The Two-Stage Decision Process During Online Purchasing-Empirical and Observational Field Studies on Cognitive, Affective and Behavioral Outcomes

---

AM Fachbereich Rechts- und Wirtschaftswissenschaften
an derTechnischen Universität Darmstadt

DISSERTATION

vorgelegt von

Wenyan Zhou
Dipl.Inf(FH), M.Sc.
geboren am 06.15.1987 in Zhejiang, China

zur Erlangung des akademischen Grades
Doctor rerum politicarum (Dr. rer. pol.)

Erstgutachter:	Prof. Dr. Oliver Hinz
Zweitgutachter:	Prof. Dr. Alexander Benlian
Tag der Einreichung:	30.11.2017
Hochschulkennziffer:	D17

Wenyan, Zhou: The Two-Stage Decision Process During Online Purchasing- Empirical and Observational Field Studies on Cognitive, Affective and Behavioral Outcomes
Darmstadt, Technische Universität Darmstadt,
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
Tag der mündlichen Prüfung: 04.04.2018

Veröffentlicht unter CC-BY-NC 4.0 International – Creative Commons Attribution
https://creativecommons.org/licenses/

Ehrenwörtliche Erklärung

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig angefertigt habe. Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Die Arbeit wurde bisher nicht zu Prüfungszwecken verwendet und noch nicht veröffentlicht.

Darmstadt, den 30. November 2017

Abstract

Of the past 20 years, online shopping has gained growing importance in both consumers' daily life and the retail industry. E-retailers such as Amazon create new sales records yearly. However, there is a big challenge for e-retailers to maintain impressive sales and meanwhile handle the increasing returns (i.e., consumers returning products to get their money back). Although several researchers have identified a need for more research regarding why and how consumers make series decisions to purchase and return items, limited research has been conducted in this context. Utility theory is widely used to explain consumers' purchase and return decisions separately. However, less research considers emotions as part of consumers' decision-making process.

Against this backdrop, this thesis presents three studies of consumers' purchase decisions, return decisions and the integrated reaction process. The first draws on the random utility model by using the choice based conjoint analysis method. A laboratory experiment reveals that providing comprehensive services could efficiently enhance consumers' purchase intentions in the electric vehicle market. The second study is from the perspective of small- and middle-size e-retailers. We analyzed the transactional data from a Chinese e-platform, and found that return policies' credibility has a positive influence on both sales and returns. The third study focuses on identifying the processes involved in consumers' return decisions in online purchasing markets by considering both cognitive reactions (utility-based) and affective reactions (emotion-based). The results of two experiments and a field study show that comprehensive services such as a gift and a colorful package can reduce consumers' return intentions by increasing the perceived utilities and creating positive emotions. By demonstrating why and how businesses can influence purchase and return decisions, this thesis enriches the consumer behavior research. E-retailers may use the results to understand

consumer behavior better and provide advanced services and policies to attract purchases and meanwhile avoid unnecessary returns.

.

Zusammenfassung

In den letzten 20 Jahren hat das Online-Shopping sowohl im täglichen Leben der Konsumenten als auch im Einzelhandel gravierend an Bedeutung gewonnen. Internethändler wie Amazon erzielen jährlich neue Verkaufsrekorde. Eine große Herausforderung besteht jedoch für Internethändler darin, beeindruckende Umsätze beizubehalten und gleichzeitig die steigende Zahl der Retouren (d.h. Produkte, die zum Erhalt von Rückerstattung des Kaufpreises zurückgesandt werden) zu bewältigen. Obwohl mehrere Forscher auf den Bedarf an mehr Forschungsarbeit bezüglich der Frage, warum und wie Verbraucher aufeinanderfolgende Kauf- und die Rückgabe-Entscheidungen treffen, hingewiesen haben, wurden in diesem Zusammenhang nur begrenzte Untersuchungen durchgeführt. Die Nutzentheorie wird weithin verwendet, um Kauf- und Retouren-Entscheidungen der Verbraucher separat zu erklären. Die Auswirkung von Emotionen auf den Entscheidungsprozess der Konsumenten wurde bislang jedoch weniger untersucht.

Vor diesem Hintergrund werden in dieser Arbeit drei Studien vorgestellt, welche die Entscheidungen zum Kauf und der Rückgabe wie auch den integrierten Reaktionsprozess untersuchen. Die erste Studie stützt sich auf das Random-Utility-Modell unter Verwendung der auf Entscheidungen basierenden Conjoint-Analyse-Methode. Eine experimentelle Studie hat ergeben, dass sich die Bereitstellung umfassender Dienstleistungen effizient auf die Kaufabsichten der Verbraucher auf dem Markt für Elektrofahrzeuge auswirken könnte. Die zweite Studie bezieht sich auf die Perspektive kleinerer und mittlerer Internethändler. Eine Analyse der Transaktionsdaten einer chinesischen E-Plattform ergab, dass die Glaubwürdigkeit von Rückgaberichtlinien einen positiven Einfluss auf Verkauf und Retouren hatte. Der Fokus der dritten Studie lag auf der Identifikation der Prozesse, die bei Rückgabeentscheidungen von Konsumenten auf Online-Beschaffungsmärkten maßgeblich

sind, wobei sowohl kognitive (auf dem Nutzwert basierende) Reaktionen als auch affektive (emotionsbasierte) Reaktionen berücksichtigt wurden.

Zwei Experimente und eine Feldstudie ergaben, dass umfassende Dienstleistungen, wie kleine Kundengeschenke und farbenfrohe Verpackungen zur Reduzierung der Rückgabeabsichten bei Verbrauchern führen können, indem sie den wahrgenommenen Nutzen erhöhen und positive Emotionen erzeugen. Indem sie aufzeigt, weshalb und auf welche Weise Unternehmen Einfluss auf Kauf- und Rückgabe-Entscheidungen nehmen können, leistet die vorliegende Arbeit einen bereichernden Beitrag zur Verbraucherverhaltensforschung. Internethändler können die Resultate zum besseren Verständnis des Verbraucherverhaltens nutzen und fortschrittliche Dienstleistungen und Richtlinien anbieten, um den Umsatz zu steigern und unnötige Retouren zu vermeiden.

.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
AVE	Average Variance Extracted
CR	Composite Reliability
DISE	Dynamic Intelligent Survey Engine
EV	Electric Vehicles
PA	Pleasure and Arousal Model
PLS	Partial Least Squared
RP	Revealed Preference
RUT	Random Utility Theory
SD	Standard Deviation
SP	Stated Preference
U.S.	Unites States
V2G	Vehicle-to-Grid
ZINB	Zero-inflated Negative Binomial Regression
ZIP	Zero-inflated Poisson Model

Chapter 1: Synopsis

1.1 Motivation and Research Questions

Online shopping is a popular online activity, which was worth 1.9 trillion United States (U.S.) dollars in the year 2016. This activity constitutes of 8.7% of the global retail sales (statista 2017). Sustained growth in sales are well foreseen. Statista (2017) shows 4.06 trillion U.S. dollars in growth by 2020. On Asia's biggest online shopping platform, Alibaba, the singles' day (Nov. 11th, 2017) sale amassed 25.3 billion U.S. dollars in revenues (Forbes 2017). This information paints a rosy picture of online shopping, but it does not consider the large number of returns that are hidden on the dark side.

In the U.S., the average return rate of online shopping reached 33% in 2012 and is projected to increase even further in the future (Banjo 2013). Hence, when the costs of reverse logistics and/or the extra labor and management costs for handling the returns are considered (Petersen and Kumar 2009), the net profits of the online shopping industry may not be that impressive. Thus, both merchants and scholars should keep in mind that online shopping includes both sales and returns.

From the behavioral perspective, the key characteristics that differentiate online shopping from the more traditional offline shopping is the two-stage decision-making process (Minnema et al. 2016; Wood 2001): the decision to order (the purchasing stage) and the decision to keep or return the ordered product (the post-purchasing stage). The purchasing

decision itself is a more time- and effort- consuming process than the keep or return decision. Consumers typically spend a great deal of time collecting and processing information from several e-retailers and evaluating an item according to its prices, product attributes (demonstrated on sites), e-retailers' reputation and users' feedback (as their expectations). In contrast, consumers usually make the keep or return decision quickly. After receiving the delivered package, consumers re-evaluate the item, based on the received/real product attributes and the previous purchasing information (e.g., price). Because consumers may not visit the respective website or check old e-mails to reacquaint themselves with actual purchasing information, the previous evaluation may be effective only to the extent their memory allows.

Information asymmetry and a time delay exist between the two stages as well. Information asymmetry occurs because consumers cannot touch, feel and try out the item during the online purchasing process (Wood 2001). The period of delay caused by the delivery process may allow the consumers' memory about the purchasing details to fade and allow other unconfirmed information (e.g., competitor's ads) to come to fore (Bechwati and Siegal 2005).

In summary, the judgment made during the post-purchasing stage might be based on factors that differ from those used during the purchasing stage, namely a very vague memory about the purchasing details, related fresh memory about the unconfirmed information and the perceived value of the real product. This situation leads to a mismatch, or an expectation

gap, between the two evaluations (Bechwati and Siegal 2005). The bigger the expectation gap, the higher is the likelihood to return the product (Anderson et al. 2009).

Against this background, merchants and scholars have attempted to narrow down the consumer expectation gap. This is generally accomplished in three ways. First, e-retailers improve their sites by using more detailed demonstrations (e.g., 3D display technology) and building a complete online customer review system. This strategy can offer more reference information about products, including other users' feedback (Sahoo et al. 2015). Second, they research and apply suitable return policies. Return policies can shorten consumers' pre- and post-purchasing search and deliberation time (Wood 2001). There policies can also positively influence the evaluations during both stages of the process (Kim and Wansink 2012). Third, e-retailers offer better complementary service (e.g., fast delivery to shorten the time delay (Jiang and Rosenbloom 2005) and personalized packages to positively influence the post-purchasing evaluation and long-term loyalty).

The majority of studies in the field of online shopping have only focused on purchasing decisions or return decisions. A limited number of studies consider the pre- and post-purchasing stage in the entire online shopping process. Few studies explore the decision mechanism behind these joint decisions. Thus, to bridge the current research gaps, the primary goal of this dissertation is to enhance our understanding regarding the working mechanism behind online purchasing and return behavior. This dissertation is guided by the two following two overarching research questions:

- RQ1: How does the return policy influence the balance between sales and returns?
- RQ2: What is the work mechanism underlying the decisions to keep or return products?

1.2 Theory Background

The theory behind decision-making studies is evaluated, following the changes in the perspective of human rationality at various points in time. The original studies considered decision making to be an entirely rational process. In these studies, the most commonly applied model was the "Utility Theory", emphasizing decision making on the basis of maximize the expected outcomes (Thaler 1985). The utility model considers consumers to be bounded/limited by rationality, according to the limitations of computational capabilities and the availability of information (Simon 1972). Subsequent developments consider emotions as an integral part of the decision-making process (Buchanan and O Connell 2006).

1.2.1Utility Theory and Its Expending

In the purchasing stage, utility theory assumes that consumers choose the option that maximizes their utility which is the perceived benefit received from the ordered items and the purchasing price (cognitive responses) (Thaler 1985). More specifically, a customer "h" decomposes his or her utility $U(purchase)_{h,i}$ for the alternative "i" into a deterministic utility $\mu_{h,i}$ and a stochastic component $\varepsilon_{h,i}$ (See Equation 1). In other words, the perceived utility by consumers can be decomposed into components that are directly related to the

attributes and levels $X_{h,i}$ shown to the consumers and components that are only known to the consumer (See Equitation 2).

$$U(purchase)_{h,i} = \mu_{h,i} + \varepsilon_{h,i} \quad (1)$$

where:

$$\mu_{h,i} = \beta'_{h,i} X_{h,i} \quad (2)$$

Considering that a consumer's purchasing choice is among a choice set from the same or different e-retailers, the random utility theory is developed. This theory is based on limited rationality and assumes consumer "h" would only purchase product "i" if the difference of $U(purchase)_{h,i}$ and $P_{h,i}$ is positive and higher than that of the other alternatives "j".

$$Pr_{h,i} = Prob(\mu_{h,i} + \varepsilon_{h.i} - P_{h,i} > \mu_{h,j} + \varepsilon_{h.j} - P_{h,j};\ \forall j \neq i) \quad (3)$$

Similarly, researchers have used the utility theory to build the return decision (Anderson et al. 2009). They assume consumer h's utility, at the post-purchasing stage, $U(return)_{h,i}$ consists of the components of the second-round deterministic perceived utility $\mu'_{h,i}$ focused on the received item i, the standard econometric error $\varepsilon_{h,i}$ and the return cost $R_{h,i}$.

$$U(return)_{h,i} = \mu'_{h,i} + \varepsilon'_{h,i} - R_{h,i} \quad (4)$$

$$Re_{h,i} = Prob(\mu'_{h,i} - R_{h,i} + \varepsilon_{h.i} < P_{h,i}) \quad (5)$$

Consumer "h" would only return the product "i" if the difference between $U(return)_{h,i}$ and $P_{h,i}$ is negative. Here, $\mu'_{h,i}$ might significantly differ from the $\mu_{h,i}$ evaluated in the purchasing stage, which generates the expectation gap ("$\mu'_{h,i} - \mu_{h,i}$", could be negative or positive) as mentioned in Session 1.1. The range of $\mu'_{h,i}$ could between 0 (which means the received item is useless for the consumer) or much higher than $\mu_{h,i}$ (which means the

received item is beyond the consumer's expectation). The return cost $R_{h,i}$ includes not only the delivery fee, but also the effort required to send the items back. When the negative expectation gap appears and the return cost is low enough, consumer "h" would be more likely to return the product and get his or her money back.

1.2.2 Two-stage Reactions Process

Bettman et al. (1998) pointed out that the utility based model (cognitive reaction) only provides a partial explanation for a response to an external stimulus. The affective reaction (i.e., emotions) also plays a significant role in the consumer decision-making process. However, other than viewing the cognitive view and the affective reactions, as separate and independent effects (i.e., Rook 1987), a rich body of work understands them as sustainable outcomes during the online shopping process (Aydinli 2014; Suwelack 2011; Amanda 2012).

There are two main research streams for the combinations of affective and cognitive reactions. The first research stream uses the affective-cognitive reaction order (Bechara et al. 2000). Studies in this stream consider emotions as an influencing factor, because the affective responses are effortless, and therefore, primed by a default whenever a behavioral opportunity presents itself. In contrast, the cognitive response requires more processing resource, and thus, turns to play a secondary place during the entire reaction process (Aydinli et al. 2014). The other research stream applies the cognitive-affective reaction order. This considers the affective reactions that can arise in a relatively controlled, post cognitive manner from a deeper higher-order processing of the incoming information. For instance,

consumers might feel regret when they realize they could have made a different and better purchase decision (Suwelack, 2011), or when they find that the return policy offered by an e-retailers are made lacks of fairness (Amanda 2012).

The key differences between these two separated research streams are the two types of affective reactions (Berkowitz 1993; LeDoux 1995; Shiv and Fedorikhin 1999). The one estimated in the affective-cognitive process is the "low-road" affective reaction, which is evoked rapidly and automatically in the limbic systems of the brain. The second affective reaction that is evaluated in the cognitive-affective process is the "high-road" affective reaction, which arises from the outcomes of the second process (Shiv and Fedorikhin 1999) and occurs relatively slowly, when compared to the low-road affective reactions.

Many studies have evaluated the affective-cognitive reaction process during the purchasing decision process and the cognitive-affective reaction process in the field of a post-purchasing period (i.e., service failures) (Aurier and Guintcheva 2014; Ladhari 2007; Oliver 1993). Service failures and product returns entail a similar decision-making process, insofar as the affected consumers can apply for a refund, or compensation, when they are not satisfied with the offered service/product. In the context of service failures, the type and degree of a service failure --- in tandem with the offered remedial measures --- can sway consumers' emotions. In line with the Appraisal-Tendency framework (Scherer et al. 2001), these affective responses can then influence consumers' post-purchasing decisions (e.g., loss or maintenance of access to goods/services), and ultimately, their satisfaction (Hibbert et al. 2012) and loyalty (DeWitt et al. 2008). More specifically, when consumers have more

negative emotions towards a service failure or are more aware of it, they are less likely to be satisfied and might not maintain their purchasing decisions (Choi and Mattila 2008).

However, we found no literature referring to the multiple-stage reaction process in a product return context—not even studies solely addressing the affective reactions.

1.3 Structure of the Dissertation

Overall, the dissertation consists of four chapters which contain three published research articles. Table 1-1 provides an overview of the three articles including research goal, study type, dataset, analytical method and publication status. Chapter 2 discusses Article 1, which focus on customers' purchase decisions in electric vehicles markets, from the perspective of the influences from complementary mobility services. Chapter 3 (Article 2) considers both purchase and return behavior in online purchase markets, by investigating the influences of various return policies. Afterward, Chapter 4 (Article 3) further explores customers' return/keep decision-making process after the online purchase, with a careful consideration of a package-opening process. In the following, each research is briefly summarized, including the main motivation, method, data set and findings.

Chapter 2 (Article 1): Fostering the Adoption of Electric Vehicles by Providing Complementary Mobility Services: A Two-step Approach using Best-Worst Scaling and Dual Response

Motivation: A broader adoption of electric vehicles (EVs, including Plug-in electric cars, hybrid electric cars, and hydrogen vehicles) can relieve both the CO2 emissions related climate problem and the problem of the oil demand gap. Previous research has noted the importance of price (including purchase price and recharging costs), charging time and driving ranges in the (non-)adoption of EVs (Beggs and Cardell 1981; Bunch et al. 1993), but has mostly been neglected the potential positive influences from complementary mobility services.

Method and Data Set: To address this gap, we employ a two-step approach utilizing a hybrid stated preference method. After investigating 251 participants by using Best-Worst Scaling (Study 1), we firstly identify the top three (i.e., IT-based parking and payment, intelligent charging stations and augmented reality services via head-up displays) options from the nice complementary mobility services consumers would prefer with an electric vehicle. In Study 2, we investigate additional 327 participants by a Dual Response experiment, and explore analyzes the importance of these three services relative to other technological and economic factors of electric vehicles.

Findings: Our results offer evidence that complementary mobility services may significantly foster electric vehicle adoption. To be specific, offering these top three services could increase the purchase probability to 9.42%, which is a strong improvement compared to the former market share predictions of 2.85%. Moreover, low purchase prices are less important than low recurring costs, such as electricity costs. This finding should be considered, and it might be beneficial to offer EVs for a higher purchase price and subsidize

certain recurring costs, such as electricity costs, through the purchase price increase. Finally, a segmentation strategy may be fruitful because, e.g., men are more attracted by technological advantages than women and elderly consumers have a higher preference for services that offer convenience.

Chapter 3 (Article 2): Determining Profit-optimizing Returning Policies: A Two-step Approach on Data from Taobao.com

Motivation: As an online purchase is a two-stage process, selecting an optimal return policy should requires taking into account two effects: respectable increases in sales; secondly, however, a higher return rate could, in turn, lead to substantial costs in terms of reverse logistics, depreciation and additional labor effort. Previous researched use mathematic models to explore the influence' balance between sales and returns (Anderson et al. 2009), but field studies are very limited.

Method and Data Set: Our study has collected 592 e-retailors' transactional data from the most important online platform in China for May 2012, and proposes a two-step model to investigate the influence of various return policies (customer friendly return policies and guarantee credibility) on both sales and returns. Specifically, we first utilize a robust regression to explain purchase behavior, and then apply a zero-inflated negative binominal regression to model the return behavior.

Findings: Referring to sales, we find that the adoption of return policies results in increased sales, while reputation works as a moderator in this process. For returns, good

reputation and traditional customer-friendly return policies (like the Seven-Day Return policy) can significantly increase the number of returns, while more guarantee credibility (enhanced by Guarantee Money) is related to fewer returns. Taken together, both the Seven-Day Return policy (profit increase of +0.29%) and Guarantee Money (profit increase of +0.016% per Yuan guarantee) ultimately increase firms' profits.

Chapter 4 (Article 3): The Impact of the Package-opening Process on Product Returns

Motivation: High product return rates are an increasingly pressing challenge for many e-retailers around the world. To address this problem, this paper offers a new perspective by focusing on the critical moment of the package-opening process. Going beyond previous research, which has primarily focused on website information (Roggeveen et al. 2014; Sahoo et al. 2015)and the product itself (Anderson et al. 2009; Wood 2001), we examine the effects of the outside appearance (i.e., the color of the delivery package) and the content of the delivery package (i.e., extra gifts, coupons, and preprinted return labels) on consumer return behavior.

Method and Data Set: We firstly conducted a well-designed experiment online to further explore the behavioral mechanism at work (Study 1, a representative 320-sample of the target population). To address the remaining concerns that especially questioned the difference between intentions and actual behavior, we present a second experiment (Study 2, including 195 students) in the real world as a robust test. Finally, we used a field study (Study

3, including 108 e-retailors' transactional data) to test whether the assumed influences of package design on return behavior can be proved in real business.

Findings: Our findings across one observational field study and two experimental studies show that a well-considered package design, including colorful packaging and extra gifts, significantly lowers consumers' return intentions and actual returns. Further analyses explore the process of consumers' cognitive–affective reactions after opening a delivery package. During this two-stage reaction process, pleasure plays a crucial role in the consumer's return choice.

Table 1-1 Dissertation Articles

	Article's Title	Research Goal	Study Type and Data Set	Analytical Method	Publication
Chapter 2	Fostering the Adoption of Electric Vehicles by Providing Complementary Mobility Services	Investigating the influence of complementary mobility services on EV adoption.	• Laboratory experiments. • 251 participants in Study 1, 327 participants in Study 2.	• Best-Worst Scaling • Dual Response	Published in: Journal of Business Economics 2015, 85(8), 921-951. VHB Ranking B.
Chapter 3	Determining Profit-optimizing Returning Policies	Comparing various return policies' influences on customer purchase and return behavior.	• Field study. • 592 e-retailers' transactional data from Taobao.com for May 2012.	• Robust regression and Zero-inflated negative binominal regression	Published in: Electronic Markets 2016, 26(2), 103-114. VHB Ranking B.
Chapter 4	The Impact of the Package-opening Process on Product Returns	Investigating the work mechanism of package -opening process on consumer's return decisions.	• Field study and laboratory experiments. • 108 e-retailers' transactional data from Taobao.com in Study 1, 320 participants in Study 2 and 195 participants in Study 3.	• Robust regression model • Analysis of variance • Structural equation model	Forthcoming in: Business Research. VHB Ranking B.

Chapter 2: Fostering the Adoption of Electric Vehicles by Providing Complementary Mobility Services: A Two-step Approach using Best-Worst Scaling and Dual Response

Authors: Oliver Hinz, Goethe-Universität, Gemany

Christian Schlereth, Otto Beisheim School of Management, Germany

Wenyan Zhou, Technische Universität Darmstadt, Germany

Published in: Journal of Business Economics, 85(8), 921-951

Abstract: There is a substantial gap in research regarding the adoption of electric vehicles as a strategy to remedy the climate problem and reduce oil consumption by integrating complementary mobility services. To address this gap, we employ a two-step approach utilizing a hybrid stated preference method. Study 1 uses Best-Worst Scaling and identifies the top three complementary mobility services consumers would prefer with an electric vehicle. Study 2 applies Dual Response and analyzes the importance of these three services relative to other technological and economic factors of electric vehicles. Our results offer evidence that complementary mobility services may significantly foster electric vehicle adoption. Moreover, low purchase prices are less important than low recurring costs, such as electricity costs. Finally, a segmentation strategy may be fruitful because, e.g., men are more attracted by technological advantages than women and elderly consumers have a higher preference for services that offer convenience.

Keywords: Adoption; Electric vehicles; Complementary mobility services**;** Best-Worst Scaling; Dual Response

2.1 Introduction

Fuel consumption is one of the main drivers of environmental pollution (e.g., CO_2 emissions), and a total of 16% of global anthropogenic CO_2 emissions come from road transport vehicles (Olivier et al. 2013). Thus, emissions from automobiles represent a significant challenge for researchers, representatives from the industrial community, and policymakers. With 98% of all vehicles in the world still running on gasoline or diesel, pressure on global oil supply is rapidly increasing, and peak oil demand is predicted to occur no later than 2030 (Aftabuzzaman and Mazloumi 2011). Therefore, from the perspectives of both science and politics, it is crucial to remedy the climate change problem and alleviate the problem of the oil demand gap.

One way to relieve both problems could involve a broader adoption of electric vehicles (EVs, including Plug-in electric cars, hybrid electric cars, and hydrogen vehicles), which use one or more electric engines or traction engines for propulsion. Germany has made EVs part of its long-term oil policy and aims to reduce emissions across all sectors by 40% by 2020 pursuing a massive EV strategy. This goal implies that the country aims to have one million EVs on its roads by 2020. To achieve this ambitious goal, nearly 1.5 billion Euros ($1.9 billion) have been invested in research on EV subsidization and the development of e-mobility in general (Peard 2013). Those investments obviously stimulated the supply of EVs by the e-mobility industry. However, on the demand side EVs face rather low levels of adoption. In 2012, only 4,157 out of a total of three million new vehicles registered in Germany were EVs. In other countries the proportion of newly registered EVs is similarly low.

Although this number is double than that of the previous year, the situation remains unsatisfactory (International Council on Clean Transportation 2013).

Thus, there seem to be a number of barriers to EV adoption. Previous research has noted the importance of price (including purchase price and recharging costs), charging time, and driving ranges in the (non-)adoption of EVs (Beggs and Cardell 1981; Bunch et al. 1993). However, along with the automobile industry's work in this regard, e-technologies have gradually improved (Wesseling et al. 2013). New battery technologies offer longer ranges, more power and a shorter recharging time (Tie and Tan 2013) (e.g., lithium-ion technology by Toyota; the new water-based battery by General Electric). In addition, government subsidization of this technology in the form of subsidies and/or tax relief (Gärling and Thøgersen 2001) may help solve the problem from the cost perspective.

Although complementary mobility services are likely to gain importance, they have mostly been neglected in research, or prior research has treated these services as just another factor, which might increase costs. Specifically, prior research has focused on repair and maintenance costs (Ewing and Sarigöllü 1998), service station costs (Brownstone et al. 2000), and operating costs (Shepherd et al. 2012). Other studies mainly included the availability of electricity charging stations (e.g., Bownstone et al. 2000; Mau et al. 2008; Potoglou and Kanaroglou 2007). Instead of merely focusing on costs or mandatory infrastructure elements, this research proposes that well-tailored complementary mobility services should be exclusively developed to improve the holistic driving experience of EV. Examples include "Intelligent charging stations" or "IT-based parking space and payment". Our study encourages thinking about the availability of exclusive EV services that might significantly

affect consumer behavior. Consequently, our paper aims to explore the influence of complementary mobility services on EV adoption and to formulate recommendations on how industry and policymakers could use our results to more effectively encourage the adoption of EVs.

To achieve this goal, we conduct two types of discrete choice experiments that have a firm foundation in sociology and behavioral research and are well-known for their ability to explain actual purchasing behavior, even if the studied products do not yet exist on the market (Swait and Andrews 2003). Because the number of possible complementary mobility services is quite large, we surmount the limits in the number of attributes of the discrete choice experiment by conducting a two-step approach: first, we utilize Best-Worst Scaling to assess the importance of identified potential complementary mobility services; second, we employ Dual Response, and explore the influence of the most important complementary mobility services relative to other well-studied attributes of EVs. This study shows the advantages of this novel two-step approach.

2.2 Literature Reviews

This section summarizes the attributes prior research has examined with respect to EV preferences and introduces the methodological foundation of our research. Subsequently, we also outline random utility theory, which serves as the core of our theoretical framework in the two empirical studies.

2.2.1 Research on Electric Vehicle Preferences

Since 1981, following the first published study on EV demand (Beggs and Cardell 1981), practitioners and scholars have focused on factors affecting the adoption of EVs by using revealed preference (RP) and stated preference (SP) methods. RP can be used to forecast future demand based on past real-world decisions, whereas SP can be used to incorporate new attributes of a product not yet available on the market. In this paper, we rely on SP. Without claiming completeness, Table 2-1 lists studies in the adoption of EVs using RP or SP in the last few years. Moreover, all attributes examined in previous studies can be assigned to one of four factors: technological, economic, environmental, and complementary mobility services (see Table 2-1).

