Electric Vehicle Recall Analysis Using Machine Learning (LLMs)

Bachelor's Thesis

Advisor:
Frank Ohnesorge

Chair of Quantitative Marketing and Consumer Analytics
L5, 2 - 2. OG
68161 Mannheim
Internet: www.quantitativemarketing.org

Table of Content

Abstract

This thesis explores methodologies and implications of vehicle recall analysis in academic research, while focusing particularly on electric vehicles and the application of machine learning processes within the research. The role of product recalls in ensuring product safety, customer satisfaction and regulatory as well as organizational compliance is investigated. Through a detailed literature review, this thesis examines different streams of research in the field of vehicle recall analysis, including their methodologies, distinct findings, and implications for electric vehicles. Traditional statistical methods as well as qualitative assessments are discussed, and emerging machine learning approaches are highlighted. Insights into recall timing, general recall trends, safety and quality concerns, and strategic management surrounding vehicle recalls are provided while focusing on unique challenges and considerations associated with electric vehicles. The findings suggest that machine learning applications may greatly enhance vehicle recall analysis by improving on existing statistical methods through their improved handling of large amounts of data. Lastly, the findings are discussed critically while highlighting managerial implications and providing suggestions for future research.

1. Introduction

This paper encompasses a thorough literature review of the existing academic research regarding vehicle recall analysis. The inclusion of electric vehicles and the different methodologies in recall analysis are emphasized and individually considered.

No matter if the object is ignition switches, floor mats, Barbie dolls or peanut butter: product quality and safety matter. For this reason, product recalls have become a regular occurrence among many industries; with the largest emphasis within the automotive, food, toy, and pharmaceutical industries.

Product recalls may be voluntary (initiated by the company without external influence) or mandatory (initiated by regulatory entities e.g. the NHTSA[1]). Therefore, product recalls are highly relevant in not only maintaining customer satisfaction but also product safety and ensuring organizational compliance. Especially in industries like pharmaceutical, food or automotive, organizations must consider the importance of product safety regarding both personal harm and legal implications. Issues with products may lead to injury, illness, or even death and consequentially to lawsuits and loss of confidence in the company by different stakeholders such as customers or investors.

Before discussing the literature in detail, it is first necessary to distinguish between product recalls and product harm crises. While often used synonymously in the literature, they have different meanings and need to be distinguished from one another. In his 2024 paper, Astvansh explains: “A harm crisis is an unfavourable circumstance in which a firm finds itself, whereas a recall is the firm’s corrective response to the circumstance” (Astvansh 2024, p. 32). I have made sure to use the correct respective term throughout this paper.

[1] National Highway Traffic Safety Administration, United States

Recently, methods of machine learning have surfaced in other academical areas and are therefore of interest regarding their application to vehicle recall analysis. Machine learning is a type of artificial intelligence in which computers use large amounts of data to learn how to fulfil tasks rather than being programmed to do them. This differs from statistical analysis due to statistics being inherently backward-looking, as their methods can only be applied to already existing data. With machine learning, once the computer is trained, it may fulfil tasks (e.g. do research, evaluate data) on its own in a thus forward-looking manner. Within machine learning itself there are many different methods, one example and our focus for this paper are large language models. Large language models use large amounts of data in order to understand and produce text in a way that is similar to the way humans do.

Product recalls have repeatedly been studied to understand the relevance and impact on different factors. In academia, both qualitative and quantitative methods have been used to accomplish these feats.

This includes understanding how and if recall timing matters (Eilert et al. 2017), how recalls impact marketing activities and their benefits (Bijgaart and Cerruti 2024; Cleeren, van Heerde, and Dekimpe 2013), how organizations learn and improve through recalls (Haunschild and Rhee 2004; Kalaignanam, Kushwaha, and Eilert 2013), how recalls impact financial and sales performance (Bernon et al. 2018; Gao et al. 2015) and many others discussed in this paper.

The remainder of this paper is structured as follows: First, methods of product recall analysis are introduced, and the topic is motivated. Next is an extensive discussion of existing research regarding vehicle recall analysis followed by its' implications for electric vehicles. Lastly, the findings are discussed, evaluated in a critical manner and managerial implications as well as limitations and suggestions for future research are presented.

2. Methods of Product Recall Analysis

This section describes the topic of product recall analysis and gives an overview of existing research on automotive recall analysis in general. It explains and lays out the different methodologies used and the resulting conclusions in academia. Finally, it discusses the implications for electric vehicles stemming from these insights.

2.1 Motivation for Product Recall Analysis

Product recalls have been studied for a long time, in many different industries and with countless different emphases. However, most existing research focused on solving questions regarding vehicle recalls directly without considering the motivation for the analysis itself.

Apart from the usually high attention to large recalls in the media and by the public – as shown by many examples such as the Takata airbag disaster or General Motors' ignition switch scandal – there are additional reasons to investigate. Such quality issues can further have detrimental effects on companies as shown by Boeing's recent airplane door incidents. The company's stock price dropped almost 10% within the first week of the 737 incident and has since lost more than 20% at the time of submission of this paper.

Within the mainstream literature, many researchers focus on problems or cases that have clear implications either for managers, policy makers or both. One example is the work of Shah, Ball and Netessine (2017), who examined the impact of different factors directly related to production on product recalls in the automotive industry. They found that both increases in product variety (i.e. how many different models are built in the same plant) and plant utilization (i.e. percentage of capacity) lead to additional manufacturing-related recalls and therefore costs (Shah, Ball, and Netessine 2017, p. 2439). While also quoting concrete figures, Shah, Ball and Netessine provide clear guidance and impactful information for managers through their work.

Highly valuable insights for policy makers are provided by Kalaignanam, Kushwaha and Eilert. They related past product recalls in the automotive industry to future injuries and recall frequency through statistical analysis. Their results show: product recalls significantly reduce future injuries and lead to less subsequent recalls (Kalaignanam, Kushwaha, and Eilert 2013, p. 53). Cautioning policy makers and firms not to start recalling vehicles for the sake of it, Kalaignanam, Kushwaha and Eilert provide valuable insights informing future policy decisions.

