UNIVERSITY OF SOUTHAMPTON

FACULTY OF BUSINESS, LAW AND ART

Southampton Business School

Machine Learning and Econometrics Models for Behavioural Decision-Making

by

Olanrewaju Orimoloye

Thesis for the degree of Doctor of Philosophy

April 2021

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF BUSINESS, LAW AND ART

SOUTHAMPTON BUSINESS SCHOOL

Doctor of Philosophy

Machine Learning and Econometrics Models for Behavioural Decision-Making

By Olanrewaju Orimoloye

Machine Learning (ML) and Econometrics models are a powerful tool for developing and testing theories by way of prediction, causal explanation and description. In many businesses, the priority and focus are; do I predict, or do I explain? Nearly all decisions are based on predictions and causal explanation, whether more intuitive or deliberative. This thesis, which is divided into three papers, explores the use of prediction methods and causal explanation, with applications to financial markets and marketing/e-commerce.

In paper one, we compared the accuracy of deep and shallow architectures by predicting thirty-four different stock indices across different time horizons (daily, hourly, minute and tick level) using financial market data. We contribute to the ML literature by exploring the degree to which is possible to predict stock price indices across different time horizon and markets.

Paper two explore the use of behavioural data from a large retail organisation to understand cross device browsing behaviour. These allows us to fill an important gap that exists in marketing/e-commerce literature, about which platform offers higher conversion rates and how online consumer browsing and buying behaviour differs among these platforms. Despite conversion rate across devices, a large proportion of consumers leave items in their shopping cart without completing a purchase in a session. This is referred to as cart abandonment. Industry report shows the rate of cart abandonment across all sectors of 75.6%.

In paper 3, we developed a unified framework using a recursive bivariate probit (RBP) model to explain the differences in online shopping cart abandonment across mobile and non-mobile devices. To our knowledge, extant research has not examined online shopping cart abandonment across mobile and non-mobile devices with field data or e-commerce click stream data. Our framework used features such as when shoppers have high basket values, browsing in the evening, if reading reviews on mobile vs non-mobile channel and numbers of attempted credit card failures on mobile vs non-mobile channel to understand device differences.

In summary, this thesis offers insights into decision-making using prediction methods at the algorithmic and individual level.

Table of Contents

List of Tables

List of Figures

Academic Thesis: Declaration of Authorship

I **Olanrewaju Orimoloye** declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Machine Learning and Econometrics Models for Behavioural Decision Making

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed:

Date: 25th April, 2021

Acknowledgements

I'd like to thank my amazing and wonderful wife – Shana and daughters (Lara & Naomi) for their patience and continual support throughout my studies. I'd also like to thank my mum and siblings for all their encouragement.

I am massively grateful to my supervisors Tiejun Ma, Ming-Chien Sung, Johnnie Johnson, for their invaluable advice and time during my PhD. This work couldn't have been done without their unwavering support, encouragement and feedback.

I have been incredibly lucky to have the opportunity to pursue a PhD and I have given it my all.

Chapter 1 Introduction

1.1.1 ML Prediction - Financial Market

Accurately predicting changes in stock price indices is potentially profitable but also a difficult challenge due to the high degree of uncertainty involved. Proponents of the efficient market hypothesis (EMH) argue that it is impossible to generate abnormal returns through 'more informed' investment decisions (Titan, 2015). However, the accelerated use of advanced algorithms has led to the identification of opportunities to trade profitably on model predictions (Chang et al., 2009; de Oliveira et al., 2013; Huang et al., 2008; Schumaker & Chen, 2009). Many techniques have been developed which can aid with this prediction task. For example, ML techniques have been employed due to their ability to handle non-linear data (Krauss et al., 2016). The most widely used ML models, such as ANN and SVM (Zhang & Wu, 2009; Chang et al., 2009; Mittermayer, 2004; Huang et al., 2005; Hsu et al. , 2016) however suffer from limited architectural depth, making it difficult to generalize on the data thereby providing inefficient and biased estimators. Deep feedforward Neural Networks (DNN), which are increasingly seen as an ML tool, have had important empirical successes in a number of traditional AI applications (Bengio et al., 2009).

Krauss et al. (2016) were the first to compare DNN, gradient-boosted trees and random forests with several ensemble models in the context of statistical arbitrage. Each model was trained on lagged returns of all stocks in the S&P 500 financial market, after elimination of survivor bias. They used sliding window methods to forecast one-day-ahead trading signals. Using out of sample data, they found that the ensemble algorithm was able to generate 0.45 percent return on a daily basis. However, when they compared the predictive accuracy of the three algorithms, random forests outperformed DNN and gradient-boosted trees. They noted that careful hyper-parameter tuning of DNN could have improved its performance. Fischer et al. (2017) applied a subset of Recurrent Neural Networks (i.e. Long Short-term Memory, LSTM, networks) to financial time series predictions using a single stock market index. They found that LSTM networks outperform memory-free classification methods, i.e., random forest (RAF) and a logistic regression classifier (LOG). In addition, Eunsuk et. al. (2017) constructed a DNN using stocks returns from the KOSPI market (the major stock market in South Korea) to examine the effect of three unsupervised feature extraction methods (principal component analysis, autoencoder, and the restricted Boltzmann machine) on the network's overall ability to predict future market behaviour. They found that DNN can extract additional information from the residuals of the autoregressive model to improve prediction performance.

A review undertaken by Hsu et al. (2016) indicated that shallow architectures, such as SVM and one-layer NN, are currently the main methods employed for predicting movements in stock price indices. As indicated above, only a handful of studies have employed deep architectures. This may be linked to the fact that DNN in general can handle different classes of structured and unstructured datasets better (used for images, text and time series data: see LeCun et al., 2015). In addition, support vector machines (SVM) and/or one-layer NN may be the most widely used methods for financial market forecasting, because of their ability to recognize patterns in nonlinear, dynamic time series data (Chang et al., 2009; Lee, 2009; Zbikowski, 2015).

Hsu et al. (2016)'s extensive survey covered 28 markets, including 12 in the US and seven in Taiwan. The most popular prediction method in these papers was ANN while only four studies considered both ANN and SVM. One of the interesting findings of the Hsu et al. (2016) survey was that very few studies evaluate prediction models in a dynamic fashion i.e. with a sliding window approach. Rather, the prevailing approach, used in 25 out of 28 previous studies, is to split a financial time series into a training and a hold-out test set. We refer to this approach as ‘static’ because it uses the same prediction model throughout the whole testing period, without updating.

To increase predictive accuracy using SVM and/or one-layer NN, some authors have suggested the use of more dimensions to augment existing stock price data (Pan et al., 2005; Kara et al., 2011). Pan et al., (2005) presented a computational approach for predicting the Australian stock market index (AORD), using multi-layer feed-forward neural networks from the AORD time series data and various interrelated markets. They noted that effectively selecting the data that is fed into a NN can improve the predictive accuracy. They showed that by including additional dimensions to augment existing datasets, they could produce an 80% prediction accuracy. Others have proposed a multi-stage approach to financial market forecasting, by first selecting an optimal feature extraction before developing ensemble models (Huang et al., 2008).

Our review of the stock price index prediction literature has highlighted four key gaps. First, the effectiveness of DNN ML methods for predicting stock price indices has not been compared to the predictive performance of existing methods which employ a shallower architecture across several major financial markets. Second, a holistic performance evaluation of two of the mostly used activation functions – RELU & TANH for predicting stock price indices has not been undertaken. In developing a deep feedforward NN, one needs to select the most appropriate activation function for the data. Much of the theoretical literature suggests that RELU (Bengio et al., 2007) greatly accelerates the convergence of stochastic gradient descent (SGD) compared with other functions (Krizhevsky et al., 2012). Unfortunately, RELU units can be fragile during training and can ‘die’ or vanish (He et al., 2016) when using the backpropagation algorithm. Jarett et al. (2009), working with the Caltech-101 dataset (a dataset of digital images available online for public use for a computer vision task) found that the nonlinear TANH activation function worked particularly well with their type of contrast normalization, followed by local average pooling. It is likely that different data will respond particularly well to the use of either the RELU or TANH activation function because RELU zeroes out negative values while TANH does not. Third, a comparison of the accuracy of a range of ML methods when predicting stock prices indices across a variety of time horizons (i.e. daily, hourly, minute and tick level) has not been undertaken. Most literature exploring the degree to which is possible to predict stock price indices has focused on one specific time horizon across one or two markets (see the survey in Hsu et al., 2016). This prevents firm conclusions being reached concerning the most appropriate methods for predicting a wide variety of stock indices (which may operate under different market conditions) across different time horizons. Fourth, no comprehensive comparative study of the predictive accuracy of DNN and alternative ML methods for predicting stock price indices in developed vs. emerging markets has been conducted. DNNs have shown some promising results using tick data in some emerging (cf. developing) markets, but SVM has been shown to produce more accurate predictions in other emerging markets (Alexandre, Pedro & Sabino, Sarah & Albuquerque, Pedro, 2015). Therefore our first paper, "**Comparing the effectiveness of deep feedforward neural networks and shallow architectures for predicting stock price indices**", helps fill these important gaps in the literature.

We have focused on predicting national stock indices because these indices were used in the majority of previous ML studies that predict direction of price changes (e.g. Bodyanskiy & Popov, 2006; de Oliveira et al., 2013; Huang et al., 2008; Huang et al., 2005; Pan, Tilakaratne, & Yearwood, 2005; Qian & Rasheed, 2007). We included 34 financial indices measured over six years from 32 countries to cover both developed and emerging markets, since one of the aims of the study is to compare the performance of DNN in emerging vs. developed markets. In addition, to ensure comparability, we wanted to use the same period for each market. This restricted the sample period because the availability of intraday data was limited and, for many markets, was only available from 2008 onwards. This required us to choose the data period where intraday data was available for most markets between 1 Feb 2008 - 19 Feb 2014 (6-year period with 1500 trading days) with the exception of the Brazilian market for which we only had data for 4 years between 1 Feb 2010 – 19 Feb 14. The raw data contains prices for each market at tick level. We were then able to transform/expand the tick-level data into minute, hourly and daily level data across multiple markets to explore the sensitivity of DNN algorithms to various time horizons and to different degrees of financial market complexity.

Our results showed that when employing daily and hourly data, shallow architectures (SVM and one-layer NN) produce more accurate stock index predictions than DNN. These results are in line with earlier studies which showed shallow architectures produce better predictions when the dataset is small. It has been suggested that this arises because the data does not have a complex structure (Bengio et al., 2007).

As expected, we found that the accuracy of stock index price predictions using a DNN model were significantly better than that, achievable using other commonly employed ML techniques (SVM and one-layer NN) when the data size increased significantly (i.e., when using minute level data). However, we cannot accept the hypothesis that predictive accuracy is higher using DNN (cf. SVM and one-layer NN) for all intraday time horizons, because, when using tick data (over 10 million cf. 800,000 observations for minute level data)), the predictive accuracy of DNN models using both types of activation function were no better than that of SVM. This may appear surprising, as previous studies have suggested that the performance of DNN improves as data size increases. As part of our paper peer review process with the Journal and before publishing paper 1, we compared the latest ML techniques, such Recurrent Neural Networks (RNN), with DNN. Our results still hold concerning DNN’s accuracy.

In addition, Cho et al. (2016) found that an increase in the size of training data for DNN led to improvements in predictive accuracy up to a certain point (an increase from 5 to 50 training images), after which accuracy did not change substantially, regardless of the training size. This might explain the behaviour we observed using this data.

Our results showed that the RELU activation function produces more accurate stock price index predictions than TANH across all time horizons. According to Krizhevsky et al., (2012), the biggest advantage of RELU is the non-saturation of its gradient, which greatly accelerates the convergence of the stochastic gradient descent compared to the TANH functions. Another nice property is that, compared to TANH neurons that involve expensive operations, the RELU can be implemented by simply thresholding a matrix of activations at zero. All independent variables used in our analysis had values of $x > 0$ which solved the issue of vanishing gradients - unlike TANH which produced dead neurons during computation.

Our results do not fully support the hypothesis that the predictive accuracy of DNN is generally greater for emerging markets than developed markets. In fact, there was no significant difference in predictive accuracy in emerging markets when using DNN and SVM for daily, hourly, and minute level data. However, we did observe higher predictive accuracy when using DNN (cf. SVM) to forecast whether a stock price index rises or falls using tick level data for three specific emerging markets: Thailand, Malaysia and the Czech Republic. Sutsarun et al. (2014) studied the impact of coups d'état in Thailand in 2006 and examined the effect on both short-run and long-run dynamics return, volatility, liquidity, and liquidity risk of returns on the Stock Exchange of Thailand index over the period of 1 January 1996 to 31 December 2011. They used tick data in their analysis and found that the immediate reaction to the coup was more evident in the stock market with a reduction in stock return. This period overlaps with the data used for our analysis and could explain why DNN was able to capture the complexity better than shallow architecture in this market. Similarly, it has been found that

Central and Eastern European countries (CEEC), such as the Czech Republic, suffer from the spill-over effect transmitted from major stock market turmoil in the USA and China (Deltuvaitė Vilma, 2016). The collapse of Lehman Brothers bank in United States in 2008 was the most significant shock transmitted to stock markets. Their empirical results also suggest that the transmission of other systemic shocks (e.g. the Middle East financial markets crash (May 2006), the Greek debt crisis (April 23, 2010), Portugal's debt crisis (May 16, 2011)) to some of the CEECs countries was also observed. This also overlaps with the data we observed for this market and could explain why shallow architectures were not able to match the performance of DNN. In Malaysia, Leow and Evelite (2015) examined the presence of a political cycle in Malaysia stock market returns and volatilities between February 1982 and April 2012. They claimed that Malaysia stock market tends to overreact to unexpected political events, such as the removal of Deputy Prime Minister and the resignation of Prime Minister. Their study further showed that the presence of a political cycle in Malaysia stock market volatilities is statistically significant, which indicates that investors use asymmetric treatments for the election information and the government policy. Using DNN in such markets with a RELU activation function was in turn able to capture the sudden change in market dynamics.

1.2.1 Causal Inference - Marketing/e-Commerce

Marketing managers, retail practitioners, and advertisers alike are facing a problem of low browsing-to-buying conversion rates. The vast majority of online browsing does not convert to a sale during that online session. An immediate online conversion is more desirable to retailers as it precludes channel-switching and provider-switching. Understanding customer browsing behaviour on different devices, including their smartphone, tablet, or personal computer (PC, which includes both laptops and desktops) is important for scholars and marketers alike. To address low conversion rates, managers seek knowledge about cross-platform online consumer behaviour. This managerial problem concerns conversion rates among device types in electronic commerce (hence, e-commerce), which broadly refers to online shopping via , such as PCs. It also concerns a more specific form of e-commerce -mobile commerce (hence, m-commerce), which informally refers to online shopping from consumers' mobile devices, such as smartphones and tablets such as iPads. Managers may ask questions, such as which platform offers higher conversion rates and how online consumer browsing and buying behaviours differ among these platforms.

In addition to being a contemporary managerial problem, this topic has a global economic significance as seen by industry trends. For example, eMarketer estimated that US consumers will spend $709.78 billion on ecommerce in 2020, an annual increase of 18% (eMarketer, 2020). According to retail analysts at GlobalData, over half of UK consumers shop online, with an increase of 29.6% forecasted between 2019 and 2024. In the most recent survey by Capgemini IMRG eRetail, a quarterly sales index using data from 210 UK retailers, over 40% of online sales were made via phones in the fourth quarter of 2018/2019.

Consistent with the IMRG industry report, some scholars found that mobile devices are the medium of choice when shopping online (de Haan et al., 2018; Xu, Kaiquan et al., 2017). Yet, a preference for mobile devices when shopping does not necessarily mean they will be used to make the most purchases. Few scholars have looked at the roles of mobile devices and purchase patterns (Kannan and Li, 2017; Shankar et al., 2016; Verhoef et al., 2017).

Despite the rising importance of m-commerce, it is worth understanding shopping from a PC because PCs may have a higher conversion rate compared to mobile platforms. There may be a conversion gap in e-commerce, which is a significant discrepancy in browsing vs. purchasing on one e-commerce platform compared to another e-commerce platform. Extant literature in this domain falls into three main categories: m-commerce, cross-device consideration in e-commerce, and electronic shopping carts.

- M-commerce refers to a subset of e-commerce. Formally defined, mobile commerce is a phenomenon that meets all the following five characteristics (Balasubramanian, Peterson, and Jarvenpaa, 2002). One, it involves communication (one-way or interactive) among humans or with humans and a device. Two, the ability to communicate does not require being at a fixed physical location. Three, the ability to communicate must be able to be maintained continuously during substantial physical movement between locations (e.g., Luo et al., 2014; Fong et al., 2015; Andrews et al., 2016; Dube et al., 2017; Li et al., 2017). Four, communication signals must be carried by electromagnetic waves. Five, there is economic benefit from the communication.

- Cross-device consideration: Studies, such as Xu et al. (2017) and de Haan et al. (2018), fall under this category. Xu et al. (2017) were the first to investigate the cross-device impact on conversion rates. Specifically, the authors examined any complementary and substitution impact of the tablet channel on the smartphone and PC channels. They leveraged a dataset from Alibaba—the largest e-commerce firm in the world—and exploited a natural experiment via the iPad app introduction. Their results showed that users' adoption of tablets enhanced the overall growth of Alibaba's e-commerce market, with an annual increase of approximately US $923.5 million. Their approach is limited to cross-device browsing "as instances where users browse on two different devices within a one-hour time window" (p. 1486). De Haan et al. (2018) was only the second paper to analyse browsing patterns across three main devices; PCs, smartphones and tablets. They investigate the role of device switching on online purchasing and find that the increased penetration of mobile devices has a significant impact on customers' online shopping behaviour, with customers frequently switching between mobile and fixed devices on the path to purchase. The authors examined device switching by analyzing clickstream data from a large e-tailer and applying propensity score matching to account for self-selection in device switching. De Haan et al. (2018) also found that when customers switch from a mobile device, such as a smartphone, to a stationary device, such as a desktop, their conversion rate is significantly higher.

However, there are opportunities to build on this work. Specifically, since then, mobile devices have become more prevalent in the life of many consumers. Mobile applications have also improved, which might lead to changes in previous research findings. Furthermore, they focused on purchase completion as a binary outcome. Building on Xu et al. (2017) and de Haan et al. (2018), our work explored the nature of cross-device browsing and offered new insights concerning the impact of cross-device browsing on purchase conversion and frequency. In particular, our research considered session level pattern browsing and conversion among three different devices, PCs, smartphones and tablets. This represents a more comprehensive approach to examining cross-device browsing. Last, we went a step beyond considering purchase completion as a "zero-one" variable, and investigated frequency of purchase completion (or buying) within a session, to allow for the possibility of a consumer completing multiple orders within the same shopping session.

- The third existing relevant research stream pertains to consumer online shopping cart use and abandonment, with abandonment representing a break in the link between placing an item in an e-cart and purchasing. Drivers of e-cart use that were identified included utilitarian motivations, such as completing a current purchase, organization of items of interest and research/information search, as well as hedonic motivations, such as entertainment (Close and Kukar-Kinney, 2010). The fact that not all motivations are goal-directed towards purchasing can partially explain the low conversion rate from shopping to buying. Kukar-Kinney and Close's work (2010) went a step further by studying the determinants of e-cart abandonment specifically and identified drivers, such as frustration with the shopping process, loading time, and total cost of the order, which contribute to the break in the conversion from online shopping to buying. However, this early work was not specific to m-commerce, nor does it examine break in conversion across different device types. Huang, Korfiatis and Chang (2017) extended the work by Close and Kukar-Kinney (2010) and Kukar-Kinney and Close (2010) to bring in the concepts related to hesitation on the mobile platform. They examined consumers' attitudes toward mobile shopping including the roles of conflicts, hesitation, and ambivalence in consumers' mobile shopping cart abandonment. While the context is mobile, the investigation lumps all mobile device together without specifying the focus (e.g., tablet or smartphone).

In particular, we considered session-level pattern browsing and conversion among PCs, smartphones and tablets. This represents a more comprehensive approach to examining cross-device browsing than has been done in the past. Last, we went a step beyond considering purchase completion as a zero-one variable, and also investigated frequency of purchase completion (or buying) within a session to allow for the possibility of a consumer completing multiple orders within the same shopping session. Our second paper, "**A cross-device examination of conversion rates from online browsing to online buying: the online path to purchasing**" made novel contributions in three ways. First, this was the first study to empirically study cross-device e-commerce along with different touch points across various stages of the online path to purchase. Second, we expanded the operationalization of cross-platform e-commerce to include PCs, tablets and smartphones. Third, we expanded knowledge on electronic shopping cart use by being the first to examine the effect of existing items in a consumer's cart from past shopping sessions on purchase frequency.

We used individual-level clickstream data from a large multinational online retailer to develop an empirical model to explain consumers' browsing and purchasing behaviour across different device types. The online retailer is a British sport brand, clothing, footwear and home products retailer with a strong multi-national presence. For the study, we used observations from customers who engaged in at least two sessions to determine how the previous session influenced the current session. The online retailer defined a session as one continuous period during which the customer is active on the website. An online shopping session starts when the customer enters the online retailer's website and ends when the customer actively leaves the website or when the customer is inactive for 30 minutes or more.

We used data from registered customers because registration is necessary to complete the purchase and it is not possible to capture device switching by unregistered users. We used sessions that belong to the same path to purchase as defined by Li and Kannan (2014); that is, as multiple touch points that a customer makes before conversion. If, for example, a customer uses a smartphone in one session but does not convert and returns in the next session using a PC and eventually converts, the two sessions were considered to belong to the same path to purchase. Our data contained device unique IDs, which allowed us to track and link a consumer identifier to the devices used on the website during each visit. The final sample included 55,500 unique customers who engaged in 958,859 usable sessions. For each session, the data included information regarding which device was used to visit the website. The dependent variable of interest was the frequency of purchase completion within a single shopping session, which is novel compared to past work that only captures if a sale was made or not. Some researchers have investigated purchase quantities (Boatwright et al., 2003), but the factors that impact frequency of purchasing per session across multiple device types have not been explored.

Endogeneity in marketing models can lead to biased coefficient estimates (Germann, Ebbes, & Grewal, 2015; Papies, Ebbes, & Van Heerde, 2017). As consumers were not randomly assigned to the treatment (device type) but self-selected the treatment, different device users could vary systematically across the different device type groups. To account for this potential self-selection bias (Garnefeld et al., 2013), we employed propensity score matching (PSM) and created an artificial control group. First, in a binary logistic regression, we calculated each customer's propensity to use a particular device (PC vs mobile - including Smartphone/Tablet) to purchase an item. Second, the matching procedure linked each customer in the treatment condition with a statistical twin from the control group who did not purchase using a particular device but statistically had the same propensity to. With a caliper matching procedure, we matched each treatment case to its nearest neighbour only if two propensity scores fell within a pre-set tolerance zone (Wangenheim and Bay, 2007). Limiting the scores to differ by a maximum of .001 – much below the recommended tolerance zone of .008 (Silverman, 1986) – we matched 150,228 customers from the treatment customers. Third, following Bommaraju and Hohenberg (2018), we computed the standardized differences in means before and after matching.

Our empirical model needs to account for the infrequent behaviours of consumer buying process in a session. Data with abundant zeros are common in studies that count the occurrence of online purchases. These types of data (i.e., count data) have values that are usually non-negative with a lower bound zero and typically have excessive zeros and over-dispersion (i.e., more variability than expected). To model the frequency of orders completed per online shopping session, we considered the Poisson regression model, the negative binomial regression model, and the zero-inflated negative binomial. To determine which of the models is the most appropriate model based on fit, we used the minimum AIC - Akaike information criterion (Bozdogan 2000). The usage of AIC is in line with existing literature (de Haan et al., 2018). Based on model fit statistics, the negative binomial (NB) was selected.

Our analyses showed that customers who shopped using tablet and who had visited clearance pages were more likely to complete an order in a shopping session, followed by those who shopped using PC. Smartphone was the least used device for conversion. Our findings also revealed that customers with items in their e-cart from a previous session using PC were more likely to complete an order in a current session compared to those using other devices (i.e., smartphones, tablets). Customers who read reviews of items while browsing on PC were more likely to complete an order than when other devices were used. Our analysis also considered device switching during the purchase process. When the device used during the previous session was a PC, we found that the conversion rate was higher for PC (6%). When the device used during the previous session was a tablet, we found that the conversion rate was higher for PC (5%). However, when the device used in the previous session was a smartphone, the conversion rate for smartphone was higher (8%). The main theoretical contribution is in offering a way to extend the theory of the path to purchase by taking an e-commerce/m-commerce focus with a proposed online path to purchase. Analyses with big data helped us generate a way to explain online consumer behaviour with actual behaviour, and the stages that consumers take in their online path to purchase, as well as the device type considerations.

To a lesser extent, our findings can have implications for the theory of affordances (Gibson 1977). Applied here, technology affordance is what a consumer with a purpose is afforded to do with technology, and consumers have functions and features whether or not consumers recognize or use them (Majchrzak & Markus, 2012). Thus, we suggest that consumers have a technology affordance by use of their mobile device which helps them fulfil their purpose to browse, purchase or both. Our findings provide empirical evidence of access and locational affordances of the channel, particularly smartphone, for browsing anytime and anywhere. Despite the rate of traffic and conversion across smartphones, tablets and PCs, the rate of leaving items in a device-shopping cart without completing a purchase was alarming to both marketers and academics. This behaviour is referred to as cart abandonment, which is the theme of our third paper. An industry report by Stephan Serrano (2020), an expert on ecommerce personalization tools, showed a cart abandonment rate across all sectors of 75.6%. There are some indications that the cart abandonment rate when using a mobile (cf. non-mobile) channel may be higher (Kibo, 2016) resulting in higher economic losses for the retailer.

Although online shopping cart abandonment has been studied, the differing factors affecting mobile and non-mobile shoppers' decisions to abandon their cart remain relatively unexplored. The causes of mobile shopping cart abandonment may not be the same as those in the non-mobile context because mobile purchasing has unique limitations and merits. For example, the small screen makes the device light and portable, but service providers may limit information search flexibility. Specifically, Ghose, Goldfarb, and Han (2013) found that the small screen size of mobile phones amplifies ranking effects, as users incur a higher cognitive load from the information chunking. Furthermore, small screen sizes can negatively influence navigation and inputting capabilities (Chae & Kim, 2004), making it hard to locate information pertinent for making online shopping decisions. According to Xu et al. (2017), the PC channel is characterized by constrained access capabilities, as it restricts Internet usage and access to places that have suitable hardware and Internet connections (Bang et al., 2013). On the other hand, ubiquitous Internet access offered by the mobile channel overcomes this limitation. This supports time critical activities, providing instantaneous information access, and facilitating immediate transactions (Jung et al., 2014; Venkatesh, 2003). The ubiquitous nature of the mobile channel also generates more travel-related discretionary time for commuters (Ghose & Han, 2011) who tend to consume mobile video content to pass the time, manage solitude, and disengage from others (O'Hara et al., 2007). Despite these benefits, limitations stemming from the processing capability and storage capacity of mobile devices, along with bandwidth constraints of the mobile network, can degrade the access quality to Internet services (Napoli and Obar, 2014).

An empirical study conducted by Xu et al. (2017) showed a tendency to switch from mobile to non-mobile largely during the hours users are at home. They also showed an increase in browsing activities on mobile phones during commuting hours, indicating that the complementary impact of mobile adoption works by spurring users to increase their browsing activities on their mobile phones while on the move. Past the evening commute hours, the share of browsing frequency attributable to mobile (vs. non-mobile) devices falls, suggesting that users tend to switch to non-mobile devices for browsing when they are at home.

There are important reasons underscoring the need to study what motivates the differences in the nature and degree of cart abandonment between mobile and non-mobile channels. Previous work has investigated the motivations for online shopping cart use (Close & Kukar-Kinney, 2010; Close, Kukar-Kinney & Benusa, 2012) and the determinants of online shopping cart abandonment (Albrecht, Hattula & Lehmann, 2017; Huang, Korfiatis & Chang, 2018) separately. However, there is a need for research to employ a model that jointly examines online shopping devices and cart abandonment whilst examining the nature of the differences in online shopping cart abandonment across mobile and non-mobile devices. One of the reasons for this is the simultaneous nature of the cart abandonment process. Before a consumer abandons their online cart, their journey usually starts by deciding whether to use a mobile or non-mobile device. Consequently, formulating this as a two-choice process is appealing.

To our knowledge, extant research has not examined online shopping cart abandonment across mobile and non-mobile devices with field data or e-commerce click stream data. In sum, despite the established managerial and economic importance of cart abandonment, there is a dearth of literature which reports empirical research concerning consumer behaviour related to online shopping cart abandonment and, specifically, the behavioural factors that differentially affect mobile and non-mobile cart abandonment. An important gap in the literature our paper fills is developing a unified framework that helps to understand how cart abandonment differs on mobile and non-mobile channels when shoppers have high basket values; when they browse in the evening; if reading reviews on mobile vs non-mobile channels reduces cart abandonment; and the security and privacy risk on mobile vs non-mobile channels. Specifically, we filled these gaps by examining click-stream data with a modelling approach and offering an addition to online consumer inhibitors (Howard and Sheth, 1969). These inhibitors include lack of availability, high price, financial status and time pressure. Here we extend these inhibitors to factors associated with mobile and non-mobile devices and examine the differences in online shopping cart abandonment when these two types of device are used. One key limitation to previous work was the use of survey data, which may not be as robust as using large data.

Several researchers have suggested that structuring the shopping environment to appeal to people that experience stress during shopping may reduce shopping cart abandonment. In particular, Albrecht, Hattula and Lehmann (2017) examined the relationship between consumer shopping stress and cart abandonment and found that a consumer's response to shopping stress depends on their motivational orientation. More specifically, the greater the in-store stress, the more likely task-oriented consumers are to abandon the trip without making purchases. The results of their four related studies showed that for customers with a task-oriented motivation, a monotonic relationship exists between shopping stress and purchase abandonment, consistent with their perception of stress as a threat to their purchase goal. However, for recreation-oriented customers, this followed a curvilinear, inverted U-shaped relationship: purchase abandonment first increased as levels of shopping stress rose, but then decreased at higher levels. Their results offered an alternative explanation for why people buy or not and suggested approaches to structuring the shopping environment to appeal to both types of consumers.

In-store and online shopping offer customers different experiences. Despite placing items in shopping carts, online shoppers frequently abandon them across devices, an issue that perplexes online retailers and has yet to be explained by scholars. A significant amount of past research (Close & Kukar-Kinney, 2010; Close, Kukar-Kinney & Benusa, 2012) has investigated various reasons why consumers abandon their cart, but no papers have directly compared the differences in online cart abandonment associated with mobile and non-mobile devices. Examining the differences of cart abandonment across mobile and non-mobile devices holds important managerial implications since it may help firms develop a multi-device strategy to increase shopping to buying conversion rates.

Oliver et al. (2003) employed the online survey data of 206 respondents to determine if redemption of online vs. offline coupons (in the form of "promotion codes") differs. They found that online coupons have a significant effect on the way shoppers react to cart abandonment. While offline coupons, such as instore vouchers, are seen as customer-initiated, online coupon internet shoppers are usually prompted to enter a code towards the conclusion of the checkout process. This prompting may influence shoppers' perceptions and behaviours (e.g., becoming irritated) leading to the propensity to increase shopping cart abandonment. In addition to the impact of coupons, transaction inconvenience, risk and waiting time have also been found to increase the chances that consumers will abandon their cart (Rajamma et al., 2009). Kukar-Kinney et al. (2010) employed survey data from 289 online consumers to explore the reasons for cart abandonment. They discovered that one of the key drivers for abandonment was that consumers used online carts for entertainment or as shopping research and an organizational tool, and this may induce them to buy at a later session or via another channel. One key limitation of using survey data in earlier studies is that they may not be as robust as using large data.

Xu and Huang (2015) presented a model of why and how cart abandonment occurs in the online shopping process without taking into consideration the differences that exist between device channels. An online survey of 210 people was conducted via the online shopping website of a communication company in China. The results demonstrated that cart abandonment was directly and positively related to shoppers browsing external websites for product comparisons. Another important factor they identified as a key driver of cart abandonment was the cost of the products in the cart.

Huang, Korfiatis and Chang (2018) were the first to relate cart abandonment directly to the use of a mobile device. They conducted two separate studies to understand why consumers hesitate to use mobile channels for shopping and, thus, abandon their mobile shopping carts. Their study used data from 232 responses to a survey posted for 30 days on an online forum. For the second study, they obtained a dataset from a marketing research provider of 226 US consumers. Their results specifically showed that emotional ambivalence is the reason why consumers hesitate at the checkout stage, leading to cart abandonment. While this set of extant works has uncovered some of the reasons for cart abandonment in general, little is known about online shopping cart abandonment and how that relates to the mobile and non-mobile channel employed.

Given the managerial importance of online shopping cart abandonment and the gap in the literature, the aim of this research was to contribute to the knowledge of consumers' behaviour with respect to online shopping carts and device channels. As such the third paper, "**Differences in online shopping cart abandonment across mobile and non-mobile devices**" sought to understand differences in the nature and degree of cart abandonment between mobile (Smartphone/Tablet) and non-mobile (PC) devices. In doing so, we formulated the problem as a two-choice process using a recursive bivariate probit (RBP) model: the first step was the choice of device to employ for online shopping (mobile vs non-mobile); the second concerned whether to abandon the cart. We leveraged a unique dataset from a large multinational clothing, footwear and home products retailer involving over 165,613 unique customers over 10 browsing sessions. To account for the lack of demographic information in our data, we conducted an experiment to determine if the lack of demographic variables will have any effect on our result. We selected consumers that used the same device throughout their online journey without switching devices and compared the results with our final model. We accounted for other additional control variables using active sessions on the website in a similar manner to Mallapragada et al. (2016). The online retailer defined a session as one continuous period in which the consumer is active on the website. A session starts when the consumer enters the online retailers' website and ends when the consumer actively leaves the website or when the consumer is inactive (i.e. they have not visited a new page on the retailer's website or have not clicked on a link on the website for more than 30 minutes). We were only interested in active sessions that belong to the same path to purchase or cart abandonment ("the multiple touches a customer makes before a conversion, as defined by Li and Kannan (2014)). Introducing the active session as a control variable allowed the model to take into consideration only the activities undertaken while the consumer is active on the website and disregarded any idleness or inactivity. We added a series of control dummy variables that denote the "sequence" or stage of each session in the purchase process. This was an indicator of how late a customer is in their online journey/funnel. In addition, similar to the findings of Mallapragada et al. (2016), we added other control variables, such as the number of product pages visited by the consumer after reading product reviews to control for reviews as a whole. Having an existing product in the cart before the current session might influence the likelihood of purchase, since prior research might have been done by the shoppers before the current session (de Haan et al., 2018). We controlled for consumers that had existing items in their cart before the current session.

