ESSAYS ON FORECASTING DEMAND AND PREFERENCES FOR CARS IN EMERGING MARKETS: THE CASE OF CHINA

Lixian Qian

B.Sc., Master of Management in Management Science

Fudan University, China

A Thesis Submitted in Part Fulfilment of the Requirements for

the Degree of Doctor of Philosophy

Department of Management Science

Lancaster University

January 2012

ACKNOWLEDGEMENTS

First and foremost, I would like to express my utmost gratitude to my supervisor, Dr. Didier Soopramanien, whose constructive advices and thoughtful suggestions have been proved invaluable for my research. It would be impossible for me to finish this thesis without Didier's consistent guidance and supervision throughout the past three and half years. My appreciation also goes to my second supervisor in the first year, Professor Ruud Teunter.

Second, I would like to thank the faculty members in the Department of Management Science, including Professor Robert Fildes, Professor David Brown, Professor Stein Wallace, Dr. Dave Worthington, Dr. Zhan Pang, and Dr. Sven Crone, for their insightful comments and suggestions for my research and teaching. I am also thankful to the secretaries in the department, including Gay Bentinck, Christine Fletcher and Lindsay Newby for their assistance and help during my PhD journey.

Last but not least, the biggest thanks will have to go to my parents and in-laws for their constant encouragement and emotional support for my study. My wife, Xian, is deserved my deepest gratitude for her unconditional love and enormous support. My daughter, Xuqing, brings us much happy time and makes my doctoral study a more enjoyable journey. I dedicate this thesis to all of them.

DECLARATON

The thesis is my own work and has not been submitted in substantially the same form for the award of a higher degree elsewhere.

ESSAYS ON FORECASTING DEMAND AND PREFERENCES FOR CARS IN EMERGING MARKETS: THE CASE OF CHINA

ABSTRACT

Lixian Qian

The emerging markets (EMs) have been increasingly important in the global economy, especially during the recession. These markets have different characteristics from the developed markets such as high level of market heterogeneity (Burgess & Steenkamp, 2006; Sheth, 2011). This thesis explores how to forecast the demand for cars in a market context that has been experiencing significant and fast growth. Car sales in 2002 were only 1.25 million units in China, while the figure increased to 11.27 million by 2010. Research on car market demand in developed economies is well established, but little attention has been paid to the emerging car markets and the challenges that researchers face when they have to predict the demand or preferences for cars in the EMs. By using the Chinese car market as the market context, this thesis explores how to tackle specific problems associated with forecasting the demand for cars in an emerging market.

The thesis contributes to the literature in the following ways. We apply some of the well-known techniques that have been applied in other domains and assess how they fare in predicting the demand and preferences for cars in a new market context. We also take into account that preferences and the way in which consumers make choices in some markets may require a different methodological approach. We demonstrate

the importance of understanding local consumer behaviour when it comes to not only collecting the data but we also show that this may mean that we have to modify or reject some of the approaches that have been used in more mature markets. The thesis also proposes novel modelling approaches that are inspired by the specific problems of predicting car demand in China, but these proposed methods can also be replicated and tested for other products in other new emerging market economies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

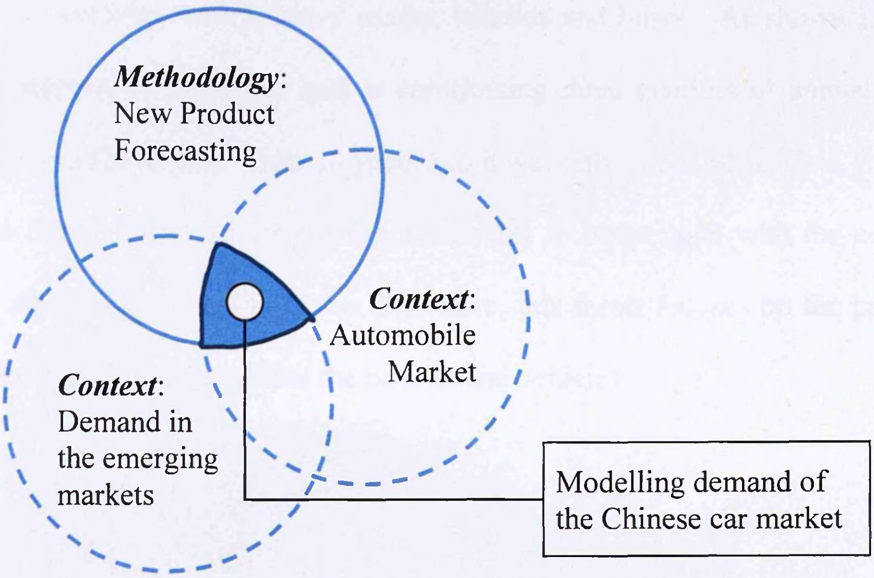
1.1 Research Background

The emerging markets (EMs, hereafter) have an important role in contributing to global economic growth especially in the context of the recent financial crisis. The importance of the EMs is particularly prominent when major developed economies have been suffering weak demand caused mainly by the economic recession during the recent global financial crisis of 2007-2012. Before the crisis, the EMs contributed 45% of global economic growth from 2000 to 2006, while their contribution rate sharply rose to more than 80% in 2007 and 2008 (O'Neill & Stupnytska, 2009). It is further predicted that the four leading EMs, namely Brazil, Russia, India and China (BRICs), will be as big as a group of seven highly industrialised economies (G7) by 2032. Research has shown that emerging markets differ from the more developed markets on the following key attributes: market heterogeneity, socioeconomic status, culture and regulation (Alden, et al., 2006; Burgess & Steenkamp, 2006; Sheth, 2011). Given the increasing significance of the EMs and their different context specific characteristics, it is crucial for both manufacturers and governments to use marketing and forecasting models to forecast demand and better understand local consumer behaviour in the EMs, so that better decisions can be made to either effectively fulfil the demand or further stimulate the growth. However, there is little attention devoted to forecasting and marketing problems in emerging markets. The research context of the thesis is an emerging market, China. We look at the marketing modelling challenges that arise in such a market and illustrate our propositions by using the case of car demand. The existing literature, particularly about forecasting and modelling car market demand, has paid little attention to the emerging markets, so it is important that we critically explore whether such existing approaches which have been used in other markets can be applied in the context of an emerging market. This thesis aims to fill in this gap and shed some light on how we can better forecast demand and understand local consumer behaviour in the EMs through appropriately accounting for the contextual characteristics and the local consumer preferences. By doing this, we also propose new modelling approaches that suit the context and the marketing problems and challenges.

Among the various economic sectors, the automobile industry is one of the most influential sectors with significant impacts on the whole society. It not only contributes a large amount of revenue to national economy but also has the potential to transform social behaviour and consumer lifestyles. Thus many countries consider the automobile industry as one of the economic pillars. When studying the automobile industry, there are many interesting topics, such as new product development and introduction strategies in the EMs, the innovations in car marketing practice, and the design and implementation of incentive policies among governments, car markets and consumers. All of them have important implications in the current context of promoting the post recession market demand. Furthermore, the transport sector, which is largely based on automobiles, is faced with important issues about both oil demand and environmental issues. The transport sector accounts for 97% of global oil demand increases (International Energy Agency, 2009). The world oil consumption ratio in the transport sector increased from 45.3% in 1973 to 61.40% in 2008 (International Energy Agency, 2010a). At the same time, car use is a significant contributor of air pollution (Fenger, 1999). Currently, the transport sector accounts for 23% of global CO_2 emissions, which are estimated to rise to 32% by 2035 (International Energy Agency, 2010b). Moreover, the emerging car markets led by the BRIC economies are expected to be the main source of global car sales growth in the next decade and China alone will account for 42% of this increase (Burgstaller, et al., 2009). Such huge growth potential has strengthened public concerns about security of oil supplies and environmental impacts. The International Energy Agency (2009, 2010b) warns that China will account for more than half of the global increase of CO_2 emissions by 2030 if it continues on current energy path and China will also contribute to almost half of global oil demand growth in next 2.5 decades, mainly driven by the growth of local car market. Therefore, given the important implications for the automobile industry and more importantly for the sustainable development of the whole society, this thesis is positioned in the larger context of modelling market demand and local consumer behaviour in the emerging car market, as illustrated by the overlapped areas of two dashed circles in Figure 1-1.