It is therefore sensible to study product recalls in a detailed manner to inform organizational decision making, shed light on popular issues and aid policy making and regulatory entities.

In academia, product recalls are generally either qualitatively or quantitatively evaluated. In my research regarding vehicle recall analysis, these practices are extended by machine learning methodology. These different methodologies will be discussed in detail in the next section.

2.2 Methodologies of Vehicle Recalls

In the methodology of vehicle recall analysis, there are three main fields.

Firstly, most researchers have focused their studies on quantitative aspects such as the time between recall discovery and recall issue or stock price developments. As these matters are usually well quantifiable, statistical analysis is used. This includes correlation analysis, descriptive statistics, and coefficient analysis. One example is the work of Ni and Huang. In their 2018 study, they analyse the 6 largest automakers in the US regarding their time between the discovery of a defect and the issuance of a recall. They find this period extends when a) recalls originate externally (e.g. complaints, policy), b) recalls stem from issues with supplied parts, c) recalls arise from design rather than manufacturing mistakes, d) multiple models are involved in the recall (Ni and Huang 2018, p. 89). Being the least complicated way of evaluating quantitative issues, statistical analysis is the most popular form of product recall analysis in the literature.

Secondly, there is a stream of research focusing on more qualitative topics including recall trends or organizations' general behaviour around product recalls. Correspondingly, more qualitative methods tend to be used, while statistical analysis often still plays a supporting part. In a study on recall trends, Murphy et al. first qualitatively assess the changing technology in the automotive industry with the rise of autonomous vehicle technology. Then, they combine their assessment with a statistical analysis component to find a CAGR of product recalls regarding issues with ECUs[2], software or sensors of 23,9% (Murphy et al. 2019, p. 12). This way, qualitative assessments that play a large role in product recall research are made empirically valid.

Lastly, a small but emerging stream of research is focusing on using methods of or related to machine learning. Ramya and Ganapathy compared two different machine learning algorithms to evaluate their performance in a task related to in-field vehicle quality performance (Ramya and Ganapathy 2022). Similar works have been provided in related fields, e.g. for car insurance policies (Alshamsi 2014) or traffic accident detection (Zhang et al. 2015).

Making use of these different methodologies, researchers found many crucial insights. These results and their respective implications for electric vehicle strategy, management and recall behaviour will be discussed in the following section.

---

[2] Electronic Control Unit used to control specific vehicle functions

3. Implications for Electric Vehicles

The existing research provides plenty of implications for electric vehicles. It can be divided into two main arms: research regarding electric vehicles in general using machine learning and research regarding vehicle recalls in general using different methodologies and setting various points of focus.

In the research regarding electric vehicles using machine learning, most authors have focused on crash prevention and safety (e.g. Balan et al. 2022, Zhu et al. 2022) or battery health and charging (e.g. Basnet and Hasan Ali 2020, Hecht, Figgener, and Sauer 2021, Naresh, Ratnakara Rao, and Prabhakar 2024, Sheikh et al. 2020). This leads to the conclusion that, at least in recent years, research regarding product recalls of electric vehicles specifically has been of less interest. As electric vehicles push the boundaries of technologic innovation in the automotive industry, it is sensible to evaluate them in terms of crash prevention and safety. Many electric cars come with autonomous driving functionalities leading to new challenges regarding crash prevention. Additionally, there is a great public interest in the safety of electric cars due to extensive media coverage regarding incidents of e.g. electric vehicle batteries catching fire, crashes et cetera. Focusing further on batteries, the charging infrastructure and potential depreciation of their cars due to battery health is a large concern for many consumers and therefore also for many manufacturers (CNBC 2024). Correspondingly, academia has focused on these areas for machine learning applications before tackling electric vehicle recalls specifically.

The vehicle recall-related research – as differentiated from general electric vehicle research using machine learning – may firstly be divided broadly into ex-ante (before recall issuance) and ex-post (after recall issuance) focused work.

The ex-ante focused research is much smaller in comparison. One key work was provided by Zhang et al. in 2015. They evaluated activity in an internet forum for Toyota owners. Using text mining and dictionary-based analysis methods, they found “overlapping components in user generated contents and official recall notices” and were able to correctly predict 50% of recalls before they occurred (Zhang et al. 2015, p. 49). This research was encouraged by multiple prior works, including Bae and Benítez-Silva (2011) and is still of strong interest in following research (Ni and Huang 2018; Shah, Ball, and Netessine 2017). While limited, the work by Zhang et al. shows promising results and could spur further research on the topic. By including more than one forum, multiple car brands and more sophisticated methods such as machine learning, the research may be significantly advanced in this area. It may also be expanded to include or specifically target electric vehicles and how their recalls may be examined ex-ante, especially under the following main areas of interest regarding vehicle recalls.

Most of the remaining recall-related research is focused on analysing vehicle recalls and product-harm crises ex-post. Within this academic area, either traditional methodologies like statistical analysis or newer approaches related to machine learning were used to analyse vehicle recalls.

In the area of machine learning methodologies, one paper by Maione, Kaminski and Baraldi tested different machine learning methodologies (clustering[3], classification[4], regression[5]) to evaluate their performance regarding vehicle recall analysis, specifically the general increase in vehicle recalls and the number of vehicles involved. They were able to analyse over 12,000 recalls for their work, much more than traditionally done in academia.

[3] Automatic identification of data groups by the machine

[4] The machine is done being trained with training data and now evaluated on test data

[5] The machine is used to predict continuous results

Their findings attested machine learning and its related methods a very high potential for future research (Maione, Kaminski, and Baraldi 2023, p. 1).

In a similar study published in 2022, Ramya and Ganapathy focused on the performance of two different machine learning algorithms, namely a Random Forest algorithm and K-Nearest-Neighbor (KNN) algorithm[6]. Their goal was to evaluate the performance of each algorithm and compare them to find out which was more effective at evaluating vehicle quality. They found the Random Forest algorithm to be superior to the KNN algorithm. It scored a 7% higher value regarding accuracy, 0.7% higher value regarding recall and proved to be 3% more precise than the KNN algorithm (Ramya and Ganapathy 2022, p. 550).