Our results showed that: i) 1% of mobile users that abandoned their cart switched to non-mobile device during the session when they started from a mobile device. However, 5% of non-mobile users that abandoned their cart switched to mobile device during the session when they started from non-mobile device; ii) The basket value per device on mobile was similar to that on non-mobile and its effect on cart abandonment; iii) 76% of cart abandonments were accounted for by those employing a mobile channel for shopping; (iv) The cart abandonment rate was 44% higher on mobile (vs. non-mobile) devices when consumers visited the site during the evening; (v) Consumers who read positive reviews were equally likely to abandon their cart using mobile (vs. non-mobile) devices; (vi) Consumers using mobile devices were more likely to abandon their cart if they experienced a high number of failed attempts to use a credit card than when experiencing the same number of failed attempts to use a credit card on non-mobile devices.

Looking at the key findings from the RBP model, our field data showed that the time of a website visit has an impact on cart abandonment. Specifically, consumers using mobile devices in the evening (i.e., between 1800 and 2300hrs) were more likely to abandon their cart than those using non-mobile devices in the evening. Interestingly, when we tested different hours of the day, such as before work (i.e., between 0600 to 0800hrs), during lunchtime (i.e. between 1200 to 1300hrs), after work (i.e., between 1700 to 1800hrs), with device type as a moderator, their effects on cart abandonment were not significant.

Our findings did not support the earlier findings of Rob et al. (2018). Their results from analyzing click stream data of browsing behaviour of hotel website users showed an increase in abandonment during the evening using non-mobile devices. Our findings, however, support the earlier work of Close et. al. (2010) that found consumers browsed on mobile device for convenience. Mobile devices also have the good dimensions of usability and being ubiquitous as their core features, which distinguish them from non-mobile devices (Ghose et al., 2012).

We examined the impact of devices on customers' cart abandonment behaviour using empirical data. We found that there were no differences in the basket value of items on mobile (vs non-mobile device) and their effects on cart abandonments. This was also the case when we examined basket quantity for both devices. Extant literature (such as Xu et al., 2017; Jen - Hui Wang et al., 2015) suggested that, due to the flexibility, convenience of usage, portability and ubiquity of mobile devices, consumers tend to have higher basket value and quantity than when using non-mobile devices. Surprisingly, in our sample data, the average size of basket value and basket quantity per device on non-mobile were higher than mobile device. One reason for this finding in our data might be that low spenders indulge in the habits of purchasing items more frequently on their mobile device while non-mobile devices are preferred for high valued items for security purposes.

Retailers often add positive product reviews in an effort to decrease cart abandonment. However, our field data shows no evidence to suggest that consumers that read positive product reviews using their mobile device were less likely to abandon their cart than those using non-mobile devices. This contrasts with earlier findings by Grewal et al. (2019), who found that reading positive reviews on mobile devices affects consumer attitudes and increases purchase intentions. The proposed explanation is that a large number of customer reviews increases information overload and choice ambiguity, and could lead to consumers avoiding having to make a choice (Maslowska et al., 2017; Scheibehenne et al., 2010).

In addition, while it may be easy to shop and place items of interest in the cart on a mobile device, consumers appear to be more reluctant to complete purchases using a mobile device. Given the "on the go" nature of mobile usage and the sheer ease of adding items to the cart when on mobile, consumers may want to take more time to fully consider the purchase before committing to buy. Our field data indeed showed that, as attempted card failure increases, cart abandonment is higher on mobile (vs. non-mobile). The need to enter detailed payment information without making mistakes may make non-mobile devices more suitable than a mobile device.

Chapter 2 Comparing the effectiveness of deep feedforward neural networks and shallow architectures for predicting stock price indices

ABSTRACT

Many existing learning algorithms suffer from limited architectural depth and the locality of estimators, making it difficult to generalize from the test set and providing inefficient and biased estimators. Deep architectures have been shown to appropriately learn correlation structures in time series data. This paper compares the effectiveness of a deep feedforward Neural Network (DNN) and shallow architectures (e.g., Support Vector Machine (SVM) and one-layer NN) when predicting stock price indices in both developed and emerging markets. An extensive evaluation is undertaken, using daily, hourly, minute and tick level data related to thirty-four financial indices from 32 countries across six years. Our evaluation results show a considerable advantage from training deep (cf. shallow) architectures, using a rectifier linear (RELU) activation function, across all thirty-four markets when 'minute' data is used. However, the predictive performance of DNN was not significantly better than that of shallower architectures when using tick level data. This is because in some cases shallow Neural Networks are able to learn deep functions using the same number of parameters as the original DNN. Hence, it is clear that the function learned by that DNN does not really have to be deep. This result suggests that when training a DNN algorithm, the predictive accuracy peaks, regardless of training size. We also examine which activation function works best for stock price index data. Our results demonstrate that the RELU activation function performs better than TANH across all markets and time horizons when using DNN to predict stock price indices. This is due to the fact that TANH produces dead neurons during computation while RELU eliminates the vanishing gradients problems observed in TANH activation function.

Keywords: Financial time series forecasting; deep feedforward Neural Network; market efficiency; machine learning.

Comparing the effectiveness of deep feedforward neural networks and shallow architectures for predicting stock price indices

2.1 Introduction

Accurately predicting changes in stock price indices is potentially profitable but is a difficult challenge due to the high degree of uncertainty involved. Proponents of the efficient market hypothesis (EMH) argue that it is impossible to generate abnormal returns through 'more informed' investment decisions (Titan, 2015). However, the accelerated use of advanced algorithms has led to the identification of opportunities to trade profitably on model predictions (Chang, et al., 2009; de Oliveira, et al., 2013; Huang, et al., 2008; Schumaker & Chen, 2009). Many techniques have been developed which can aid this prediction task. For example, machine learning (ML) techniques have been employed due to their ability to handle non-linear data (Krauss et al. (2016). The two most widely used ML algorithms for predicting stock price indices are ANN and SVM (See Zhang and Wu., 2009; Chang, et al., 2009; Mittermayer, 2004; Huang, et al., 2005; Hsu et al. (2016).

It has been suggested that learning multiple levels of distributed representations offers considerable advantages over shallow architectures (Bengio et al., 2009) and empirical work suggests that it is difficult to train shallow architectures to be as accurate as DNN. For vision tasks, a study on deep convolutional nets suggested that deeper models are preferred under a parameter budget (Mathieu et al., 2013). Furthermore, Lecun et al., 2013 trained shallow nets on scale invariant feature transform (SIFT) to classify a large-scale ImageNet dataset and found that it was difficult to train large, high-accuracy, shallow architectures. In addition, Seide et al., (2011) show that deeper models are more accurate than shallow models in speech acoustic modelling and Abdel-Hamid et al., (2012) recorded a 5% increase in predictive accuracy in speech recognition using a deep (vs. shallow) architecture using one million labelled points on test data. Ba et al., (2014) provide some pointers into why deep feedforward NN might perform better than shallow architectures. He noted that a network with a large enough single hidden layer of sigmoid units can approximate any decision boundary.

Consequently, in attempting to predict changes in stock price indices across thirty-four markets using different time horizons, we deploy several powerful methods inspired by the latest trends in machine learning. First, we compared the effectiveness of DNN and shallow architectures (e.g., Support Vector Machine (SVM) and one-layer NN).

In developing a deep feedforward NN, one needs to select the most appropriate activation function for the data. Much of the theoretical literature suggests that RELU (Bengio et al., 2007) greatly accelerates the convergence of stochastic gradient descent (SGD) compared with other functions (Krizhevsky et al., 2012). Unfortunately, RELU units can be fragile during training and can 'die' or vanish (He et al., 2016) when using the backpropagation algorithm. Jarett et al. (2009), working with the Caltech-101 dataset (a dataset of digital images available online for public use for a computer vision task) found that the nonlinear TANH activation function worked particularly well with their type of contrast normalization, followed by local average pooling. It is likely that different data will respond particularly well to the use of either the RELU or TANH activation function because RELU zeros out negative values while TANH does not. Consequently, an important contribution of this paper is that we examine which activation function works best for stock price index data.

With large positive returns and low correlation with returns in developed financial markets, emerging financial markets provide theoretically an ideal environment for international portfolio diversification. Consequently, investors seeking high returns and good diversification often turn to these markets (Berger, Pukthuanthong and Yang, 2011). A shallow algorithm, such as a one layer-NN and other non-linear algorithms, might not be able to fully account for the complex relationships in modelling price changes in these markets because they have no level of abstraction. It is therefore difficult to capture complexities inherent in the data structure. We examined whether DNN could better capture these dynamics, thereby providing the basis of a more accurate prediction model.

Most literature exploring the degree to which is possible to predict stock price indices have focussed on one specific time horizon across one or two markets (see the survey in Hsu et al., 2016). This prevents firm conclusions being reached concerning the most appropriate methods for predicting a wide variety of stock indices (which may operate under different market conditions) across different time horizons. Consequently, to fill this important research gap, we develop a comparison of the accuracy of deep and shallow architectures for predicting thirty-four different stock indices across different time horizons (daily, hourly, minute and tick level). This analysis provides a means of comparing the performance of deep and shallow architectures under different market and prediction environments.

The remainder of this paper is organized as follows: In section 2.2, we review related work in the financial and ML literatures. In section 2.3, we develop three hypotheses, based on the literature review, concerning how different methodological factors affect predictive accuracy. In section 2.4, we outline our experimental design. In section 2.5, we present our empirical results and in section 2.6 we discuss these in the light of existing literature. We conclude in section 2.7 with a summary of the main findings, the implications the work and suggested areas for further research.

2.2 Literature review

Increasingly, the DNN approach is seen as an important ML tool. DNN have had important empirical successes in a number of traditional AI applications, such as computer vision and natural language processing (Bengio, 2009; Bengio et al., 2013). As a result, DNN are attracting much attention from the academic and industrial communities. Companies like Google, Microsoft, Apple, IBM and Baidu are investing in DNN, with the first widely distributed products being used by consumers aimed at speech recognition. DNN is also used for object recognition (Google Goggles), image and music information retrieval (Google Image Search, Google Music), as well as computational advertising (Le et al., 2012). Furthermore, a DNN building block (the restricted Boltzmann machine, or RBM) was used as a crucial part of the winning entry in a million-dollar ML competition (the Netflix competition) (Salakhutdinov et al., 2007; Toscher et al., 2009).

There have been several studies demonstrating the effectiveness of DNN methods in a variety of application domains such as image and speech processing (Huang & LeCun, 2006) but it has rarely been employed for forecasting problems. DNN have achieved very low error rates on various image databases (the MNIST database of handwritten digits, 2014) (Fukushima, 2003) and in this domain it has been reported that the learning process of DNN methods are ‘surprisingly fast’ (Hinton, Osindero, & Teh, 2006). DNN have also been shown to be more effective than conventional algorithms when employed for electroencephalographic signal processing (Ren & Wu, 2014). Much of this work was based on a paper by Deng et. al, (2013), which reported a major breakthrough in DNN speech recognition. Given the success of DNN in both image and speech processing, it is surprising that DNN has, as outlined below, rarely been used to predict stock price indices and its performance across a range of markets and forecast time horizons has not been explored.

Krauss et al. (2016) were the first to compare DNN, gradient-boosted trees and random forest with several ensemble models in the context of statistical arbitrage. Each model was trained on lagged returns of all stocks in the S&P 500 financial market, after elimination of survivor bias. They used sliding window methods to forecast one-day-ahead trading signals. Using out of sample data, they found that the ensemble algorithm was able to generate 0.45

percent return on a daily basis. However, when they compared the predictive accuracy of the three algorithms, random forest outperformed DNN and gradient-boosted trees. They noted that careful hyper-parameter tuning of DNN could have improved its performance. Fischer et. al. (2017) applied a subset of Recurrent Neural Networks (i.e. Long Short-term Memory (LSTM) networks) to financial time series predictions using a single stock market index. They found that LSTM networks outperform memory-free classification methods, i.e., random forest (RAF) and a logistic regression classifier (LOG). In addition, Eunsuk et. al. (2017) constructed a DNN using stocks returns from the KOSPI market (the major stock market in South Korea) to examine the effect of three unsupervised feature extraction methods (principal component analysis, autoencoder, and the restricted Boltzmann machine) on the network's overall ability to predict future market behaviour. They found that DNN can extract additional information from the residuals of the autoregressive model to improve prediction performance.

A review undertaken by Hsu et al. (2016) indicates that shallow architectures, such as SVM and one-layer NN, are currently the main methods employed for predicting movements in stock price indices. As indicated above, only a handful of subsequent studies have employed deep architectures. This may be linked to the fact that DNN in general can better handle different classes of structured and unstructured datasets (used for images, text and time series data: see LeCun et al., 2015). In addition, support vector machines (SVM) and/or one-layer NN may be the most widely used methods for financial market forecasting, because of their ability to recognize patterns in nonlinear, dynamic time series data (Chang et al., 2009; Lee, 2009; Zbikowski, 2015).

Hsu et al (2016)'s extensive survey covered 28 markets, of which the US accounted for 12 and Taiwan 7. The most popular prediction method in these papers was ANN and only four studies considered both ANN and SVM. One of the interesting findings of the Hsu et al (2016) survey was that very few studies evaluate prediction models in a dynamic fashion i.e. sliding window approach. Rather, the prevailing approach, used in 25 out of 28 previous studies, is to split a financial time series into a training and a hold-out test set. We refer to this approach as 'static' because it uses the same prediction model throughout the whole testing period, without updating.

To increase predictive accuracy using SVM and/or one-layer NN, some authors have suggested the use of more dimensions to augment existing stock price data (Pan et. al., 2005, Kara et al., 2011). Pan et. al., (2005) present a computational approach for predicting the Australian stock market index (AORD), using multi-layer feed-forward neural networks from the AORD time series data and various interrelated markets. They noted that effectively selecting the data that is fed into a NN can improve the predictive accuracy. They showed that by including additional dimensions to augment existing datasets, they could produce an 80% prediction accuracy. Others have proposed a multi-stage approach to financial market forecasting, by first selecting an optimal feature extraction before developing ensemble models (Huang et al., 2008).

Our review of the stock price index prediction literature has highlighted four key gaps. First, that the effectiveness of DNN ML methods for predicting stock price indices has not been compared to the predictive performance of existing methods which employ a shallower architecture across several major financial markets. Second, a holistic performance evaluation of two of the mostly used activation functions – RELU & TANH for predicting stock price indices has not been undertaken. Third, a comparison of the accuracy of a range of ML methods when predicting stock prices indices across a variety of time horizons (i.e. daily, hourly, minute and tick level) has not been undertaken. This is the case, despite the fact that there is some evidence that in other domains that whilst DNN performs best when datasets are big, there may be some limits, beyond which the predictive performance does not improve. Fourth, no comprehensive comparative study of the predictive accuracy of DNN and alternative ML methods for predicting stock price indices in developed vs. emerging markets, has been conducted. DNN have showed some promising results using tick data in some emerging (cf. developing) markets, but SVM has been shown to produce more accurate predictions in other emerging markets (Alexandre, Pedro & Sabino, Sarah & Albuquerque, Pedro, 2015). We aim to help fill these important gaps in the literature.

2.3 Hypothesis

Previous stock price index prediction studies using ML, have predominantly employed daily time horizon data (Hsu et al., 2016). This could be attributed to fact that these data are free whereas this is not the case for intraday data. For example, Cenesizoglu T. and Timmermann A. (2008) found that stock returns prediction models based on daily or longer data intervals generally underperform models predicting over a shorter time interval.

Bearing this in mind, DNN with its superior representation space compared to shallow models, should out-perform other algorithms when predicting stock indices using intraday data. This is likely to arise because DNN may effectively employ the significantly larger and more complex dataset to create greater accuracy (Abdel-Hamid et al., 2012). Cho et al. (2016) conducted experimental studies on deep learning systems to explore how much data is needed to train a medical image to achieve high accuracy with a low standard deviation. They found that an increase in the training data from 5 to 50 training images significantly improved accuracy but does not improve significantly from 100 to 200. In particular, they found that beyond a certain training set size the accuracy hardly improved. In addition, Angelidis et al., (2008) compared the performance of various models using intra-day vs. inter-day. They investigated their forecasting performance for three European equity indices and found that intra-day model clearly produced the most accurate variance forecasts.

To explore the view that DNN are more accurate than shallow models when employing larger datasets, we test the following hypothesis:

H1: Predictive accuracy is higher using DNN for intraday (i.e. minute/tick) time horizons than SVM and one-layer NN.

DNN requires activation functions to map the output of a node given an input or set of inputs. The use of Rectified Linear Unit (RELU) has become very popular in the last few years. It computes the function $f(\mathrm{x}) = \max(0, \mathrm{x})$. In other words, the activation has a threshold of zero (i.e. takes only positive numbers). Krizhevsky et al. (2012) found that employing RELU greatly accelerated the convergence of stochastic gradient descent (cf. when employing TANH functions). They argued that this is due to its linear, non-saturating form. Compared to TANH neurons that involve expensive operations (exponentials etc.), the RELU can be implemented by simply thresholding a matrix of activations at zero. Nair et al. 2010 used CIFAR- 10 data (consisting of 60000 32x32 colour images in 10 classes, with 6000 images per class) to build a four-layer convolutional network and concluded that RELU was several times faster and produced more accurate predictions than TANH. However, it is likely that different datasets will possess different features that will lead to either RELU or TANH producing more accurate predictions. To our best knowledge, a comparison of the predictive accuracy of the two most widely used activation functions (RELU or TANH) for stock price indices, has not been undertaken. Consequently, based on the findings from the other domains discussed above, we test our second hypothesis:

H2: Using the RELU activation function produces more accurate stock price index predictions than using the TANH activation function across all time horizons.

The financial economics literature suggests that informational efficiency differs between established and emerging financial markets (Griffin et al., 2010; Karemera & Ojah, 1999). In particular, there is some evidence that the EMH, which posits that asset prices reflect all relevant information (Fama, 1970 & 1991), may not hold in emerging markets because of extraneous factors (financial, economic, political, social and cultural environment). These factors can lead emerging markets to suffer greater volatility and to experience more sudden stock market declines than developed markets.

It has been found that using a Gaussian distribution to model price changes or stock market returns in emerging markets is not appropriate because the prices and returns are fat-tailed (i.e. large price changes can be expected relatively often) (see, for example, Harvey, 1995; Bekaert and Harvey, 1997). A shallow algorithm, such as a one layer-NN and other non-linear algorithms, might not be able to fully account for the complex relationships in modelling price changes in these markets because they have no level of abstraction and it is therefore difficult to capture complexities inherent in the data structure. Taken together, these factors motivate the following hypothesis:

H3: The accuracy of stock price indices predictions using a DNN is greater in emerging markets than in developed markets and is greater than that arising from employing a model with shallower architecture.

The proposed set of hypotheses, enable us to examine the extent to which various factors, such as forecasting horizon, activation functions, selected models (i.e. DNN vs. SVM vs one-layer NN), and data from emerging (vs. developed) markets influence the accuracy of financial time series forecasting. The testing of these hypotheses should be helpful in determining if DNN have a place in stock market predictions.

[Table 2.1 about here]

2.4 Experimental design

Table 2.1 summarizes the main factors we considered (forecast time horizon, activation functions and market classification) which could influence the predictability of stock price indices. Thirty-four financial time series from a wide cross section of the world's major markets were used to test our hypotheses. Our design is similar to that of Gerlein, McGinnity, Belatreche, and Coleman's (2016) study in which we performed repeated experiments to clarify the influence of several factors on the prediction performance of simple ML classifiers. In addition, we conducted robustness checks of our results by testing different levels of parameters settings and obtained the best results for each ML algorithm. This provides confidence that reliance can be placed on the conclusions regarding the comparative effectiveness of the shallower architecture ML and deeper architecture DNN methods.

In the following subsections, we discuss the factors we examined and motivate our choices of individual factor levels.

2. 4.1.1 Data, variables, and forecasting time horizon

The dataset employed in this study was obtained from TickWrite Data Inc., including time series of financial indices from around the world. We have focused on predicting national stock indices because these indices were used in the majority of previous ML studies that predict direction of price changes (e.g. Bodyanskiy & Popov, 2006; de Oliveira et al., 2013; Huang et al., 2008; Huang et al., 2005; Pan, Tilakaratne, & Yearwood, 2005; Qian & Rasheed, 2007). We included as many markets as possible to cover both developed and emerging markets, since one of the aims of the study is to compare the performance of DNN in emerging vs. developed markets. In addition, to ensure comparability, we wanted to use the same period for each market. This restricted the sample period because the availability of intraday data is limited and, for many markets, is only available from 2008 onwards. This requires us to choose the data period where intraday data is available for most markets between 1 Feb 2008 - 19 Feb 2014 (6-year period with 1500 trading days) with an exception of Brazilian market for which we only have the data for 4 years between 1 Feb 2010 – 19 Feb 14. The raw data contains prices for each market at tick level. We then are able to transform/expand the tick-level data into minute, hourly and daily level data across multiple markets to explore the sensitivity of DNN algorithms to various time horizons and to different degrees of financial market complexity.

2.4.1.2 Market classification

The term emerging markets was first used in the 1980s by Antoine van Agtmael, a World Bank economist. The classification of countries as emerging markets is carried out and reviewed on a regular basis by a range of international financial institutions. They all use different categories, methodologies and degrees of granularity. It is, therefore, not surprising that there is no agreed consensus on what constitutes an emerging market (Colm Keaney (2012)). The Financial Times Stock Exchange (FTSE), for example, categorizes emerging markets into 9 'advanced' and 13 'secondary' emerging markets, whereas Bloomberg's Morgan Stanley Capital International (MSCI) emerging market index comprises 26 countries, split into three regions. For this paper, we only count a market as emerging if it is classified by the FTSE and MSCI as emerging. In table 2.2 we list the markets which the data restrictions outlined in section 2.4.1.1 enabled us to include and we identify which of these are categorized as 'emerging' by the MSCI and the FTSE.

[Table 2.2 about here]

2.4.1.3. Forecasting methods

The most widely used ML methods discussed in the literature for financial time series forecasting are SVM and NN (Ballings, Van den Poel, Hespeels, & Gryp, 2015). One of the explanations for their popularity (cf. other advanced techniques, such as bagged or boosted decision trees), is that they are better suited to handling continuous covariates (Hsu et al., 2016). These are the type of covariates which occur frequently in financial time series.

SVMs belong to the general category of kernel methods (Shawe -Taylor et. al., 2004; Scholkopf et al., 2002), that depend on the data only through dot-products. When this is the case, the dot product can be replaced by a kernel function, n which computes a dot product in some possibly high-dimensional feature space. This has two advantages: First, the ability to generate nonlinear decision boundaries using methods designed for linear classifiers. Second, the use of kernel functions allows the user to apply a classifier to data that has no obvious fixed-dimensional vector space representation (Ben-Hur et al., 2010). Modelling using SVM involves the construction of a hyperplane as the decision surface such that the margin of separation between positive and negative examples is maximized (Xu, Zhou, & Wang, 2009). For a training set of samples, with input vectors $x_i \in R^d$ and corresponding labels $y_i \in \{+1, -1\}$, SVM learns how to classify objects into two classes.

We face a two-class problem, with label +1(close stock price index increases) and -1(close stock price index decreases). We define x as a vector with components x_i; in our case openindex (opening price of index), closingindex (closing price of index), highestindex (highest value the index reaches in the period) and lowestindex (lowest value the index reaches in the period). The notation x_i denote the ith vector in a dataset composed of n labelled examples (x_i, y_i) where y_i is the label associated with x_i. The objects x_i are called inputs. To define a linear classifier, we use the dot product between the vectors, also referred to as the inner product or scalar product. A linear classifier is based on a linear discriminant function of the form:

$$f(x) = w^T x + b \qquad (2.1)$$

The vector w is known as the weight vector, and b is called the bias. Consider the case b=0. The set of points x_i such that w^T=0 are all points that are perpendicular to w and go through the origin - a line in two dimensions, a plane in three dimensions and more generally, a hyperplane. A hyperplane divides the space into two according to the sign of the discriminant function $f(x)$. The boundary region, classified as positive and negative, is called the decision boundary of the classifier. A classifier with linear decision boundary is called a linear classifier and if the decision boundary depends on the data in a non-linear way, the classifier is said to be non-linear (Ben-Hur et al., 2010).

The parameters of SVM include the type of kernel function, the degree of kernel function (d: for a polynomial kernel; γ for a radial basis kernel) and a regularization constant c. Tables 2.3a, 2.3b and 2.3c shows the numbers of levels tested in different parameter settings for all the algorithms compared in this paper. To determine the parameters efficiently, we followed the approach used by Patel et al. (2015). Namely, we test five levels on d, ten levels of γ and 4 to 5 levels of c in the parameter setting experiments. These parameters and the levels which are tested, are summarized in Table 2.3a. For one stock index, these settings of parameters yield a total of 20 and 40 treatments for the SVMs employing polynomial and radial basis kernel functions, respectively. We selected the best polynomial kernel SVM model and the best radial basis SVM model models (using parameter combinations that resulted in the best average of the training and holdout prediction performances) for comparison with NN and DNN.

[Table 2.3a about here]

NN is a nonlinear prediction method and is one of the most popular algorithms for stock index prediction (Kara et. al., 2011). NNs are often referred to as feedforward NNs or multilayer perceptron (MLP). Inspired by the functioning of biological NNs, NNs are a dense network of inter-connected neurons which get activated based on inputs. A one-layer feed-forward NN is employed in our study.

Inputs for the network are four simple variables (Openindex/Closingindex/Highestindex/Lowestindex) which are represented by four neurons in the input layer. We did not to use technical indicators since previous studies question their usefulness (Lesmond et al., 2004) and an extensive simulation by Hsu et al. (2016) found that technical indicators offer little advantage over simple covariates in terms of accuracy or ROI. We employ an output layer with a single neuron, with log sigmoid as the transfer function, resulting in a continuous value output between 0 and 1. A threshold of 0.5 is used to determine whether we predict an increase (> 0.5) or decrease (<0.5) in the stock price index. We conducted robustness checks on the use of 0.5 as a threshold, but found that using 0.5 produced the most accurate predictions. In addition, this approach is in line with that employed in the majority of previous studies (e.g. Bodyanskiy & Popov, 2006; de Oliveira et al., 2013; Huang et al., 2008; Huang et al., 2005; Pan, Tilakaratne, & Year- wood, 2005; Qian & Rasheed, 2007; Hsu et al.,2016). We used tan sigmoid as the transfer function for the hidden layer since it was the setting that produced the most accurate predictions. Gradient descent with momentum was used to adjust the weights; at each epoch, weights were adjusted so that a global minimum could be reached.

The NN model parameters include the number of hidden layer neurons (n), the value of the learning rate (lr), the momentum constant (mc) and the number of epochs (ep). To determine appropriate levels in an efficient manner, ten levels of n, nine levels of mc and ten levels of ep were tested in the parameter setting experiments. Initially, the value of lr was fixed to 0.1. The parameters and the levels which were tested are summarized in Table 3b. We selected the best three NN models for comparison with DNN and SVM by selecting the three parameter combinations that resulted in the highest average prediction accuracy in the training and holdout data sets. For these top performing models, the learning rate lr was varied in the interval of [0.1, 0.2].

[Table 2.3b about here]

A DNN consists of an input layer, two or more hidden layers, and an output layer. DNNs are made up of many different functions (Dixon et al., 2016). For example, we might have four functions $f^{(1)}$,$f^{(2)}$,$f^{(3)}$ and $f^{(4)}$ all connected in a chain, to form $f(x) \approx f^{(4)}, f^{(3)}, f^{(2)} f^{(1)}(x))$. In this case, $f^{(1)}$ is called the first layer, $f^{(2)}$ the second layer, and so on; the final layer of a feedforward network being referred to as the output layer. The output layer is either a classification or regression layer, to match the output space. The overall length of the network of the chain gives the depth of the model and deep learning arises from this terminology.

NNs rely on linear functions and these have limitations when training large and complex data (Bengio et al., 2007). The appeal of DNNs (vs. NNs), therefore, stems from their ability to learn non-linear relationships.

We consider the classic feedforward architecture as shown in (2.2), where a neuron in the previous layer l is fully connected with all neurons in the subsequent layer l+1 via directed edges, each representing a certain weight. Also, each non-output layer of the net has a bias unit, serving as an activation threshold for the neurons in the subsequent layer. As such, each neuron receives a weighted combination α of the n_l outputs of the neurons in the previous layer l as an input, as follows:

$$\alpha = \sum_{i=1}^{n_l} w_i x_i + b, \tag{2.2}$$

with w_i denoting the weight of the output x_i and b the bias. The weighted combination α is transformed via an activation function f, so that the output signal $f(\alpha)$ is relayed to the neurons in the subsequent layer. Following Krizhevsky et al. (2012), we used both rectified linear units (which computes the function $\alpha = \sum_{i=1}^{n_l} w_i x_i + b$) and TANH, which constricts a real-value into the range [-1, 1]).

For the entire network, let W be the collection $\cup_{l=1}^{L-1} W_l$, with W_l denoting the weight matrix that connects layers l and l+1 for a network of L layers. Analogously, let B be the collection B= $\cup_{l=1}^{L-1} b_l$, with c_l denoting the column vector of biases of layer l. The collections of W and B fully determine the output of the entire DNN as shown in (2.3). Learning takes place by adapting these weights in order to minimize the error using the training data. In particular, the objective is to minimize some loss function $L(W,B|j)$ for each training set j. We are trying to solve a classification problem (i.e. the stock price index falls or rises), which means we can use the following cross-entropy loss function:

$$L(W,B|j)=-\sum_{y \epsilon o} (\ln \left(o_y^{(j)} \right) t_y^{(j)} + \ln (1- o_y^{(j)})(1- t_y^{(j)})) \tag{2.3}$$

where y represents the output units, o the output layers, $o_y^{(j)}$ the prediction and $t_y^{(j)}$ the actual output for example j. This loss function is minimized by stochastic gradient descent. The gradient of the loss function $L(W,B|j)$ is calculated via backpropagation. We implemented this using SAS Viya in-memory architecture, taking advantage of the cloud analytics service (CAS) for efficient and fast processing.

Since the input space is very simple (i.e. four variables), we used a straightforward architecture. The design of an NN is more of an art than a science and tuning parameters is often undertaken via computationally intensive hyper-parameter optimization routines and cross-validation. We tried multiple hidden layers (from 2 to 10). A popular rule is to set the number of neurons in the first hidden layer of a feedforward network so as to use as many neurons as there are inputs (Krizhevsky et al., 2012). We followed this approach. Our network is relatively small (4 inputs and 160 weight parameters) but deep learning allows for large-scale models with thousands of features and millions of parameters, offering significant potential for further studies with larger numbers of inputs and weight parameters. The number of levels tested in different parameter settings for DNN are summarised in table 2.3c.

[Table 2.3c about here]

2.4.2 Performance measurement

2.4.2.1 Predictive accuracy

We used accuracy to evaluate the performance of the proposed models. Computation of these evaluation measures requires us to estimate Precision and Recall. These are evaluated from True Positive (TP), False Positive (FP), True Negative (TN) and False. We define these parameters, as follows:

$$\text{Precision}_{\text{positive}} = \frac{\text{TP}}{\text{TP+FP}} \tag{2.4}$$

$$\text{Precision}_{\text{negative}} = \frac{\text{TP}}{\text{TP+FP}} \tag{2.5}$$

$$\text{Recall}_{\text{positive}} = \frac{\text{TP}}{\text{TP+FN}} \tag{2.6}$$

$$\text{Recall}_{\text{negative}} = \frac{\text{TN}}{\text{TN+FP}} \tag{2.7}$$

Precision is the weighted average of precision positive and negative, while Recall is the weighted average of recall positive and negative. Accuracy is estimated using the following:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{2.8}$$

2.4.2.2 Model evaluation

Lo (1991) and Sadique et al. (2001) conducted extensive experiments using various time horizons in different markets and concluded that stock market returns exhibit long-term memory. This suggests that a 'static' approach for predicting stock price indices (i.e. the same prediction model is used throughout the whole testing period, without updating) is the one most likely to yield higher predictive accuracy. This was confirmed by Hsu et al.'s (2016) study, which compared the static approach and a dynamic approach involving a sliding-window cross-validation performed in model training and evaluated multiple times using smaller chunks of sequentially ordered data (suggested by Lessmann et al., 2011). Hsu et al. (2016) demonstrated that the static outperformed the dynamic approach. Consequently, the static approach is employed in this study. To test the accuracy of the forecasting models, we used the standard static approach, whereby we split a financial time series chronologically into three non-overlapping sets: 50, 25 and 25 percent for training, validation and testing, respectively. The training data is used to estimate the parameters of the forecasting model and the validation set is used to tune meta-parameters (e.g., the regularization parameter, miniBatchsize, learning rate etc.) by means of empirical experimentation. For example, using daily data, of the 1500 trading days for each of the 34 financial time series in our data, we use the first 50% (750) of trading days to train the forecasting models with alternative meta-parameter settings (see Tables 2.3a,b,c) and assess their accuracy on the following 375 trading days (our validation sample). The validation sample predictions reveal the best meta-parameter setting for a given time-series. The fully-specified models with fixed optimal meta-parameters are trained and validated on the first 75% (1125) of trading days in the dataset. These models are then used to generate predictions in the test set (the remaining 375 trading days). To measure forecast accuracy, the predictions are compared to the actual values of the target variable in the test set. This allows us to compare models (e.g., SVM vs. NN vs DNN) in terms of their predictive accuracy on hold-out data. We call this approach a static evaluation, because the same model is used to forecast all observations in the test set (de Oliveira et al., 2013; Kara et al., 2011).

2.4.2.3 ANOVA (Analysis of variance) for experimental design

We used ANOVA to test for differences among the means of levels of model accuracy across four time horizons. To qualify these differences, we used Tukey's Standardized Range (HSD) test. HSD is a single-step multiple comparison procedure and statistical test. It can be used on raw data or in conjunction with an ANOVA (post-hoc analysis) to find means that are significantly different from each other. If there are two treatment levels, this analysis is equivalent to a t-test comparing two group means. The assumptions of analysis of variance (Steel and Torrie 1980) are that treatment effects are additive and experimental errors are independently random with a normal distribution that has mean zero and constant variance.

To test if our data is consistent with the null hypothesis, we calculate the F −test statistic (using the ANOVA), f_0, using the ratio of the variability between groups (S_B^2) to the variability within groups (S_W^2). The F −test statistic has an F −distribution with df_1=k-1 and df_2= n_{tot}-k, where k is the number of groups and n_{tot} is the total number of observations.

2.5. Results

2.5.1. Factors influencing the accuracy of forecasts of stock price indices.

The experimental design includes three main factors: forecast time horizon, activation functions and market classification. We obtained 544 individual prediction results: across 34 financial markets x 4 forecast time horizons x 4 model specifications (SVM, NN, DNN (with both the RELU and TANH activation functions; indicated, hereafter, as DNN (RELU) and DNN (TANH), respectively). The results enable us to compare the forecasting accuracy with different time intervals.

