impact of a Product-Harm Crisis on Marketing Effectiveness," *Marketing Science*, 26 (2), 230-43.

Wieder, Ben (2011), "NHTSA Doesn't Analyze its Recall Data to ID Worst Car Manufacturers, GAO Says," (accessed December 9, 2011), [available at www.iwatchnews.org/node/4925].

Table 1
LITERATURE OVERVIEW ON PRODUCT RECALLS

Study	Performance Outcomes	Recall Timing	Pre-recall/ Post-recall	Industry	Relevant Findings
Borah and Tellis (2016)	✓		Post-recall	Automotive	Product recalls create negative chatter for not only the affected brand but also other brands of the same firm or in the same category (perverse halo). This chatter negatively impacts firm performance.
Chen, Ganesan, and Liu (2009)	✓		Post-recall	Consumer goods	The stock market responds negatively to recalls announced before the firm receives any reports of injuries. Firms with higher brand quality are less likely to announce a proactive recall.
Cleeren, van Heerde, and Dekimpe (2013)	✓		Post-recall	FMCG	When firms acknowledge blame in a crisis, they are more likely to experience decreases in advertising effectiveness and increases in price sensitivity. Crisis publicity increases advertising effectiveness for the brand and category.
Dawar and Pillutla (2000)			Post-recall	Soft drinks	While strong expectations buffer the impact of a crisis on brand equity (relative to weak expectations), consumers punish those brands more for a stonewalling response than response that is ambiguous or supportive.
Hora, Bapuji, and Roth (2011)		✓	Post-recall	Toys	The lag between product sold and recall is influenced by the type of product defect and entity announcing the recall.
Liu and Shankar (2015)	✓		Post-recall	Automotive	Recalls negatively impact brand sales and advertising effectiveness. The reaction is more negative when the recall is severe, faces greater attention, and involves higher quality brands.
Rubel, Naik, and Srinivasan (2011)	✓		Post-recall	Automotive	To recover from a crisis, firms should reduce advertising spending before and increase it after the crisis. This recommendation is supported using data from the Ford Rollover crisis.
van Heerde, Helsen, and Dekimpe (2007)	✓		Post-recall	Food	Product harm crisis reduces a brand's performance and the effectiveness of its marketing instruments. The brand also becomes more vulnerable to competitive effects.
This study	✓	✓	Pre-recall	Automotive	In an opened safety investigation, the severity of the problem enhances time to recall. However, this effect is weaker when brands have 1) higher reliability and 2) greater intensity of recalls in the past. In contrast, the effect is stronger for brands that are diverse. Finally, the time to recall is negatively related to stock market performance.

Table 2
OVERVIEW OF RATIONALE FOR HYPOTHESES

Main effects/ Moderator effects	Ability to Recall Faster	Motivation to Recall Faster	Net Effects on Time to Recall	Rationale
Problem severity (H1)	**-**[a]	**-**[a]	*Increase* [b]	• Triggers problemistic search • Recall is more likely to be punished • Need to determine internal accountability
Problem severity x brand reliability (H2)	+	n.a.	*Decrease*	• Late recall can damage reputation • More likely to have systems in place to determine root cause of failure
Problem severity x brand diversification (H3)	-	**-**	*Increase*	• Need to test multiple products for presence of defect • Lower ability to integrate complex knowledge
Problem severity x past brand recall intensity	-	+	**?**	• More likely to have knowledge stock • Likely to have resource constraints, diverted attention

Plus (negative) signs represent strengthening (weakening) of the effect in question. The net effect on time to recall is given in the fourth column.
[a] Problem severity will lower ‘Ability to Recall Faster’ and ‘Motivation to Recall Faster’.
[b] The net effect of problem severity is an ‘Increase’ in Time to Recall.