Within the research relying on methodologies of statistical analysis, the main areas of interest are in order of magnitude a) general recall trend analysis, b) reliability, accidents, quality and safety, c) brand preference, firm value and advertising, d) recall timing, e) lobbying and compliance, d) voluntary versus involuntary recalls, e) plant operations.

3.1 General Recall Trend Analysis

In general recall trend analysis, mostly statistical analysis was used. The main trends discovered were that recalls are expensive for manufacturers, their frequency and severity is increasing, and they differ strongly between manufacturers.

Ahsan found in both 2010 and 2013, that costs incurred in a product recall tend to outsize those incurred in a regular forward distribution setting by a factor of two to three (Ahsan 2010, p. 1123, 2013, p. 14). Considering that – as of 2024 – electric vehicles tend to be more expensive than internal combustion engine cars (University of Michigan 2024), this

[6] Both machine learning algorithms are used for classification and regression related issues

supports the assumption that product recalls for electric vehicles may also increase in costs and may be more expensive than those for internal combustion engine cars.

In recent developments, both frequency and severity of recalls have been increasing (Murphy et al. 2019, p. 14,23). Two main factors influencing this have been identified in the literature. On the one hand, cars – and especially electric cars – are becoming increasingly technologically advanced and complicated. Modern cars have hundreds of millions of lines of code, while a Boeing 787 makes due with only 6.5 million lines (Pancik et al. 2018, p. 1). On the other hand, this increased complexity leads to increasing customer complaints. More customer complaints in return lead to more product recalls (Ahsan 2010, p. 1127). Considering the findings of Yongqin et al. – an increase in voluntary and a decrease in involuntary recalls – this leads to the assumption that automotive manufacturers are acutely concerned with customer satisfaction (Yongqin et al. 2022, p. 1). Looking into the future of electric cars, this implies a closer relationship between manufacturer and customer. New technological developments may also allow for over-the-air updates or fixes that are made possible with electric vehicle technology but were not possible with internal combustion engine cars in the past. This may also aid with the fluctuating recall completion rate of 50-70% in the US (Yongqin et al. 2022, p. 5).

Strong differences in number of recalls can be found between manufacturers. In their 2007 study, Bates et al. found that „East Asian producers comprise one-third of the dataset, but represent eight out of the 10 companies with the lowest recall rates.“ (Bates et al. 2007, p. 209). Ahsan also found that the top seven car manufacturers in the US, of which only Honda, Toyota and Nissan are Asian, make up for 70% of all car recalls (Ahsan 2010, p. 1127). It can therefore be assumed that this trend will continue into the aera of electric vehicles, with East Asian manufacturers producing more reliable and less failure-prone cars and the largest car manufacturers making up most of the recall traffic.

The hypothetical main factors for recall completion are vehicle age, recall campaign size, severity, multitude of recalls on the same vehicle and vehicle type (Malec, Smith, and Smuts 2021, p. 38). In analysing these, Malec, Smith, and Smuts found that completion rates increase with multitude and severity and are higher for luxury vehicles versus regular consumer vehicles (Malec, Smith, and Smuts 2021, pp. 53–54). Their findings illustrate the importance of research on reliability, accidents, quality, and safety.

3.2 Reliability, Accidents, Quality, and Safety

In the literature, four main constituents have been found regarding the impact of product recalls in the automotive industry on accidents, future recalls, and injuries. In this context, product recalls are a large force of good. Mainly, recalls reduce the number of accidents, the severity of accidents, the number of future recalls and the number of injuries.

Bae and Benitez-Silva investigated the effect of vehicle recalls on the number of accidents in the US-market using a synthetic panel data model in combination with statistical analysis. They found a reduction of accidents between 7.8 and 16 percent following a recall (Bae and Benítez-Silva 2011, p. 823). Later, they turned to study the severity of accidents in 2013. In their consecutive work, they discovered that “if a recall for a new-year model is issued, then the severity of injuries of accidents continuously diminishes during the first year after the recall […] because defects are repaired over time but also because drivers react by driving more carefully until the defects are fixed” (Bae and Benítez-Silva 2013, p. 1232). They therefore separated the impact of automotive product recalls into two effects: drivers’ behavioural response, i.e. being more careful when aware of potential threats, and the elimination of hazardous defects through the recall process itself.

The reduction of numbers of recalls and numbers of injuries was found by Kalaignanam, Kushwaha, and Eilert in 2013. Using statistical correlation analysis and univariate transfer function analysis they found that “increases in recall magnitude lead to decreases in future number of injuries and recalls. This effect […] is partially mediated by future changes in product reliability. The […] relationship […] is (1) stronger for firms with higher shared product assets and (2) weaker for brands of higher prior quality” (Kalaignanam, Kushwaha, and Eilert 2013, p. 41). Considering ongoing concerns regarding the safety of electric vehicles, mainly due to fires but also due to certain drive systems failing, these results are very promising. Manufacturers of electric vehicles may focus on improving quality control while at the same time issuing recalls where necessary, thus improving trust and customer safety while saving costs in the long run through a reduced number of future recalls. Especially brands in the luxury and high-quality segment need to tread lightly however, as will be discussed in following segments.

3.3 Brand Preference, Firm Value, and Advertising

Both direct (e.g. vehicle repairs) and indirect costs (e.g. loss of sales) related to product recalls and those recalls’ effects on brand preference, firm value and advertising are of interest to many researchers.

Two papers headed by Texas A&M Professor Yan Liu found a negative effect of product recalls in the automotive industry on both brand preference and firm value. In 2015, they evaluated the effect on brand preference in detail. They found increasing negative effects with brand quality, media coverage and recall severity (Liu and Shankar 2015, p. 2533). Following prospect theory, when brands are more reputable, customers’ expectations are higher and therefore more easily and more strongly disappointed in the event of a recall. Consecutive research followed in 2017, where more emphasis was placed on advertising and firms’ approaches to mitigate negative effects of product recalls. Then, Liu, Shankar and Yun found a negative effect of product recalls in the automotive industry on firm value. By working with short-term abnormal returns analysis and long-term calendar-time portfolio analysis (CTAR) in combination with statistical correlation, they found a lagging negative effect over time. This effect may be mitigated by the firm through initiating the recall voluntarily and smart advertising strategy (Liu, Shankar, and Yun 2017, p. 45). Here, it is important to focus on promotion advertisement over brand advertisement during the product recall, while inversing that focus after the recall is concluded. If an automaker advertises the affected name plate (e.g. Camry) or a promotion (e.g. 0% finance) during the product harm crisis, the effects on brand preference and firm value are positive in the short-term but negative in the long-term (Liu, Shankar, and Yun 2017, p. 44). On the contrary, brand advertising (e.g. Toyota) has a positive effect at least on firm value in the long run (Liu, Shankar, and Yun 2017, p. 31).