The null hypothesis is that the means of all groups are identical. Figure 1 displays the distribution of model accuracy across all time horizons among different model specifications. The $F-$test for the daily data for example suggests that there are differences among the accuracy of model performance i.e. the variation of model accuracy among the four models (numerator) is much larger than the variation accuracy within the four models ($F = 19.04$, $p < 0.0001$). However, this does not show if they are significantly different from each other. To examine this, we perform a Tukey's Standardized Range (HSD) test. The HSD table is usually denoted by letters to represent if the mean value are significantly different or not. If the same letter is shown, this denotes that the mean is not significantly different while a different letter shows that there is a significant difference in some of the mean values.

A number of conclusions emerge from the result presented in Table 4. For daily data of 1,500 observations, SVM performs significantly better than all the other algorithms (HSD grouping is 'A'). The second best model is NN, which is not significantly better than DNN (RELU) since both HSD groupings are 'B' but is significantly better than DNN (TANH) with HSD groupings of 'C'. Figure 2 shows a box plot of prediction accuracy for all four model specifications. These results are not surprising, because the daily data sample size is relatively small and it has generally been found that DNN only performs better than shallow architectures when using large datasets (Abdel-Hamid et al., 2012).

We increased the sample size by using hourly data with 15,000 observations. Figure 3 shows a box plot of prediction accuracy for all four model specifications. Using this data, the F-test produces an insignificant result ($F = 2.49$, $p > 0.0635$), suggesting that there are no differences in the predictive accuracy of the different model specifications. Table 5 shows HSD groupings of all the same letter 'A,' which confirms our findings of no differences in accuracy between the model specifications. Once again, these results are not surprising, since most studies only point to DNN producing superior results when far larger datasets are employed

(Abdel-Hamid et al., 2012; Srivastava, Hinton, Krizhevsky, Sutskever, and Salakhutdinov 2014). However, the fact that with hourly data, DNN now produces accuracy figures more similar to the shallower architecture models, suggests that the increased data size improves DNN's accuracy.

[Figure 2.1 about here]

[Table 2.4 ~ 2.7 about here]

[Figure 2.2 & 2.3 about here]

When employing the data at minute level, the number of observations increased substantially (to 800,000). An F-test then suggested that there were significant differences in the predictive accuracy of the four models (F = 9.36 $p < 0.0001$). Table 2.6 shows that the predictions from the two DNNs (using RELU and TANH activation functions) are significantly more accurate than those employing the SVM and NN models (see Figure 2.4). The two DNNs also have an HSD groupings of 'A' while SVM and NN have an HSD grouping of 'B' to stress these differences. SVM is slightly more accurate than NN, but not significantly so.

Finally, we explored the prediction accuracy of the four model specifications using tick level data (10 million observations for the training and validation). Table 2.7 shows that there was no significant difference between the performance of all four models (F = 0.75, $p > 0.5234$). The HSD groupings are all same letter – 'A'. Figure 2.5 shows a box plot of prediction accuracy for all four model specifications. It might have been expected that DNN with either activation function would perform better than both SVM and NN, given DNN's more complex architecture. However, our results are in line with those of Cho et al. (2016), who found that increasing the size of training data leads to an increase in predictive accuracy of DNN up to a point, but beyond that the accuracy does not change much regardless of training size. He attributed this to the fact that the training sets does not have a balanced class to reach the desired accuracy. This shows that the performance of DNN is not only influenced by sample size but by other combinations of factors (Ba et al., 2014). In particular, DNN can be very powerful because it can employ more parameters, it can learn more complex functions given the same number of parameters and it has better inductive bias (and, thus, learns more interesting/useful functions). However, it has also been shown in some cases shallow neural networks can learn these deep functions using the same number of parameters as the original deep models. For example, on the TIMIT phoneme recognition and CIFAR-10 image recognition tasks, shallow networks were trained to perform similarly to complex, well-engineered, deeper convolutional architectures (Ba et al., 2014). In summary, the result presented does not support the hypothesis H1, that predictive accuracy is higher using DNN for intraday (i.e. minute/tick) time horizons than SVM and one-layer NN.

[Figure 2.4 ~ 2.6 about here]

To test the second hypothesis, we explored whether RELU is better than TANH for stock price index predictions. Many practitioners have suggested that when employing DNN, one should start by using RELU, as this is regarded as a good approximator (Nair et al., 2010). Figure 2.6 shows boxplots of the predictive accuracy of DNN models using both RELU and TANH activation functions across all four-time horizons. Our results show that RELU is a better activation function for predicting stock price movement. This was expected, as many previous studies have found RELU to be far better than TANH (e.g., Krizhevsky et al., 2012). However, our results suggest that there are significant differences between the predictive accuracy which can be achieved when employing the RELU and TANH activation functions in DNN functions across the four-time horizons ($F = 16.87$, $p < 0.0001$) (see Figure 2.6). Whilst the RELU activation function leads to better predictions across all time horizons, it was far better than TANH when employing intraday data, especially minute data, across all markets (see Table 2.6). Our findings support our second hypothesis that using the RELU activation function produces more accurate stock price index predictions than using the TANH activation function across all time horizons.

To test our third hypothesis, we examined whether the predictive performance of DNN was better when applied to emerging markets and whether it outperforms both SVM and NN in these markets. We believed this may be the case because of its ability to approximate complex functions (e.g., using intraday and daily stock price data, DNN can learn time shifted correlations among stock prices: Ibikule et al. (2017) and Krauss et al. (2016)) and the behaviour of emerging markets is considered to be more complex (cf. developed markets) (Berger, Pukthuanthong and Yang, 2011).

Figure 2.7, compares the accuracy of the DNN model in predicting prices in emerging vs developed markets across all four time frames (daily, hourly, minute, tick). The results suggest that there was no difference in the predictive accuracy achieved between emerging and developed markets using DNN for stock price index prediction when using daily, hourly or minute level data. However, when employing tick level data the predictive accuracy of a DNN model applied to emerging markets appears to be greater than that achieved for developed markets (i.e. superior accuracy of 10%). Emerging markets for which DNN produced particularly high accuracy (i.e. >70%) using tick level data include Thailand, Malaysia and Czech Republic.

[Figure 2.7 about here]

To formally test whether DNN produced more accurate predictions for emerging markets than SVM or NN, we examined the statistical significance of observed mean differences across factor levels when using regression analysis as shown in (2.9). In particular, we estimate the following models to explain predictive accuracy:

$$Accuracy = \alpha + \beta_{DR1} DR1 + \beta_{DR2} DR2\ \varepsilon \ , \tag{2.9}$$

where DR1 and DR2 are dummy variables taking value 1 (for DNN (RELU) and 0 for SVM. DR2 is taking value 1 (for DNN (RELU) and 0 for NN. Fitting various algorithms as a factor variable in the model allows the response to vary in a nonlinear way and requires more parameters than a linear term. Factor variables are typically fitted using 'treatment contrasts' which means that one level of the factor variable forms the baseline and the remaining levels of the factor have corresponding coefficients.

The results of the regression analyses comparing the accuracy of DNN (RELU), SVM and NN are displayed in Table 2.8. The F-statistics and p-values confirm the statistical significance of the regressions. The adjusted R^2 values suggest that the independent variables explain about 19% of the observed predictive accuracy. This value may appear rather low. However, it is important to remember that the development of prices in financial markets is driven by a multitude of factors, many of which are not considered in this study.

In interpreting the result, it is important to remember the reference model that forms the basis of the comparison (i.e. the DNN (RELU) model). Considering the regression coefficient of the intercept in the accuracy regression model, such a model produces a directional accuracy regression model of 65%. The coefficients of SVM and NN are significantly different from DNN (see Table 2.8) and the negative signs shows a reduction in accuracy compared to DNN. We reject H3 since the accuracy of DNN in both emerging markets and developed markets were similar across all time horizons. When comparing DNN with SVM and NN, few markets appears to be better which might suggests why the coefficient of our regression for DNN is higher.

[Table 2.8 about here]

2.6 Discussion

Our results show that when employing daily and hourly data, shallow architectures (SVM and one-layer NN) produce more accurate stock index predictions than DNN. These results are in line with earlier studies which show shallow architectures produce better predictions when the dataset is small. It has been suggested that this arises because the data does not have a complex structure (Bengio et al., 2007).

As expected, we find that the accuracy of stock index price predictions using a DNN model are significantly better than that achievable using other commonly employed ML techniques (SVM and one-layer NN) when the data size increases significantly (i.e. when using a minute level data). However, we cannot accept hypothesis H1, that the predictive accuracy is higher using DNN (cf. SVM and one-layer NN) for all intraday time horizons, because when using tick data (over 10 million cf. 800,000 observations for minute level data)), the predictive accuracy of DNN models using both types of activation function were no better than SVM. This may appear surprising, as previous studies have suggested that the performance of DNN increases as data size increases. However, Schindler et al. (2016) compared shallow architecture versus DNN for music genre classification for example. They also carried out extensive comparative evaluations on four well-known datasets of different sizes including the application of two audio data augmentation methods. The system used is based on a parallel CNN architecture where separate CNN Layers are optimized for processing and recognizing music relations in the frequency domain and to capture temporal relations. Their result shows that for smaller datasets shallow models seem to be more appropriate since deeper models showed no significant improvement. They also found that deeper models performed slightly better in the presence of larger datasets, but a clear conclusion that deeper models are generally better could not be drawn. Their results clearly supports our findings concerning tick level data as well.

In addition, Cho et al. (2016) found that an increase in the size of training data for DNN led to improvements in predictive accuracy up to a certain point (an increase from 5 to 50 training images), after which accuracy did not change substantially, regardless of the training size. This might explain the behaviour we experience using this data. On the other hand, Ba et al. (2014) suggested that in some cases shallow neural networks are able to learn deep functions using the same number of parameters as the original deep learning. He used TIMIT phoneme recognition and CIFAR-10 dataset to demonstrate this. In addition, when he trained shallow neural networks directly on the original labelled training data with the same number of parameters as DNN, he noted that shallow neural networks learned to mimic a DNN with high fidelity. Hence, it is clear that the function learned by that DNN does not really have to be deep.

Our results offer support for H2, namely, that the RELU activation function produces more accurate stock price index predictions than TANH across all time horizon. This further explains why Nair et al. (2010) used CIFAR- 10 data to build a four-layer convolutional network and concluded that RELU was several times faster and better than TANH. RELU is also known to eliminate the vanishing gradients problems observed in TANH activation functions.

According to Krizhevsky et al., (2012), the biggest advantage of RELU is non-saturation of its gradient, which greatly accelerates the convergence of stochastic gradient descent compared to the TANH functions. Another nice property is that compared to TANH neurons that involve expensive operations, the RELU can be implemented by simply thresholding a matrix of activations at zero. Glorot et al., (2011) carried out various experiments comparing the performance of RELU vs TANH using four image classification datasets of different scales. They noted that RELU can find local minima of greater or equal quality than those obtained with its smooth counterpart, TANH. They also noted that RELU is computationally efficient.

RELU truly proves itself when used to solve non-zero-centre problem such as our case. All independent variables used in our analysis have values of x > 0 which solves the issue of vanishing gradients unlike TANH which produces dead neurons during computation.

Our results do not fully support H3, that the predictive accuracy of DNN is generally greater in emerging markets than in developed markets. In fact, there was no significant difference in predictive accuracy in emerging markets when using DNN and SVM for daily, hourly and minute level data. However, we did observe higher predictive accuracy when using DNN (cf. SVM) to forecast whether a stock price index rises or falls using tick level data for three specific emerging markets, Thailand, Malaysia and Czech Republic. Sutsarun et. al (2014) studied the impact of coup de'etats in Thailand in 2006 and examine the effect on both short-run and long-run dynamics return, volatility, liquidity, and liquidity risk of returns on the Stock Exchange of Thailand index over this period of 1 January 1996 to 31 December 2011. They used tick data in their analysis and found that the immediate reaction to the coup is more evident in the stock market with a reduction in stock return. This period overlaps with the data used for our analysis and could explain why DNN was able to capture the complexity better than shallow architecture in this market. Similarly, it has been found that Central and Eastern European countries (CEEC) such as Czech Republic suffers from the spillover effect transmit from major stock market turmoil in the USA and China (Deltuvaitė Vilma, 2016). The collapse of Lehman Brothers bank in United States in 2008 was the most significant shock transmitted to stock markets. Their empirical results also suggest that the transmission of other systemic shocks (e.g. the Middle East financial markets crash (May 2006), Greek debt crisis (April 23, 2010), Portugal's debt crisis (May 16, 2011)) was also observed on some of the CEECs countries. This also overlap with the data we observed for this market and could explain why shallow architecture were not able to match the performance of DNN. In Malaysia (Leow and Evelite (2015) examines the presence of a political cycle in Malaysia stock market returns and volatilities over the period of February 1982 to April 2012. They claimed that Malaysia stock market tends to overreact to unexpected political events such as removal of Deputy Prime Minister and resignation of Prime Minister. Their studies further shows that the presence of a political cycle in Malaysia stock market volatilities is statistically significant which indicates that investors take asymmetric treatments to the election information and the government policy. Using DNN in such markets with RELU activation function was in turn able to capture the sudden change in market dynamics.

2.7 Conclusions

Deep learning is making major advances in solving problems that have resisted the best attempts of the artificial intelligence community for many years, such as in computer vision and natural language processing. It has turned out to be very successful at discovering intricate structures in high-dimensional data and is, therefore, applicable to many domains of science, business and government. In addition, DNNs have been predicted to have many more successes in the near future, largely because they require very little engineering by hand. Consequently, DNNs can easily take advantage of increases in the amount of available computational ability and data. With so many successful applications of DNN, it is perhaps surprising that few studies have employed DNN to forecast financial time series (Krauss et al., 2016). Our paper is the first, to the best of our knowledge, to use DNN in the context of predicting the direction of stock price movements across multiple markets in order to understand to what extent this novel algorithm is sensitive to sample sizes. We set out to clarify if this is the case by comparing the predictive accuracy of widely used shallow methods with DNN using different forecast time horizons (daily/hourly/minute/tick).

As expected, our results show that data size has an effect on the relative predictive performance of DNN. SVM and one-layer NN outperformed DNN when daily and hourly data was used. However, the predictive accuracy of DNN was significantly better than that of SVM and one-layer NN using minute level data. Previous research has suggested that DNNs perform particularly well when they can learn the underlying structure using large datasets. It is likely that this explain our results, since there is a significant increase in data size from the 1,500 observations for daily prediction to the 800,000 observations for minute predictions (Bengio et al., 2007).

Interestingly, the accuracy of stock index price predictions at the tick level based using DNN was not significantly better than predictions based on shallower architectures. Since many studies have advocated the use of DNN when confronted with complex, big data, one would have expected DNN to have a better predictive performance than other methods at the tick price level. Some studies have suggested that this is not always the case (e.g., Cho et al., 2016) and further research is needed, employing data other than that related to stock markets, to see if similar a finding applies.

Taken together, our results suggest that practitioners who are looking to include DNN as one of the algorithms they use when predicting stock price movement, should first consider the complexity of the decision space, the size of the dataset and how balanced is the target class. A suitable future research avenue might be to develop an ensemble model which combines the power of SVM and DNN (Krauss et al., 2016). Research into DNN is going through a lot of transformation and other areas of exploration is by looking at different variants such as DeepAR, fuzzy methods or probabilistic methods.

A second important contribution of the research is that we have demonstrated that RELU activation function outperforms the TANH activation function when employing DNN across all forecast time horizons (daily, hourly, minute and tick level). Carefully tuning the hyper parameter optimization may still yield advantageous results for both activation function. This is subject to further research using more stock markets data. In case a practitioner is oblivious on which activation function to apply and if cheap computing power is available, RELU should be explored extensively to gain better predictive accuracy.

A third contribution of the paper is the discovery that DNN outperforms shallower architectures at the tick level in some emerging markets. In particular, we found that predictions based on DNN at the tick level were significantly more accurate than those based on shallower architectures for the Thai, Czech and Malaysian markets. We believe DNN was able to capture the recent upward trend in their economic growth which was also reflected in the data. The high levels of predictive accuracy achieved for these markets (79%, 84% and 79%, respectively) suggest significant exploitable inefficacies in these markets. Furthermore political issues such as sudden change of leadership in these countries and economic turmoil might suggest why shallow architectures were not able to perform as well as DNN. Since stock markets are rapidly changing in emerging markets, practitioners should always compare both shallow architecture and DNN before taking any decision. Further research should explore datasets from sub-markets within emerging markets using DNN with shallow architectures.

Overall, this paper has implications for financial economics and professional finance practitioners. In particular we provide empirical evidence that the most widely celebrated machine learning techniques (DNN) does not necessarily outperform SVM and NN in all cases using large data from many major financial markets. We show that RELU should be leveraged widely by practitioners as this was better in comparison to TANH when analysing stock market data. Lastly, none of the algorithms we tested were superior for predicting stock price indices in emerging vs developed markets, even though all the methods offer the prospect of identifying inefficiency in the pricing in these markets.

Chapter 3 A Cross-Device Examination of Conversion Rates from Online Browsing to Online Buying: The Online Path to Purchasing

Abstract

Marketing managers seek to understand cross-device consumer behavior in online path to purchase. Building on de Haan et al. (2018), we model the effect of online browsing across device types (PC, smartphone and tablet) on frequency of shopping orders per session. A dataset from a large multinational retailer is used. Due to device type endogeneity and sample bias, propensity score matching is employed. We find that frequency of completed orders is highest via tablet. When the device used during the previous session was PC or tablet, the conversion rate is higher for PC (6% and 5%). However, when the device used previously was smartphone, the conversion rate for smartphone is higher (8%). Consumers use tablet for viewing clearance pages and PC for product reviews. Tablets bring higher cart value, while PCs increase positive effects of having an existing cart and reading reviews on conversion. Smartphones and tablets experience more traffic than PCs before and after work hours, while PCs experience more traffic during lunch time. Findings contribute to understanding of browsing patterns, time spent shopping, and time of visit across device types. An emerging explanation of the online path to purchasing and managerial implications for practitioners in e-commerce/m-commerce are offered.

Keywords: Online consumer behavior, e-commerce, big data, mobile commerce

Marketing managers, e-tail practitioners, and advertisers alike are facing low browsing-to-buying conversion rates. The vast majority of online browsing is not converted to a sale during that online session. Understanding online consumer behavior on different devices, including their smartphone, tablet, or personal computer (PC, which includes both laptops and desktops) is important for scholars and marketers alike. To address low conversion rates, managers seek knowledge about cross-platform online consumer behavior. This managerial problem concerns conversion rates among device types in electronic commerce (hence, e-commerce), which broadly refers to online shopping via technology such as PCs. It also concerns a more specific form of e-commerce—mobile commerce (hence, m-commerce), which informally refers to online shopping from consumers' mobile devices such as smartphones and tablets such as iPads. Managers may ask questions, such as which platform offers higher conversion rates and how online consumer browsing and buying behavior differs among these platforms.

In addition to being a contemporary managerial problem, this topic has a global economic significance as seen by industry trends. For example, E-Marketer estimated that US consumers would spend $709.78 billion on ecommerce in 2020, an annual increase of 18% (eMarketer 2020). According to retail analysts at GlobalData, over half of UK consumers shop online, with an increase of 29.6% forecasted between 2019 and 2024. In the most recent survey by Capgemini IMRG eRetail, a quarterly sales index using data from 210 UK retailers, over 40% of online sales were made via phones in the fourth quarter of 2018/2019.

Consistent with the IMRG industry report, scholars find that mobile devices are the medium of choice when shopping online (de Haan et al. 2018; Xu, Kaiquan et al. 2017). Yet, preference for mobile devices when shopping does not necessarily mean they will be used to to purchase. Surprisingly few scholars have looked at the roles of mobile devices and purchase patterns (Kannan and Li 2017; Verhoef et al. 2017) relative to the massive "digital transformation" that deserves multidisciplinary inquiry (Verhoef et al. 2021). Despite the rising importance of m-commerce, shopping from a PC is worthy of understanding because PCs may have a higher conversion rate compared to mobile platforms. There may be a conversion gap in e-commerce, which is a significant discrepancy in browsing vs. purchasing on one e-commerce platform compared to another e-commerce platform.

Thus, our objectives are twofold. The first objective is to model the effect of consumers' online browsing across device types (PC, smartphone and tablet) on the frequency of orders completed per online shopping session. The second objective is to consider how the online path to purchase process may be distinct for each device type.

The research offers three contributions to the marketing literature: First, this is the first study to empirically study cross-device e-commerce along with different touch points across various stages of the online path to purchase. Second, we examine PCs, tablets and smartphones, where most of the literature focuses on just the smartphone. Third, we expand knowledge on online shopping cart use by studying the effect of existing items in a cart from past shopping sessions on purchase frequency. A further contribution is an addition of three factors to extend work by de Haan et al. (2018): 1) browsing patterns across various pages visited by consumers, 2) time spent on each shopping session, and 3) time of day of visit.

We fill a knowledge gap by determining the role of mobile devices in the online path and value to conversion, focusing on the moderators of any device type effect on shopping behavior patterns. This supplements only two other studies in marketing that use a cross-device approach (de Haan et al. 2018 and Xu et al. 2017). Our findings show that customers who shop using tablet and who have visited clearance pages are more likely to complete an order in a shopping session, followed by those who shop using PC. We show how the smartphone is the least used device for conversion. Our findings also reveal that customers with items in their e-cart from a previous session using PC are more likely to complete an order in a current session compared to those using other devices (i.e., smartphones, tablets). We also explain why customers who read reviews of items during their browsing trip on PC (vs. phone or tablet) are more likely to complete the order. In sum, we demonstrate that tablet is the device of choice, leading to highest conversion and purchase frequency overall, while PC is still an extremely important device that leads to high conversion/purchase frequency.

The rest of this paper is organized as follows. Next, we provide a synthesis of literature on the role of device type in e-commerce and m-commerce. An overview of the empirical context containing the contextual setting for the data follows. The subsequent section provides a description of the empirical models used in the analyses, followed by the results and robustness checks. We then explore potential underlying theoretical mechanisms and boundaries. We conclude with a discussion of results, implications, limitations, and a suggested future research agenda.

3.1 Literature Review

Table 3.1 synthesizes the relevant literature in marketing and delineates where this research is positioned while illuminating the void in the literature on e-commerce conversions by device modality. It is clear from Table 3.1 that there are only two works that examine e-commerce from a cross-device perspective, and they have conflicting results; as such, in lieu of hypotheses we take a modelling approach and interpret the findings with respect to those papers while offering a start towards emergent explanation of the online path to purchase.

[Table 3.1]

A common pattern among scholarship examining online purchasing using mobile devices is that they examine only one device (namely the mobile phone) in isolation (e.g., Luo et al. 2014; Fong et al. 2015; Andrews et al. 2016; Dube et al. 2017; Li et al. 2017). Again, the two notable exceptions are de Hann et al. (2018) and Xu et al. (2017) which examine cross-device browsing behaviors and conversion rates. Xu et al. (2017) were the first to investigate the cross-device impact on conversion rates. Specifically, they examined any complementary and substitution impact of the tablet channel on the smartphone and PC channel. They leveraged a dataset from Alibaba—the largest e-commerce firm in the world—and exploited a natural experiment via the iPad app introduction. Their results show that consumers' adoption of tablets substantially enhanced the overall growth of Alibaba's e-commerce market. This impressive dataset illuminated the importance of the tablet. Their study also examined cross-device browsing, which was limited to "as instances where users browse on two different devices within a one-hour time window" (Xu et al. 2017, p. 1486).

de Haan et al. (2018) is only the second paper to analyse browsing patterns across PCs, smartphones and tablets. They investigate the role of device switching on online purchasing and find that the increased adoption of mobile devices significantly impacts online shopping behavior, with customers at times switching between mobile and fixed devices on the path to purchase. Device switching is analyzed using clickstream data from a large e-tailer. Propensity score matching accounts for self-selection in device switching. The authors also find that when customers switch from a mobile device (e.g., a smartphone) to a stationary device (e.g., a desktop), their conversion rate is significantly higher.

There are opportunities to build on the seminal work of de Haan et al. (2018). Specifically, since then, mobile devices have become even more prevalent in the life of many consumers. E-commerce technology has also improved, potentially leading to changes in previous research findings. Furthermore, de Haan et al. (2018) focused on purchase completion as a binary outcome. Thus, building on Xu et al. (2017) and de Haan et al. (2018), our work will offer insights concerning the impact of cross-device browsing on purchase conversion and frequency. In particular, our research will consider session-level pattern browsing and conversion among PCs, smartphones and tablets. This represents a more comprehensive approach to examining cross-device browsing than has been done in the past.

Last, we will go a step beyond considering purchase completion as a zero-one variable, and also investigate frequency of purchase completion (or buying) within a session to allow for the possibility of a consumer completing multiple orders within the same shopping session.

3.2 Methods

3.2.1 Data Description and Variable Operationalization

We use individual-level clickstream data from a large multinational online retailer to develop an empirical model to explain consumers' browsing and purchasing behavior across different device types. The online retailer is a British sport brand, clothing, footwear and home products retailer with a strong multi-national presence. For the study, we use observations from customers who engaged in at least two sessions to determine how the previous session influences the current session. The online retailer defines a session as one continuous period in which the customer is active on the website. An online shopping session starts when the customer enters the online retailer's website and ends when the customer actively leaves the website or when the customer is inactive for 30 minutes or more.

We use data from registered customers because registration is necessary to complete the purchase and it is not possible to capture device switching by unregistered users. We use sessions that belong to the same path to purchase as defined by Li and Kannan (2014) as multiple touch points that a customer makes before conversion. If, for example, a customer uses a smartphone in one session but does not convert and returns in the next session using a PC and eventually converts, the two sessions belong to the same path to purchase. Our data contains device unique IDs which allows us to track and link a consumer identifier to the devices used on the website during each visit. The final sample includes 55,500 unique customers who engaged in 958,859 usable sessions. For each session, the data include information regarding which device was used to visit the website. The dependent variable of interest is frequency of purchase completion within a single shopping session, which is novel compared to past work that only captures if a sale was made or not. Some researchers have investigated purchase quantities (Boatwright et al. 2003), but the factors that impact frequency of purchasing per session across multiple device types have not been explored.

We categorize the sessions by the device used, as shown in Table 2 with the descriptive statistics per device type. We can see that more customers prefer to continue the online path to purchase on the same device in subsequent sessions rather than switch devices. We find that the conversation rate for tablet per session is 8.58%, which shows that consumers convert on average more via tablet than via PC (7.96%) and smartphone (5.50%). However, when the frequency of orders per session is greater than one, smartphone is the leading device for conversion with 2.78%, followed by PC (1.82%) and tablet (1.36%). When device used during the previous session is PC, we find that the conversion rate is higher for PC (6%). When device used during the previous session is tablet, we find that the conversion rate is higher for PC (5%). However, when the device used in previous session is smartphone, the conversion rate for smartphone is higher (8%). Table 3.3 shows summary statistics of variables across all devices. In addition, the correlation matrix is in Table 3.4. Please note that for device types tablet and smartphone, PC is serving as a baseline comparison. We next discuss the independent variables included in the model and the reasons for their choice.

[Table 3.2, Table 3.3, and Table 3.4]

Online retailers have strategic interest in the e-cart value of a transaction, because their profits are directly tied to the shipping costs associated with the purchases (Boatwright, Borle, and Kadane 2003). E-cart value is used as one of the predictors in our analyses, as the total monetary value of all items in the e-cart. Related to e-cart shopping behavior it indicates whether the consumers started the shopping session with a new (i.e., empty) cart or had any items in an existing cart. We used a variable named existing cart to denote a use of a cart containing item(s) from a previous shopping session.

Consumers may spend time browsing pages and gathering information before purchasing. Past research shows that consumer online purchase behavior can be better predicted by incorporating browsing characteristics as covariates (Montgomery et al. 2004) and that browsing characteristics such as page views and visit duration are jointly determined within a session (Bucklin and Sismeiro 2003). As such, we integrate browsing pages and active seconds spent per session into the model. Additional browsing characteristics include indicators about whether consumers visited a clearance portion of the shopping website and whether they browsed through any customer product reviews. To investigate the moderating role of device type, interactions of device type with 1) browsing clearance sites, 2) reading product reviews, and 3) having an existing cart, are also included. We also incorporate time spent shopping and time of the day when shopping into our model as independent variables.

3.2.2 Controlling for Endogeneity Using Propensity Score Matching and Control Variables

Endogeneity in marketing models can lead to biased coefficient estimates (Germann, Ebbes, and Grewal 2015; Papies, Ebbes, and Van Heerde 2017). Because consumers were not randomly assigned to the treatment (device type) but self-selected the treatment, different device users could vary systematically across the different device type groups. To account for this potential self-selection bias (Garnefeld et al. 2013), we employ propensity score matching (PSM) and create an artificial control group. First, in a binary logistic regression we calculate each customer’s propensity to use a particular device (PC vs mobile (including SmartPhone/Tablet)) to purchase an item (see Table 3.5, Panel A). Second, the matching procedure links each customer in the treatment condition with a statistical twin from the control group who did not purchase using a particular device but has statistically the same propensity to. With a caliper matching procedure, we match each treatment case to its nearest neighbor only if two propensity scores fall within a pre-set tolerance zone (Wangenheim and

Bay on 2007). Limiting the scores to differ by a maximum of .001 – much below the recommended tolerance zone of .008 (Silverman 1986) – we match 150,228 customers from the treatment customers. Third, following Bommaraju and Hohenberg (2018), we compute standardized differences in means before and after matching (Table 3.5, Panel B).

In addition to controlling for the endogeneity of device used, another way to correct for endogeneity involves using instrumental variables. Instrumental variables allow for unbiased estimates, which are implemented through two-stage least squares (2SLS) models that use estimated values from a first-stage regression for a possibly endogenous variable (Germann, Ebbes, and Grewal 2015). Current recommendations stress to first use control variables in the data sets to control for unobserved effects before deciding to use instrumental variables (Germann, Ebbes, and Grewal 2015; Papies, Ebbes, and Van Heerde 2017). Thus, we also use control variables. To allow for a fair comparison of the potential impact of user interfaces, screen sizes of the device types represent control variables in our model. We also control for alternative explanations that might arise due to variation in individual differences. Six dummy variables account for continents or regions where the consumer is browsing from, with Asia as the baseline (Africa, North America, South America, Europe and Australia) using the categorization provided by Statista 2020 reports on internet traffic percentage of total web traffic by region. We also include dummy variables to account for variation across time of the day. We add a series of control dummy variables that denote the "sequence" of each session. Thus, appropriate measures have been undertaken to address endogeneity. Next, we overview methods beginning with the empirical model and analyses.

3.2 Empirical Model and Analyses

Recall, our primary objective is to understand the effect of consumers' online browsing across device types on the frequency of orders completed per shopping session. To develop the most appropriate model, we need to account for the infrequent behaviors of consumer buying process in a session. Data with abundant zeros are common in studies that count the occurrence of online purchases. Count data have values that are usually non-negative with a lower bound zero and typically have excessive zeros and over-dispersion.

A typical way of analyzing count data includes specification of a Poisson distribution with a log link (the log of the expectation of a response variable is predicted by the linear combination of covariates, i.e., predictors) in a Poisson regression. Another more formal way is to use a negative binomial (NB) regression. Each of these models belong to the family of generalized linear models (McCullagh and Nelder 1989). Marketing literature analyzing online purchase behaviors has generally not considered the mean-variance relationship and proportions of zeros in such data, making such model bias and results from such models not as reliable in those cases. The frequency distribution of orders completed per session in the data indicates an extremely large number of zeros. Using Ordinary Least Square (OLS) estimation, the most widely used regression method, is not suitable due to bias in such coefficient estimates. Therefore, to model the frequency of orders completed per online shopping session, we considered the Poisson regression model, the negative binomial regression model, and the zero-inflated negative binomial. Each of these are overviewed next.

3.2.1 Poisson Regression Model

In the Poisson regression model, the frequency of orders completed per session is modelled as follows:

$$\Pr(Y_i = y_i) = \frac{e^{-\mu}\mu^{y_i}}{y!}, y_i = 0, 1, 2 \ldots \ldots \ldots \quad (3.1)$$

where μ is the count mean.

Let $X = (X_1, X_2, \dots \dots, X_p)$ be a vector of covariates, in this case, device type, seconds spent, read customer review, e-cart value, existing cart, product views, time of visits, and clearance page visits. Let $\beta = (\beta_1, \dots \dots, \beta_p)$ be a vector of regression parameters.

The logarithm of μ is assumed to be a linear combination of p covariates of the form

$$E(Y|X) = \mu = \exp(X\beta)$$

The conditional mean and conditional variance are equal for the Poisson regression model, that is $E(Y|X) = Var(Y|X) = \mu$. The greater the mean, the greater the variability of the data.

Another way of dealing with over-dispersion is to use the mean regression function and the variance function from the Poisson GLM, but to leave the dispersion parameter ϕ unrestricted. Thus, ϕ is not assumed to be fixed at 1, but is estimated from the data. This strategy leads to the same coefficient estimates as the standard Poisson model, but inference is adjusted for over-dispersion. Consequently, both models (quasi-Poisson and adjusted Poisson) adopt the estimating function view of the Poisson model and do not correspond to models with fully specified likelihoods.

3.2.2 Negative Binomial Regression Model (NB**)**

The assumption that the variance of counts is equal to the mean also implies that the variability of the outcomes sharing the same covariates values (a population has the same values for $(X_1, X_2, \dots \dots X_p)$) is equal to the mean. If this fails to be true, estimates of the regression coefficients can still be consistent using Poisson regression, but the standard errors can be biased. They tend to be too small and increase the rate of Type I error (false positive results) (Hilbe 2014). The negative binomial distribution has the following form:

$$\Pr(Y_i = y_i) = \frac{\Gamma(k + y_i)}{\Gamma(k) y_i} \left(\frac{k}{k+\mu}\right)^k \left(\frac{\mu}{k+\mu}\right)^{y_i} \qquad (3.2)$$

Where μ is the mean and k is the dispersion parameter. The variance of the above distribution is $\mu + \mu^2/k$, and hence decreasing values of k correspond to increasing dispersion levels. As

k approaches infinity, a Poisson distribution is obtained. The negative binomial regression model captures the over-dispersion in count data that the Poisson model cannot.

3.2.3 Zero-inflated Negative Binomial (ZINB)

Zero-inflated negative binomial (ZINB) (Mullahy 1986; Lambert 1992) is another model class capable of handling excess zeros. ZINB can handle zero-inflation and over-dispersion simultaneously. An advantage of using models that deal with zero-inflation is that they reduce bias from extreme non-normality and have the ability to model the effect on consumers' susceptibility and magnitude simultaneously. The probability is given as:

$$\Pr(Y = 0) = \pi + (1 - \pi)(1 + \alpha\mu)^{-\frac{1}{\alpha}} \tag{3.3}$$

Where mean is μ and α is the over-dispersion parameter.

