

# **Plant Operations and Product Recalls in the Automotive Industry: An Empirical Investigation**

Rachna Shah, Associate Professor\*\*\*  
Supply Chain and Operations Department  
Carlson School of Management  
University of Minnesota  
321 Nineteenth Avenue South  
Minneapolis, MN 55455-0438  
E-mail: [shahx024@umn.edu](mailto:shahx024@umn.edu)

George Ball, Ph.D. Student  
Supply Chain and Operations Department  
Carlson School of Management  
University of Minnesota  
E-mail: [ball0197@umn.edu](mailto:ball0197@umn.edu)

Serguei Netessine  
INSEAD  
Email: [Serguei.NETESSINE@insead.edu](mailto:Serguei.NETESSINE@insead.edu)

\*\*\* Corresponding author

# Plant Operations and Product Recalls in the Automotive Industry: An Empirical Investigation

---

While there is overwhelming evidence of the negative consequences of product recalls, empirical evidence of plant-level drivers of recalls is non-existent. We examine potential plant-level recall causes for North American automotive manufacturers by combining production-line data over a 7-year period with auto-recall data, while differentiating between manufacturing- and design-related recalls. We find that 1) increasing variety increases design recalls but only increases manufacturing recalls in the presence of high utilization or low focus; 2) high utilization has a negative effect on both types of recalls; and 3) focus reduces manufacturing recalls. In quantitative terms, we show that 5% increase in utilization leads to an average increase of about 13.6% in manufacturing- and 9.8% in design-recalls. Installing three additional options, an increase of roughly one standard deviation, results in an increase of 30% in design recalls. Further, moving from an unfocused to a focused factory can lead to a 74% reduction in manufacturing-recalls.

---

## 1.0. Introduction

Recent years have seen a significant increase in product recalls, both in terms of the number of recalls and the number of units impacted per recall. This trend is observed across industries and product categories, ranging from food and toys to automobiles. According to the Consumer Product Safety Commission, the number of consumer products being recalled from the marketplace is increasing exponentially. Recalls related to food products using peanut butter in 2009 and the massive lead related toy recall in 2007 are etched in consumers' psyches due to heavy media coverage. The National Highway Traffic and Safety Administration (NHTSA) has reported similar trends in the automotive industry. According to NHTSA, the average number of auto recalls per million registered U.S. vehicles has risen steadily, from 3.10 in 1980s to 8.25 in the 1990s and 11.79 between 2000 and 2010. This is significant, since each auto recall is associated with a potential economic consequence of \$20 million or more<sup>1</sup>. Equally significant, five of the top 10 largest auto recalls in U.S. history occurred in the last 10 years (www.nhtsa.gov).

Product recalls like these are associated with considerable present and future costs for a company. These include inconvenience to consumers, drops in financial measures of performance, and even the loss of human life. The well-publicized Toyota recall in 2009 and 2010 that resulted in serious accidents and loss of life may have cost the company over \$1 billion<sup>2</sup>. Topps, one of the largest makers of hamburgers in the United States, went into bankruptcy after the recall of 21.7 million frozen hamburgers in September 2007. And although Merck, one of the largest pharma manufacturers, avoided a similar fate, its stock

---

<sup>1</sup> Jarrell and Peltzman (1985) estimate an average per car recall cost of \$200. This cost includes repair, replacement, and lost sales. An average of 100,000 cars were involved in each auto recall from 1980 to 2013. <http://www-odi.nhtsa.dot.gov/downloads/>.

<sup>2</sup> <http://www.investopedia.com/financial-edge/0210/the-cost-of-an-auto-recall.aspx>

price plummeted from \$45.07 to \$33 when it recalled a medication called Vioxx on September 30, 2004—it was the company’s largest single-day price drop ever. Academic research also highlights the negative impact of product recalls documented in business press. For example, product recalls reduce shareholder wealth (Jarrell and Peltzman, 1985), stock prices (Thirumalai and Sinha, 2011), and organizational learning (Haunschild and Rhee, 2004), and suggest poor future product safety, higher injury rate, and increases in recall frequency (Kalaiganam et al., 2013).

Product recalls have been studied in marketing, economics, strategy, and operations management. An extensive and careful review of the related literature highlights that most of the existing research has focused on the consequences of product recalls in which the researcher models product recall as the independent variable to examine its impact on firms’ financial and market performance. These studies unequivocally demonstrate that product recalls are negatively associated with various measures of firm performance; at the same time, they point to the paucity of studies examining causes of product recalls.

Two noteworthy exceptions to “consequence-focused” research include Haunschild and Rhee (2004) and Thirumalai and Sinha (2011), both of which model product recalls as a dependent variable. While Haunschild and Rhee (2004) examine whether the initiator of a past product recall (firm or a regulatory agency) impacts the likelihood of future recalls, Thirumalai and Sinha (2011) analyze the impact of a firm’s R&D focus on future recalls. These two studies represent a first step towards highlighting the importance of examining drivers of product recalls. Even so, neither study focuses on operational drivers of product recalls—they conceptualize the causes at a high level of abstraction.

Further, in our comprehensive literature review, we did not find a single study investigating the operational characteristics of the production process as the cause or contributing causes of product recalls. Additionally, our literature review demonstrated that all existing studies considered product recalls as homogenous but in practice, product recalls are frequently attributed either to the manufacturing process or to the design process. It is important to make this distinction because the same operational causes may have a differing relationship with the two types of recalls. In the current study, we address these two gaps in the literature by investigating the impact of operational characteristics of the production process and product attributes on manufacturing- and design-related recalls.

It is critical to understand operational causes because managers face significant and often contradictory challenges. They are mandated to help firms increase revenue and market share by producing multiple end-product configurations to satisfy diverse customer needs and they must do so by using key resources efficiently so that the firm becomes or remains cost competitive. This bifurcated focus on variation and cost-savings is substantiated by the dramatic increases in variety in every product category, from toothpaste to running shoes to automobiles (Gopal et al., 2013; Shah and Shin, 2007; Fisher and Ittner, 1999; Moreno and Terwiesch, 2013) and by the worldwide popularity and

implementation of “lean” tools and practices over the last two decades (Staats et al., 2011; Shah and Ward, 2007). It is well established that increasing variety on the production floor as opposed to *delayed variegation* (e.g., postponement—Ramdas, 2003) increases complexity of the production process (MacDuffie et al., 1996). Similarly, managers are often hesitant to make new labor and capital investments, opting instead to work current employees harder when demand increases. This decision can “take a toll on the body”, be “grueling” for employees (Rogers, 2013), and increase the potential for errors (Adler et al., 1997).

In this study, we first examine whether operational characteristics such as variety and efficiency, individually or jointly, impact manufacturing- and design-related product recalls. We then identify a strategic lever that firms can use to mitigate the negative relationship between variety and product recall: the decision to allocate major product lines to its production facilities. Specifically, a firm must decide whether to manufacture single or multiple product lines in a production facility, an approach popularized under the “focused factory” umbrella (Skinner, 1974). The operational benefits of focused facilities are well established in conceptual and empirical research in manufacturing, healthcare, and the airline industry (McDermott and Stock, 2011; Tsikritsis, 2007). In this study, we examine whether focus can also offset the negative impact of variety on the incidence of product recall. We study our research questions in the context of the automotive industry using a unique dataset compiled by linking automotive recalls to operational data collected from North American automotive assembly lines.

We make three contributions to existing research. First, we identify operational causes of manufacturing- and design-related product recalls and examine them using plant-level data. While the relationships are of significant theoretical interest, they are also important to the managerial community, given the increasing frequency and scope of product recalls and their potentially devastating financial effects. Second, for managers facing contrasting goals, we demonstrate how an increase in focus *or* a decrease in utilization can ameliorate the negative effects of variety on recalls. Finally, we provide empirical support for *direct* relationships between operational variables and higher levels of firm performance, relationships that are frequently hypothesized, but rarely empirically identified. Previous researchers suggest that this may be because superior operational performance is essential to appropriate market and financial benefits from structural and infrastructural decisions (Scherer and Ross, 1990). Establishing a direct relationship is thus of significant theoretical and empirical interest.

## **2.0. Literature review and research hypotheses**

### **2.1. Research context**

We chose the automotive industry as the context for our study for three reasons. First, an increase in product variety and manufacturing efficiency has become the norm in the auto industry, making it a

fertile setting to study their individual and joint effects. While mergers have reduced the number of dominant manufacturers, end-product configurations have proliferated because each manufacturer (e.g., Toyota) produces multiple brands (e.g., Toyota, Lexus) with multiple models (e.g., Camry, LS 600), platforms (e.g., sedan, SUV), and body styles (e.g., 2-door vs. 4-door sedan). Each facility turns out a large number of distinctly different end products, all with hundreds of unique options (e.g., manual vs. automatic transmission, 6-speed vs. 8-speed automatic transmission). In 2008, for instance, Ford Motor Company reported that some of their vehicles had over one billion possible option combinations that could be installed during production and assembly process (autonews.com, 8/19/2008). We examine these factory-installed options because this metric of product variety is most commonly used in practice and widely studied in operations management literature (MacDuffie et al., 1996; Fisher and Ittner, 1999).

A second reason for selecting the auto industry as the setting for our study is the fact that, despite a profusion of end-product configurations, the production process, especially the assembly line, has not changed significantly during the study period and is relatively similar across auto manufacturers. This reduces spurious sources of heterogeneity in our data. The auto industry is highly-regulated and regularly inspected, and safety rules and regulations are applied consistently and systematically, resulting in greater conformance in broad operating policies and procedures across manufacturers. This makes both the plant-level and recall data more comparable across different manufacturers.

Finally, the auto industry has a broad economic and psychological footprint, making it one of the most important industries in the U.S. Because of its potential to impact many other industry sectors, it has been studied extensively both by academics and by popular media (Goyal et al., 2006; Olivares and Cachon, 2009; Cachon and Olivares, 2010; Cachon et al., 2012; Moreno and Terwiesch, 2012, 2013). This long and deep history enables researchers to compare and validate empirical results and the implications of these results with greater subtlety. The unintended consequence of such close scrutiny is that secondary data are readily available.