Due to insufficient research to date regarding electric vehicles specifically, these implications are as of now only valid for internal combustion engine cars. As they are very general however, it seems probable that they will also apply to electric vehicles. The surroundings of a recall (i.e. media coverage) should ideally be identical for both internal combustion engine and electric vehicles. Additionally, stakeholders do not change in either case. Customers, investors, media etc. are involved either way. The insights are therefore very valuable to managers involved in automotive product recalls for all types of drive systems. The reasons for the product recall and customer concerns may differ between electric vehicles and internal combustion engine cars. However, it may still be sensible to consider negative effects on brand preference and firm value while focusing on one’s own firms’ reputation and advertising or general product recall mitigation strategy.

3.4 Recall Timing

When developing a new product, firms must weigh the costs and benefits of going to market early versus taking more time for product testing and safety assurance. This has become especially important in recent years with electric vehicles changing established manufacturing processes, challenging known operations, and posing new safety questions. A game-theoretic model was used by Kim, Koenigsberg and Ofek in 2022 to evaluate this trade-off. They discovered the driving forces behind the decision making of firms in this type of situation. Namely: recall avoidance, profiting from first mover advantages, avoidance of late mover disadvantages, and avoidance of making the same move as the competition, thus intensifying it (Kim, Koenigsberg, and Ofek 2022, p. 8905). These forces pose plenty of implications for electric car manufacturers. With average costs of $12 million for each recall, avoiding such is crucial (Singh and Grewal 2023, p. 745). These numbers were established considering data related to mainly internal combustion engine cars between 2008 and 2016. Therefore, recall costs may be even higher for electric vehicles, as discussed before. On the other hand, especially for manufacturers that were not building electric vehicles in the past, first versus late mover concerns may be very pressing. Examining Hyundai Motor Group for example, that are dominating their domestic market of South Korea with over 90% market share (Statista, 2023). Had they not decided to move into electrifying some of their vehicles and offering new or improved models by the early 2010s, it might have cost them significant market share to date.

With quality concerns at the forefront, the timing of recalls and firms' decisions regarding them has been studied in the literature. As described before, Ni and Huang (2018) analysed the 6 largest automakers in the US regarding their time between discovery of a defect and the issuance of a recall. They find this period extends when a) recalls originate externally (e.g. complaints, policy), b) recalls stem from issues with supplied parts, c) recalls arise from design rather than manufacturing mistakes, d) multiple models are involved in the recall (Ni and Huang 2018, p. 89). Similarly, Eilert et al. found longer time to recall for more severe issues and more diversified brands. When a brand is well known for being reliable and/or has had major recalls recently, time to recall tends to fall (Eilert et al. 2017, p. 111).

Considering the development of electric cars and the negative implications of product recalls for branding and financial performance discussed before, automotive firms should be careful when diversifying their product portfolio. More models lead to more complexity, more suppliers, a longer time to recall and therefore higher recall-associated costs. This is especially important for brands that focus on both internal combustion engine cars and electric vehicles. Luxury vehicle manufacturers should be especially careful not to water down their image and not to be incurring higher costs than their low-cost competitors due to shorter recall times. When pursuing such a strategy, these risks need to be taken into consideration.

3.5 Lobbying and Compliance

Product recalls may be voluntary or involuntary. In the case of involuntary recalls, they are dictated by regulatory entities, such as the NHTSA in the United States or the European Commission. In the EU, lobbying and related actions of members of the parliament are regulated by the Code of Conduct for Members of the European Parliament with respect to financial interests and conflicts of interest.

The influence of lobbying and therefore the automotive industry on those regulatory entities has been studied by Singh and Grewal in 2023. Via a statistical regression analysis, they found two significant figures. On the one hand, there is a direct influence of lobbying expenditure on recall volume. For an additional lobbying spend of $404,367, one fewer voluntary recall is observed. Similarly, for $1.66 million invested into lobbying, one fewer involuntary recall is observed. On the other hand, with an assumed cost of $50 per vehicle involved in the recall, Singh and Grewal estimate almost $12 million in savings for each recall that is avoided through lobbying expenses (Singh and Grewal 2023, p. 745).

Weighing up half or up to one and a half million dollars in costs versus up to $12 million in savings, an increase in lobbying expenditure may be the logical decision for managers. According to Sedgwick (2022), electric vehicles were twice as likely to experience a recall than internal combustion engine cars between 2017 and the first half of 2022. Considering this increased risk for electric vehicle manufacturers, managers may opt to increase lobbying expenditure in line with an increase in production and sales of electric vehicles.

3.6 *Voluntary Versus Involuntary Recalls*

The role of volition in vehicle recalls has furthermore been considered by Haunschild and Rhee. They investigated how voluntary versus involuntary recalls lead to organizational learning effects. Learning in this case meaning reduction of future product recalls. They found that “voluntary recalls result in more learning than mandated recalls” (Haunschild and Rhee 2004, p. 1545). This implies that firms which focus on recalling their products when necessary - without waiting for regulatory influence or only considering recalls with regulatory pressure - tend to suffer less recalls in the future.

For electric vehicle manufacturers specifically, this may motivate a culture of ownership. In this relatively new field for most automotive manufacturers, focusing on owning eventual manufacturing or design mistakes and issuing recalls where appropriate should foster strong customer relations, save costs, and strengthen the company’s market position in the long run. However, this should not lead to issuing product recalls prematurely. Learning is indeed higher from voluntary recalls, but involuntary recalls also have positive effects in reducing future recalls and should thus not be disregarded.