3.3 Analysis

All analysis was conducted using R software. To determine which of the models is the most appropriate to use, we estimated the fitted count regression model. Table 6 includes a comparison of five considered models (Poisson, Poisson adjusted, quasi Poisson, negative binomial, and zero-inflated negative binomial).

[Table 3.6]

As a first attempt to understand the effect of online browsing across device types on the frequency of orders *completed* per shopping session, we fit a basic Poisson regression model. Table 6 shows that all coefficient estimates are highly significant. Yet, over-dispersion is present; thus, we compute the Poisson model using sandwich standard errors. Some independent variables are still significant, but the standard errors are more robust. The quasi Poisson model deals with over-dispersion (excess zeros) in a formal way. The coefficients are similar to the sandwich standard error, leading to similar parameter estimates. Next, we consider a negative binomial (NB) model. The NB has an advantage in that it is associated with a formal likelihood so that information criteria are available. The coefficients and standard error appear to be more stable in comparison to the Poisson models. A different way of augmenting the NB model with additional probability weight for zero counts is a zero-inflated negative binomial (ZINB) regression, considered next. In Table 6, the output for ZINB contains negative binomial regression coefficients for each of the variables along with standard errors. A second block follows that corresponds to the inflation model. This includes logit coefficients for predicting excess zeros along with their standard errors.

3.3.1 Model Selection Criteria

To select the preferred model, we first used the minimum AIC (Akaike information criterion) (Bozdogan 2000). AIC is given by:

$$AIC = -2logL(\theta) + 2c \qquad (3.4)$$

where L(θ) is the maximized likelihood function for the estimated model and −L(θ) offers summary information on how much discrepancy exists between the model and the data, where c is the number of free parameters in the model. The negative binomial is the preferred model based on the smallest AIC value (i.e., 415,673).

Next, we ran the Vuong non-nested test, which is based on a comparison of the predicted probabilities of two models that do not nest. A large and positive test statistic provides evidence of the superiority of model 1 over model 2, while a large and negative test statistic provides evidence of the superiority of model 2 over model 1. Table 3.7 shows the model comparisons using Vuong non-nested tests. The model comparisons include Poisson vs. NB model; NB vs. ZINB model; and Poisson vs. ZINB model. Vuong test statistics and *p*-values are shown. The first comparison (Poisson versus NB model), with a Vuong test statistic of -47.91 and $p < 2.22e\text{-}16$, indicates that the NB model is preferred. The preferred model is then compared with the next model (ZINB). After a series of tests and model comparisons, NB remains the preferred model of choice as the Vuong test statistic shows in the second column of Table 3.7. Hence, results presented next are based on the NB model.

[Table 3.7]

3.4 Results

The results focus on the effects of the following factors: device type on frequency of completed orders, items in e-carts from prior online shopping session on multiple orders by device type, number of pages viewed and time of day on frequency of orders completed, time spent online on frequency of orders completed, viewing product reviews on frequency of orders completed, and visiting clearance pages on frequency of orders completed. Next, we overview the findings and compare them to the two other studies that examine cross-device considerations on either purchase conversion (de Haan et al. 2018) or sales (Xu et al. 2017).

3.4.1 Key Findings

The effect of device type on frequency of completed orders. In line with the key research objective, we find that consumers' device type has a significant effect on the frequency of orders completed. Particularly, the estimate of the coefficient of frequency of completed orders on smartphones is 0.43 lower ($p < .001$) than for those using PC. However, for those using tablet, the estimate of the coefficient of the frequency of completed orders is significantly higher than for PC by 0.20 ($p < .001$). Thus, the conversion and purchase rate is highest when consumers shop via tablet, followed by PC and then smartphone. The finding that smartphones have lowest frequency of orders completed is consistent with Xu et al. (2017), who did not study conversion rates but similarly found that tablet adoption increases online sales. Xu et al. (2017) also showed that tablets are considered as complementary device to smartphones, as they are both platforms for m-commerce. Thus, our results showing the strength of tablets are distinct yet complementary to Xu et al. (2017) in that they both demonstrate that tablets are the strongest device type for online sales and conversions.

Our finding also makes sense when considering complementary work on consumers' purchase intentions with respect to using mobile devices. Studies by Bart, Stephen, and Sarvary (2014) and Grewal and Stephen (2019) find a positive impact of mobile devices on purchase intentions. Another complementary empirical study about mobile devices finds a positive role of mobile devices (Lou, Andrews, Fang and Phang 2013) on purchase of a promoted movie. Furthermore, two conceptual papers about the role of mobile devices in consumer behavior also suggest that mobile platform should have a positive impact (Shankar et al. 2010 and Grewal et al. 2016).

Despite this, our finding contradicts the finding from de Haan et al. (2018), which is the only other paper besides Xu et al. (2017) that also examined cross-device factors on online shopping with behavioral data. de Haan et al. (2018) found that a less mobile device such as a desktop has a significantly higher conversion rate compared to mobile devices (tablet, smartphone). Other complementary papers examining mobile devices also found a negative impact of mobile devices. Namely, Ghose, Goldfarb and Han (2013) found less clicks from a mobile device and Marz, Schubach, and Schumann (2017) and Ransbotham, Lurie, and Liu (2019) found consumers report feelings of lower value via mobile.

Thus, while there is limited literature examining tablets, some research on the net effect of mobile devices in general in marketing studies is mixed. To provide further support of our suggestion that there is a positive effect for marketers when tablets are used, we find that e-cart value is highest with tablets. In addition to the tablet being associated with the highest conversion to purchase rate, the average total e-cart value of shoppers using tablet (£33.00 or $40.71) is higher than of those using PC (£29.84 or $36.81) or smartphone (£24.24

or $29.90). Thus, consumers tend to have the highest valued e-cart when shopping on tablets and lowest when on smartphones (with close to a ten-dollar difference between the two).

The effect of items in e-carts from prior online shopping session on multiple orders by device type. Our results in Table 6 (NB column) also show that having an existing cart containing item(s) from an earlier shopping session overall positively affects frequency of orders during the current session (.475, $p < .001$). While there is no formal research to which we can compare this finding, it has face validity in that a consumer who begins a shopping session with a product or products from their past online shopping session is more likely to convert to the sale. This could be explained by work of Close and Kukar-Kinney (2010) who note that putting an item in an e-cart in itself can be explained by a latent motivation to purchase that product. Furthermore, Kukar-Kinney and Close (2010) note that the role of dynamic carts (carts that keep the unpurchased items in the cart after the end of a shopping session) may be important to help conversions to sales at a later point.

Further, as seen in Figure 3.1A, there is a significant interaction between device type and having an existing cart. The results suggest that customers who started a shopping session on a PC and had items in their cart from a previous online shopping session completed more purchases in a session than when smartphones or tablets were used. This suggests that customer returns to the previously used cart are most profitable when done on a PC, possibly indicating a more purchase goal-driven motivation, rather than a more random or coincidental browsing. Studying effects of existing cart and its interactions contributes to the growing literature on consumer e-cart use.

[Figure 3.1]

The effect of number of pages viewed and time of day on frequency of orders completed. We find the number of pages viewed to have a significant positive effect on frequency of purchase completion (.069, $p < .001$): the more webpages consumer has seen, the greater the frequency of ordering within the session. This result supports a finding by Mallapragada et al. (2016). Extending their work, our data allow us to also monitor consumers' shopping activities during the time of day. Specially, we find that shopping before working hours has a significant positive effect on frequency of completed orders (0.073, $p < .01$). Thus, early morning shopping is most likely to result in a purchase. Next, shopping during lunch hours also has a positive significant effect on frequency of completed orders (0.043, $p < .01$). This means that those who shop during the lunch hours tend to have a good chance of purchasing. However, we find that shopping in the evening has a negative effect (-0.082, $p < .01$) on the frequency of orders made. For every visit made during the evening, the frequency of orders drops. This means that shopping online at night, regardless of device type, has the lowest conversion. We see that consumers tend to *browse* online more so than *purchase* online at night. Our study is the first to link frequency of purchase made across all devices with time spent during different times of the day.

The effect of time spent online on frequency of orders completed. On average, customers spend more shopping time on their smartphone before working hours (between 6-8am) and on their tablet after working hours (between 7-11am). However, this is not significantly different from time spent on other devices. Time spent online (in seconds) to search for items has a significant positive impact on purchasing online; yet the coefficient is very small (0.0001, $p < .01$) due to the unit of measurement (seconds). Our study is the first to link frequency of purchase made as a function of the time spent on the website.

The effect of viewing product reviews on frequency of orders completed. A further result concerns the impact of viewing product reviews on online purchase behavior. There is no significant effect of product reviews on frequency of orders completed (0.001, $p > .05$) overall. This is contrary to Liu et al. (2019), who suggested there may be an effect of product reviews on online consumer purchase behavior and Allard et al. (2020), who found via lab experiments that exposure to product reviews ultimately results in positive consumer responses. However, the latter does not necessarily mean purchases. Interestingly, and as depicted in Figure 3.1B, even though the main effect of product reviews is not significant in our research, we identify a significant interaction effect of device type and reading product reviews. Thus, device type serves as a moderator of the effect of product reviews. Customers who read other consumers’ reviews of items before purchase on PC completed more orders than when any other channels (i.e., smartphones and tablets) were used. Further, reading product reviews on a smartphone is less likely and does not drive purchase behavior. Thus, using big data, we provide evidence to show that viewing product reviews increases the frequency of completed online shopping orders, but only for PCs.

Consumers likely use smartphones for convenience and shorter shopping sessions (de Haan 2018), rather than conducting extensive product research and information search, such as reading through customer reviews. While there is a rich literature on product reviews, there is little if any that provides evidence for the effect of product reviews on frequency of orders completed for different devices using behavioral click-stream data.

The effect of visiting clearance page on frequency of orders completed. The results (Table 3.6, NB column futher indicate that visiting a clearance page has a significant and positive main effect on frequency of purchasing (0.265, $p < .01$). Customers who visit a clearance page are more likely to complete purchases than those who do not visit it. Our analyses further show a moderating role of device type on the effect of visiting a clearance page. This moderation effect is depicted in Figure 3.1C. Specifically, customers who shop using their tablet and who have visited a clearance page complete more purchases in a session, followed by those who shop using PC. Smartphone is the least used device to encourage purchasing upon visiting a clearance page.

The effect of e-cart value in e-carts on frequency of orders completed. Last, e-cart value has a significant and negative impact on frequency of orders completed (-0.00003, $p <$ .01). This means that the higher the total monetary amount of the items in the e-cart, the less likely that order will be completed. While this is a unique finding, it can be linked to an earlier finding by Close and Kukar-Kinney (2019) that consumers may place items in e-carts in part as a motivation to see the total price for the items of interests. A potential explanation for the negative effect is that when consumers have a big ticket item or high value items in their e-cart, they may be reluctant to complete the transaction due to sticker shock (seeing a high total price) and the financial risk of spending a large amount. This may be especially true with online shopping, where the consumer in many cases has not physically seen or touched the product(s).

4.2 Robustness Analysis

To enhance the validity of the results and to improve the confidence in the estimation approach, we conducted robustness analysis via a Bayesian framework, which provides multiple benefits. It allows probabilistic statements about the parameters for each consumer, provides richer information on the model's parameters by using a posterior distribution (rather than point estimates), and allows for more direct inferences about effects (Rossi and Allenby 2003). Dyachenko and Allenby (2015) examined the traditional approach to mediation analysis based on the statistical significance of the coefficients in the regression models. The results showed that these models are neither necessary nor insufficient when it comes to claim the presence of mediation for specific individuals. The Bayesian framework allows us to incorporate heterogeneity into the mediating behavior. To obtain the parameter estimates, we used the Bayesian Markov Chain Monte Carlo (MCMC) method which draws from the posterior distribution of the parameters obtained and used as a basis for inference.

Using Gibb samplers with conjugate priors (Rossi and Allenby, 2003), we estimated a joint posterior distribution of the parameters for our model. The conjugate prior allowed us to analytically derive and directly sample the conditional posteriors of the state parameters. Conversely, given the state parameters, the system became a multivariate probit model with hierarchical priors, thus we sampled the data directly using the standard MCMC method (Rossi and Allenby 2003). Our parameter estimates are similar to the NB estimates in Table 6, giving confidence in the robustness of the results. Based on these results that focus on the cross-device aspects, we now offer a more wholistic explanation of the theoretical process.

3.5 Towards a Theory of Online Path to Online Purchasing

This proposed explanation is a step towards extending theory in marketing on the consumer path to purchase model (Batra and Keller 2016) by considering the process specific to online paths to online purchasing. Specifically, while the "awareness" and "consideration" stages of the path to purchase are triggered on any device, the findings indicate that the "search/learn" phase of the process, as indicated by searching reading product reviews, is particularly relevant when consumers are using PC or tablet, and helps them make the purchase decision. Further, the "seeing value, being willing to pay" phase (Batra and Keller 2016) is found to be most relevant to consumers using tablets, as indicated by highest clearance site visits, and their largest effect on purchase completion via tablets.

To show how the online shopping session tends to progress based on these data, we offer a theoretical explanation of the online path to purchase, which helps explain the process that consumers may follow when shopping online. Figure 3.2 depicts this process by device type (A-C). Consistent with Batra and Keller (2016), consumers do not necessarily proceed through all of the steps of the purchase process and complete the purchase. For instance, consumers often will stop after merely browsing. These steps are depicted in Figure 3.2A-3.2C.

[Figure 3.2A-3.2C]

3.5.1 An Explanation of the Online Path to (Online) Purchase

Step 1. Enter an e-commerce or m-commerce site. In step 1 of the online path to online purchase, the consumer enters the e-commerce site using their selected device type. Because going on to the online shopping site is the first step in the online path to purchase, and this is where any click-stream modelling begins for an e-tailer, 100% of consumers enter the site and thereby browse one page. This page is the page that the consumer first views, and is considered their landing page, which may differ by the way in which they came to the page and their respective search terms or what they entered into their browser. While most consumers go to step 2 and browse other pages, for some consumers who do not want to look at other options, they may skip step 2 and go straight to step 3 and view the product details.

Step 2. Browse pages. In step 2 of the online path to purchase, the consumer begins to browse. In this step, the consumer leaves the first step (their landing page) and clicks on to another page on the website. Based on our data, 14.97% of online consumers entered step 2 and browsed two or more pages on a given e-commerce website using PC as shown in Figure 3.3A. On the other hand, the corresponding percentage for those using smartphone was 39.7% and for tablet it was 12.4% Other pages visited are most commonly other products, but two other types of pages stand out as important. One is browsing the clearance pages on a website; we find that 0.03% of consumers viewed a clearance page using PC, while 0.05% used smartphone and 0.02% were on tablet when viewing clearance page. The second notable type of page is a product review page. Despite the great interest in the marketing literature on online review pages (e.g., Grewal and Stephen 2019; Mars, Schubach, and Schumann 2017; Ransbotham, Lurie, and Liu 2019), it is still not known how visiting review pages impacts online consumer journey. Accessing a product review page is not necessarily viewing product details, which is step 3.

Step 3. View product. In step 3, the consumer views the product. Here, they are able to see photos of the product and product details and information. Note, that this is different from the product review page, which may be linked from the product page, but it is on a separate page. In our data, 8.01% of online shoppers who entered a retail site clicked on the product using PC. However, 0.99% used smartphone while 7.09% were on tablet as shown in Figure 3.3B (smartphone) and Figure 3.3C (tablet). A large drop in the percentage of customers who view products using smartphone indicates that in many cases smartphone is used only for the initial steps in the customers' online journey, but that consumers quickly lose interest in pursuing the shopping task. This drop is smallest for tablet, followed by PC.

Step 4. Add to e-cart. In step 4, the consumer places one or more items in their e-cart. In the vast majority of e-tail websites it is necessary to place the item in an e-cart to be able to purchase it. Hence, the role of the e-cart is one of the most important aspects in all of e-commerce. In our research, 1.14% of consumers who began an online shopping session on a retail site placed one or more items in their e-cart on PC. On smartphone, the corresponding value was 0.11% and on the tablet it was 1.11%. Similar to step 3, the drop in percentage of customers proceeding to this step from step 3 is the largest for smartphone users, and the smallest for tablet, with PC performing close to, but not quite as well as tablet.

While it is typically necessary to place an item into an e-cart to purchase it, not all consumers keep an item in their cart. Of PC shoppers, 0.41% will remove the item from their e-cart, therefore exhibiting an early form of shopping hesitation which may lead to later e-cart abandonment. E-cart abandonment occurs when consumers simply do not complete the purchase of any items within the same shopping session, and leave or log off without completing the purchase (Close and Kukar-Kinney 2010). On smartphone 0.03% of shoppers will remove an item, while on tablet, the corresponding percentage is 0.40%. Interestingly, in comparison with the percentage of customers who add item(s) to e-cart, the percentage of customers who remove items from e-cart is smallest for smartphone, with tablet and PC resulting in a larger proportion of those who remove items.

Step 5. Place order. The fifth step in the online path to purchasing is to place the order. This entails purchasing online and ordering the product or products that were in a customer's e-cart. At times, the item may have been placed in a past online shopping session, and at other times, during the current shopping session. Regardless, we show that 0.29% of all online shopping sessions started using PC will ultimately end up with at least one order. On smartphone, we found the corresponding value to be 0.02%, while it is 0.27% on tablet. In the vast majority of cases, the customer online journey will end here. However, some customers will also proceed to a sixth step in the online path to purchasing.

Step 6. Place multiple orders. In the final step, a consumer who has already completed an order (see step 5) continues to shop during the same session and orders something else, thus completing multiple purchase orders during the same session. This entails going back through the steps again in many cases, as the consumer is still in the online shopping session, viewing an additional page or pages, viewing the item, placing it in the cart, and completing a second (or further) order. For PC and tablet users, .02% of all started shopping sessions conclude in multiple orders, while for smartphone users this number is .001%. However, relative to all orders completed, multiple orders represent 7.4% of orders for tablet, 6.8% for PC, and 5% for smartphone, thus not a negligible amount. The findings show that tablet is not only the most effective device to stimulate conversion, but also multiple orders within the same session. Next, we discuss the contributions and implications of the present research.

3.6. Discussion

Findings show that customers who shop using their tablet and who have visited a clearance page are more likely to complete an order in a shopping session, followed by those who shop using PC. Smartphone is the least used device for conversion to sale. Furthermore, the value of the items in the cart is greatest for shoppers who are using their tablet, and the lowest for shoppers who are using their smartphone. Thus, in addressing our key research question, tablets are of greatest value to help convert shopping into buying and to stimulate multiple orders during the same shopping session.

That said, there is still something impressive and perhaps unexpected with respect to the power of those who shop from their PC when considering multiple orders. Customers with items in their cart from a previous session using PC are more likely to purchase in a current session in comparison to other devices (i.e., smartphones, tablets). Further, customers who read reviews of items during their browsing session via PC are more likely to purchase than when other devices are used. Thus, even though our research findings demonstrate that tablet is the device of choice and leads to the highest increase in the likelihood of multiple orders per session, PC is still very important especially when consumers exhibit more goal-directed behaviors. Returning to an existing cart can be seen as a more purchase-driven behavior and may be an indication of a stronger interest in immediate purchase, especially when conducted on a PC. While multiple orders within a session are fairly rare, they are of great interest for marketers who are interested in behavioral loyalty. Some e-tailers may also be able to bundle multiple orders together into the same package to save on shipping costs.

3.6.1 Contributions for Marketing Scholarship

This research makes novel contributions in three ways. First, this is the first work to study cross-device e-commerce and m-commerce browsing and online shopping along with the impact of viewing clearance pages, existing items in e-carts carried over from the previous session, or reading customer review pages from a cross-device consideration. We found significant findings for each of these novel considerations on the frequency of online orders completed. Namely, we expand knowledge on the electronic shopping cart use by being the first to examine the effect of existing items in consumers' cart from past shopping sessions. This is referred to as a dynamic cart, in that the cart does not necessarily remove the items upon cessation of an online shopping session. While that term was introduced in earlier research (Kukar-Kinney and Close 2010), which was published back when static carts (carts that deleted items upon cessation of an online shopping session) were the norm, the impact of dynamic carts had not been empirically addressed for online conversions.

A second contribution over past work is that we add three new considerations to build on de Haan et al. (2018): 1) browsing patterns across various pages such as clearance pages and product reviews visited by consumers, 2) time spent on each shopping session in seconds, and 3) time of day of visit. With respect to these three areas, we find that viewing clearance pages and product reviews leads to higher frequency of completed orders for either specific devices (product reviews for PCs) or all devices (clearance). Next, we find a significant impact in that when online shoppers spend more time during their shopping session they have greater frequency of completing the order. Furthermore, we found that morning shoppers most frequently complete their orders, followed by lunch-time shoppers. Both morning and lunch-time shoppers have positive and significant effects on orders completed. Meanwhile, interestingly we find that night-time online shopping has a significant and negative impact on order completion. Thus, night-time shoppers have a significantly lower chance of order completion and are more likely to simply be browsing at night.

Last, we expand the operationalization of cross-platform e-commerce and m-commerce to include PCs, tablets and smartphones and use big data for development of the empirical model. Previous work, again with the exception of Xu (2017) and de Haan (2018), tends to focus on one of the three platforms, not comparing them or considering how they work potentially in tandem or as substitutes in online shopping.

Theoretically, we help explain the process of the online path to purchase across different device types (see Figure 2 for an overview). Analyses with big data help us generate a way to explain online consumer behavior with actual behavior, and the stages that consumers take in their online path to purchase as well as the device type considerations.

3.6.2 Managerial Implications

Our work addresses a hot topic for managers who are keenly interested in using big data (Verhoef et al. 2016) to understand the online path to purchase in e-commerce and the role of consumers' device types when they shop. These findings should lead to updated strategies with respect to their e-commerce and m-commerce offerings and integrated brand promotions of such in the following areas: device type, dynamic e-carts, time of day for advertising, product review pages and clearance pages, and e-cart value.

Device type. Our findings suggest that marketers and managers should place emphasis on the tablet, as advertising to consumers who shop from tablets may be especially valuable. This recommendation is based on our finding that conversation rate is highest when consumers shop via tablet, followed by PC and then smartphone, as well as the fact that the value of the items in the e-carts are highest for tablets.

Dynamic e-carts. Our findings suggest that marketers should use dynamic carts in their e-commerce platforms rather than static carts (which have the item or items disappear upon the online shopping session ending). This suggestion is based on our finding that customers with items in their e-cart from the last online shopping session using a PC completed more multiple orders in a session than any other device (i.e., smartphones, tablets).

Time of day to advertise for e-commerce. Our findings suggest that marketers should send promotional ads on online shopping platforms in the evening, as this is the time of day where consumers seem to need an extra nudge to go from browsing to purchasing. This implication is based on our finding that for every visit made during the evening, the frequency of orders drops. Again, shopping online at night, regardless of device type, has the lowest conversion as consumers tend to browse online at night. Thus, evening media investments are suggested where appropriate. Specifically, time-limited deals available for a very short time period (e.g., an hour) may help spur the night shoppers into action.

Product review pages and clearance pages. Our findings suggest that marketers can encourage consumers to read product reviews, especially from more fixed devices such as PCs. Overall, without taking device type into account, there is no significant behavioral evidence with big data that viewing product reviews increases the frequency of orders purchased. However, when consumers read the reviews from a PC, the effect of reading product reviews is intensified for conversions. Related, our findings suggest that marketers can further encourage views of clearance webpages. While this seems counter to standard advice to encourage higher end pricing, our results demonstrate that visiting a clearance page has a significant positive effect on frequency of purchases made.

E-cart value. Our findings provide mixed evidence with respect to encouraging high value of items for online shopping. A caveat to that traditional wisdom is that when the value is high, the conversion rate for the online order completion is lower, possibly because of a “sticker shock” when seeing the total price for the cart. This means that managers should be aware that the larger the worth of the e-cart, the less likely that order will be completed.

3.7 Limitations and Areas of Future Research

While the present work builds on past work and generates original findings, it does have limitations. One, while past work on e-commerce has shown the motivations for online shopping (Wolfinbarger and Gilly 2003) and e-cart use (Close and Kukar-Kinney 2010), the behavioral data shows the actions and the online consumer actual behaviors, but does not include aspects such as attitude. Another important topic that is worth including in future work is the role of device types on omnichannel aspects (Verhoef et al. 2005; 2007). Further, while we used big data based on a multi-national sample spanning hundreds of brands and several different countries, the data does not include purchasing services online. Findings could be different if we were modelling the online path to purchase for services or experiential goods, such as sporting tickets or vacations. Future research can replicate this work in the context of services or experiential goods subject to data availability. It would be managerially and theoretically relevant to examine if any differences in findings would occur when a tangible product is not what is being browsed for or purchased online.

Chapter 4 Differences in Online Shopping Cart Abandonment across Mobile and Non-Mobile Devices

Abstract

A significant amount of past research has investigated why consumers abandon their cart but no papers have examined the differences in cart abandonment between mobile and non-mobile devices. Examining the differences in cart abandonment across mobile and non-mobile devices can help firms to formulate a strategy to decrease abandonment and increase profitability. In this paper, we seek to understand the differences in the nature and degree of cart abandonment between mobile (smartphone/tablet) and non-mobile (PC) devices. In doing so, we formulate the problem as a two-choice process using a recursive bivariate probit (RBP) model. The first step is the choice of device to employ for online shopping (mobile vs non-mobile) and the second concerns whether to abandon the cart or not. We leveraged a unique dataset from a large multinational clothing, footwear and home products retailer involving over 165,613 unique customers across 10 browsing sessions. Our results show the following: i) 1% of mobile users that abandon their cart switch to a non-mobile device during the session when they start from a mobile device. However, 5% of non-mobile users that abandon their cart switch to a mobile device during the session when they start from a non-mobile device; ii) The basket value per device on a mobile is similar to on a non-mobile and its effect on cart abandonment; iii) 76% of cart abandonments were accounted for by those employing a mobile channel for shopping; iv) The cart abandonment rate was 44% higher on mobile (*vs*. non-mobile) devices when the consumers visited the site during the evening; (v) Consumers who read positive reviews were equally likely to abandon their cart using mobile (*vs*. non-mobile) devices and (vi) Consumers using mobile devices were more likely to abandon their cart if they experienced a high number of failed attempts to use a credit card than when experiencing the same number of failed attempts to use a credit card on non-mobile devices. Based on the findings, we discuss

how organizations can enhance their online multi-device strategy to increase shopping to buying conversion rates.

Keywords: Online consumer behavior, cart abandonment, e-commerce, mobile device, non-mobile device, retail, recursive bivariate probit model

4.1 INTRODUCTION

A number of papers have explored the motivations for online shopping cart use (Close and Kukar-Kinney 2010; Close, Kukar-Kinney and Benusa 2012) and the determinants of online shopping cart abandonment (Albrecht, Hattula and Lehmann 2017; Huang, Korfiatis and Chang 2018; Kukar-Kinney and Close 2010). However, none of them have examined the factors influencing the extent to which the cart abandonment is different depending on whether a mobile or non-mobile channel is used.

Device shopping refers to the purchase of goods or services from devices such as smartphones, tablets or a PC via a wireless network (Wu and Hisa, 2004). These forms of shopping have become very popular among shoppers. De Haan et al (2018) explained that the rapid increase in the use of mobile devices has changed how consumers behave and shop online. With the increasing sales and penetration of smartphones and tablets, combined with the declining sale of PCs (mobile devices were employed for 65% of the total time that consumers spent on digital media in 2016), the PC is fast becoming a “secondary touch point” for an increasing number of digital users (Sterling 2016).

Cart abandonment behavior, specifically leaving items in a device-shopping cart without completing a purchase session, is a vital aspect of device shopping that interests both marketers and academics. An industry report by Stephan Serrano (2020), an expert on ecommerce personalization tools, shows a cart abandonment rate across all sectors of 75.6%. There are some indications that the cart abandonment rate when using a mobile (cf. non-mobile) channel may be higher (Kibo, 2016) resulting in higher economic losses for the retailer.

Our results confirm that mobile devices account for 76% of all abandonments. Although online shopping cart abandonment has been studied, the differing factors affecting the mobile and non-mobile shoppers' decisions to abandon their cart remain relatively unexplored. The causes of mobile shopping cart abandonment may not be the same as those that apply in the non-mobile context because mobile purchasing has unique limitations and merits. For example, the small screen makes the device light and portable but service providers may limit the information search flexibility. Specifically, Ghose, Goldfarb and Han (2013) found that the small screen size of mobile phones amplifies ranking effects as users incur a higher cognitive load from the information chunking. Furthermore, small screen sizes can negatively influence navigation and inputting capabilities (Chae and Kim 2004), making it hard to locate pertinent information when making online shopping decisions.

According to Xu et al (2017), the PC channel is characterized by constrained access capabilities as it restricts Internet usage and access to places that have suitable hardware and an appropriate Internet connection (Bang et al. 2013). On the other hand, ubiquitous Internet access offered by the mobile channel overcomes this limitation. This supports time critical activities, providing instantaneous information access and facilitating immediate transactions (Jung et al. 2014; Venkatesh 2003). The ubiquitous nature of the mobile channel also generates more travel-related discretionary time for commuters (Ghose and Han 2011) who tend to consume mobile video content to pass time, manage solitude and disengage from others (O'Hara et al. 2007). Despite these benefits, the limitations stemming from the processing capability and storage capacity of mobile devices, along with the bandwidth constraints of the mobile network, can degrade the access quality to Internet services (Napoli and Obar 2014).

An empirical study conducted by Xu et al (2017) showed a tendency to switch from mobile to non-mobile devices largely during the hours when the users are at home. They also show an increase in browsing activities on mobile phones during commuting hours, indicating that the complementary impact of mobile adoption works by spurring users to increase their browsing activities on their mobile phones while on the move. Past the evening commute hours, the share of browsing frequency attributable to mobile (vs. non-mobile) devices falls, suggesting that users tend to switch to non-mobile devices for browsing when they are at home.

There are important reasons underscoring the need to study what motivates the differences in the nature and degree of cart abandonment between mobile and non-mobile channels. Previous work has investigated the motivations for online shopping cart use (Close and Kukar-Kinney 2010; Close, Kukar-Kinney and Benusa 2012) and the determinants of online shopping cart abandonment (Albrecht, Hattula and Lehmann 2017; Huang, Korfiatis and Chang 2018) separately. However, there is a need for research to employ a model that jointly examines online shopping devices and cart abandonment while examining the nature of the differences in online shopping cart abandonment across mobile and non-mobile devices respectively. One of the reasons for this is the simultaneous nature of the cart abandonment process. Before a consumer abandons their online cart, their journey usually starts by deciding whether to use a mobile or non-mobile device. Consequently, formulating this as a two-choice process is appealing.

To our knowledge, the extant research has not examined online shopping cart abandonment across mobile and non-mobile devices with field data or e-commerce click stream data. Despite the established managerial and economic importance of cart abandonment, there is a dearth of literature reporting the empirical research concerning consumer behavior related to online shopping cart abandonment and the behavioral factors that differentially affect mobile and non-mobile cart abandonment specifically.

Given the managerial importance of online shopping cart abandonment and the gap in the literature, the aim of this research is to contribute to the knowledge of consumer behavior regarding online shopping carts and different device channels. In order to achieve this aim, a recursive bivariate probit (RBP) model was employed to capture the consumers' needs in order to deal with two choices. The first is which device channel to employ for shopping (mobile *vs*. non-mobile) and second is whether or not to abandon their cart. We leveraged a unique dataset from a large multinational clothing, footwear and home products retailer involving over 165,613 unique customers across 10 browsing sessions.

We found that the rate of cart abandonment is significantly higher when using mobile *vs.* non-mobile devices. In addition, timing plays an important role in determining when a cart is likely to be abandoned. In particular, we have examined the temporal aspects of the users' browsing patterns and its effect on cart abandonment across different devices. We found the rate of cart abandonment to be higher among those who visited the website on their mobile (*vs.* non-mobile) in the evening. In addition, because of mobile enhanced digital mobility, mobile devices allow customers to engage with a retailer through wireless sessions under all temporal and spatial situations (Shankar et al. 2010). This flexibility due to using a mobile device through wireless technology suggests that shoppers on mobile device will have higher basket value over those using a non-mobile device, thereby resulting in higher cart abandonment on a mobile (*vs.* non-mobile) device. We did not find there to be any differences in the basket value per device on mobiles (*vs*. non-mobile) and its effect on cart abandonment. Our finding suggests that as the rate of attempted credit card failure increases, the rate of cart abandonment becomes higher on a mobile (*vs.* non-mobile).

Unexpectedly, we did not find evidence to suggest that shoppers who read positive product reviews using mobile devices are less likely to abandon their cart than those using non-mobile devices as suggested by the extant literature (Ananthakrishnan et al., 2020; Chevalier and Mayzlin 2006). We discuss how our findings facilitate the development of a multi-device strategy for online retailers to use to help reduce cart abandonment across mobile and non-mobile devices.

The rest of this paper is organized as follows. In section 4.2, we examine both the theoretical and empirical literature related to cart abandonment. In section 4.3, we provide an overview of the contextual settings for our data and we also describe the empirical models used in our analyses. The results, alongside various robustness checks, are presented in section 4.4. In section 4.5, we discuss our key findings and highlight the managerial insights that can be derived from these findings. In section 4.6, we outline some of the opportunities for further research.

4.1 Relationship with the Extant Research

We examined the papers in leading marketing journals for all years until 2020 on the following topics: online cart abandonment, e-commerce, mobile devices and non-mobile devices. Table 1 synthesizes the relevant literature in marketing and delineates where our paper is positioned. The purpose of this review was to determine the gaps in the existing literature. It is clear from Table 4.1 that there is only one paper that has examined cart abandonment on mobile devices. This employed survey data. In addition, cart abandonment on non-mobile channels has received virtually no attention in the literature.

Several researchers have suggested that structuring the shopping environment to appeal to people that experience stress during shopping may reduce the rate of shopping cart abandonment. In particular, Albrecht, Hattula and Lehmann (2017) examined the relationship between consumer shopping stress and cart abandonment and found that a consumer's response to shopping stress depends on their motivational orientation. In particular, the greater the in-store stress, the more likely task-oriented consumers are likely to abandon the trip without making purchases. The results of their four related studies showed that for customers with task-oriented motivation, a monotonic relationship exists between shopping stress and purchase abandonment. This is consistent with their perception of stress as a threat to purchase goal. However, for recreation-oriented customers, this follows a curvilinear inverted U-shaped relationship. Purchase abandonment first increases as the level of shopping stress rises but then decreases at higher levels. Their results offer an alternative explanation why people buy or not and it also suggests approaches to structuring the shopping environment to appeal to both types of consumer.

In-store and online shopping offers customers different experiences. Despite placing items in shopping carts, online shoppers frequently abandon them across devices, an issue that perplexes online retailers and has yet to be explained by scholars. A significant amount of the previous research (Close and Kukar-Kinney 2010; Close, Kukar-Kinney and Benusa 2012) has investigated the various reasons why consumers abandon their cart. However, no papers have directly compared the differences in online cart abandonment across mobile and non-mobile devices. Examining the differences in cart abandonment across mobile and non-mobile devices has important managerial implications since it may help firms to develop a multi-device strategy to increase their shopping into a higher buying conversion rate.