Table 3
DATA SOURCES AND OPERATIONALIZATION

Variable	Operationalization	Data Sources
Time to recall *(TIMETORECALL)*	Time between opening date of investigation and the date of recall (in days).	NHTSA
Stock market reaction *(CAR)*	Cumulative abnormal returns in the [-2, 2] window around the recall announcement.	CRSP
Problem severity (*PROBSEV)*	Principal component score of number of a) complaints, b) crashes/fires, c) injuries, and d) fatalities. Higher scores indicate more severe problems	NHTSA
Brand reliability *(RELIABILITY)*	3-year average of problem scores of all models of a make (5=fewest problems, 1=most problems) (Rhee and Haunschild 2006). Higher scores indicate more reliable brands.	Consumer Reports
Brand diversification (*DIVERSE*)	Principal component score of a) number of models and b) variation in engine sizes across models (Rhee and Haunschild 2006). Higher scores indicate more diverse brands.	Ward's Automotive Yearbook Automotive News Market Data Book Consumer Reports
Past recall intensity (*PASTRECINT)*	Count of the total number of vehicles recalled by the brand in the previous year. Higher numbers indicate greater past recall intensity.	NHTSA
Firm profitability *(ROA)*	Return on assets= Net income/ Total assets	Compustat
Recall size *(RECSIZE)*	Number of vehicles investigated for the defect.	NHTSA
Past publicity (*PASTPUBLICITY)*	Count of the number of recalls in the past year that received coverage in the *Wall Street Journal.*	Lexis-Nexis
Investigation type (*INVTYPE)*	SQ = 1 if investigation opened as service query, 0 otherwise, RQ = 1 if investigation opened as recall query, 0 otherwise, EA = 1 if investigation upgraded to engineering analysis, 0 otherwise.	NHTSA
Brand sales (*BRANDSALES)*	Sales of the brand involved in the recall.	Compustat

Table 4

MODEL COMPARISONS WITH DIFFERENT DISTRIBUTIONS OF TIME TO RECALL

Distributions	Log Likelihood	BIC Statistic
Exponential	-323.22	728.53
Log normal	-317.71	717.51
Log logistic	-342.47	767.03
Weibull	*-312.40*	*706.89*

Table 5a

DESCRIPTIVE STATISTICS & CORRELATION MATRIX

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1.TIMETORECALL	1									
2. PROBSEV	.47	1								
3. RELIABILITY	-.17	-.04	1							
4. DIVERSE	.09	.04	-.05	1						
5. PASTRECINT	.12	.18	.09	.21	1					
6. ROA	-.07	-.01	-.02	-.11	-.06	1				
7. RECSIZE	.18	.30	-.01	.16	.07	.01	1			
8. PASTPUBLICITY	-.07	.04	.07	.08	.47	-.06	-.01	1		
9. BRANDSALES	.05	.13	.14	.66	.39	-.09	.24	.25	1	
10. CAR[-2,2]	-.18	-.13	.03	.10	.06	-.34	-.04	.04	.03	1
Mean	295	.07	4.12	.01	2.51	.02	182197	5.22	996217	-.6
S.D	248	1.37	.35	1.14	6.46	.12	291971	20.80	1024140	.4

Table 5b

MEAN TIME TO RECALL BY MANUFACTURERS

Manufacturer	Time to Recall (days)
Chrysler	293.89
Ford	360.95
General Motors	375
Honda	265.5
Hyundai	249.28
Nissan	265.71
Toyota	213.91
Volkswagen	229.61
Volvo	285

Only manufacturers with at least 10 recalls are included.