Considering the emerging car markets such as China's, as most consumers have no prior experience of using cars, the analogy here is that the car is a new product concept. Therefore, this thesis is mainly based on the methodology of new product forecasting as shown with the solid line circle in Figure 1-1. We elaborate further on the specific research questions in section 1.3.

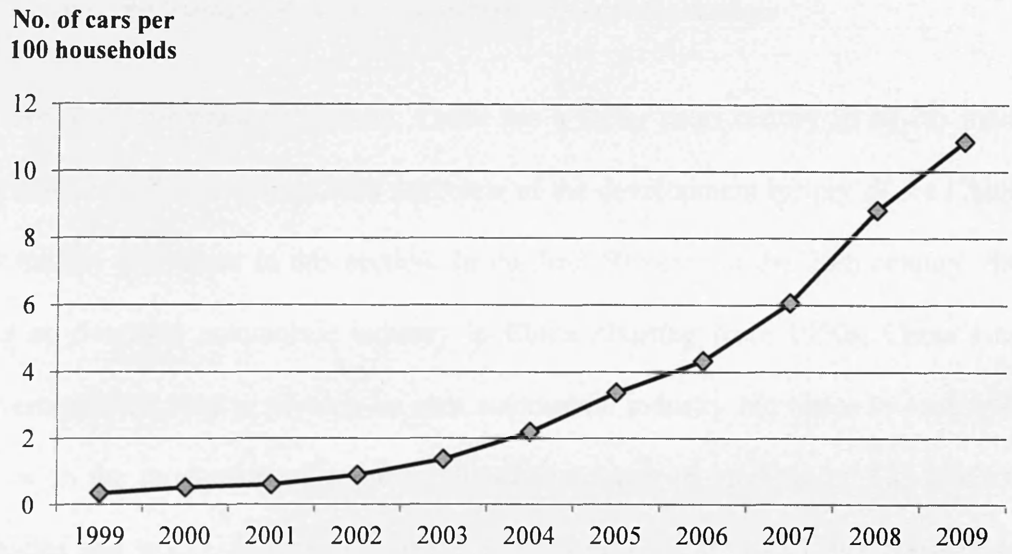
Figure 1-1: Position of the thesis - market context and methodology



1.2 An Overview of the Chinese Car Market

Let us now elaborate on the context of the Chinese car market. On the one hand, the Chinese car market is a typical emerging market with the short development history and low household car ownership level. Figure 1-2 shows the urban household car ownership level in China from 1999 to 2009. Every 100 Chinese urban households only owned 10.89 cars by the end of 2009 in spite of nearly 10 years' fast growth, which is much lower than the world average level. On the other hand, China has become one of the most important automobile markets in the world. The fast development of the Chinese automobile market is mainly driven by the massive growth in the passenger car market since 2002. Similar to other markets, the category of the passenger cars in China is defined as the motor vehicles with at least four wheels, used for people-carrying purpose and comprising no more than 9 seats including the driver's seat, while the commercial vehicles mainly include light commercial vehicles (LCV), heavy trucks, coaches and buses[1]. As shown in Figure 1-3, the segment of passenger cars is contributing three quarters of annual vehicle productions in China now, while 10 years ago it was only about 30%. Thus, it is more important for car manufacturers and governments to better cope with the consumer demand for passenger cars in China. Therefore, this thesis focuses on the passenger car market in China and excludes the commercial vehicles.

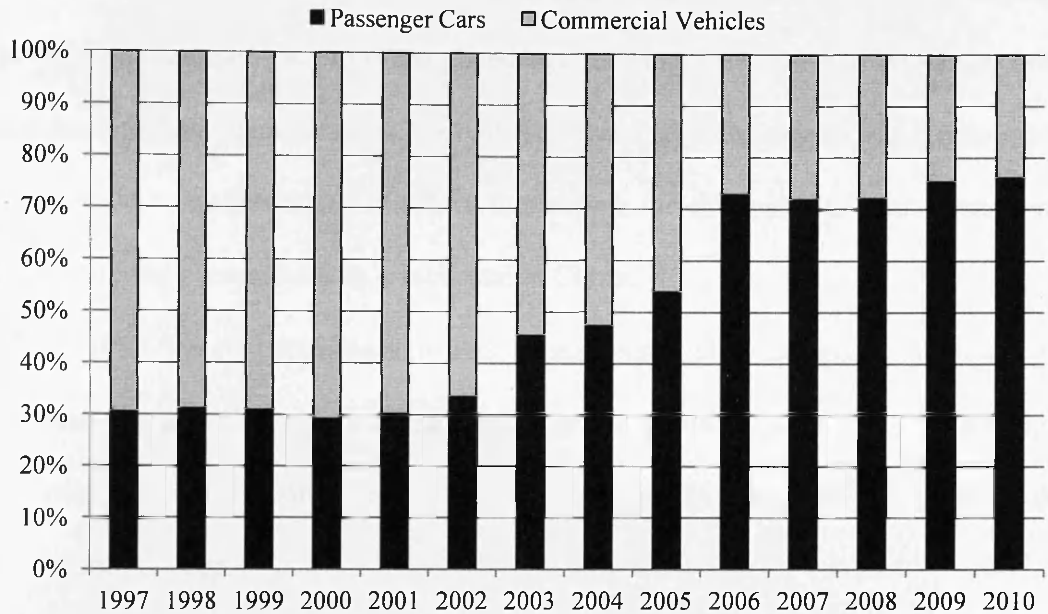
Figure 1-2: Household car ownership level in China (1999-2009)



Note: Data is for urban and township households only.
Data source: China Statistical Yearbooks (2003-2010)

[1] See detailed vehicle type definitions on the website of the International Organization of Motor Vehicle Manufacturers (http://oica.net/wp-content/uploads/stats-definition.pdf).

Figure 1-3: Market shares of passenger cars and commercial vehicles in China (1997-2010)



Data source: the International Organization of Motor Vehicle Manufacturers

As we have discussed previously, China has a fairly short history in its car market development. We provide a brief overview of the development history of the Chinese car market as follows in this section. In the first 50 years of the 20th century, there was no domestic automobile industry in China. Starting from 1950s, China's new government decided to develop its own automobile industry but chose to exclusively focus on the production of commercial vehicles, such as medium or heavy loading vehicles and trucks, to meet the demand for large scale infrastructure construction. The development of Chinese passenger car manufacturing did not start until the establishment of the reform and opening policy in late 1970s. Here we briefly review the development history of the Chinese car market over the past three decades. For more historical reviews of the Chinese automobile industry and car market development, please refer to Chin (2010) and Holweg, Luo, & Oliver (2009).

1. 1980s and early 1990s

In the early 1980s, in order to initiate the domestic passenger car market, the Chinese government changed its policy to encourage the domestic firms to cooperate with multinational car manufacturers, so that the advanced technologies, equipments and funds could be introduced to China. In this period, the first batch of joint venture car manufacturing companies was established in China.

- In 1983, the first Sino-foreign joint venture automobile company, Beijing Jeep, was established between American Motors Corporation (AMC, which was taken over by Chrysler later) and Beijing Automotive Works (BAW) to produce the Cherokee Jeep.
- In 1985, the first Sino-foreign joint venture car manufacturing company, Shanghai Volkswagen, was formed between Shanghai Automobile Industry Corporation (SAIC) and Volkswagen to produce the Santana sedan. In the same year, Guangzhou Peugeot Automobile Company was setup between Guangzhou Automotive Works and Peugeot Corporation to produce the Peugeot 504 light truck and the Peugeot 505 sedan.
- In 1991, Volkswagen's second joint venture company in China was established in Changchun with First Automotive Works (FAW). One year later, Dongfeng Automobile Corporation setup its first joint venture company, Dongfeng Citroen (Shenlong Automobile), with Citroen from France.