## **2.2. Product Recalls**

Recalls in the auto industry occur when a product fails to meet Federal Motor Vehicle Safety Standards or when there is a safety-related defect in the vehicle or equipment (<http://www-odi.nhtsa.dot.gov/recalls/recallprocess.cfm>). In the U.S., the NHTSA, a regulatory agency created in 1966, is responsible for enforcing federal motor vehicle safety standards and ensuring driver safety. When an automotive model does not comply with safety standards, NHTSA might mandate that it be recalled from the market and fixed of its defects by the automaker at no additional cost to the consumer. Alternately, an automaker may voluntarily recall its models from the market, a trend that has seen a significant increase in recent years.

Recalls occur for many reasons, but two predominate: a correctly designed part was not installed properly on the assembly line or an incorrectly designed part that was installed properly but the defect went undetected. We refer to these as *manufacturing*- and *design*-related recalls respectively. A faulty weld leading to a recall is an example of a manufacturing-related recall. For example, in 2004, the Ford Taurus was recalled because certain vehicles equipped with power adjustable seats may have been inadequately welded ([www.nhtsa.com](http://www.nhtsa.com)). In this example, the seats themselves were correctly designed, but the manufacturing task of installing them to the chassis on the assembly line was performed improperly. In contrast, Honda recalled its 2004 Acura TL because it found that certain operating conditions resulted in heat build-up between the countershaft and secondary shaft second gears in the automatic transmission, and that heat could eventually lead to gear breakage. In this design-related recall, engineers had failed to provide an appropriate cooling mechanism or adequate insulation—no amount of manufacturing prowess could have prevented the recall.

Although most existing research does not differentiate between manufacturing- and design-related recalls, we make this distinction because the operational conditions that lead to these two types of recalls differ significantly. We therefore investigate them concurrently. We also want to underscore the difference between recalls and product quality, concepts that are closely related but not equivalent. Recalls are a consequence of a quality system failure. More specifically, product recalls occur because a defect was not detected until the product was in use, making them akin to an external failure in the quality literature (Fine, 1986). Product quality, on the other hand, can signify objective or subjective attributes of a product that can be measured internally or externally to the factory walls. This distinction is salient, as it helped guide our literature review.

A review of the relevant literature revealed interesting commonalities among existing studies. We present a summary of the most relevant studies in Table 1 and highlight a few common themes here. We found that the primary research focus has been examining the consequences of product recall on firm performance (Cheah et al., 2007; Jarrell and Peltzman, 1985). In a majority of studies, recall is modeled as an independent, exogenous event. Using longitudinal secondary data and event study analysis, researchers have shown that recalls have a significant negative impact on firm performance as measured primarily through stock price (Chen et al., 2009; Jarrell and Peltzman, 1985). The negative relationship is consistent across industry sectors and the measures of firm performance used, and it is robust to various empirical specifications (Davidson and Worrell, 1992; Cheah et al., 2007).

For instance, using longitudinal secondary data from the automotive and pharma industries, Jarrell and Peltzman (1985) find that shareholders of recalling firms bear substantial losses compared to the actual cost of recalls incurred by the recalling firm. Cheah et al. (2007) and Chen et al. (2009) also identify negative stock price movements following recall announcements within the consumer product

and pharmaceutical industries, and Davidson and Worrell (1992) perform a similar analysis across a 20-year sample (1968 to 1987) for all publicly traded, non-automotive U.S. companies and find negative stock value changes following government-initiated recalls.

Bromiley and Marcus (1989) and Thirumalai and Sinha (2011), however, do *not* find support for a negative stock market reaction following a recall announcement in the automotive or medical device industries. Bromiley and Marcus (1989) contend that, in the previous automotive recall research (e.g., Jarrell and Peltzman, 1985), the time window following the recall announcement was too short; stock prices generally rebounded after an initial decline, but the short window of study concealed that effect. Further, Thirumalai and Sinha (2011) conjecture that capital markets already account for the risky nature of medical devices in stock prices, so they do not drive down prices when recalls occur. Their results, then, can be attributed to the specific context of their study: the medical device industry.

In addition to changes in stock price, some research studies have investigated recalls' impact on less tangible measures, such as customer loyalty. Archer and Wesolowsky (1996) investigate customer loyalty to the manufacturer and dealer following automotive recalls and determine that customers view manufacturers' recalls for known safety defects as a heightened concern for consumer safety, so there is, surprisingly, an increase in post-recall consumer loyalty. In contrast, Hartman (1987) observes that the broader secondary market may not share such positive sentiments, and a brand's used car market value falls once a model is linked to a recall.

Our literature review also revealed that understanding related to the organizational and operational causes of recalls is in its infancy, as gleaned from the scarcity of research positing product recalls as the dependent variable. Using automotive industry data spanning 33 years, Haunschild and Rhee (2004) have shown that the impact of past recalls on a firm's future recalls is contingent upon whether the first recall was voluntary or mandated. The authors showed that the incidence of subsequent recalls reduces following a voluntary first recall, indicating that firms learn more broadly when they recall voluntarily. In the medical device industry, Thirumalai and Sinha (2011) find that, as firms expand their product scope, the likelihood of a recall increases. And we found one study that used organizational variables such as recall strategy (proactive or reactive), type of defect (manufacturing or design), and the firm's location within the supply chain (proximity to customer) in the toy industry (Hora et al., 2011). However, the dependent variable in the study was recall timeliness, measured as time from initial market release to recall announcement, instead of the recall *occurrence*. While the results from Hora et al. (2011) provide interesting insights, the authors did not focus on operational characteristics associated with product recall. Our study seeks to rectify the gap by empirically identifying plant-level causes of manufacturing- and design-related recalls in the automotive industry and therefore provide actionable insights to managers wishing to reduce recall instances.

### 2.3. Variety and Recalls

A relationship between product variety and recall has not been explicitly posited or examined in the previous literature, but product variety has been studied in association with many other performance measures. We can draw two conclusions: in general, product variety has a negative impact on operational performance (Fisher and Ittner, 1999; Ton and Raman, 2010) and it has a positive impact on market and financial performance of the firm (Kekre and Srinivasan, 1990).

For instance, MacDuffie et al. (1996) find that a greater number of options is negatively associated with productivity as measured by hours per vehicle, and Fisher and Ittner (1999) find that rework in an automotive assembly line increases with increasing variety. Ton and Raman (2010) use retail store data to show that product variety leads to more “phantom products” on store shelves, a signal of poor internal quality similar to rework on a manufacturing line. One exception is Kekre and Srinivasan (1990), who determine that variety leads to higher market share and higher profitability (implying lower costs) across a broad set of industries. The nature of these relationships makes sense. Increasing product variety exacerbates complexity on the production line and increases potential for error, consequently resulting in lower labor productivity and first pass yield. At the same time, it provides more options for customers, resulting in increased sales revenue and profit.

In automotive industry, variety is frequently conceptualized on the assembly line in terms of factory-installed options. It is measured as *option content* variety. MacDuffie et al. (1996) showed that option content variety has an adverse effect on labor hours per car—it takes longer to make each car when there are more options as to the finished product. Using variability in option content as a predictor, Fisher and Ittner (1999) also observe a positive relationship with vehicle rework. Increasing option content variety creates variation and possible confusion for both assembly line employees and inspection personnel, increasing the likelihood of error and defect identification. Factory-installed options have an impact within and across production steps: each option not only drives incremental processing from the responsible assembly line employee, but also impacts that employee’s self-inspection process and subsequent assembly and inspection steps downstream. As options are added incrementally, formalized work processes need to be modified to include additional steps, simultaneously reducing the available self-inspection time available to that assembler and increasing the need for such inspection. As the vehicle progresses down the assembly line, subsequent assembly steps are affected.

Final inspectors responsible for verifying vehicle quality upon assembly completion are required to test and inspect additional aspects of the vehicle. For instance, when a vehicle incorporates side air bags, mechanical and electrical modifications are required for proper door assembly. Such changes may impact electrical wiring required to install a subsequent option, such as power door locks. Equivalently, when anti-lock brakes are installed, it may modify installation steps for traction control or limited slip

differential. When such a vehicle is inspected by either the assembler or a final inspector, the inclusion of each additional option drives up the time, attention, and accuracy required for a thorough inspection. As options are added, the assembly and inspection processes become more time consuming, complex, and error-prone. In sum, we expect that as the number of factory-installed options increase, so does the risk of a manufacturing related recall.

*Hypothesis 1a: Increase in factory-installed options is associated with an increase in manufacturing-related recalls.*

Variety has frequently been associated with manufacturing performance problems in the literature, though research linking it with design performance is limited. Most of this research shows only that factory-installed options often exist in the form of or contain several shared components. For example, the components required for an anti-lock braking system or the electronics that enable power door locks are not designed uniquely for each model, but are shared across many models. Component sharing is shown to negatively impact design performance. Using brake rotor sharing data in the automotive industry, Ramdas and Randall (2008) find that component failures in the marketplace decrease when a component is designed specifically for a model, but failures may increase as the component is shared across an increasing number of models. The authors conclude that increased component sharing leads to a higher risk of *lack of fit* between a component and the final product. As this design fit decreases, failure risks increase. Consequently, because manufacturers do not redesign each factory-installed option for every model, some level of lack of fit risk must exist within each factory-installed option. When this risk develops into an actual design failure, manufacturing personnel and equipment may perform with perfection, but it is all in vain: the product is bound to fail. Each installed option should increase the design-related failure risk of the final product. We therefore hypothesize:

*Hypothesis 1b: Increase in factory-installed options is associated with an increase in design-related recalls.*

#### **2.4. Utilization and Recalls**

Maximizing manufacturing equipment and personnel utilization is an important objective of managers and has been associated in the literature with many benefits, including deeper process understanding, manufacturing responsiveness, higher productivity, and lower costs (Lieberman and Demeester, 1999; Womack and Jones, 2003; Schonberger, 2008). However, highly utilized manufacturing lines can also lead to overstress and potential failure (Adler et al., 1997).