3.7 Plant Operations

In terms of the impact of plant operations on product recalls, clear conclusions can be drawn. Shah, Ball and Netessine examined the influence of different factors directly related to production on vehicle recalls in the automotive industry. They evaluated product variety (i.e. how many different models are built in the same plant), plant variety (i.e. in how many different plants the same model is produced) and capacity utilization of North American automotive manufacturers and NHTSA recall data. They found that both increases in product variety and plant utilization (i.e. percentage of capacity) lead to additional manufacturing-related recalls and therefore costs. As does an increase in plant variety combined with increased capacity utilization (Shah, Ball, and Netessine 2017, p. 2439). Direct figures are also provided: firstly, if four options for a certain model are added, this increase in product variety statistically leads to two additional recalls and $46.2 million in associated costs. Secondly, products produced in a plant exceeding 100% capacity experience eight additional recalls estimated at $167 million (Shah, Ball, and Netessine 2017, p. 2439).

For electric vehicle manufacturers, this may encourage cautious operations. More options may satisfy consumer demand and lead to more sales. They may also however increase recalls and therefore costs in the long run. This trade-off must be carefully considered. Moreover, with some car manufacturers producing internal combustion engine cars in the same plants as electric vehicles (e.g. BMW, The Business Times 2024), capacity utilization might become an issue. Exceeding maximum capacity may lead to more productivity and enable manufacturers to satisfy both customers of internal combustion engine cars and electric vehicles but may also cause costly future vehicle recalls for both types of vehicles. Therefore, it is imperative in both cases to weigh up if the potential gains outweigh the potential losses.

Returning to the initial distinctions, the existing research is split up into the following categories. Firstly, general research regarding electric vehicles using machine learning, which recently was focused mainly on safety- and battery-related issues. This is differentiated from vehicle recall analysis using mixed methodologies.

Within the mixed methodology research, a distinction between ex-post and ex-ante focused research can be made. Finally, within the large body of ex-post focused research, papers can be categorized by their methodologies (some machine learning, mostly statistical analysis and a few qualitative assessments) and points of interest, i.e. general recall trends, reliability, accidents, quality and safety, brand preference, firm value and advertising, recall timing, lobbying and compliance, voluntary versus involuntary recalls and plant operations.

These findings and their implications will be discussed in detail in the following final section of this paper.

4. Discussion

This section discusses the findings regarding electric vehicle recall analysis using machine learning including large language models (LLMs) in academia. It aims at evaluating results in a critical manner, emphasizing managerial implications, acknowledging limitations, and suggesting possibilities for future research.

4.1 Critical Evaluation

In evaluating the findings within the existing literature regarding electric vehicle recall analysis using machine learning (LLMs) specifically, one must first acknowledge that such explicit research does not yet exist. This may be due to different reasons. Firstly, the field of machine learning in general and its use in research applications is very young. Compared to statistical analysis, which has been around for almost a whole century, researchers are entering uncharted territory in use of machine learning and large language models (LLMs) and may therefore be inclined to stick to tried and tested methods instead.

Additionally, electric vehicles have only gained significant traction and mainstream success in the past decade. Before, many automotive manufacturers were already investigating potential alternative drive systems using prototypes such as hydrogen combustion engines (BMW Hydrogen 7 1989), hybrid electric vehicles (Toyota Prius 1997), or fully electric vehicles (Mitsubishi i-MiEV 2009) with varying amounts of success. Only with Tesla pioneering the market with their flagship Model S in 2012 and large automakers like Renault or Smart following suit, have electric vehicles started to find their place in the mainstream.

The interplay of these factors likely leads to the lack of specific research regarding electric vehicle recall analysis using machine learning.

Broadening the scope and considering general vehicle recall analysis using machine learning, some research can be found. Specifically, the works by Maione, Kaminski and Baraldi (2023) and Ramya and Ganapathy (2022) focused on applying machine learning concepts to research in the automotive industry, with Maione, Kaminski and Baraldi qualitatively targeting vehicle recalls specifically and Ramya and Ganapathy investigating the performance of different algorithms in a more quantitative manner. Hopefully, more research on vehicle recall analysis in general and electric vehicle recall analysis specifically will follow in the future making use of machine learning methodologies.

By also including general vehicle recall analysis using different methodologies, the scope could be expanded even further. Through this lens, different streams of research with many different works over the years could be found. Namely: general recall trend analysis, reliability, accidents, quality and safety, brand preference, firm value and advertising, lobbying and compliance, voluntary versus involuntary recalls and plant operations. Depending on the type of research, mainly statistical analysis was used, sometimes combined with other qualitative methods.

The managerial implications of the findings of this broad spectrum of research will be discussed next.

4.2 Managerial Implications

The managerial implications stemming from academia are plentiful. Considering general recall trend analysis, researchers found that recalls are expensive for manufacturers, their frequency and severity is increasing, and they differ strongly between manufacturers (Ahsan 2010, p. 1123, 2013, p. 14; Bates et al. 2007, p. 209; Pancik et al. 2018, p. 1).

In the research regarding reliability, accidents, quality and safety, conclusions were that product recalls are in context a large force of good. They reduce the number of accidents, the severity of accidents, the number of future recalls and the number of injuries (Bae and Benítez-Silva 2011, p. 823, 2013, p. 1232; Kalaignanam, Kushwaha, and Eilert 2013, p. 41). If managers are concerned with a positive brand image, customer satisfaction and avoidance of legal and other costs, they should try to avoid any unnecessary recalls but at the same time handle those necessary ones as well as they can, i.e. being accommodating with customers and cooperative with regulatory entities, no matter if they are voluntary or involuntary. This may include a short time to recall, efficient communication both internally and directly with customers, and remedial marketing activities.

Considering the effect of product recalls on brand preference and firm value and their interplay with firms' advertising, researchers found that product recalls negatively affect both brand preference and firm value. Advertising may be used strategically by managers to mitigate both effects in the long and short term (Liu and Shankar 2015, p. 2533; Liu, Shankar, and Yun 2017, pp. 44–45).