Better technologies improve purchase intentions by increasing consumers' perceived ease-of-use and perceived usefulness. For EVs, technological superiority is expressed by attributes such as charging time, range per charge, motor power, acceleration, top speed, and multiple-fuel capability. Early in the 1980s, Beggs et al. (1981) documented that limited range and long recharging time were the most significant barriers to the adoption of EVs. Moreover, because the development of technology itself is a dynamic process, the relationship between the adoption of a new technology (such as EVs) and its popularity in the market has also proven to be dynamic (social contagion effect) (Axsen et al. 2009). Customers' choices also can change the development of technology (Mau et al. 2008). Because technological attributes continue to be frequently used in recent studies on EV adoption (Dagsvik et al. 2002; Lieven et al. 2011), we also consider them in our study. However, we do not devote much attention to

Table 2-1: Electric Vehicle Studies using RP and SP after 2000

Study	Data	Method (Econometric model)	Factors contributing to the adoption of EVs				Conclusions
			Technol o-gical	Econo-mical	Environ-mental	Complementary mobility services	
Bownstone et al. (2000)	7,387 households in California	Joint SP/RP data (multinomial and mixed logit)	×	×		× Electricity charging stations	Large heterogeneity in preference for fuel types could be due to respondents' different information sets and fundamental uncertainty.
Ewing and Sarigöllü (2000)	881 respondents in Canada	SP discrete choice experiment (multinomial logit)	×	×	×		Relative vehicle prices and performance levels as well as differential commuting costs and times had modest effects on vehicle choice.
Dagsvik et al. (2002)	622 Norwegian residents	SP discrete choice experiment (random utility models for ranking)	×	×			Alternative fuel vehicles appear to be fully competitive alternatives. Driving range is an important attribute.
Potoglou and Kanaroglou (2007)	902 respondents in Canada	SP discrete choice experiment (nested logit model)	×	×	×	× Electricity charging stations	Reducing monetary costs, purchase tax relief and low emissions rates encourage households to adopt a cleaner vehicle.
Mau et al. (2008)	1935 respondents in Canada	SP discrete choice experiment (multinomial logit)	×	×		× Electricity charging stations	Consumers' preferences in choosing between conventional and new technologies can change with market conditions.
Axsen et al. (2009)	535 Canadians and 408 Californians	Joint SP/RP data (multinomial logit)	×	×	×		Investigates the 'neighbor effect' and uses SP and RP choice research to improve the behavioral realism of an energy–economy model.
Hidrue et al. (2011)	3,029 respondents in the US	SP discrete choice experiment (latent class random utility model)	×	×			Estimates willingness to pay for EV attributes.
Hackbarth and Madlener (2013)	711 people in Germany	SP discrete choice experiment (multinomial logit)	×	×	×	× Electricity charging stations; policy incentives	Conventional vehicles maintain dominance in market; increasing driving range to that of conventional vehicles has same effect as multiple policy incentives.

SP: Stated Preferences, RP: Revealed Preferences; "×" indicates factors investigated in this paper.

the social contagion effect, which has been examined in other domains for new product adoption (Hinz et al. 2014).

Another important factor considered in all studies is the attribute "costs", which consist of EV purchase price plus mileage-dependent operating costs. The mileage-dependent operating costs in the context of EVs imply electricity costs or recharging costs. Previous research has shown that both the purchase price and the mileage-dependent operating costs may significantly affect consumers' decisions about adopting EVs, but no unified conclusion about the optimal pricing strategy has been reached. Other economic factors, such as parking, commuting, and repair and maintenance costs, have also been discussed in later research, but their influence on choosing a vehicle is considered rather weak (Ewing and Sarigöllü 1998). As is well-known, price is always directly connected to the level of performance, and studies have recently begun to estimate the willingness to pay not only for an EV but also for high-tech features of the vehicle (Hidrue et al. 2011). In our study, we focus on the two most important cost factors: purchase price and electricity costs.

Because EVs are supposed to solve emission problems, the environmental component is regarded as an important aspect of the purchase. Scholars consider not only CO_2 emissions (Shepherd et al. 2012) and the reduction of pollution (Potoglou and Kanaroglou 2007), but also consumers' level of environmental consciousness (Ewing and Sarigöllü 1998). Ewing and Sarigöllü (1998) found that over one third of respondents were willing to pay CAN $1000 more for a vehicle if it caused substantially lower emissions.

Today, services have become key not only to customer satisfaction but also to promoting new products in the vehicle market (Fassnacht et al. 2011). Unfortunately, there is little

research on this topic, especially for EVs, a gap this study aims to close. To the best knowledge of the authors, previous research has mainly integrated the availability of electricity charging stations, which is a mandatory infrastructure element for the success of EV adoption. Hackbarth and Madlener (2013) predict that multiple policy incentives, such as permission for EVs to drive in bus lanes or vehicle tax reductions, will raise the shares of EVs markedly. Although policy incentives are important factors, their implementations tend to be exogenous to most car manufactures. Therefore, we propose that research should keep pace with the development of the newest services complementing consumers' driving experience. We will introduce such services in detail later in Section 2.3.

Overall, even though all four main factors can have a significant influence on the adoption of EVs, in this study we focus more on complementary mobility services. We take on the idea of examining technological and economic factors which can help us to better understand additional routes to foster the adoption of electric vehicles.

2.2.2 Theoretical Foundation

Because transactional data on complementary mobility services are not available, or likely do not exist, we turn our attention to stated preference methods. These methods collect data using surveys, which cost less than setting up test-markets and which simplify data analysis because factors of interest can be experimentally manipulated and tested in a controlled research environment. According to Rao (2014), we classify stated preference methods into four classes: (rating- or ranking-based) conjoint analysis, discrete choice experiments, self-explicated methods, and hybrid methods (see

Figure 2-1).

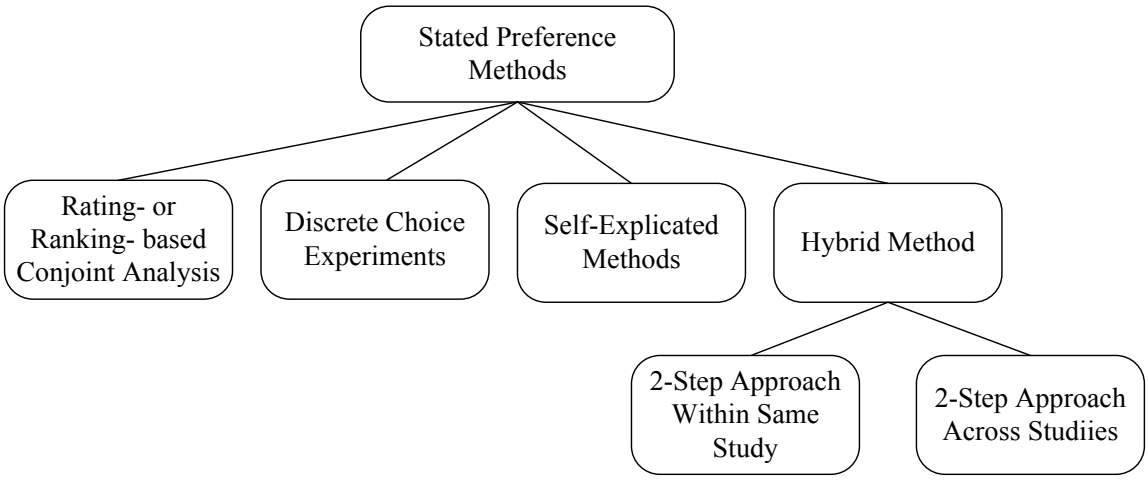


Figure 2-1 Classification of Stated-preference Methods

Rating- or ranking-based conjoint analysis, frequently used in the 1970s and 1980s, allows respondents to rate or rank products, which are described by the particular attributes selected for the study (Green et al. 2001). Those observations are then used to calculate respondents' preferences by estimating the contribution of each attribute and level to the observed outcome. The origins of these methods lay in psychology and are principally associated with research dealing with ways to mathematically represent the behavior of rankings observed as an outcome of systematic factorial manipulation (i.e., known as "factorial designs") of independent factors (also known as "attributes"). As Louviere et al. (2010) outline, these methods rely on formal proofs about the mathematical (algebraic) representations of rank orderings of orthogonal arrays (originally complete factorial arrays). The effort required to perform these methods is rather low; however, the number of selected attributes and levels should not be too high in order to keep the number of products, and thus respondents' cognitive burden, at a feasible level. In addition, deriving managerial

recommendations, e.g., through what-if analysis or counterfactual simulation, is only possible under strong and theoretically unsupported assumptions because the observations do not readily translate into choices. Consequently, observations cannot be analyzed consistently with neoclassical economic theory to simulate respondents' choices, i.e., researchers must apply arbitrary selected probability models (e.g., first choice, probabilistic choice, or logit), which substantially differ in their managerial recommendations, though none of them are theoretically supported.

Instead of relying on ratings or rankings that are artificially translated into preferences, discrete choice experiments allow respondents to repeatedly make choices between a set of alternative products. Given the similarity to real-world purchase decisions, discrete choice experiments are able to explain actual purchasing behavior well, and they have a firm foundation in sociology and behavioral research (Swait and Andrews 2003). More specifically, with random utility theory, these experiments are backed-up by a long-standing, well-tested theory of choice behavior that can take inter-linked behaviors into account (see McFadden 2001). Louviere et al. (2010), in their comparison of discrete choice experiments with conjoint analysis, conclude that random utility theory provides an explanation of the choice behavior of humans, not numbers.[1] However, because each observed choice only provides a marginal amount of information, some researchers recommend conducting discrete choice experiments only in the case of a few attributes (e.g., 3-8), which might become difficult as the complexity of the studied product increases.

[1] See the subsequent section 0.

Self-explicated methods use direct assessments of attributes and their importance or apply various types of adaptive models involving ratings of pairs of alternatives, often on the basis of a partial set of the attributes (see Schlereth et al. 2014). Although most studies could not empirically demonstrate the superiority of discrete choice experiments over conjoint or self-explicated methods (Schlereth et al. 2011), the same argument as against conjoint analysis applies here, namely, that these methods also lack a direct link to respondents' actual choices.

For the purpose of this research, we use a hybrid approach that also addresses the issue of a large number of attributes and levels by limiting their number before presenting them to respondents. Essentially, hybrid approaches involve two steps that can be made within one study (so that a respondent has to go through both steps) or separated into two studies. An appealing aspect of the hybrid approach is that the researcher is free to choose the most suitable stated preference method in each step.

In our study, we rely on two types of discrete choice experiments, given their firm theoretical and behavioral foundations, and given that we are particularly interested in consumer choices of an EV, when adding complementary mobility services. Step 1 is used to determine the most important attributes. Because sufficient knowledge about product attributes of EVs are already available (see Section 2.2.1), we can concentrate our attention on the selection of consumers' most important complementary mobility services. Step 2 is a discrete choice experiment using a limited set of attributes (i.e., the most important complementary mobility services combined with the product attributes) to study consumer

preferences for any of them. By this way, the limitation with respect to the number of attributes and levels of conjoint analysis and discrete choice experiments is mitigated.

Random Utility Theory as the Core of Modeling Consumer Decision Making

In both types of discrete choice experiments, later presented in the two empirical studies, respondents inspect two or more alternatives and are assumed to choose the one which maximizes their utility. The observed choices provide researchers with rich insights into consumer preferences. In marketing research as well as in economics random utility theory is applied on such observations (McFadden 1974; Thurstone 1927). Random utility theory assumes that respondent "*h*" decomposes his or her utility $u_{h,i}$ for alternative "*I*" into a deterministic component $v_{h,i}$ and a stochastic component $\varepsilon_{h,i}$, i.e., $u_{h,i} = v_{h,i} + \varepsilon_{h,i}$. This means that, respondents' utility can be decomposed into components that are directly related to the attributes and levels shown to the respondent and components that are known to the respondent, but they cannot be observed by the analyst. The stochastic component accounts for Thurstone (1927) realization that respondents make errors in their choices, which means that they are not necessarily always choosing the alternative with the highest deterministic utility. McFadden (2001) generalization of Thurstone's RUT model provides tractable, closed-form models that accommodate choices from sets of three or more alternatives (see Train 2009). More formally, the probability of respondent "*h*" choosing alternative "*I*" is:

$$\mathrm{Pr}_{h,i} = \mathrm{Prob}(\mathrm{v}_{h,i} + \varepsilon_{h,i} > \mathrm{v}_{h,j} + \varepsilon_{h,j}; \forall j \neq i), \tag{6}$$

We can rewrite that to:

$$\Pr_{h,i} = \text{Prob}(\varepsilon_{h,j} < v_{h,i} - v_{h,j} + \varepsilon_{h,i}; \forall j \neq i). \tag{7}$$

Assuming a Gumbel distribution for the stochastic component $\varepsilon_{h,j}$, its density is $\text{f}(\varepsilon_{h,j}) = e^{-\varepsilon_{h,j}} \cdot e^{-e^{-\varepsilon_{h,j}}}$ and its cumulative distribution is $F(\varepsilon_{h,j}) = e^{-e^{-\varepsilon_{h,j}}}$. By assuming independence between all error terms $j \neq i$, we can rewrite Equation 7) as the cumulative distribution given the error term $\varepsilon_{h,i}$:

$$\Pr_{h,i} \mid \varepsilon_{h,i} = \prod_{j \neq i} e^{-e^{-(v_{h,i} - v_{h,j} + \varepsilon_{h,i})}}. \tag{8}$$

Of course, observing $\varepsilon_{h,i}$ is not possible, but weighting all possible values of $\varepsilon_{h,i}$ in Equation (8) by their density results in the following integral:

$$\Pr_{h,i} = \int_{-\infty}^{+\infty} \left(\prod_{j \neq i} e^{-e^{-(v_{h,i} - v_{h,j} + \varepsilon_{h,i})}} \right) \cdot e^{-\varepsilon_{h,i}} \cdot e^{-e^{-\varepsilon_{h,i}}} \, d\varepsilon_{h,i}. \tag{9}$$

Finally, we can apply algebraic manipulations and obtain the probability $\Pr_{h,i}$ of respondent h selecting product i among a set of product alternatives J in a mathematically convenient closed form:

$$\Pr_{h,i} = \frac{e^{v_{h,i}}}{\sum_{j \in J} e^{v_{h,j}}}. \tag{10}$$

2.3 Complementary Mobility Service Operationalization

In cooperation with a German consultancy firm, and based on elaborate discussions with EV experts (three workshops with three experts plus four telephone interviews with CEOs of car sharing companies) as well as the analysis of industry reports, we identified nine important complementary mobility services (see Table 2-2). The nine selected services are frequently mentioned in the media, and expectations are high that they might complement EV technology. Some of the services can increase consumers' perceived ease-of-use and usefulness by reducing time-dependent costs (e.g., saving the time required to pay and park)

or by increasing brand loyalty by forming online communities (Algesheimer et al. 2006) (e.g., social network app in car); others offer new driving experiences and make driving more exciting and intelligent (e.g., providing visual real-time updates on traffic information).

All nine complementary mobility services can be classified into two classes. One class of services is specific for electric vehicles and the other class can — in principle — also be integrated in traditional vehicles. The first class includes "Intelligent charging stations" and "Vehicle-to-Grid" (V2G), two services that certainly make only sense for EV. "Intelligent charging stations" is a demand side management instrument that could be used to improve energy efficiency, reduce time of use, allow quick demand response, and enlarge the spinning reserve (Palensky and Dietrich 2011). "Vehicle-to-Grid" (V2G) is an energy system that realizes large synergies between the vehicle fleet and the electricity system. For society, the advantages of developing V2G include an additional revenue stream for cleaner vehicles, increased stability and reliability for the electric grid, lower electric system costs, and (eventually) inexpensive storage and backup capacity for renewable electricity (Kempton and Tomić 2005). For consumers, V2G might serve as another source of income if electricity providers offer real time prices for energy; they can charge their battery, when energy costs are low, and feed in the electricity grid in phases of high demand and high prices.

The other seven services belong to the second class and can improve the utility of both, electric and conventional vehicles. For example "Augmented reality services via head-up displays", which is not restricted to electric engines, use the windshield as a projection surface for displaying virtual content and may help drivers detect and respond to traffic changes more quickly and increase navigational accuracy (Fadden et al. 1998),. Although these seven

complementary services would be or have already been used in conventional vehicles, their availability might contribute more utility for electronic vehicles than for conventional cars.

Table 2-2: Complementary Mobility Services

Complementary Mobility Services	Explanation
Intelligent charging station	Intelligent charging stations simplify charging the EV battery. They enable to automatically identify drivers and to bill energy consumption.
Vehicle-to-Grid (V2G)	To realize substantial synergies between the EV fleet and the electricity system, V2G refers to the return of electricity from the battery of an EV into the electric grid. Drivers can help mitigate peak demand shocks and earn money at the same time: They can charge the battery, when energy costs are low and feed in the electricity grid in phases of high demand and high prices.
IT-based parking and payment	IT-based parking systems directly guide drivers to parking spaces and allow them to pay easily and automatically.
Drive-through for bill payment	Bills may be authorized and paid directly from the EV for certain products or services (e.g., fuel bills, parking fees, or tolls).
Connection to mobility providers	By contracting with mobility providers, drivers may rent and switch batteries offered by mobility providers. Moreover, mobility providers offer intelligent services (such as traffic or travel information) that can also be booked.
Remote diagnostics and updated supply	The software (e.g., operating system) adopted in EVs may be remotely controlled and updated by car repair shops. Meanwhile, remote diagnostics may be offered in the event of errors or defects.
In-car apps, purely vehicle-related function	In-car apps are software applications that equip EVs with additional functions directly related to driving (e.g., driver logs, electricity cost logs).
In-car apps, not purely vehicle-related function	In-car apps that are not directly related to driving, e.g., social media or music apps.
Augmented reality services via head-up displays	Augmented reality services automatically identify and project relevant information on the windshield via a head-up display. The mentioned examples include navigation, information about electricity consumption, prices for nearby recharging stations as well as hotel and restaurant recommendations.

Taking an example, bundling new complementary services like “IT-based parking” with EVs would make the EV market more attractive. Advertisements and word-of-mouth conversations about self-parking could firmly grasp customers’ attentions and create a strong innovation image for these companies and their products. Customers always pay more attention to innovations and might regard such developments as a breakthrough that could

help them move quickly through the first stages (create product awareness) in the adoption process (Armstrong et al. 2010). Although EV development and service innovations are relatively independent, more innovative companies should have more experience and passion on new product development. In turn, that could increase the success of new products (higher sales and longer sale duration in this situation) (Cooper and Kleinschmidt 1987). Facts also supported our inferences: the world's first experimental prototypes of automatic parallel parking was developed at INRIA on a Ligier electric car in the mid-1990s (Paromtchik and Laugier 1998), and BMW announced in January 2014 the following for their new "IT-based parking" service that is first exclusively available for their i-Series (Brigl 2014).

Moreover, complementary mobility services could be aligned with the special needs of EVs. Strongly market-oriented car manufacturers are well-advised to not only sell their EVs at a low price and to advertise the technology itself but to also think about improving the holistic driving experience (Armstrong et al. 2010). For example, the combination of "Augmented reality services via head-up displays" and EV can create a better holistic driving experience. Obviously, the limited driving range is one of the largest barriers for EVs adoption. "Augmented reality services via head-up displays" could provide access to vehicle and environment-related information, such as the current driving conditions and battery charge, route guidance and nearest charging stations and potentially, with which even could educate drivers in their energy efficient driving skills.

2.4 Study 1: Identification of Most Important Complementary Mobility Services

Table 2-3 shows that the number of potentially interesting complementary mobility services is quite large. Since many traditional methods like conjoint analysis cannot deal with too many attributes, we address the problem of studying preferences of complementary mobility services in conjunction with other EV attributes as follows: Study 1 applies Best-Worst Scaling to identify the most important complementary mobility services among those listed in Table 2-3, which will serve later as an input for Study 2. We also test whether simply counting how often a complementary mobility service has been chosen as the best or worst alternative will lead to results similar to those of more sophisticated estimation methods that have recently been proposed in the literature.

2.4.1 Best-Worst Scaling (Case 1)

Best-Worst Scaling, introduced in 1992 by Finn and Louviere, has recently grown in popularity. In case 1, respondents view multiple choice sets that consist of a subset of attributes (in our study, the list of complementary mobility services from Table 2-3) and are repeatedly asked to choose their least and most preferred attributes.[2] Thus, the Best-Worst Scaling forces respondents to trade off attributes of varying attractiveness, which have binary levels (e.g., exists and does not exist or applies and does not apply). Researchers can then determine the preference for an attribute by comparing how frequently respondents have chosen that attribute relative to other attributes in the choice set. Table 2-3 provides an illustrative example of a choice set employed in Study 1.

[2] For a conceptual framework of Best-Worst Scaling, see Louviere et al. (2013). There are alternative methods with supplement “case 2” and “case 3”, in which respondents either choose the most and least preferred level of a product (case 2) or the most and least preferred alternative described by its attributes and levels (case 3).

Table 2-3: Example Choice Set in Study 1

Most preferred	Complementary mobility services	Least preferred
X	IT-based parking space and payment	
	Intelligent charging station	
	Augmented reality services via head-up displays	X

Compared with other preference measurement methods, Best-Worst Scaling offers several unique advantages. For example, compared with ranking methods, which are known to yield low accuracy and consistency when there are more than seven attributes (e.g., Bettman, Johnson, and Payne, 1990), Best-Worst Scaling can be applied to substantially more attributes. Best-Worst Scaling also avoids the assumption of equal differences between two subsequently ranked attributes, which leads to more realistic results. Compared with verbal measurement scales (e.g., rating tasks or Likert scales), respondents do not use artificially numerical, subjectively interpretable representatives for their preferences (e.g., agree-disagree scaling) but instead choose among decisions that are easy to understand and can be made quickly. Results from this method provide a higher degree of discrimination (Lee et al. 2008), and the interpretation of responses is consistent across respondents. Therefore, it is a suitable method for cross-cultural studies and studies with respondents of heterogeneous backgrounds or educational skills.

Another strength of Best-Worst Scaling is that its observations are easy to analyze because simply counting best and worst choices is sufficient to obtain either individual or aggregate sample preference estimates (Finn and Louviere 1992; Mueller Loose and Lockshin

2013). Imagine, for example, nine attributes and twelve choice sets, each consisting of three of these attributes. In the case of a balanced design, each of the attributes appears four times (= 12 x 3 / 9). Consequently, an attribute may generate Best-Worst scores ranging between -4 and +4, depending on how frequently this attribute was chosen as the best attribute (+1), as the worst attribute (-1), or not chosen at all (+0). Adding +5 to all Best-Worst scores will transform them into a range between 1 and 9, which is the response we would observe in a nine-point rating task. The properties of this range depend on the number of repetitions of each attribute. Analysis is simple and can be conducted without any proprietary software.

As an alternative to the count analysis, Marley et al. (2008) propose proper probabilistic choice models, such as the *MaxDiff model*, which have their foundations in random utility theory for estimation. The probability $BW_C(j',j'')$ of choosing the pair of attributes *j*' and *j*'' in choice set *C* as the best and worst attributes is calculated as follows:

$$BW_C(j',j'') = \frac{\exp(v_{j'} - v_{j''})}{\sum_{r,s \in C \setminus \{r \neq s\}} \exp(v_r - v_s)} \quad (j' \neq j'') \tag{11}$$

In this model, $v_{j'}$ is the deterministic utility of attribute *j*' on the aggregate level, i.e., one parameter representing the utility of all respondents. The probability $BW_C(j',j'')$ is calculated by maximizing the differences between any chosen pair of the best and worst attributes *j*' and *j*'' through ($v_{j'}$-$v_{j''}$) relative to the sum of all possible combinations of attributes *r* and *s* in a choice set. The aggregate level utilities are sufficient for the purpose of our study. Nevertheless, estimating the individual preferences is also possible.

We estimate Equation (6) using a Maximum Likelihood estimator, which was programmed in Matlab. The denominator of Equation (6) uses permutations to obtain every combination of alternative *r* and *s* within a choice set.

2.4.2 Set-up

The questionnaire for study 1 consisted of three sections: 1) brief explanations of EVs, which were followed by the presentation of the nine complementary mobility services listed in Table 2-2, each explained using brief textual and pictorial descriptions; 2) the choice sets of Best-Worst Scaling; and 3) demographic and socio-economic questions.

We presented the nine attributes in 12 choice sets consisting of three attributes each. Using a balanced incomplete block design (see Table A2-14 in the Appendix), each attribute appeared four times with a pair frequency of one. We described the complementary mobility services and provided pictures such that every respondent could easily understand the services. Respondents were asked to choose the most and least preferred complementary mobility service in each of the 12 choice sets.

2.4.3 Data

Study 1 was conducted in the first quarter of 2013 with a total of 251 completed questionnaires. We recruited respondents through different channels, such as postings in specialized forums on car-related topics and inviting colleagues, friends of the authors, and students from two major German universities to participate in the study. As an incentive, we

offered entry in a lottery for 3 gift vouchers valued at 20€ each. Therefore, we consider study 1 to be a convenience sample rather than a representative sample. Table 2-4 summarizes the respondents' demographic characteristics.

Table 2-4: Demographic Characteristics in Study 1

Gender	Age	Occupation
Male (68.5%)	18-24 (25.9%)	Unemployed (1.2%)
Female (31.5%)	25-34 (44.6%)	Employee (57.0%)
	35-44 (11.6%)	Workers (2.8%)
	45-54 (13.1%)	Civil Servants (0.8%)
	55+ (4.8%)	Pensioners (0.4%)
		Freelancers (4.0%)
		Students (31.9%)
		Others (2.0%)

N=251.

2.4.4 Results

Table 2-5 reports the average Best-Worst scores, their standard deviations over individuals in parentheses and the results of the maximum likelihood estimation. "IT-based parking space and payment" and "Intelligent charging station" are by far the most desirable complementary mobility services. "Augmented reality services via head-up displays" and "Remote diagnostics and update supply" follow at third and fourth place, respectively. "Drive-through payment" and "In-car apps not for purely vehicle-related functions" are not preferred by our respondents.

A graphical comparison of the Best-Worst scores against the maximum likelihood estimates that we present in Figure 2-2 Best-Worst Scores (y) vs. Maximum Likelihood

Estimates (x) reveals that both estimation results are proportional to one another, which indicates high robustness.

Table 2-5: Results of the Best-Worst Scaling Analysis

Ranking	Additional mobility service	Best –Worst (Differences between Best and Worst Score)	Average Best-Worst Scores	Maximum Likelihood Estimates
1	IT-based parking space and payment	375	1.49 (1.96)	0.83 (0.70)
2	Intelligent charging station	330	1.31 (2.22)	0.74 (1.00)
3	Augmented reality services via head-up displays	190	0.76 (2.38)	0.46 (0.89)
4	Remote diagnostics and update supply	129	0.51 (2.30)	0.25 (0.87)
5	In-car apps for purely vehicle-related functions	74	0.29 (1.99)	0.13 (0.76)
6	Vehicle-to-Grid	-145	-0.58 (2.62)	-0.37 (1.44)
7	Connection to mobility agents	-208	-0.83 (2.45)	-0.49 (1.13)
8	Drive-through payment	-345	-1.37 (2.16)	-0.74 (0.92)
9	In-car apps not for purely vehicle-related functions	-400	-1.59 (2.10)	-0.86 (0.77)

N=251.

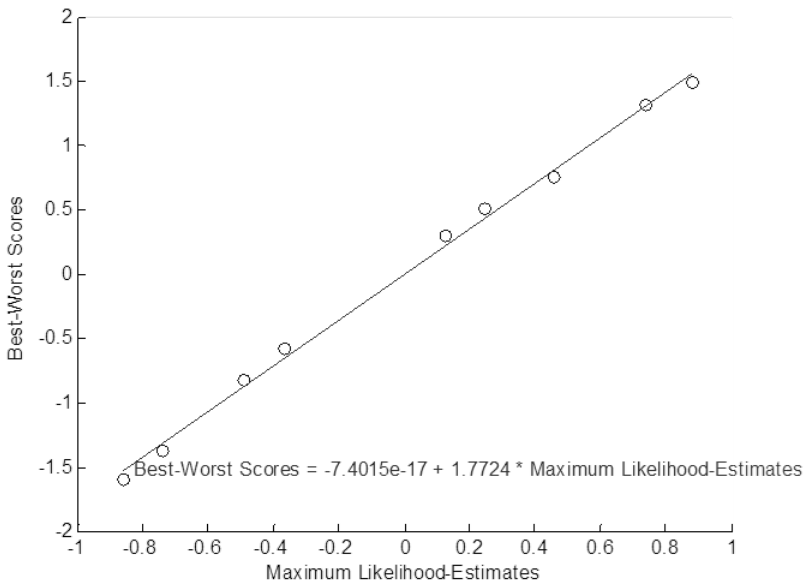


Figure 2-2 Best-Worst Scores (y) vs. Maximum Likelihood Estimates (x)

This proportionality demonstrates that the simple count analysis is sufficient to derive managerial recommendations on the ranking of the complementary mobility services and underlines the strength of this simple trade-off-based stated preference measurement method. However, it remains unclear, how the most preferred complementary mobility services are perceived relative to technological and economic factors of EV adoption. We address this gap in the subsequent study.