Literature shows that, as assembly line utilization increases, higher cognitive and physical burdens are placed upon the assembler and can have undesired ramifications. For example, Adler et al. (1997) demonstrate the effects of high utilization rates on manufacturing and inspection processes. In a

joint Toyota-GM plant (NUMMI) case study on a 1993 product launch, the authors describe how production personnel often force-fitted poorly designed connections between components to keep up with high capacity utilization mandates. March (1981) highlights the concern that as excess capacity decreases, process improvement is mired: assemblers and support staff have less free time to observe problems, ponder solutions, and implement fixes. De Treville and Antonakis (2006) propose that if *leanness* becomes excessive, manufacturing personnel may stop following standard operating procedures and begin taking shortcuts. Those can impair quality. These authors acknowledge that “One of the more controversial lean production practices attempts to increase utilization by providing employees with less time than needed to accomplish a given task (e.g., by increasing the speed of the production line without adding more workers)” (De Treville and Antonakis, 2006). If an assembler or inspector is truly given less time than needed to accomplish their assigned task, accuracy of both the assembly and the inspection is unquestionably at risk.

These studies demonstrate that, as utilization rates rise, assembly personnel may take *shortcuts* or *force-fit* connections; actions that would likely bring about product failures. Moreover, we expect the self-inspection and final inspection steps on the assembly line to be equally impacted; the ability to vigilantly detect errors may be thwarted with declining excess capacity. Less time available to inspect one’s own work may force an assembler to either shortcut their self-inspection or simply miss manufacturing flaws as they struggle to keep up with the line pace. The same should be true for a final inspector.

We expect that as utilization rates increase on an assembly line, manufacturing defect creation will increase and defect detection will decrease. From a design-related recall standpoint, we do not hypothesize such a relationship, because design recalls should be unaffected by the manufacturing environment. We therefore hypothesize:

*Hypothesis 2a(b): Increase in utilization is positively (not) associated with manufacturing- (design-) related recalls.*

## **2.5. Focus and recalls**

Focus has been studied extensively in management theory. It is achieved by narrowing the range of demands placed on an individual, a process, and a facility, thus eliminating trade-offs inherent in systems facing competing demands (Skinner, 1974). The benefits accrue because the elements of the operating system can be better aligned to a narrowed set of demands. A statistically significant positive link between focus and superior performance is well established in diverse industry settings. Mukherjee et al. (2000) find that both labor productivity and conformance quality suffer as a manufacturing line defocuses its operations and takes on more heterogeneous products. Using broad secondary data from U.S. plants between 1972 and 1982, Brush and Karnani (1996) found a positive association between

focus and productivity. Clark and Huckman (2012) looked to healthcare to determine that focused operations in a hospital are associated with lower in-hospital mortality, a frequently used measure of quality. KC and Terwiesch (2011), and McDermott and Stock (2011) also provide evidence that focus leads to improved quality in healthcare settings, and a similar relationship is seen in the airline industry (Lapr  and Tsikriktsis, 2006; Tsikriktsis, 2007).

However, focus is also associated with reducing a firm’s ability to produce multiple end-product configurations economically and effectively. Focus is, thus, an expensive strategy for firms in capital-intensive industries (such as automotives) in which managers feel immense pressure to produce multiple models on the same assembly line to better leverage enormous fixed costs. In choosing between a focused or flexible strategy, managers need to carefully consider the benefits of focus against reduced product-mix flexibility.

In a plant that produces only one model, all personnel attention (from the plant manager to each assembly line employee) is aligned on a common set of goals. In a one-model plant, all the attention is paid to the common set of errors that plague the vehicle in the assembly process, and this attention and focus will likely lead to faster reaction to problems on the production line, higher quality assembly processes, and a finely calibrated awareness of defects for both self and final inspection. In a multi-model plant, this alignment is diminished. Different actors must be attentive to quality, yield, and productivity goals for the plant, reducing specific attention to each individual model’s performance.

While focus has been associated with higher productivity, lower costs, and improved quality, its impact on product recalls has not been studied. Focus in automotive manufacturing sharpens the ability of employees to assemble and inspect, and the repetitive nature of such environments should allow manufacturing in accordance with familiar, well-understood specifications. Focus should increase employees’ ability to identify defects quickly and accurately, reducing the likelihood of manufacturing defects occurring, let alone “escaping” into the marketplace. Design defects, by definition, should be unaffected by focus as their prevalence for failure should not be altered by any operational characteristics. The design risk of failure remains even when a product is manufactured to perfection. We would expect that a focused environment would have no impact upon design-related defects that could result in a recall. Therefore, we hypothesize:

*Hypothesis 3a(b): Focused manufacturing is negatively (not) associated with manufacturing- (design-) related recalls.*

## **2.6. Variety, Utilization, and Recalls**

As more factory-installed options are included on a vehicle, variability increases on the assembly line. According to Hopp and Spearman (2008), a central tenet of *Factory Physics* is that variability in a

production system should be buffered by more inventory, more capacity, or more time to perform the task. However, manufacturers who increase utilization levels while attempting to provide greater variety without instituting another appropriate buffering mechanism seem to violate that rule. The same authors note that “In realistic lines containing variability; pushing utilization close to 1 will seriously degrade other measures” (Hopp and Spearman, 2008). We believe that recalls are one such measure impacted by the combination of high utilization and high variety. As more variants of a model progress down the manufacturing line, personnel and equipment require a minimum level of excess capacity to changeover assembly and inspection systems that are both mental and mechanical. As utilization rises, the time available to perform the actual operation, a self-inspection or final inspection and to pause between steps decreases. All this when time is *precisely* what is needed to accurately install options and inspect a highly customized vehicle. As variety increases, a proportional amount of capacity may be required to compensate. When this does not occur, a manufacturing line mismatch may be detrimental to the final product. As manufacturers strive to increase variety and increase utilization, we expect that manufacturing defect creation should rise and defect detection should fall. This relationship however should not be present for design-related recalls. We therefore hypothesize:

*Hypothesis 4a(b): Increase in utilization worsens (does not impact) the positive relationship between variety and manufacturing- (design-) related recalls.*

## **2.7. Variety, Focus, and Recalls**

The harmonization of goals, objectives, and tactics that are the consequence of a focused plant creates an environment that is quick to identify problems, encounters fewer unexpected defects, and benefits from highly repeatable processes. This setting may abate the complexity and confusion of multiple factory-installed options. A key driver of focusing a factory, the simplification of complexities and work processes (White et al., 1999), should have a dampening effect upon sources of complexities that create disturbances within manufacturing. We expect, then, that equipment and personnel can absorb high levels of factory-installed options more effectively in a focused factory and a focused environment should reduce the number of recalls that may occur due to these options. As we have defined design-related recalls, the inherent nature of such problems is not likely to be impacted by a narrowing of focus in manufacturing.

*Hypothesis 5a(b): Focused manufacturing dampens (has no impact on) the positive relationship between variety and manufacturing- (design-) related recalls.*

### **3.0. Data and Empirical Strategy**

#### **3.1 Data Collection**

We investigate recall occurrences of all automobiles made in North America from 2000-2006 using three sources: The Harbour Report, Ward's Automotive ([www.wardsauto.com](http://www.wardsauto.com)), and the NHTSA. The Harbour Report provides annualized plant-level measures on North American auto plants, with data provided voluntarily by all major manufacturers. From Harbour, we capture utilization for each model-year and the number of models produced in each plant. For variety and volume data, we utilize Ward's Automotive. Both Harbour and Ward's have been used extensively in past automotive related research (Gopal et al., 2013; Cachon and Olivares, 2010; Fisher and Ittner, 1999). The number of recalls for each model-year is obtained from the NHTSA website.

We chose the model-year as our unit of analysis for two reasons. First, recalls occur at the model-year, not at the plant- or firm-level. Additionally, while other recall-related research was conducted at the firm- or business unit-level (Haunschild and Rhee, 2004), our objective to identify operational causes of recalls requires the finest unit of analysis possible. For instance, most automakers (e.g., Chevrolet or Dodge) build cars in multiple plants; hence we must study our questions *below* the automaker-level. There are also multiple models produced at certain plants, and they do not all experience a simultaneous recall; again, the plant is not an appropriate unit of analysis for our work.

We can classify automotive plants in our data into three categories based on a Model\_Plant description (Table 2) which will be necessary for calculating plant utilization. More specifically, we use the number of plants in which a model is built and the number of other models built within those same plants to develop our classification scheme. For example, in 2002, the Chevrolet Corvette was only built in one plant (Bowling Green, KY), and it was the only model built in that plant that year. Thus, the 2002 Corvette is assigned the Singleplantmodel\_Singlemodelplant category. Conversely, in 2006, the Ford GT, Lincoln LS, and Lincoln Town Car were all produced in the Wixom, MI plant. These three models were not produced in any other plant, so they are designated as Singleplantmodel\_Multiplemodelplant. Finally, the 2004 Ford Focus was built in the Wixom, MI plant and in the Hermosillo, MX plant, in which other cars were also built. The Focus earns the designation Multipleplantmodel\_Multiplemodelplant. Our sample spans seven years (2000-2006), and consists of eight manufacturers (General Motors, Ford, Chrysler, Toyota, Honda, Nissan, Mazda, and Mitsubishi), 32 production plants, and 80 unique models resulting in a sample size of 232 car model-years.

## 3.2 Variables

### 3.2.1. Dependent Variable: Recalls

We operationalize recalls separately as *Manufacturing* and *Design recalls*, and measure each as the total number of recalls that occurred in a given model-year. The data was obtained from NHTSA's website in December 2012. In addition to recall date, the reported data includes vehicle make, recall type, model-year, production start and end date for recalled vehicles, approximate number of units affected, and defect description. Thus, we capture all recalls that occurred until the end of 2012 for all models produced between 2000 and 2006. This allows for a reasonable amount of time (between 0 and 12 years) for defects and recalls to surface, reasonable because the average time between production and recall for the top 10 auto recalls in history was a little over 5 years (<http://editorial.autos.msn.com/10-largest-auto-recalls-in-history>).