Researchers found the driving forces behind the decision making of firms in situations surrounding product recalls using qualitative methods. They include recall avoidance, profiting from first mover advantages, avoidance of late mover disadvantages, and avoidance of making the same move as the competition, thus intensifying it (Kim, Koenigsberg, and Ofek 2022, p. 8905). Managers should consider these factors and the possibility of their competition being aware of these factors when making decisions around product recalls. When concerned about recalling their products quickly, managers may need to take the origin of the recall, supplier relations, design and manufacturing mistakes, and the amount of models potentially involved in the recall into account, as suggested by Ni and Huang (Ni and Huang 2018, p. 89). Other factors to examine may be the brands' amount of diversification, i.e. how many different models are produced and the severity of the issues (Eilert et al. 2017, p. 111).

The direct research regarding lobbying and compliance implies for managers to simply increase lobbying expenditure as a direct relationship between more lobbying expenses and less future recalls has been found (Singh and Grewal 2023, p. 745). This should be taken with reservations however, as automotive manufacturers firstly may not be able to reduce their product recalls to zero simply by spending as much money as possible on lobbying. Secondly, product recalls are an important contributor to customer welfare and public safety and should therefore be carried out when necessary.

Similarly, researchers found that "voluntary recalls result in more learning than mandated recalls" (Haunschild and Rhee 2004, p. 1545). Learning in this case meaning reduction of future product recalls, managers should consider these positive effects of product recalls, namely reducing the number of future recalls, and not only focus on the negative short-term consequences, e.g. costs.

Finally, researchers found the manufacturing-related factors contributing to an increase in product recalls. They found increases in product variety (i.e. how many different models are built in the same plant) and plant utilization (i.e. percentage of capacity) as well as increases in plant variety (i.e. in how many different plants the same model is produced) combined with increased capacity utilization contributing to an increase in future product recalls to varying degrees (Shah, Ball, and Netessine 2017, p. 2439). Managers should tread lightly when considering an increase in any of these areas and always weigh up potential costs and benefits.

In the following section, this paper's limitations are considered and suggestions for future research are identified.

4.3 Limitations and Future Research

The literature research done in this paper is not without flaw. Potential limiting factors may firstly be an incomplete overview of all existing literature regarding electric vehicle recall analysis using machine learning (LLMs). An evident opportunity for future research is expanding existing literature regarding electric vehicle recall analysis using machine learning (LLMs) directly. As the availability of data is crucial especially for machine learning methodologies, the easily available Western data may help with the first steps in this area.

As there was no such explicit research to be found on this topic specifically, the scope had to be broadened. The broader the scope, the more research may be relevant. Therefore, the more research is relevant, the more likely the chance that papers which may be deemed relevant by some are overlooked.

Considering the literature that was covered in this paper, additional limitations arise. First and foremost, most studies focus on the US market for different reasons, one being easily obtainable data from the NHTSA. Up until 2009, the US was the largest automotive market in the world before being overtaken by China (BBC News 2010). Therefore, it makes sense to use US data for automotive product recall analysis. However, European and Asian markets are close to on par with the US in terms of size, but rarely considered in most studies covered in this paper. As especially Asian countries tend to strongly differ in terms of culture from western markets such as Europe and the US, future research with a special focus on Asian countries may shed light on new considerations regarding automotive recalls.

Additionally, most researchers used statistical analysis in their studies. While this is not a limit per se, it should be uncomplicated for future researchers to expand on their research questions by simply replacing statistical methods with those of machine learning, including but not limited to large language models (LLMs).

Lastly, large automakers were exclusively considered in most research covered in this paper. While this is only a small limitation, future research may expand existing research questions by also considering smaller automotive manufacturers to see if there are any new or differing considerations.