Table 6

WEIBULL AFT REGRESSION RESULTS OF THE PREDICTORS OF TIME TO RECALL

Variable	Column I (Random firm effects)	Column II (Fixed firm effects)	Column III (Problem severity measure excluding complaints)
PROBSEV (α_1)	.84 ***	.90 **	.18 **
	[.32]	[.37]	[.07]
RELIABILITY (α_2)	-.29 **	-.20 *	-.38 **
	[.12]	[.11]	[.15]
DIVERSE (α_3)	.06	.10 **	.05
	[.04]	[.04]	[.04]
PASTRECINT (α_4)	-.008	-.002	-.004
	[.0072]	[.005]	[.006]
*PROBSEV * RELIABILITY (α_5)*	-.18 **	-.19 **	-.037
	[.08]	[.09]	[.08]
*PROBSEV * DIVERSE (α_6)*	.08 **	.09 **	.10 ***
	[.03]	[.04]	[.03]
*PROBSEV * PASTRECINT (α_7)*	-.0074 ***	-.0079 ***	-.006 ***
	[.0011]	[.0021]	[.0009]
ROA (α_8)	-.13	-.14	-.09
	[.12]	[.32]	[.12]
RECSIZE (α_9)$^{\$}$	.025	.082	.06 *
	[1.08]	[1.25]	[.03]
PASTPUBLICITY (α_{10})	-.006 ***	-.006 **	-.006 ***
	[.001]	[.003]	[.001]
Investigation Type Fixed Effects ($\alpha_{11\text{-}13}$)	2 significant	2 significant	2 significant
Unobserved Firm Effects	Clustered standard errors	Fixed effects One significant	Clustered standard errors
Shape Parameter	1.76 ***	1.79 ***	1.70 ***
	[.12]	[.09]	[.09]
LL	-312.40	-306.19	-318.16
Number of Observations	352	352	352

* *p*<.10, ** *p*<.05, *** *p*<.01, Clustered robust standard errors in parentheses in columns I & III, and robust standard errors in column II. The dependent variable is *TIMETORECALL*.

$ Unit is ten million vehicles

Table 7
RESULTS OF THE IMPACT OF TIME TO RECALL ON SHORT-TERM CUMULATIVE ABNORMAL RETURNS

Variable	Column I CAR[-2,2]	Column II CAR[-1,1]	Column III All variables from first stage (no control for endogeneity)	Column IV CAR[-2,2] Bootstrapped SE
TIMETORECALL(η1)	-4.11×10^{-5} **	-1.34×10^{-5} **	-4.8×10^{-5} **	-4.11×10^{-5} **
	[1.71×10^{-5}]	[4.7×10^{-6}]	[1.81×10^{-5}]	[1.76×10^{-5}]
RECSIZE (η2)	-2.6×10^{-5} **	-9.83×10^{-6}	-.061	-2.6×10^{-5} **
	[1.6×10^{-5}]	[1.75×10^{-5}]	[.11]	[1.81×10^{-5}]
BRANDSALES (η3)	-.004	-.004	-.026 *	-.004
	[.005]	[.004]	[.016]	[.007]
PROBSEV (η4)	-.001	-.0026 **	-.0028 *	-.001
	[.001]	[.0003]	[.0017]	[.001]
RESIDUALFS (η5)	-7.28×10^{-6}	-3.25×10^{-6}		-7.28×10^{-6}
	[1.74×10^{-5}]	[5.12×10^{-6}]		[1.81×10^{-5}]
RELIABILITY (η6)			-.013	
			[.017]	
DIVERSE (η7)			.0049	
			[.0058]	
PASTRECINT (η8)			-.0017	
			[.0042]	
*PROBSEV * RELIABILITY (η9)*			.007	
			[.008]	
*PROBSEV * DIVERSE (η10)*			.0008	
			[.004]	
*PROBSEV *PASTRECINT (η11)*			-9.21×10^{-5}	
			[1.54×10^{-4}]	
ROA (η12)			.097	
			[.10]	
PAST PUBLICITY (η13)			-.0004	
			[.001]	
Investigation Type Fixed Effects (η_{14-15})			None significant	
Unobserved Firm Effects	Clustered standard errors	Clustered standard errors	Clustered standard errors	Clustered & bootstrapped standard errors (500 replications)
Number of Observations	73	73	73	73
R- Square	.09	.09	.28	.09

* *p*<.10, ** *p*<.05, *** *p*<.01, Clustered robust standard errors in parentheses.
The dependent variable is cumulative abnormal returns (CAR) over different windows.