We made the following exclusions within our sample: we removed duplicate recalls reported on identical defects within the same model-year. As our goal is to understand plant-level drivers of recalls, we also exclude recalls the NHTSA has attributed to either post-market replacement parts used by dealers or due to supplier manufactured components (excluding 242 recalls). To classify recalls into manufacturing or design categories, we comprehensively reviewed each recall description using the following rules: If the recall description mentioned that the defect was caused by manufacturing, installation, heat treatment, soldering, torqueing, tightening, forming, mis-location, damage during manufacturing, inadequate sealing, or mis-manufacturing, it was categorized as a manufacturing-related recall. If the recall description stated that the defect was caused by product degradation, inadequate design, malfunction, or other non-manufacturing related criteria, it was categorized as a design-related recall. Our final sample consists of 182 manufacturing- and 285 design-related recalls (467 total).

### 3.2.2. Independent Variables

**Utilization.** We operationalize *Utilization* as assembly line utilization at the plant in which the model-year was manufactured. We obtained assembly line utilization from Harbour, which reports it as the number of units produced as a ratio of maximum capacity at the plant. Manufacturers can report utilization greater than 100% because maximum capacity does not include overtime. Our main analysis includes only those models that are built in one plant (the first two rows of Table 2, totaling 80 models), because this plant-level utilization estimate is a close proxy for the utilization at the model level. For the 16 models built in multiple plants (row three in Table 2), we computed average utilization across all of the plants in which the model is made. This utilization measure is fairly noisy; nevertheless, we used it for robustness check and we report the results separately.

**Variety.** We measure *Variety* by computing total number of factory-installed options on a given model-year. These options include items such as automatic air conditioning, electronic suspension, side air bags, traction control, and anti-lock brakes. Thirty possible options were available for factory installation on the models in our panel when including the number of options that were utilized at least once within that model-year. This data is obtained from Ward's Automotive.

**Focus.** We measure *Focus* with an indicator variable to signify if the model is only built in one plant, and that there are no other models built in that plant. Per Table 2, all model-years in the *Singleplantmodel\_Singlemodelplant* category are assigned a 1 for this variable, and all other model-years are assigned a 0. We used Harbour to create this variable.

### 3.2.3. Control Variables

**Volume.** We take the natural logarithm of the volume (*Ln volume*) for each model-year to control for the effect of volume in the occurrence of a recall. This data is obtained from Ward's Automotive.

**Employees.** Larger automotive plants are typically associated with higher level of labor and capital resources, which may reduce the likelihood of a recall. To control for this possible association, we measure the natural log of the total number of direct and indirect employees at the plant (*Ln employees*) using data obtained from Harbour Reports.

**Lagged recalls.** We create *Lagged manufacturing recalls* and *Lagged design recalls* as one method of correcting for potential omitted variable bias. It is possible that certain auto models have inherently high recall risks that are not captured in our dataset. By controlling for past recalls on each model in the previous model year, we mitigate some potential bias in our analysis. We summed the previous year's manufacturing or design recalls and used this as a control variable. This data was obtained from NHTSA.

**Time since production.** The more time that has elapsed following production of a vehicle, the more opportunities there are for manufacturing and design defects to surface and lead to recalls. To control for this effect, we create a *Time since production* variable based on the time difference between 2012 and the model-year of the vehicle. The average time since production in our dataset is 9.06 years.

**Manufacturer.** To control for the effect of ownership on recall, we classify manufacturers into three categories based on geography: U.S. manufacturers, Japanese manufacturers, and Toyota. Following other researchers, we kept Toyota in its own separate category because of their reputation for high quality and distinction for lean production (Lieberman and Demeester, 1999). The manufacturer of each model-year was obtained from Ward's Automotive.

### 3.3. Empirical Approach

We estimate the following linear model:

$$R_{it} = \alpha C_{it} + \beta V_{it} + \lambda U_{it} + \gamma F_{it} + \rho(V * U)_{it} + \delta(V * F)_{it} + \mu_{it} \quad (1)$$

where we use  $R_{it}$  to denote the number of recalls that occur for model  $i$  in year  $t$ ,  $C_{it}$  is the vector of control variables,  $V_{it}$  is the number of options installed in a model,  $U_{it}$  is the utilization level at the plant in which the model was built,  $F_{it}$  is an indicator variable for focus in the plant,  $(V * U)_{it}$  is the interaction between the number of options and the utilization, and  $(V * F)_{it}$  is the interaction between the number of options and the focus indicator variable.

We use generalized estimated equations (GEE) to examine the variation in recalls per model-year across a 7-year unbalanced panel. GEE models estimate the marginal effect of covariates averaged across units and can be interpreted in our case as the overall effect of utilization, variety, and focus on the number of recalls *across* the population (e.g., population-averaged model) (Rhee et al., 2006, Sine et al., 2003, Liang and Zeger, 1986). In contrast, fixed and random effects models analyze the impact of an independent variable on the dependent variable *within* each model-year. We use a GEE model for three reasons. First, it allows for correction for autocorrelation, specification of the distribution of the dependent variable, as well as the use of robust standard errors in an unbalanced panel. Because recalls are count data, a Gaussian distribution is inappropriate. A Poisson regression is correct only when the mean and variance of the dependent variable are equal. Our data demonstrates overdispersion, indicating the need for a negative binomial model. We use a negative binomial distribution with a log linear link function and an exchangeable working correlation structure, which models a shared correlation between observations within a group. Robust standard error can be computed correctly and is valid even when the correlation structure is mis-specified in a GEE model (Rhee et al., 2006; Greene, 2012).

We do not use a fixed effects model because it is unable to estimate time invariant effects. We have multiple models that did not experience any recalls during the 7 years; eliminating these observations would likely bias our results (Sine et al., 2003). Moreover, Hausman test indicates that a random effects model is preferred to a fixed effects model. Therefore, we check for robustness later with a random effect negative binomial model.

### 4.0. Results

The correlation matrix of the variables is presented in Table 4. We observe a strong positive correlation between utilization and both manufacturing and design recalls. There is also a strong negative relationship between focus and manufacturing recalls. We tested for multi-collinearity in our data and

found Variance Inflation Factors (VIF) were below 2.2. VIF above 10 is frequently viewed as problematic to standard error calculations (Kutner et al., 2005), so we report no evidence of multi-collinearity.

The results from the negative binomial GEE analysis are reported in Table 5. We conducted separate analyses for manufacturing (columns 1 through 5) and design (columns 6 through 10) recalls. For each of our dependent variables, we first include all our control variables (columns 1 and 6). To test for Hypotheses 1, 2, and 3, we include variety, utilization, and focus and examine the significance of their effect on each type of recall; the results are presented in columns 2 and 7. Finally, we include the two interaction terms in separate steps for the two dependent variables (columns 3, 4 and 8, 9). We also present a full model with control variables, main effects, and interaction effects for comparison purposes (columns 5 and 10).

Beginning with manufacturing-related recalls, we find that time since production is positively and significantly related to recalls, suggesting that the longer a model has been on the market, the higher its likelihood of recall. Incorporating the main effects of variety, utilization, and focus reveal interesting relationships. While variety has no significant effect on manufacturing recalls (no support for Hypothesis 1a), both utilization and focus are significant, supporting Hypotheses 2a and 3a, respectively. Because the relationships are modeled using a negative binomial log link function using standardized predictors, the effects need to be interpreted multiplicatively. More specifically, a one standard deviation change in an independent variable increases the number of recalls by a factor of  $exp^{\beta}$  (Gardner et al., 1995). Thus, one standard deviation increase in plant utilization (25.44 percent) results in a 68% ( $exp^{0.52} = 1.68$ ) increase in manufacturing recalls. Because focus is a categorical variable, moving from a non-focused (focus=0) to a focused plant (focus=1) results in a 74% ( $exp^{-1.36} = 0.26$ ) decrease in the number of manufacturing recalls.

We also find that both interaction terms are significant at  $p < .05$  levels. *Variety\*Utilization* coefficient ( $\beta=0.22$ ) is positive, providing support to Hypothesis 4a. This implies that an increase in utilization has an aggravating effect on the relationship between variety and manufacturing recalls. The interaction plot (Figure 1) confirms that manufacturing recalls are significantly higher for high utilization plants (dashed line with squares) both for low and high variety conditions as compared to the low utilization plants (solid line with diamonds) and that moving to high variety increases recalls *only* in a high utilization plant. We further examined this relationship with a conditional effects plot, frequently used to determine the changing relationship of one variable with another when it is dependent upon a third variable (Kutner et al., 2005). Figure 2 demonstrates that, below the 25<sup>th</sup> percentile of utilization, there is almost no relationship between variety and manufacturing recalls. However, the relationship clearly strengthens above the 75<sup>th</sup> percentile.

Finally, the coefficient for *Variety\*Focus* is negative and demonstrates that focus ameliorates the negative impact of variety on manufacturing recalls, supporting Hypothesis 5a. The interaction plots among variety, focus, and manufacturing- and design-recalls (Figure 3) show that manufacturing recalls are considerably lower for high focus plants (dashed line with squares) with both low and high variety conditions when compared to the low focus plants (solid line with diamonds). These results imply that variety leads to recalls when assembly line utilization is high or plant focus is low.

Moving to design-related recalls, we find that controls for volume, time since production, and lagged design recalls are significant. As expected, higher volume models are more likely to experience a design-related recall. Time since production is positive and significantly related to design recalls; more time on the market increases the likelihood of a design failure. Lagged recalls are also a positive predictor of current model-year design recalls. This is intuitive, as manufacturers do not always alter designs with each model-year, and a previous year's design-related recall may indicate a higher likelihood of design recalls in the current year.

In evaluating our main effects, we find that variety is a significant predictor of design recalls in the hypothesized direction, providing support for Hypothesis 1b. A one standard deviation increase in the number of options (3.4 options per model-year) will increase the number of design recalls by 30% ( $exp^{0.26} = 1.30$ ). Hypothesis 2b is rejected, as we observe that design recalls unexpectedly increase 49% ( $exp^{0.40} = 1.49$ ) with a one standard deviation increase in plant utilization. This is counter to our expectation that utilization should not affect design recalls. Finally, we see that focus plays no part in design-related recalls, supporting Hypothesis 3b.