Oliver et al (2003) employed online survey data for 206 respondents to determine whether the redemption of online *vs.* offline coupons (in the form of "promotion codes") differed. They found that online coupons have a significant effect on the way that shoppers react to cart abandonment. While offline coupons such as in-store vouchers are seen as customer-initiated, online coupon internet shoppers are usually prompted to enter a code towards the conclusion of the checkout process. This prompting may influence the shoppers' perception and behavior (e.g. becoming irritated), leading to the propensity to increase shopping cart abandonment. In addition to the impact of coupons, transaction inconvenience, risk and waiting time have also been found to increase the chance of consumers abandoning their cart (Rajamma et al., 2009). Kukar-Kinney et al's (2010) study employed survey data from 289 online consumers to explore the reasons for cart abandonment. They discovered that one of the key drivers for abandonment is consumers using online carts for entertainment or as a shopping research and an organizational tool. This may induce them to buy in a later session or via another channel. One key limitation of using survey data as found in the earlier studies is that it may not be as robust as using large datasets.

Xu and Huang (2015) presented a model of why and how cart abandonment occurs in the online shopping process without taking into consideration the differences that exist between the device channels. An online survey of 210 people was conducted via the online shopping website of a communication company in China. Their results demonstrated that cart abandonment was directly and positively influenced by the shoppers browsing external websites for product comparisons. Another important factor that they identified as a key driver of cart abandonment was the cost of the products in the cart.

Huang, Korfiatis and Chang (2018) were the first to relate cart abandonment directly to the use of a mobile device. They conducted two separate studies to understand why consumers hesitate to use mobile channels for shopping and thus abandon their mobile shopping carts. Their study used data from 232 responses to a survey posted for 30 days on an online forum. For the second study, they obtained a dataset from a marketing research provider totaling 226 US consumers. Their results specifically showed that emotional ambivalence is the reason why consumers hesitate at the checkout stage, leading to cart abandonment. While this set of extant works has uncovered some of the reasons for cart abandonment in general, little is known about the online shopping cart abandonment and how that relates to the mobile or non-mobile channel employed.

Consequently, we take on a hypothesis and modeling approach and interpret the findings with respect of those papers. An important gap in the literature is work on cart abandonment employing a unified framework that helps to understand how cart abandonment differs between mobile and non-mobile channels when shoppers have high basket values, are browsing in the evening, whether reading reviews on mobile vs non-mobile channels reduces cart abandonment and the number of attempted credit card failures on mobile vs non-mobile channels. It is our intention to fill in these gaps by offering additional online consumer inhibitors (Howard and Sheth 1969). These inhibitors include a lack of availability, high price, financial status and time pressure. Here we extend these inhibitors to the factors associated with mobile and non-mobile devices and examine the differences in online shopping cart abandonment when these two broad types of device are used.

[Table 4.1]

Hypothesis

Jen - Hui Wang et al (2015) theorized that mobile devices are an effective platform for customers to develop habitual interactions with a retailer because they provide convenience. This reinforces the customers' psychological and experiential state of being in a relationship with the firm. Customers who interact with a firm via mobile devices integrate its products or services into their routine. Repeated behavior has been shown to form Verplanken 2006). Mobile technology has been shown to be a superior platform for a firm to engender habitual interaction from its customers because mobile devices themselves are an integral part of the customers' daily routines (Jen - Hui Wang et al., 2015). Habitual interactions can be beneficial for a brand in a competitive environment. As customers become dependent on their habits, they rely on automatic thinking and cease to consider alternatives (Fazio, Ledbetter, and Lowles-schwen 2000). Other research has also shown that touch interfaces on mobile devices enhance the users' perceived sense of product ownership and thus increase their purchase intention (Brasel and Gips 2014). Additionally, habitual behavior is reinforced by contexts and past performances (Neal et al. 2012). Thus the more adept the customers are at interacting with a firm, the more likely they will continue to do so in the future. When customers interact with a retailer via mobile devices in addition to non-mobile channels, the firm gains increased opportunities to engage with them. The extant literature shows that customers who shop from multiple channels have a higher customer lifetime value (Ansari, Mela, and Neslin 2008; Kumar and Venkatesan 2005). One explanation for this positive impact on the retailers' revenue is that cross-channel availability increases access. Because of their enhanced digital mobility, mobile devices allow customers to engage with a retailer through wireless sessions under all types of temporal or spatial situations (Shankar et al., 2010). As such, customers can utilize a retailer's offerings as they go about their daily lives, increasing the value of the items in their basket. They can shop and modify their baskets during the additional sessions that are made available through wireless technology. We therefore posit the following:

H1: *Cart abandonment is higher on mobile (cf. non-mobile) devices due to high basket value on mobile device.*

Time of the day has been shown to be an important driver of cart abandonment because it helps to determine which devices are used (de Haan et al., 2018). An empirical study conducted by Xu et al (2017) showed that the substitutive impact of non-mobile devices over the mobile channel largely manifests during the hours in which users are at home. Specifically, their empirical findings showed that during commuting hours (0700 to 09:00 hrs and 1700 to 1900 hrs), mobile devices are associated with the highest conversion rates while non-mobile devices are associated with higher abandonment. They found that in the evening, the share of non-mobile device use increased, suggesting that users tend to use non-mobile devices when they are at home. Interestingly, they saw an increase in the mobile browsing frequency from midnight to 0300 am. In addition, Rob et al (2018) analyzed a click stream dataset on the browsing behavior of hotel website users. They showed that users access the website during working hours using non-mobile devices. During the evening, they reported that the number of users that abandoned their cart on a non-mobile device is 50% higher than those who use mobile devices. Online retailers need to know the time of the day when customers are likely to be on a mobile or non-mobile device so then they can send an email reminding the consumer about the products in their shopping basket. This approach could help retailers to reach consumers at the right moment on the right device.

In light of the existing evidence that more shoppers abandon their carts using non-mobile devices in the evening than when using mobile devices in the evening, we tested the following hypothesis:

H2: *Cart abandonment is higher on non-mobile (cf. mobile) devices when shoppers browse in the evening.*

Given the overwhelming prevalence of mobile technology, gaining a broad understanding of how the use of mobile devices is influencing people's perceptions of the content that they view online is an increasingly important research objective. However, despite the substantial proliferation of devices, relatively little is known about the relationship between devices and consumer behavior. A common use of devices is creating user-generated content (UGC) and disseminating it through online platforms. This includes posts on social networks, sharing photos and videos through apps, and rating and reviewing products and services on online review sites to increase sales. However, instead of considering the differences that arise in the actual content that consumers produce based on whether or not they use devices to write the UGC (e.g. the emotionality of UGC; Melumad, Inman, and Pham 2019; Ransbotham, Lurie, and Liu 2019), Grewal et al (2019) considered how the knowledge crafted on a device affects consumer attitude and purchase intention. This is practically relevant because some of the popular platforms (e.g., TripAdvisor) explicitly make consumers aware whether a user posted a review from a mobile device or not (with a "via mobile" icon adjacent to the review). Although this type of cue might seem innocuous, they found that this knowledge can positively influence the consumers' evaluation of an online review. Specifically, they found that knowing whether a review was written on a mobile (vs. non-mobile) device - holding the actual content of the review constant - can lead to a higher purchase intention. In addition, the previous work has found there to be significant economic effects due to online mobile reviews in a variety of settings, including the effect of reviews on product sales and purchase decisions (Ananthakrishnan et al., 2020; Chevalier and Mayzlin 2006), the effect of ratings on hotel bookings (Ghose et al. (2012)) and the effect of online reviews on movie box office performance (Chintagunta et al. (2010) & Dellarocas et al. (2007)). Upon recognizing the importance of online reviews in consumer decision-making, subsequent papers on the information systems and marketing have developed algorithmic techniques to mine and predict the usefulness and subjectivity of online reviews based on various parameters. This provides context to our understanding regarding which reviews have a higher probability of influencing a customer's purchase decision (Ghose et al. (2012) & Archak et al. (2011)). Previous research has shown that the act of simply reading positive reviews on a mobile (*vs. non-mobile*) device can lead to a greater purchase intention. Consequently, it is surprising that to the best of our knowledge, no research has been conducted to explore the impact that reading positive product reviews on a mobile device can have on cart abandonment. Consequently, we tested the third hypothesis:

H3: *Reading positive product reviews on a mobile channel (cf. non-mobile) reduces the rate of cart abandonment.*

Consumers seek to ensure that they are making the right purchase and therefore they strive to reduce the risk of making an incorrect or inappropriate purchase online (Moorthy, Ratchford, and Talukdar 1997). In this stage in the decision process, the information processing needs are more fine-grained, detailed and deliberate in order to lower the risks involved and to make the right decision (Moe 2003). Fine-grained information processing needs might require the larger and higher-resolution screen of a non-mobile device. Indeed, the need to input detailed payment information without making a mistake may also make non-mobile devices more suitable. Similarly, the increased perception of security and privacy risks when using a mobile device (Ananthakrishnan et al., 2020) may lead consumers to choosing a non-mobile device for completing their purchase transactions. In addition, situational factors such as being outdoors or in a public place may increase the perception of risk such as the risk of someone reading over their shoulder while the shipment and billing information or credit card details are being entered. If the transaction risk on a mobile device is high, the customer may well wait to access a device that is perceived as being more secure (Chin et al., 2012). Consumers using a mobile device are more likely to abandon their cart if they experience a high number of failed attempts to use a credit card than when experiencing the same number of failed attempts to use a credit card on a non-mobile device. Consequently, we tested the following hypothesis:

H4: *Cart abandonment is higher on mobile (cf. non-mobile) devices due to high number of failed attempts to use a credit card on mobile devices.*

In the next section, we discuss the data, the modeling approach and how we account for endogeneity and heterogeneity in our model framework.

4.3. Methodology

4.3.1 Data Description

To test our hypotheses, we employed a large dataset consisting of individual level click stream session data from a large European online retailer. The online retailer is a British clothing, footwear and home products retailer. Specifically, the data comes from its online retail site and includes over 165,613 unique customers across 10 sessions. For each session, the data includes detailed information regarding which device was used to visit the website and on which device cart abandonment occurred. It also included a unique ID number for each session, session date, browser name, basket value, browsing behavior across time, browsing pattern (e.g. read reviews) and credit card failure in addition to the number of attempts across all devices. The online retailer defines ‘a session’ as one continuous period in which the customer is active on the website. A session starts when the customer enters the online retailer's website and ends when the customer actively leaves the website or when the customer is inactive (e.g. has not visited a new page on the retailer's website or has not clicked on a link on the website for more than 30 minutes).

4.3.2 Methods

4.3.2.1 Variables

The dependent/target variable employed in this study is a binary indicator and took on a value of 1 if the customer did not complete an order (shopping cart abandonment) and 0 otherwise. For the analysis, we used the following tracked behaviors as our IVs: (a) if customer used a mobile *vs.* non-mobile device, (b) the basket value of the customers on a mobile *vs.* non-mobile device, (c) if a customer visited the site during the evening on a mobile *vs.* non-mobile device, (d) if a customer read positive product reviews on a mobile *vs.* non-mobile device and (e) the number of credit card failed attempts during payment on a mobile *vs.* non-mobile device. We also included the interaction terms between variables (a) and (b), (c), (d) and (e) respectively.

4.3.2.2 Econometric Model

When modeling two jointly determined binary choices, we used recursive the bivariate probit (RBP) model to test our hypothesis, similar to Bell, Corsten and Knox (2011). The RBP is a system of two probit equations that allows the errors terms to be correlated and the binary dependent choice in one equation is an endogenous regressor in the other equation. In our case, we employed a mobile *vs*. non-mobile binary endogenous conditioning factor and utilized the

RBP model as if the consumers face two choices. First, the device to use (mobile *vs*. non-mobile) and second, whether to abandon the cart.

Consider $device\_type^{*}_{1i}$ and $cart_abandonment^{*}_{2i}$as latent variables representing the probability of cart abandonment and the browsing using a particular device channel (i.e. mobile *vs*. non-mobile device). $device\_type^{*}_{1i}$ and $Cart\_abandonment^{*}_{2i}$ are dichotomous variables that observe the following rule:

$device\_type^{*}_{1i}$ = 1 if $device\_type^{*}_{1i} > 0$
0 if $device\_type^{*}_{1i} \leq 0$

$cart\_abandonment^{*}_{2i}$ = 1 if $cart\_abandonment^{*}_{2i} > 0$
0 if $cart\_abandonment^{*}_{2i} \leq 0$

We estimated the RBP model using the Full Information Maximum Likelihood (FIML) method. Our system of equations was as follows:

$$cart\_abandonment^{*}_{2i} = \beta_1 \text{ X } device\_type^{*}_{1i} + \beta' OtherIndependentvariables \quad (4.1)$$

$$device\_type^{*}_{1i} = \beta' \; Other\ Independent\ variables (2)$$

All other IVs and $device\_type^{*}_{1i}$ are the vectors of the exogenous variables. Here, the set of $Other\ Independent\ variables$ regressors are partly common to both $cart_abandonment^{*}_{2i}$and $device\_type^{*}_{1i}$. The error terms are assumed to be independent and identically distributed (iid) as bivariate normal.

Choosing such a model implies that there is a correlation between the error terms. We imposed the conditions Var (u_{1i}) $=$ I and Var (u_{2i}) $=$ I, where I is the unity matrix. Cov (u_{1i} , u_{2i}) = ρ I $\neq$ 0, with ρ reflects the decisions that online shoppers face i.e. which device to use and whether to abandon the cart or not. The operationalization of the variables used in our case is similar to how De Haan et al., (2018) and Moe (2003) did the same.

To enhance the validity of our results and to improve the confidence in our estimation approach, we conducted robustness analysis using a Bayesian framework. The Bayesian framework provided multiple benefits. It allowed us to make probabilistic statements about the parameters for each individual, provided richer information on the model's parameters by utilizing a posterior distribution rather than point estimates as well as made more direct inferences about the effects (Rossi and Allenby, 2003). Dyachenko and Allenby (2015) examined the traditional approach to mediation analysis based on the statistical significance of the coefficients in the regression models. The experimental results showed that these models are neither necessary nor insufficient when it comes to claim the presence of mediation for specific individuals. The Bayesian framework allows us to incorporate heterogeneity into the mediating behavior. To obtain the parameter estimates of the Bayesian model, we used the Bayesian Markov Chain Monte Carlo (MCMC) method which draws from the posterior distribution of the parameters obtained and used as a basis for inference.

Using Gibb samplers with conjugate priors (Rossi and Allenby, 2003), we estimated a joint posterior distribution of the parameters for our model. The conjugate prior allowed us to analytically derive and directly sample the conditional posteriors of the state parameters. Conversely, given the state parameters, the system became a multivariate probit model with hierarchical priors, thus we sampled the data directly using the standard MCMC method (Rossi and Allenby, 2003).

4.3.2.3 Controlling for Endogeneity and Consumer Heterogeneity

It is widely known that endogeneity in marketing models can lead to biased coefficient estimates (Germann, Ebbes, and Grewal 2015; Papies, Ebbes, and Van Heerde 2017). A common way to correct for endogeneity involves using independent variables (IVs). In an ideal case, IVs allow for unbiased estimates that can be implemented through either two-stage least squares (2SLS) models that used estimated values from first-stage regression for the possibly endogenous variables (Angrist and Pischke 2009; Germann, Ebbes, and Grewal 2015; Wooldridge 2010) or via an equivalent control function approach that includes the first-stage regression residuals as the control variables in the main model (Petrin and Train 2010). The current recommendations stress that researchers should first carefully exploit the control variables and panel structures in the data sets to control for unobserved effects before deciding on the use of IVs (Germann, Ebbes, and Grewal 2015; Papies, Ebbes, and Van Heerde 2017). Recursive Bivariate Probit (RBP) is a system consisting of two probit equations that allow the error terms to be correlated. The binary dependent choice in one equation is the endogenous regressor in the other equation. This allows us to account for the endogeneity of the treatment ("mobile" vs "non-mobile") in our case (Filippini et al., 2018). To account for the lack of demographic information in our data, we conducted an experiment to determine if the lack of demographic variables will have any effect on our result. We selected consumers that used the same device throughout their online journey without switching devices and compared the results with our final model. We accounted for other additional control variables using active sessions on the website in a similar manner to Mallapragada et al (2016). The online retailer defines a session as one continuous period in which the consumer is active on the website. A session starts when the consumer enters the online retailers' website and ends when the consumer actively leaves the website or when the consumer is inactive (i.e. they have not visited a new page on the retailer's website or have not clicked on a link on the website for more than 30 minutes). We are interested only in active sessions that belong to the same path to purchase or cart abandonment ("the multiple touches a customer makes before a conversion, as defined by Li and Kannan (2014)). Introducing active session as a control variable allows the model to take into consideration only the activities undertaken while the consumer is active on the website to disregard any idleness or inactivity. We added a series of control dummy variables that denote the "sequence" or stage of each session in the purchase process. This is an indicator of how late a customer is in their online journey/funnel. In addition, similar to the findings of Mallapragada et al (2016), we added other control variables such as the number of product pages visited by the consumer after reading product reviews to control for reviews as a whole. Having an existing product in the cart before the current session might influence the likelihood of purchase since prior research might have been done by the shoppers before the current session (de Haan et al., 2018). We controlled for consumers that had existing items in their cart before the current session.

4.4. Empirical Results

4.4.1 Descriptive Statistics

In Table 4.2, we provide details of the descriptive statistics for the IVs when mobile and non-mobile devices are employed. When categorizing consumers by the device used, as we do in Table 4.2, the mean number of shoppers that abandon their cart per device is higher on mobile devices. In addition, the mean number of shoppers that visit during the evening per device is also higher on mobile devices as shown in the second row of Table 4.2. Interestingly, the earlier research by Rob et.al (2018) analyzed the click stream data of browsing behavior of hotel website users. They reported that the number of users using non-mobile devices in the evening is much higher than the use of mobile devices.

In the context of user-generated content (UGC), mobile devices have been considered an easier way of allowing consumers to review products in a timely manner (Xu et al, 2017). Our summary statistics of the variables in Table 4.2 shows that the mean number of positive product reviews read per device on non-mobile devices is higher than on mobile devices. One would have imagined the opposite as consumers assume that mobile reviews are more physically effort-heavy to craft and subsequently they equate this greater perceived effort with the credibility of the review.

The descriptive statistics in our data in Table 4.2 shows that the mean number of credit card attempted failures per device on mobile device is higher than on a non-mobile device. Lastly, the mean basket value per device on a non-mobile is higher than on a mobile device.

[Table 4.2]

In Table 4.3, we present the results of calculating the correlations between the IVs. These results indicate that the correlation between these variables is low, suggesting that multicollinearity (the variance inflation factors in our models are below 10) is not a problem.

[Table 4.3]

4.4.2 Model Results

In order to test our hypotheses, we created a base model (i.e., Model 1) with our control variables (active seconds, product pages, existing cart, cart use and session dummy variables) and then compared that with the performance of the proposed model which incorporated all of the IVs alongside the control variable (Model 2). Our analysis is similar to that of Yli-Renko et al (2008). We first centered the variables of the interaction terms to reduce the multicollinearity (Aiken and West 1991). We have provided the model fit criteria and details regarding model comparison in Table 4.4. We present the estimates obtained from proposed model (M2 (RBP)) in Table 4.5.

The results focus on the effects of device type on cart abandonment in addition to basket value, time of visit, positive reviews read and the number of failed credit card attempts.

Main Effects on Cart Abandonment: The main effect of basket value is significant and positively associated (0.119, $p < .01$) with cart abandonment as shown in Table 4.5. All else being equal, it seems that shoppers with a high basket value are more likely to abandon their cart. Our results show strong significant effects for the consumers browsing in the evening (between 18 -23 hrs) related to cart abandonment with a significant coefficient (0.065, $p < .01$). A further result concerns the impact of reading positive product reviews on the choice of device channels. There is no significant effect due to reading positive product reviews on cart abandonment overall. This is contrary to Ananthakrishnan et al (2020) who suggested that there might be an effect from positive product reviews on online consumer purchase behavior. Our analysis is the first to link the number of credit card failure attempts and cart abandonment. This is reasonable since credit card payments are blocked when the card PIN is entered incorrectly 3 times for security reasons. This is to safeguard the consumer's funds in case they lose their card and someone tries to use it without their authorization. As shown in Table 5, all consecutive credit card failed transaction attempts irrespective of whether a mobile or non-mobile device is used for payment do not have a significant effect on cart abandonment.

Moderating Role of Device Type on Cart Abandonment: The interaction effects of device type and basket value on cart abandonment are not significant which does not support our H1 as shown in Table 4.5. In our sample data, the average size of basket value per device on a non-mobile was £33.204 while it was £26.079 on a mobile device. Even though our findings suggest there to be no significant interaction effects between device type and basket value on cart abandonment, one would expect a mobile device to have a higher basket value per device in comparison to non-mobile devices due to the ubiquitous nature of a mobile device (Xu et al., (2017). We also identified a significant negative interaction effect in relation to device channel and visiting during the evening (-0.011, $p < .01$). Device channel can therefore serve as a moderator of the effect of visiting during the evening. In Figure 4.1A, the interaction plots show that on average, cart abandonment is higher on mobile devices than non-mobile devices when shoppers browse in the evening which does not support the H2 hypothesis. This is different from the findings of the earlier research by Rob et al (2018) which concluded that cart abandonment is higher on non-mobile devices in the evening. An alternative explanation for our finding is that even though a significant number of consumers might decide to substitute their non-mobile device for a mobile device during the evening as suggested by Xu et al (2017), other consumers might continue on their mobile device journey for convenience (Close et al., 2010). We also identified a non-significant interaction effect due to device type and reading product reviews as shown in Figure 4.1B which does not support H3. Thus device type does not serve as a moderator of the effect of product reviews. Using the current field data, we can provide evidence that shows that viewing positive product reviews either on a mobile or non-mobile device does not affect cart abandonment. However, device type has a significant interaction effect with credit card failed attempts. In Figure 4.1C, the interaction plots suggest that as the attempted card failure rate increases, the rate of cart abandonment becomes higher on mobile (*vs.* non-mobile). This supports H4.

The earlier findings by Ananthakrishnan et al (2020) suggest that the reason why consumers choose a non-mobile device for completing their purchase transactions is the increased perception of security and privacy risks. In addition, situational factors such as being outdoors or in a public place may increase the perception of risk on a mobile such as the risk of someone reading over their shoulder while the shipment and billing information are entered.

[Figure 4.1]

4.4.3 Robustness Analysis

When a large quantities of data is analyzed, it is clearly important that the effects are both statistically significant and managerially substantive (Verhoef, Kooge, and Walk 2016). To enhance the validity of our results and to improve the confidence in our estimation approach, we estimated an alternative model using a Bayesian framework. This serves as a robustness check for the RBP model as it takes a strict view of mediation by employing the property of conditional independence to construct the model likelihood. The results of estimating this model are shown in Table 4 and our parameter estimates are displayed in Table 4.5. The results of this MCMC analysis are similar to those from the estimation of the RBP model discussed. In particular, the model estimates are similar in terms of direction and magnitude. This provides confidence in our results.

4.4.4 Variable sensitivity

To further validate our main results, we conducted an additional test to assess the plausibility of the alternative explanation of our results when we operationalized the variable "card failure attempt" as binary rather than counted as shown in Table 4.6. We tested the effects on cart abandonment and device type using both the RBP and MCMC models. The coefficient estimate as shown in Table 6 for card failure attempts is not significant (-0.096, $p > .01$). This coefficient is similar to when we operationalized the variable as counted as shown in Table 4.5. We selected the model with "card failure attempt" as counted because we can properly examine the incremental interaction effects of the credit card failure attempts on device type and cart abandonment. As shown in Figure 4.1C, we were able to visualize the incremental interaction effects which suggests that as the rate of attempted card failure increases, cart abandonment is higher on mobile (*vs.* non-mobile).

[Table 4.4, Table 4.5 and Table 4.6]

4.4.5 Controlling for Demographic Variables

To account for the lack of demographic information in our data, we conducted an experiment to determine if the lack of demographic variable will have any effect on our results. We selected consumers that used same device throughout their journey without switching devices and compared the results with our selected model. A sample of 10,513 consumers was used for this experiment. We have presented the results in Table 4.7.

Main Effects: We compared the results for the main effects of the selected model (Table 4.5) with our experimentation model in Table 4.7 in order to see if the effects are similar or different. All coefficient effects have similar signs and effects except for consumers visiting in the evening, which did not have a significant effect on cart abandonment.

Moderating Effects: All of our interaction effects have similar coefficient effects as shown in Tables 4.5 and 4.7 for comparison. These results confirm the robustness of our final model and also allowed us to control for the lack of demographic variables in our data.

[Table 4.7]

4.5. Discussion

While preparing to shop online, consumers encounter a range of inhibitors (Howard and Sheth 1969) which may trigger them to abort the process and abandon their cart. These inhibitors include a lack of availability, high price, financial status and time pressure (Howard and Sheth 1969). Here we found there to be additional factors associated with these inhibitors that are in turn associated with the differences in online shopping cart abandonment when mobile and non-mobile devices are used. Our field data shows that the time of a website visit has an impact on cart abandonment. Specifically, consumers using a mobile device in the evening (between 1800 to 2300 hr) were found to be more likely to abandon their cart than those using a non-mobile device in the evening. Interestingly, when we tested different hours of the day such as before work (between 6 to 8 hr), during lunchtime (between 1200 to 1300 hr) and after work (between 1700 to 1800 hr) with the device type as a moderator, the effect on cart abandonment was not significant.

Our finding does not support earlier findings of Rob et.al (2018). Their results from analyzing click stream data on the browsing behavior of hotel website users showed an increase in abandonment during the evening when using non-mobile devices. Our findings support the earlier work of Close et al (2010) that consumers browse on a mobile device for convenience. Mobile device also have a good dimensions in terms of usability and are ubiquitous as their core features distinguish them from a non-mobile device (Ghose et al, 2012). We examined the impact of device type on customer cart abandonment behavior using empirical data. We found that there are no differences in the basket value of items on mobile (vs non-mobile) devices and their effect on cart abandonment. The extant literature (Xu et al., 2017, Jen - Hui Wang et al, 2015) suggests that due to the flexibility, convenience of usage, portability and ubiquity of mobile devices, consumers tend to have a higher basket value than non-mobile devices. Surprisingly in our sample data, the average size of the basket value per device on a non-mobile is higher than that on a mobile device. One reason to suggest this is that the participants in our data might be low spenders that might indulge in the habit of purchasing items more frequently on their mobile device while a non-mobile device is preferred for higher valued items for security purposes.

Retailers often add positive product reviews in an effort to decrease cart abandonment. However, our field data shows no evidence to suggest that consumers that read positive product reviews using a mobile device are less likely to abandon their cart than those using a non-mobile device. This contrasts with the earlier findings by Grewal et al (2019) who found that reading positive reviews on a mobile device affects the consumer attitude and increases their purchase intention. The proposed explanation is that a large number of customer reviews increases the rate of information overload and choice ambiguity which could potentially lead to consumers avoiding having to make a choice (Maslowska et al. 2017; Scheibehenne et al. 2010).

In addition, while it may be easy to shop and place items of interest in the cart on a mobile device, consumers appear to be more reluctant to complete purchases using a mobile device. Given the “on the go” nature of mobile usage and the sheer ease of adding items to the cart when on a mobile, the consumers may want to take more time to fully consider the purchase before committing to buy. Our field data shows that as the rate of attempted card failure increases, the level of cart abandonment is higher on mobile (*vs.* non-mobile). The need to enter detailed payment information without making mistakes may make non-mobile devices more suitable than mobile devices.

4.5.1. Conclusions and Implications

To our knowledge, this is the first attempt to examine the nature of differences in the degree of cart abandonment across mobile and non-mobile device channels within the same framework. In particular, a unique dataset from a large multinational clothing, footwear and home products retailer was secured and employed to help fill in this important research gap. We achieved this by leveraging the recursive bivariate probit (RBP) model which explicitly accounted for endogeneity. The flexibility of our approach allowed us to jointly model two choices in one model with the first involving the choice of device (mobile *vs*. non-mobile) and the second concerning whether to abandon the cart or not.

One behavioral implication is based around the basket value per device. We show that there is no difference in the value of the items placed in the basket value on mobile (vs non-mobile) platforms and their effect on cart abandonment. Retailers should not focus only on a mobile channel when formulating a cart abandonment strategy but have a multi-device strategy for cart abandonment in place.

We have shown that non-mobile device traffic increases during the evening as consumers continue their online journey at home. Online retailers should make it easy for consumers to pick up from where they ended the previous session irrespective of the device used.

Our research is the first to use a large click stream data on the failed attempts to use credit cards on mobile *vs*. non-mobile devices. Fraudsters and hackers use various techniques to gain access to consumer information by virtue of impersonation. Even the best security systems are able to prevent most, but not all, attacks (Bolton and Hand 2002). Organizations periodically evaluate the effectiveness of their systems against new forms of abuse and make improvements.

We found there to be a higher occurrence of failed attempted credit card payments on mobile *vs*. non-mobile devices. Consumers using mobile devices are more likely to abandon their cart if they experience a high number of failed attempts when using a credit card than when experiencing the same number of failed attempts on non-mobile devices. Our research seeks to inform both the academic and managerial understanding of this issue. Consequently, to reduce the rate of cart abandonment, retailers should ensure that they subscribe to apps that pass rigorous approval processes and are trusted by the consumer.

Retailer-provided information in the form of product information is used to decide which products to purchase. However, it appears that customer-provided information, in the form of consumer reviews, is used by consumers as a screening tool to decide which products may not perform as expected and thus, which products not to purchase. One of the limitations of the existing literature is the inability of the data to enable researchers to link cart abandonment to a user exploring online product reviews on mobile devices. Mallapragada et al (2016), for example, identified this as a gap in the literature. Our unique data allowed us to use real click-stream data to link the reviews read with cart abandonment. We do not find any evidence that consumers that read positive product reviews using mobile devices are less likely to abandon their cart than those using non-mobile devices. According to recent research by

Northwestern University's Spiegel Research Center (2017, pg., 10), it was noted that: *"Readers are skeptical of reviews that are too positive and, in many cases, a negative online review is seen as more credible. Additional research by social commerce specialist Revoo indicates that consumers spend four times as long on a site when they interact with negative reviews, with a 67% increase in conversion rate"*.

The lesson from our findings is that even if positive reviews could be used to boost sales, having more reviews is not necessarily better after a point. As online commerce evolves, it is important for firms that adopt this media to understand the role of the factors influencing cart abandonment across the different device channels. The findings presented here should go some way to helping these organizations develop effective strategies to reduce cart abandonment.

4.6. Limitations and Future Research

We investigated the differences in online cart abandonment behavior across mobile and non-mobile devices using session level data. While the field nature of the dataset allows us to track actual shopping and browsing behavior, providing an advantage over the previous research that relied on the consumers' self-reported survey data (for example, see Table 4.1), the nature of the data also presents some limitations. In particular, our data does not contain demographic information. However, we were able to control for this limitation by conducting an experiment to check if it had any effect on our results.

Future research might look at comparing the impact of mobile *vs*. non-mobile on cart abandonment using the sales outcomes. Another area to explore is the role of review valence for mobile *vs*. non-mobile on cart abandonment.

From the extant literature, the personalization of promotion offers has been shown to reduce cart abandonment. As competition in the retailing industry becomes intense, most retailers are massively discounting their products online, often to the point of harming revenues. Future research work could allow retailers to understand how smart promotion strategies can be applied based on the profile of the consumer while remaining competitive. This may help reduce abandonment across the device channels.

Chapter 5 Conclusion

This section concludes the thesis by summarizing the main contributions of the three papers and the importance of these contributions in the wider literature. Then, our discussion showed how the three papers came together to form part of a larger totality that provides new knowledge and understanding within marketing and financial markets domain area. Finally, some limitations and future research areas are identified.

In management research, it is imperative to clarify the objectives of one's research (what is the problem we want to address?) and to explicitly distinguish between issues of prediction or causality before deciding whether ML is an appropriate tool for examining this research question.

This work filled a very important gap in marketing/ecommerce and finance literature by using ML prediction and causal inference model to address these problems.

5.1.1 ML Prediction

Deep learning is making major advances in solving problems that have resisted the best attempts of the artificial intelligence community for many years, such as in computer vision and natural language processing. It has turned out to be very successful at discovering intricate structures in high-dimensional data and is, therefore, applicable to many domains of science, business and government. In addition, DNNs have been predicted to have many more successes in the near future, largely because they require very little engineering by hand. Consequently, DNNs can easily take advantage of increases in the amount of available computational ability and data. With so many successful applications of DNN, it is perhaps surprising that few studies have employed DNN to forecast financial time series (Krauss et al., 2016). Paper 1 was the first, to the best of our knowledge, to use DNN in the context of predicting the direction of stock price movements across multiple markets in order to understand to the extent to which this novel algorithm is sensitive to sample sizes. We set out to clarify if this is the case by comparing the predictive accuracy of widely used shallow methods with DNN using different forecast time horizons (daily/hourly/minute/tick).

Our results showed that data size has an effect on the relative predictive performance of DNN. SVM and one-layer NN outperformed DNN when daily and hourly data was used. However, the predictive accuracy of DNN was significantly better than that of SVM and one-layer NN when using minute level data. Previous research has suggested that DNNs perform particularly well when they can learn the underlying structure using large datasets. It is likely that this explains our results, since there was a significant increase in data size from 1,500 observations for daily prediction to the 800,000 observations for minute predictions (Bengio et al., 2007).

Interestingly, the accuracy of stock index price predictions at the tick level using DNN was not significantly better than predictions based on shallower architectures. Since many studies have advocated the use of DNN when confronted with complex, big data, one would have expected DNN to have a better predictive performance than other methods at the tick price level. Some studies have suggested that this is not always the case (e.g., Cho et al., 2016) and further research is needed, employing data other than that related to stock markets, to see if similar a finding applies.

Taken together, our results suggested that practitioners looking to include DNN as one of the algorithms they use when predicting stock price movement should first consider the complexity of the decision space, the size of the dataset and how balanced the target class is.

A second important contribution of the research is that we demonstrated that RELU activation function outperforms the TANH activation function when employing DNN across all forecast time horizons (daily, hourly, minute and tick level). Carefully tuning the hyper parameter optimization may still yield advantageous results for both activation functions. This is subject to further research using more stock markets data. In case a practitioner is oblivious regarding which activation function to apply and if cheap computing power is available, RELU should be explored extensively to gain better predictive accuracy.