Car market demand during this period primarily came from various governmental organisations, public sectors and commercial companies. The Chinese households in 1980s couldn't afford the private cars due to their fairly low income. Instead, the Chinese urban households at that time were mainly intended to adopt home appliances, typically including televisions, refrigerators and washing machines, which were known as three "must have" items for the marriages in 1980s in China[2].

2. Mid 1990s to 2000

In mid 1990s, the Chinese government started to transform its car market policy from supporting the car demand in public or business sectors to encouraging private car adoption. The *Automotive Industry Policy* published in July 1994 is regarded as a watershed policy (Chin, 2010), because it not only established the automotive industry to be one of the "pillar industries" of the national economy in China, but also explicitly encouraged the household car ownership and identified the passenger cars to be one of the development priorities thereafter.

During this period, several new car manufacturing companies were established in China, who became the important market players in the new century. Shanghai General Motors and Guangzhou Honda opened in 1997 and 1998 respectively as the new members of Sino-foreign joint venture car makers. At the same time, a few domestic enterprises, such as Chery and Geely, entered the car manufacturing field in late 1990s to produce cars with their own brands.

In spite of the improvement of government policy and the establishment of more car manufacturers, many problems still existed in the Chinese car market, which restricted the private car market demand in this period. First of all, various taxes and fees above the car price overcharged potential consumers. Local protectionism was another issue at that time. Regional governments set up entry barriers for the cars not produced in

[2] "Chinese wedding costs spiraling", 14/10/2011, China News Service (CNS), http://www.ecns.cn/life-style/2011/10-14/3014.shtml

their own regions through imposing additional fees or unfair regulations. Another critical problem in this period is the limited variety of products. Santana (Shanghai Volkswagen), Jetta (FAW Volkswagen) and Fukang (Dongfeng Citroen) dominated the car market in 1990s, and thus they were popularly called "Old Three" car models in China[3].

3. First decade of the new century

The first significant event in the new century was the China's entry to the WTO in January 2002, from which the domestic car market in China was expected to suffer severe impacts from the imported cars. In order to ensure the stable development of the car market, the Chinese government implemented a number of practical policies in advance, including:

- In February 2001, the Ministry of Public Security revised the Motor Vehicle Registration Procedure to facilitate the development of the car rental business and the used car market.
- In May 2001, the State Planning Commission decided to remove the price control on domestically produced cars, which provided the car manufacturers the right to independently determine their products' prices in the market.
- In August 2001, the State Economic and Trade Commission, Ministry of Finance, State Administration of Taxation and State Environmental Protection Administration jointly issued a notification to reduce 30% of consumption tax for the cars that meet the emission standard of "Euro II".

[3] "Breaking the rules: Moments of truth in the car buying process", AC Nielson China, 2007, http://jp.nielsen.com/industry/Breakingtherule_China.pdf.pdf

- In November 2001, the State Administration of Taxation adjusted the base prices to calculate the vehicle purchase tax, which indirectly reduced the expenditure of the consumers when buying new cars.

All these supporting policies provided a solid base for the massive domestic car market growth in following years without negative impacts due to the entry into the WTO. According to the statistics from the China Association of Automobile Manufacturers (CAAM), both production and sales of the domestic-made cars exceeded 1.2 million units in 2002, which increased by 59.18% and 62.43% respectively in comparison to in the previous year. It was the highest car market growth rate achieved since 1993 in China, which demonstrated the start of large-scale household car adoption in China. Thus, the year of 2002 was called the first year of the car popularisation era in China.

In the following years, the Chinese passenger car market continued this fast expanding trend with more than 15% annual growth rate. As we have shown in Figure 1-2, the household car ownership level starts its "exponential" increase since 2002. It shows that every 100 Chinese urban households owned less than 1 car with little growth until 2002, but the rate increased faster thereafter and reached over 10 cars per 100 households by 2009.

In 2008, the Chinese car market was seriously impacted by the global financial crisis and its growth rate dropped to below 10%. In order to ensure the stable development of the car market, the State Council of China approved the *Automobile Industry*

Adjustment and Revitalisation Plan in early 2009[4]. The major point is to further boost the private automobile market. One critical stimulus is to halve the purchasing tax in 2009 for passenger vehicles with no more than 1.6L engine size and provide the allowance to rural consumers to buy the cross-over vans or light trucks. The incentive policy demonstrated the immediate effect in the market where the domestic passenger car sales in 2009 achieved an outstanding growth rate of 52.93% compared to in 2008, which helped China to become the world largest automobile market in terms of annual sales by the end of that year.

At the same time, the Chinese government started to promote the green car market to address the issues of oil shortages and environmental impacts due to the conventional petrol cars. In January 2009, a public sector-focused "Ten cities and Thousand vehicles" program was initiated in China, which aims to select 10 cities each year and introduces 1000 hybrid, electric or hydrogen fuel cell vehicles for taxis, bus and other public services in each city (Huo, et al., 2010). In June 2010, the Chinese government announced a pilot subsidy policy for household green car buyers in 5 selected cities (Shanghai, Changchun, Shenzhen, Hangzhou and Hefei). This policy has been designed for the period between 2010 and 2012 and mainly supports plug-in hybrid and electric cars. The governmental subsidies for each hybrid and electric car can go up to RMB 50,000 and RMB 60,000 respectively, depending on the different battery capacity.

[4] "Stimulus plan gives auto sector a shot in the arm", 37/03/2009, China Economic Net, http://en.ce.cn/Insight/200903/27/t20090327_18635392.shtml

1.3 Research Objectives and Contributions

With the development of the Chinese car market, different types of marketing problems have gradually emerged, which have significantly shaped the research questions. Figure 1-4 shows the main framework of this thesis. In the context of the Chinese car market, we discuss here how we develop this thesis framework to effectively address four different but interrelated research problems or challenges (shown as the small rounded rectangles under the main research question) by employing different approaches.