2.5 Study 2: Influence of Complementary Mobility Services Relative to Other Electric Vehicles Attributes

In study 2, we explore the influence of complementary mobility services relative to well-studied factors of EV adoption. We employ Dual Response (Brazell et al. 2006; Dhar and Simonson 2003), which enables researchers to also estimate preferences for the levels (e.g., 25,000€ compared to 30,000€) of attributes (here, purchase price). We subsequently describe Dual Response before presenting the study set-up, data collection, results and additional managerial insights obtained from a counterfactual simulation.

2.5.1 Dual Response

Dual Response is a modification of the traditional discrete choice experiment (also frequently referred to as Choice-Based Conjoint). Instead of combining product alternatives and a no-purchase option in a choice set, Dual Response repeatedly asks two types of questions: First, which is the most attractive product alternative in a choice set in a forced choice question without any no-purchase option? Second, in a free-choice question, would they buy

the chosen product alternative? This method is particularly useful if a high proportion of no-purchase decisions is expected because researchers could then additionally observe the trade-off decision among various products, thus yielding more accurate parameter estimates and more stable preferences (Brazell et al. 2006).

For the analysis, we apply random utility theory again and decompose respondent h's utility $u_{h,i}$ for product i into a deterministic component $v_{h,i}$ and a stochastic component $\varepsilon_{h,i}$, i.e., $u_{h,i} = v_{h,i} + \varepsilon_{h,i}$. Assuming a Gumbel distribution for the stochastic component, we express the probability $\Pr_{h,i}$ for making the observed choices in the forced and free choice questions, as follows:

$$\Pr_{h,i} = \frac{\exp(v_{h,i})}{\sum_{i' \in I} \exp(v_{h,i'})} \cdot \frac{\exp(v_{h,i})}{\exp(0) + \exp(v_{h,i})}. \tag{12}$$

The first term describes the probability of choosing product "*i*" as the most preferred product among the set of products "*i*", and the second term models the probability of purchasing it. Each $v_{h,i}$ assumes an additive functional form, i.e., $v_{h,i} = \beta_h \cdot X_i$, where beta is a vector of preferences of respondent "*h*" for all attributes and a constant. X_i is a vector specifying the attributes levels of each attribute in product "*i*".

For the estimation, we employ Hierarchical Bayes, which is a powerful instrument that delivers the distributions of the parameter estimates for respondents at the individual level despite the low number of observations per respondent. The term "hierarchical" refers to the technique of iterating over the lower individual level (i.e., respondents' individual parameters) and the higher population level, assuming that the respondents' parameters are described by a multivariate normal distribution. In each iteration, it draws candidate values from the

posterior distribution (i.e., the estimates) and "borrows" information from the distribution of the population level to make predictions about each respondent's parameters.

We use a normal distribution for all parameters except the price parameter. Here, we assumed a log-normal distribution, which we multiplied afterwards by (-1) to ensure negative values. Thereby, we employed a vector model for the attribute price on the deterministic utility and an effect-coded partworth model for all other attributes. In particular, using a partworth model for electricity cost per 100km instead of a vector model as for purchase price was motivated by the respective cost per 100km, which were close to 0. We employ standard diffuse priors on the parameters with means 0.1 and standard deviations of 5, which imply vague prior knowledge about the parameters. We obtain information about the posterior distribution based on 10,000 iterations that we obtained after discarding a sufficient number of (burn-in) iterations (also 10,000). Convergence was assessed by examining the trace plot of the posterior's likelihood. The estimator is programmed in Matlab and is an extended version of the code provided by Train (2009).[3] For a detailed description of the estimator, Bayesian concepts, and an introduction in prior and posterior distributions, we refer to Section 12 in Train (2009).

2.5.2 Set-up

Study 2 compares the relative importance of complementary mobility services to the previously studied attributes of the technological and economic categories. A literature review and experience reports on EVs were used to identify the most important attributes and

[3] See http://eml.berkeley.edu/~train/software.html.

respective attribute levels (see Table A2-13in the Appendix). The attribute levels we chose (see Table 2-6) cover the largest part of the current electric vehicles market but also incorporate likely technological improvements and slightly declining prices. The technological attributes are "range per charge", "charging time", and "motor power". The economic attributes are "purchase price" and "electricity costs per 100 km". The price is slightly lowered (but nevertheless reasonable according to Table A2-13 in the Appendix) compared with average market prices to reflect expectations about future price developments. We then add the top three complementary mobility services to the list to explore their effects on the purchase-decision process.

Table 2-6: Attributes and Attribute Levels Included in Our Main Study

	Unit	Range	Levels
Range per charge	km	4	100; 175; 250; 325
Charging time	h	2	1; 4
Motor power	kW	2	40; 80
Purchase price	€	4	15,000; 20,000; 25,000; 30,000
Electricity cost per 100 km	€	4	1; 3; 5; 7
IT-based parking space and payment	[ ][4]	2	supported; not supported
Intelligent charging station	[ ]	2	supported; not supported
Augmented reality services via head-up displays	[ ]	2	supported; not supported

Employing the techniques in Street and Burgess (2007), we created a D-optimal (4·2·2·4·4·2·2·2) fractional factorial design with 14 choice sets, i.e., 12 choice sets for the estimation and 2 for the holdouts (see Table A2-14 in the Appendix). These designs are known for their high efficiency and their suitability for a diverse range of research questions. Each choice set shows three different EVs and subsequently asks the respondent in a separate

[4] The complementary mobility services are recorded as dummy variables, so there is no unit.

question whether he or she would buy the most preferred EV (see example as Table A2-15 in the Appendix).

The questionnaire of study 2 consisted of two sections. First, we collected demographic information on gender and age and presented basic information about EVs. At the end of this section, the respondents indicated their interest on a four-point rating scale. Only the respondents who answered "Yes, I can imagine purchasing an electric vehicle" continued with the remainder of the survey; for the other respondents, we assumed that their willingness to pay was below the minimum price level in the discrete choice experiment. The second part continued with the discrete choice experiment.

Table 2-7: Sample's Interest in Purchasing an EV

Question	Can you imagine purchasing an electric vehicle?	Number of participants
Four-point rating scales	Yes, I can imagine purchasing an electric vehicle.	168 (51.4%)
	No, but I can imagine leasing an electric vehicle.	24 (7.3%)
	No, but I can imagine using an electric vehicle as part of a car-sharing services.	57 (17.4%)
	No, I cannot imagine using an electric vehicle at all.	78 (23.9%)

2.5.3 Data

We hired a market research firm that collected a representative sample of the German population with respect to gender and age in April 2013. We obtained 327 completed questionnaires. A total of 168 of the respondents (51.4%, see Table 2-7) reported having sufficient interest in purchasing an EV in the future and entered the analysis, which is quite high compared with the 16% completion rate in California in 1998 and the 30% completion rate in a large Swedish city in 2001 (Gärling and Thøgersen 2001). We excluded the fastest 10% of the respondents (i.e., 18 respondents) to ensure that only respondents who did not click

through the survey entered the analysis. Therefore, a total of 150 completed questionnaires were considered for further evaluations. Table 2-8 summarizes the respondents' demographic characteristics.

Table 2-8: Demographic Characteristics in Study 2

Gender	Age	Occupation
		Unemployed (1.3%)
	18-24 (10%)	Employee (45.3%)
Male (56%)	25-34 (18.7%)	Workers (7.3%)
Female (44%)	35-44 (15.3%)	Civil Servants (2.0%)
	45-54 (23.3%)	Pensioners (20.0%)
	55-64 (22.7%)	Freelancers (8.7%)
	65-74 (8%)	Students (8.0%)
	75-84 (2%)	Pupils (0.7%)
		Others (5.3%)
		Unspecified (1.3%)

N=150.

2.5.4 Results

Based on the Dual Response choices, we estimate the parameter values and derive importance weights of the attributes (see Table 2-9). Signs and magnitudes of the parameter values are reasonable and provide face validity. Internal validity is high, with a first choice-hit rate of 92.3% in the within-sample choice sets. Predictive validity is also high, with 75.8% of the hold-out choice sets being correctly predicted.

The importance weights averaged across all respondents and listed in Table 2-9 demonstrate that electricity cost (for 100 km) has the highest importance (25.03%) which means that the recurring cost is the most important attribute when customer consider buying an EV. The aggregated parameter values of electricity cost gradually decrease, and the maximum decline of 1.58 occurs when electricity costs per 100 km rises from 5€ to 7€. Range

per charge closely follows as second most important attribute (21.85%). We observe a substantial increase in utility when the range per charge increases from 100 km to 175 km. In that case, aggregated parameter values increase by 1.92. There is no substantial increase (only 0.43) in utility when the range increases from 250 km to 325 km. Surprisingly, purchase price only ranks fifth (10.37%); however, it has one of the highest standard deviations (i.e., 11.56%), indicating heterogeneous preferences.

These numbers differ from results for traditional vehicles markets where purchase price almost means everything. This is mainly due to two reasons: First, drivers preferring electric vehicles are likely to be environment-friendly people or technology enthusiasts, who might be more concerned about emission rates or technology innovation than about purchase price (Potoglou and Kanaroglou 2007). Second, the high fuel price is probably one of the main reasons why people consider purchasing an EV (Hidrue et al. 2011). With this in mind, they definitely pay more attention to electricity costs than to purchase price.

Importance weights of all three complementary mobility services together ("IT-based parking space and payment", 10.38%; "Intelligent charging station", 10.42%; "Augmented reality services via head-up displays", 6.67%) exceed the importance of electricity cost and reach 28.01%. In line with Study 1, we observe that "IT-based parking space and payment" and "Intelligent charging station" share approximately the same importance and that there is a substantially larger gap with respect to the third complementary mobility service, "Augmented reality services via head-up displays". We conclude that our results are robust over different methods.

Table 2-9: Parameter Estimates of Study 2

Attributes	Levels	Aggregated parameter values (Standard Deviation)	Average Importance weights (Standard Deviation)
Constant		-0.88 (5.96)	
Range per charge100 km	100 km	-1.87 (1.21)	21.85% (9.73%)
	175 km	0.05 (0.64)	
	250 km	0.70 (0.69)	
	325 km	1.13 (0.85)	
Charging time	1 hour	0.29 (0.49)	5.65% (4.68%)
	4 hours	-0.29 (0.49)	
Motor power	40 kW	-0.48 (0.83)	9.63% (9.33%)
	80 kW	0.48 (0.83)	
Purchase price	(per 1,000€)	-0.13 (0.24)	10.37% (11.56%)
Electricity cost per 100 KM	1€	1.54 (1.23)	25.03% (11.59%)
	3€	0.86 (0.70)	
	5€	-0.41 (0.75)	
	7€	-1.99 (1.51)	
IT-based parking space and payment	supported	0.80 (0.66)	10.38% (7.38%)
	not supported	-0.80 (0.66)	
Intelligent charging station	supported	0.78 (0.61)	10.42% (7.93%)
	not supported	-0.78 (0.61)	
Augmented reality services via head-up displays	supported	0.38 (0.57)	6.67% (5.78%)
	not supported	-0.38 (0.57)	

N=150

Furthermore, we conducted an analysis of variance (One-way ANOVA) that indicated that gender has an influence on the importance weight of “Augmented reality via Head-up displays” (see Table 2-10). We can reject the hypothesis that there are no differences among 6 age groups on the importance weight of “IT-based parking space and payment” and “Intelligent charging station” (see Table 2-11). Men are more attracted than women to technological innovations (such as “Augmented reality via head-up displays”), and older people prefer complementary mobility services that offer additional convenience more than

younger people. Other socio-demographic characteristics have no significant effect on the level of importance weights.

Table 2-10: Influence of Gender on Importance Weights of Properties

	Female N=66	Male N=84	Significant differences between groups
Augmented reality via head-up displays (Average weights)	5.65%	7.47%	$p < 0.1$

N=150

Table 2-11: Influence of Age on Importance Weights of Properties

	18-24 year N=15	25-34 year N=28	35-44 year N=23	45-54 year N=35	55-64 year N=34	65-84 year N=15	Significant differences between groups
IT-based parking space and payment (Average weights)	8.64%	9.14%	10.37%	10.19%	9.00%	18.03%	$p < 0.001$
Intelligent charging station (Average weights)	10.94%	10.58%	7.36%	8.77%	12.04%	14.48%	$p < 0.1$

N=150

2.6 Counterfactual Simulation

Using a counterfactual simulation, we highlight the managerial insights emerging from the results of our two studies. Managers usually want to acquire better knowledge about how prices or technical capabilities of EVs affect adoption rates and other relevant variables to make informed decisions.

In a stylized scenario, we consider current market offers of EVs, which we label as follows: status quo, i.e., purchase price of 40.000€; charging time of 4 h; range per charge of 175 km; and electricity costs of 5€ per 100 km (see Appendix). For ease of illustration, we do not differentiate between different brands of EVs, we consider static prices that do not change

over time, and we neglect the associated costs within each improvement. We predict market share s_i of buying an EV i using Equation 13). For each respondent h, we calculate their deterministic utility $v_{h,i}$ of purchasing a specific EV i. We average this probability (where $|H|$ indicates the number of respondents) and multiply the value with 51.4%, i.e., the share of respondents $s_{\text{sufficientinterest}}$ who had sufficient interest, to account for the fact that not everyone entered the discrete choice experiment.

$$s_i = s_{\text{sufficient interest}} \cdot \frac{1}{|H|} \cdot \sum_h \frac{\exp(v_{h,i})}{\exp(0) + \exp(v_{h,i})}. \tag{13}$$

The result helps us to assess the change in probability after improving each of the attributes one by one. We also test the change in the probability when introducing the complementary mobility services. These predictions might help EV manufacturers prioritize technical innovations and analyze potential increases in adoption rate relative to the costs of these innovations. The results are summarized in Table 2-12.

We obtain market share predictions of 2.85% for EVs. Decreasing purchases prices by 25% only increases purchase probability by 1.16%. This is substantially lower than the 2.00% probability for EV manufacturers who are able to develop new technologies that enable consumers to decrease recurring cost (the electricity costs per 100 km from 5€ to 3€). However, this increase in purchase probability is rather low compared to the impact of adding our three complementary mobility services. The latter adds 6.57% to purchase probability. Support for “IT-based parking space and payment” increases the share by 2.30%, and both “Intelligent charging station” and “Augmented reality services via head-up displays” are able to foster EV adoption comparable to a 10,000€ price cut. Therefore, complementary mobility

services might be a good lever for fostering the adoption of EV. Still, only a mix of improvements in the attributes can yield the anticipated success.

Table 2-12: Results of the Counterfactual Simulation

	Status Quo	40.000€ → 30.000€	5€ per 100 km → 3€ per 100 km	175 km range → 250 km range	
Purchase probability	2.85%	4.01%	4.85%	3.50%	
Change in probability compared to status quo		+1.16%	+2.00%	+0.65%	
	Status Quo	IT-based parking space and payment	Intelligent charging station	Augmented reality services via head-up displays	With all three mobility services
Purchase probability	2.85%	5.15%	4.33%	3.93%	9.42%
Change in probability compared to status quo		+2.30%	+1.48%	+1.08%	+6.57%

At first sight, this result might be surprising and seem counterintuitive, given that complementary mobility services will improve the share of purchases more strongly than substantial price cuts. Nevertheless, we believe the findings are justified for the following reason: We acknowledge that prices of approximately 30.000€ are rather high and exceed Germany's 2012 average prices for a new car (i.e., 26,780€) and for a second-hand car (i.e., 12,730€).[5] Therefore, even with a 10,000€ discount, EVs would still be considered as premium products. The segment of premium product buyers is known to be less price sensitive. Instead, potential buyers will be more interested in improvements in the holistic driving experience. Thus, adding unique capabilities, such as the proposed complementary mobility services, might be a better lever of EV adoption than just tweaking the specifications or prices.[6]

[5] Source: de.statista.com.

[6] As a robustness test, we use another status quo with a purchase price of 30,000€ and electricity costs of 3€ per 100 km.

2.7 Conclusions

Building upon a two-step approach using Best-Worst Scaling and Dual Response, we not only estimate the importance of different complementary mobility services but also enhance the general state of science in the fields of EV adoption and survey methodologies. First, our findings show that the adoption rate for EVs is expected to be higher than in previous years. Specifically, 51.4% of the respondents could actually imagine buying an EV. Second, "IT-based parking and payment", "Intelligent charging stations", and "Augmented reality services via head-up displays" are the top three services preferred by consumers. These services can indeed foster the adoption of EVs along with other technological and economic factors. Third, instead of the purchase price, recurring cost (electricity cost) is the most important attribute individuals consider when thinking about adopting an EV. These results are very encouraging and useful for the electric automobile industry. Fourth, hybrid stated preference methods have proved to be an effective and efficient survey methodology in the adoption field.

The influences of specific complementary mobility services on the adoption of EVs are examined together with technological and economic factors, which make the EV demand literature more comprehensive and abundant. The top three complementary mobility services we selected from Best-Worst Scaling (study 1) have high importance weights ("IT-based parking and payment": 10.38%, "Intelligent charging stations": 10.42%; and Augmented reality services via head-up displays: 6.67 %), as the Dual Response (Study 2) shows. Offering these top three services could increase the purchase probability to 9.42%, which is a

The general insights do not change, as reported in the Appendix.

strong improvement compared to the former 2.85%. Thus, our results confirm that these services may significantly affect the adoption rate of EVs and should thus be carefully considered by policymakers and the automobile industry.

As previous studies have concluded, electricity costs and range per charge are two of the most important factors that foster or hinder EV adoption. According to the results of our study, the range per charge should be 175 km or more, and electricity costs should be reduced as much as possible. However, unlike previous studies, we found that the purchase price plays a minor role for the respondents in our German sample. Compared to importance weights of electricity costs (25.03%) and the range per charge (21.85%), the importance weight of purchase price, at 10.37%, appears to be relatively low (see Table 2-9 for full information). This finding should be considered, and it might be beneficial to offer EVs for a higher purchase price and subsidize certain recurring costs, such as electricity costs, through the purchase price increase. It might also make sense to use the best and most expensive technologies to reduce electricity costs. Prospective buyers pay more attention to recurring costs than to purchase price. Moreover, price elasticity is rather low compared to offering complementary mobility services. Decreasing purchase price from 40,000 € to 30,000 only increases purchase probability by 1.16% which is even less than the increase from adopting the “IT-based parking space and payment” and “Intelligent charging” services. This result emphasizes that consumers consider the holistic driving experience as more important than a lower purchase price. Thus, instead of a price-cutting strategy, offering complementary mobility services seems to be a promising strategy.

Moreover, the analysis of the data obtained in this study reveals that for automakers and other (potential) market participants, potential consumers exhibit a wide range of demand due to a heterogeneous preference structure. For example, men are significantly more interested in EVs and electric mobility in general and show more passion in augmented reality services via head-up displays compared to women, whereas older consumers are attracted by high-quality parking and charging services. Therefore, a segmentation strategy may be fruitful in the EV market.

From a methodological point of view, our paper's contribution is to summarize the classification of stated-preference method and then to propose and demonstrate a two-step approach (hybrid methods) using the Best-Worst Scaling method prior to Dual Response. The first step serves to identify a subset of the most preferred complementary mobility services. Best-Worst scaling is particularly helpful when preferences need to be captured by a relatively small sample. Making the required decisions simply requires selecting the best and worst attributes and thus is fairly simple for respondents. The interpretation of responses is consistent across respondents, even in the presence of heterogeneity with respect to knowledge and cultural background. Also, the data analysis is simple and does not require complex statistical knowledge. As we demonstrate with our study, the simple count analysis using Best-Worst scores provides results consistent with the more sophisticated random utility theory aligned MaxDiff model. In a second step, we integrated the most preferred complementary mobility services in Dual Response to study their impact on purchase decisions relative to other attributes frequently considered in prior research. By doing that, we reduced the number of choice sets and alternatives to an acceptable level. Notably,

conclusions we obtained from the Best-Worst Scaling and Dual Response are consistent in rank and in distances, although it is unlikely that the same respondent participated in both studies. This again supports the feasibility of this two-step approach. We thus expect that this two-step approach will also be useful in other research domains.

Future studies might deep-dive beyond our reasoning in Section 2.5 and investigate further into the causes of why the impact of purchase price was relatively low compared to the complementary mobility services. One way might be to test for interaction effects between complementary mobility services and purchase price. Such tests require different design generation processes that explicitly incorporate potential interaction effects, though. In addition, we could not test, whether the number of levels or their ranges affected the high relative importance of complementary mobility services. For example, we used 4 levels for price and only 2 levels for each complementary mobility service (i.e., available and not available). However, here for example, De Wilde et al. (2008) would predict that the higher number of levels for price would rather increase its importance compared to the 2 levels attributes. Moreover, along with the development of this era, the importance of various attracts would be changed. Especially for complementary services, new services will appear and the old ones will whether be popular or be eliminated. As a result, further research should keep pace with the times.

Appendix

Table A2-13: Marketing Research on Electronic Vehicles

Brand	Model	Purchasing price (€)	Range per charge (km)	Top speed (km/h)	Electric cost[7] (€/100 km)	Charging time (h)	Motor power (kW)
Tazzarri	Zero	24,499.00	140	100	2.20	9	15
Renault	Kangoo Z.E.	23,800.00	160	130	3.88	3.96	44
Renault	Zoe	20,600.00	210	135	3.65	2.60	65
Volvo	C30 Electric		145	130	4.15	3.90	80
Mitsubishi	i-Miev	34,990.00	160	130	3.13	7	47
Citroen	C-Zero	34,164.50	150	130	3.15	6	49
Opel	Ampera	42,900.00	55	161	3.38	4	111
Chevrolet	Volt	41,950.00	55	161	3.38	6.50	111
Peugeot	iOn	29,393.00	150	130	3.15	3.51	49
Ford	Focus Electic	40,000.00	160	136	3.85	4.13	107
Citroen	Berlingo First Electic	59,694.00	95	95		5.46	28
Buddy	Pure Mobility AS	26,989.00	120	80	2.20	8	13
Nissan	Leaf	35,000.00	160	145	3.75	3.90	80
Tesla	Roadster	118,000.00	395	201		3.30	215
BMW	Mini E	34,950.00	175	152	3.50	3.5	150
Daimler	Smart Electric Drive	23,000.00	115	112	3.00	4.29	30
Ford	Transit Connect Electric	53,544.00	130	120	5.43	5.46	105
S.A.M. Group	Sam EV II	16,600.00	100	90	2.00	5	11.6
DFM Mini Auto	Van EQ 6380	15,988.00	120	85	5.00	4.03	

Collected from: 1. http://www.adac.de/infotestrat/autodatenbank/suchergebnis.aspx

2. http://electric-car-database.com/de/?h=n

[7] Electric cost is calculated by electric consumption (kWh/100km) times the electricity price in Germany (0.25€ per kWh).

Table A2-14: Design of Study 1 (Each Row Represents a Choice Set and Each Cell the Attribute Index of Each Alternative)

Alternative 1	Alternative 2	Alternative 3
2	4	8
3	5	9
4	7	9
1	2	3
2	5	7
3	4	6
2	6	9
1	8	9
1	4	5

Table A2-15: Design of Study 2 (Each Row Represents a Choice Set and Each Cell the Level-Index of Each Attribute per Alternative)

Alternative 1								Alternative 2								Alternative 3							
2	1	0	2	0	1	0	1	0	2	1	1	1	0	1	0	1	0	0	0	1	0	0	1
1	0	1	2	0	1	1	0	2	2	0	1	1	0	0	1	3	3	1	3	0	1	1	0
2	0	1	1	0	0	1	1	3	1	1	3	1	1	1	1	0	3	0	0	0	0	0	0
1	2	0	3	1	1	0	0	3	3	0	2	1	1	0	0	0	1	1	0	0	0	1	1
1	3	1	1	0	1	0	1	3	2	0	0	0	1	1	1	2	1	0	2	1	0	1	0
1	3	0	1	0	1	1	1	3	1	1	0	1	0	0	0	0	0	1	3	1	0	0	0
2	0	0	3	0	0	1	0	3	1	0	1	0	0	0	0	0	2	1	2	1	1	0	1
*1	2	0	2	0	0	0	0	0	0	0	3	1	1	1	1	2	3	1	0	1	1	1	1
2	2	1	0	0	1	0	0	1	1	1	3	0	0	0	1	3	3	0	2	1	0	1	1
3	2	0	3	0	0	0	1	1	1	1	0	1	1	1	0	0	0	1	1	1	1	1	0
1	2	0	0	1	0	1	1	2	3	1	2	1	0	0	1	0	0	0	1	0	1	1	0
3	0	1	2	0	1	0	0	0	3	0	3	1	0	1	1	2	1	1	1	0	1	0	0
*0	1	0	0	0	1	0	1	3	0	0	1	1	1	0	1	1	2	1	2	0	0	1	0
2	2	0	2	0	1	1	1	1	0	1	3	1	1	0	1	0	3	0	0	0	0	1	0

* Choice sets that were used for holdout predictions.

Chapter 3: Determining Profit-optimizing Return Policies – A Two-Step Approach on Data from Taobao.com

Authors: Wenyan Zhou, Technische Universität Darmstadt, Germany

Oliver Hinz, Goethe-Universität, Gemany

Published in: Electronic Markets, 26(2), 103-114

Abstract: Selecting an optimal return policy requires taking into account two effects: the potential positive effect on sales and the potential negative effect of higher costs. We propose a two-step model, in which we first utilize a robust regression to explain purchase behavior, and then apply a zero-inflated negative binominal regression to model the return behavior. We apply this model to data from the most important online platform in China and obtain three main findings. First, the adoption of return policies results in increased sales, while reputation works as a moderator in this process. Second, good reputation and traditional customer friendly return policies (like the Seven-Day Return policy) can significantly increase the number of returns, while more guarantee credibility (enhanced by Guarantee Money) is related to fewer returns. Taken together, both the Seven-Day Return policy (profit increase of +0.29 %) and Guarantee Money (profit increase of +0.016 % per Yuan guarantee) ultimately increase firms' profit.

Key words: Return policy; Return behavior; Zero-inflated negative binominal regression; Seven-Day Return policy; Guarantee credibility

3.1 Introduction

High return rates are a global problem for online retailers, threatening their business model in the long run. According to the Wall Street Journal, a third of all Internet transactions are returned by shoppers and the return rate is still increasing (Banjo, 2013). High return rates cause substantial costs (reverse logistics cost, product depreciation, management of return process and so on) (Blanchard, 2005). ASOS Chief Executive Nick Robertson stated that a 1 percent decrease in return rates would immediately increase profits by 10 million pounds ($16 million, approximately 30% of their net income in 2012) (Thomason, 2013). In the U.S., product returns cost manufacturers and retailers approximately $100 billion annually due to lost sales and reverse logistics, reducing profits by 3.8% on average per retailer or manufacturer (Blanchard, 2007).

We analyse the effect of return policies for small and medium-sized online shops. The situation is especially challenging for them because they typically have only limited liquidity and relatively high labour costs which make them more fragile and high return rates constitute a significant challenge for them. High uncertainty in online environments (compared to traditional brick-and-mortar shops) and intense competition force them to offer attractive return policies to attract customers. For these reasons, it is crucial for small and medium-sized online shops to evaluate whether their return policies are beneficial and which policies are better with respect to profits. As our analyses rests on data provided by Taobao.com, the results can directly be applied by retailers that are active on Taobao.com.

Moreover the results may also hold for small and medium-sized online shops that face a fierce competition.

The online purchase process in general can be modeled as two separate decisions (Bechwati & Siegal, 2005; Wood, 2001): customers' decision to order and, upon receipt, their decision to keep or return. The effects of return policy can persist from the pre-purchase to the post-purchase phase (Wood, 2001). In the first phase, customers cannot experience and assess the actual quality of the ordered products, which constitutes an information asymmetry. Return policies aim at countering this asymmetry and have given online retailing a huge boost (Banjo, 2013). Tolerant return policies act as a signal that induces customers to perceive higher quality and lower risk (Glover & Benbasat, 2010) in a product. As a result, customers could spend less time on considering whether to buy or not, which ultimately eases their purchase decision.

In the second phase (post-purchase), return policies can increase customers' satisfaction and maintain long-lasting relationships (Pizzutti & Fernandes, 2010), because they can easily get their money back if they are unsatisfied. Return behavior can be attributed to two causes: one is the gap between expectations and actual product quality (fit problem), and the other is opportunistic or planned behavior. The first one can be improved by offering accurate descriptions as well as services like user-generated product evaluations. But the second cause depends more on customer personality and the return policies themselves. Prior research has shown that relatively restricted return policies – for example, charging "hassle" fees (Davis, Hagerty, & CGerstner, 1998) and restocking fees (Shulman, Coughlan, & Savaskan, 2009),

or granting only conditional return guarantees (i.e., only solving verifiable problems) (Chu, Gerstner, & Hess, 1998) – can effectively reduce the return rate.