As predicted, none of our interactions affect design recalls. An inherently faulty design should not be impacted by the context in which the model is assembled, supporting Hypotheses 4b and 5b.

#### **4.1. Robustness Checks**

We conducted several robustness checks in our analysis (Table 6). For comparative ease, we present final results from Table 5 (columns 5 and 10) in columns 1 and 2 of Table 6.

Prior automotive recall research has shown that recalls frequently affect different models manufactured by the same automaker (Haunschild and Rhee, 2004). To address the shared vs. unique recalls concern, we divided manufacturing and design recalls into those that were unique to an individual model and those that were shared across models. We observe that even with a much reduced sample size, the results for unique manufacturing recalls are consistent for the variables of interest (column 3). In examining the relationships for shared recalls, we find that the main effects are consistent with column 1, but the interactions are no longer significant (column 4). Columns 5 and 6 display the results of unique and shared design recalls. While variety is significantly and positively related to shared recalls, it is not

significant for unique recalls. This implies that the relationship between variety and design recall in our main model (column 2) may be driven solely by shared design recalls. In other words, variety is only a predictor of design related recalls when we consider those recalls that are shared across multiple models.

Columns 7 and 8 incorporate the `Multipleplantmodel_Multiplemodelplant` category of model-years. This category was excluded from our main analysis because our measure for utilization required us to compute average utilization across multiple plants. Even though the computed average utilization measure is less precise, our results show that all substantive relationships are the same in sign and significance. Finally, results using a negative binomial random effects regression specification, similar to ordinary least square regression, are consistent with GEE findings (columns 9 and 10).

Previous automotive research has shown that launching a new vehicle impacts plant productivity (Gopal et al., 2013) and may impact other plant performance measures (like utilization, which impacts recalls). Using Harbour data, we developed a launch indicator variable which equals 1 if a launch occurred at the plant during the model-year of production and 0 otherwise. When included in our models, this variable had no significant correlation with any of our variables and did not alter any of our findings. It is possible that a plant is chosen for a launch because of its utilization levels, which would create a selection bias and possibly impact recalls for the newly launched vehicle. Gopal et al. (2013), using a dataset that overlaps our panel years, also did not find any significant difference in utilization between launch and non-launch plants, supporting our findings. Finally, we ran a panel probit model to determine which of our variables might predict the launch indicator variable. None of our substantive variables of interest predict the launch indicator variable with significance below 0.2, signifying a reduced chance of selection bias in relation to new launches within our plants.

Finally, we examined the relationship between utilization and manufacturing recalls to identify utilization ranges in which the relationship differs in a significant manner. To do so, we parsed utilization into quartiles and reanalyzed the data using GEE specification. We report the utilization quartile ranges in Table 7 and GEE results in Table 8. The results show that the dummy for utilization level in the 3<sup>rd</sup> and 4<sup>th</sup> quartile have a positive and significant impact on manufacturing recalls at  $p < 0.10$  and  $p < 0.01$ , respectively. Utilization at the 1<sup>st</sup> and 2<sup>nd</sup> quartile are not significant predictors of manufacturing recalls. These findings imply that utilization's impact upon recalls is primarily driven by very high levels of utilization representing the 4<sup>th</sup> quartile. Specifically, utilization rates over 98% are particularly harmful for a plant's recall performance.

Endogeneity (the presence of correlation between independent variables and the error term) creates bias in model coefficients under three circumstances: omitted variables, simultaneity, and measurement error (Wooldridge, 2010). Possible omitted variables in our study may be model-specific attributes that impact the chances of a recall, such as the experience of the design engineers or plant

workforce. We attempt to capture some of this bias by including a lagged recall control variable in our analyses, but clearly there are other possible variables we have been unable to capture in our dataset. To determine the likelihood of omitted variable bias, we use a Hausman test, which compares the fixed and random effects coefficients in a panel model. A statistically insignificant difference between these two coefficients indicates that the assumption of orthogonality between the independent variables and any unobserved time-constant effects holds. With an insignificant Hausman test ( $p$  value = 0.42) we can conclude there is only a low risk of omitted variable bias.

Simultaneity exists when there is a causal loop between the dependent variable and at least one independent variable. The structure of our data minimizes the chances of simultaneity. Specifically, a recall on a certain model-year must occur after the manufacture of the model (when the variety and utilization measures are recorded); hence within a certain model-year, recalls cannot cause our independent variables. Finally, since measures for utilization, variety and focus were collected through well-established sources with transparent and verified methods, it makes measurement error unlikely.

## **4.2. Discussion**

First and foremost, our empirical results indicate that high variety results in a significant increase in design recalls and contrary to conventional wisdom, it does not impact manufacturing recalls. The results are robust to model and measure specifications, though with slightly different effect sizes. The differential impact of variety on manufacturing and design recalls is noteworthy for several reasons. First, our findings corroborate previous research showing that failure risk of a component increases as design specificity decreases (Ramdas and Randall, 2008). Perhaps each factory-installed option contributes significant inherent design-related failure risk to the vehicle, and this potential lack of fit between the option and the vehicle presents a more serious risk than the effort of actually installing the option. Our robustness checks confirm that this variety-recall relationship is primarily driven by shared design recalls (those recalls that cross multiple models and have a shared root cause). Design engineers should consider this risk in comparison to the costs of designing a unique option for a given vehicle.

Second, while previous researchers have consistently shown a negative relationship between variety and manufacturing performance, our findings suggest that categorizing performance into finer grained levels may help us understand the relationship more accurately and enable managers to manage operations more effectively. At a more practical level, our results show that manufacturing personnel's resistance to increased variety in assembly operations might be needless.

Our results also provide support for our hypotheses linking utilization and recalls. Although we had argued that increasing utilization should result in increasing manufacturing recalls but not design recalls, our results show that utilization is positively associated with both manufacturing and design

recalls. This implies that an increase in utilization is universally bad for product recalls. One possible explanation is that a highly utilized line more easily exposes design risks that would lay dormant under standard utilization rates.

In quantitative terms, our data show that each 5% increase in utilization leads to an average increase of about 13.6% in the incidence of manufacturing- and 9.8% in design- recalls respectively. Similarly, installing three additional options, an increase of roughly one standard deviation, results in an increase of 30% in the incidence of design recalls. Any increase in recalls will result in a significant increase in total firm costs (Jarell and Peltzman, 1985).

Further, while variety does not directly affect manufacturing recalls, high variety and high utilization together result in significantly higher levels of manufacturing-related product recalls. The underlying logic is that variety creates a need for excess capacity, yet manufacturers ignore this relationship because they are incentivized to increase both variety and utilization. When this occurs, assemblers and inspectors are unable to absorb the cognitive burdens. Not only do they create manufacturing defects, but they also overlook those defects. When choosing capacity requirements on an assembly line, managers need to consider not only demand, but also variety.

Our study contributes to the literature documenting benefits of focus and adds another important benefit to the list: reducing recalls. The relationship between focus and manufacturing related recalls is stronger than we expected. In fact, a focused environment can greatly reduce manufacturing related recalls, with a 74% reduction in such problems. Focus also mitigates the relationship between variety and manufacturing recalls. When a factory focuses on a single model, personnel are able to assemble and inspect high variety reliably. An unfocused factory exposes and compounds the complexity of variety, while a focused factory diminishes it. This positive aspect of focus in reducing recalls must be, of course, traded off with positive aspects of flexibility with respect to product launch (Gopal et al., 2013) and pricing (Moreno and Terwiesch, 2012).

Managers will also be interested in our findings regarding the relationships between variety and utilization, and variety and focus. Because utilization rates can be varied effectively in the short term, managers can use capacity as a tactical lever to absorb complexity created by increased variety on the line. Alternately, focus, long considered a strategic tool, can be used to create a finer-tuned manufacturing assembly and inspection system that is able to operate successfully and diminish recalls in high variety situations. When unable to direct their highest variety models to focused factories, managers might create capacity cushions in multi-model plants.

The main effects of utilization and focus lend support to previously assumed relationships, but highlight aspects that have not been well understood. Running an assembly line at high utilization reduces manufacturing costs, but creates risks. As employees work faster, mistakes can occur in assembly and

quality inspection. Our research indicates a precise utilization level beyond which an automotive assembly line experiences setbacks. We can draw a number of conclusions from the results in Table 8. We see that manufacturing is capable of running reliably at very high levels of utilization (up to 98%): assemblers and inspectors seem able to function effectively in an environment of almost maximal efficiency. Overtime, frequently used to respond to fluctuations in demand without permanently increasing staffing levels, brings flexibility to managers, but may also bring more risks than previously understood (Kesavan et al., 2012). Once a manufacturing line moves above 98%, the quality systems that enable quality product to be built and inspected seemingly break down. In an overtime situation, either the plant is working their current staff extra hours or they temporarily hire an extra shift to increase capacity. Either case is fraught with challenges that are difficult for manufacturers to overcome. Personnel fatigue may prevent high quality assembly and inspection on overtime, and temporary staff may be less equipped to perform appropriate processes accurately. It seems practitioners would be better off running at comparatively lower levels of utilization by hiring full-time personnel that can surge when needed to increase volume, but still keeping the overall utilization of the plant below overtime levels. This finding is especially relevant in today's economy, which rewards manufacturers who rely on overtime with revised union contracts that pay lower overtime rates (Rogers, 2013).

While the strategic choice to focus a factory requires substantial capital investments, the previously understood benefits to managers revolved primarily around efficiency and costs. Our work uncovers a potentially more important effect: lower recall rates. Looking across our panel, we see that, of the focused models, there are four made by General Motors, three by Ford, two by Chrysler, and one by Mazda. Interestingly, Toyota, which has prioritized flexibility through lean practices, does not appear on this list. U.S. manufacturers are much more likely to focus a factory on just one model than their Japanese counterparts. This indicates a divergence in strategy that bears relevance in recall performance. Flexibility at the model level is clearly required by all manufacturers due to the large number of factory-installed options, but flexibility at the plant-level is less a customer requirement than a firm choice. This choice seems to be pertinent, and manufacturers may require plant-level focus to absorb model-level flexibility.