A third contribution of the paper was the discovery that DNN outperforms shallower architectures at the tick level in some emerging markets. In particular, we found that predictions based on DNN at the tick level were significantly more accurate than those based on shallower architectures for the Thai, Czech and Malaysian markets. We believe DNN was able to capture the recent upward trend in their economic growth which was also reflected in the data. The high levels of predictive accuracy achieved for these markets (79%, 84% and 79%, respectively) suggested significant exploitable inefficacies in these markets.

Overall, this paper has implications for financial economics and professional finance practitioners. In particular, we provided empirical evidence that the most widely celebrated machine learning technique (DNN) does not necessarily outperform SVM and NN in all cases using large data from many major financial markets. We showed that RELU should be leveraged widely by practitioners as this was better in comparison to TANH when analyzing stock market data. Lastly, none of the algorithms we tested were better for predicting stock price indices in emerging vs developed markets, even though all the methods offer the prospect of identifying inefficiency in pricing in these markets.

5.1.2 Causal Inference

Understanding customer browsing behaviour on different devices, including their smartphone, tablet, or personal computer (PC, which includes both laptops and desktops) is important for scholars and marketers alike. To address low conversion rates, managers seek knowledge about cross-platform online consumer behaviour. This problem of low-conversation rates across devices has been unaddressed by scholars in previous literature (Close & Kukar-Kinney, 2010; and Kukar-Kinney & Close, 2010).

To address this problem, Paper 2 filled the gap by examining shopping cart impacts with click-stream data and added a cross-device perspective, as suggested by more recent marketing studies (de Haan et al. 2018; Xu et al. 2017). Our results showed that customers who shopped using their tablet and who visited a clearance page were more likely to complete an order in a shopping session, followed by those who shop using PC. A Smartphones were the least used device for conversion to sale. Furthermore, the value of the items in the cart was greatest for shoppers who were using their tablet, and the lowest for shoppers who were using their smartphone. Thus, in addressing our key research question, tablets were of greatest value to help convert shopping into buying. That said, there is still something impressive and perhaps unexpected with respect to the power of those who shop from their PC when considering multiple orders. The findings revealed that customers with items in their cart from a previous session using PC were more likely to complete an order in a current session in comparison to other devices (i.e., smartphones, tablets). Further, customers who read reviews of items during their browsing session via PC were more likely to complete an order than when other devices were used. Thus, even though our research findings demonstrated that tablets were the device of choice and led to the highest increase in the likelihood of multiple orders per session, PCs were still very important especially when consumers exhibited more goal-directed behaviours. Returning to an existing cart can be seen as a more purchase-driven behaviour and may be an indication of a stronger interest in immediate purchase, especially when conducted on a PC. While multiple orders within a session were fairly rare, they are of great interest for marketers who are interested in behavioural loyalty.

The paper makes novel contributions in three ways. First, this was the first paper to study cross-device e-commerce and m-commerce browsing and online shopping along with the impact of viewing clearance pages, existing items in e-carts carried over from the previous session, or reading customer review pages from a cross-device consideration. Each of these novel considerations yielded significant findings relating to the frequency of online orders completed. Namely, we expanded knowledge on the electronic shopping cart use by being the first to examine the effect of existing items in consumers' cart from past shopping sessions. This is referred to as a dynamic cart, in that the cart does not necessarily remove the items upon cessation of an online shopping session. While that term was introduced in earlier research (Kukar-Kinney & Close, 2010), which was published back when static carts (carts that deleted items upon cessation of an online shopping session) were the norm, the impact of dynamic carts has not been empirically addressed for online conversions. A second contribution to past work is that we add three new considerations: 1) browsing patterns across various pages, such as clearance pages and product reviews visited by consumers; 2) time spent on each shopping session in seconds; and 3) time of day of visit. With respect to these three areas, we found that viewing clearance pages and product reviews lead to a higher frequency of completed orders for either specific devices (product reviews for PCs) or all devices (clearance). Next, we found in that when online shoppers spent more seconds during their shopping session, they will have greater frequency of completing the order. Furthermore, we found that morning shoppers most frequently completed their order, followed by lunch time shoppers. Both morning and lunch time shoppers had positive and significant effects on orders completed.

Meanwhile, interestingly, we found that night-time online shopping had a significant and negative impact on completing the order. Thus, night-time shoppers had a significantly lower chance of order completion and were more likely to simply be browsing at night. Last, we expanded the operationalization of cross-platform e-commerce and m-commerce to include PCs, tablets and smartphones and used big data for empirical work. The main theoretical contribution was offering a way to extend theory of the path to purchase by taking an e-commerce/m-commerce focus with a proposed online path to purchase. Analyses with big data help us generate a way to explain online consumer behaviour with actual behaviour, and the stages that consumers take in their online path to purchase as well as the device type considerations. To a lesser extent, our findings can have implications for the theory of affordances (Gibson, 1977). Applied here, a technology affordance is what a consumer with a purpose is afforded to do with technology, and consumers have functions and features whether or not consumers recognize or use them (Majchrzak & Markus, 2012). Thus, we suggest that consumers have a technology affordance by their use of their mobile device which helps them fulfil their purpose to browse, purchase or both.

Interestingly, as marketing managers, retail practitioners, and advertisers alike face the problem of low browsing-to-buying conversion rates, understanding the determinant of online shopping cart abandonment is of paramount importance. This has been widely studied in the literature (Albrecht, Hattula & Lehmann, 2017; Huang, Korfiatis & Chang, 2018; Kukar-Kinney & Close, 2010). However, none of them have examined the factors influencing the extent to which the cart abandonment is different depending upon whether a mobile or non-mobile channel is used. There are important reasons underscoring the need to study what motivates the differences in the nature and degree of cart abandonment between mobile and non-mobile channels. Previous work has investigated the motivations for online shopping cart use (Close & Kukar-Kinney, 2010; Close, Kukar-Kinney & Benusa, 2012) and the determinants of online shopping cart abandonment (Albrecht, Hattula & Lehmann, 2017; Huang, Korfiatis & Chang, 2018) separately. However, there is a need for research to employ a model that jointly examines online shopping devices and cart abandonment whilst examining the nature of the differences in online shopping cart abandonment across mobile and non-mobile devices. One of the reasons for this is the simultaneous nature of the cart abandonment process. Before a consumer abandons their online cart, their journey usually starts by deciding whether to use a mobile or non-mobile device. Consequently, formulating this as a two-choice process is appealing.

Paper 3 addressed this problem as the first to examine the nature of differences in the degree of cart abandonment across mobile and non-mobile device channels within the same framework. In particular, a unique dataset from a large multinational clothing, footwear and home products retailer was secured and employed to help fill this important research gap. We achieved this by leveraging the recursive bivariate probit (RBP) model, which explicitly accounts for endogeneity (Filippini et al., 2018). The flexibility of our approach allowed us to jointly model two choices in one model, the first involving the choice of device (mobile vs. non-mobile) and the second, concerning whether to abandon the cart.

One behavioural implication is around basket value per device. We showed that there were no differences in the value of items placed in the basket value on mobile (vs non-mobile) and their effect on cart abandonment. Retailers should not focus only on mobile channels when formulating cart abandonment strategy but have a multi-device strategy for cart abandonment. We have shown that non-mobile device traffic increases during the evening as consumers continue their online journey at home. Online retailers should make it easy for consumers to pick up from where they ended in the previous session irrespective of device used. Our research is the first to use large click stream data of failed attempts to use credit cards on mobile vs. non-mobile devices. Fraudsters and hackers use various techniques to gain access to consumer information by virtue of impersonation. Even the best security systems are able to prevent most, but not all, attacks (Bolton and Hand 2002) and organizations periodically evaluate the effectiveness of their systems against new forms of abuse and make improvements. We found a higher occurrence of failed attempted credit card payments on mobile vs. non-mobile devices. Consumers using mobile devices are more likely to abandon their cart if they experience a high number of failed attempts to use a credit card than when experiencing the same number of failed attempts to use a credit card on non-mobile devices. Our research seeks to inform both the academic and managerial understanding of this issue. Consequently, to reduce rates of cart abandonment, retailers should ensure they subscribe to apps that pass through rigorous approval processes and are trusted by the consumer. Retailer-provided information, in the form of product information, is used to decide which products to purchase. However, it appears that customer-provided information, in the form of consumer reviews, is used by consumers as a screening tool to decide which products may not perform as expected, and, thus, which products not to purchase. One of the limitations of the existing literature is the inability of the data to enable researchers to link cart abandonment to a user exploring online product reviews on mobile devices. Mallapragada et al. 2016, for example, identified this as a gap in the literature. Our unique data allowed us to use real click-stream data to link reviews read with cart abandonment. We did not find evidence that consumers that read positive product reviews using mobile devices were less likely to abandon their cart than those using non-mobile devices. Recent research from Northwestern University's Spiegel Research Center, (2017, pg., 10) noted that: "readers are skeptical of reviews that are too positive and, in many cases, a negative online review is seen as more credible. Additional research by social commerce specialist Revoo indicates that consumers spend four times as long on a site when they interact with negative reviews, with a 67% increase in conversion rate". The lesson from our findings was that even if positive reviews could be used to boost sales, having more reviews is not necessarily better after a certain point. As online commerce evolves, it is important for firms that adopt these media to understand the role of the factors influencing cart abandonment across device channels. The findings presented here could help these organizations develop effective strategies to reduce cart abandonment.

5.2.1 Limitations and future research

5.2.1.1 ML Prediction

A suitable future research avenue might be to develop an ensemble model which combines the power of SVM and DNN (Krauss et al., 2016). Research into DNN is undergoing a lot of transformation and other areas of exploration are emerging by looking at different variants, such as DeepAR, fuzzy methods or probabilistic methods.

Furthermore, political issues, such as sudden changes of leadership in these countries and economic turmoil, might suggest why shallow architectures were not able to perform as well as DNN. Since stock markets are rapidly changing in emerging markets, practitioners should always compare both shallow architecture and DNN before taking any decision. Further research could explore datasets from sub-markets within emerging markets using DNN with shallow architectures.

5.2.1.2 ML Causal Inference

While the present work built on past work and generated original findings, it does have limitations. One, while past work on e-commerce has shown the motivations for online shopping (Wolfinbarger & Gilly, 2003) and e-cart use (Close & Kukar-Kinney, 2010) using causal inference, behavioural data showed the actions and online consumers' actual behaviours, but did not include aspects such as attitude. Another important topic that is worth including in future work is the role of identity-signalling, as suggested by Grewal et al. (2019).

Further, while we used big data based on a multi-national sample spanning hundreds of brands and several different countries, the data did not include purchasing services online. The findings could be different if we were modelling the online path to purchase for services or experiential goods, such as sporting tickets or vacations. Future research can replicate this work in the context of services or experiential goods subject to data availability. Managerially and theoretically, it would be relevant to examine if any differences in findings would occur when the product being browsed for or purchased online is not tangible.

We focused on business to consumer e-commerce and m-commerce. There may be different outcomes if the project was based on business-to-business e-commerce or m-commerce. As a great portion of online sales are in fact by businesses and not at the consumer level, this presents another important opportunity for future research.

While the field nature of the dataset allowed us to track actual shopping and browsing behaviour, providing an advantage over previous research that relied on consumers' self-report survey data (for example, see Table 4.1), the nature of the data also presents some limitations. In particular, our data does not contain demographic information; however, we were able to control for this limitation by conducting an experiment to check if it had any effect on our results.

Future research might look at comparing the impacts of mobile vs. non-mobile on cart abandonment using sales outcomes. Another area to explore is the role of review valence of mobile vs. non-mobile on cart abandonment.

From the extant literature, the personalization of promotion offers have been shown to reduce cart abandonment. As competition in the retailing industry becomes intense, most retailers are massively discount their products online, often to the point of harming revenues. Future research work could allow retailers to understand how smart promotion strategies can be applied based on the profile of the consumer while remaining competitive. This may help reduce abandonment across device channels.

Chapter 6 Appendices

Appendix A:

Table 2.1: Summary of experimental settings

Hypotheses	Experimental Factor	Factor Levels
H1	Forecast horizon	Intraday *vs.* Daily
H2	Activation functions	TANH *vs.* RELU
H3	Market classification	Emerging *vs*. developed market

Table 2.2 Classification of the financial markets employed in this study as 'emerging' and 'developed'

Emerging market classification		Economy	Index
FTSE	MSCI		
		Spain	IBEX 35
		Switzerland	Swiss Market Index
		Finland	OMXH25
Y	Y	Turkey	ISE-100
		France	CAC 40
		Sweden	OMX ALL - SHARE Stockholm Index
		Norway	OSE All Share Index
Y	Y	Brazil	Brazilian Bovespa Futures
		Hong - Kong	Hang Seng index
Y	Y	Czech	Prague Stock Exchange Index
		US	Dow Jones Industrial Average
		Denmark	OMX Copenhagen Index
		Latvia	Riga Index
		Austria	ATX
		US	S&P 500
Y	Y	South Africa	FTSE/JSE Africa Top40
Y	Y	Hungary	BUX
		UK	FTSE 100
Y	Y	Thailand	Thai Stock exchange MAI securities index
		Germany	DAX
		Korea	KOSPI 200 Index
		Singapore	Straits times index
		Netherland	AEX
		Belgium	BEL20
		Japan	Nikkei 225
		Canada	SP TSX composite index
		US	NASDAQ-100
Y	Y	China	ShangHai SE composite index

Y	Y	Indonesia	Jakarta composite index
		Italy	FTSE MIB Index
		Portugal	PSI-20
		Estonia	OMX Tallinn index
Y	Y	Malaysia	FTSE Bursa Malaysia KLCI index
		Lithuania	OMX Vilnius index

Table 2.3a: The numbers of levels tested in different parameter settings for SVM

Parameters	Levels (polynomial)
Degree of kernel function (*d*)	1,2,…..5
Gamma in Kernel function (γ)	0,0.1,0.2,….,5.0
Regularization parameter (*c*)	1,10,100

Table 2.3b: The numbers of levels tested in different parameter settings for NN

Parameters	Levels(s)
Number of neurons (*n*)	10,20,……,100
Epochs (*ep*)	1,2,…..,10
Momentum constant (*mc*)	0.1,0.2,…..,0.9
Learning rate (*lr*)	0.1,0.2

Table 2.3c: The numbers of levels tested in different parameter settings for DNN

Parameters	Levels(s)
Number of neurons (*n*)	10,20,……,100
Epochs (*ep*)	1,2,…..,10
Momentum constant (*mc*)	0.1,0.2,…..,0.9
Learning rate (*lr*)	0.1,0.2
Hidden layers	2,3,……,10
Activation functions	RELU/TANH
MiniBatchSize	5,6,7 ……,15

Table 2.4: Tukey's Standardized Range (HSD) test: Comparing the Accuracy of Models using Daily Data

Alpha	0.05
Error of degree of freedom	132
Error mean square	0.005
Critical value of Studentized Range	3.679
Minimum Significant Difference	0.045

Tukey Grouping	Mean	Methods
A	0.578	SVM
B	0.522	NN
B	0.489	DNN(RELU)
C	0.452	DNN(TANH)

Table 2.5: Tukey's Standardized Range (HSD) test for Accuracy of Models using Hourly Data

Alpha	0.05
Error of degree of freedom	132
Error mean square	0.005
Critical value of Studentized Range	3.69
Minimum Significant Difference	0.045

Tukey Grouping	Mean	Methods
A	0.557	SVM
A	0.549	DNN(RELU)
A	0.538	DNN(TANH)
A	0.514	NN

Table 2.6: Tukey's Standardized Range (HSD) test for Accuracy using Minute Data

Alpha	0.05
Error of degree of freedom	132
Error mean square	0.008
Critical value of Studentized Range	3.68
Minimum Significant Difference	0.057

Tukey Grouping	Mean	Methods
A	0.645	DNN(RELU)
A	0.617	DNN(TANH)
B	0.558	SVM
B	0.547	NN

Table 2.7: Tukey's Standardized Range (HSD) test for Accuracy of Models using Tick Data

Alpha	0.05
Error of degree of freedom	132
Error mean square	0.009
Critical value of Studentized Range	3.68
Minimum Significant Difference	0.062

Tukey Grouping	Mean	Methods
A	0.607	DNN(RELU)
A	0.585	SVM

	A	0.582	DNN(TANH)
	A	0.573	NN

Table 2.8: Regression analysis of Predictive accuracy – DNN (RELU) vs SVM vs NN

Predictive accuracy	Estimated Coefficient	Std. Error	t value	Pr (>|t|)
(Intercept)	0.64559	0.01489	43.358	< 2e-16***
as.factor(NN)	-0.09824	0.02106	-4.665	9.67e-06***
as.factor(SVM)	-0.08765	0.02106	-4.162	6.73e-05***
Residual standard error	0.08682	df	99	
R-Squared	0.2094	Adjusted R - Squared	0.1935	
F- Statistic	13.11	(on 2 and 99 DF)		
p-value	8.87E-06			

Appendix B:

Table 3.1. Synthesis of Literature on Device Type in Mobile Commerce and Electronic Commerce

Source	Area	Settings	Scope	Consider types of websites visited	Time of the day of visit	Time spent on shopping session in seconds	Frequency of order made / session	Consider all three device type (PC/Smartphone/Tablet)	Findings
The current research	e-commerce browsing to sales conversions	Click stream	All devices	Y	Y	Y	Y	Y	Tablet leads to highest e-cart value and highest conversion rate overall, followed by PC, and last, smartphone. Visiting clearance page leads to highest conversion for tablet. Existing cart and reading product reviews lead to highest conversion for PC. Smartphone is used for browsing, but less for completing purchase.
de Haan et al. (2018)	Risk to product category	Click stream	All devices	N	N	N	N	Y	The increased penetration of mobile devices has a significant impact on customers' online shopping behavior, with customers frequently switching between mobile and fixed devices on the path to purchase.
Xu et al. (2017)	e-commerce sales	e-commerce	All devices	N	Y	N	N	Y	Users' adoption of tablets enhanced the overall growth of Alibaba's e-commerce market, with an annual increase of approximately US$923.5 million.
Li et al. (2017)	Mobile promotion	Field experiment	Mobile users	N	N	N	N	N	Purchase responses to promotions were higher and faster in sunny weather relative to cloudy weather, and were lower and slower in rainy weather.
Andrews et al. (2016)	Mobile ads	Survey	Mobile users	N	Y	N	N	N	On average, the purchase rates measured 2.1% with fewer than two people per square meter and increased to 4.3% with five people per square meter, after controlling for peak and off-peak times, weekdays and weekends, mobile use behaviours, and randomly sending mobile ads to users.
Luo et al. (2014)	Mobile promotion	Survey	Mobile users	N	N	N	N	N	Temporal targeting and geographical targeting individually increase sales purchases.

Table 3.2. Descriptive Statistics per Device (N = 958,859 Sessions)

Descriptive Statistics per Device (N = 958,859 Sessions)			
Variable	***PC***	***Tablet***	***Smartphone***
Number of sessions	212,431	161,923	584,505
Previous PC device session	170,500	28,000	100,201
Previous tablet device session	20,000	130,600	45,000
Previous smartphone device session	26,102	29,000	511,400
Pages per session	8.150	8.074	6.070
Conversion rate	7.96%	8.58%	5.50%
Conversion rate with previous session on PC device	6%	1%	3%
Conversion rate with previous session on tablet device	5%	2%	3%
Conversion rate with previous session on smartphone device	1%	1%	8%
Conversion rate > 1 per session	1,145 (1.82%)	856 (1.36%)	1,749 (2.78%)

Table 3.3. Summary Statistics of Variables Across Devices

Variables	Description	PC				Smartphone				Tablet			
		Mean	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**	**Mean**	**SD**	**Min**	**Max**
Order completed	Frequency of orders completed in session	0.08	0.3	1	3	0.06	0.25	1	3	0.09	0.32	1	4
e-cart value	Total value of items place in the e-cart	29.84	208.35	2	1993 2	24.24	122.89	1	6907	33	156.85	2	9900
Existing cart	If a customer has a product in their cart on the first page of the session or not	0.09	0.29	0	1	0.15	0.36	0	1	0.15	0.35	0	1
Pages seen	# of pages viewed	7.39	16.55	1	2623	6.1	10.27	0	819	8.03	11.92	1	919
Product views	# of products viewed	1.86	28.69	1	1301 8	0.07	0.83	1	85	2.29	5.17	1	137
Visit before work	Visit count between (Hour >=6 and Hour<=8)	0.12	0.33	0	1	0.14	0.34	0	1	0.12	0.33	0	1
Visit during lunch	Visit count between (Hour >=12 and Hour<=13)	0.14	0.35	0	1	0.1	0.3	0	1	0.09	0.29	0	1
Visit after work	Visit count between (Hour>=17 and Hour<=18)	0.11	0.31	0	1	0.13	0.33	0	1	0.15	0.35	0	1
Visit during evening	Visit count between (Hour>=19 and Hour<=23)	0.4	0.8	0	1	0.56	0.9	0	1	0.6	0.92	0	1
Seconds spent (in '000)	Total number of seconds the entire session was active	286	629	1	3973 1	244	490	1	2173 5	436	761	1	18368
Read review	Number of product reviews read	0.08	0.53	0	34	0	0.06	0	12	0.12	0.64	0	31
Clearance	Flag if customer has visited clearance section of site	0.15	0.35	0	1	0.14	0.34	0	1	0.19	0.39	0	1

Table 3.4. Correlation Matrix

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) Order completed	1											
(2) e-cart value	0.121***	1										
(3) Existing cart	0.104***	0.249***	1									
(4) Pages seen	0.422***	0.167***	0.089***	1								
(5) Product views	0.051***	0.029***	0.006***	0.279***	1							
(6) Visit before work	0.005***	0.001	0.012***	-0.0129***	-0.005***	1						
(7) Visit during lunch	0.002	-0.000	-0.007***	-0.003**	0.001	-0.134***	1					
(8) Visit after work	0.002*	0.002*	-0.002	0.004***	0.002*	-0.146***	-0.131***	1				
(9) Visit during evening	-0.0170***	-0.002*	0.001	0.005***	0.002*	-0.232***	-0.208***	-0.226***	1			
(10) Seconds spent	0.380***	0.182***	0.097***	0.713***	0.180***	-0.017***	-0.016***	0.011***	0.025***	1		
(11) Read review	0.098***	0.043***	0.021***	0.209***	0.088***	-0.005***	0.006***	0.004***	-0.002	0.221***	1	
(12) Clearance	0.086***	0.027***	0.083***	0.134***	-0.000	0.013***	-0.011***	-0.015***	-0.009***	0.252***	0.009***	1

*p<0.05, **p<0.01, ***p<0

Table 3.5. Propensity Score Matching Results

Determinants of Device Types Propensity	Exogenous Variable	Coefficients
	Constant	-4.982***
	e-cart value	0.999***
	Existing cart	1.722***
	Pages seen	0.993***
	Product views	0.904***
	Visit before work	1.394***
	Visit during lunch	0.875***
	Visit after work	1.289***
	Visit during evening	1.227***
	Seconds spent	0.000001***
	Read review	0.703***
	Clearance	0.967***

Mean Before Matching			Exogenous variable: Customer Behaviour Before Treatment	Means After Matching			
Control Group (n=362,134)	Treatment Group(n = 150,231)	Mean comparison ρ -value		Control Group (n= 150,228)	Treatment Group(n =150,228)	Mean comparison ρ-value	Percentage reduction in bias (PRB)*
23.53	18.45	0.00	e-cart value	21.22	21.45	0.93	99
0.12	0.11	0.00	Existing cart	0.1	0.12	0.69	83
5.6	4.8	0.00	Pages seen	4.4	4.5	0.88	98
0.7	0.6	0.00	Product views	0.6	0.63	0.81	95
0.13	0.12	0.00	Visit before work	0.13	0.14	0.64	93
0.1	0.2	0.00	Visit during lunch	0.25	0.27	0.55	93
0.12	0.09	0.00	Visit after work	0.11	0.13	0.45	85
0.54	0.61	0.00	Visit during evening	0.6	0.69	0.78	87
28,703	22,796	0.00	Seconds spent	24,498	24,721	0.56	99
0.03	0.01	0.00	Read review	0.12	0.14	0.68	86
0.15	0.12	0.00	Clearance	0.14	0.16	0.74	88

*p<0.05, **p<0.01, ***p<0.001; *In line with Garnefield et al. (2013), we calculated the PRB using a formula from Rosenbaum and Rubin (1983).

Table 3.6. Summary of Fitted Count Regression Model

Note: Device types smartphone and tablet are compared to PC which serves as the baseline.

Variables	Poisson	Poisson adjusted	Quasi	NB	ZINB	MCMC Parameter estimate
(Intercept)	-5.135*** (0.010)	-5.135*** (0.170)	-3.638*** (0.012)	**-5.135*** (0.171)**	- 3.356*** (0.009)	-4.231 (0.008)
as.factor(smartphone)	-0.002*** (0.010)	-0.002*** 0.20754	-0.002*** 0.087	**-0.427*** 0.013**		-0.304 (0.014)
as.factor(tablet)	0.224*** (0.014)	0.224*** (0.286)	0.224*** (0.117)	**0.203*** (0.017)**		0.891 (0.005)
Seconds spent	0.000***	0.000***	0.000***	**0.000*** **		0.000
Read reviews	0.046*** (0.002)	0.046*** (0.036)	0.046*** (0.037)	**0.001*** (0.008)**	0.186*** (0.007)	0.086 (0.019)
e-cart value	0.000*** (0.000)	0.000*** (0.000)	0.000*** (0.000)	**-0.000*** (0.000)**	0.001*** (0.000)	-0.000 (-0.000)
Existing cart	0.959*** (0.016)	0.959*** (0.288)	0.959*** (0.252)	**0.475*** (0.022)**	- 0.159*** (0.013)	0.825 (-0.001)
Product views	0.000*** (0.000)	0.000*** (0.000)	0.0002*** (0.000)	**-0.065*** (0.000)**		-0.032 (-0.005)
Visit before work	1.755*** (0.01)	1.755*** (0.78)	1.755*** (0.165)	**0.073*** (0.014)**	-0.02 (0.015)	0.054 (0.002)
Visit during lunch	1.896*** (0.012)	1.896** (0.859)	1.896*** (0.189)	**0.043*** (0.015)**	0.017 (0.016)	0.091 (0.004)
Visit after work	2.985*** (0.003)	2.985*** (1.05)	2.985*** (0.044)	**0.009 (0.015)**	0.016 (0.015)	0.317 (0.008)
Visit during evening	0.844*** (0.004)	0.844*** (0.412)	0.844*** (0.071)	**-0.082*** (0.006)**	- 0.068*** (0.006)	-0.371 (0.019)
Pages seen	0.006*** (0.000)	0.006*** (0.000)	0.006*** (0.000)	**0.069*** (0.000)**		0.015 (0.007)
Clearance	-0.664*** (0.014)	0.664*** (0.599)	0.664*** (0.215)	**0.265*** (0.021)**	0.363*** (0.011)	0.276 (0.022)
Clearance X device type(smartphone)	-0.064*** (0.022)	-0.121*** (0.592)	-0.032*** (0.259)	**-0.028*** (0.024)**		-0.182 (0.155)
Clearance X device type(tablet)	-0.118*** (0.025)	-0.181*** (0.592)	-0.181*** (0.259)	**-0.423*** (0.012)**		-0.142 (0.115)
Existing cart X device type(smartphone)	-0.142*** (0.22)	-0.144*** (0.032)	-0.138*** (0.002)	**-0.031*** (0.012)**		-0.022
Existing cart X device type(tablet)	-0.383*** (0.018)	-0.383*** (0.318)	-0.383*** (0.285)	**-0.128*** (0.025)**		-0.546 (0.222)
Read Reviews X device type(smartphone)	0.019*** (0.004)	0.019*** (0.032)	0.019*** (0.063)	**0.109*** (0.011)**		0.022(0.011)
Read Reviews X device type(tablet)	-0.012*** (0.001)	0.011*** (0.001)	0.018*** (0.011)	**0.022*** (0.011)**		0.123(0.011)
Screen size '360x640'	-0.071***	-0.071***	-0.071***	**-0.228*** **		-0.008 (0.002)

	(0.014)	(0.014)	(0.014)	**(0.017)**		
Screen size '768x1024'	0.619***	0.619***	0.619***	**2.436*****		0.514
	(0.167)	(0.167)	(0.167)	**(0.209)**		(0.127)
Screen size '320x568'	-0.021	-0.021	-0.021	**0.008**		-0.011
	(0.017)	(0.017)	(0.017)	**(0.019)**		(0.002)
Screen size '1366x768'	0.689***	0.689***	0.689***	**2.128**		0.429
	(0.161)	(0.161)	(0.161)	**(0.204)*****		(0.132)
Screen size '414x736'	-0.083***	-0.083***	-0.083***	**-0.070*****		-0.073
	(0.019)	(0.019)	(0.019)	**(0.022)**		(0.004)
Africa	-1.215**	-1.215**	-1.215**	**-1.648*****		-1.442
	(0.372)	(0.372)	(0.372)	**(0.432)**		(0.241)
North America	-0.945**	-0.945**	-0.945**	**-0.852***		-0.821
	(0.359)	(0.359)	(0.359)	**(0.412)**		(0.421)
South America	-0.802	-0.802	-0.802	**-0.899**		-0.437
	(0.558)	(0.558)	(0.558)	**(0.163)**		(0.121)
Europe	0.319	0.319	0.319	**0.245**		0.562
	(0.333)	(0.333)	(0.333)	**(0.386)**		(0.212)
Australia	-1.513***	-1.513***	-1.513***	**-1.256****		-2,173
	(0.427)	(0.427)	(0.427)	**(0.477)**		(0.821)
Session dummies	Included	Included	Included	**Included**	Included	Included
Time dummies	Included	Included	Included	**Included**	Included	Included
(Intercept)					-5.272***	
					(0.079)	
Read review					-0.240***	
					(0.041)	
Cart quantity					0.174***	
					(0.004)	
e-cart value					0.002***	
					(0.000)	
Existing cart					2.502***	
					(0.063)	
Visit before work					-0.041	
					(0.069)	
Visit during lunch					0.051	
					(0.079)	
Visit after work					-0.092	
					(0.074)	
Visit during evening					-0.103***	
					(0.029)	
Clearance					-0.087	
					(0.053)	
Observations	150,228		150,228	**150,228**	150,228	
Number of parameters	27	27	27	**27**	20	
Log Likelihood	-187339			**-337,100**	-219809	
AIC	374745			**337168**	439661	
BIC	375128			**415886**	439909	
$\sum \hat{f}_i(0)$	77062			**77886**	94405	

Significant levels for variables * p<0.05, ** p<0.01, *** p<0.001 Negative binomial (NB) is in bold as the preferred model.

Table 3.7. Model Comparison Using Vuong Non-Nested Tests

Model comparison	Vuong Test Statistic	p	Preferable Model
Poisson vs. NB	-47.91	< 2.22e-16	NB
NB vs. ZINB	39.52	< 2.22e-16	NB
Poisson vs. ZINB	-36.48	< 2.22e-16	ZINB

Appendix C

Table 4.1: How the Current Study Extends the Shopping Cart Use and Abandonment Literature

Article	Focus	Data	Method	Includes Online Cart Use	Includes Mobile Devices	Includes Field Data or Behavioral DV	Includes Non-Mobile Devices
The current research	**Differences in Online Shopping Cart Abandonment across Model and Non-Mobile devices**	**165,613 unique shoppers click stream data at a large online retailer**	**Field study**	✓	✓	✓	✓
Huang, Korfiatis and Chang (2018)	Conflicts, ambivalence and hesitation in mobile cart abandonment	232 mobile shoppers from Taiwan & 226 of US consumer	Surveys		✓		
Albrecht, Hattula and Lehmann (2017)	Shopping stress and task vs. recreation orientation on purchase abandonment	Recruited 883 participants using online panel	Survey, Experiment			✓	
Close, Kukar-Kinney and Benusa (2012)	Online shopping cart abandonment	289 online shoppers from 44 States in the US	Conceptual	✓			
Xu and Huang (2015)	Determinants of cart abandonment in China	Online survey of 210 people in China	Survey				

Kukar-Kinney and Close (2010)	Determinants of cart abandonment	289 online shoppers from 44 States in the US	Surveys					
Rajamma and Paswan (2009)	Perceived waiting time, risk and transaction inconvenience in cart abandonment	720 survey data from business undergraduate students	Survey					
Oliver and Shor (2003)	Role of promotion codes in digital redemption of promotion codes or abandonment	206 respondents using online surveys	Survey					

Table 4.2 Variable Names, Descriptions and Descriptive Statistics of Cart abandonment by Consumers

Variables	Description	Mean		SD		Min		Max	
		Non-mobile	Mobile	Non-mobile	Mobile	Non-mobile	Mobile	Non-mobile	Mobile
Cart abandonment	Shopping cart abandonment: 1 -Abandoned cart without purchase, 0 - Completed purchase	0.52	0.938	0.259	0.24	0	0	1	1
Visit during evening	1 if visit between 1800 to 2300hr	0.197	0.452	0.398	0.322	0	0	1	1
Read positive review	Read positive product review or not	0.078	0.027	0.526	0.307	0	0	1	1
Card failure	Card failure / decline during payment	1.061	3.091	0.059	0.054	0	0	5	6
Basket value (£)	Cart value	33.204	26.079	216.108	131.314	0	0	19932	9900

Table 4.3: Variable Correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) cart abandonment	1						
(2) device type	0.029***	1					
(3) read review	-0.098***	-0.067***	1				
(4) visit during evening	0.0144***	0.092***	-0.002***	1			
(5) card failure	-0.005***	-0.001***	-0.007***	-0.001***	1		
(6) basket value	-0.127***	-0.019***	0.043***	0.002***	-0.001***	1	
(7) basket quantity	-0.320***	-0.013***	0.082***	-0.011***	0.002***	0.417***	1

Significant levels for variables * p<0.05, ** p<0.01, *** p<0.001

Table 4.4 Model Comparison

Model	Main effect of theoretical variable	Log-Likelihood	Akaike Information Criterion	Bayesian Information Criterion
M1	No	-531,300	1,062,632	1,062,818
M2 (RBP)	**Yes**	**-469,576**	**939,225**	**939,643**

Notes: The model with the best fit is indicated in boldface.