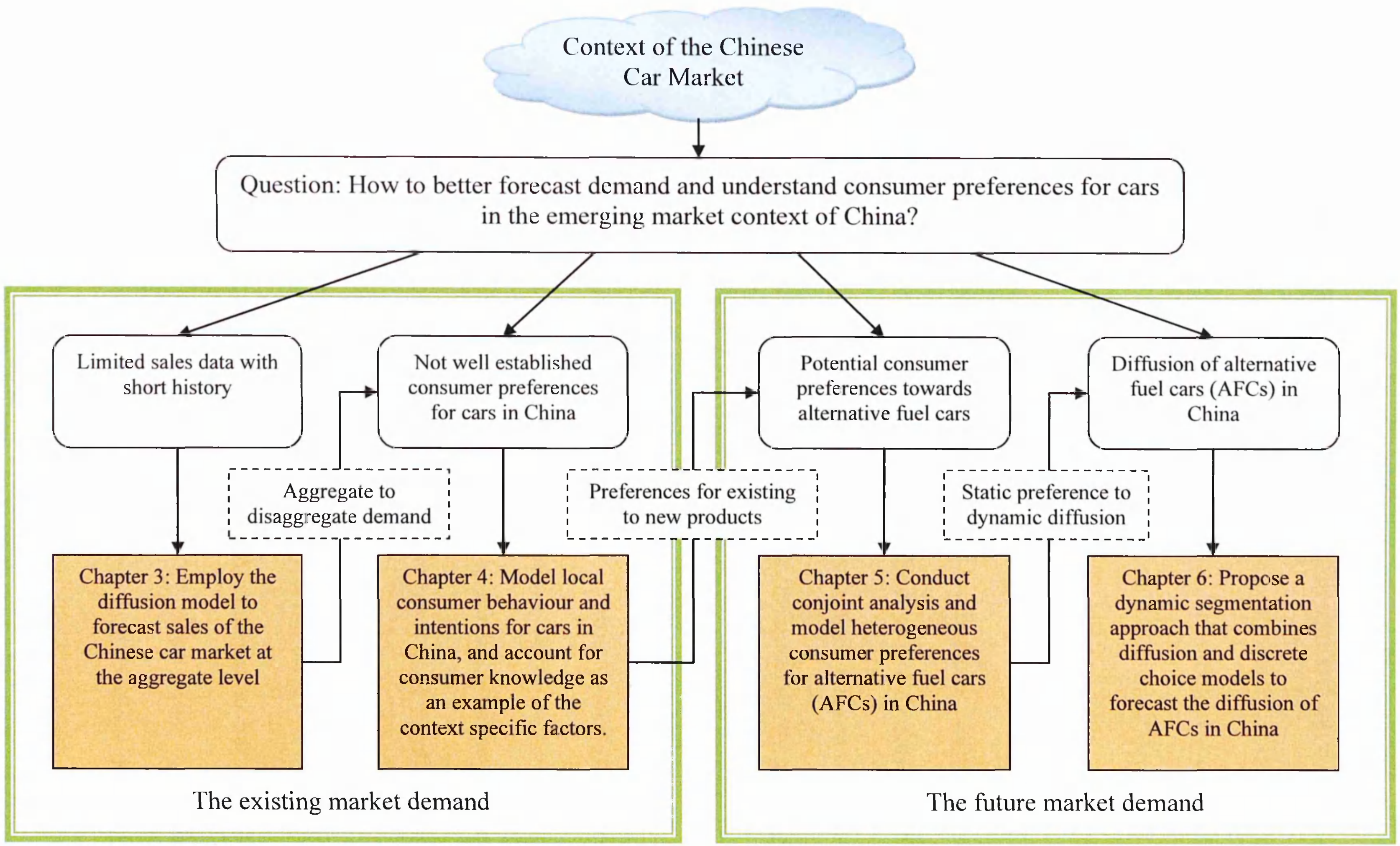
This thesis tackles two broad research questions as illustrated with two large rectangles in Figure 1-4: the current and future market demand respectively in the Chinese car market. Regarding the existing demand in China, the world car market is a typical sector that has been deeply affected by the financial crisis during the past three years. Before I started my PhD, I could not expect that China would grow so fast and became the world largest automobile market in terms of annual sales in 2009. The rapid growth of the car market not only means the increase in the sales figures but this has also led to changes in living style and local consumer behaviour in China. Therefore, the research questions about the existing market demand tackle modelling demand at both aggregate and disaggregate levels[5]. At the aggregate level, we consider how to better forecast sales at the aggregate level for car manufacturers who want to further invest in this market and for the local government who cares about the economic growth, oil supply and environmental issues resulted from the car sales growth. However, the short history of household car ownership in China means that

[5] The terms of "disaggregate level", "individual level" and "micro level" are used interchangeably in this thesis. The terms of "aggregate level" and "macro level" are also used interchangeably.

only limited sales data is available to forecast sales. The approach that we propose to address this challenge is to use the diffusion model that is usually employed for modelling car ownership to forecast car sales. We find this approach can outperform various benchmark models including time series and linear econometric models. At the disaggregate level, we explore the challenge of understanding local consumer behaviour and how this affects the adoption decision and the types of cars that consumers want to buy. This is important for the market players such as car manufacturers and governments because it provides them with more practical implications about how to effectively influence consumer behaviour and thus the market demand. The thesis explicitly compares the local consumer behaviour for car ownership and car type choice in China with the general findings in other more mature markets. Importantly, given the significantly different characteristics of the Chinese car market, we account for context specific factors, such as consumer knowledge, on local consumer behaviour. Consumer knowledge is found to have significant effects in many aspects of car market consumer behaviour, including car ownership, car type choice and future purchase intentions.

The second aspect of the thesis focuses on the future market demand, which is about the new products that are to become available in the future. During the period of recession, governments in major economies consistently carried out supporting or incentive policies to encourage the development and adoption of green cars or alternative fuel cars. China has been no exception and the Chinese government has defined the development of alternative fuel vehicles (such as hybrid and electric cars) as one of the national strategies to support the sustainable development of China. Therefore, it is important that we understand the consumers' potential preferences towards these green cars in China. This is important for car manufacturers as they need to know if consumers are going to buy such types of cars. This issue is also important for the government to see what types of policies they should be implementing to encourage more consumers to buy environmentally friendly types of cars. Given that these green cars are brand new products, we conducted a conjoint analysis and developed different specifications of discrete choice models to explore the potential heterogeneity of consumer preferences for them. Furthermore, with the identified heterogeneous preferences for the green cars, the next question emerging out of the research is how to forecast the dynamic diffusion of these green cars in China. The answer to this question is that we develop a dynamic segmentation approach to better capture the heterogeneous consumer preferences across the segments of car owners and non-car owners. In addition, we account for another important context characteristic of continuously increasing car ownership level in China. By comparing with the non-segmentation approach that is usually employed in the literature, we empirically demonstrate how different forecasts of the green cars are generated when the discrete choice models that better capture heterogeneity are applied in each segment. Importantly, in such an emerging market where the segmentation (whether households own cars or not) change for over time is appropriately accounted for.

Figure 1-4: The Thesis Framework



The key research objective of this thesis is to explore marketing approaches and models that can forecast demand in a rapidly changing market environment of the Chinese car market by accounting for local consumer preferences and the contextual characteristics. Most marketing models are developed in more mature markets and have been widely applied in these markets. It is important for researchers to assess how these approaches fare in the EMs. Therefore, this thesis explores different types of demand modelling approaches in an emerging market by looking at aggregate level sales forecasting to micro level consumer behaviour, from revealed preferences and purchase intentions towards existing products to stated preferences towards new products (green cars), and from modelling static preferences to forecasting the dynamic diffusion of the green cars. At the same time, this thesis is not just about the application of existing methods. We develop or extend the methods by properly accounting for the context specific characteristics. For example, we do not simply use the diffusion model to model the car ownership growth as appearing in most literature, but develop a sales forecasting approach based on the diffusion model to better capture the nonlinear growth pattern of sales. Furthermore, the context specific variables, such as the consumer knowledge, are accounted for when understanding local consumer behaviour for cars in China, in addition to the typical variables that have been included in the studies based on the developed markets. Moreover, the discrete choice models developed to understand the stated preferences for the green cars also take into consideration the fact that most Chinese consumers have no car ownership experience so that we do not hold any prior assumption about the correlations between any types of cars. Also, the dynamic segmentation approach proposed to forecast the diffusion of the green cars also considers the context features of China, which include the high level of consumer heterogeneity and the dynamic change of car ownership level over time in China. In summary, all these studies are linked together and form the main contribution of this thesis, which is that we explore the effectiveness of various demand modelling approaches in the context of the emerging Chinese car market and importantly by accounting for the contextual characteristics of this market. The thesis is organised around the following specific contributions which are connected to the marketing modelling challenges of understanding and predicting the demand for cars.

The first specific contribution of this thesis is that we explore the challenge of obtaining market forecasts at the aggregate level with limited historical sales data. As Meade & Islam (2006) point out, the empirical comparison of the diffusion models and how they are used to forecast sales against other models generally receives very little attention in the literature. This thesis proposes an approach that employs the diffusion models, which are usually used to predict car ownership levels in the literature (De Jong, et al., 2004), to forecast car sales in the Chinese car market. The study compares quarterly and annual car sales forecasts from three basic diffusion specifications with the time trend only (Bass, Gompertz and Logistic models), two extended specifications with both time and GDP per capita as independent variables (Gompertz and Logistic models) and three benchmark models of exponential smoothing, ARIMA and linear econometric models. The study also demonstrates the value of rolling forecasting approach in emerging markets and particularly so in the Chinese car market where there is the rapidly changing demand.

The second specific contribution of this thesis is to thoroughly investigate the consumer preferences in the Chinese car market in comparison to those found in other car markets. Here we mainly compare revealed preferences based on the car adoption behaviour using household survey data collected in China. We investigate the key determinants of car ownership and car type choice respectively in China and compare them with the general findings in other markets. This is particularly important for car manufacturers based in other markets that are going to manufacture and sell cars in a new market. Then we demonstrate why it is important to account for variables that may be specific to a particular context. We demonstrate this by showing how consumer knowledge as an example of context specific variables impacts on the choice to adopt cars and the types of cars that consumers buy. We segment the market based on levels of consumer knowledge and further demonstrate that if the heterogeneity of consumer knowledge is not properly accounted for, we obtain biased estimates for some of the parameters especially those that relate to preferences. We also examine the purchase intentions for cars in China and find the important effect of the consumer knowledge on consumers' future purchase intentions.