### **4.3. Conclusions and Limitations**

Product recalls have attracted significant attention from researchers, though little literature links operational causes to recalls. Our study brings renewed attention. Using a combination of econometric techniques and data from the North American automotive industry, we show that plant-level characteristics can significantly impact recalls, but the nature of impact depends upon the type of recall. Specifically, while variety increases design recalls, it has no effect on manufacturing recalls. Higher utilization negatively impacts both manufacturing and design recalls. And, while managers might be able

to mitigate the negative effects of higher variety on manufacturing recalls by manipulating utilization level and focus, these same operational tools do not effectively mitigate the relationship between variety and design recalls. Higher specificity of design for factory-installed options is revealed as one possible mechanism to alleviate this negative relationship.

The structure of the data creates a few limitations. Many of the variables are annualized measures; while the number of models built in a plant does not significantly change within a given model-year, variety and utilization can fluctuate more frequently. Product recalls may not necessarily be created within an annual time window, even though NHTSA reports recalls at the model-year. To address this concern, we subtracted production start and end times for each recall in our dataset and found that the average number of days of production for a recall is 444. This suggests that annualizing our measures is reasonable, though a monthly analysis would allow for an even finer-grained study. The study is also limited to North American auto plants. While such a subset of automotive manufacturers represents the industry leaders in volume, a global study would substantiate the effects observed in this current work.

Finally, we measured recalls as manufacturing- and design-related recalls, excluding two less frequently occurring categories: supplier-manufactured and dealer-installed components. Future research may examine the drivers of these recalls and whether errors in those components spill over into manufacturer recalls. Potential future work can also expand the sources of data to include additional design related recall causes beyond plant level measures. Shortcomings notwithstanding, our paper takes an important step in documenting the plant level drivers of recalls in the automotive industry.

---

## **References:**

- Adler, P. S., Goldoftas, B., & Levine, D. I. (1997). Ergonomics, employee involvement, and the Toyota production system: A case study of NUMMI's 1993 model introduction. *Industrial and Labor Relations Review*, 416-437.
- Archer, N., & Wesolowsky, G. (1996). Consumer response to service and product quality: A study of motor vehicle owners. *Journal of Operations Management*, 14(2), 103-118.
- Bromiley, P & Marcus, A. (1989). The deterrent to dubious corporate behavior: Profitability, probability and safety recalls. *Strategic Management Journal*, 10(3), 233-250.
- Brush, T., & Karnani, A. (1996). Impact of plant size and focus on productivity: An empirical study. *Management Science*, 42(7), 1065-1081.
- Cachon, G., Gallino, S., & Olivares, M. (2012). Severe weather and automobile assembly productivity. Available at SSRN 2099798.
- Cachon, G. P., & Olivares, M. (2010). Drivers of finished-goods inventory in the US automobile industry. *Management Science*, 56(1), 202-216.
- Cheah, E. T., Chan, W. L., & Chieng, C. L. L. (2007). The corporate social responsibility of pharmaceutical product recalls: An empirical examination of US and UK markets. *Journal of Business Ethics*, 76(4), 427-449.

- Chen, Y., Ganesan, S., & Liu, Y. (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), 214-226.
- Clark, J. R., & Huckman, R. S. (2012). Broadening focus: Spillovers, complementarities, and specialization in the hospital industry. *Management Science*, 58(4), 708-722.
- Davidson, W. N., & Worrell, D. L. (1992). Research notes and communications: The effect of product recall announcements on shareholder wealth. *Strategic Management Journal*, 13(6), 467-473.
- De Treville, S. D., & Antonakis, J. (2006). Could lean production job design be intrinsically motivating? contextual, configurational, and levels-of-analysis issues. *Journal of Operations Management*, 24(2), 99-123.
- Fisher, M. L., & Ittner, C. D. (1999). The impact of product variety on automobile assembly operations: Empirical evidence and simulation analysis. *Management Science*, 45(6), 771-786.
- Gardner, W., Mulvey, E., & Shaw, E. (1995). Regression analysis of counts and rates: Poisson, overdispersed poisson, and negative binomial models. *Psychological Bulletin*, 118(3), 392-404.
- Gopal, A., Goyal, M., Netessine, S., & Reindorp, M. (2013). The impact of new product introduction on plant productivity in the North American automotive industry. *Management Science*, doi:10.1287/mnsc.2013.1709.
- Goyal, M., S. Netessine & T. Randall. (2006). Deployment of Manufacturing Flexibility: An Empirical Analysis of the North American Automotive Industry. Working Paper, INSEAD.
- Hartman, R. S. (1987). Product quality and market efficiency: The effect of product recalls on resale prices and firm valuation. *The Review of Economics and Statistics*, 69(2), 367-372.
- Haunschild, P. R., & Rhee, M. (2004). The role of volition in organizational learning: The case of automotive product recalls. *Management Science*, 50(11), 1545-1560.
- Hopp, W. J., & Spearman, M. L. (2008). *Factory physics*. Vol 2. McGraw-Hill/Irwin, New York.
- Hora, M., Bapuji, H., & Roth, A. V. (2011). Safety hazard and time to recall: The role of recall strategy, product defect type, and supply chain player in the US toy industry. *Journal of Operations Management*, 29(7), 766-777.
- Jarrell, G., & Peltzman, S. (1985). The impact of product recalls on the wealth of sellers. *The Journal of Political Economy*, 93(3), 512-536.
- Kalaiganam, K., Kushwaha, T., & Eilert, M. (2013). The impact of product recalls on future product reliability and future accidents: Evidence from the automobile industry. *Journal of Marketing: A Quarterly Publication of the American Marketing Association*, 77(2), 41-57.
- Kesavan, S., Staats, B. R., & Gilland, W. (2012). Volume Flexibility in Services: The Costs and Benefits of Flexible Labor Resources. Working paper, UNC at Chapel Hill.
- Kekre, S., & Srinivasan, K. (1990). Broader product line: A necessity to achieve success? *Management Science*, 36(10), 1216-1232.
- KC, Diwas and Terwiesch, C. (2011). The Effect of Focus on Performance: Evidence from California Hospitals, *Management Science*, 57(11), 1897-1912.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* McGraw-Hill/Irwin, New York.
- Lapr e, M. A., & Tsikriktsis, N. (2006). Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Science*, 52(3), 352-366.

- Lieberman, M. B., & Demeester, L. (1999). Inventory reduction and productivity growth: Linkages in the Japanese automotive industry. *Management Science*, 45(4), 466-485.
- MacDuffie, J. P., Sethuraman, K., & Fisher, M. L. (1996). Product variety and manufacturing performance: Evidence from the international automotive assembly plant study. *Management Science*, 42(3), 350-369.
- March, J. G. (1981). Footnotes to organizational change. *Administrative Science Quarterly*, 26(4), 563-577.
- McDermott, C. M., & Stock, G. N. (2011). Focus as emphasis: Conceptual and performance implications for hospitals. *Journal of Operations Management*, 29(6), 616-626.
- Moreno, A. & C. Terwiesch. (2012). Pricing and Production Flexibility: An Empirical Analysis of the U.S. Automotive Industry. Working Paper, The Wharton School.
- Moreno, A. & C. Terwiesch. (2013). The Effects of Product Line Breadth: Evidence from the Automotive Industry. Working Paper, The Wharton School.
- Mukherjee, A., Mitchell, W., & Talbot, F. B. (2000). The impact of new manufacturing requirements on production line productivity and quality at a focused factory. *Journal of Operations Management*, 18(2), 139-168.
- Olivares, M., & Cachon, G. P. (2009). Competing retailers and inventory: An empirical investigation of General Motors' dealerships in isolated US markets. *Management Science*, 55(9), 1586-1604.
- Ramdas, K., & Randall, T. (2008). Does component sharing help or hurt reliability? An empirical study in the automotive industry. *Management Science*, 54(5), 922-938.
- Rogers, C. (2013). Open All Night: America's Car Factories. *The Wall Street Journal*, August 17-18, 2013(1).
- Scherer, F. M., & Ross, D. (1990). Industrial market structure and economic performance. 3<sup>rd</sup> ed. Boston: Houghton Mifflin.
- Schonberger, R. J. (2008). *World class manufacturing*. 3<sup>rd</sup> ed. New York: Simon & Schuster Inc.
- Shah, R., & Shin, H. (2007). Relationships among information technology, inventory, and profitability: An investigation of level invariance using sector level data. *Journal of Operations Management*, 25(4), 768-784.
- Shah, R., & Ward, P. T. (2007). Defining and developing measures of lean production. *Journal of Operations Management*, 25(4), 785-805.
- Skinner, W. (1974). *The focused factory* Harvard Business Review.
- Staats, B. R., Brunner, D. J., & Upton, D. M. (2011). Lean principles, learning, and knowledge work: Evidence from a software services provider. *Journal of Operations Management*, 29(5), 376-390.
- Thirumalai, S., & Sinha, K. K. (2011). Product recalls in the medical device industry: An empirical exploration of the sources and financial consequences. *Management Science*, 57(2), 376-392.
- Ton, Z., & Raman, A. (2010). The effect of product variety and inventory levels on retail store sales: A longitudinal study. *Production and Operations Management*, 19(5), 546-560.
- Tsikriktsis, N. (2007). The effect of operational performance and focus on profitability: A longitudinal study of the US airline industry. *Manufacturing & Service Operations Management*, 9(4), 506-517.
- Womack, J. P., & Jones, D. T. (2003). *Lean thinking: Banish waste and create wealth in your corporation*, revised and updated. 3<sup>rd</sup> ed. New York: Simon and Schuster Inc.