Overall, tolerant return policies can generate respectable increases in sales, but a higher return rate could, in turn, lead to substantial costs in terms of reverse logistics, depreciation and additional labor effort. Therefore, both scholars and practitioners have tried to optimize return policies and find the balance between increased sales and higher return rates using models or experiments. However, empirical studies are limited. The only related study at shop level (using shops' operating data) is Davis et al. (1998), who investigated 133 retailers with various return policies and found that retailer return policies vary with how quickly a product is consumed, the salvage value (i.e., the retailer's ability to resell the product or obtain credit from its suppliers) of returned merchandise, and whether there are opportunities to cross-sell or substitute other items when returns occur. Other related studies explore the balance using models. For example, Anderson, Hansen, and Simester (2009) established a model to measure the purchase and return of apparel items from the customer perspective and estimated the model with 987 customer records from a mail-order catalog company.

In this paper, we focus on the research question: How do various return policies (customer friendly return policies and guarantee credibility) affect customer purchase and return behavior for small and medium-sized online shops? For this purpose we suggest a two-step approach: First we determine the impact of the return policy on sales, and second we examine the influence of the return policy on return behavior. We collected data for 600 shops on the most important online marketplace in China, Taobao.com. The combined gross

merchandise volume of Taobao Marketplace and the affiliated Tmall.com exceeds 1 trillion yuan (~132 billion EUR) in 2013(Alibaba, 2013). The data offer the opportunity to analyze the effect of return policies as the sellers can choose from various return policies on Taobao.

Due to the characteristics of our data, we apply a robust regression model to deal with minor concerns about the potential failure to meet assumptions, such as normality, heteroscedasticity, or observations that exhibit large residuals, leverage, or influence. Further, we estimate a zero-inflated negative binomial regression (ZINB) to deal with excessive zeroes in the return rate. We also control for guarantee credibility (Suwelack, Hogreve, & Hoyer, 2011) and reputation (Roggeveen, Goodstein, & Grewal, 2014) and their potential interaction effects.

3.2 Literature Review on Return Policies

A service guarantee is a promise by a company to compensate the customer in some way if the defined level of service delivered is not fully met (Sum, Lee, Hays, & Hill, 2002); examples include money-back guarantees and lowest-price guarantees. Prior research (e.g., Su and Zhang (2009)) has suggested that return policies can be analyzed in three different dimensions: 1) return deadlines; 2) consumer effort required (in terms of bringing back original receipts and filled-in return forms); and 3) extent of return coverage (extent of money back due to shipping charges, inventory holding charges and re-stocking fees).

During the purchase decision process, a tolerant return policy – which entails longer deadlines, less required effort for returning and more coverage – provides an effective signal that reduces uncertainty (Heiman, McWilliams, & Zilberman, 2001) and heightens perceived

quality (Boulding & Kirmani, 1993; Moorthy & Srinivasan, 1995). Signaling theory (Kirmani & Rao, 2000) proposes an explanation for the conditions under which a guarantee reliably signals quality and influences consumer choice (Boulding & Kirmani, 1993; Erevelles, Roy, & Yip, 2001). That is, when consumers are confronted with information asymmetry about true product quality, they try to assess the magnitude of the penalty faced by the seller when the product's actual quality is lower than its promised quality (Boulding & Kirmani, 1993). This penalty results from both direct return costs (amount of return, expenses) and indirect reputational implications. Further penalties involve the return rate, which is influenced by the perceived quality of the market mechanism used to detect misleading applications of the signal (e.g., ease of understanding, interest in comparison) (Heiman et al., 2001). Meanwhile, Desmet (2013) found that this relationship is moderated by brand, price, and the relationship between customer and retailer.

While previous research on return policy's marketing signals has mainly focused on quality signals, Suwelack et al. (2011) revealed that the credibility of a return policy (guarantee credibility) is a key mediator between the return policy and customer purchasing behavior. Specifically, the more credibility customers feel the more purchase intention will they have. Guarantee credibility also can work as a trust signal (Mavlanova & Benbunan-Fich, 2010) which can influence customer return behavior.

In the return decision process, a tolerant return policy however also increases product return rates for customers in remote purchase environments (Wood, 2001). Anderson et al. (2009) used utility theory to point out that customers return items only if the net utility (Net

utility = Deterministic utility + Fit of transaction + Return cost) is negative. Specifically, according to the two-step process, the utility of the product (both pre-purchase utility and post-purchase utility) for a specific customer is derived from three parts: (1) Deterministic utility, which is known to the firm and the customer at the time of purchase; (2) fit of transaction, which is unknown to the firm but known to the customer after the purchase (Petersen & Kumar, 2009); and (3) return cost, which includes the shipping fee (Frischmann, Hinz, & Skiera, 2012), time and labor cost. More lenient policies impose less return costs for the customer, thereby increasing return intention. But other scholars have treated the service guarantee (return policy) as a positive factor for return behavior in the long run, namely in offering a continuous positive effect on employees' motivation and ability to learn from service failure, which can increase service quality and indirectly reduce customer intention to return indirectly (Dutta, Biswas, & Grewal, 2007).

While there exists a vast array of literature on the optimal way to design return policies (e.g., Padmanabhan and Png (1997)), only a few papers (e.g., Che (1996)) have examined the impact of return policies on both the pre-purchase and the post-purchase steps (Janakiraman & Ordóñez, 2012), and the empirical research on the second step is especially lacking. Table 3-1 shows the related empirical studies. We can conclude that 1) most empirical studies work on the customer level (using customers' purchase history or performing behavior experiments), examining the behavior of a number of customers from one company or platform; 2) most of the studies apply simple OLS regressions when analyzing data; and 3) there is no study that focuses on the effects of guarantee certainty on return behavior. Our

paper aims to close these gaps by analyzing real data using a zero-inflated negative binominal regression, for the purpose of measuring the influence of return policies on both, sales and returns.

3.3 Hypotheses and Conceptual Model

To investigate the return policy's effect on the whole purchase process, we propose a two-stage conceptual model. We analyze the effect of return policies and guarantee credibility on both, the purchase and the return stage (See

Figure 3-1). We control for reputation as another important quality signal for online shopping (Zhang, Luo, & Li, 2012).

Customer friendly return policies work as quality signal and risk control measure when customers make their purchase decision. Because of high return costs, the adopted customer friendly return policy is a costly indicator for the seller's confidence in their own product quality. Prospective buyers assume that sellers do not adopt policies that lead to losses. So the products' quality is expected to be higher or closer to their description on the websites, which can lead to a higher fit of expectations and actual product quality. Moreover, customer friendly return policy can reduce the cost

Table 3-1 Summary of Prior Empirical Research Focusing on Customers' Product Return Behavior

Studies	*Method*	*Data*	*Main contributions*
Hess and Mayhew (1997)	Hazard model and empirical evidence	1000 customers during 4 years	Using hazard model to predict return behavior.
Davis et al. (1998)	OLS regression	133 stores in the Sacramento, California area	Retailers prefer a low-hassle return policy when 1) its products benefit are only realized in long term; 2) there are opportunities for cross selling; and 3) the salvage value from returned merchandise are high.
Wood (2001)	Conjoint experiments	68 undergraduate students in the main study	Lenient return policy increases purchase rates and product return rates via signal effect.
Bechwati and Siegal (2005)	Labor experiments	87 undergraduate students in Study 1; 117 undergraduate students in Study 2	Introduces a framework of the mechanisms underlying product returns. Customer return choices are different when they facing disconfirming information.
Anderson et al. (2009)	Developing econometric model and empirical evidence	987 customers of a mail-order catalog company	Provide empirical evidence that return policy gives customers an option value that is measurable; add the option value to model how different return policies affect firm profits.
Petersen and Kumar (2009)	Seemingly unrelated regression, Tobit Model and empirical evidence	Transactions information of 1572 customers between January 1998 and August 2004 in a B2C company	Empirically demonstrate the role of product returns in the exchange process and show how product return behavior affects future customer and firm behavior.
Bonifield, Cole, and Schultz (2010)	Regression; Field experiments	141 of the e-retailers listed on BizRate.com; 290 consumers at an e-tail site	Show a correlation between perceived quality of online retailers and product return policy leniency. The shopping experience moderates the relationship between leniency and consumer reaction.
Janakiraman and Ordóñez (2012)	Labor experiment	245 participants from a large public university in the southwestern U.S.	Decreasing the product return deadline has the counterintuitive effect of leading to an increase in product return rates in some cases.

of retracting a bad decision and thus enables consumers upfront to purchase while maintaining some flexibility. If consumers have the choice between two equivalent alternatives, they are likely to choose the alternative with the better return policy.

The return policy also influences the after-purchase phase. Return policies without shipping and restocking fees directly decrease return costs and return policies with lenient return deadlines giving the buyers more time to reconsider the purchase. Customer friendly return policies can thus reduce absolute return costs (money, time and labor cost). As Anderson et al. (2009) pointed out that a customer will choose to return products if the net utility is negative and lower return costs ultimately lead to a lower net utility. We thus state hypothesis 1 and 2.

H1: Customer friendly return policies are positively related with sales.

H2: Customer friendly return policies are related with more returns.

We define guarantee credibility in our paper as the credibility of a return policy which can certainly influence the buyers' two-stage decisions. Usually, credibility is difficult to improve or measure (Zhuo, Wei, Liu, Koong, & Miao, 2013), but online shops can easily boost the credibility of return policies directly via support measures – for example, by adopting Guarantee Money (e.g., in Taobao.com), which ensures that the return policy can be enforced or sellers can be punished if they do not implement their return policies correctly. At Taobao.com sellers can allocate an arbitrary amount of money (Guarantee Money) that the platform operator can use to compensate unsatisfied buyers if there is dispute between buyer and seller. On the one hand, guarantee credibility is thus a strong signal for high quality and credibility of the return policy of the focal sellers. Thus, we suggest that guarantee credibility could also be a factor that directly affects customer purchasing behavior. For example, large firms could use guarantee policies that small firms cannot afford to imitate. If low-quality

firms follow though, they will suffer significant financial losses when they are caught cheating. On the other hand, guarantee credibility can effectively propel return policies forward which makes it easier for buyers to return their products. Similar as return policy, guarantee credibility could significantly reduce the absolute value of return cost. So we derive hypotheses 3 and 4 as follows:

H3: Guarantee credibility is positively related with sales.

H4: Guarantee credibility is related to more returns.

Reputation is another important quality signal (Zhang et al., 2012). Purohit and Srivastava (2001) provided a classification scheme to differentiate between transient and durable signals. Durable signals, such as reputation, are classified as “high-scope” signals, meaning that the cue has evolved over time and cannot be changed easily. More transient cues, such as return policies and guarantee credibility, are classified as “low-scope” signals, meaning that they are fairly easy to change and weigh relatively less as a signal of quality in contrast to a high-scope signal (Purohit & Srivastava, 2001). Meanwhile, Roggeveen et al. (2014) found that moderately incongruent signals can be combined to enhance evaluations. In particular, if a firm's reputation and policies are complementary, this can positively moderate the effect of return policies on customer behavior. However, reputation only stands for the past records of a firm; it cannot guarantee the firm’s future behavior. Based on these insights we propose hypotheses 5 and 6 as follows.

H5: Reputation moderates the effect of return policy on sales.

H6: Reputation moderates the effect of guarantee credibility on sales.

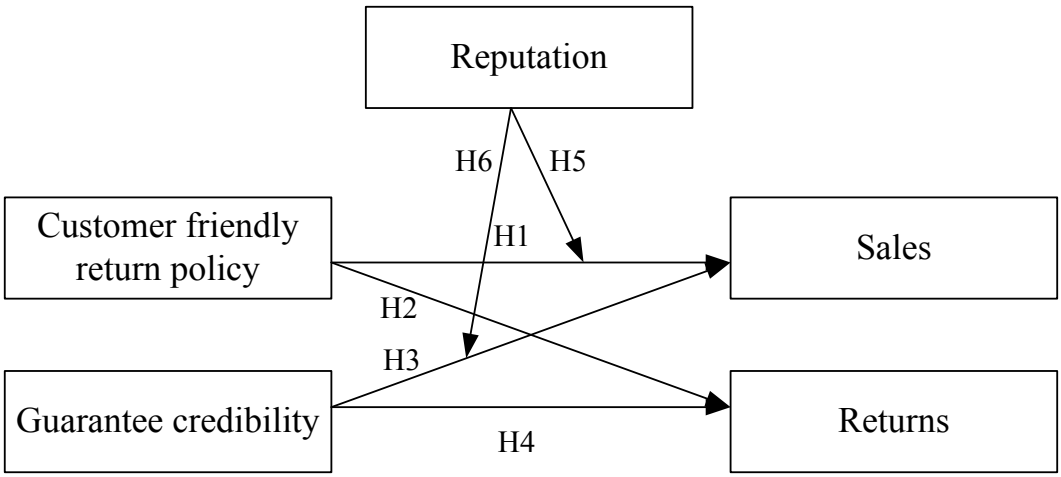


Figure 3-1 Conceptual Model

3.4 Data and Model

Taobao.com belongs to the Alibaba group and is the largest global online B2C and C2C platform. The China based platform started their business in May 2003 and had 500 million registered users by the end of 2013(Ye, Xu, Kiang, Wu, & Sun, 2013). Taobao is a Bilateral Market, and also can be categorized as a third-party market place (Timmers, 1998). Like on eBay, registered users can act both, as sellers and buyers, on Taobao. The platform offer unified ordering, delivery and payment system, but sellers have flexibility to choose their own guarantee strategy. Another characteristic of Taobao.com is that most of sellers are small and medium-sized shops in China who do not sell well-known brands and only focus on one special category like clothing and shoes. Although Taobao has started to offer its service in other countries as well, the main market is still China. According to the three-month Alexa traffic rankings, Taobao.com is ranked 9th and ebay.com is ranked 21th worldwide8. The blossoming of the online market in China has stimulated much interest among marketers.

[8] Retrieved on Dec. 8, 2014

However, there is still limited research that attempts to understand customer behavior or example on Taobao.com.

On Taobao, customers can leave reviews (e.g., comments: good, normal, negative; scores (1-5, 5 =very good) for description, service and delivery) after purchasing; the default value is "good" if customers do not leave feedback within 15 days, but scores have no default values. All information can be found on the ratings pages, including customer reviews and each shop's return rate. We randomly choose 600 online shops, but we could not retrieve all the necessary data for eight of them, leaving a total of 592 shops in our sample. The shops' categories are mainly clothing, shoes, bags, digital products, books and others (see Table 3-2). To control the effect of product types on customer behavior (Bae & Lee, 2011), we use industry data (e.g. Industry Return Rate) as a control variable in our models.

The return rate is the proportion of returns in relation to sales in the observation period (in May, 2012). We also calculated the return rate in March 2014 and found that the changes are not substantial (see Table 3-2). The two main return policies adopted by sellers on the platform are "Consumer Guarantee" and "Seven-Day Return". "Consumer Guarantee" means that customers can return or change products within 15 days if there is a quality problem or if the product description does not match the received product. Moreover, "Guarantee Money" includes a special form of a money back guarantee: Taobao.com administrators guarantee that the shops offering "Consumer Guarantee" will return the customer's funds immediately– a feature that makes the return process both smoother and more credible. The "Seven-Day Return" policy means that customers can return the product for any reason within seven days.

We use two dummy variables to model whether shops adopt these two return policies or not. Table 1-1explains all variables used in this paper.

Table 3-2 Descriptive Statistics

	Min.	1st Qu.	Median	3rd Qu.	Max.	Mean
Sales	0	4	21	73	8324	116.7
Return rate	0.000	0.000	0.015	0.067	1.500[9]	0.070

Return rate	0%	5%	10%	20%	30%	50%	Mean
Accumulation	45.5%	67.8%	84.5%	91.3%	95.7%	97.7%	7%

Guarantee	Seven-Day Return	Consumer Guarantee	Guarantee Money
Ratio of adoption	47.7%	88.7%	76.6% (Range: 0 – 10,000 Yuan)

Industry	Home Accessory	Health Food	Jewel Accessory
Return rate in 2012	0%	1.77%	2.28%
Return rate in 2014	1.62%	2.72%	3.03%
Industry	**Digital Products**	**Book and Media**	**Clothes and Shoes**
Return rate in 2012	2.31%	5.31%	5.53%
Return rate in 2014	4.02%	5.48%	6.58%

To examine the tradeoff between the impact of return policies on sales and the one on returns, we suggest two steps. In the first step, we apply a robust regression model to estimate the influence of return policies on sales. In the second step, we estimate a zero-inflated negative binomial regression (ZINB) to examine the impact of return policies on the return rate, a dependent variable with excessive zeroes. The robust regression model is as follows:

$$Sales_i = \alpha_0 + \alpha_1 * Reputation + \alpha_2 * Seven\ Day\ Return + \alpha_3 * \ Guarantee\ Money$$

[9] Return behavior can happen within 7 days after purchasing, so there can be lagged effects. A return rate>1 means that customers returned products that they have bought in the last 7 days of the previous month which can in this case exceed the sales in the focal month. For example, shop A sold 10 products in May, one of them was returned in May, three of them were returned in June. Then in the next month, Shop A only sold 2 products which the buyers kept but as a result the return rate in June is 1.5.

$$+ \alpha_4 * Customer\ Guarantee + \alpha_5 * Description + \alpha_6 * Industry\ Return\ Rate$$

$$+ \alpha_7 * Return\ Rate + \alpha_8 * Reputation * Seven\ Day\ Return$$

$$+\alpha_9 * Reputation * \ Guarantee\ Money + \alpha_{10} * Reputation * \ Customer\ Guarantee + \ \varepsilon \tag{14}$$

Table 3-3 Measurement of Variables

	Measurement
Reputation	Total amount of good comments since the business' start
Seven-Day Return	Customers can return or change their products for any reason within 7 days
Customer Guarantee	Customers can return or change their products within 15 days if there are quality or service (exp. description misunderstandings) problems
Guarantee Money	Collected by platform operator in advance to pay money back to customers. Sellers can choose the amount of Guarantee Money they allocate to the platform operators.
Description Score	Scores provided by customers after receiving products to describe whether the information offered by shoppers matches the real product(s); the range of the score is 1 (very bad) to 5 (very good)
Return Rate	Number of returns divided by sales in the specific month
Return Number	Returns in absolute numbers
Return100	Return rate * 100
Sales	Number of sales (operationalized by the number of good, normal and bad comments)
Industry Return Rate	Average return rate in this retailer's main business' industry

In the second step, we need to account for the fact that the return rate is heavily skewed (see Figure 3-2), with the majority of the shops experiencing very few returns. About 45.5% of the shops encounter no return behavior, and for two main reasons: 1) there are some new shops that have no transaction records at all; and 2) some shops have customers who are so satisfied that they do not make use of the return option. As a result, we cannot use a regular regression model, so we instead apply a zero-inflated negative binomial regression (ZINB). ZINB models have two stages: one stage is used to estimate the zero-values (non-returns in this case) and the other is used to estimate the actual returns in absolute numbers (Mwalili,

Lesaffre, & Declerck, 2008). All cases are used in both analyses, but they are weighted based on the results of the model's logistic component (see Equation 6). Using this model, we can well explain the zero-inflation, so long as two conditions of ZINB are considered: 1) target behavior is not always happening (we have zeroes and non-zeroes in the dataset); and 2) target behavior has to be any integer, including zero (Heilbron, 1994). To meet these conditions, we use the integer of return rate times 100 (we labeled "Return100", e.g. if the return rate is 7.034%, the "Return100" is 7.) and use return rate times sales (we labeled "Return Number") as two separated dependent variables.

While the data in this study are not true count data, the chosen model is appropriate for two reasons: first, the data are distributed exclusively on the non-negative integers and tend to show heteroskedasticity (exactly like true count data); second, the data appear to be a result of mixture models (i.e., two separate parts cause the need for zero-inflation) (Simons, Neal, & Gaher, 2006). As such, even though the data are technically not generated by a count process, the resulting distribution has the expected characteristics of a count process, and thus a count model is appropriate.

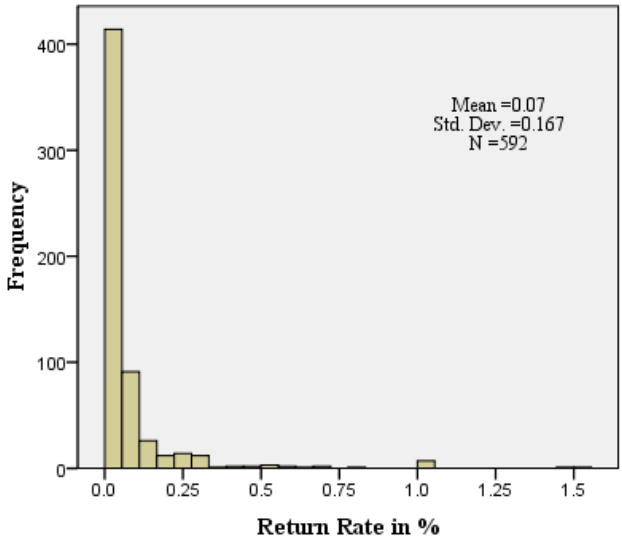


Figure 3-2 Histogram of Return Rate in %

The ZINB distribution is a mixture distribution assigning a mass of p to "extra" zeroes and a mass of (1−ω) to a negative binomial distribution, where $0 \leq \omega \leq 1$. McLachlan and Peel (2004) noted that the negative binomial distribution is a continuous mixture of Poisson distributions, which allows the Poisson mean to be gamma distributed; in this way, over-dispersion is modeled. We will compare the results of the ZINB with the outcomes of a zero-inflated Poisson model (ZIP) later on to demonstrate why we do not use a ZIP model in our study. The negative binomial distribution is given by

$$P(Y = 0) = \omega \tag{15}$$

$$P\big(Y \sim \text{Negative Binomial } (\lambda, \alpha)\big) = 1 - \omega \tag{16}$$

yielding the following distribution of counts:

$$P(0) = \omega + (1 - \omega) * F(0|\lambda) \tag{17}$$

$$P(k) = (1 - \omega) * F(k|\lambda) \tag{18}$$

where F represents the reference distribution (negative binomial with fixed parameter α), represents the predicted probability of being always-zero, modeled by the logistic component

of the model, and represents the predicted mean of the negative binomial component of the model.

Reputation, Seven-Day Return, Guarantee Money, Customer Guarantee, Description and Industry Return Rate were included as independent variables in both components of the models. Thus, the two-part model was parameterized as

$$\omega_i = \frac{e^{\beta_0 + \beta_1 * Reputation + \beta_2 * Seven\ Day\ Return + \beta_3 * Gratantee\ Money + \beta_4 * Customer\ Guarantee + \beta_5 * Description + \beta_6 * Industry\ Return\ rate}}{1 + e^{\beta_0 + \beta_1 * Reputation + \beta_2 * Seven\ Day\ Return + \beta_3 * Gratantee\ Money + \beta_4 * Customer\ Guarantee + \beta_5 * Description + \beta_6 * Industry\ Return\ rate}} \tag{19}$$

$$\lambda_i = e^{\gamma_0 + \gamma_1 * Reputation + \gamma_2 * Seven\ Day\ Return + \gamma_3 * Gratantee\ Money + \gamma_4 * Customer\ Guarantee + \gamma_5 * Description + \gamma_6 * Industry\ Return\ rate} \tag{20}$$

We then use the software package R to estimate both the robust regression and ZINB model.

3.5 Results

We start with a Vuong test and find that a zero-inflated negative binomial (ZINB) regression model is superior over a zero-inflated Poisson model (ZIP) (t-value = 19.16929, $p<0.001$) and a standard negative binomial model (t-value= 4.945271, $p<0.001$). The estimated parameters and the p-values of the models are presented in Table 5 and Table 6. It is clear that return policies can significantly affect both sales and returns, but various policies function in different ways. Reputation is another important factor working in both a direct and an indirect way.

Columns 2 and 3 in Table 3-4 indicate the main factors influencing sales, while columns 5 and 6 present a model that includes the interaction between the shops' reputations and their

return policies. According to the data, good reputation and the Seven-Day Return policy significantly increase sales ($p<.05$) which supports H1. Other return policies like Guarantee Money and Customer Guarantee also have positive effects, but this result is not robust across Model 1 and 2. Hence, we only find partly support for H3. With regard to the moderator effects, better reputation reduces the influence of traditional return policies (Seven-Day Return policy and Customer Guarantee) on sales ($p<.05$), but bolsters the influence of guarantee credibility (Guarantee Money) on sales ($p<.05$), thus supporting H5 and H6. The reason is that traditional return policies and reputation both work as quality signals, thereby acting as substitutes. From a customer's perspective, sellers with a better reputation are likely to offer generous return policies, and it follows that sellers offering better return policies must be more confident about their services and products. However, guarantee credibility extends the reliability of quality signals, thereby serving as a complement to the historical reputation. Guarantee Money is specifically a support measure for future behavior, guaranteeing that buyers can get their money back immediately if necessary.

Table 3-5 presents the results derived from fitting the ZINB regression model to the return data. The R-square is relatively low because we cannot consider customer-specific factors like the customers' personality traits.

Table 3-4 Impact of Return Policies on Sales

Meethod	Robust Regression				
Dependent Variable	Sales				
	Model 1 (Without interaction)		Model 2 (With interaction)		
	Coef.	P-value		Coef.	P-value
Intercept	23.018	0.271	α_0	-5.6236	0.872
Reputation	0.005***	0.000	α_1	0.0410***	0.000
Seven-Day Return	8.970***	0.001	α_2	16.757***	0.000
Guarantee Money	0.019***	0.000	α_3	0.001	0.656
Customer Guarantee	-0.582	0.887	α_4	11.520**	0.005
Description	-4.646	0.543	α_5	0.225	0.974
Industry Return Rate	111.770	0.241	α_6	135.383	0.126
Return Rate	-13.347 ˙	0.059	α_7	-13.738*	0.035
Reputation* Seven-Day Return			α_8	-0.007***	0.000
Reputation* Guarantee Money			α_9	0.000***	0.000
Reputation* Customer Guarantee			α_{10}	-0.039***	0.000
AIC	8,862.610			8,389.974	
R-square	0.4398			0.7323	

*** p<0.001, ** p<0.01, * p<0.05, ˙ p< 0.1. Significant parts are marked in bold, n=592.

Return number, Reputation and Guarantee Money remain significant in both parts of the ZINB model. Specifically, good reputation can reduce the possibility of certain non-returns (γ_1= -7.038e-04, p<.01) and increase the return number if shops already have return records ($ß_1$ =4.813e-05, p<.01). The reason is that sellers with a better reputation (more past sales) are more likely to face an opportunist (e.g., purchase, use and return of clothing on purpose). And then due to the large number of sales, even with the same return rate, the absolute return

number increases. So in both, the zero-inflation and count model, reputation – as we have operationalized it – increases the return number. Meanwhile, it appears that a higher amount of Guarantee Money reduces the number of returns, likely because it speaks to the return policy's credibility. Specifically, if a shop were to increase its Guarantee Money by one Yuan, the expected return number in one month would decrease by a factor of $\exp(ß_3)=0.998$ ($p<.1$) while holding all other variables constant. But at the same time, the odds that a shop has certain non-returns would decrease by a factor of $\exp(\gamma_3) = 0.997$ ($p<.01$). In other words, the higher the Guarantee Money, the less likely shops have certain non-returns. On the other hand, the Seven-Day Return policy only works in the negative binomial part. By adopting this return policy, shops' expected return number is $\exp(ß_2)=1.414$ ($p=0.110$) times the expected return number for other shops. These results definitely support H2 and partly supported H4.

To our surprise, the Industry Return Rate influences return behavior only in the zero-inflation model ($\gamma_6=4.620e+01$, $p=0.008$). In other words, whether shops have return records largely depends on industry characteristics, while specific return numbers depend mainly on the shops' service or customer fit (as outlined before in our paper).