The third specific contribution of this thesis is that we investigate the potential heterogeneity of Chinese consumer preferences towards green cars. As we have discussed previously, it is important for the government and the car manufacturers to know whether Chinese consumers are going to buy the environmentally friendly cars given that they also face the option of non-environmentally friendly cars. This thesis investigates the stated preferences for different types of cars in China, i.e. one type of conventional petrol cars available in the market and two types of alternative fuel cars (hybrid and electric cars) considered as brand new products. Unlike previous studies, we do not impose any prior assumption about how consumers perceive the different types of cars in China. By comparing 4 different choice structures where each of them reflects how consumers choose between alternative fuel cars and conventional types of cars, we find that consumers in China do differentiate between the types of alternative fuel cars and are more likely to consider switching from petrol fuel vehicles to hybrid than to electric cars. Through taking into consideration that most households in China are non-car owners, further segmentation analysis shows that the car owners do consider the hybrid and electric cars to be similar, while the non-car owners perceive that the hybrid cars are more correlated with the petrol cars.

The fourth specific contribution of this thesis is that we present a preference-based demand forecasting approach for the new products in the Chinese car market. The proposed dynamic segmentation approach combines the well-specified discrete choice models for each segment with the segmentation diffusion model. Compared with the existing studies in the literature that employ one specific specification of discrete choice model for all consumers (Eggers & Eggers, 2011; Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006), the proposed approach helps account for not only the segmentation dynamics at the aggregate level but also two types of preference heterogeneity at the disaggregate level. The first type of preference heterogeneity can be identified across segments using the appropriate segmentation approach. Additional preference heterogeneity within each segment can be accounted for by employing appropriate and different choice structures for each segment. By applying the proposed approach to forecast the demand for the green cars in China, we empirically demonstrate how different forecasts of new product demand for such types of green cars are generated when the discrete choice models that better capture heterogeneity are applied in each segment and, importantly, in such markets, there is a need to consider the segmentation changes over time.

1.4 Summary and Thesis Outline

This chapter has presented the research background and context of this thesis. On the one hand, the importance of the EMs is becoming increasingly prominent, particularly in the context of the global financial crisis. On the other hand, different institutional characteristics are demonstrated in the EMs in comparison to in the developed markets. However, little attention has been paid to the forecasting and marketing problems in the EMs, as most existing marketing approaches were developed and applied in the context of developed markets. This chapter has discussed that how this thesis investigates whether these existing approaches can be applied in the context of an emerging market of China for car demand. More importantly it extends these approaches through appropriately accounting for the contextual characteristics and the local consumer preferences. Therefore this chapter has also set the scene to explore the marketing modelling challenges in this emerging market, through providing an overview of the Chinese car market and reviewing the latest market problems and challenges that emerge during its development.

As illustrated in Figure 1-4 of the thesis framework, this thesis tackles four closely related marketing problems or challenges in the context of the Chinese car market. Then this chapter has discussed in detail the main contribution of this thesis and the specific contributions of the four contributing chapters that address different marketing problems or challenges. After this chapter of setting the context of the research, the structure of the thesis is organised as follows.

The next chapter presents a brief review of new product forecasting methodologies, data and data collection methods used in the thesis. Please note that the literature review that relates to the specific research questions is reviewed in each contributing chapter. In this chapter we review the fundamental approaches that can be used for new product forecasting, specifically the diffusion model at the aggregate level and the conjoint analysis as well as discrete choice models at the individual level. In Chapter 2, we also discuss the data and data collection methods we use in the thesis. Both secondary and primary data are collected in the thesis. In particular, we discuss in detail the questionnaire design and importantly how we organised and implemented the survey in China.

The following 4 chapters after Chapter 2 are the main contributing chapters of this thesis and their specific contributions have been discussed in section 1.3. Chapter 3 presents an aggregate sales forecasting approach on the basis of the diffusion model to address the challenge of how to provide better sales forecasts with limited historical data. Chapter 4 investigates the Chinese consumers' revealed preferences at the disaggregate level in terms of car ownership and car type choice by comparing results from China to the general findings in the developed car markets and importantly accounting for the context specific variables such as the consumer knowledge. In addition, this chapter also models the consumers' future purchase intentions and identifies the important effects of consumer knowledge on their purchase intentions. Chapter 5 investigates the stated preferences of the Chinese consumers towards different types of green cars and the conventional petrol car based on a choice-based conjoint analysis. Using the static preference information explored in Chapter 5, Chapter 6 proposes a dynamic segmentation approach and empirically demonstrates that it can be used to better forecast demand for green cars in the emerging market context of China through appropriately accounting for the market specific characteristics.

Chapter 7 is the concluding chapter of the thesis with a summary of the major research proposition, key contributions and the possible generalisation of this research for other products in other EMs. This chapter also presents some possible directions for future research.

CHAPTER 2.

AN OVERVIEW OF NEW PRODUCT FORECASTING METHODS AND DATA

2.1 Introduction

In this chapter, we provide a review of the generic approaches that can be used to model and forecast demand for new products and in the thesis we have applied and modified some of these approaches to tackle the specific problems in the contributing chapters: Chapters 3 to Chapter 6. Please note that we will review the literature that relates to each specific problem in the respective contributing chapters. In addition, this chapter also provides an overview of data that has been used in this thesis. Both secondary and primary data are collected and used in different ways. We introduce in particular how we designed the questionnaire and organised the survey to collect the data in China. This chapter effectively constitute the methodology chapter of the thesis in that it documents the data as well as the empirical modelling approaches that have been applied to tackle the specific problems.

The car is a typical example of new products in China, due to its short adoption history and low car ownership level in this market. There are two general types of new product forecasting models based on different model specifications: the diffusion model and the individual behaviour-based adoption or choice model (Wind, 1981). Diffusion models are based on the time series data and assume a sigmoid-shaped growth curve of the product penetration levels (Bass, 1969; Mahajan, et al., 2000; Meade & Islam, 2006). On the other hand, the choice models are based on individual level data to investigate the consumer preferences for different characteristics of the products and how this will affect the choice of different options presented to the consumer (Greene, 2009; Train, 2003). For new products without any sales history, conjoint analysis based on hypothetical scenarios is usually employed to collect individual's potential attitudes or preferences towards the new products before applying the choice model (Green, et al., 2001; Gustafsson, et al., 2007; Louviere, et al., 2000). Recently, there have been some empirical studies that combine diffusion models and choice models to forecast new product demand (Jun & Park, 1999; Kumar, et al., 2002; Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006). As this combination approach is built upon two fundamental new product forecasting methods (i.e. the diffusion and choice models) that we mention above, we will not review this approach here, but leave it until Chapter 6 where we propose a better combination approach that accounts for consumer preference heterogeneity and car ownership dynamics over time when forecasting the diffusion of green cars in China.

The remainder of this chapter is organised as follows. The next section reviews the different forms of diffusion model, their estimation methods and the applications to forecast new product demand. Section 2.3 briefly reviews the conjoint analysis and its key elements. Section 2.4 discusses the different specifications of discrete choice models. Section 2.5 presents the data and data collection method. This chapter ends with a summary of new product forecasting methods and highlights their linkages with different contributing chapters in this thesis.

2.2 Diffusion Models

Diffusion models have been developed since 1960s to model and forecast the diffusion of technology innovations and new durable products (Meade & Islam, 2006). Researchers recommend to use diffusion models for new product forecasting when only early sales data is available (Wind, et al., 1981) and individual level data is limited (De Jong, et al., 2004).

There are different specifications of diffusion models. Three well-known diffusion types of models Gompertz, Logistic and Bass diffusion models are briefly presented below with some of their applications. For more diffusion model specifications, please refer to the Appendix section in Meade & Islam (2006), where the authors summarise eight types of sigmoid-shaped diffusion models.