Table 4.5 Results: Predictors of Cart abandonment & Device Channels

Variable category	Variables	Probability of Cart Abandon	MCMC Parameter Estimate
	Intercept	-1.357***(0.480)	8.762(0.842)
	Mobile	0.414(0.150)	0.141(0.266)
Browsing behaviour across time	visit during evening	0.065***(0.214)	0.299(0.206)
	visit during evening x device type	-0.011***(0.061)	-0.121(0.226)
Browsing pattern	read review	-0.022(0.041)	-0.101(0.096)
	read review x device type	0.066(0.033)	0.090(0.153)
Card failure	card failure attempt 1	5.621(3.244)	-6.562(3.062)
	card failure attempt 2	-5.596(2.090)	7.188(2.053)
	card failure attempt 3	0.051(1.099)	-
	card failure attempt 4	-5.101(3.090)	5.716(1.361)
	card failure attempt 1 x device type	-1.115(4.351)	-
	card failure attempt 2 x device type	-1.858***(3.734)	-

	card failure attempt 3 x device type	-2.328***(3.734)	-
	card failure attempt 4 x device type	-2.778***(3.734)	-
Basket value	Log(basket value)	0.119***(0.027)	0.021(0.010)
	Log(basket value x device type)	0.056(0.037)	0.044(0.021)
Control variables	Log (active seconds)	-0.261***(0.038)	-0.124(0.021)
	product page	0.007***(0.001)	0.002(0.001)
	existing cart	0.339***(0.068)	0.226(0.021)
	cart use	-2.251***(0.313)	-3.221(0.313)
	session2	-0.104**(0.035)	-0.214(0.021)
	session3	-0.149**(0.048)	-0.126(0.016)
	session4	-0.096(0.066)	-0.013(0.044)
	session5	-0.164*(0.0771)	-0.129(0.011)
	session6	-0.194*(0.093)	-0.342(0.001)
	session7	-0.351***(0.096)	-0.271(0.021)
	session8	-0.271*(0.132)	-0.321(0.110)
	session9	-0.036(0.148)	-0.022(0.108)
	session10	-0.037(0.158)	-0.017(0.112)

Variable category	Variables	Probability of Mobile vs Non-mobile	MCMC Parameter Estimate
	Intercept	-1.359***(0.009)	'-1.687(0.576)
Browsing behaviour across time	visit during evening	0.401***(0.020)	0.429(0.091)
Browsing pattern	read review	-0.065***(0.016)	-0.246(0.077)
Card failure	card failure attempt 1	-0.433(0.589)	-0.233(0.519)
	card failure attempt 2	0.100(0.556)	0.129(0.551)
	card failure attempt 3	-6.034(3.781)	-
	card failure attempt 4	-6.130(5.319)	-5.230(3.319)
Basket value	Log(basket value)	-0.014(0.018)	-
Control variables	Log (active seconds)	0.098***(0.009)	-
	product page	0.007***(001)	-
	existing cart	0.124***(0.032)	0.112(0.032)
	cart use	-0.179***(0.046)	-0.179(0.046)
	session2	0.084***(0.033)	0.072(0.010)
	session3	0.168***(0.041)	0.140(0.012)
	session4	0.233***(0.050)	0.218(0.011)
	session5	0.263***(0.062)	0.153(0.071)
	session6	0.305***(0.075)	0.207(0.021)
	session7	0.311***(0.085)	0.218(0.012)
	session8	0.491***(0.104)	0.462(0.121)
	session9	0.412***(0.116)	0.332(0.111)
	session10	0.362**(0.126)	0.521(0.162)

*p < .10 **p < .05 ***p < .01. Standard errors are in parenthesis

Table 4.6 Results: Variable Sensitivity

Variable category	Variables	Probability of Cart Abandon	MCMC Parameter Estimate
	Intercept	8.584***(0.614)	8.476(0.877)
	Mobile device	0.423(1.150)	0.336(0.567)
Browsing behaviour across time	visit during evening	0.063***(0.214)	0.264(0.200)
	visit during evening x device type	-0.009***(0.061)	-0.183(0.178)
Browsing pattern	read review	-0.023(0.041)	-0.063(0.107)
	read review x device type	0.066(0.033)	0.115(0.136)
Card failure	card failure attempt	-0.096(0.244)	-1.633(1.079)
	card failure attempt x device type	-6.284***(0.175)	-4.828(0.352)
Basket value	Log(basket value)	0.119***(0.027)	0.211(0.011)
	Log(basket value x device type)	0.056(0.037)	0.125(0.017)
Control variables	Log (Active seconds)	-0.595***(0.013)	-0.282(0.041)
	product page	0.007***(001)	0.022(001)
	existing cart	0.339***(0.067)	0.311(0.133)
	cart use	-2.252***(0.312)	-3.2511(0.112)
	session2	-0.104**(0.035)	-0.111(0.225)
	session3	-0.149**(0.048)	-0.412(0.019)
	session4	-0.096(0.066)	-0.012(0.031)
	session5	-0.163*(0.077)	-0.191(0.069)
	session6	-0.194*(0.093)	-0.214(0.021)
	session7	-0.351***(0.096)	-0.321(0.710)
	session8	-0.259*(0.132)	-0.165(0.187)
	session9	-0.035(0.147)	-0.031(0.123)
	session10	-0.036(0.157)	-0.002(0.276)

Variable category	Variables	Probability of Mobile vs Non-mobile	MCMC Parameter Estimate
	Intercept	-1.359***(0.119)	-1.671(0.564)
Browsing behaviour across time	visit during evening	0.401***(0.020)	0.437(0.095)
Browsing pattern	read review	-0.023(0.041)	-0.259(0.072)
Card failure	card failure attempt	-0.261(0.180)	-0.225(0.319)
Basket value	Log(basket value)	-0.097***(0.023)	-0.021(0.023)
Control variables	Log (active seconds)	-0.595***(0.013)	-0.211(0.018)
	product page	0.007***(0.001)	0.071(0.002)
	existing cart	0.123***(0.032)	0.233(0.012)
	cart use	-0.179***(0.046)	-0.139(0.041)
	session2	0.084***(0.033)	0.072(0.010)
	session3	0.168***(0.041)	0.140(0.012)
	session4	0.233***(0.050)	0.218(0.011)

	session5	0.263***(0.062)	0.153(0.071)
	session6	0.305***(0.075)	0.207(0.021)
	session7	0.311***(0.085)	0.218(0.012)
	session8	0.491***(0.104)	0.462(0.121)
	session9	0.412***(0.116)	0.332(0.111)
	session10	0.362***(0.126)	0.521(0.162)

*p < .10 **p < .05 ***p < .01. Standard errors are in parenthesis

Table 4.7 Results: Controlling for Demographic Variable

Variable category	Variables	Alternative estimates
	Intercept	-5.656***(0.543)
	Mobile	0.685(1.051)
Browsing behaviour across time	visit during evening	-0.097(0.167)
	visit during evening x device type	-0.030(0.084)
Browsing pattern	read review	-0.004(0.042)
	read review x device type	0.009(0.055)
Card failure	card failure attempt 1	6.356(3.06)
	card failure attempt 2	-6.367(2.835
	card failure attempt 3	-
	card failure attempt 4	-4.503(0.038)
	card failure attempt 1 x device type	-
	card failure attempt 2 x device type	-
	card failure attempt 3 x device type	-
	card failure attempt 4 x device type	-
Basket value	Log(basket value)	0.159***(0.048)
	Log(basket value x device type)	0.031(0.055)
Control variables	Log (Active seconds)	-0.316***(0.086)
	product page	0.013***(0.002)
	existing cart	0.301***(0.149)
	cart use	-1.968*(0.532)
Variable category	Variables	
	Intercept	-4.042***(0.392)
Browsing behaviour across time	visit during evening	0.372***(0.042)
Browsing pattern	read review	-0.029(0.030)
Card failure	card failure attempt 1	5.861(6.329)
	card failure attempt 2	6.165(4.471)
	card failure attempt 3	-
	card failure attempt 4	-6.023(6.653)

Basket value	Log(basket value)	-0.009(0.026)
Control variables	Log (Active seconds)	0.304***(0.029)
	product page	-0.008***(0.002)
	existing cart	0.172**(0.056)
	cart use	-0.171(0.089)

*p < .10 **p < .05 ***p < .01. Standard errors are in parenthesis

Figure 2.1: Boxplot, comparing the forecasting accuracy of all four model specifications (SVM, NN, DNN (RELU), DNN (TANH)), across all four time horizons (daily, hourly, minute, tick)

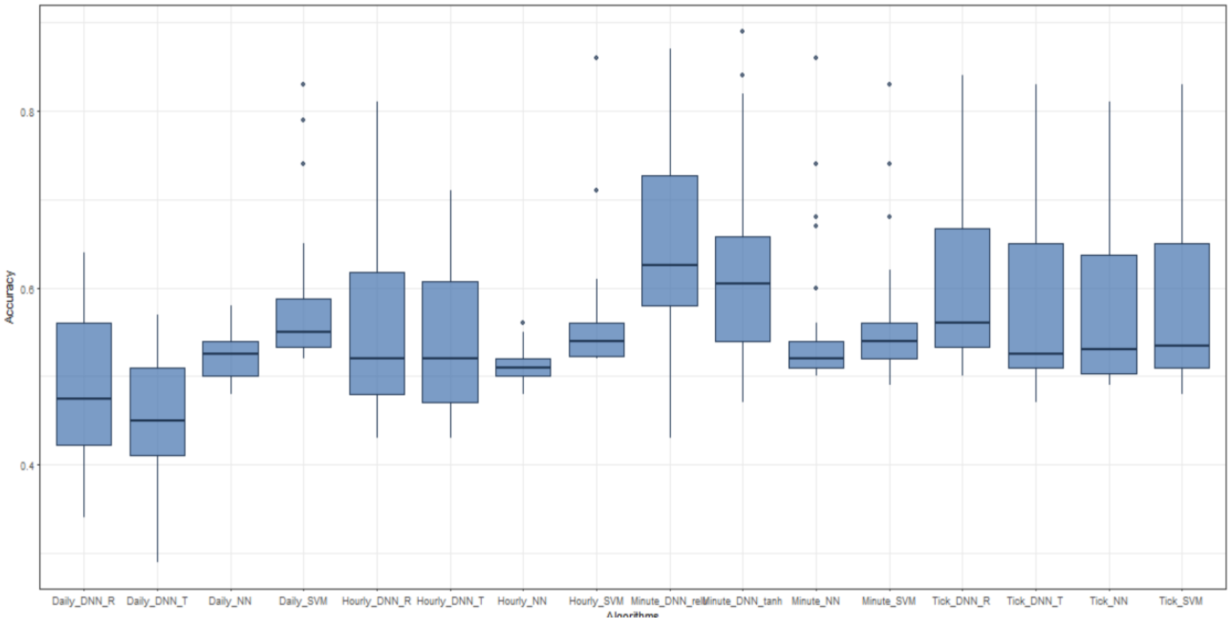


Figure 2.2: Box plot of prediction accuracy for all four model specifications (SVM, NN, DNN (RELU), DNN (TANH)), using daily data

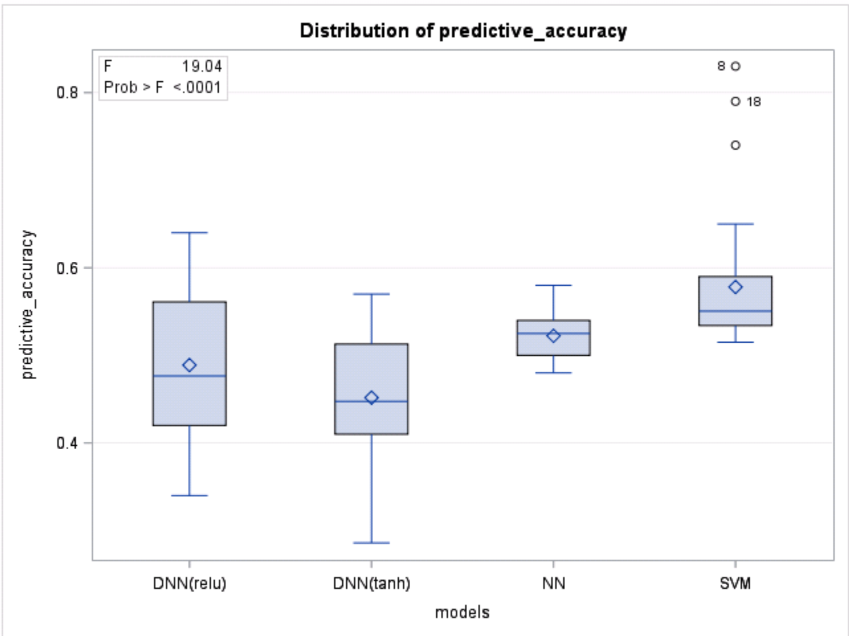


Figure 2.3: Box plot of accuracy for each all four model specifications (SVM, NN, DNN (RELU), DNN (TANH)), using hourly data

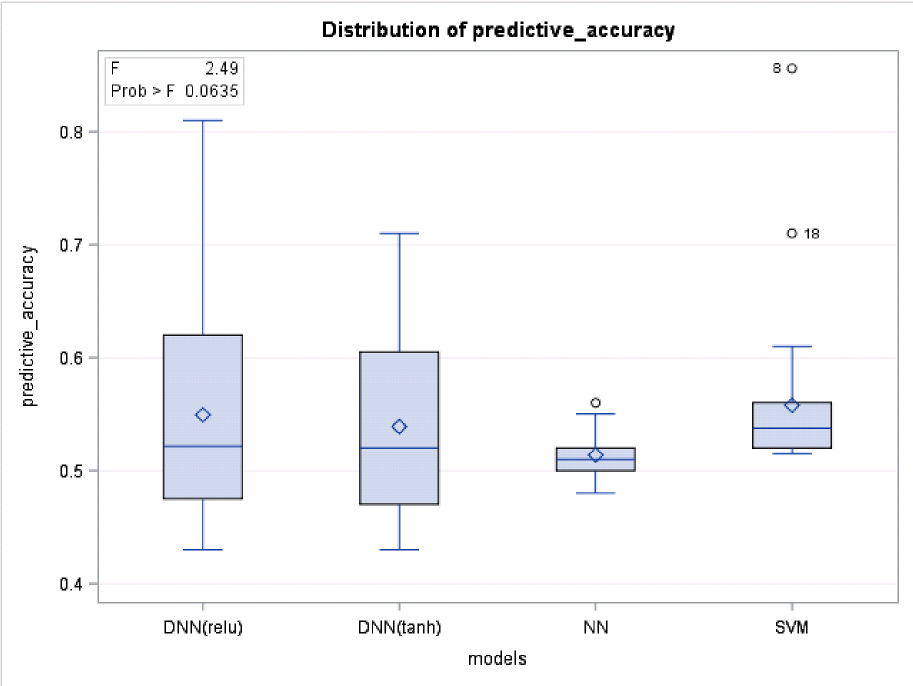


Figure 2.4: Box plot of accuracy for each of the four model specifications (SVM, NN, DNN (RELU), DNN (TANH)), using minute data

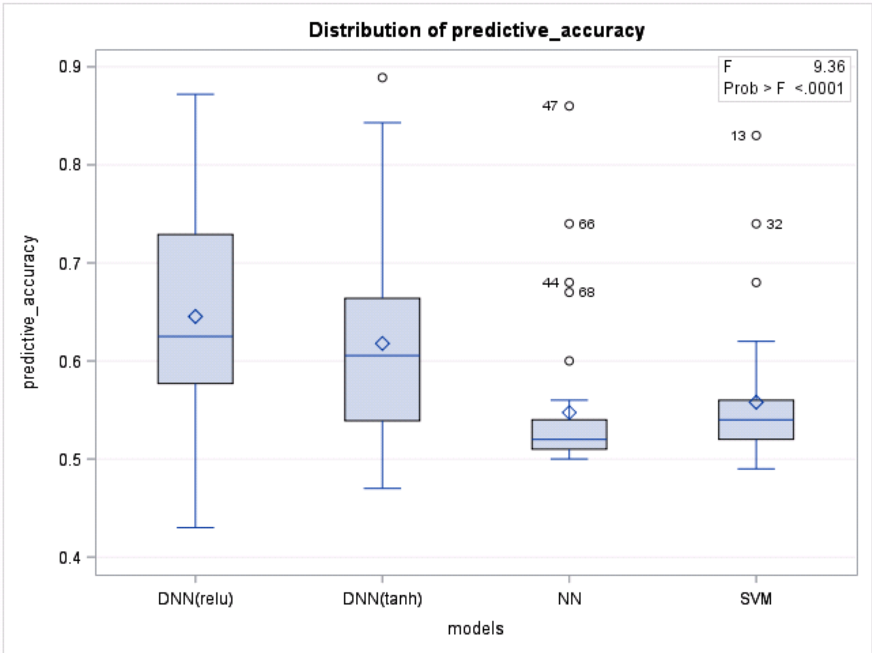


Figure 2.5: Box plot of accuracy all four model specifications (SVM, NN, DNN (RELU), DNN (TANH)), using tick data

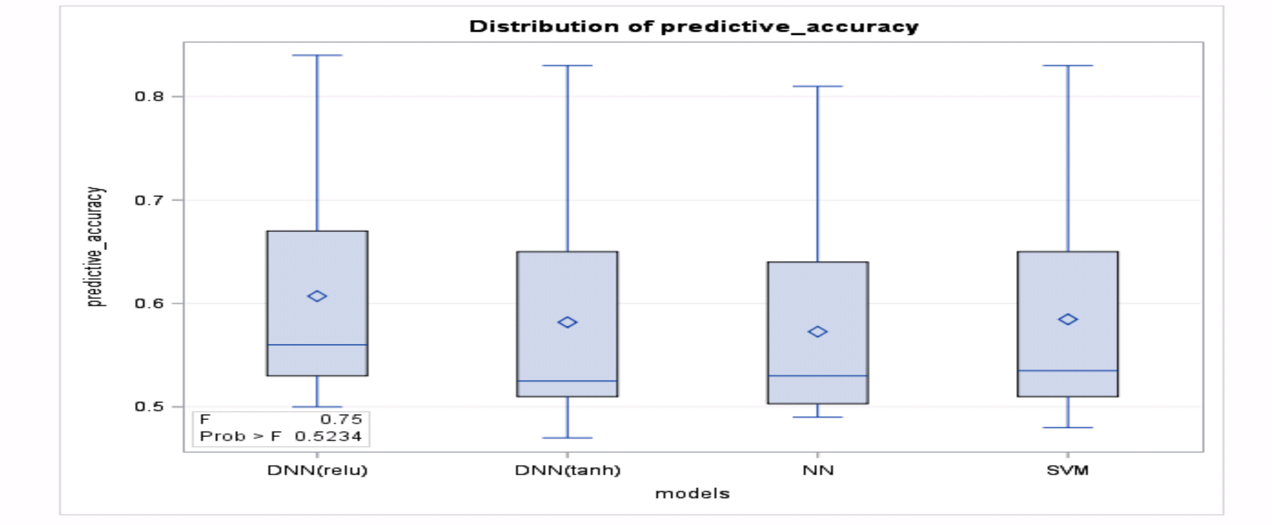


Figure 2.6: Box plot of accuracy for DNN using RELU/TANH activation functions across all four time horizons (daily, hourly, minute, tick).

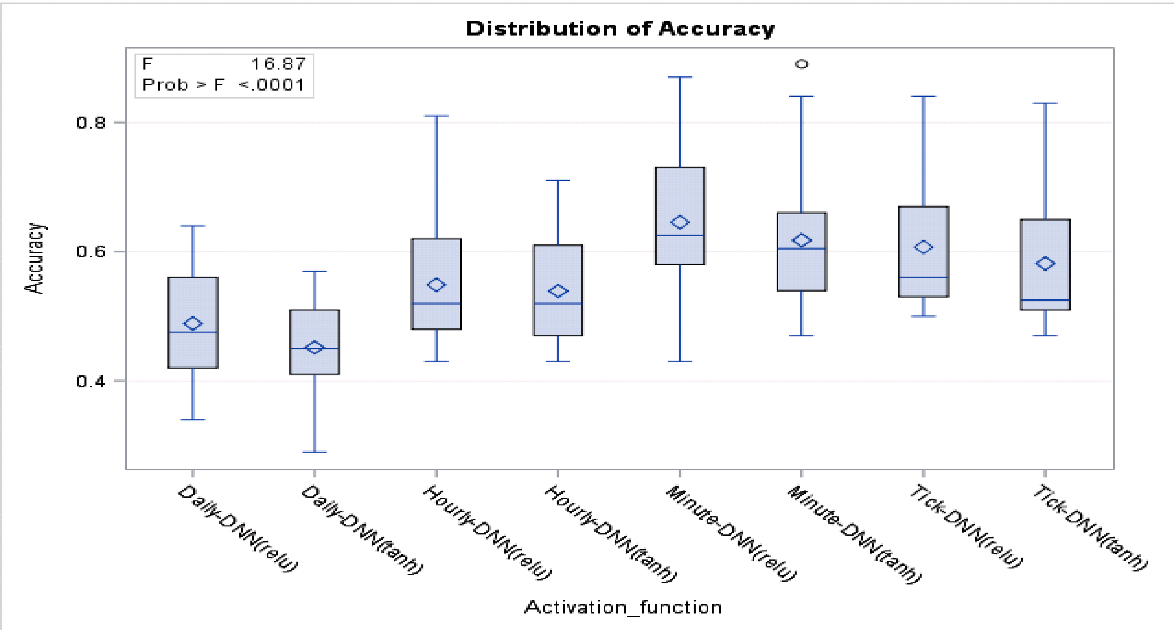


Figure 2.7: Predictive accuracy of emerging markets VS developed markets using DNN

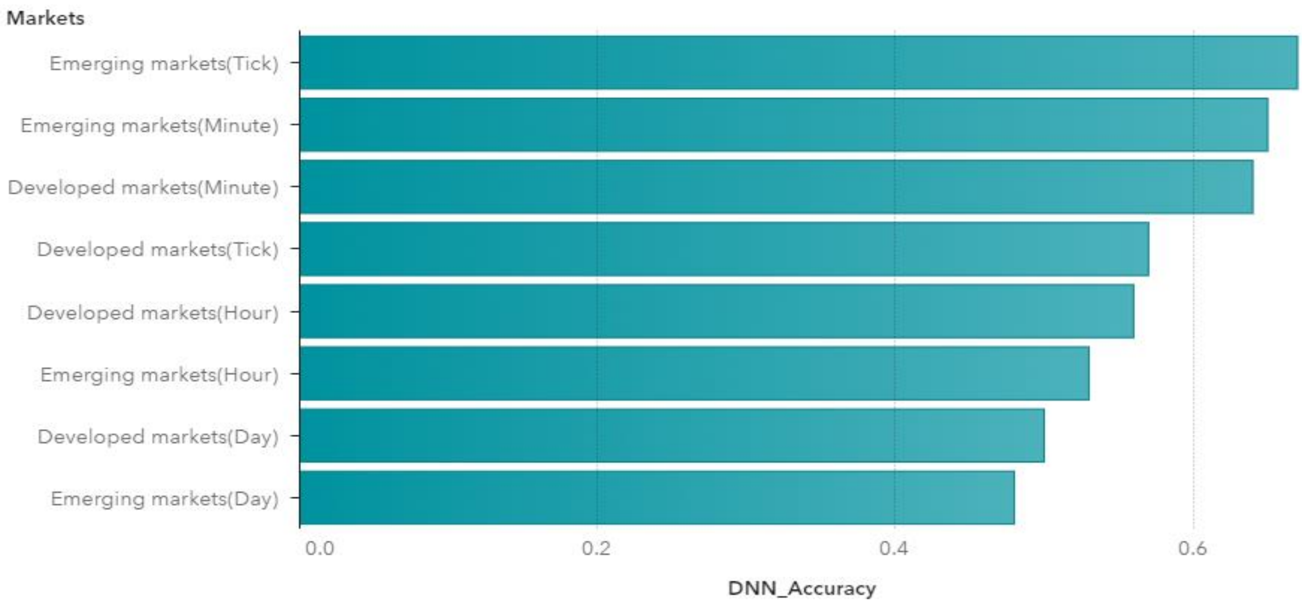
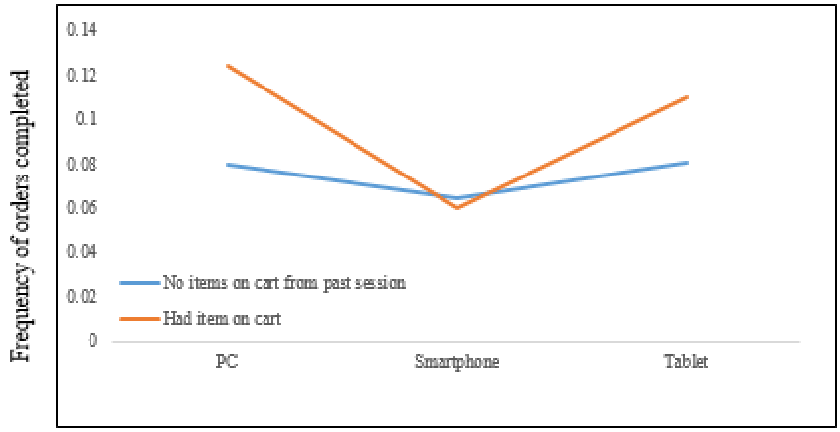
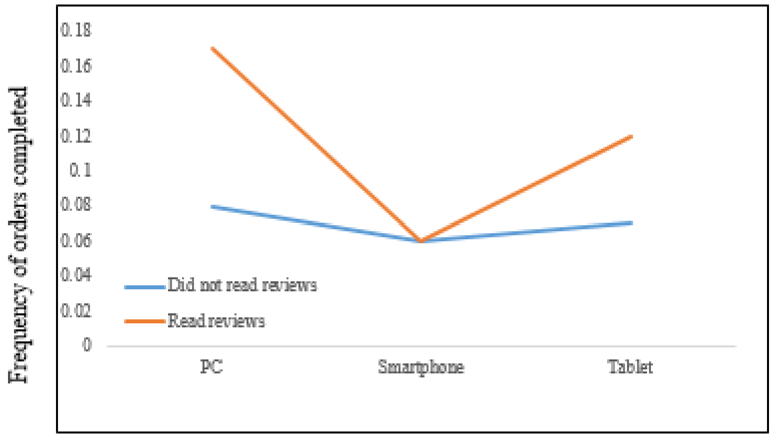


Figure 3.1. Interaction Plots

A. Interaction Between Existing Cart and Device Types



B. Interaction Between Read Reviews and Device Types



C. Interaction Between Clearance and Device Types

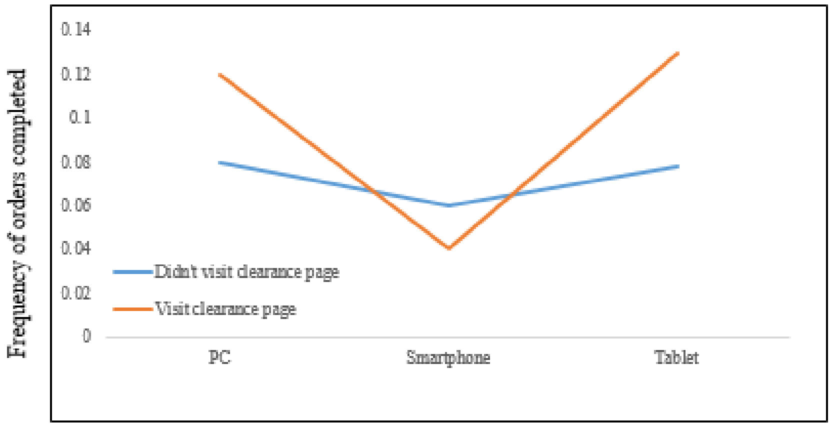
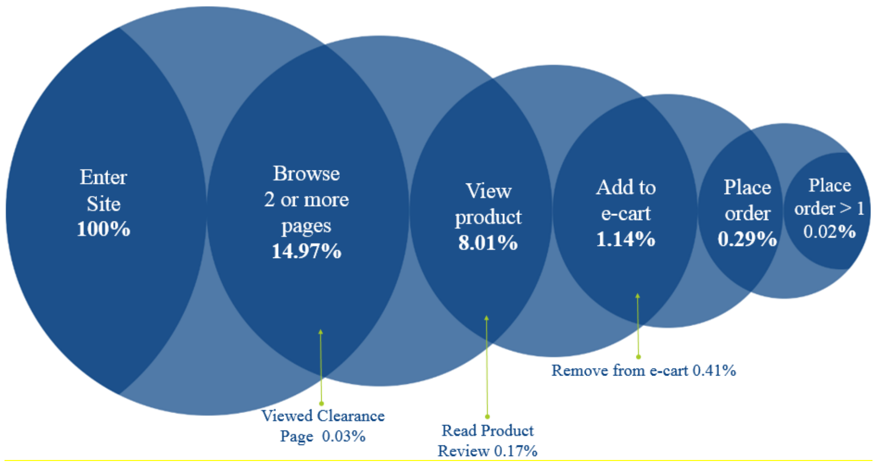
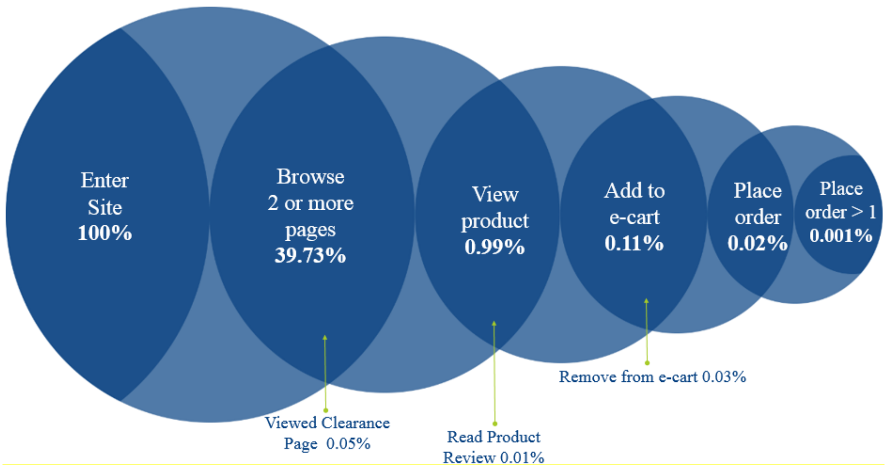


Figure 3.2. The Online Path to Purchase Across Device Types

A. The Online Path to Purchase from PCs

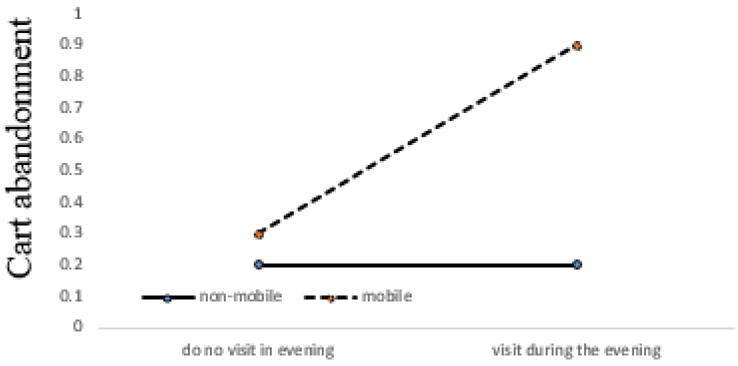


B. The Online Path to Purchase from Smartphones

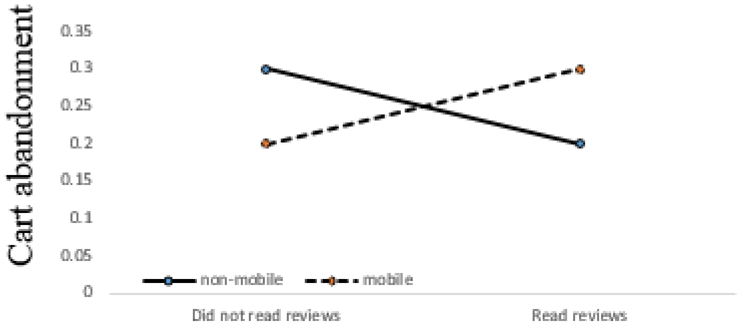


C. The Online Path to Purchase from Tablets

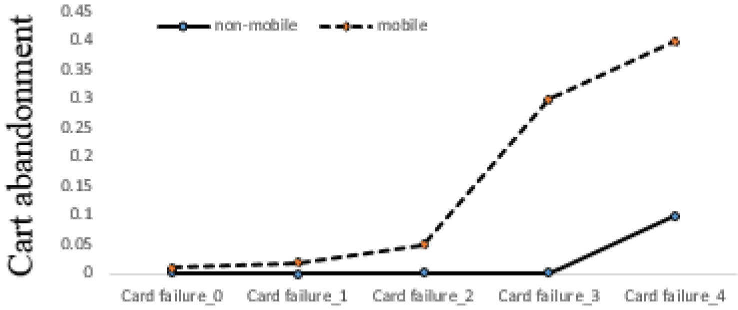
Figure 4.1: Interaction Plots
A. Interaction Between Evening Visit and Device Types



B. Interaction Between Read Positive Reviews and Device Types



C. Interaction Between Card Failure and Device Types



Chapter 7 References

Abdel-Hamid, O., Mohamed, A-r.,Jiang, H., & Penn, G. (2012). Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*. 4277-4280.

Ackley, D.H., Hinton, G.E., & Sejnowski,T.J (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, 9 (1): 147-169.

Aiken, Leona S. and Stephen G. West (1991). Multiple Regression Testing and Interpreting Interactions. Newbury Park, CA: Sage Publications.

Albrecht, Carmen-Maria, Stefan Hattula, and Donald R. Lehmann (2017), "The relationship between consumer shopping stress and purchase abandonment in task-oriented and recreation-oriented consumers," *Journal of the Academy of Marketing Science*, 45(5), 720–740.

Alexandre, P., Sabino, & S.,Albuquerque, P. (2015). Portfolio Selection with Support Vector Machines in Low Economic: Perspectives in Emerging Markets. *Economic computation and economic cybernetics studies and research / Academy of Economic Studies.* 49(4):261.

Allard, Thomas, Lea Dunn, and Katherine White (2020), "Negative Reviews, Positive Impact: Consumer Empathetic Responding to Unfair Word of Mouth," *Journal of Marketing*, 84 (4), 86-108.

Ananthakrishnan, Uttara, Beibei Li and Michael D. Smith (2020), "A Tangled Web: Should Online Review Portals Display Fraudulent Reviews?" *Information Systems Research*.

Andrews, Michelle, Xueming Luo, Zheng Fang and Anindya Ghose (2016), "Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness," *Marketing Science*, 35 (2), 218-233.

Angelidis, T., & Degiannakis, S. (2008). Volatility forecasting: Intra-day versus inter-day models. *Journal of International Financial Markets, Institutions and Money*. 18: 449-465. Doi:10.1016/j.intfin.2007.07.001.

Angrist, Joshua, and Joern-Steffen Pischke (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton, NJ: Princeton University Press.

Ansari, Asim, Carl F. Mela and Scott A. Neslin (2008), "Customer Channel Migration," *Journal of Marketing Research*, 45 (2), 60–76.

Archak, Nikolay, Anindya Ghose, Panagiotis G Ipeirotis. (2011), "Deriving the pricing power of product features by mining consumer reviews," *Management Science*, 57(8) 1485–1509.

Ba, L.J., & Caruana, R. (2014, December). Do deep nets really need to be deep? Proceedings of the 27th International Conference on Neural Information Processing Systems, 2: 2654-2662.

Ballings, M., Van den Poel, D., Hespeels, N. , & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. *Expert Systems with Applications*, 42 (20): 7046–7056.