Table 3-5 Result of Zero-inflated Negative Binominal Regression Model

		Zero-inflated Negative Binomial Regression					
Dependent Variable		**Model 3: Return Number**			**Model 4: Return100**		
		Count model coefficients					
		Coef.	**P-value**	**e $^{\beta \text{ or } \gamma}$**	**Coef.**	**P-value**	**e $^{\beta \text{ or } \gamma}$**
Intercept	β_0	25.87e+01***	0.000	1.719e+11	7.004e+00 ˙	0.086	1.101e+03
Reputation	β_1	4.813e-05***	0.000	1.00e+00	-1.016e-05	0.852	9.999e-01
Seven-Day Return	β_2	3.464e-01	0.110	1.414e+00	-4.989e-01*	0.005	6.072e-01
Guarantee Money	β_3	-2.297e-04 ˙	0.074	9.998e-01	-3.060e-04*	0.011	9.997e-01
Customer Guarantee	β_4	-1.249e-02	0.969	9.876e-01	6.307e-02	0.807	1.065e+00
Description	β_5	-5.101e+00***	0.000	6.092e-03	-8.347e-01-	0.291	4.340e-01
Industry Return Rate	β_6	4.885e+00	0.512	1.322e+02	-4.533e-01	0.879	3.130e-01
		Zero-inflation model coefficients					
Intercept	γ_0	7.484e+00	0.104	1.778e+03	-1.130e+00	0.747	3.231e-01
Reputation	γ_1	-7.038e-04***	0.000	9.993e-01	-6.138e-04***	0.000	9.994e-01
Seven-Day Return	γ_2	-4.643e-01	0.472	6.286e-01	-3.945e-01	0.198	6.740e-01
Guarantee Money	γ_3	-2.486e-03***	0.000	9.975e-01	-1.649e-03***	0.000	9.984e-01
Customer Guarantee	γ_4	-4.334e-01	0.520	6.483e-01	-1.193e-01	0.756	8.875e-01
Description	γ_5	-9.311e-01	0.327	3.941e-01	8.154e-01	0.256	2.260e+00
Industry Return Rate	γ_6	-4.620e+01**	0.008	8.614e-21	-3.635e+01***	0.000	1.634e-16
AIC		2,475.544			2,931.722		
R-square		0.0339			0.0142		

*** p<0.001, ** p<0.01, * p<0.05, ˙ p< 0.1. Significant parts are marked in bold, n=592.

We also examine the return rate (Return100) (see Model 4 in Table 5, R-square = 0.0142). Although the ZINB model for the return number (see Model 3 in Table 5, R-square = 0.0339) fits the data better (according to AIC and R-square value), examining the return rate does yield some valuable information. Specifically, Reputation and Seven-Day Return policy stand out as the major differences: we find that the effect of reputation on returns is not robust. The second one shows that the adoption of the Seven-Day Return policy would increase the return number, but at the same time decrease the return rate ($ß_2$= -4.989e-01, $p<0.01$). In other words, this return policy increases sales and returns concurrently, but the influence on sales is stronger than on returns, thereby creating the possibility for higher profits.

3.6 Counterfactual Simulation

Using a counterfactual simulation, we highlight the managerial insights that emerge from the results of our two-step analysis. Managers usually want to acquire better knowledge about how return policies affect both profit and return cost in order to make informed decisions. According to McKinsey's report in 2013, Chinese e-retailers realize margins of 8-10% of earnings before interest, taxes, and amortization, which are slightly higher than the average margin for physical retailers (Dobbs et al., 2013). Meanwhile, return behavior can reduce profits by 3.8% on average (Blanchard, 2007). By integrating the results from model 1 in Table 3-4 and model 3 in Table 3-5, we can calculate the margin for return policies and their impact on profits. Adopting the Seven-Day Return policy, for instance, can produce an additional margin gain of +0.29% compared to shops without this policy (see the Appendix

for the detailed calculation). Meanwhile, adopting the Guarantee Money policy would increase profits. Increasing Guarantee Money by 1 Yuan would increase profits by +0.016%. Although the effect is most likely not truly linear, we carefully conclude that an additional 100 Yuan of Guarantee Money could increase profits by +1.6%. As a result, we suggest that retailers on online platform like Taobao should choose effective return policies, both traditional policies (e.g., Seven-Day Return policy) and polices that increase the guarantee credibility.

3.7 Conclusions

To help small and medium-sized online shop optimize their return policy, our paper proposes a holistic view on the decision process and take the double-sided influence (both sales and returns) of return policies into account. And we find that both adopting customer-friendly return policies and increasing guarantee credibility can significantly increase profits for small and medium-sized online shops.

Although some effects seem obvious, especially those pertaining to the first purchasing phase, the existence of a distinct second phase, i.e. the decision to return or keep a product, leads to more complex effects. The adoption of return policies and higher guarantee credibility results in increased sales, while reputation works as a moderator that decreases the influence of traditional return policies and increases the influence of guarantee credibility on sales. In the second step, we applied a zero-inflated negative binominal regression model and found that reputation, guarantee credibility, and the average return rate in the particular

industry contribute to whether shops have return records at all. For those shops with return records, good reputation and traditional return policies (like Seven-Day Return policy) can significantly increase the return number, while higher guarantee credibility and a better product description could reduce it. It is also interesting to note how the impact of return policies differs between the return number and the return rate. For the non-zero part, reputation and return policies enhance the return number, but not the return rate, which means that sellers would benefit from enhancing their reputations and adopting any available return policies.

Our results are relevant for the large number of actors on Taobao and can potentially be generalized to small and medium-sized online shops that face a similar situation like actors on Taobao. Moreover other platform operators like eBay could revisit the portfolio of return policies that sellers on these platforms can offer.

3.7.1 Theoretical Implications

The theoretical contributions of this work are threefold. First, guarantee credibility not only works as a mediator between return policy and perceived risk, as it ultimately increases the customer's willingness to pay (Suwelack et al., 2011), but seems to also be a new dimension of return policies that directly affects customers' purchasing or returning behavior. Dimensions of return policy should thus not be restricted to return deadline, customer effort and return coverage (Posselt, Gerstner, & Radic, 2008; Su & Zhang, 2009), but should also include guarantee credibility.

Second, the results of this study show that the return policy and guarantee credibility work in different ways. In this paper, we treat reputation as a moderator and find opposing moderating effects relative to return policy and guarantee credibility. This result shows that traditional return policies work as a quality signal akin to reputation, while guarantee credibility works as signal quality that is complementary to reputation because it is a guarantee for future behavior.

Third, our study adopts a zero-inflated negative binominal regression model to explain online return behavior. We find evidence that this model is better suited than other regular regression models because ZINB can analyze both zero-values (current non-returns in this case) and counts parts among actual returns separately. The existence of differences between analyzing these two parts shows that this model might be useful for obtaining deeper insights into return behavior.

3.7.2 Managerial Implications

Our paper suggests that operators of online platform like ebay.com and Taobao.com should offer both traditional return policies and policies that increase the guarantee credibility. In this way, small and medium-sized online sellers on these platforms can increase sales as well as decrease costs incurred from customer return behavior, thereby ultimately increasing profits. Now, most global online platforms do not offer guarantee credibility policies, although such policies provide clear benefit for all parties. For example, Guarantee Money is a “win-win-win” policy insofar as it offers better service to customers, stimulates sales, and

increases the turnover for platform operators. In this paper we suggest a two-step approach for examining the influence of return policies on sales and returns. The proposed approach combines robust regression and ZINB regression, and allows users to easily estimate the results with available transactional data. This approach proves to be valuable for making informed decisions about the optimal return policies for each shop. Finally, we suggest that managers do not focus only on return numbers, but also pay more attention to changes in the return rate because some return policies might increase the absolute return number but not the return rate itself.

For small and medium-sized online sellers without a platform, customer friendly return policies and guarantee credibility would be more important. Because lacking of the protection (umbrella-functionality) of the third-party, customers have higher level of uncertainty. As a result, more tolerant return policies should be fully implemented.

3.7.3 Limitations and Future Research

This study only examines return policies and guarantee credibility in the online environment. Significant differences between online and offline markets limit the generalizability of our results. For example, offline purchases do not suffer from severe information asymmetry and the fit between expectation and actual outcomes is higher. This leads to lower return intentions after purchases in brick-and-mortar stores. So our results cannot be transferred blindly to offline markets. Future research could compare the effect of return policies between online and offline markets.

Moreover, the insights generated by our results may be limited due to cultural differences. Almost all the shops on Taobao run their business in China and the buyers typically come from China as well, so a cultural bias may exist in our study too. We plan to test for culture influences on customer return behavior in the future.

Finally, the R-Squares of our models are relatively low. This might be due to the missing information on customer personality which likely also effects return behavior. A combination of data on the shop and on buyers' level seems promising.

Appendix

Change for sales (**ΔS**): Change due to return policy [α_2 or α_3] /Average number of sales [116.726]

Change for returns (**ΔR**): Change due to return policy [$ß_2$ or $ß_3$]* Average return rate (with return record) [0.134]

Change for return probability (**RP**): Change due to return policy [γ_2 or γ_3]* Non-return rate [45.5%]

Profit (**P**) = (1+**ΔS**)*[1-**ΔR*(1- RP)**]* Margin - **ΔR*(1- RP)*** Return cost

where Margin is 10% (Dobbs et al., 2013) of revenue, while Return cost is 3.8% (Blanchard, 2007) of revenue.

Profit of normal sellers ($\mathbf{P_{normal}}$) = 1*(1-7.4%)*10% - 7.4%*3.8%=8.9788%

Profit of sellers using Seven-Day Return ($\mathbf{P_{7\text{-}day}}$) = (1+0.0768) * [1-0.1895 * (1-45.7%)] * 10% - 0.1895 * (1- 45.7%)*3.8% = 9.269%

Increase in Profit from by Seven-Day Return Policy ($\Delta\mathbf{P_{7\text{-}day}}$) = $\mathbf{P_{7\text{-}day}}$ - $\mathbf{P_{normal}}$ = 0.29%

Profit of sellers using Guarantee Money ($\mathbf{P_{money}}$) = (1+0.00016) * [1-0.13397 * (1-45.58%)] *10% - 0.13397 * (1-45.58%) *3.8% = 8.995%

Increase in Profit from by Guarantee Money Policy ($\Delta\mathbf{P_{money}}$) = $\mathbf{P_{money}}$ - $\mathbf{P_{normal}}$ = 0.016%

Chapter 4: The Impact of The Package -opening Process on Product Returns

Authors: Wenyan Zhou, Technische Universität Darmstadt, Germany

Oliver Hinz, Goethe-Universität, Gemany

Alexander Benlian, Technische Universität Darmstadt, Germany

Published in: Business Research

Abstract: High product return rates are an increasingly pressing challenge for many e-retailers around the world. To address this problem, this paper offers a new perspective by focusing on the critical moment of the package-opening process. Going beyond previous research, which has primarily focused on website information and the product itself, we examine the effects of the outside appearance (i.e., the color of the delivery package) and contents of the delivery package (i.e., extra gifts, coupons, and preprinted return labels) on consumer return behavior. Our findings across two experimental studies and an observational field study show that a well-considered package design, including colorful packaging and extra gifts, significantly lowers consumers' return intentions and actual returns. We also explore the process of consumers' cognitive–affective reactions after opening a delivery package. During this two-stage reaction process, pleasure plays a crucial role in the consumer's return choice.

Keywords Online purchase; Consumer return behavior; Color; Delivery packages; Free gifts

4.1 Introduction

Rising online return rates pose a serious threat to e-retailers around the world. One recent survey showed that 36.1% of German e-retailers selling fashion and accessories experienced a return rate of 20% or more in 2014 (Institut für Demoskopie Köln 2015). In the United States, the average return rate reached 33% in 2012 and is projected to increase even further in the future (Banjo 2013). To compound matters, 72% of e-retailers bear the costs of delivery and/or the extra labor and management costs for handling the returns, which can ultimately render a lenient return policy very expensive. Researchers and managers are currently seeking ways to mitigate this problem, but our understanding of what drives high return rates remains limited.

Researchers agree that online purchasing can be viewed as a two-stage decision process (Wood 2001): the decision to order (the purchase decision) and the decision to keep or return the ordered product (the post-purchase decision). The purchase decision itself is more time-consuming than the post-purchase decision: Consumers typically spend a great deal of time collecting and processing information from several e-retailers and comparing different products, brands, and prices. In contrast, they usually make the post-purchase decision quickly, sometimes within a few seconds after opening the package. Therefore, it is critical to determine the factors that influence the post-purchase decision during this brief yet decisive period.

A vast array of literature on consumer return behavior pertains to website information and the product itself (see Figure 4-1), primarily using signaling (Janakiraman and Ordóñez

2012), utility (Wood 2001), and expectation (Bechwati and Siegal 2005) theories. In general, research indicates that a good reputation (Zhou and Hinz 2016), a high review score (Sahoo et al. 2015), and higher-quality products (Anderson et al. 2009b) can reduce consumer return intentions by increasing perceived utility and signaling higher quality. Other studies show that a lenient return policy can increase returns because of lower return costs for the consumer (Wood 2001), while an accurate product description on purchase websites can decrease return intentions by narrowing consumers' expectation gap (Heiman et al. 2001).

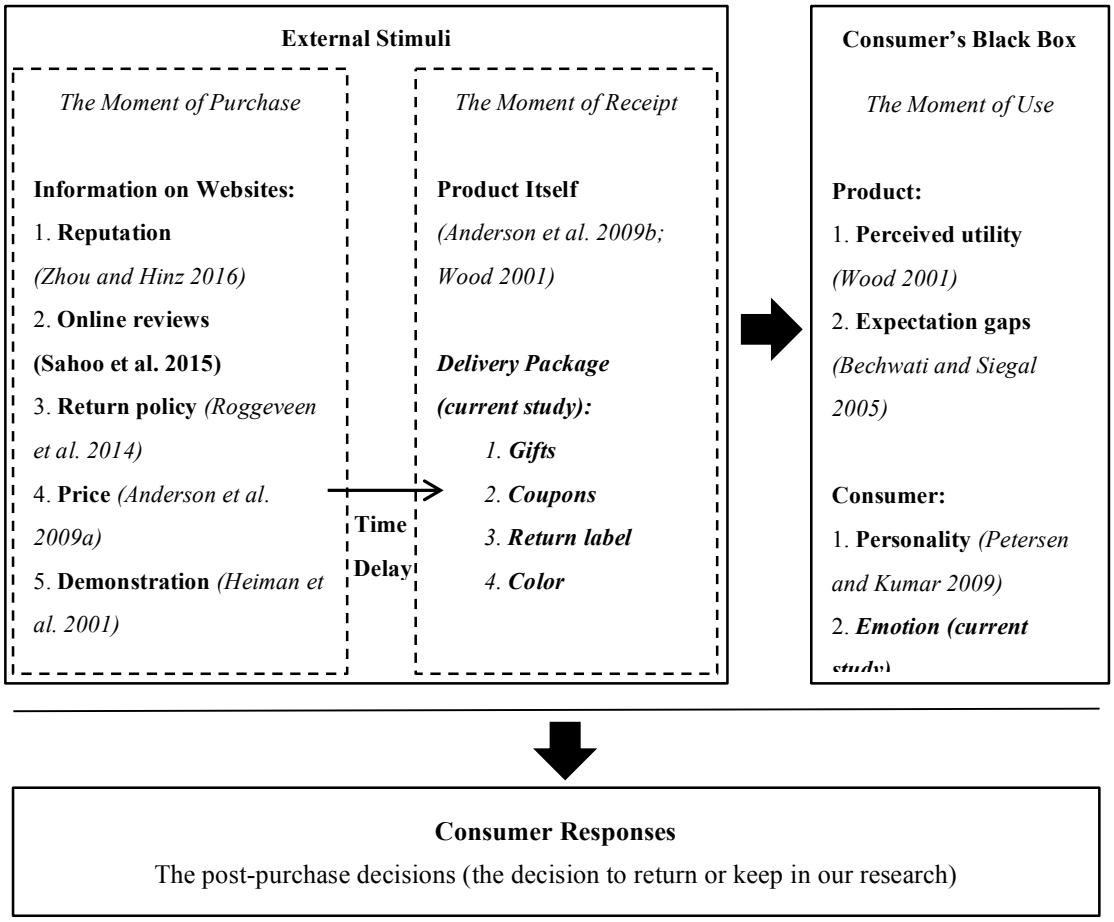


Figure 4-1 Factors that can Influence the Post-purchase Decisions in Different Moments

However, because of the time delay between purchasing and receiving a product ordered online, consumers may not clearly remember all the purchase details at the moment of receipt

and may not visit the respective website or check old e-mails to reacquaint themselves with actual purchase information or return policies. As a result, above and beyond purchase details, delivery package cues are likely one of the last opportunities for e-retailers to influence consumers' product return behavior (Garretson and Burton 2005). The information carried by the delivery package is probably clearer and fresher than what remains in consumers' memories (i.e., the stimuli shown at the moment of purchase). Thus, to address the problem of high return rates, both e-retailers and researchers need to recognize the importance of delivery package design, even though little, if any, extant research has done so.

The composition of the delivery package design includes the outside appearance (e.g., color, shape, size) and its contents (e.g., extra gifts, coupons, return labels, fillers, and receipts). We focus on color, extra gifts, coupons, and return labels as the four most critical aspects in our study, because they do not heavily rely on the characteristics of ordered products (i.e., we do not consider shape, size, and fillers) and can potentially influence consumer behavior (see Section 4.2). Moreover, e-retailers can easily manipulate all these factors.

Against this background, we pose the following research question: *How does the package-opening process influence consumers' return decision and why?* The question includes two parts: consumers' responses to external stimuli (i.e., the delivery package design) at the moment of opening the package and their reaction process when they make return decisions. Answering these questions is crucial for research and practice alike. Theoretically, the study advances research on the drivers of consumer return behavior and sheds more light

on the return decision process. A new perspective on the package-opening process would also aid in analyzing other consumer post-purchase decisions. Practically, addressing this research question identifies several controllable factors that e-retailers can alter to reduce return rates.

This article proceeds as follows: In Section 4.2, we review the literature on consumer return behavior, as well as the potential impact that designed packages can have on consumer behavior. In Section 4.3, we present an experiment (wherein we imitate the purchasing and package-opening process using stop-motion videos) that tests the impact of package design on consumers' return intentions (Study 1). We further apply partial least squares (PLS) regressions to understand consumers' cognitive and affective reactions during the opening process. Section 4.4 describes Study 2, which serves as a robustness test and extends the analysis to real return behavior in an experimental setting. In Section 4.5, we use transactional data from China's largest online platform (Taobao.com) to examine the relationship between package design and a web shop's return rates (Study 3). Section 4.6 concludes with a discussion of the implications and future research avenues.

4.2 Literature Review and Conceptual Framework

Our study aims to clarify the entire process between consumers' reaction during the package opening period to their final decision to return or keep. Thus, we consider three types of processes that occur during one's exposure to a stimulus event (Berkowitz 1993; LeDoux 1995; Shiv and Fedorikhin 1999), which is the opening of a delivery package. The first is the 'low-road' affective[10] processes, which are evoked rapidly and automatically in the limbic

[10] The 'low-road' affective process is highly evolutionary and is designed to protect individuals from life-threatening danger, and to elicit defensive responses without conscious thought. The 'high-road' affective/cognitive processes, by contrast,

systems of the brain. The second is the 'high-road' cognitive[11] processes that involve the cortical systems of the brain. In this stage, consumers collect and process information by thinking, reasoning, and being aware of their surroundings. The third is the 'high-road' affective processes, which arise from the outcomes of the second process (Shiv and Fedorikhin 1999), and occur relatively slowly compared to low-road affective reactions. All these affective and cognitive reactions can influence consumer return behavior and will be discussed in the next sections.

4.2.1 Cognitive Reactions in Consumer Return Behavior

Prior literature usually considers consumer return behavior a consequence of cognitive reactions. According to utility theory, consumers return products only if the net perceived utilities (i.e., the utilitarian utility and hedonic utility) are negative. The net perceived utilities are evaluated by a customer's perceived utilities at the purchase stage, the perceived fit of the transaction (e.g., physical fit or sensory related to color) and the return costs (Anderson et al. 2009b) at the moment of the receipt. In the situation of online shopping, the perceived utilities (including utilitarian and hedonic utilities) are generated by product information given on websites and stored in consumers' memories. The perceived utilities can be increased by lower purchase prices (due to a higher customer surplus) (Anderson et al. 2009a), a good reputation, and a lenient return policy (i.e., longer deadlines, less required return effort, and more coverage (Roggeveen et al. 2014)). Return costs are the fees that

---

involve an indirect pathway to the amygdala. In this case, thalamic info is transmitted to the sensory cortex where it is further processed and evaluated for level of threat prior to being sent to the amygdala (LeDoux 1995).

consumers must bear if they decide to return the ordered products, such as restocking fees and frictional costs for bringing the parcel to a post office (Frischmann et al. 2012), which could be reduced by a lenient return policy (e.g., a preprinted return label).

The expectation gap (also called fit of transaction in some articles) is a term frequently mentioned in the consumer return behavior literature (Anderson et al. 2009b; Petersen and Kumar 2009). An expectation gap exists when the perceived utilitarian and hedonic utilities at the moment of receipt are not equal to those perceived by the consumer at the moment of purchase. This phenomenon can be attributed to the following two characteristics of the two-stage decision process: (1) the time delay between payment and receipt, which allows disconfirming information (e.g., advertisements from competitors) to provide a potential negative influence (Bechwati and Siegal 2005), and (2) the information asymmetry between sellers and buyers (Anderson et al. 2009b), which can result in a difference between the expectations created by website information and the perceived utility of the actual product. Larger negative expectation gaps at the moment of product receipt can increase consumers' return intentions. There are basically two solutions to narrow the negative expectation gap: One way is calibrating consumers' perceived utility at the moment of purchase. E-retailers can create reliable expectations by offering more detailed and accurate descriptions, as well as services like user-generated product evaluations (Zhou and Hinz 2016). Alternatively, e-retailers can enhance consumers' perceived utilities at the moment of receipt, e.g., through a well-designed package.

Although prior research has paid adequate attention to the cognitive reactions in consumers' return behavior, no study has, to our knowledge, considered the potential impact of the delivery package. A delivery package with a suitable color, extra gifts, or coupons might also enhance the perceived utilities of the entire transaction, which could ultimately decrease consumers' return intentions and actual return behavior.

Color and cognitive reactions

Color is an aesthetic stimulus widely used in marketing (i.e., advertisement, logo, brand personality, package, and atmospherics) to grab consumers' attention, enhance purchase intention, and increase perceived service quality (Labrecque et al. 2013; Puccinelli et al. 2013). Color can convey both embodied meaning and referential meaning (Zeltner 1975). Referential meaning emerges from the network of semantic associations and, as such, is a dynamic and reflexive process (Elliot and Maier 2007). According to memory models, people store semantic information in a complex network comprising conceptual nodes (external stimuli) and links (pathways between nodes) (McClelland 1988). As nodes become activated (e.g., through stimulation by colors), the activation spreads to additional nodes through links. For instance, many premium brands (e.g., IBM, Walmart, Volkswagen) use blue in their logos and product package designs because consumers link blue with competence; the color is further associated with intelligence, communication, trust, efficiency, duty, and logic. Thus, a brand with blue hues can positively affect consumers' perceptions of the brand's competence (Labrecque and Milne 2012), which can lead to greater perceived utilities. However, these associations can differ among various cultures (Madden et al. 2000). For

instance, the color “red” in China signals auspiciousness, while in the western world, “red” is related to anger (Jacobs et al. 1991).

The delivery package design can also serve as part of a brand image. Thus, suitably colored delivery packages can activate the node of a competent brand and/or high-quality products, and thereby increase consumers’ evaluations of the entire transaction.

Gift, coupons and cognitive reactions

The two main categories of sales promotions are nonmonetary (e.g., extra gifts) and monetary (e.g., coupons), which provide consumers with an array of hedonic and utilitarian benefits (Chandon et al. 2000). Specifically, the hedonic benefits include value expression (i.e., the expression and enhancement of the self-concept and personal values) entertainment, and exploration, while utilitarian benefits include savings, quality, convenience, and value expression. While many aspects of Chandon et al.'s (2000) “benefit congruency framework” do not apply to the post-purchase stage, typical sales promotions can still exert positive influences (Liu and Chou 2015). Suitable extra gifts can increase utilitarian benefits by saving money that would have gone to additional products (e.g., accessories to the products purchased) or by providing a convenient way to try new products (e.g., a trial product). Moreover, extra gifts can create hedonic benefits by offering entertainment functions (e.g., a whistle as a gift with a football jersey).

Coupons bring more utilitarian benefits than hedonic ones (Chandon et al. 2000). Specifically, coupons could result in money savings by offering discounts to be redeemed with the next purchase.

In sum, extra gifts and coupons in delivery packages may increase perceived utilities by creating additional utilitarian and hedonic experiences, which can ultimately reduce product returns.

4.2.2 Affective Reactions in Consumer Return Behavior

Consumers' affective reactions have two main dimensions: pleasure and arousal (PA model) (Chang et al. 2014; Ladhari 2007; Mazaheri et al. 2014; Mazaheri et al. 2012). Pleasure is the degree to which consumers feel happy, good, contented, or joyful (Mehrabian and Russell 1974); arousal is the degree to which consumers feel excited, stimulated, awake, or active (Mehrabian and Russell 1974). Both pleasure and arousal are important affective responses and can be influenced by a well-considered package design. In contrast to the widespread concern about cognitive reactions, research attention to the affective reactions (e.g., emotions) in consumer return behavior has been limited to date.

We found no literature referring to the multiple-stage reaction process in a product return context—not even studies solely addressing affective reactions. The only relevant stream of literature we found is situated in the context of service failures (Aurier and Guintcheva 2014; Ladhari 2007; Oliver 1993). Service failures and product returns entail a similar decision-making process, insofar as affected consumers can apply for a refund or compensation when they are not satisfied with the offered service/product. In the context of service failures, the type and degree of a service failure, in tandem with the offered remedial measures, can sway consumers' emotions. In line with the Appraisal-Tendency framework

(Scherer et al. 2001), these affective responses can then influence consumers' post-purchase decisions (e.g., loss or maintenance of access to goods/services) and ultimately their satisfaction (Hibbert et al. 2012) and loyalty (DeWitt et al. 2008). Specifically, when consumers have more negative emotions towards a service failure or are more aware of it, they are less likely to be satisfied and might not maintain their purchase decisions (Choi and Mattila 2008).

It is worth noting that the emotions mentioned in previous service failures research are "high-road" affective reactions (i.e, anger and regret), as they are the outcomes of cognitive reactions (i.e., appraisals about service failure) (Bonifield and Cole 2007; Choi and Mattila 2008; Gelbrich 2010; Smith and Bolton 2002). However, marketing researchers have also explored "low-road" affective reactions, e.g., emotions evoked by color, gift and coupon. As we mentioned before (in Section 4.2.1), colors convey an embodied meaning, which is not learned but rather driven by stimulation that is embodied in colors (Meyers-Levy and Peracchio 1995). Thus, marketers commonly use a long-wavelength color (e.g., red) in the pre-purchase stage (e.g., advertisement) to stimulate arousal (Mehta and Zhu 2009) (e.g., excitement) and tend to use a short-wavelength color (e.g., blue) in the post-purchase stage to lower arousal (e.g., relaxation). Gifts or coupons, as an effective promotion strategy, can create a positive surprise when they appear in the package (Heilman et al. 2002). This positive surprise may cause consumers to experience greater perceived pleasure and perhaps even arousal.

However, as the multiple-stage reaction process for return or keep decisions passes quickly, it is difficult to clearly distinguish "low-road" affective reactions from "high-road" ones. Nonetheless, given that return behavior requires deliberation, we decided to focus on a two-stage "high-road" cognitive-affective reaction process, in line with other studies on service failure (Chitturi et al. 2007; Schoefer 2010).

Finally, researchers have also identified personality (Petersen and Kumar 2009), demographic variables, and industry characteristics (Anderson et al. 2009b) as influential factors in consumer return behavior; therefore, we controlled for these factors. Figure 4-2 shows the resulting conceptual framework of our study.

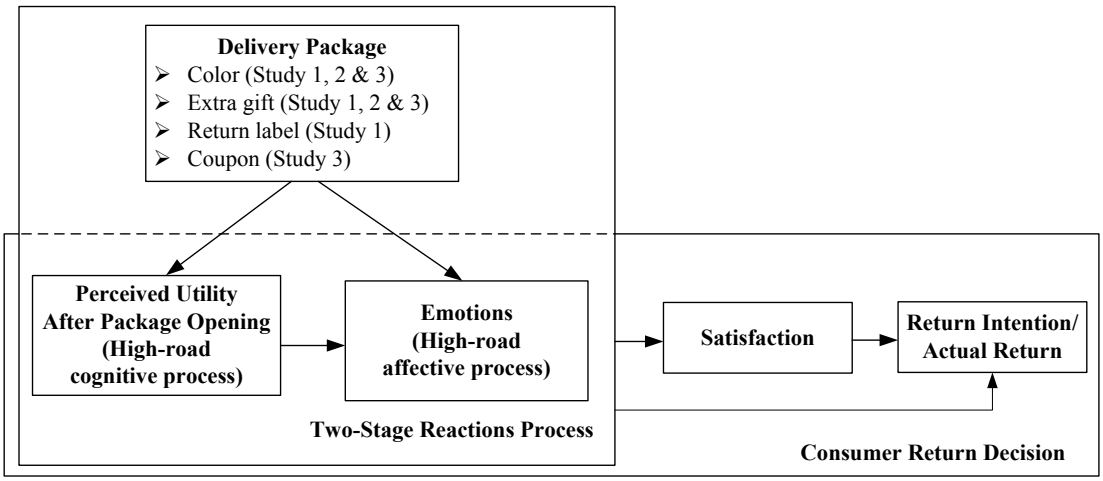


Figure 4-2 Conceptual Framework

We conducted three empirical studies to estimate the impact of various package designs on consumer return behavior (see Figure 4-2). Studies 1 and 2 are controlled experiments. Study 1 focuses on the impact of the package-opening process on consumers' return intentions, while Study 2 observes consumers' real return behavior after receiving packages in different designs. Study 3 is a field study, exploring the relationship between package designs and e-retailers' return rates in a real online market. In both Study 1 and 2, we used

blue as the hue for a colorful package and brown for the control group. In Study 3, we compared the commonly used color for packages (i.e., brown) to all the other colorful hues (e.g., red, blue and black). The consistent results of all three studies provide convincing evidence that the package-opening process can influence consumers' return decisions.