2.2.1 Gompertz model

The Gompertz model (Gregg, et al., 1964) has been used to study the car market and car ownership. Using the Gompertz model, the aggregate car ownership level at time t is defined as:

$$N_t = M \cdot \exp\big(-\alpha \cdot \exp(-\beta \cdot t)\big) \qquad (2\text{-}1)$$

where N_t is defined as number of cars per 1000 people, M is the saturation or equilibrium car ownership level in the long term where market growth is stagnating and car ownership N_t reaches a plateau, and α and β are two positive parameters that define the shape of the growth curve. One characteristic of the Gompertz model is that its inflection point occurs before half of the market has adopted the product (Meade & Islam, 1995), which implies a slow diffusion speed and longer diffusion duration[6]. Other than the time variable which is typically included in the basic specification the model as in equation (2-1), other explanatory variables can be added into a Gompertz model. Therefore, the Gompertz model can be generalised as:

$$N_t = M \cdot \exp\left(-\alpha \cdot \exp\left(-\beta' \cdot X\right)\right) \tag{2-2}$$

where X is the explanatory variable vector and β is the parameter vector.

2.2.2 Logistic model

The logistic formulation of the diffusion model can be expressed as:

$$N_t = \frac{M}{1 + \alpha \cdot \exp\left(-\beta \cdot t\right)} \tag{2-3}$$

where α and β are parameters to determine the initial level and growth speed, and M is the saturation level of car ownership. Compared to the Gompertz formulation, the logistic model is symmetrical about its inflection point, which means the growth of the market slows down after half of the market adopts the product. Just like the Gompertz model, the logistic model has been generalised to include other dependent variables:

$$N_t = \frac{M}{1 + \alpha \cdot \exp\left(-\beta' \cdot X\right)} \tag{2-4}$$

where X is the explanatory variable vector and β is the parameter vector.

[6] The exact inflection point of the Gompertz curve is at 1/e, which is about 36.8% of total market potential.

2.2.3 Bass model

The Bass model (Bass, 1969) classifies potential adopters of new products into 2 different groups: *innovators* and *imitators*. In the Bass model, the probability for an individual who does not own a car who will purchase at time t is

$$f_t = (p + qF_t)(1 - F_t)$$

where F_t is the cumulative distribution function (CDF) of Bass model, standing for the ratio of car ownership level at time (t) against the saturation level, and p and q are coefficients of innovation and imitation respectively. Based on the initial condition of no adoption before the diffusion process, i.e. $F_0 = 0$, the cumulative distribution function of Bass model can be derived as

$$F_t = \frac{N_t}{M} = \frac{1 - \exp(-(p + q)t)}{1 + \left(\frac{q}{p}\right) \exp(-(p + q)t)}.$$

Therefore, the car ownership level based on the Bass model is expressed as

$$N_t = M \cdot \frac{1 - \exp(-(p + q)t)}{1 + (\frac{q}{p})\exp(-(p + q)t)} \tag{2-5}$$

Essentially, the Bass model is built based on the generalised logistic curve (Mahajan & Muller, 1979; Meade & Islam, 1995).

The Bass model has been generalised to accommodate the effects from other marketing variables (Bass, et al., 1994):

$$f_t = (p + qF_t)(1 - F_t)x_t$$

where x_t stands for the marketing effort that can be modelled in a following equation to account for various marketing mix variables such as price and advertising.

$$x_t = 1 + \beta_1 \frac{\mathrm{Pr'}_t}{\mathrm{Pr}_t} + \beta_2 \frac{\mathrm{Ad'}_t}{\mathrm{Ad}_t}$$

where Pr_t and Ad_t are the price and advertising at time *t*, and Pr'_t and Ad'_t are the change rate of price and advertising respectively at time *t*.

Regarding the model estimation, although the ordinary least square (OLS) is employed in Bass (1969) to estimate the parameters of the Bass model, it has been found that the OLS method is not an optimal choice, as it may lead to the wrong signs of the estimated parameters and large bias of the estimated amount (Meade & Islam, 2006; Putsis & Srinivasan, 2000). Alternative methods for the Bass model estimation are nonlinear least square (NLLS) (Srinivasan & Mason, 1986) and maximum-likelihood estimation (MLE) (Schmittlein & Mahajan, 1982). As discussed in Meade & Islam (2006) and Putsis & Srinivasan (2000), the performance difference between NLLS and MLE is not significant, but it is clear that both of them are superior over OLS when estimating the Bass model. In addition, the NLLS method is also applicable to estimate non-Bass diffusion models, such as the Logistic and Gompertz models (Meade & Islam, 1995). It is worth noting that Van den Bulte & Lilien (1997) point out that the NLLS estimation method may underestimate *m* and *p*, and overestimate *q* for the Bass model. One solution suggested by them to avoid the estimation bias is to exogenously set different levels of *m* and then investigate the sensitivity of other parameters. We follow this approach in the thesis.

The diffusion models have been widely applied to model and forecast the diffusion of many products, including automobile (Bouachera & Mazraati, 2008; Dargay & Gately, 1999; Dargay, et al., 2007; Kobos, et al., 2003; Tanner, 1958, 1975), telecommunication (Robertson, et al., 2007; Sundqvist, et al., 2005; Wu & Chu, 2010), and many other durable goods (Bass, 1969; Bottomley & Fildes, 1998; Tsai, et al., 2010). In particular, in the EMs where data are very limited, the diffusion models are thought to be the only available method to model the diffusion of cars, which is usually measured as the car ownership levels (De Jong, et al., 2004). More importantly, Meade and Islam (2006) point out that more attention should been paid to the empirical performance comparisons of diffusion models with other sales forecasting methods, because it is important for researchers to justify their choice of forecasting models by comparing the forecasting capabilities of different models (Fildes, et al., 2008). Therefore, in the next chapter, we develop a sales forecasting method based on the diffusion model and compare the sales forecasts from such method against those produced by three benchmark models, which are Exponential Smoothing, ARIMA and linear econometric models that employ sales data directly.

2.3 Conjoint Analysis

Conjoint analysis is one of the most useful marketing research methods for analysing consumer tradeoffs between two or more products with different profiles (Green, et al., 2001). It was introduced into the marketing research domain in early 1970s with the seminal work from Paul Green and Vithala Rao (Green & Rao, 1971; Wind & Green, 2004). Since then, conjoint analysis has been extensively used to investigate not only consumer preferences or intentions to buy existing products, but also how consumers may react to potential changes in the existing products or to new products to be introduced to the market later (Cattin & Wittink, 1982; Green, et al., 2001; Wittink & Cattin, 1989). In this thesis, the conjoint analysis technique is applied in Chapter 5 to investigate the consumers' potential preferences towards the green cars in China, and in Chapter 6 to forecast the diffusion of the green cars by accounting for heterogeneous consumer preferences across segments.

In their review of this approach, Hauser & Rao (2004) propose that conjoint analysis has five basic elements

- Decomposing the alternative (product or service) into a set of attributes (factors). This also includes the definition of different levels of each attribute to be explored. The choice of different numbers of attributes and levels of each attribute is worth careful considerations. On the one hand, by including more attributes and more levels on each attribute, researchers can get more information about consumer preferences. On the other hand, the number of combinations of attributes, called product profiles, will increase massively with more attributes and/or levels included in the conjoint analysis, which leads to the increased respondent burden and data collection issue. Therefore, balanced definitions of attributes and their levels are required when conducting the conjoint analysis.
- Representation of alternatives. It involves how to introduce the product or service as well as each attribute, so that respondents can fully understand them. In addition to presenting the value of each attribute, the representation can utilise verbal description, pictorial illustration (Hensher & Greene, 2003), and more vivid multimedia or virtual reality techniques (Rogers & Soopramanien, 2009) to present the products.
- Fractional design of experiments. Considering a relatively small experiment with 5 attributes and 3 levels on each attribute, the complete factorial design

will involve 243 ($=3^5$) total combinations. This suggests that the complete factorial design is usually impractical for the conjoint analysis except for some cases extremely limited attributes and levels (Louviere, et al., 2000). A feasible solution is the fractional factorial design which selects a subset of the complete design based on different sampling methods. One of the most popular fraction designs is the orthogonal design, where all attributes are statistically independent with each other (Hensher, et al., 2005; Louviere, et al., 2000).