Bang, Youngsok, Lee, Dong-Joo, Han, Kunsoo, Hwang, Minha, Ahn, Jae-Hyeon (2013), "Channel Capabilities, Product Characteristics, and Impacts of Mobile Channel Introduction," *Journal of Management Information Systems*, 30(2) 101-125.

Barrilliance, Stephan Serrano, 2020, Complete List of Cart Abandonment Statistics: 2006-2020, Available at https://www.barilliance.com/cart-abandonment-rate-statistics/

Bart, Yakov, Andrew T. Stephen, and Miklos Sarvary (2014), "Which Products are Best Suited to Mobile Advertising? A Field Study of Mobile Display Advertising Effects on Consumer Attitudes and Intentions," *Journal of Marketing Research*, 51 (3), 270-85.

Batra, Rajeev and Kevin Keller (2016), "Integrating Marketing Communications: New Findings, New Lessons and New Ideas," *Journal of Marketing*, 80 (6), 122–145.

Bekaert, G., & Harvey, C. R. (1997). Emerging equity market volatility. *Journal of Financial Economics*, 43 (1): 29-77.

Bell, David & Corsten, Daniel & Knox, George & Zhang, Xinmei & Dai, Yongge. (2011), "From Point of Purchase to Path to Purchase: How Pre-shopping Factors Drive Unplanned Buying," *Journal of Marketing*, 75(1), 22-2429.

Bengio, Y. (2009). Learning Deep Architectures for AI. *Journal Foundations and Trends in Machine Learning archive*, 2(1): 1-127.

Bengio, Y. (2013). Deep Learning of Representations: Looking Forward. *Statistical Language and Speech Processing. SLSP. Lecture Notes in Computer Science*, 7978. https://doi.org/10.1007/978-3-642-39593-2_1.

Bengio, Y., & Lecun, Y. (2007). Scaling learning algorithms towards AI. *In L. Bottou, O. Chapelle, D. DeCoste, & J. Weston (Eds.), Large-scale kernel machines MIT Press.*

Ben-Hur, A., & Weston, J. (2010). A user's guide to support vector machines. *Methods in Molecular Biology series*. 2010; 609:223-39. https://doi.org/10.1007/978-1-60327-241-4_13.

Berger, D., Pukthuanthong, K. J., & Yang, J. (2011). International Diversification with Frontier Markets, *Journal of Financial Economics*, 101(1), 227-242.

Boatwright, Peter, Borle Sharad, and Joseph Kadane (2003), "A Model of the Joint Distribution of Purchase Quantity and Timing," *Journal of the American Statistical Association*, 98, 564–572.

Bodyanskiy, Y., & Popov, S. (2006). Neural network approach to forecasting of quasiperiodic financial time series. *European Journal of Operational Research*, 175 (3), 1357–1366.

Bolton, Richard, David, Hand (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3) 235–249.

Bommaraju, Raghu and Sebastian Hohenberg (2018), "Self-Selected Sales Incentives: Evidence of Their Effectiveness, Persistence, Durability, and Underlying Mechanisms," *Journal of Marketing*, 82 (5), 106–24.

Bordes, A., Chopra, S., & Weston, J. (2014). Question answering with subgraph embeddings. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*. http:// arxiv.org/abs/1406.3676v3.

Boser, B.E., Guyon, I. M., & Vapnik,V.N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory*: 144-152.

Bozdogan, Hamparsum (2000), "Akaike's Information Criterion and Recent Developments in Information Complexity," *Journal of Mathematical Psychology*, 44 (1), 62–9.

Brasel, S. Adam and James Gips (2014), "Tablets, Touchscreens and Touch-pads: How Varying Touch Interfaces Trigger Psychological Ownership and Endowment," *Journal of Consumer Psychology*, 24 (April), 226–33.

Bucklin, Randolph and Catarina Sismeiro (2003), "A Model of Web Site Browsing Behavior Estimated on Clickstream Data," *Journal of Marketing Research*, 40 (3), 249-267.

Capgemini IMRG online eRetail Index 2020. Accessed at https://www.capgemini.com/gb-en/imrg-capgemini-online-retail-index/ on June 27, 2020.

Celis, E. E., & Shen, L. J. (2015). Political Cycle and Stock Market - The Case of Malaysia. *Proceedings of the Second Asia-Pacific Conference on Global Business, Economics, Finance and Social Sciences*. http://globalbizresearch.org/Vietnam_Conference/pdf/V564.pdf

Cenesizoglu, T., & Timmermann, A. (2008). Is the Distribution of Stock Returns Predictable? http://dx.doi.org/10.2139/ssrn.1107185.

Chae, Minhee, Kim, Jinwoo (2004), "Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones," *Behaviour & Information Technology*, 23(3) 165-181.

Chang, P.-C., Liu, C.-H., Lin, J.-L., Fan, C.-Y. , & Ng, C. S. (2009). A neural network with a case based dynamic window for stock trading prediction. *Expert Systems with Applications*, 36: 6889–6898.

Chevalier, Judith A, Dina Mayzlin. (2006), "The effect of word of mouth on sales: Online book reviews," *Journal of marketing research*, 43(3) 345–354.

Chin, Erika, Adrienne Porter Felt, Vyas Sekar, and David Wagner (2012), "Measuring User Confidence in Smartphone Security and Privacy," *SOUPS '12: Proceedings of the Eighth Symposium on Usable Privacy and Security*, July 11–13, Washington, DC.

Chintagunta, Pradeep K, Shyam Gopinath, Sriram Venkataraman. (2010), "The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets," *Marketing Science*, 29(5) 944–957.

Cho, J., Lee, K., Shin, E., Choy, G., & Do, S. (2016). How much data is needed to train a medical image deep learning system to achieve necessary high accuracy? https://arxiv.org/abs/1511.06348.

Close, Angeline and Monika Kukar-Kinney (2010), “Beyond Buying: Motivations Behind Consumers' Online Shopping Cart Use,” *Journal of Business Research*, 63 (9-10),

Close, Angeline and Monika Kukar-Kinney (2010), “Beyond Buying: Motivations behind Consumers' Online Shopping Cart Use,” *Journal of Business Research*, 63(9-10), 986-992.

Close, Angeline, Monika Kukar-Kinney, and Timothy K. Benusa (2012), “Toward a Theory of Consumer Electronic Shopping Cart Behavior: Motivations of e-Cart Use and Abandonment,” *Online Consumer Behavior: Theory and Research in Social Media, Advertising and e-tail*. New York, NY: Routledge.

Collobert, R. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research* 12: 2493–2537.

Cortes, C. & Vapnik, V. (1995). Support-Vector Networks. Machine Learning, 20: 273-297. https://doi.org/10.1023/A:1022627411411.

de Haan, Evert, P.K. Kannan, Peter Verhoef, and Thorsten Wiesel (2018), “Device Switching in Online Purchasing: Examining the Strategic Contingencies,” *Journal of Marketing*, 82 (5), 341-349.

de Oliveira, F. A., Nobre, C. N., & Zárate, L. E. (2013). Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index-case study of PETR4, *Petrobras, Brazil. Expert Systems with Applications*, 40 (18): 7596–7606.

Dellarocas, Chrysanthos, Xiaoquan Michael Zhang, Neveen F Awad (2007), “Exploring the value of online product reviews in forecasting sales: The case of motion pictures,” *Journal of Interactive marketing*, 21(4) 23–45.

Deltuvaitė, V. (2016). Transmission of Shocks through Stock Markets Channel: The Case of the CEECs. *Procedia Economics and Finance*. 39: 292-297

Deng, L., Li, J., Huang, J.-T., Yao, K., Yu, D., Seide, F., Seltzer, M., Zweig, G., He, X., & Williams, J. (2013). Recent advances in deep learning for speech research at Microsoft in Acoustics, *Speech and Signal Processing (ICASSP), IEEE International Conference*, 8604–8608.

Distribution of Purchase Quantity and Timing,” *Journal of the American Statistical Association*, 98, 564–572.

Dyachenko, Tatiana and Allenby Greg, “Bayesian analysis of heterogeneous mediation”, under review at the *Journal of Consumer Research*.

Eunsuk, C., Chulwoo, H., & Parka, F.C. (2017). Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies. *Expert Systems with Applications* 83: 187–205.

Evert, de Haan, P.K. Kannan, Peter, Verhoef and Thorsten, Wiesel (2018), “Device switching in online purchasing: Examining the strategic contingencies,” *Journal of Marketing*, 82 (5), 341-349.

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

Fama, E. F. (1991). Efficient capital markets: II. *The Journal of Finance*, 46 (5): 1575–1617.

Farabet, C., Couprie, C., Najman, L., & Lecun, Y. (2013). Learning hierarchical features for scene labeling, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35(8):1915-29. doi: 10.1109/TPAMI.2012.231.

Fazio, Russell H., Janet E. Ledbetter and Tamara Lowles-schwen (2000), "On the Costs of Accessible Attitudes: Detecting That the Attitude Objects Has Changed," *Journal of Personality and Social Psychology*, 78,197–210.

Filippini, Massimo, Greene, Willian, Kumar, Nilkanth and Martinez-Cruz, Adan. (2018), "A note on the different interpretation of the correlation parameters in the Bivariate Probit and the Recursive Bivariate Probit," *Economics Letters*, 167 (2018) 104-107.

Fischer, T., & Krauss, C. (2017). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2) 10.1016/j.ejor.2017.11.054.

Focuseconomics, (2018). https://www.focus-economics.com/blog/growing-and-contracting-economies-of-2018/ Accessed September 3, 2018.

Fukushima, K. (2003). Neocognitron for handwritten digit recognition. *Neurocomputing,* 51:161–180.

Garnefeld, Ina, Andreas Eggert, Sabrina V. Helm, and Stephen S. Tax (2013), "Growing Existing Customers' Revenue Streams Through Customer Referral Programs," *Journal of Marketing*, 77 (4), 17–32.

George, G., Osinga, C. E., Lavie, D., & Scott, B. (2016). Big data and data science methods for management research. *Academy of Management Journal*, 59, 1493–1507. https:// doi.org/10.5465/amj.2016.4005. Ghahramani, Z. (2015).

Gerlein, E. A., McGinnity, M., Belatreche, A. , & Coleman, S. (2016). Evaluating machine learning classification for financial trading: An empirical approach. *Expert Systems with Applications,* 54, 193–207.

Germann, Frank, Peter Ebbes, and Rajdeep Grewal (2015), "The Chief Marketing Officer Matters," *Journal of Marketing*, 79 (3), 1–22.

Germann, Frank, Peter Ebbes, and Rajdeep Grewal (2015), "The Chief Marketing Officer Matters," *Journal of Marketing*, 79 (3), 1–22.

Ghose, Anindya and Sang Pil Han (2011), "An Empirical Analysis of User Content Generation and Usage Behavior on the Mobile Internet," *Management Science*. 57(9), 1671-1691.

Ghose, Anindya, Avi Goldfarb, and Sang Pil Han (2013), "How is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research*, 24 (3), 613-631.

Ghose, Anindya, Goldfarb, Avi and Sang Pil Han (2013), "How is the Mobile Internet Different? Search Costs and Local Activities," *Information Systems Research*. 24(3), 613-631.

Ghose, Anindya, Panagiotis G Ipeirotis, Beibei Li. (2012), "Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content," *Marketing Science*,31(3) 493–520.

GlobalData (2020), Accessed at https://www.globaldata.com/one-fifth-of-uk-retail-spend-set-

Glorot, X., Bordes, A., Bengio, Y. (2010). Deep Sparse Rectifier Neural Networks. *Journal of Machine Learning Research*, 15, 513-520

Goodfellow, I. J., Warde-Farley, D., Mirza, M., Courville, A. C., & Bengio, Y. (2013). Maxout Networks, *International Conference on Machine Learning* – 28: III,1319-III,1327.

Greene, William (1998) "Gender economics courses in liberal arts colleges: Further results," *The Journal of Economic Education*. 29(4), 291–300.

Greene, William (2003) "*Econometric Analysis. 5th Edition",* Prentice Hall, Upper Saddle River.

Grewal, Dhruv, Bart Yakov, Marin Spann, and Peter Pal Zubcsek (2016), "Mobile Advertising: A Framework and Research Agenda," *Journal of Interactive Marketing*, 34 (May), 3-14.

Grewal, Dhruv, Carl-Philip Ahlbom, Lauren Beitelspacher, Stephanie M. Noble, and Jens Nordfält (2018), "In-Store Mobile Phone Use and Customer Shopping Behavior: Evidence from the Field," *Journal of Marketing*, 82 (4), 102-126.

Grewal, Lauren, and Andrew T. Stephen (2019), "In Mobile We Trust: The Effects of Mobile versus Nonmobile Reviews on Consumer Purchase Intentions," *Journal of Marketing Research*, 56 (5), 791-808.

Grewal, Lauren, Andrew T. Stephen, and Nicole Verrochi Coleman (2019), "When Posting About Products on Social Media Backfires: The Negative Effects of Consumer Identity Signalling on Product Interest," *Journal of Marketing Research,* 56 (2), 197-210.

Grewal, Lauren, Stephen, Andrew (2019) "In Mobile We Trust: The Effects of Mobile Versus Nonmobile Reviews on Consumer Purchase Intentions," *Journal of Marketing Research*. 56(5), 791-808.

Griffin, J. M., Kelly, P., & Nardari, F. (2010). Do Market Efficiency Measures Yield Correct Inferences? A Comparison of Developed and Emerging. *Markets Review of Financial Studies,* 2 (8): 3225-3277.

Griswold, Mike (2006), "Out of Stock," *Forbes,* http://www.forbes.com/2006/12/14/amr-supply stock-biz-logistics-cx_mg_1214stock.html.

Harvey, C. R. (1995). Predictable Risk and Returns in Emerging Markets. *Review of Financial Studies*, 8(3): 773–816.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV*, 770-778. doi: 10.1109/CVPR.2016.90.

Helena Yli-Renko and Ramkumar Janakiraman (2008), "How Customer Portfolio Affects New Product Development in Technology-Based Entrepreneurial Firms," *Journal of Marketing*, 72(5), 131-148.

Hilbe, Joseph. (2014), *Negative Binomial Regression*. Cambridge, UK: Cambridge University

Hinton, G.E., Osindero, S., & Teh, Y. (2006). A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, 18: 1527–1554.

Hsu, M-W., Lessmann, S., Sung, M-C., Ma, T., & Johnson, J.E.V. (2016). Bridging the divide in financial market forecasting: machine learners vs. financial economists. *Expert Systems with Applications*, 61: 215-234.

Huang, C.-J., Yang, D.-X., & Chuang, Y.-T. (2008). Application of wrapper approach and composite classifier to the stock trend prediction. *Expert Systems with Applications*, 34, 2870–2878.

Huang, F. J., & LeCun, Y. (2006). Large-scale learning with SVM and convolutional nets for generic object categorization. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, CVPR (Vol. 1, pp. 284-291). [1640771] DOI: 10.1109/CVPR.2006.164.

Huang, Guei-Hua, Nikolaos Korfiatis, and Chun-Tuan Chang (2018), "Mobile Shopping Cart Abandonment: The Roles of Conflicts, Ambivalence, and Hesitation," *Journal of Business Research*, 85, 165-174.

Huang, W., Nakamori, Y., & Wang, S.-Y. (2005).Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32: 2513–2522.

Huanhuan, Y., Rongda, C., & Zhang, G. (2014). A SVM stock selection model within PCA. *Procedia Computer Science*, 10.1016/j.procs.2014.05.284.

Hubert, Marco, Markus Blut, Christian Brock, Christof Backhaus, and Tim Eberhardt (2017), "Acceptance of Smartphone-Based Mobile Shopping: Mobile Benefits, Customer Characteristics, Perceived Risks, and the Impact of Application Context," *Psychology and Marketing*, 34(2), 175-194.

Ibikunle, G., Möws, B. (2017). How random are intraday stock prices? Evidence from deep learning. Edinburgh Business School, Working paper. Available at: https://www.research.ed.ac.uk/portal/files/32816714/Ibikunle_and_M_ws_2017.pdf (Accessed: 11 November 2018)

International Conference on International Conference on Machine Learning: 807-814.

Jaakkola, T., & Haussler, D. (1998). Exploiting Generative Models in Discriminative Classifiers, *In Advances in Neural Information Processing Systems*, II, 487-493.

Jarrett, K., Kavukcuoglu, K., Ranzato, M., & LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? *IEEE 12th International Conference on Computer Vision, Kyoto*: 2146-2153, doi: 10.1109/ICCV.2009.5459469.

Kannan, P.K. and Hongshuang (Alice) Li (2017), "Digital Marketing: A Framework, Review and Research Agenda," *International Journal of Research in Marketing*, 34 (1), 22–45.

Kannan, P.K., and Hongshuang (Alice) Li (2017), "Digital Marketing: A Framework, Review and Research Agenda," *International Journal of Research in Marketing*, 34 (1), 22–45.

Kara, Y., Acar Boyacioglu, M., & Baykan, Ö. K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications*, 38: 5311–5319.

Karemera, D., Ojah, K., Cole, J.A. (1999). Random Walks and Market Efficiency Tests: Evidence from Emerging Equity Markets. *Review of Quantitative Finance and Accounting* (1999) 13: 171. https://doi.org/10.1023/A:1008399910942.

Kearney, C. (2012). Emerging Markets Research: Trends, Issues and Future Directions. *Emerging Markets Review*, 13 (2): 159-183.

Kibo (2016), "eCommerce Performance Index," http://www.marketlive.com/in-the-news/press/performance-index-v31.html.

Krauss, C., Anh Do, X., & Huck, N. (2016). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 10.1016/j.ejor.2016.10.031.

Krizhevsky, A., Ilya, S., & Hinton, G.E. (2012). ImageNet classification with deep convolutional neural networks. *International Conference on Neural Information Processing Systems*, 1: 1097-1105.

Kukar-Kinney, Monika and Angeline Close (2010), "The Determinants of Consumers 'Shopping Cart Abandonment," *Journal of the Academy of Marketing Science*, 38 (2), 240-250.

Kumar, V. and Rajkumar Venkatesan (2005), "Who are the Multichannel Shop-pers and How Do They Perform?: Correlates of Multichannel Shopping Behavior," *Journal of Interactive Marketing*, 19 (2), 44–62.

Lambert, Diane (1992), "Zero-Inflated Poisson Regression, With an Application to Defects in Manufacturing," *Technometrics*, 34 (1), 1-14.

Law, Rob, Elise Wong, Dimitrios Buhalis, and Richard Hatter (2018), "Time-Varying Browsing Behavior of Hotel Website Users," *E-Review of Tourism Research*.

Le, Q.V., Ranzato, M.A., Monga, R., Devin, M., Chen,K., Corrado, G.S., Dean, J., & Ng, A.Y.(2012). Building high-level features using large scale unsupervised learning. *International Conference on Machine Learning. Proceedings of the 29th*: 507-514.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521 (7553): 436–444. http://dx.doi.org/10.1038/nature14539 10.1038/nature14539.

Lecun, Y., Farabet, C., Couprie C., & Najman, L. (2013). Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 35(8):1915-29 doi: 10.1109/TPAMI.2012.231.

Lesmond, D. A., Schill, M. J., & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics*, 71(2): 349–380.

Lessmann, S., Sung, M., & Johnson, J. E. (2011). Towards a methodology for measuring the true degree of efficiency in a speculative market. *Journal of the Operational Research Society*, 62: 2120–2132.

Leung, M. K., Xiong, H. Y., Lee, L. J., & Frey, B. J. (2014). Deep learning of the tissue regulated splicing code. *Bioinformatics* 30: i121–i129.

Li, Chenxi, Xueming Luo, Cheng Zhang, and Xiaoyi Wang (2017), "Sunny, Rainy, and Cloudy with a Chance of Mobile Promotion Effectiveness" *Marketing Science*, 36 (5), 762–779.

Li, Hongshuang (Alice), and P.K. Kannan (2014), "Attributing Conversions in a Multichannel Online Marketing Environment: An Empirical Model and a Field Experiment," *Journal of Marketing Research*, 51 (1), 40–56.

Liu, Zhen, Shao-hui Lei, Yulang Guo, and Zhi-an Zhou (2020), "The Interaction Effect of Online Review Language Style and Product Type on Consumers' Purchase Intentions," *Palgrave Communications*, 6 (1), 11.

Lo, A. W. (1991). Long-Term Memory in Stock Market Prices. *Econometrica*, 59 (5), 1279–1313.

Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Phang (2014), "Mobile Targeting," *Management Science*, 60 (7), 1738-1756.

Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E., & Svetnik, V. (2015). Deep neural nets as a method for quantitative structure-activity relationships. *Journal of Chemical Information and Modeling*, 55: 263–274.

Maddala GS (1983), "*Limited Dependent and Qualitative Variables in Econometrics,"* Cambridge University Press: Cambridge.

Mallapragada, Girish, Chandukala, Sandeep and Liu, Qing (2016), "Exploring the effects of "what" (product) and "where" (website) characteristics on online shopping behavior," *Journal of Marketing*. 80(2), 21-38.

Mallapragada, Girish, Sandeep Chandukala, and Qing Liu (2016), "Exploring the Effects of Manufacturing," *Technometrics*, 34 (1), 1-14.

Marketing in the Retailing Environment: Current Insights and Future Research Avenues," *Journal of Interactive Marketing*, 24 (2), 111-120.

Marz, Armin, Sebastian Schubach, and Jan H. Schumann (2017), "Why Would I Read a Mobile Review?" Device Compatibility Perceptions and Effects on Perceived Helpfulness," *Psychology & Marketing*, 34 (2), 119-37.

Maslowska, Ewa, Edward C. Malthouse, and Stefan F. Bernritter (2017), "Too Good to Be True: The Role of Online Reviews' Features in Probability to Buy," *International Journal of Advertising*, 36(1), 142-163.

Mathieu, M., LeCun, Y., Fergus, R., Eigen, D., Sermanet, P., & Zhang, X. (2013). OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. *International Conference on Learning Representations (ICLR2014)*, http://arxiv.org/abs/1312.6229.

McCullagh, Peter and John Nelder (1989), Generalized Linear Models. 2nd ed. London: Chapman & Hall.

Melumad, Shiri, Jeffrey Inman, and Michel Tuan Pham (2019), "Selectively Emotional: How Smartphone Use Changes User-Generated Content" *Journal of Marketing Research*. 56(4), 259-275.

Michelle Andrews, et al. (2017), "Consumer Connectivity in a Complex, Technology-Enabled, and Mobile-Oriented World with Smart Products," *Journal of Interactive Marketing*, 40, 1–8.

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and e-Business Management*, 16, 547–578. https://doi.org/10.1007/s10257-017-0362-y.

Mittermayer, M.-A. (2004). Forecasting Intraday stock price trends with text mining techniques. *Proceedings of the 37th annual Hawaii international conference*: 10.
Mobile Word of Mouth: How are Mobile Reviews Different?," *Marketing Science*, 38 (5), 773-792.

Moe, Wendy W. (2003), "Buying, Searching, or Browsing: Differentiating Between Online Shoppers Using In-Store Navigational Clickstream," *Journal of Consumer Psychology*, 13 (1/2), 29–39.

Montgomery, Alan, Shibo Li, Srinivasan, Kannan and John Liechty (2004), "Modelling Online Browsing and Path Analysis Using Clickstream Data," *Marketing Science*, 23 (4), 579-595.

Moorthy, Sridhar, Brian T. Ratchford, and Dababrata Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research*, 23 (4), 263–77.

Mullahy, John (1986), Specification and Testing of Some Modified Count Data Models," *Journal of Econometrics*," 33, (3), 341-365

Nair, V., & Hinton, G.E. (2010). Rectified linear units improve restricted boltzmann machines. *International Conference on International Conference on Machine Learning*: 807-814.

Napoli, Philip and Obar, Jonathan (2014), "The Emerging Mobile Internet Underclass: A Critique of Mobile Internet Access," *The Information Society*. 30(5) 323-334.

Neal, David T., Wendy Wood, Jennifer S. Labrecque and Phillippa Lally (2012), "How Do Habits Guide Behavior? Perceived and Actual Triggers of Habitsin Daily Life," *Journal of Experimental Social Psychology,* 48 (March), 492–8.

Northwestern University's Spiegel Research Center. 2017, How Online Reviews Influence Sales. Available at https://spiegel.medill.northwestern.edu/online-reviews/

O'Hara, Kenton, Mitchell, April and Vorbau, Alex (2007), "Consuming video on mobile devices" *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* 857-866.

Oliver, Richard and Shor, Mikhael (2003), "Digital redemption of coupons: Satisfying and dissatisfying effects of promotion codes," *Journal of Product and Brand Management.* 12(2), 121.

Pan, H., Chandima, T., & Yearwood, J. (2005). Predicting Australian Stock Market Index Using Neural Networks Exploiting Dynamical Swings and Intermarket Influences. *Journal of Research and Practice in Information Technology,* 37(1) 10.1007/978-3-540-24581-0_28.

Papies, Dominik, Peter Ebbes, and Harald J. van Heerde (2017), "Addressing endogeneity in marketing models," *in Advanced Methods for Modeling Markets*. Cham, Switzerland: Springer.

Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques. *Expert Systems with Applications,* 42(1), 259-268.

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

Porteus, Evan (2002), *Foundation of Stochastic Inventory Theory*. Stanford, CA: Stanford Business Books.

Qian, B., & Rasheed, K. (2007). Stock market prediction with multiple classifiers. *Applied Intelligence,* 26: 25–33 http://doi.org/ . doi: 10.1007/s10489- 006- 0001- 7.

Rajamma, Rajamma, Paswan, Audhesh, and Hossain, Muhammad (2009), "Why do shoppers abandon shopping cart? Perceived waiting time, risk, and transaction inconvenience*,"* *Journal of Product and Brand Management*, 18(3), 188–197.

Ransbotham, Sam, Lurie, Nicholas, and Liu, Hongju (2019), "Creation and Consumption of Mobile Word of Mouth: How Are Mobile Reviews Different?" *Marketing Science*, 38, 773-792.

Ransbotham, Sam, Lurie, Nicholas, and Liu, Hongju (2019), "Creation and Consumption of Mobile Word of Mouth: How Are Mobile Reviews Different?" *Marketing Science*, 38, 773-792.

Rebecca Jen-Hui Wang, Edward C. Malthouse, Lakshman Krishnamurthi (2015), "On the Go: How Mobile Shopping Affects Customer Purchase Behavior," *Journal of Retailing*, 91(2), 217-234.

Ren, Y., & Wu, Y. (2014). Convolutional deep belief networks for feature extraction of EEG signal. *International Joint Conference on Neural Networks (IJCNN), Beijing*, 2850-2853. doi: 10.1109/IJCNN.2014.6889383.

Retail Sales, Office for National Statistics, GB (2020), Accessed at https://www.ons.gov.uk/

Rossi, Peter E. and Greg M. Allenby (2003), "Bayesian Statistics and Marketing," *Marketing Science*, 22 (3), 304–328.

Rumelhart, D. E., Hinton, G. E., &Williams, R. J. (1986). Learning internal representations by error-propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition.* Volume 1, Issue 6088: 318-362.

Saad, Lydia (2015), "Most Smartphone Users Still Rely on Computer for Web Purchases," (accessed July 1, 2020), http://www.gallup.com/poll/184037/smartphone-users-rely-computerweb-purchases.aspx.

Sadique, S., & Silvapulle, P. (2001). Long-Term Memory in Stock Market Returns: International Evidence. *International Journal of Finance & Economics*. (6) 59-67. 10.1002/ijfe.143.

Salakhutdinov, R., Mnih, A., & Hinton, G. (2007). Restricted Boltzmann machines for collaborative filtering. *Proceedings of the 24th International Conference on Machine learning.*791-798.

Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd (2010), "Can There ever Be Too Many Options? A Meta-analytic Review of Choice Overload," *Journal of Consumer Research*, 37(3), 409-425.

Schindler, A., Lidy, T., Rauber, A. (2016). Comparing shallow versus deep neural network architectures for automatic music genre classification. *In Proceedings of the 9th Forum Media Technology (FMT2016), St. Poelten, Austria.*

Scholkopf, B., & Cristianini, N. (2002). Support vector machines and kernel methods: The new generation of learning machines. *Ai Magazine,* Volume 23 Issue 3: 31-41.

Schumaker, R. P., & Chen, H. (2009). Textual analysis of stock market prediction using breaking financial news: The AZFin text system. *ACM Transactions on Information Systems* (TOIS), 27(12).

Seide, F., Li, G., Xie, C., & Dong, Y. (2011). Feature engineering in Context-Dependent Deep Neural Networks for conversational speech transcription. *IEEE Workshop on Automatic Speech Recognition & Understanding*. 10.1109/ASRU.2011.6163899.

Shankar, Venkatesh, Alladi Venkatesh, Charles Hofacker, and Prasad Naik (2010), "Mobile Marketing in the Retailing Environment: Current Insights and Future Research Avenues," *Journal of Interactive Marketing*, 24 (2), 111–12.

Shawe-Taylor, J., & Cristianini, N. (2004). Kernel Methods for Pattern Analysis. *Cambridge university press.*

Shopping and Social Media: Friends or Foes?," *Journal of Marketing*, 81 (6), 24-41.

Silverman, Bernard W. (1986), *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., & Salakhutdinov R (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research Volume* 15 Issue 1: 1929-1958.

Steel, R.G.D., Torrie, J.H. (1980). *Principles and Procedures of Statistics, Second Edition, New York: McGraw-Hill Book Co.*

Sterling, Greg (2016), "All Digital Growth Now Coming from Mobile Usage—comScore," *Marketing Land*, http://marketingland.com/digital-growth-now-coming-mobile-usage-comscore-171505.

Suited to Mobile Advertising? A Field Study of Mobile Display Advertising Effects on Consumer Attitudes and Intentions," *Journal of Marketing Research*, 51 (3), 270-85.

Sutsarun, L., Sirimon, T., Wee, M., & Brooks, R (2014). Thai Financial Markets and Political Change. *Journal of Financial Management, Markets and Institutions*, Issue 1: 5-26.
The Impact of Tablets on Digital Commerce," *Management Science*, 63 (5), 1469–92.

Timmermann, A., & Granger, C.W.J. (2004). Efficient market hypothesis and forecasting. *International Journal of Forecasting* Volume 20 Issue 1:15–27.

Titan, A. G. (2015). The Efficient Market Hypothesis: review of specialized literature and empirical research. *Procedia Economics and Finance* 32(2015) 442 – 449.
to-be-online-by-2024/ on June 27, 2020.

Tompson, J., Stein, M., LeCun, Y., & Perlin, K., (2014). Real-time Continuous Pose Recovery of Human Hands Using Convolutional Networks. *ACMTrans. Graph*. doi:10.1145/2629500.

Töscher, A., Jahrer, M., & Bell, R. M. (2009).The Big Chaos solution to the Netflix grand prize. http://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf.

Verhoef, Peter C., Andrew T. Stephen, P.K. Kannan, Xueming Luo, Vibhanshu Abhishek, Michelle Andrews, et al. (2017), "Consumer Connectivity in a Complex, Technology-Enabled, and Mobile-Oriented World with Smart Products," *Journal of Interactive Marketing*, 40, 1–8.

Verhoef, Peter C., Edwin Kooge, and Natasha Walk (2016), *Creating Value with Big Data Analytics: Making Smarter Marketing Decisions*. London: Routledge.

Verhoef, Peter C., Pallassana K. Kannan, and J. Jeffrey Inman (2015), "From Multi-Channel Retailing to Omni-Channel Retailing: Introduction to the Special Issue on Multi-Channel Retailing," *Journal of Retailing* 91 (2), 174-18.

Verhoef, Peter C., Scott A. Neslin, and Björn Vroomen (2007), "Multichannel Customer Management: Understanding the Research-Shopper Phenomenon," *International Journal of Research in Marketing* 24 (2), 129-148.

Verhoef, Peter C., Thijs Broekhuizen, Yakov Bart, Abhi Bhattacharya, John Qi Dong, Nicolai Fabian, and Michael Haenlein (2021), "Digital Transformation: A Multidisciplinary Reflection and Research Agenda," *Journal of Business Research* 122 (1), 889-901.

Verplanken, Bas (2006), "Beyond Frequency: Habit as Mental Construct," *British Journal of Social Psychology*, 45 (September), 639–56.

Wangenheim, Florian v. and Toma´s Bay´on (2007), "Behavioral Consequences of Overbooking Service Capacity," *Journal of Marketing*, 71 (4), 36–47.

What (Product) and Where (Website) Characteristics on Online Shopping Behavior" *Journal of Marketing*, 80 (2), 21-38.

Williams, C. K. I., & Rasmussen, C.E. (1996). Gaussian Processes for Regression. *In Advances in Neural Information Processing Systems,* 8. eds. D. S. Touretzky, M. C. Mozer, M. E. Hasselmo, MIT Press.

Wolfinbarger, Mary and Mary C. Gilly (2003), "eTailQ: Dimensionalizing, Measuring and Predicting Etail Quality," *Journal of Retailing*, 79 (3), 183-198.

Wooldridge, Jeffrey M. (2010), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

Wu, Jen-Her, & Hisa, Tzyh-Lih (2004), "Analysis of e-commerce innovation and impact: A hypercube model," *Electronic Commerce Research and Applications*, 3(4), 389–404.

Xu, Kaiquan, Jason Chan, Anindya Ghose, and Sang Pil Han (2017), "Battle of the Channels: The Impact of Tablets on Digital Commerce," *Management Science*, 63 (5), 1469–92.

Xu, Q., Zhou, H., Wang, Y., & Huang, J. (2009). Fuzzy support vector machine for classification of EEG signals using wavelet-based features. *Journal of Medical Engineering & Physics*, 31(7):858-65. doi: 10.1016/j.medengphy.

Xu, Yin and Jinsong Huang (2015), "Factors Influencing Cart Abandonment in the Online Shopping Process," *Social Behavior and Personality: An International Journal*, 43(10), 1617-1627

Yang, Kenneth (2005), "Exploring factors affecting the adoption of mobile commerce in Singapore," *Telematics and Informatics*, 22(3), 257–277.

Yang, Kiseol, and Forney, Judith (2013), "The moderating role of consumer technology anxiety in mobile shopping adoption: Differential effects of facilitating conditions and social influences," *Journal of Electronic Commerce Research*, 14(4), 334–347.

Żbikowski, K. (2015). Using Volume Weighted Support Vector Machines with walk forward testing and feature selection for the purpose of creating stock trading strategy. *Expert Systems with Applications*, 42(4): 1797-1805.

Zhang, B. G., Patuwo, E, & Hu,Y. M (1998). Forecasting with artificial neural networks: *The state of the art. International Journal of Forecasting*, 14 (1): 35-62.

Zhang, Y., & Wu, L. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications*, 36(5): 8849-8854.

Zhang, Yuchi, Michael Trusov, Andrew T. Stephen, and Jamal Zainab Jamal (2017), "Online Shopping and Social Media: Friends or Foes?," *Journal of Marketing*, 81 (6), 24-41.