4.3 Study 1: An Experiment on the Package-opening Process with Soccer Jerseys

In Study 1, we designed an experiment with eight different treatments and invited 375 (43 for a pretest and 332 for the main experiment) subjects to participate. We first used analysis of variance (ANOVA) and analysis of covariance (ANCOVA) to examine the influence of colorful packages, extra gifts, and preprinted return labels on consumers' return intentions. We then estimated a PLS model to better understand how these effects operate through consumers' cognitive–affective reaction processes.

4.3.1 Methods

We based our choice of package stimuli on related research and the pretest results. For the pretest, we invited 43 German participants (50% female; average age 36.1 years [SD = 10.20]) to provide feedback on different manipulations. Figure 4-3 shows the final treatments for both the product and package stimuli. The results of the pretest also show that our scales achieved good reliability.

Product

Four criteria guided our selection of a product: (1) A high proportion of our target population should be interested in buying this product, (2) the product should belong to an industry whose return rate is relatively high, (3) the product should have both utilitarian and hedonic value, and (4) a product defect can be easily manipulated to enhance consumer return intention after the package opening. Keeping all these criteria in mind, we selected a jersey of the German national football team and added a 5mm × 5mm black stain on the back. We decided to use only one black stain because in our pretest, as more than one mark led to extremely high return rates (>90%), strongly reducing the variance in our dependent variable. To ensure that people noticed the stain and assessed the problem similarly, we stated that "You have no idea what caused the stain, but you notice that you might not be allowed to return the jersey after washing it."

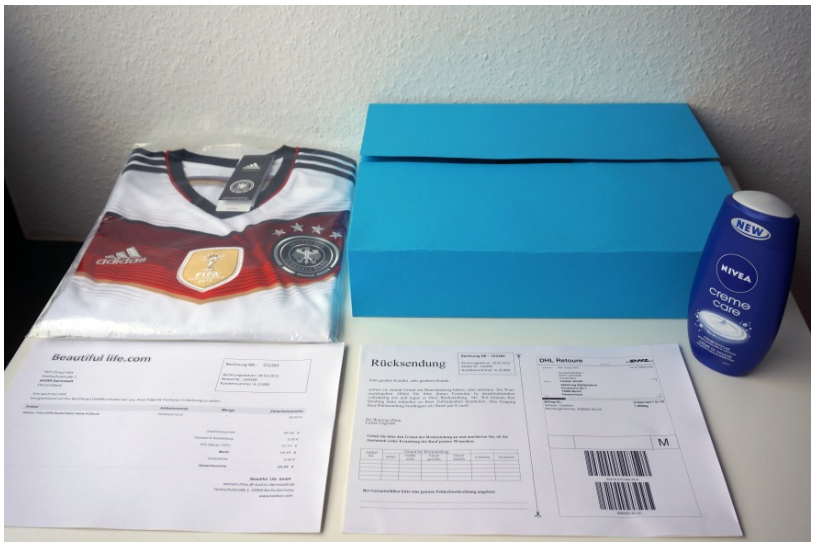


Figure 4-3 Product and Package Stimuli in Study 1

Color

According to a report in 2001, 40% of Germans' favorite color is blue, followed by red (19%) and then green (18%) (Institut für Demoskopie Allensbach 2001). In our pretest, both men

and women indicated that the color of an ideal delivery package, other than standard brown, was blue. Thus, we chose blue-colored delivery packages for our experiment. Crowley (1993) documents that blue has a strong impact on shopping in terms of both evaluation and activation, which meets the requirements of our research goal. The control group received a delivery package in standard brown.

Gift

In the pretest, we also tested the estimated price of various extra gifts. In line with the results, we selected Nivea Creme Care as the extra gift for our main test. The price (approximately €2.5) is 3% of the price of a soccer jersey, and both men and women can use it.

Preprinted return label

We placed a preprinted and prepaid DHL label with a return shipping address into the package. To return the package, participants needed only to glue this return label to the original delivery package and bring it to a post station. In the control group, participants needed to log in to their accounts, complete several forms, and then print the return document themselves. We reasoned that a preprinted, prepaid DHL label could significantly reduce return costs and thus, according to utility theory, increase consumers' return intentions.

Design and procedure

We employed a 2 (colorful vs. not colorful) × 2 (gift vs. no gift) × 2 (preprinted return label vs. no preprinted return label) between-subjects design on the online survey platform

Dynamic Intelligent Survey Engine[12]. In step 1, we randomly assigned participants to one of the eight experimental conditions and asked for demographic information (i.e., age, gender, and career). In step 2, we simulated an online purchase process. Participants were asked to imagine that they had decided to buy a jersey of the German national football team for the upcoming World Cup and then to specify their size and gender in order to obtain the appropriate jersey.

In step 3, we clarified that they were to imagine that they paid for their selected jersey, and then we asked for their emotions (pleasure and arousal) toward and perceived utility (utilitarian and hedonic) of the jersey. For step 4, we needed to create an artificial time delay between the payment and the virtual receipt. Thus, we employed a filler task in which participants answered questions about their online shopping experience and personality by identifying the extent of their extroversion, agreeableness, conscientiousness, neuroticism, and openness (a 10-item short version of the Big Five Inventory in German) (Rammstedt and John 2007). Afterward, participants learned that "after 3 days, you receive your order." Subsequently, in step 5, we told participants, "Please assume that you were the person who opened the package in the video" and then used a 30-second stop-motion animation to show the entire opening process. In stop-motion (also known as stop-frame) animation, an object (in this case, the package) is moved in small increments between individually photographed frames, creating the illusion of movement when the series of frames is played as a continuous

[12] The Dynamic Intelligent Survey Engine, or DISE, is a powerful survey engine built to support and facilitate a variety of advanced data collection methods. See http://www.dise-online.com/.

sequence. This technique allowed us to control the timing and method of package opening. The eight videos in the eight experimental groups were exactly the same except for our manipulations. The gift and/or the preprinted return label appeared for approximately five seconds (six photos for the process taking the items from the package, two photos for a full-screen display of the details, and another two photos for putting the items down; for details, see Table 4-1). We used an amplification process for the gift and preprinted return label to ensure that every participant could recognize each stimulus clearly. Participants could not move to the next step until they finished watching the whole video.

In step 6, we surveyed participants' current emotions and the perceived utility of the whole package, along with their satisfaction and return intentions. To keep the package in participants' minds, we placed a picture of the package, showing all the items, at the top of the questionnaire (see Figure 4-3). In step 7, in order to match their return intention to real return behavior, we communicated that every participant had a chance to win the real package shown in the video (with extra gift/colorful package/preprinted return label and a jersey with a stain) and that they could send the jersey back for a new, flawless one. We then asked whether they would really return their jersey in that case. As an additional motivation and to increase realism, we asked participants to voluntarily give their contact information and jersey size.

Table 4-1 Details of the Stop-motion Animation Video

Manipulation	Number of photos	Display time (Seconds)
Colorful package	60	25s
Extra gift	10	5s
Preprinted return label	10	5s
Dirty-stained jersey	30	10s (5s for the dirty stain)
Total	80[a]	approximately 30s[a]

[a] Some photos include more than one stimulus (e.g., participants saw the extra gift in the colorful package).

In the final step 8, we randomly chose 5 participants and sent them the package exactly as shown in the video of their treatment group and asked them whether they would like to return the flawed jersey. If they wanted to return, they had to bring the parcel to the post office and had to wait until they received their flawless jersey. This additional step allowed us to observe their real return decision and examine whether their answers (return intentions) in the experiment matched their real behavior. Figure 4-4 Experimental Procedure of Study 1summarizes the entire experimental procedure.

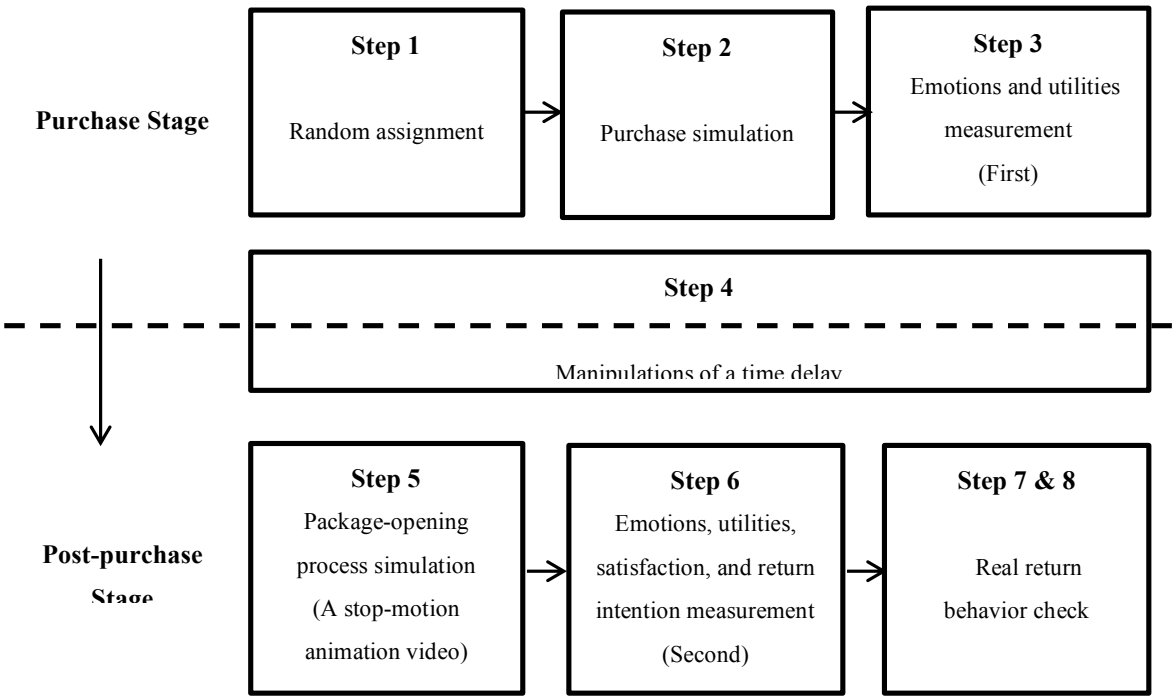


Figure 4-4 Experimental Procedure of Study 1

Scales

We adapted our items for measuring the constructs from prior marketing research (see Table 1-1) using multi-item Likert-type scales for each. We assessed perceived utility using the hedonic/utilitarian scale proposed by Voss et al. (2003). This scale includes eight-point semantic differential items, but we decided to use only seven points according to the Cronbach's α results (>.7). Moreover, we measured emotions using the PA model (including three items for pleasure and three items for arousal) from Mehrabian and Russell (1974). We measured perceived utility and emotion twice—once after participants' purchase decisions and again after the package-opening process. Note that the perceived utility tested following the package opening pertains to the whole package. For consumer satisfaction, we adopted Finn's (2005) three-item scale, which is widely used in marketing research.

To assess consumer return intention, we used the Net Promoter Score (NPS), which is based on an 11-point Likert scale (0 = "not at all likely" and 10 = "very likely") introduced by Reichheld (2003) and widely used to measure attitudes or behavioral intentions (Samson 2006). The NPS is calculated with a single question, in our case, "How likely is it that you would return the package?" We identified participants who responded with a score of 9 or 10 on the NPS as package returners and those who responded with a score of 0 to 6 as package keepers.

In the real return behavior check (Step 7 and 8), we coded participants' answers with a dummy variable equal to 0 if they claimed to keep the whole package shown in the video and 1 if they opted to send it back to get a new one. Although receiving a gift is different from a

real purchase, the return decision is similar in our simulated case. Thus, we believe participants' choice of gift return can proxy for their actual behavior after receiving a product with a small flaw. We then compared participants' return intention (0–6 for non-return, 9–10 for return) and their real return choice (0 for non-return, 1 for return); these two answers were highly correlated ($p < .01$).

Table 4-2 Scale Items

Construct	Item	Cronbach's α	CR	AVE
Utility				
Utilitarian utility	Ineffective/effective	.929	.949	.824
	Helpful/unhelpful			
	Functional/not functional			
	Practical/impractical			
Hedonic utility	Fun/not fun	.918	.948	.859
	Delightful/not delightful			
	Enjoyable/unenjoyable			
Emotion				
Pleasure	Happy/unhappy	.941	.962	.895
	Pleased/annoyed			
	Contented/melancholic			
Arousal	Relaxed/stimulated	.749	.846	.648
	Calm/excited			
	Dull/jittery			
Satisfaction	I feel comfortable with the package.	.943	.963	.898
	The package is satisfying to me.			
	The package is worth the time and money I spent on it.			

All survey items were presented in the respondents' native language (German). We pretested the final questionnaire with doctoral students and university employees to identify unclear and ambiguous questions. The convergent and discriminant validity for the constructs exceeded all critical values (see Table 4-2).

4.3.2 Sample

After conducting a pretest with 43 participants who came from our target population of native Germans with Internet access, we employed a professional market research company to collect a representative sample for our main study in March 2015. Our initial sample for our main study included 332 participants, all of whom had recent online shopping experience. To keep our sample representative within each experimental group, we set quotas for age and gender according to Europe's 2014 online shopping consumer report (Eurostat 2014). To verify the validity of the responses, we checked each participant's response patterns and completion time.

Table 4-3 Descriptive Statistics

	Percentage	Std. Dev.
Age (years)		
16–24	22.5%	
25–34	31.8%	10.21
35–44	25.6%	
45–54	20.0%	
Number of online shopping experiences in the past six months.		
0	.93%	.77
1–3	22.50%	
4–10	41.88%	
>10	34.69%	
Individual return rate after online shopping.		
0%–5%	34.06%	
5%–10%	45.31%	.84
10%–20%	15.00%	
>20%	5.63%	

We excluded five questionnaires that were completed in less than five minutes, six questionnaires that exhibited a visible pattern of the same response on all the Likert scales, and one questionnaire from a participant who reported that his computer was unable to play

the video. The final sample thus consisted of 320 completed surveys (see Table 4-3). An ANOVA revealed no significant differences in participants' age, gender, occupation, and soccer preference among the eight experimental groups, which indicates that our randomization worked as intended.

4.3.3 Common Method Bias Analysis

We strived to design the questionnaire carefully, which entailed ensuring participants' anonymity, using a random order for survey items, providing concrete survey instructions, and asking participants to answer the questions as honestly as possible (Podsakoff et al. 2003). Nevertheless, self-reported data can suffer from common method biases, such as consistency motifs or social desirability concerns (Podsakoff et al. 2003). Thus, we adopted the marker variable approach (Rönkkö and Ylitalo 2011) to test whether a common method bias confounded our results.

We performed the marker variable method (Rönkkö and Ylitalo 2011) with two marker items (two items for Openness, which the ANCOVA in Table 4-4 shows to be unrelated to the dependent variables) taken from our empirical data set; these items were not included in our research model and lack an explicit theoretical influence on the constructs in our research model. Following Rönkkö and Ylitalo's (2011) method, we found relatively low correlations between the marker items and study items (the mean values of the correlation coefficients were .046 and .061) and determined that these low correlations must have been caused by the method. Next, we included the marker items as additional latent variables in our PLS analysis

model and compared the results between the original research model (without the marker variables) and the common method bias test model (with marker variables). The results indicate that the marker variables had no significant effects on the dependent variables (satisfaction and return intention) or on other effective endogenous variables (utilitarian utility, hedonic utility, and pleasure) (see Table 4-10). In any case, only one relationship between the marker variable and arousal was significant; however, because arousal was non-significant (see Section 4.3.6), this finding does not influence our main conclusions. In addition, the path coefficients between all main contrasts and consumer behavior did not significantly differ between these two models. Therefore, we can conclude that a common method bias did not likely distort the main results of our study.

4.3.4 Measurement Model Validation

Our research model contains seven reflective multi-item constructs and six one-item constructs. The quality of the reflective measurement models depends on convergent validity and discriminant validity (Bagozzi and Yi 1988).

To analyze convergent validity, we determined indicator reliability and internal consistency. All the indicator loadings of the reflective multi-item constructs were, at a minimum, significant at the .01 level. For the internal consistency assessment, we examined the composite reliability (CR), Cronbach's alpha, and average variance extracted (AVE) (see Table 4-2) (Teo et al. 2003). All the CR indices, as well as the Cronbach's alpha values, met the threshold of .7 (Nunnally et al. 1967). Furthermore, for AVE, all reflective multi-item

constructs met Fornell and Larcker's (1981) suggested critical level of .5. In summary, the constructs satisfied all criteria for indicator reliability and internal consistency, in support of convergent validity.

We also analyzed the constructs' discriminant validity by examining whether the square root of the indicators' AVE within any construct was higher than the correlations between it and any other construct (Son and Benbasat 2007). All included constructs met this criterion, thus evidencing discriminant validity (see Table 4-11 in Appendix). Moreover, none of the correlations between any pair of constructs were higher than the threshold value of .9 (Son and Benbasat 2007), and there was no evidence of critically high cross-loadings between the main constructs (see Table 4-11 in Appendix). Therefore, we can conclude that the reflective constructs possessed discriminant validity.

4.3.5 Results from ANOVA and ANCOVA

We first used ANOVA to test the significant differences in satisfaction and return intention among different package design groups (color, gift, and return label). We then added participants' demographics and personality to the analysis model as covariates (ANCOVA) to test the stability of the results (see Table 4-4 and Figure 4-5).

The results of both analyses showed that an extra gift can significantly influence consumer satisfaction and return intention, while a colorful package only has a significant impact on consumers' return intentions. More specifically, a colorful package significantly reduced consumers' return intentions ($\bar{R}_{\text{color}}$ = 9.076 vs. $\bar{R}_{\text{no color}}$ = 9.662, see Figure 4-5b; F =

3.66, $p < .1$, see Table 4-4) compared with a standard brown package, but had no significant impact on consumer satisfaction. Meanwhile, an extra gift in the package increased consumer satisfaction ($\bar{S}_{\text{gift}} = 2.576$ vs. $\bar{S}_{\text{no gift}} = 2.072$, see Figure 4-5a; F = 10.685, $p < .001$, see Table 4-4) and reduced return intentions ($\bar{R}_{\text{gift}} = 9.050$ vs. $\bar{R}_{\text{no gift}} = 9.648$, see Figure 4-5c; F = 4.417, $p < .05$, see Table 4-4). These results offer initial evidence for the impact of package design on consumer return behavior.

Table 4-4 Results of ANOVA and ANCOVA

Variables	ANOVA		ANCOVA	
	Satisfaction F-value	Return intention F-value	Satisfaction F-value	Return intention F-value
Color (C)	1.787	3.124*	1.764	3.660 *
Gift (G)	9.607***	5.036**	10.685***	4.417 **
Return label (R)	.360	.940	.274	.994
C×R	2.147	.005	1.822	.000
C×G	1.577	.687	1.529	.659
G×R	.181	1.430	.265	1.004
C×G×R	.047	.018	.041	.001
Covariates				
Gender			2.339	.224
Age			.018	1.297
Extraversion			4.157**	1.815
Agreeableness			5.622**	7.516***
Conscientiousness			4.414**	.002
Neuroticism			3.802**	1.148
Openness			.228	.223
Soccer preference			.199	.400
F-value	2.230***	1.593	2.568***	1.788**
R-square	.048	.034	.112	.081

*p < .1; **p < .05; ***p < .01; N = 320.

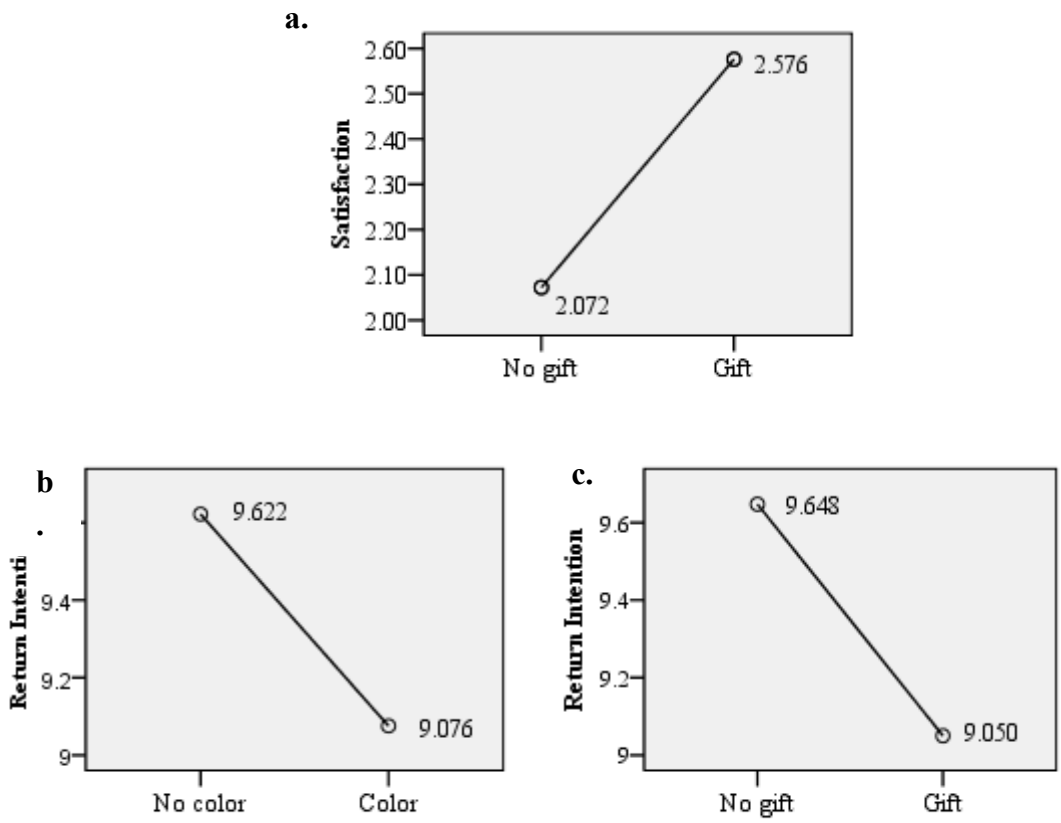


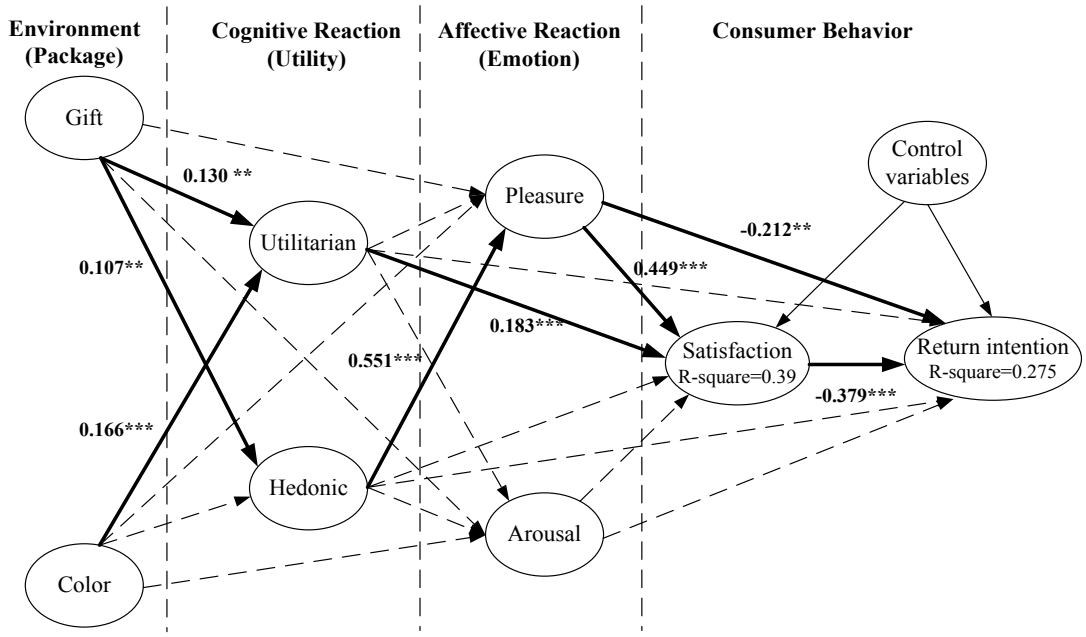
Figure 4-5 Influences of the Package Design on Consumers' Satisfaction and Return Intentions.

Our results further showed that a preprinted return label had no significant effect on consumer satisfaction or return intentions (p >.1, see Table 4-4). However, this result might have occurred because European consumers know that their return rights are highly protected by the Consumer Protection Law, and thus the 14-day return policy is already deeply rooted in their decision processes. The other possible reason is that the preprinted return label does not significantly reduce return costs. We also tested the interactions among color, gift, and return label, but none of them were significant.

4.3.6 Results from PLS analysis

To analyze the package-opening process more thoroughly, we operationalized our model as a structural equation model and estimated it using Smart PLS (v.3.2.1) (Ringle et al. 2015). This method is well suited for exploratory research and shares the modest distributional and sample size requirements of ordinary least squares linear regression. We also used two models to individually test the cognitive process (without affective reactions) or affective process (without cognitive reactions); the results can be found in Appendix (Table 4-10). To reduce common method bias, we included common control variables for our main dependent variables: age, gender, soccer preference, and personality. The main results appear in Figure 4-6.

The squared multiple correlations (R^2) of .39 for satisfaction and .28 for consumers' return intention are high, which means 39% of the variance in satisfaction and 28% of the variance in return intention can be explained by the chosen constructs (Glantz and Slinker 1990). To assess the significance of the path coefficients, we used the bootstrapping procedure implemented in Smart PLS with 1,000 resamples. Figure 4-6 displays the results, with continuous lines representing significant path coefficients and dashed lines indicating non-significant paths.



*p < .1; **p < .05; ***p < .01, solid arrows indicate significant paths, while dashed arrows indicate insignificant paths.

Figure 4-6 PLS Model and Results

Package color can positively influence consumer return decisions, as we expected, but surprisingly, our data indicate it only works through the cognitive process via perceived utilities. These results confirm Chebat and Morrin's (2007) major finding that, in the realm of consumer behavior, the influence of colors is largely facilitated by cognitive rather than affective mechanisms. Specifically, we found that the perceived utilitarian utility of the blue delivery package is relatively higher than the standard brown package (.166, p < .01). In other words, the blue hues associated with a high-value brand can enhance consumers' evaluation of packaged products.

The extra gift significantly increased both the utilitarian utility (.107, p < .05) and the hedonic utility (.130, p < .05) of the whole package, but showed no significant direct impact on arousal and pleasure. The reason might be that because e-retailers commonly offer extra gifts, consumers may not feel special when receiving one. At the same time, consumers can

easily recognize the utility benefits of extra gifts. When comparing the relative impact of gifts and color, the former works more effectively, but the costs of the latter are significantly lower.

Our results also show that utilitarian and hedonic utility impact the consumers' post-purchase decisions in various ways. Higher utilitarian utility increases consumer satisfaction (.183, $p < .01$), which is consistent with previous empirical findings (e.g., Anderson et al., 2009b). In contrast, hedonic utility is positively and strongly related to pleasure (.551, $p < .01$).

In line with our expectations, satisfaction is negatively related to consumer return intention (–.379, $p < .01$). In short, the more satisfied consumers are after opening the package, the less return intention they exhibit. The results also indicate that pleasure plays the most crucial role in consumers' post-purchase decisions. Pleasure is the only factor in our research model that can directly increase satisfaction (.449, $p < .01$) and simultaneously decrease return intention (–.212, $p < .01$). However, arousal did neither influence satisfaction nor return intention. Indeed, the PLS results revealed that arousal had no significant relationship to any other constructs in our research model.

Furthermore, by using the bootstrapping procedure as a mediation test (Suwelack et al. 2011), we found significant indirect effects of the package design (i.e., extra gifts and colorful packages) on emotions and return intentions (see Table 4-5), emphasizing the cognitive-affective reactions process. Specifically, we found that extra gifts invoke more pleasure by increasing hedonic utility (.072, $p< .05$, see Table 4-5). In turn, pleasure can

directly and indirectly (via satisfaction, -0.171, p< .01, see Table 4-5) reduce return intentions. In addition, only a colorful package (.030, *p*< .1, see Table 4-5) can indirectly lead to higher consumer satisfaction, namely by increasing the utilitarian utility. Satisfaction is thus an important mediator, through which utilitarian utility (-.073, *p*< .05, see Table 4-5) and pleasure (-.171, *p*< .01, see Table 4-5) can significantly reduce consumer return intention.