- Conjoint data collection. In early applications of conjoint analysis, a ranking method was usually used, such as in Green & Wind (1975), which requires respondents to rank order the different profiles of products presented to them. Later, researchers in both academia and industry found they could collect rating data with scales about consumer preferences, which can contribute to very robust analysis (Hauser & Rao, 2004; Louviere, 1988) and reduce respondent burden by requiring less judgements or comparisons from each respondent (Raghavarao, et al., 2011). The potential issue of the rating method is that it is usually weak to capture the competitions or the tradeoffs made by the respondents between different product profiles. Furthermore, when using the rating data to model consumer choice probabilities, Guyon & Petiot (2011) point out that three typical model specifications, including the multinomial logit (MNL) model (McFadden, 1974), Bradley-Terry-Luce (BTL) model and the Green and Krieger model (Green & Krieger, 1988) are not able to avoid the independence from irrelevant alternative (IIA) property. Therefore, an alternative way to collect the conjoint analysis data is the choice-based conjoint analysis, which can capture the competitions or tradeoffs between several products and have the flexibility to hold the IIA property (Elrod, et al., 1992). In the choice-based conjoint analysis, respondents only make a single choice from several product profiles presented to them simultaneously (see a choice scenario example in Figure 5-2). The choice-based conjoint analysis is also known as the stated preference (SP) or stated choice (SC) experiment in the domain of choice modelling (Hensher, et al., 2005; Louviere, et al., 2000; Raghavarao, et al., 2011). In addition, a hybrid or adaptive technique has been developed to customise the product profiles for different respondents based on the previous information collected from them, and it usually requires the respondents to consider a subset of the full-profile scenarios (Green & Krieger, 1996).

- Modelling methods. When evaluating consumer preferences for the presented product, a part-worth model is usually employed (Green, et al., 2001), which defines the consumer preference for each product as the summation of part-worth function values of different attributes at the selected levels on each attribute (i.e. the rate for each attribute). With the choice-based conjoint analysis data, discrete choice models (Train, 2003) have developed a large variety of model specifications to account for different correlation patterns between alternatives. In the next section, we review the most common discrete choice models, including the MNL model, the nest logit (NL) model, the multinomial probit (MNP) model and the mixed logit model.

A typical example of the conjoint analysis is discussed in Train (2003) to obtain the stated preference data on residential customers' choice of energy suppliers. In the experiments, each respondent was presented with 8-12 choice scenarios, where each scenario consists of four hypothetical energy suppliers. The attributes of each supplier include three types of energy price (fixed price, price based on different time of date, or seasonal price), the length of the contract, and whether the supplier is a local utility company, whether it is a well known energy supplier or an unfamiliar company. Therefore, a generic utility function can be developed to include both alternative attributes (X) and individual characteristics of the respondents (Z)

$$U_{in} = \alpha_n + \beta' X_{in} + \gamma_n' Z_i + \varepsilon_{in},$$

where α_n is the alternative specific constant and γ_n' is the effect of individual characteristics on different alternatives. Importantly, the coefficient vector of alternative's attributes, β', is the weights the respondents place on each attribute, which essentially shows how each attribute influences the respondents' utility and valuation on each alternative and thus their choices.

2.4 Discrete Choice Models

Discrete choice models have been widely applied since 1960's due to the rapidly increasing availability of survey data on individual behaviour as well as the growing use of computers for complex analysis (McFadden, 2001). Considering a choice environment, a decision maker faces a choice set with a finite number of alternatives. Each alternative provides "utility" or a level of satisfaction to the consumer. This can be defined as U_{in}, depending on both the individual's characteristics and the attributes of the alternative. Following *utility maximization* criteria where the consumer chooses the alternative with the highest level of utility, the choice probability for the decision maker n to chooses alternative i from choice set J_n is

$$P_{in} = \Pr(U_{in} > U_{jn}, \text{for all } j \in J_n \text{ and } j \neq i) \tag{2-6}$$

Since the utility cannot be completely explained by measurable/observable characteristics, it is decomposed into the observed portion (V_{in}) and the random portion (ε_{in}). Thus equation (2-6) can be rewritten as:

$$P_{in} = \Pr\left(V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn}, \forall j \neq i\right)$$

$$= \int I\left(\varepsilon_{jn} - \varepsilon_{in} < V_{in} - V_{jn}, \forall j \neq i\right) f(\varepsilon_n) \mathrm{d}\varepsilon_n \tag{2-7}$$

where $f(\varepsilon_n)$ is the joint density of the random vector $\varepsilon_n = \{\varepsilon_{in}, \cdots, \varepsilon_{In}\}$ and $I(\cdot)$ is the indicator function whose value equals 1 when the term within the parentheses is true and 0 otherwise. The choice of different specifications of the density $f(\varepsilon_n)$ will directly determine the different structures of following discrete choice models and thus the assumed choice behaviour to make the decisions. Due to the space constraint, we mainly review the multinomial logit (MNL) model and the nested logit model which are going to be employed in this thesis. The multinomial probit (MNP) and the mixed logit models are briefly discussed thereafter.

2.4.1 The multinomial logit model

The multinomial logit (MNL) model, proposed by McFadden (1974), is the most widely used model in the discrete choice methods. The random term of the MNL model follows i.i.d. type I extreme value (Gumbel) distribution with the probability density function (pdf) defined as

$$f(\varepsilon_{in}) = \exp(-\varepsilon_{in}) \cdot \exp\left(-\exp(-\varepsilon_{in})\right),$$

and cumulative distribution function (CDF) as

$$F(\varepsilon_{in}) = \exp\big(-\exp(-\varepsilon_{in})\big).$$

Therefore, from equation (2-7), the MNL choice probability for decision maker n to choose alternative i can be derived as a closed form as

$$P_{in} = \frac{\exp(V_{in})}{\sum_j \exp(V_{jn})} \quad (2\text{-}8)$$

Because the MNL model has the closed form probability in equation (2-8), the traditional maximum likelihood technique is commonly used to estimate the parameters of the MNL model.

A well known characteristic of the MNL model is the independence from irrelevant alternatives (IIA) property. This property implies that all choices are mutually independent of each other ignoring the possibility that in some situations some alternatives are similar to each other. The Hausman Test is proposed to test whether the IIA property is violated or not when applying a MNL model (Hausman & Mcfadden, 1984). When applying the Hauman Test, the original choice set is first reduced by removing one alternative and then the MNL model is estimated again to get estimation results of a restricted model. Then Hausman test statistic is calculated based on both unrestricted (original) and restricted model results, which should follow chi-square distribution with the degree of freedom to be the number of estimated parameters (Hensher, et al., 2005). Given the IIA property of the MNL model, many extension models have been proposed to allow for different correlations between alternatives, such as the nested logit (NL) model, the multinomial probit (MNP) model and the mixed logit model.

2.4.2 The nested logit model

The most famous extension of the MNL model is the nested logit (NL) model. The NL model groups alternatives available in the choice set into several subsets, called "nests", and allows the correlations between the utilities of pairs of alternatives within a nest while no correlation for alternatives in different nests (Daly & Zachary, 1978; McFadden, 1978; Williams, 1977). A common way to understand the NL model is the tree structure, where each branch stands for a subset of alternatives and each leaf on the branch stands for the alternative. Importantly, the NL model relaxes the IIA property of the MNL model is following way. For any two alternatives in the same nest, their ratio of probabilities remains independent of the changes of other alternatives out of the nest, while for two alternatives in the different nests, their probability ratio depends on all alternatives in both nests. Therefore, in the NL model, the IIA property is held within each nest but is not held across nests (Train, 2003).