Appendix A: Literature Table

Year of Publication & Authors	Title	Methodologies	Research quesiton	Key findings
Liu and Shankar 2015	The Dynamic Impact of Product-Harm Crises on Brand Preference and Advertising Effectiveness: An Empirical Analysis of the Automobile Industry	state space model, Kalman filter	How do product harm crises affect brand preference in the short- and long-term? What (mitigating) effects does advertising have on these effects?	product harm crises affect brand preference negatively both long- and short-term
Liu, Shankar, and Yun 2017	Crisis Management Strategies and the Long-Term Effects of Product Recalls on Firm Value	short-term abnormal returns analysis and long-term calendar-time portfolio analysis, statistical correlation	How do product recalls affect firm value in the short- and long-term? How can firm mitigate these effects?	product recalls negatively affect firm value both short- and long-term
Shah, Ball, and Netessine 2017	Plant Operations and Product Recalls in the Automotive Industry: An Empirical Investigation	statistical correlation analysis	How do different manufacturing-related factors affect recall frequency?	increases in product variety and plant utilization lead to an increase future recalls
Kalaignanam, Kushwaha, and Eilert 2013	The Impact of Product Recalls on Future Product Reliability and Future Accidents: Evidence from the Automobile Industry	statistical correlation analysis, univariate transfer function analysis	How do product recalls affect future accidents and future recall frequency?	product recalls reduce future injuries and future recall frequency
Bae and Benítez-Silva 2011	Do vehicle recalls reduce the number of accidents? The case of the U.S. car market	synthetic panel data model approach, statistical analysis	Do vehicle recalls reduce the number of accidents?	recalls reduce the number of accidents
Bae and Benítez-Silva 2013	The Effects of Automobile Recalls on the Severity of Accidents	statistical analysis	Do vehicle recalls reduce the severity of future injureies?	recalls reduce the severity of future injuries
Singh and Grewal 2023	Lobbying and Product Recalls: A Study of the U.S. Automobile Industry	statistical regression analysis	Do lobbying expenses reduce volumes of future recalls?	increasing lobbying expenditures leads to a reduction of both voluntary and involuntary recalls
Eilert et al. 2017	Does it Pay to Recall your Product Early? An Empirical Investigation in the Automobile Industry	statistical analysis, short-term event study methodology	How (quickly) do firms and the stock market each react to product recalls?	problem severity and brand diversification increase time to recall, stock markets punish recall delays
Ni and Huang 2018	Discovery-to-Recall in the Automotive Industry: A Problem-Solving Perspective on Investigation of Quality Failures	event history analysis, statistical correlation analysis	What factors affect firms' swiftness in reacting to product recalls?	firms take longer to recall vehicles when recalls originate externally, recalls stem from issues with supplied parts, recalls arise from design rather than manufacturing mistakes, multiple models are involved in the recall
Haunschild and Rhee 2004	The Role of Volition in Organizational Learning: The Case of Automotive Product Recalls	statistical correlation analysis	How do automakers learn from voluntary versus involuntary recalls?	voluntary recalls reduce future recalls more than involuntary recalls
Ahsan 2010	Understanding trends of car recalls	statistical analysis	Who initiates product recalls and how do their costs compare to forward distribution?	top seven car makers initiate 70% of recalls, recall costs are a multiple of 2-3 of regular forward distribution costs
Ahsan 2013	Trend analysis of car recalls: evidence from the US market	statistical analysis	How are product recalls initiated and what are their costs?	product recalls tend to be initiated by customer complaints, 70% are initiated by the top six car makers, product recalls are more expensive than regular forward distribution
Murphy et al. 2019	The impact of autonomous vehicle technologies on product recall risk	qualitative assessment, statistical analysis	How has recall frequence evolved over the past years and what impact does technology have on this development?	more recalls over the past years, increasng share of technology-related recalls
Yongqin et al. 2022	Analysis on the Trends and Characteristics of Vehicle Recalls in the United States	statistical correlation analysis	How has recall frequency of influenced and uninfluenced recalls evolved over the past years?	uninfluenced recalls are increasing, influenced recalls declining
Malec, Smith, and Smuts 2021	Recall and Vehicle Characteristics Associated with Vehicle Repair Rates	qualitative assessment, statistical analysis	What factors affect vehicle repair rates? How many vehicles are repaired?	key factors influencing vehicle repair rates are age/recall campaign size/severity/multitude of recalls on same vehicle/vehicle type, recall completion rates are higher for severe defects, vehicles with more than one recall, vehicles made by luxury manufacturers
Bates et al. 2007	Motor vehicle recalls: Trends, patterns and emerging issues	statistical analysis	What are key recall trends and factors in recent years?	recalls increase over time, substantial differences between manufacturers in recall rates
Maione, Kaminski, and Baraldi 2023	The automotive recall data search and its analysis applying machine learning	machine learning	How can machine learning be used to investigate worldwide recall trends?	machine learning has a high potential in recall research with different methodologies
Ramya and Ganapathy 2022	Evaluation of Vehicle Quality Performance using Random forest in Comparison with KNN to measure the Accuracy, Recall, and Precision	machine learning	How does a Random Forest algorithm perform in vehicle quality evaluation?	Random Forest algorithm performs better than KNN algorithm, best at evaluating quality
Zhang et al. 2015	Predicting Vehicle Recalls with User-Generated Contents: A Text Mining Approach	machine learning (text mining)	How can text mining be used to predict vehicle recalls?	text mining model was able to predict 50% of recalls correctly
Kim, Koenigsberg, and Ofek 2022	I Don't "Recall": The Decision to Delay Innovation Launch to Avoid Costly Product Failure	game-theoretic model	What are the key influences on firms' decision surrounding product recalls?	the key factors are recall avoidance, profiting from first mover advantages, avoidance of late mover disadvantages, and avoidance of making the same move as the competition, thus intensifying it

Appendix B: Comparative Literature Table

	Methodologies			Research area							
Year of Publication & Authors	statistical methods	qualitative methods	machine learning	brand preference & advertising	reliability, accidents, quality, safety	lobbying & compliance	timing of recalls	general recall trends	volition	plant operations	methodology evaluation
This paper	x	x	x	x	x	x	x	x	x	x	x
Liu and Shankar 2015	x			x							
Liu, Shankar, and Yun 2017	x			x							
Shah, Ball, and Netessine 2017	x									x	
Kalaignanam, Kushwaha, and Eilert 2013	x				x						
Bae and Benítez-Silva 2011	x				x						
Bae and Benítez-Silva 2013	x				x						
Singh and Grewal 2023	x					x					
Eilert et al. 2017	x						x				
Kim, Koenigsberg, and Ofek 2022		x					x				
Ni and Huang 2018	x						x				
Haunschild and Rhee 2004	x								x		
Ahsan 2010	x							x			
Ahsan 2013	x							x			
Murphy et al. 2019	x	x						x			
Yongqin et al. 2022	x							x			
Bates et al. 2007	x							x			
Malec, Smith, and Smuts 2021	x	x			x						
Maione, Kaminski, and Baraldi 2023			x					x			x
Ramya and Ganapathy 2022			x		x						x
Zhang et al. 2015			x				x				x

References

Ahsan, K. (2010), "Understanding trends of car recalls," in *2010 IEEE International Conference on Industrial Engineering and Engineering Management*, 1123–27.

Ahsan, Kamrul (2013), "Trend Analysis of Car Recalls: Evidence from the US Market," *International Journal of Managing Value and Supply Chains*, 4, 1–16.

Alshamsi, Asma S. (2014), "Predicting car insurance policies using random forest," in *2014 10th International Conference on Innovations in Information Technology (IIT)*, 128–32.

Astvansh, Vivek (2024), "What Is (and Isn't) a Product Recall?," SSRN Scholarly Paper, Rochester, NY.

Bae, Yong-Kyun and Hugo Benítez-Silva (2011), "Do vehicle recalls reduce the number of accidents? The case of the U.S. car market: Journal of Policy Analysis & Management," *Journal of Policy Analysis & Management*, 30 (4), 821–62.

Bae, Yong-Kyun and Hugo Benítez-Silva (2013), "The Effects of Automobile Recalls on the Severity of Accidents: Economic Inquiry," *Economic Inquiry*, 51 (2), 1232–50.

Balan, Gunapriya, Singaravelan Arumugam, Suresh Muthusamy, Hitesh Panchal, Hossam Kotb, Mohit Bajaj, Sherif S. M. Ghoneim, and Undefined Kitmo (2022), "An Improved Deep Learning-Based Technique for Driver Detection and Driver Assistance in Electric Vehicles with Better Performance," *International Transactions on Electrical Energy Systems*, 2022, e8548172.