Table 4-5 Indirect Effects

Indirect effects	Coefficient	SD	T-value
Gift->Utilitarian->Satisfaction	.020	.014	1.375
Gift->Hedonic->Pleasure	.072 **	.033	2.204
Color->Utilitarian->Satisfaction	.030 *	.016	1.869
Utilitarian->Satisfaction->Return intention	-.073 **	.032	-2.194
Hedonic->Pleasure->Return intention	-.111 **	.048	-2.454
Pleasure->Satisfaction->Return intention	-.171 ***	.047	-3.534

*p < .1; **p < .05; ***p < .01

Moreover, we also tested the models that solely included cognitive or affection reactions. The results (see Table 4-10) show that extra gifts and colorful packages can have direct effects on perceived utilities, but not on emotions. The only significant direct effect on emotions is the one from extra gifts on pleasure (.097, *p*< .1), but that might be a result of cognitive reactions like hedonic utility (.072, *p*< .01, see Table 5). Furthermore, we tested the model in reverse order (i.e., affective-cognitive reaction process) and found that our package manipulations did not directly influence affective user reactions (pleasure and arousal). Thus, a cognitive-affective reaction process seems more plausible based on our data.

Among the control variables, only agreeableness had a significantly negative effect on return intentions. In other words, consumers who are kind, sympathetic, cooperative, warm, and considerate are more tolerant of product defects, as might be expected.

4.3.7 Real Return Behavior Check

Following the experiment's completion, we randomly drew five winners (2 men and 3 women, average age 25.9 years [SD=10.78]) from the final sample of 320 participants. They received the package as shown in the video of their experimental group. The participants who did not receive a pre-paid DHL label were allowed to email us for a free DHL label (a PDF file). Four of the winners returned the slightly flawed jersey to get a new one and one kept it, which was exactly in line with their stated survey response. This small-number sample may serve as initial evidence that the measured return intention is a reasonable and valid proxy for actual return behavior. This point will be further corroborated in our third study.

4.4 Study 2: A Robustness Test with Respect to Actual Return Behavior with Chocolate Bars

To challenge our findings on return intentions, we conducted a robustness test in January 2017 using an experiment with actual return behavior. We invited 394 students from a German university to a controlled, on-campus experiment. This study was designed to estimate the influence of two factors from Study 1 (i.e., extra gifts and colorful packages) on *actual* return behavior.

First, we presented the participants with a short survey where they rated 6 chocolate flavors of a well-known German brand on 5-point Likert scales. As a reward, they received a numbered voucher to get the chocolate bar of his/her favorite flavor (e.g., flavor A). Participants had to wait two more hours before they could redeem the voucher (we launched

the survey before 11 am and allowed participants to redeem their vouchers after 1 pm). Participants would then receive an envelope containing a chocolate bar but of a non-favorite flavor (e.g., flavor F while flavor A was favored, according to the number on the voucher) in order to create a manipulated product issue. The envelopes were randomly varied by color (brown or blue) and extra gift (with or without a small candy, i.e., mini bag with 5-6 gummi bears). After opening the envelopes, they could find a small note inside saying: "Hi, you can come back for another flavor or receive 1 Euro instead. Have a nice day!"

This study simulates the whole online purchasing process—from ordering (i.e., the choice of their favorite flavor they would get later), a time delay (i.e., two hours), package receipt (i.e., an envelope), a product issue (i.e., a wrong flavor) and actual return behavior (i.e., change for another flavor or a "refund" in the form of 1 Euro) (see Figure 4-7).

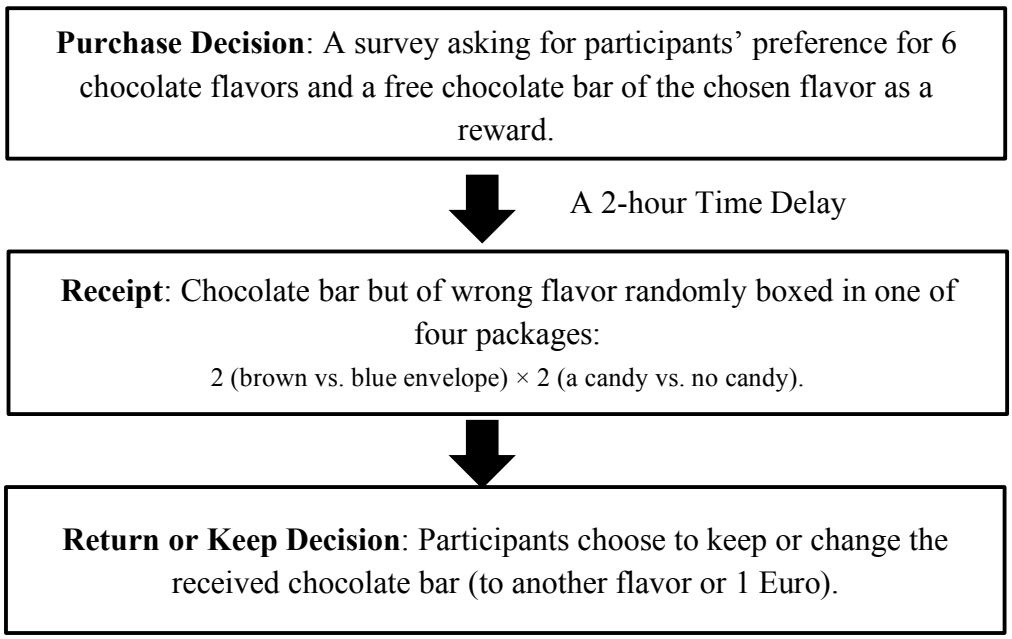


Figure 4-7 Experimental Procedure of Study 2

4.4.1 Method

Of the initial 394 students, 195 (78 females and 117 males) came to redeem their vouchers, while 62 of them returned for a change (see Table 4-6). Aside from asking for the participants' preference in the survey, we also collected information about participants' gender, the degree of liking chocolate ("Like", 5-point Likert scale with 1=very little and 5=very much). We then calculated the standard deviation of the rating scores for 6 flavors ("Variance") and the difference between the chosen flavor and the received flavor's scores ("Gap"). We coded the package that participants received with "Color" (0 means brown envelope; 1 means blue envelope) and "Gift" (0 means without extra candy; 1 means with extra candy). We coded the participants' actual return behavior as a dummy variable: 0 means the participant kept the "wrong" chocolate bar, while 1 means the participant returned it for another chocolate bar or 1 Euro.

Table 4-6 Participants in Study 2 and Their Return Behavior

	Returns	Participants	Return rate
Brown envelope without Gift	22	51	43.1%
Brown envelope with Gift	14	45	31.1%
Blue envelope without Gift	14	50	28.0%
Blue envelope with Gift	12	49	24.5%
Total	62	195	31.8%

We estimated the following equation with cluster-robust (per day of the experiment) errors: Equation 1 shows the estimated logistic regression:

$$\text{Prob}(Return\ Behavior_i = 1) = F(a'X_i) \qquad (21)$$

where Return Behavior$_i$ is a binary outcome variable with observed values 0 (keep) or 1 (return) which stands for the i_{th} participant's keep or return choice, X_i is a vector collecting

related independent variables ("$Color_i$", "$Gift_i$", "$Gender_i$", "$Like_i$", "$Variance_i$", "Gap_i", "Chocolate $Chosen_i$" and "Chocolate $Received_i$") and a potential interaction between "Color" and "Gift" (i.e., "$Color_i \times Gift_i$") to X_i, and α are the coefficients. F() denotes the logistic distribution.We use Maximum Likelihood (ML) to estimate the model. The odds ratio corresponding to the i_{th} coefficient is calculated by $\psi_i = \exp(\alpha_i)$ and it is approximated with the delta method.

4.4.2 Results and Discussion

Table 4-7shows the regression results. These results confirm the positive influences of a colorful package (α_1=-.456, $p < .05$; $ß_1$= -.789, $p < .01$) and an extra gift (α_2=-.462, $p < .01$; $ß_2$= -.819, $p < .01$) on consumers' actual return behavior. Specifically, when we presented a blue package or an extra candy, the odds of returning decreased by about -37% in both cases. Moreover, there was no interactive effect between a colorful package and an extra gift ($ß_3$=.737, $p > .1$), which confirms the ANOVA analysis result of Study 1 (see Table 4-4).

In sum, the results of Study 2—which includes components of ordering, unpacking and actual return behavior—fully support the initial findings from Study 1. Thus, we can claim that our conclusions are not just based on artificial effects, but rather demonstrate good robustness and high validity in other experimental settings.

Table 4-7 Regression Results of Study 2

Return Behavior (0/1)	Measurement	Without interaction			With interaction		
		Coef.	Odds Ratio	Std. Err. of Coef.	Coef.	Odds Ratio	Std. Err. of Coef.
Intercept	0/1	.677	1.969	.785	.892	2.439	.797
Color	0/1	-.456	.634	.192**	-.789	.454	.209***
Gift	0/1	-.462	.630	.067***	-.819	.441	.303***
Color * Gift	0/1				.737	2.089	.476
Gender	0/1	-.082	.921	.342	-.034	.967	.362
Like	Degree of liking chocolate (scale 1-5)	-.179	.836	.134	-.216	0.806	.133
Variance	S.D. of 6 Flavors' scores	.123	1.130	.775	.167	1.181	.729
Gap	$Score_{choose}$-$Score_{received}$ (range 0-4)	.103	1.108	.087	.104	1.109	.094
Chocolate Chosen	Categorical variable						
Flavor A		Fixed	1		Fixed	1	
Flavor B		-.285	.752	.111**	-.305	.737	.104***
Flavor C		-.974	.378	.258***	-1.044	.352	.266***
Flavor D		-.717	.488	.629	-.778	.459	.676
Flavor E		-.785	.456	.794	-.743	.476	.785
Flavor F		-1.612	.200	1.187	-1.511	.221	1.208
Chocolate Offered	Categorical variable						
Flavor A		Fixed	1		Fixed	1	
Flavor B		-2.190	.112	.438***	-2.267	.104	.480***
Flavor C		-.705	.494	.723	-.740	.477	.732
Flavor D		-1.239	.290	.929	-1.309	0.270	1.002
Flavor E		.243	1.275	.555	.265	1.303	.536
Flavor F		-.149	.862	.345	-.202	.817	.305
Pseudo R-square			.117			.122	

*p < .1; **p < .05; ***p < .01; N=195.

4.5 Study 3: An Analysis of Transactional Data on Taobao.com

As a follow-up study, we analyzed transactional field data to examine whether package design (i.e., color, gift, and coupon) is related to e-retailers' return rates. We analyzed transactional data from the largest online shopping platform in China, Taobao.com, and used web shops' return rates as the dependent variable.

4.5.1 Method

Taobao.com was founded in 2003 and currently belongs to the Alibaba group. The site had 500 million registered users by the end of 2013 (Ye et al. 2013). We drew a random sample of 400 e-retailers from a large sample pool (2,826 e-retailers from Taobao.com) in August 2014 from the categories "digital products" and "clothing". We used these industries because of their relatively high return rates and the substantial variation in their e-retailers' delivery package designs. With the links stored in our database, we also collected information about these e-retailers' delivery packages.

We obtained information about each e-retailer's delivery package, including the color, extra gifts, and coupons, by using numerous consumer comments (both text and photo) and retailer product demonstrations. At Taobao.com, it is normal for consumers to upload photo reviews of their received products, including both the outside appearance and its contents. Meanwhile, fierce competition on the platform encourages e-retailers to describe as many product details as they can, including delivery packages, coupons and gifts. Thus, we were able to obtain detailed package information from many e-retailers' websites. We restricted our sample to e-retailers where complete package information was obtainable. We coded "Gift" as 0 if there was no gift included and "Gift" as 1 if the particular e-retailer provided at least one gift. We coded "Color" as 0 if the package box was the typical light brown and 1 if the package box was a different color (e.g., red, blue, pink). We excluded e-retailers that used plastic bags to ship orders. We also coded "Coupon" as 0 if there was no coupon and 1 if there was at least one coupon in the delivery package. Two Chinese students coded the data,

which resulted in a sample of 108 shops for which we could reliably reconstruct the delivery package design (see Table 4-8). We should note that e-retailers do not provide preprinted return labels in China, so we were not able to examine the effect of return costs in this study.

Table 4-8 Descriptive Statistics of the Valid Sample in Study 3

Variable (Unit)	Return rate (%)	Guarantee money (Chinese Yuan)	Service score (Range: 1–5)
Mean	6.33	65,400	4.77
SD	9.29	38,663	.083

	Gift	Coupon	Colorful package
Number of shops with gift/ coupon/ colorful package	64	15	28
Number of shops without gift/ coupon/ colorful package	44	93	80

N = 108.

Our dependent variable is the e-retailer's return rate, which is calculated by the platform using the number of completed returns (calculated by the platform's internal return system) divided by the number of deals (calculated by the platform's trading system) in the observation month. This value is automatically generated by Taobao.com and is displayed on every e-retailers' public rating page. We also considered other related factors as control variables: The leniency of the return policy is measured by "Guarantee Money", which platform operators collect in advance in order to enforce e-retailers' compliance with their promised return policy and offer refunds in case of disputes between e-retailers and consumers. "Review score" is measured by a customer-generated score (ranging from 1 = "very bad" to 5 = "very good") that rates a shop's service quality in the recent half year. The "Industry Return Rate" measures the average return rate across a given industry. The variable "Market" represents the two markets served on Taobao.com: a business-to-consumer market

coded as 1 and a consumer-to-consumer market coded as 0. Based on the range of return rates (0 to 1), we estimated a regression model for fractional response variables with logit distribution (called "fractional logistic model" by STATA) with robust standard errors to avoid a strong influence of outliers. We thus estimated the following equation:

$$\text{Prob}(Return\ Rate_i) = F(\beta' Z_i) \qquad (22)$$

where Return $Rate_i$ stands for the i-th e-retailer's return rate (an aggregate percentage figure, i.e. between 0% and 100%), Z_i represents a vector for all related independent variables ("Gift", "Coupon", "Color", "Gurantee Money", "Review Score", "Industry Return Rate" and "Market"), and β' is the corresponding vector of the coefficients. We use the maximum quasi-likelihood[13] method to estimate the regression model with fractional dependent variable. The odds ratio corresponding to the i_{th} coefficient is $\psi_i = \exp(\beta_i)$ which is approximated with the delta method.

4.5.2 Results and Discussion

However, our estimation results indicate that coupons do not significantly reduce return rates ($\gamma 2$ = .049, n.s.). We surmise that coupons may only be related to consumers' repurchase decisions and not to their return decisions. Another possible explanation is that the prevalence of e-coupons on the Internet makes coupons less useful. In addition, the data show that

[13] The quasi-likelihood function was introduced by Robert Wedderburn in 1974 to describe a function which has similar properties to the log-likelihood function, except that a quasi-likelihood function is not the log-likelihood corresponding to any actual probability distribution. Quasi-likelihood models can be fitted using a straightforward extension of the algorithms used to fit generalized linear models.

improving service and product quality ($\gamma 5 = -2.417$, $p < .1$) is the most effective way for e-retailers to decrease return rates.

Table 4-9 shows the regression results, which reveal that extra gifts and colorful packages have significant and negative impacts on shops' return rates ($\gamma_1 = -.721$, $p < .05$; $\gamma_3 = -.843$, $p < .01$). In other words, a well-designed delivery package can help reduce return intentions.

However, our estimation results indicate that coupons do not significantly reduce return rates ($\gamma_2 = .049$, n.s.). We surmise that coupons may only be related to consumers' repurchase decisions and not to their return decisions. Another possible explanation is that the prevalence of e-coupons on the Internet makes coupons less useful. In addition, the data show that improving service and product quality ($\gamma_5 = -2.417$, $p < .1$) is the most effective way for e-retailers to decrease return rates.

Table 4-9 Results of Regression Model for Fractional Response Variables

Return Rate	Coef.	Odds Ratio	Std. Err. of Coef.	$p > \|z\|$
Intercept	7.759	2341.707	6.217	0.212
Gift	–.721**	0.486	0.317	0.023
Coupon	0.049	1.050	0.327	0.882
Color	–0.843***	0.430	0.242	0.000
Guarantee Money	–0.00001	1.000	0.000	0.167
Review Score	–2.417*	0.089	1.340	0.071
Industry Return Rate	0.282***	1.326	0.058	0.000
Market	0.573	1.775	0.554	0.301
Pseudo R-square	0.0765			

N = 108. *p <.1; **p < .05; ***p < .01.

4.6 Discussion and Implications

Previous research has acknowledged that various factors (e.g., product quality, product demonstration, return policies), working through a cognitive reaction process, can

significantly impact consumer return behavior during the two-stage decision process of online purchasing (Anderson et al. 2009b; Frischmann et al. 2012; Petersen and Kumar 2009; Wood 2001). However, no study has yet examined how consumer return decisions are influenced by a delivery package's appearance (e.g., the color) and contents (e.g., extra gifts, preprinted return labels), or the affective action process. To close these gaps and identify useful packaging strategies, we combined the synergistic properties of two controlled experiments and an observational field study, which together offer robust insights into how the delivery package affects return intentions.

Our studies offer three key findings. The first is that the appearance (i.e., colorful package) and composition (i.e., extra gift) of the delivery package can positively influence consumers' return decisions, at least when said package is a part of the integral purchasing experience. This conclusion was initially supported by the results of a controlled experiment (Study 1), then confirmed by the results of an experiment involving real return behavior (Study 2), and finally supplemented by our analysis of a cross-sectional, transactional data set (Study 3). Second, our results reveal a two-stage reaction process ("high-road" cognitive–affective) when consumers open the delivery package. These cognitive–affective reactions can also more thoroughly explain how the delivery package influences consumers' return behavior. Third and finally, we show that perceived pleasure is the only reaction that can directly influence both consumer satisfaction and return intentions.

4.6.1 Theoretical Contributions

To the best of our knowledge, this paper is the first to examine the entire chain of the two-stage, "high-road" cognitive–affective reaction process in the context of consumer return behavior, which contributes an integrated and holistic perspective to the return behavior research field. In prior studies, researchers have mainly focused on the cognitive process at the neglect of the affective process. In this study, we confirm that perceived utility has an impact on return intention, but more in an indirect way, i.e. via satisfaction (i.e., utilitarian utility). However, the "high-road" affective reaction (i.e., pleasure perceptions), rather than the cognitive reaction, is directly responsible for consumers' return decisions. Moreover, the hedonic utility which with utilitarian utility constitutes an integral part of the perceived utility (Chandon et al. 2000)) has an indirect impact on return intention only via pleasure. Thus, we recommend that future research look beyond consumers' cognitive reactions (e.g., perceived utilitarian utility and expectation gap) and pay closer attention to hedonic utility and the influence of consumers' affective reactions (i.e., emotions).

In addition, this article is one of the first to highlight the importance of the package-opening process. Previous research has explored how a consumer's return decision is influenced by the purchase stage (Petersen and Kumar 2009), a delay period (Bechwati and Siegal 2005), and the final user experience, but has largely ignored the short but critical moment of unpacking. After purchasing, consumers may have a vague impression of their purchase decision, but they are still easily affected when making their return decisions. A well-designed delivery package makes a first and important impression and can ultimately

influence consumers' return decisions. As a result, we believe that models about consumers' post-purchase decision-making should factor in delivery package design.

Furthermore, on the one hand, this work adds color and gift into the long list of influential factors to consumer return behavior. We find that both, color and extra gifts, can significantly decrease consumer's return intentions by adding extra perceived utilities. On the other hand, this study also enriches color and gift research by demonstrating their positive influences on the post-purchase behavior. By integrating literatures on color psychology, promotion, and consumer return behavior, we discovered that colors and extra gifts can influence both purchase and return decisions.

4.6.2 Practical Contributions

Given our key findings, we suggest that e-retailers focus on the delivery package design to lower consumers' return intentions. Choosing a suitable color for delivery packages and offering extra gifts can potentially mitigate return rate problems. With respect to cost efficiency, colorful packages seem more promising because they are cheaper to implement than extra gifts. Specifically, the price of a colorful packaging carton is around 0.136 Euro (1 Yuan, according to Alibaba.com). There is no difference between normal (brown) cartons and colorful ones (even the designed one with words and graphs) when the order quantity exceeds 1,000. As a result, when e-retailers can order more than 1,000 cartons at one time, even a small gift (like the candy used in Study 2) is more expensive than a designed package.

In addition, color can help increase brand familiarity (Labrecque and Milne 2012). For these reasons, we recommend using colorful packages.

Moreover, e-retailers should pay more attention to consumers' emotions (e.g., pleasure perceptions). We used colorful delivery packages and extra gifts to generate positive emotions, but other add-ins (e.g., a note with jokes or wishes) might also be useful. On the flip side, e-retailers might mitigate consumers' negative emotions by offering better post-purchase service.

Granted, the positive effects of package design on emotions (i.e., positive surprise) might wear off over time or with widespread use, but the positive effects on perceived utility are sustainable. Thus, we highly encourage e-retailers to think about a suitable package design strategy.

4.6.3 Limitations

This study comes with several limitations. First, although we carefully chose the products in our experiment—a jersey of the German national football team (in Study 1) and a chocolate bar (in Study 2)—we cannot easily generatilize their effects to all other products. However, because we compared relative differences across the experimental groups, this limitation should not constitute a severe problem for our analyses.

Second, the controlled experiments only included one hue for the colorful package (i.e., blue in Study 1 and 2) and two gifts (Nivea cream in Study 1 and candy in Study 2), which limits their generalizability. For example, the PLS results showed no significant relationship

between a blue package and consumers' perceived hedonic utility, but the result might change for a yellow package. Moreover, even the same hue (i.e., blue) might have different effects in different cultures. Future research could explore these possibilities in greater deail. Still, our study provides significant evidence that a well-designed delivery package can positively influence consumer return intention. E-retailers should experiment with different colors and/or gifts in order to achieve a suitable and memorable delivery package.

Third, the between-subjects experiment design in Studies 1 and 2 led to a limited sample size for each experimental group. In both experimental studies, the number of participants in each group did not exceed 50, which may have influenced the accuracy of our study. Furthermore, the general limitations of a cross-sectional study might have biased the results of our field study (Study 3). Some unobserved factors (such as the e-retailer's brand strength) might influence e-retailers' return rate, but could not be included in our regression models. In fairness, the consistent results of three studies should imbue our conclusions with some confidence. Nevertheless, future studies should further explore our research questions by using panel data analysis. Because panel-econometric approaches can better control for confounding effects, they may be able to corroborate our findings and even deduce the influence of changing one's package strategy.

In sum, this article offers a new research perspective on consumer return behavior, one that found consistent results across Chinese and German data. However, there is a need for further research into how consumer return behavior is influenced by the product category or different color hues and gifts. Scholars could also explore the interaction between different

customer segments and package design, such as how different color hues present variable effects across cultures.

Appendix

Table 4-10 Common Method Bias Assessment

Path coefficients	Affective Model	Cognitive Model	Research Model (w/o marker variable)	CMB test model (with marker variable)
Package design				
Gift -> Arousal	-0.013		0.006	-0.001
Gift -> Hedonic		0.13**	0.13**	0.128**
Gift -> Pleasure	0.097*		0.025	0.026
Gift -> Utilitarian		0.108**	0.107**	0.105**
Color -> Arousal	-0.048		-0.038	-0.035
Color -> Hedonic		0.067	0.066	0.067
Color -> Pleasure	-0.013		-0.051	-0.051
Color -> Utilitarian		0.166***	0.166***	0.167***
Utility				
Utilitarian -> Arousal			-0.062	-0.024
Utilitarian -> Pleasure			0.002	0.003
Utilitarian -> Return intention		0.03	0.016	0.014
Utilitarian -> Satisfaction		0.181**	0.183***	0.184***
Hedonic -> Arousal			-0.033	-0.052
Hedonic -> Pleasure			0.551***	0.551***
Hedonic -> Return intention		-0.047	0.04	0.047
Hedonic -> Satisfaction		0.317***	0.08	0.083
Emotion				
Arousal -> Return intention	0.023		0.027	0.043
Arousal -> Satisfaction	-0.054		-0.035	-0.021
Pleasure -> Return intention	-0.193**		-0.213***	-0.213***
Pleasure -> Satisfaction	0.549***		0.436***	0.434***
Satisfaction				
Satisfaction -> Return intention	-0.369***	-0.47***	-0.382***	-0.386***
Control variables				
Agreeableness -> Return intention	-0.082	-0.095*	-0.081	-0.08*
Agreeableness -> Satisfaction	0.054	0.105*	0.057	0.058
Conscientiousness -> Return intention	0.016	0.016	0.017	0.006

Conscientiousness -> Satisfaction	0.087	0.109	0.086	0.08
Extraversion -> Return intention	0.045	0.055	0.048	0.037
Extraversion -> Satisfaction	-0.087	-0.089	-0.057	-0.064
Neuroticism -> Return intention	0.016	0.002	-0.013	-0.013
Neuroticism -> Satisfaction	0.087	-0.091	-0.048	-0.051
Age -> Return intention	-0.06	-0.062	-0.055	-0.054
Age -> Satisfaction	0.005	0.044	0.023	0.024
Football -> Return intention	-0.036	-0.04	-0.042	-0.043
Football -> Satisfaction	0.028	-0.019	-0.013	-0.014
Gender -> Return intention	-0.006	-0.008	-0.007	-0.017
Gender -> Satisfaction	-0.047	-0.06	-0.056	-0.062
Marker variable				
Openness -> Arousal	-	-	-	0.172**
Openness -> Hedonic	-	-	-	-0.046
Openness -> Pleasure	-	-	-	0.039
Openness -> Return intention	-	-	-	0.062
Openness -> Satisfaction	-	-	-	0.04
Openness -> Utilitarian	-	-	-	-0.037

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

Table 4-11 Discriminant Validity of the Reflective Multi-item Constructs: Construct Correlations and Square Root of AVE

	Gift	Color	Utilitarian	Hedonic	Pleasure	Arousal	Return intention	Satisfaction
Gift	1							
Color	0.012	1						
Utilitarian	0.109	0.167	0.908					
Hedonic	0.131	0.068	0.73	0.927				
Pleasure	0.097	-0.011	0.401	0.554	0.946			
Arousal	-0.006	-0.05	-0.091	-0.08	-0.066	0.805		
Return intention	-0.125	-0.098	-0.22	-0.258	-0.426	0.075	1	
Satisfaction	0.171	0.074	0.429	0.475	0.584	-0.105	-0.499	0.947
Agreeableness	0.048	-0.063	0.055	0.093	0.166	0.008	-0.179	0.145
Conscientiousness	-0.051	0.057	-0.012	0.038	0.079	0.004	-0.034	0.116
Extraversion	0.015	0.065	-0.104	-0.022	-0.026	0.033	0.058	-0.046
Neuroticism	-0.032	-0.072	-0.076	-0.119	-0.155	0.1	0.074	-0.156
Age	0.062	-0.028	-0.05	-0.074	0.039	-0.047	-0.094	0.055
Football	-0.017	0.107	0.165	0.128	0.076	0.038	-0.075	0.073
Gender	-0.006	-0.006	-0.021	-0.073	-0.048	0.197	0.052	-0.098

	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Age	Football	Gender
Gift							
Color							
Utilitarian							
Hedonic							
Pleasure							
Arousal							
Return intention							
Satisfaction							
Agreeableness	0.75						
Conscientiousness	-0.017	0.736					
Extraversion	0.071	0.319	0.862				
Neuroticism	-0.051	-0.238	-0.32	0.845			
Age	0.165	0.166	0.074	-0.119	1		
Football	0.171	0.018	0.255	-0.142	-0.011	1	
Gender	0.007	0.007	0.012	0.145	-0.059	-0.19	1

References

Aftabuzzaman, Md, and Ehsan Mazloumi. 2011. Achieving sustainable urban transport mobility in post peak oil era. *Transport Policy* 18 (5):695-702.

Algesheimer, René, Andreas Herrmann, and Marcus Dimpfel. 2006. Die Wirkung von Brand Communities auf die Markenloyalität—Eine dynamische analyse im automobilmarkt. *Zeitschrift für Betriebswirtschaft* 76 (9):933-958.

Alibaba. 2013. Alibaba group company overview. Alibaba Group. http://www.alibabagroup.com/en/about/overview. Accessed Oct. 21 2013.

Anderson, Eric T, Karsten Hansen, DI Simester, and Lei K Wang. 2009a. How price affects returns: The perceived value and incremental customer effects? *Unpublished Manuscript, Northwestern University, Evanston, IL*.

Anderson, Eric T, Karsten Hansen, and Duncan Simester. 2009b. The option value of returns: Theory and empirical evidence. *Marketing Science* 28 (3):405-423.

Armstrong, Gary, Margaret H Cunningham, and Philip Kotler. 2010. *Principles of Marketing*. Dedining a Market-Oriented Mission. New Jersey: Pearson Education.