Table 1. Relevant Recall Literature

Article	Type of data <sup>a,b</sup> sample size	Industry sector	Research Focus	Independent variables	Recall use <sup>c</sup>	Dependent variables	Control variables	Results
Jarrell & Peltzman (1985)	S; L(1974-82); N=148	Pharma; Auto	Impact of product recalls on shareholder wealth; Compare destruction of wealth to the recall costs for the firm	Recall event	IV	Stock market returns	None	Loss to shareholder wealth is substantially greater than the costs incurred by the firms.
Bromiley & Marcus (1989)	S; L(1967-83); N=91	Auto	Stock market reaction to product recalls	Recall event	IV	Stock market returns	Presidential administration	Stock market returns do not vary significantly with product recall announcements; they are not an effective deterrent to dubious corporate behavior.
Davidson & Worrell (1992)	S; L(1968-87); N=133	Auto	Stock market reaction to product recalls	Recall event	IV	Stock market returns	None	Product recalls are negatively associated with stock market returns. The effect is stronger when products are replaced rather than repaired.
Archer & Weslowski (1996)	P; X; N=659	Auto	Impact of quality and service incidents upon customer loyalty	Negative quality and service incidents including recalls	IV	Customer loyalty	Duration of vehicle ownership	Product recalls do not impact customer loyalty to manufacturer or dealer.
Haunschild & Rhee (2004)	S; L(1966-99); N=2287	Auto	Firm learning following voluntary and involuntary recalls	No. of recalls	IV, DV	# of recalls	Volume; firm size & age; presidential administration; industry competition; year	Voluntary recalls result in more learning than mandated recalls when learning is measured as reduction in subsequent involuntary recalls.
Rhee & Haunschild (2006)	S; L(1975-99); N=1853	Auto	Impact of organizational reputation on market share following a recall	No. of recalls; organizational reputation; vehicle characteristics	IV	Market Share	Volume; age of firm; country of manufacturer	Positive organizational reputation is associated with more negative market share reactions to product recalls. Having few substitutes for a product buffers this reaction.
Cheah et al., (2007)	S; L(1998-2004); N=177	Pharma	Impact of corporate social responsibility (CSR) practices on stock market reaction to product recalls	Recall event	IV	Stock market returns	Country; CSR programs, severity of recall	Product recalls are negatively associated with stock market returns. The effect is dependent upon geography, recall severity, and CSR practices.
Chen et al., (2009)	S; L(1996-2007); N=153	Consumer product	Impact of recall strategy (proactive vs. passive) on stock market reaction to product recalls	Recall event	IV	Stock market returns	Firm size; reputation; liabilities; volume; time on market; price; product category	Proactive recalls are associated with more negative stock market reactions than passive recalls.
Thirumalai & Sinha (2011)	S; L(2002-05); variable samples	Medical device	Stock market reaction to product recalls and sources of recalls	Recall event; recall class & experience; product scope; R&D intensity	IV, DV	Stock market returns; No. of recalls	Sales; year; Units recalled	Market penalties for medical device recalls are not significant. R&D focus increases likelihood of recalls.
Hora et al., (2011)	S; L(1993-2008); N=528	Toy	Impact of recall strategy, product defect and supply chain entity on time to recall	Recall strategy; source of defect, supply chain entity	DV	Time to recall	Product price; volume; type of hazard; country of manufacturer	Preventive recall strategy, design defects & manufacturers are associated with longer time to recall.

a Primary = P; Secondary = S; b Cross-sectional data = X; Secondary longitudinal data (duration) = L(19xx-20xx); c Independent variable = IV; Dependent variable = DV; Control variable = CV

Table 2. Model\_Plant Descriptions

Category Description	Model	Plant in which the model was built	# of Models
Singleplantmodel_Multiplemodelplant	Built in one plant	Multiple models built in plant	69
Singleplantmodel_Singlemodelplant	Built in one plant	No other models built in plant	11
Multipleplantmodel_Multiplemodelplant	Built in multiple plants	Multiple models built in at least one of the shared plants	16

Table 3. Description of Variables and Summary Statistics

Variable	Description	Mean	Std. dev.
<i>Mfg Recalls</i>	Number of manufacturing related product recalls	0.83	1.14
<i>Design Recalls</i>	Number of design related product recalls	1.24	1.49
<i>Volume</i>	Number of cars sold in model-year	80,933	62,094
<i>Employees</i>	Number of employees at plant at end of year	2,409	851
<i>Lagged Recalls</i>	Number of Mfg or Design recalls in previous year		
	Mfg lagged recalls	0.86	1.02
	Design lagged recalls	1.25	1.26
<i>Time since production</i>	Number of years between model-year and the year 2012	9.06	2.01
<i>US</i>	Indicator for cars built by US manufacturers	0.86	0.35
<i>Toyota</i>	Indicator for cars built by Toyota	0.03	0.18
<i>Variety</i>	Number of factory-installed options used	20.31	3.46
<i>Utilization</i>	Plant utilization during year of manufacture	77.32	25.44
<i>Focus</i>	Indicator for plants that only build one car	0.12	0.32

Table 4. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11
1 Manufacturing recalls	1										
2 Design recalls	0.32*	1									
3 Volume	0.19*	0.18*	1								
4 Employees	0.13*	0.24*	0.31*	1							
5 Lagged mfg recalls	0.25*	0.18*	0.05	0.15*	1						
6 Lagged design recalls	0.11	0.47*	0.09	0.22*	0.30*	1					
7 Time since production	0.27*	0.25*	0.01	0.18*	0.19*	0.25*	1				
8 US	-0.11	-0.11	0.04	0.06	-0.09	-0.05	0.12	1			
9 Toyota	0.08	0.02	0.05	-0.06	0.02	-0.09	-0.15*	-0.46*	1		
10 Variety	0.03	0.09	-0.32*	0.09	0.06	0.08	-0.12	0.09	-0.04	1	
11 Utilization	0.31*	0.29*	0.49*	0.56*	0.15*	0.13	0.06	-0.02	0.22*	-0.22*	1
12 Focus	-0.22*	-0.02	0.02	-0.20*	-0.15*	-0.05	0.01	0.03	-0.07	-0.22*	-0.15*

Table 5. GEE Negative Binomial Results

	<u>Manufacturing Recalls</u>					<u>Design Recalls</u>				
	1	2	3	4	5	6	7	8	9	10
Ln volume	0.17 (0.17)	0.10 (0.18)	0.09 (0.18)	0.14 (0.15)	0.14 (0.15)	0.23* (0.10)	0.15 (0.10)	0.15 (0.10)	0.17 (0.10)	0.17 (0.10)
Ln employees	0.20 (0.22)	-0.64** (0.23)	-0.75*** (0.23)	-0.76*** (0.23)	-0.85*** (0.23)	0.40 (0.25)	-0.05 (0.26)	-0.04 (0.27)	-0.08 (0.29)	-0.07 (0.30)
Lagged Recalls	0.09 (0.06)	0.13* (0.06)	0.10+ (0.05)	0.11+ (0.06)	0.09 (0.06)	0.21*** (0.06)	0.26*** (0.06)	0.26*** (0.06)	0.23*** (0.05)	0.23*** (0.05)
Time since production	0.19*** (0.05)	0.20*** (0.04)	0.21*** (0.04)	0.22*** (0.04)	0.22*** (0.04)	0.13* (0.05)	0.11* (0.05)	0.11* (0.05)	0.12** (0.05)	0.13** (0.05)
US	-0.31 (0.21)	-0.66*** (0.17)	-0.56*** (0.17)	-0.72*** (0.19)	-0.64*** (0.18)	-0.24 (0.33)	-0.41 (0.26)	-0.44+ (0.26)	-0.37 (0.27)	-0.40 (0.27)
Toyota	0.37 (0.38)	-0.44 (0.36)	-0.41 (0.31)	-0.47 (0.35)	-0.45 (0.31)	0.33 (0.42)	-0.13 (0.29)	-0.13 (0.30)	-0.08 (0.30)	-0.07 (0.31)
Variety		0.10 (0.13)	0.04 (0.12)	0.15 (0.13)	0.10 (0.13)		0.26** (0.10)	0.27** (0.09)	0.28* (0.11)	0.30** (0.11)
Utilization		0.52*** (0.14)	0.56*** (0.14)	0.52*** (0.14)	0.55*** (0.14)		0.40*** (0.11)	0.40*** (0.11)	0.40*** (0.12)	0.39*** (0.12)
Focus		-1.36*** (0.38)	-1.44*** (0.35)	-3.35*** (0.73)	-3.26*** (0.64)		0.40 (0.29)	0.41 (0.30)	0.14 (0.40)	0.13 (0.41)
Variety*Utilization			0.22* (0.10)		0.19* (0.09)			-0.04 (0.09)		-0.05 (0.10)
Variety*Focus				-1.95*** (0.47)	-1.75*** (0.40)				-0.34 (0.31)	-0.38 (0.32)
Model-years	232	232	232	232	232	232	232	232	232	232
No. of recalls	182	182	182	182	182	285	285	285	285	285
Wald Chi2	22.58	60.64	77.78	75.84	80.62	118.35	210.76	202.33	190.73	179.74