Table 8
ROBUSTNESS ANALYSES

Variable	Column I Including manufacturer headquarters dummies	Column II Including luxury brand dummy	Column IIIa: Split population duration model Incidence of product recall: Logit	Column IIIb: Split population duration model Time to recall: Weibull
PROBSEV	.82 ***	.8 **	.50 ***	.76 **
	[.36]	[.36]	[.21]	[.34]
Asia Based Manufacturer	-.015			
	[.11]			
U.S. Based Manufacturer	.04			
	[.10]			
Luxury Brand		-.039		
		[.079]		
RELIABILITY	-.29 *	-.29 **	-.34	-.17
	[.14]	[.13]	[.27]	[.18]
DIVERSE	.06	.06	.25 *	.09 *
	[.04]	[.05]	[.13]	[.05]
PASTRECINT	-.0022	-0079	.017	-.0058
	[.0052]	[.0059]	[.032]	[.0042]
*PROBSEV * RELIABILITY*	-.17 *	-.17 *	-.12 **	-.070 **
	[.09]	[.09]	[.06]	[.025]
*PROBSEV * DIVERSE*	.08 **	.09 ***	-.23	-.18 ***
	[.04]	[.03]	[.18]	[.05]
*PROBSEV * PASTRECINT*	-.0074 ***	-.0075 ***	.09 ***	-.0026 ***
	[.001]	[.0017]	[.03]	[.0009]
ROA	-.09	-.08	.15	.29 *
	[.12]	[.12]	[.13]	[.17]
RECSIZE	-.23	-.31	.29 *	-.11
	[1.12]	[1.05]	[.15]	[.13]
PAST PUBLICITY	-.006	-.006	.04 **	-.004
	[.001]	[.001]	[.02]	[.003]
Investigation Type Fixed Effects	2 of 3 significant	2 of 3 significant	None significant	2 of 3 significant
Unobserved Firm Effects	Clustered standard errors	Clustered standard errors	Clustered standard errors	Clustered standard errors
Intercept	3.27 ***	3.26 ***	.95 ***	.76 **
	[.64]	[.56]	[.33]	[.31]
Shape Parameter	1.76 ***	1.77 ***	1.23 ***	
	[.12]	[.12]	[.22]	
LL	-311.02	-312.34	-707.32	
Number of Observations	352	352	352	

Table 9
ROBUSTNESS CHECKS: LONG TERM ABNORMAL RETURNS

	Average month in (0,12)	Average month in (0,24)
Intercept (Abnormal returns)	-.0061	-.0049
Market factor (Beta)	1.2945 ***	1.2753 ***
Size factor (SMB)	-.1030	-.0880
Book to market factor (HML)	.6384 ***	.5962 ***
Adj- R^2	42.54%	42.56%

* $p<.10$, ** $p<.05$, *** $p<.01$, Standard errors are heteroskedasticity consistent.

Table 10
POST HOC ANALYSES: ASSESSING THE MANAGERIAL RELEVANCE OF THE FINDINGS

Marginal Effects of Problem Severity Under Conditions of	Predicted Time to Recall (days)	Net Present Value of Shareholder Loss
Low Brand Reliability	357	($140m)
Average Brand Reliability	309	($112m)
High Brand Reliability	259	($56m)
Low Brand Diversification	279	($34m)
Average Brand Diversification	304	($109m)
High Brand Diversification	344	($137m)
Low Past Recall Intensity	368	($156m)
Average Past Recall Intensity	256	($75m)
High Past Recall Intensity	172	($14m)

Note: Low, Average, and High Values are set at the 10th, 50th, and 90th percentile.
Net present value is computed as CAR * Market Capitalization. The average market capitalization of firms in the sample is $28 billion.
Marginal effects refers to the change in the expected value of the dependent variable for a unit change in the independent variable.