Aurier, Philippe, and Guergana Guintcheva. 2014. Using affect–expectations theory to explain the Direction of the impacts of experiential emotions on satisfaction. *Psychology & Marketing* 31 (10): 900-913.

Axsen, Jonn, Dean C Mountain, and Mark Jaccard. 2009. Combining stated and revealed choice research to simulate the neighbor effect: The case of hybrid-electric vehicles. *Resource and Energy Economics* 31 (3):221-238.

Bae, Soonyong, and Taesik Lee. 2011. Product type and consumers' perception of online consumer reviews. *Electronic Markets* 21 (4):255-266.

Bagozzi, Richard P, and Youjae Yi. 1988. On the evaluation of structural equation models. *Journal of the Academy of Marketing Science* 16 (1):74-94.

Banjo, Shelly. 2013. Rampant returns plague e-retailers. The Wall Street Journal. http://online.wsj.com/news/articles/SB10001424052702304773104579270260683155216. Accessed 15th May 2015.

Bechwati, Nada Nasr, and Wendy Schneier Siegal. 2005. The impact of the prechoice process on roduct returns. *Journal of Marketing Research* 42 (3):358-367.

Beggs, S., and S. Cardell. 1981. Assessing the potential demand for electric car. *Journal of Econometrics* 17 (1):1-19.

Berkowitz, Leonard. 1993. *Aggression: Its causes, Consequences, and Control*. Mcgraw-Hill Book Company.

Bettman, James R, Eric J Johnson, and John W Payne. 1990. A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes* 45 (1):111-139.

Blanchard, Dave. 2005. Moving forward in reverse. *Logistics Today* 46 (7):1-8.

Blanchard, Dave. 2007. Supply chains also work in reverse. Industry Week. http://www.industryweek.com/articles/supply_chains_also_work_in_reverse_13947.aspx. Accessed May 15 2014.

Bonifield, Carolyn, and Catherine Cole. 2007. Affective responses to service failure: Anger, regret, and retaliatory versus conciliatory responses. *Marketing Letters* 18 (1-2):85-99.

Bonifield, Carolyn, Catherine Cole, and Randall L. Schultz. 2010. Product returns on the internet: A case of mixed signals? *Journal of Business Research* 63 (9-10):1058-1065.

Boulding, William, and Amna Kirmani. 1993. A consumer-side experimental examination of signaling theory: Do consumers perceive warranties as signals of quality? *Journal of Consumer Research* 20 (1):111-123.

Brazell, Jeff D, Christopher G Diener, Ekaterina Karniouchina, William L Moore, Válerie Séverin, and Pierre-Francois Uldry. 2006. The no-choice option and dual response choice designs. *Marketing Letters* 17 (4):255-268.

Brigl, Silke. 2014. BMW at the Consumer Electronics Show (CES) in Las Vegas 2014. https://www.press.bmwgroup.com/global/pressDetail.html?title=bmw-at-the-consumer-electronics-show-ces-in-las-vegas-2014&outputChannelId=6&id=T0162885EN&left_menu_item=node__5238. Accessed December 2014.

Brownstone, David, David S. Bunch, and Kenneth Train. 2000. Joint mixed logit models of stated and revealed preferences for alternative-fuel. *Transportation Research Part B:*

Methodological 34 (5):315-338.

Buchanan, Leigh, and Andrew O Connell. 2006. A brief history of decision making. *Harvard Business Review* 84 (1):32.

Bunch, David S, Mark Bradley, Thomas F Golob, Ryuichi Kitamura, and Gareth P Occhiuzzo. 1993. Demand for clean-fuel vehicles in California: A discrete-choice stated preference pilot project. *Transportation Research Part A: Policy and Practice* 27 (3):237-253.

Chandon, Pierre, Brian Wansink, and Gilles Laurent. 2000. A benefit congruency framework of sales promotion effectiveness. *Journal of Marketing* 64 (4):65-81.

Chang, En-Chung, Yilin Lv, Ting-Jui Chou, Qingwen He, and Zhuozhao Song. 2014. Now or later: Delay's effects on post-consumption emotions and consumer loyalty. *Journal of Business Research* 67 (7):1368-1375.

Che, Yeon-Koo. 1996. Customer return policies for experience goods. *Journal of Industrial Economics* 44 (1):17-24.

Chebat, Jean-Charles, and Maureen Morrin. 2007. Colors and cultures: exploring the effects of mall décor on consumer perceptions. *Journal of Business Research* 60 (3):189-196.

Chitturi, Ravindra, Rajagopal Raghunathan, and Vijay Mahajan. 2007. Form versus function: How the intensities of specific emotions evoked in functional versus hedonic trade-offs mediate product preferences. *Journal of Marketing Research* 44 (4):702-714.

Choi, Sunmee, and Anna S Mattila. 2008. Perceived controllability and service expectations: Influences on customer reactions following service failure. *Journal of Business Research* 61 (1):24-30.

Chu, Wujin, Eltan Gerstner, and James D. Hess. 1998. Managing dissatisfaction how to decrease customer opportunism by partial refunds. *Journal of Service Research* 1 (2):140-155.

Cooper, Robert G, and Elko J Kleinschmidt. 1987. New products: What separates winners from losers? *Journal of Product Innovation Management* 4 (3):169-184.

Crowley, Ayn E. 1993. The two-dimensional impact of color on shopping. *Marketing Letters* 4 (1):59-69.

Dagsvik, John K, Tom Wennemo, Dag G Wetterwald, and Rolf Aaberge. 2002. Potential demand

for alternative fuel vehicles. *Transportation Research Part B: Methodological* 36 (4):361-384.

Davis, Scott, Michael Hagerty, and Eitan CGerstner. 1998. Return policies and the optimal level of "Hassle". *Journal of Economics and Business* 50 (5):445-460.

De Wilde, Els, Alan DJ Cooke, and Chris Janiszewski. 2008. Attentional contrast during sequential judgments: A source of the number-of-levels effect. *Journal of Marketing Research* 45 (4):437-449.

Desmet, Pierre. 2013. How retailer money-back guarantees influence consumer preferences for retailer versus national brands. *Journal of Business Research* 67 (9):1971-1978. doi:10.1016/j.jbusres.2013.11.001.

DeWitt, Tom, Doan T Nguyen, and Roger Marshall. 2008. Exploring customer loyalty following service recovery the mediating effects of trust and emotions. *Journal of Service Research* 10 (3):269-281.

Dhar, Ravi, and Itamar Simonson. 2003. The Effect of forced choice on choice. *Journal of Marketing Research* 40 (2):146-160.

Dobbs, Richard, Yougang Chen, Gordon Orr, James Manyika, Michael Chui, and Elsie Chang. 2013. China's e-tail revolution. McKinsey Global Institute.

Dutta, Sujay, Abhijit Biswas, and Dhruv Grewal. 2007. Low price signal default: An empirical investigation of its consequences. *Journal of the Academy of Marketing Science* 35 (1):76-88.

Elliot, Andrew J, and Markus A Maier. 2007. Color and psychological functioning. *Current Directions in Psychological Science* 16 (5):250-254.

Erevelles, Sunil, Abhik Roy, and Leslie SC Yip. 2001. The universality of the signal theory for products and services. *Journal of Business Research* 52 (2):175-187.

Eurostat. 2014. Anteil der Online-Käufer in folgenden Altersgruppen in Europa im Jahr 2014. http://de.statista.com/statistik/daten/studie/76300/umfrage/altersverteilung-der-online-kaeufer-in-deutschland/. Accessed 30th Feb. 2016.

Ewing, Gordon O, and Emine Sarigöllü. 1998. Car Fuel-Type Choice under travel demand

management and economic incentives. *Transportation Research Part D: Transport and Environment* 3 (6):429-444.

Ewing, Gordon, and Emine Sarigöllü. 2000. Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment. *Journal of Public Policy and Marketing* 19 (1):106-118.

Fadden, Steven, Patricia May Ververs, and Christopher D Wickens. 1998. Costs and benefits of head-up display use: A meta-analytic approach. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*: SAGE Publications.

Fassnacht, Martin, Christian Stallkamp, and Ludger Rolfes. 2011. Betriebsformen im Automobilhandel – Resultate einer empirischen Untersuchung. *Zeitschrift für Betriebswirtschaft* 81 (11):1181-1203.

Finn, Adam. 2005. Reassessing the foundations of customer delight. *Journal of Service Research* 8 (2):103-116.

Finn, Adam, and Jordan J Louviere. 1992. Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy and Marketing* 11 (2):12-25.

Forbes. 2017. Alibaba's Singles' Day By The Numbers: A Record $25 Billion Haul. https://www.forbes.com/sites/helenwang/2017/11/12/alibabas-singles-day-by-the-numbers-a-record-25-billion-haul/#54a237c91db1. Accessed November 2017.

Frischmann, Tanja, Oliver Hinz, and Bernd Skiera. 2012. Retailers' use of shipping cost strategies: Free shipping or partitioned prices? *International Journal of Electronic Commerce* 16 (3):65-88.

Gärling, Anita, and John Thøgersen. 2001. Marketing of electric vehicles. *Business Strategy and the Environment* 10 (1):53-65.

Garretson, Judith A, and Scot Burton. 2005. The role of spokescharacters as advertisement and package cues in integrated marketing communications. *Journal of Marketing* 69 (4):118-132.

Gelbrich, Katja. 2010. Anger, frustration, and helplessness after service failure: Coping strategies

and effective informational support. *Journal of the Academy of Marketing Science* 38 (5):567-585.

Glantz, Stanton A, and Bryan K Slinker. 1990. *Primer of Applied Regression and Analysis of Variance*. New York: McGraw-Hill Book Company.

Glover, Steven, and Izak Benbasat. 2010. A comprehensive model of perceived risk of e-commerce transactions. *International Journal of Electronic Commerce* 15 (2):47-78.

Green, Paul E, Abba M Krieger, and Yoram Wind. 2001. Thirty years of conjoint analysis: reflections and prospects. *Interfaces* 31 (3):S56-S73.

Hackbarth, André, and Reinhard Madlener. 2013. Consumer preferences for alternative fuel vehicles: A discrete choice analysis. *Transportation Research Part D: Transport and Environment* 25 (1):5-17.

Heilbron, David C. 1994. Zero-altered and other regression models for count data with added zeros. *Biometrical Journal* 36 (5):531-547.

Heilman, Carrie M, Kent Nakamoto, and Ambar G Rao. 2002. Pleasant surprises: Consumer response to unexpected in-store coupons. *Journal of Marketing Research* 39 (2):242-252.

Heiman, Amir, Bruce McWilliams, and David Zilberman. 2001. Demonstrations and money-back guarantees: market mechanisms to reduce uncertainty. *Journal of Business Research* 54 (1):71-84.

Hess, James D., and Glenn E. Mayhew. 1997. Modeling merchandise returns in direct marketing. *Journal of Interactive Marketing* 11 (2):20-35.

Hibbert, Sally A, Maria G Piacentini, and Margaret K Hogg. 2012. Service recovery following dysfunctional consumer participation. *Journal of Consumer Behaviour* 11 (4):329-338.

Hidrue, Michael K., George R. Parsons, Willett Kempton, and Meryl P. Gardner. 2011. Willingness to pay for electric vehicles and their attributes. *Resource and Energy Economics* 33 (3):686-705.

Hinz, Oliver, Christian Schulze, and Carsten Takac. 2014. New product adoption in social networks: Why direction matters? *Journal of Business Research* 67 (1):2836-2844.

Institut für Demoskopie Allensbach 2001. Welche ist Ihre Lieblingsfarbe?

http://de.statista.com/statistik/daten/studie/50/umfrage/eigene-lieblingsfarbe/. Accessed 30th Oct. 2014.

Institut für Demoskopie Köln 2015. How high is the average rate of return of goods in your shop per year? (Germany, 2014). http://www.statista.com/statistics/327314/share-of-returns-in-online-shopping-regarding-fashion-and-consumer-electronics/. Accessed 16th Aug. 2015.

International Council on Clean Transportation. 2013. European Vehicle Market Statistics Pocketbook 2013.

Jacobs, Laurence, Charles Keown, Reginald Worthley, and Kyung-Il Ghymn. 1991. Cross-cultural colour comparisons: global marketers beware! *International Mmarketing Review* 8 (3): 21-33.

Janakiraman, Narayan, and Lisa Ordóñez. 2012. Effect of effort and deadlines on consumer product returns. *Journal of Consumer Psychology* 22 (2) : 260-271.

Jiang, Pingjun, and Bert Rosenbloom. 2005. Customer intention to return online: price perception, attribute-level performance, and satisfaction unfolding over time. *European Journal of Marketing* 39 (1/2):150-174.

Kempton, Willett, and Jasna Tomić. 2005. Vehicle-to-Grid power fundamentals: Calculating capacity and net revenue. *Journal of Power Sources* 144 (1):268-279.

Kim, Junyong, and Brian Wansink. 2012. How retailers' recommendation and return policies alter product evaluations. *Journal of Retailing* 88 (4): 528-541.

Kirmani, Amna, and Akshay R Rao. 2000. No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of Marketing* 64 (2):66-79.

Labrecque, Lauren I, and George R Milne. 2012. Exciting red and competent blue: The importance of color in marketing. *Journal of the Academy of Marketing Science* 40 (5):711-727.

Labrecque, Lauren I, Vanessa M Patrick, and George R Milne. 2013. The marketers' prismatic palette: A review of color research and future directions. *Psychology & Marketing* 30 (2): 187-202.

Ladhari, Riadh. 2007. The effect of consumption emotions on satisfaction and word-of-mouth communications. *Psychology & Marketing* 24 (12):1085-1108.

LeDoux, Joseph E. 1995. Emotion: Clues from the brain. *Annual Review of Psychology* 46: 209-235.

Lee, Julie Anne, Geoffrey Soutar, and Jordan Louviere. 2008. The best–worst scaling approach: An alternative to Schwartz's values survey. *Journal of Personality Assessment* 90 (4):335-347.

Lieven, Theo, Silke Mühlmeier, Sven Henkel, and Johann F. Waller. 2011. Who will buy electric cars? An empirical study in Germany. *Transportation Research Part D: Transport and Environment* 16 (3):236-243.

Liu, Hsin-Hsien, and Hsuan-Yi Chou. 2015. The effects of promotional frames of sales packages on perceived price increases and repurchase intentions. *International Journal of Research in Marketing* 32 (1):23-33.

Louviere, Jordan J, Terry N Flynn, and Richard T Carson. 2010. Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling* 3 (3):57-72.

Louviere, Jordan, Ian Lings, Towhidul Islam, Siegfried Gudergan, and Terry Flynn. 2013. An introduction to the application of (case 1) best–worst scaling in marketing research. *International Journal of Research in Marketing* 30 (3):292-303.

Madden, Thomas J, Kelly Hewett, and Martin S Roth. 2000. Managing images in different cultures: A cross-national study of color meanings and preferences. *Journal of International Marketing* 8 (4):90-107.

Marley, AAJ, Terry N Flynn, and JJ Louviere. 2008. Probabilistic models of set-dependent and attribute-level best–worst choice. *Journal of Mathematical Psychology* 52 (5):281-296.

Mau, Paulus, Jimena Eyzaguirre, Mark Jaccard, Colleen Collins-Dodd, and Kenneth Tiedemann. 2008. The 'neighbor effect': Simulating dynamics in consumer preferences for new vehicle technologies. *Ecological Economics* 68 (1):504-516.

Mavlanova, Tamilla, and Raquel Benbunan-Fich. 2010. Counterfeit products on the internet: The role of seller-level and product-level information. *International Journal of Electronic*

Commerce 15 (2):79-104.

Mazaheri, Ebrahim, Marie-Odile Richard, and Michel Laroche. 2012. The role of emotions in online consumer behavior: A comparison of search, experience, and credence services. *Journal of Services Marketing* 26 (7):535-550.

Mazaheri, Ebrahim, Marie Odile Richard, Michel Laroche, and Linda C Ueltschy. 2014. The influence of culture, emotions, intangibility, and atmospheric cues on online behavior. *Journal of Business Research* 67 (3):253-259.

McClelland, James L. 1988. Connectionist models and psychological evidence. *Journal of Memory and Language* 27 (2):107-123.

McFadden, Daniel. 1974. *Conditional Logit Analysis of Qualitative Choice Behavior*. New York.

McFadden, Daniel. 2001. Disaggregate behavioral travel demand's RUM Side: A 30-year retrospective. In *Travel Behaviour Research: The Leading Edge*, 17-64. Amsterdam: Pergamon.

McLachlan, Geoffrey, and David Peel. 2004. *Finite Mixture Models*. New York: John Wiley & Sons.

Mehrabian, Albert, and James A Russell. 1974. *An Approach to Environmental Psychology*. Cambridge: MIT Press.

Mehta, Ravi, and Rui Juliet Zhu. 2009. Blue or red? Exploring the effect of color on cognitive task performances. *Science* 323 (5918):1226-1229.

Meyers-Levy, Joan, and Laura A Peracchio. 1995. Understanding the effects of color: How the correspondence between available and required resources affects attitudes. *Journal of Consumer Research* 22 (3):121-138.

Minnema, Alec, Tammo HA Bijmolt, Sonja Gensler, and Thorsten Wiesel. 2016. To keep or not to keep: Effects of online customer reviews on product returns. *Journal of Retailing* 92 (3):253-267.

Moorthy, Sridhar, and Kannan Srinivasan. 1995. Signaling quality with a money-back guarantee: The role of transaction costs. *Marketing Science* 14 (4):442-466.

Mueller Loose, Simone, and Larry Lockshin. 2013. Testing the robustness of best worst scaling

for cross-national segmentation with different numbers of choice sets. *Food Quality and Preference* 27 (2):230-242.

Mwalili, S. M., E. Lesaffre, and D. Declerck. 2008. The zero-inflated negative binomial regression model with correction for misclassification: An example in caries research. *Stat Methods Med Res* 17 (2):123-139.

Nunnally, Jum C, Ira H Bernstein, and Jos MF ten Berge. 1967. *Psychometric Theory*. New York: McGraw-Hill Book Company.

Oliver, Richard L. 1993. Cognitive, affective, and attribute bases of the satisfaction response. *Journal of Consumer Research*: 418-430.

Olivier, Jos G.J., Greet Janssens-Maenhout, Marilena Muntean, and Jeroen A.H.W. Peters. 2013. Trends in Global Co2 Emmissions: 2013 Report: PBL Netherlands Environmental Assessment Agency.

Padmanabhan, V, and Ivan PL Png. 1997. Manufacturer's return policies and retail competition. *Marketing Science* 16 (1):81-94.

Palensky, Peter, and Dietmar Dietrich. 2011. Demand side management: Demand response, intelligent energy dystems and smart loads. *Industrial Informatics, IEEE Transactions on* 7 (3):381-388.

Paromtchik, Igor, and Christian Laugier. 1998. Automatic parallel parking and returning to traffic maneuvers. In *Video Proc. of the IEEE Int. Conf. on Robotics and Automation*: Citeseer.

Peard, Estelle. 2013. Electric Cars Slow to Gain Traction in Germany. http://phys.org/news/2013-05-electric-cars-gain-traction-germany.html. Accessed July 2014.

Petersen, J Andrew, and V Kumar. 2009. Are product returns a necessary evil? Antecedents and consequences. *Journal of Marketing* 73 (3):35-51.

Pizzutti, Cristiane, and Daniel Fernandes. 2010. Effect of recovery efforts on consumer trust and loyalty in e-tail: A contingency model. *International Journal of Electronic Commerce* 14 (4):127-160.

Podsakoff, Philip M, Scott B MacKenzie, Jeong-Yeon Lee, and Nathan P Podsakoff. 2003.

Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology* 88 (5):879-903.

Posselt, Thorsten, Eitan Gerstner, and Dubravko Radic. 2008. Rating e-tailers' money-back guarantees. *Journal of Service Research* 10 (3):207-219.

Potoglou, Dimitris, and Pavlos S Kanaroglou. 2007. Household demand and willingness to pay for clean vehicles. *Transportation Research Part D: Transport and Environment* 12 (4):264-274.

Puccinelli, Nancy M, Rajesh Chandrashekaran, Dhruv Grewal, and Rajneesh Suri. 2013. Are men seduced by red? The effect of red versus black prices on price perceptions. *Journal of Retailing* 89 (2):115-125.

Purohit, Devavrat, and Joydeep Srivastava. 2001. Effect of manufacturer reputation, retailer reputation, and product warranty on consumer judgments of product quality: A cue diagnosticity framework. *Journal of Consumer Psychology* 10 (3):123-134.

Rammstedt, Beatrice, and Oliver P John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five inventory in english and german. *Journal of research in Personality* 41 (1):203-212.

Rao, Vithala R. 2014. *Applied Conjoint Analysis*. Theory and design of conjoint studies (ratings based methods). New York: Springer.

Reichheld, Frederick F. 2003. The one number you need to grow. *Harvard Business Review* 81 (12):46-55.

Ringle, Christian M, Sven Wende, and Jan-Michael Becker. 2015. SmartPLS 3. *Boenningstedt: SmartPLS GmbH,* http://www. smartpls. com.

Roggeveen, Anne L., Ronald C. Goodstein, and Dhruv Grewal. 2014. Improving the effect of guarantees: The role of a retailer's reputation. *Journal of Retailing* 90 (1):27-39.

Rönkkö, Mikko, and Jukka Ylitalo. 2011. PLS marker variable approach to diagnosing and controlling for method variance. Paper presented at the Proceedings of the International Conference on Information Systems Shanghai.

Sahoo, Nachiketa, Chrysanthos Dellarocas, and Shuba Srinivasan. 2015. The impact of online

product reviews on product returns. In *Boston U. School of Management Research Paper No. 2491276*.

Samson, Alain. 2006. Understanding the buzz that matters: Negative vs positive word of mouth. *International Journal of Market Research* 48 (6):647-657.

Scherer, Klaus R, Angela Schorr, and Tom Johnstone. 2001. *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Schlereth, Christian, Christine Eckert, René Schaaf, and Bernd Skiera. 2014. Measurement of preferences with self-explicated approaches: A classification and merge of trade-off-and non-trade-off-based evaluation types. *European Journal of Operational Research* 238 (1):185-198.

Schlereth, Christian, Bernd Skiera, and Agnieszka Wolk. 2011. Measuring consumers' preferences for metered pricing of services. *Journal of Service Research* 11 (4):443-459.

Schoefer, Klaus. 2010. Cultural moderation in the formation of recovery satisfaction judgments: A cognitive-affective perspective. *Journal of Service Research* 13 (1):52-66.

Shepherd, Simon, Peter Bonsall, and Gillian Harrison. 2012. Factors affecting future demand for electric vehicles: A model based study. *Transport Policy* 20 (1):62-74.

Shiv, Baba, and Alexander Fedorikhin. 1999. Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research* 26 (3):278-292.

Shulman, Jeffrey D, Anne T Coughlan, and R Canan Savaskan. 2009. Optimal restocking fees and information provision in an integrated demand-supply model of product returns. *Manufacturing & Service Operations Management* 11 (4):577-594.

Simon, Herbert A. 1972. Theories of bounded rationality. *Decision and Organization* 1 (1):161-176.

Simons, Jeffrey S., Dan J. Neal, and Raluca M. Gaher. 2006. Risk for marijuana-related problems among college students: An application of zero-inflated negative binomial regression. *The American Journal of Drug and Alcohol Abuse* 32 (1):41-53.

Smith, Amy K, and Ruth N Bolton. 2002. The effect of customers' emotional responses to service

failures on their recovery effort evaluations and satisfaction judgments. *Journal of the Academy of Marketing Science* 30 (1):5-23.

Son, Jai-Yeol, and Izak Benbasat. 2007. Organizational buyers' adoption and use of B2B electronic marketplaces: efficiency-and legitimacy-oriented perspectives. *Journal of Management Information Systems* 24 (1):55-99.

statista. 2017. E-commerce share of total global retail sales from 2015 to 2021. https://www.statista.com/statistics/534123/e-commerce-share-of-retail-sales-worldwide/. Accessed November 2017.

Street, Deborah J, and Leonie Burgess. 2007. *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. New Jersey: John Wiley and Sons.

Su, Xuanming, and Fuqiang Zhang. 2009. On the value of commitment and availability guarantees when selling to strategic consumers. *Management Science* 55 (5):713-726.

Sum, Chee-Chuong, Yang-Sang Lee, Julie M Hays, and Arthur V Hill. 2002. Modeling the effects of a service guarantee on perceived service quality using alternating conditional Expectations (ACE). *Decision Sciences* 33 (3):347-384.

Suwelack, Thomas, Jens Hogreve, and Wayne D. Hoyer. 2011. Understanding money-back guarantees: Cognitive, affective, and behavioral outcomes. *Journal of Retailing* 87 (4): 462-478.

Swait, Joffre, and Rick L Andrews. 2003. Enriching scanner panel models with choice experiments. *Marketing Science* 22 (4):442-460.

Teo, Hock-Hai, Kwok Kee Wei, and Izak Benbasat. 2003. Predicting intention to adopt interorganizational linkages: An institutional perspective. *MIS Quarterly* 27 (1):19-49.

Thaler, Richard. 1985. Mental accounting and consumer choice. *Marketing Science* 4 (3):199-214.

Thomason, Emma. 2013. Online retailers go hi-tech to size up shoppers and cut returns. Reuters. http://www.reuters.com/article/2013/09/27/net-us-retail-online-returns-idUSBRE98Q0GS20130927. Accessed May 15 2014.

Thurstone, Louis L. 1927. A law of comparative judgment. *Psychological Review* 34 (4):273.

Tie, Siang Fui, and Chee Wei Tan. 2013. A review of energy sources and energy management

system in electric vehicles. *Renewable and Sustainable Energy Reviews* 20 (1):82-102.

Timmers, Paul. 1998. Business models for electronic markets. *Electronic Markets* 8 (2):3-8.

Train, Kenneth. 2009. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Voss, Kevin E, Eric R Spangenberg, and Bianca Grohmann. 2003. Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of Marketing Research* 40 (3): 310-320.

Wesseling, J. H., E. M. M. I. Niesten, J. Faber, and M. P. Hekkert. 2013. Business strategies of incumbents in the market for electric vehicles: Opportunities and incentives for sustainable innovation. In *Business Strategy and the Environment*.

Wood, Stacy L. 2001. Remote purchase environments: The influence of return policy leniency on two-stage decision processed. *Journal of Marketing Research* 32 (2):157-169.

Ye, Qiang, Min Xu, Melody Kiang, Weifang Wu, and Fangfang Sun. 2013. In-depth analysis of the seller reputation and price premium relationship: A comparison between eBay US and Taobao China. *Journal of Electronic Commerce Research* 14 (1):1-10.

Zeltner, Philip M. 1975. *John Dewey's Aesthetic Philosophy*. Amsterdam: John Benjamins Publishing Company.

Zhang, Xianfeng, Jifeng Luo, and Qi Li. 2012. Do different reputation systems provide consistent signals of seller quality: a canonical correlation investigation of Chinese C2C marketplaces. *Electronic Markets* 22 (3):155-168.

Zhou, Wenyan, and Oliver Hinz. 2016. Determining profit-optimizing return policies–a two-step approach on data from Taobao.com. *Electronic Markets* 26 (2):103-114.

Zhuo, Jun, June Wei, LaiC Liu, KaiS Koong, and Shengtao Miao. 2013. An examination of the determinants of service quality in the Chinese express industry. *Electronic Markets* 23 (2):163-172.

Resume

Wenyan Zhou

Contact Information

Address: Chemnitzer Str. 47, 40627 Düsseldorf

Tel: +49 1726086932

Email: zhouwenyan615@gmail.com

Experience

04/2017- Present
User Researcher
trivago N.V., Düsseldorf, Germany

09/2013 - Present
Ph.D. Student
Electronic Markets in the department of Information System
Technische Universität Darmstadt, Darmstadt, Germany

09/2010 -06/2013
Master in E-Commerce
Zhongnan University of Economics and Law, Wuhan, China

09/2006 - 06/2010
Bachelor in E-Commerce
Zhongnan University of Economics and Law, Wuhan, China

09/2006 - 06/2010
Bachelor in Accounting
Wuhan University, Wuhan, China

Thesis Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisor Prof. Dr. Oliver Hinz for the continuous support of my Ph.D study and research, for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study.

Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Dr. Alexander Benlian, Prof. Dr. Alexander Kock, and Dr. Simon Emde, for their encouragement, insightful comments, and hard questions.

My sincere thanks also go to Herr. Partick Felka, Frau. Cristina Mihale-Wilson and other colleges in our E-market department for offering me incredible help and supports during my life in German and also my study.

A lot of thinks should also work for the China Scholarship Council for supporting my academic career in these four years.

Last but not the least, I would like to thank my family: my parents Chen,Jufen and Zhou, Yongpin, for giving birth to me at the first place and supporting me spiritually throughout my life.