Standard errors in parentheses

+ p<0.10, \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

Table 6. Robustness Checks

	Mfg Recalls GEE model 1	Design Recalls GEE model 2	Mfg Unique Recalls 3	Mfg Shared Recalls 4	Design Unique Recalls 5	Design Shared Recalls 6	Mfg Recalls all plants 7	Design Recalls all plants 8	Mfg Recalls Neg Bin. RE 9	Design Recalls Neg Bin. RE 10
Ln volume	0.14 (0.15)	0.17 (0.10)	0.31 (0.19)	0.07 (0.07)	0.02 (0.15)	0.12 (0.10)	0.07 (0.09)	0.07 (0.07)	0.23* (0.11)	0.20+ (0.10)
Ln employees	-0.85*** (0.23)	-0.07 (0.30)	-0.48+ (0.25)	-0.10 (0.11)	-0.13 (0.24)	-0.35* (0.17)	-0.16 (0.13)	-0.15 (0.11)	-0.17 (0.17)	0.04 (0.16)
Lagged Recalls	0.09 (0.06)	0.23*** (0.05)	-0.23+ (0.13)	0.08* (0.04)	0.06 (0.09)	0.28*** (0.07)	0.14** (0.05)	0.23*** (0.05)	0.10 (0.08)	0.27*** (0.07)
Time since production	0.22*** (0.04)	0.13** (0.05)	-0.01 (0.07)	0.19*** (0.03)	0.09 (0.07)	0.23*** (0.05)	0.17*** (0.04)	0.14** (0.04)	0.21*** (0.04)	0.11** (0.04)
US	-0.64*** (0.18)	-0.40 (0.27)	-0.75+ (0.43)	-0.10 (0.21)	-0.46 (0.73)	-0.46 (0.43)	-0.56*** (0.16)	-0.53** (0.20)	-0.37 (0.25)	-0.44* (0.20)
Toyota	-0.45 (0.31)	-0.07 (0.31)	0.26 (0.58)	-0.53* (0.22)	0.13 (0.87)	-0.29 (0.52)	-0.19 (0.20)	-0.47 (0.33)	-0.02 (0.44)	-0.03 (0.38)
Variety	0.10 (0.13)	0.30** (0.11)	-0.16 (0.22)	0.12 (0.09)	0.06 (0.23)	0.35** (0.12)	0.05 (0.10)	0.31*** (0.09)	0.08 (0.10)	0.29** (0.09)
Utilization	0.55*** (0.14)	0.39*** (0.12)	0.37+ (0.22)	0.19*** (0.05)	0.38** (0.13)	0.39*** (0.11)	0.32*** (0.09)	0.40*** (0.08)	0.37*** (0.10)	0.31*** (0.09)
Focus	-3.26*** (0.64)	0.13 (0.41)	-2.62** (0.92)	-0.51*** (0.14)	0.25 (0.46)	-1.98+ (1.03)	-3.15*** (0.70)	0.14 (0.39)	-2.89* (1.15)	0.31 (0.32)
Variety*Utilization	0.19* (0.09)	-0.05 (0.10)	0.34+ (0.17)	0.03 (0.05)	-0.07 (0.13)	0.01 (0.10)	0.15+ (0.08)	-0.09 (0.08)	0.17+ (0.09)	-0.06 (0.08)
Variety*Focus	-1.75*** (0.40)	-0.38 (0.32)	-1.52** (0.51)	-0.25 (0.17)	-0.47 (0.39)	-1.02 (0.73)	-1.61*** (0.41)	-0.39 (0.29)	-1.47+ (0.75)	-0.16 (0.33)
Model-years	232	232	232	232	232	232	287	287	232	232
No. of recalls	182	285	64	118	128	157	223	344	182	285
Wald Chi2	80.62	179.74	70.80	105.33	29.13	130.09	71.62	322.32	68.27	105.36

Standard errors in parentheses

+ p&lt;0.10, \*p&lt;0.05, \*\*p&lt;0.01, \*\*\*p&lt;0.001

Table 7. Utilization Quartiles

Utilization Quartiles	Quartile Range
1 <sup>st</sup> quartile	20.3%-62.0%
2 <sup>nd</sup> quartile	62.1%-82.6%
3 <sup>rd</sup> quartile	82.7%-98.0%
4 <sup>th</sup> quartile	98.1%-132.0%

Table 8. Utilization Quartile GEE Regression

	Mfg Recalls
Ln volume	0.14 (0.20)
Ln employees	-0.37+ (0.22)
Lagged Recalls	0.14* (0.07)
Time since launch	0.18*** (0.05)
US	-0.46** (0.16)
Toyota	-0.27 (0.36)
Variety	0.12 (0.13)
Focus	-1.37*** (0.39)
Utilization 2 <sup>nd</sup> Quartile	0.08 (0.31)
Utilization 3 <sup>rd</sup> Quartile	0.57+ (0.32)
Utilization 4 <sup>th</sup> Quartile	0.92** (0.36)
Model-years	232
No. of recalls	182
Wald Chi2	60.70

Standard errors in parentheses  
 + p<0.10, \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

Figure 1. Interaction Plot - Variety and Utilization for Manufacturing and Design Recalls

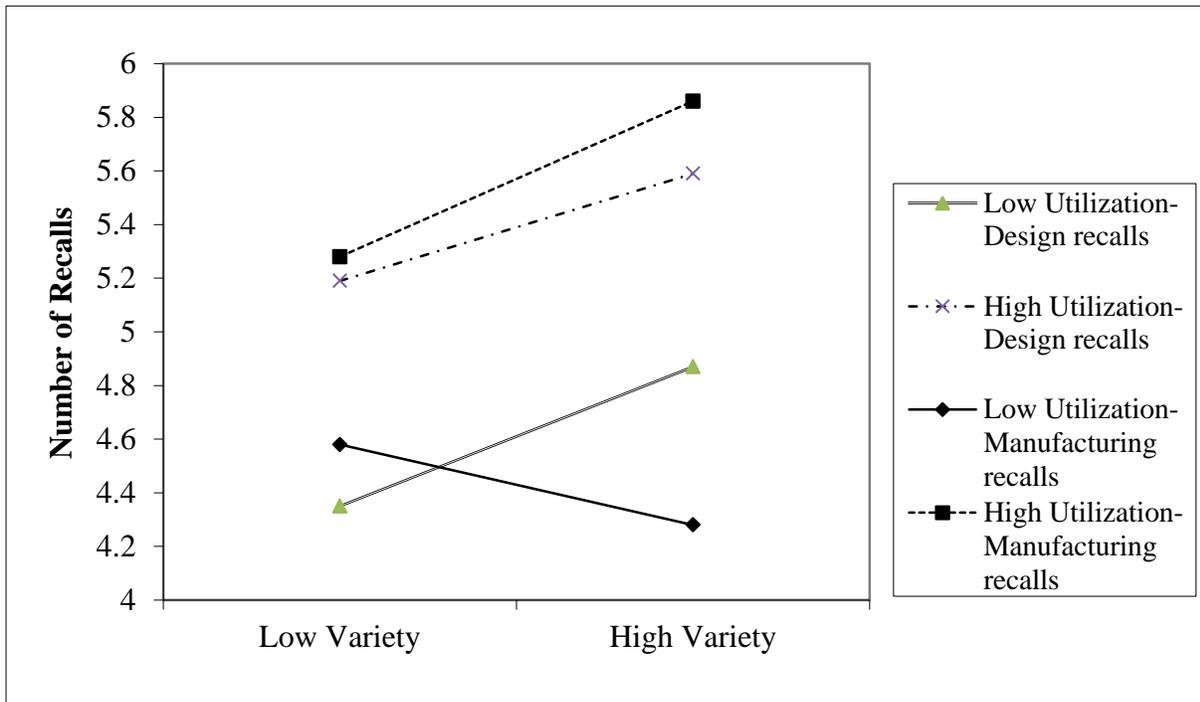


Figure 2. Conditional Effects Plot - Variety and Utilization on Manufacturing Recalls

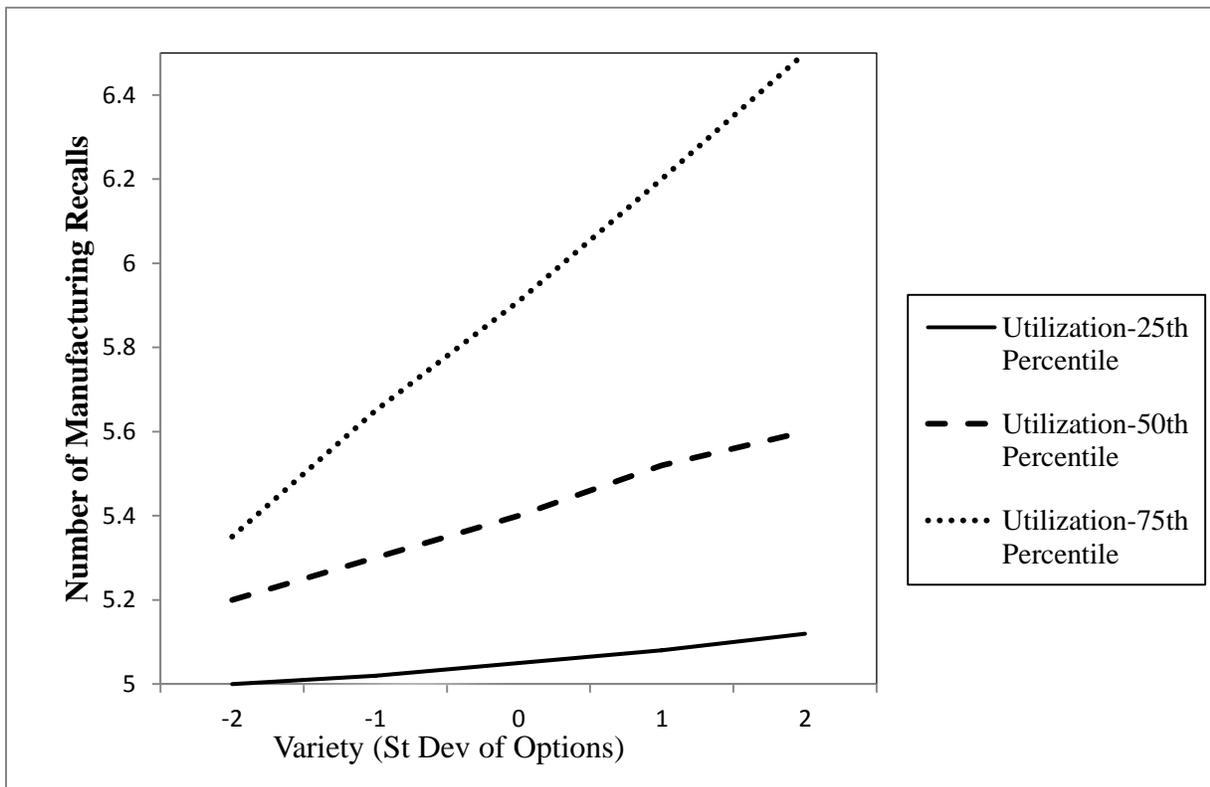


Figure 3. Interaction Plot – Variety and Focus for Manufacturing and Design Recalls