Basnet, Manoj and Mohd. Hasan Ali (2020), "Deep Learning-based Intrusion Detection System for Electric Vehicle Charging Station," in *2020 2nd International Conference on Smart Power & Internet Energy Systems (SPIES)*, 408–13.

Bates, Hilary, Matthias Holweg, Michael Lewis, and Nick Oliver (2007), "Motor vehicle recalls: Trends, patterns and emerging issues," *Omega*, 35 (2), 202–10.

Bernon, Michael, Marko Bastl, Wenqian Zhang, and Mark Johnson (2018), "Product recalls: The effects of industry, recall strategy and hazard, on shareholder wealth," *International Journal of Business Science & Applied Management (IJBSAM)*, 13 (1), 1–14.

Bijgaart, Inge van den and Davide Cerruti (2024), "The Effect of Information on Market Activity: Evidence from Vehicle Recalls," *The Review of Economics and Statistics*, 106 (1), 230–45.

Cleeren, Kathleen, 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.

Eilert, Meike, Satish Jayachandran, Kartik Kalaignanam, and Tracey A. Swartz (2017), "Does it Pay to Recall your Product Early? An Empirical Investigation in the Automobile Industry," *Journal of Marketing*, 81 (3), 111–29.

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.

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.

Hecht, Christopher, Jan Figgener, and Dirk Uwe Sauer (2021), "Predicting Electric Vehicle Charging Station Availability Using Ensemble Machine Learning," *Energies*, 14 (23), 7834.

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.

Kim, Byungyeon, Oded Koenigsberg, and Elie Ofek (2022), "I Don't 'Recall': The Decision to Delay Innovation Launch to Avoid Costly Product Failure," *Management Science*, 68 (12), 8889–8908.

Liu, Yan and Venkatesh Shankar (2015), "The Dynamic Impact of Product-Harm Crises on Brand Preference and Advertising Effectiveness: An Empirical Analysis of the Automobile Industry," *Management Science*, 61 (10), 2514–35.

Liu, Yan, Venkatesh Shankar, and Wonjoo Yun (2017), "Crisis Management Strategies and the Long-Term Effects of Product Recalls on Firm Value," *Journal of Marketing*, 81 (5), 30–48.

Maione, Bruno Fernandes, Paulo Carlos Kaminski, and Emilio Carlos Baraldi (2023), "The automotive recall data search and its analysis applying machine learning," *Production*, 33, e20220117.

Malec, Andrew M., Patricia K. Smith, and Anson E. Smuts (2021), "Recall and Vehicle Characteristics Associated with Vehicle Repair Rates: Review of Industrial Organization," *Review of Industrial Organization*, 59 (1), 37–55.

Murphy, Finbarr, Fabian Pütz, Martin Mullins, Torsten Rohlfs, Dennis Wrana, and Michael Biermann (2019), "The impact of autonomous vehicle technologies on product recall risk," *International Journal of Production Research*.

Naresh, Vankamamidi S., Guduru V. N. S. R. Ratnakara Rao, and D. V. N. Prabhakar (n.d.), "Predictive machine learning in optimizing the performance of electric vehicle batteries: Techniques, challenges, and solutions," *WIREs Data Mining and Knowledge Discovery*, n/a (n/a), e1539.

Ni, John and Xiaowen Huang (2018), "Discovery-to-Recall in the Automotive Industry: A Problem-Solving Perspective on Investigation of Quality Failures," *Journal of Supply Chain Management*, 54 (2), 71–95.

Pancik, J., A. Vemola, R. Kledus, M. Semela, and A. Bradac (2018), "Auto recalls and software quality in the automotive sector," *EAI Endorsed Transactions on Scalable Information Systems*, 5 (17).

Ramya, V. and Kirupa Ganapathy (2022), "Evaluation of Vehicle Quality Performance using Random forest in Comparison with KNN to measure the Accuracy, Recall, and Precision," in *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, 550–55.

Shah, Rachna, George P. Ball, and Serguei Netessine (2017), "Plant Operations and Product Recalls in the Automotive Industry: An Empirical Investigation," *Management Science*, 63 (8), 2439–59.

Sheikh, Shehzar Shahzad, Mahnoor Anjum, Muhammad Abdullah Khan, Syed Ali Hassan, Hassan Abdullah Khalid, Adel Gastli, and Lazhar Ben-Brahim (2020), "A Battery Health Monitoring Method Using Machine Learning: A Data-Driven Approach," *Energies*, 13 (14), 3658.

Singh, Khimendra and Rajdeep Grewal (2023), "Lobbying and Product Recalls: A Study of the U.S. Automobile Industry," *Journal of Marketing Research*, 60 (4), 728–49.

Yongqin, Feng Xiaorui, Zhang Le Liu, and Wenzhao Li (2022), "Analysis on the Trends and Characteristics of Vehicle Recalls in the United States," Atlantis Press, 1–5.

Zhang, Xuan, Shuo Niu, Da Zhang, G. Alan Wang, and Weiguo Fan (2015), "Predicting Vehicle Recalls with User-Generated Contents: A Text Mining Approach," in

Intelligence and Security Informatics, M. Chau, G. A. Wang, and H. Chen, eds., Cham: Springer International Publishing, 41–50.

Zhu, Meixin, Hao (Frank) Yang, Chenxi Liu, Ziyuan Pu, and Yinhai Wang (2022), "Real-time crash identification using connected electric vehicle operation data," *Accident Analysis & Prevention*, 173, 106708.

I hereby declare that I have written the enclosed Bachelors thesis myself and that I have not used any outside help that is not apparent from the information I have provided. I also assure that this thesis or parts thereof have not been submitted by myself or by others as a performance record elsewhere. Literal or analogous adoptions from other writings and publications in printed or electronic form are marked. All secondary literature and other sources are identified and listed in the bibliography. The same applies to graphical representations and images as well as to all internet sources and answers generated by AI-based applications. I further agree that my work may be sent and stored anonymously in electronic form for the purpose of plagiarism checking. I am aware that correction of the work may be waived if this declaration is not given.

Mannheim, June 17, 2024