In the two-level NL, the probability of decision maker n to choose alternative i is the product of marginal probability of the nest m (P_{mn}) and the conditional probability of the alternative i given to choosing nest m ($P_{i|m,n}$),

$$P_{in} = P_{i|m,n} \cdot P_{mn}$$

$$= \frac{\exp(V_{in}/\mu_m)}{\sum_j \exp(V_{jn}/\mu_m)} \cdot \frac{\exp(\mu_m \Gamma_{mn})}{\sum_{m'=1}^{M} \exp(\mu_{m'} \Gamma_{m'n})} \qquad (2\text{-}9)$$

where Γ_{mn} is the log-sum or inclusive value (IV) variable of nest m with log-sum or IV parameter μ_m, $\Gamma_{mn} = \ln(\sum_j \exp(V_{jn}/\mu_m))$. The utility maximisation assumption requires the IV parameters must be between zero and one (Greene, 2009). The IV parameter is an indicator of independence of all alternatives in the nest and $1 - \mu_m^2$ is the correlation index between pairs of alternatives within the same nest (Ben-Akiva & Lerman, 1985). When $\mu_m = 0$, it means all alternatives in the nest m are identical with perfect correlations. When $\mu_m = 1$ for all nests, the NL model collapses to the MNL model, which indicates the independence and no correlation between all alternatives in every nest. If there is only one alternative in every nest, the NL model also reduces to the MNL model.

Table 2-1: Direct- and cross-elasticity of the MNL and NL Models

Model	Direct-elasticity (change in P_i due to change in X_i)	Cross-elasticity (change in P_j due to change in X_i)
MNL	$\beta(1-P_i)X_i$	$-\beta P_i X_i$
NL	• i not in the nest $\beta(1-P_i)X_i$ • i in the nest m $\beta\left[(1-P_i)+\left(\frac{1}{\mu_m}-1\right)(1-P_{i\mid m})\right]X_i$	• i and j not in the same nest $-\beta P_i X_i$ • i and j in the same nest m $-\beta\left[P_i+\left(\frac{1}{\mu_m}-1\right)P_{i\mid m}\right]X_i$

Elasticity is an important way to understand the relationship between alternatives in the model. Table 2-1 compares the direct and cross-elasticities of the MNL and the NL models. When the alternative is not in the nest, the direct and cross-elasticities of the NL model have the same structures as those of the MNL model. However, when the alternative is in the nest, its direct-elasticity is greater than the corresponding direct elasticity for alternatives not in the nest, since the IV parameter of the nest (μ_m) is less than one following utility maximisation. Similarly, the cross-elasticity for the pair of alternatives in the same nest demonstrates greater magnitude than the corresponding cross-elasticity for alternatives not in the same nest. This implies that the competition among similar alternatives within the nest is higher than dissimilar ones across nests. The magnitudes of both direct- and cross-elasticities increase as the

IV parameter decreases from one, which is the IV parameter value for either single alternative nest or the MNL model.

The multi-level NL model is the direct extension of the two-level NL model. For instance, in three-level NL models, the choice set is partitioned into nests and then each nest is further partitioned into sub-nests. Therefore, the choice probability of the three-level NL model can be express as the product of three logit forms similar to equation (2-9).

2.4.3 The multinomial probit model

The multinomial probit (MNP) model (Daganzo, 1979; Hausman & Wise, 1978) assumes that the random term ε_n in equation (2-7) follows a joint normal distribution with mean of zero and covariance matrix Σ. This assumption allows the MNP model to hold several advantages over the MNL model: random taste variation, unrestricted substitution pattern without the IIA property, and its applicability for panel data analysis as it can allow for the unobserved utility to be correlated as a joint normal distribution over time for each decision maker (Train, 2003).

Unlike the MNL and NL models, the choice probability of the MNP model cannot be expressed as a closed form, and thus it has a major shortcoming in the high-cost estimation. Simulation is the common technique to approximate the MNP choice probability. Train (2003) explains the simulation techniques for the MNP model in detail. In addition, the normal distribution assumption of random terms is sometimes invalid as a representation of the random taste heterogeneity (Hess, 2005). Therefore, mixed logit models have been proposed to allow for more flexible distribution of the random term.

2.4.4 The mixed logit models

The mixed logit models have also gained popularity recently in that it not only holds the same advantages as the MNP model but also allows for a flexible distribution for the random portion. The mixed logit model is the direct extension of the MNL model by allowing for the random coefficients of the observed utility in the MNL model. Therefore, the mixed logit model is also called *random coefficients logit* model (McFadden & Train, 2000).

The choice probability of the mixed logit model is the MNL choice probability in equation (2-8) conditional on the coefficients of explanatory variables. The distribution selection of the random coefficients is largely based on the characteristics of the corresponding explanatory variables captured by the researcher. The normal distribution is widely used to approximate the preference randomness among people. The lognormal is useful when all people have the similar inclination for certain attributes, such as price, so the utilisation of lognormal distribution can ensure the negative sign for the coefficient of the price. Other distributions, such as uniform, triangle, and truncated normal are also illustrated in Train (2003). Due to the lack of a close-form, the estimation of the mixed logit model must rely on simulated methods, as discussed in McFadden & Train (2000) and Train (2003).

2.5 Data and Data Collection

This section discusses the data used in this thesis, which consists of both secondary data and primary data. The secondary data essentially includes various time series data such as historical car sales, GDP per capita and household car ownership level in China, and we also collected vehicle specification data from the secondary sources. The primary data is mainly collected through a consumer survey conducted in early 2010. We also discuss how we designed the questionnaire for this survey and then how we organised and implemented the survey in China.

2.5.1 Secondary data

We collected secondary data of car market sales from the China Associate of Automobile Manufacturers and the Development Research Centre of the State Council of China. It is worth noting that the concept of passenger cars in China did not formally exist until the end of 2001 when a new national standard of vehicle classification (GB/T3730.1-2001) was implemented. Therefore, the car sales data we collected starts from 2002. In addition, we also collected GDP per capita and population data since 2002 in China from either the website of the National Bureau of Statistics of China (NBSC) or the China Statistical Yearbooks published by the NBSC. These datasets are mainly used in Chapter 3 to forecast the car market sales in China using diffusion models versus the benchmark models.

Another secondary piece of data that we collected is vehicle attributes, which are used for analysing consumer preferences towards different types of cars with different specifications in Chapter 4. During our consumer survey, in order to reduce the respondents' burden and avoid their reporting errors, we did not ask for them to provide detailed specifications of their owned vehicles. Instead, we collected the information from the third party website or magazine based on their purchased car models and purchase time (year and month). In addition, when analysing consumers' car type choice preferences, we needed to construct a choice set for each respondent, which consisted of not only the selected model, but also other model alternatives[7]. Therefore, vehicle attributes must be collected for all choice alternatives in the choice set. Specifically, we collected the following vehicle attributes: car purchase price (RMB), fuel consumption rate (litre per 100 km), maximum horsepower (Kwt), minimum turning radius (metre), vehicle length (metre), the number of airbags, and vehicle country of origin. Most data of vehicle attributes except the fuel consumption rate were collected from the automobile magazine, *Orient Auto*, which regularly updates and publishes detail specifications of all available cars in China. Regarding fuel consumption rate, the magazine publishes data released by car manufacturers, which are inaccurate as they commonly assume constant driving speeds such as 90 km/h. The more accurate fuel consumption rates for all vehicle models were collected from a governmental website of car fuel consumption rates developed by the Ministry of Industry and Information Technology of China[8]

In addition, we also collected household car ownership data reported in the China Statistical Yearbooks from 2002 to 2009. This data was collected by the NBSC every year based on the annual nationwide household survey and it is reported as the number of cars owned by every 100 households. Given that only 10.89 cars were owned per

[7] For the details of how to construct the choice set randomly based on the market shares, please see Section 4.6.1 in Chapter 4.
[8] The website address is http://chinaafc.miit.gov.cn.

100 households by 2009, China is still at the early stage of car adoption and also most car-owning households only own one car. Thus we can use the household car ownership data to approximate the percentage of the car-owning households in China. The data is used in Chapter 6 to estimate a car ownership diffusion model, based on which the future segment size of car-owning households in China is then extrapolated in the forecasting horizon.

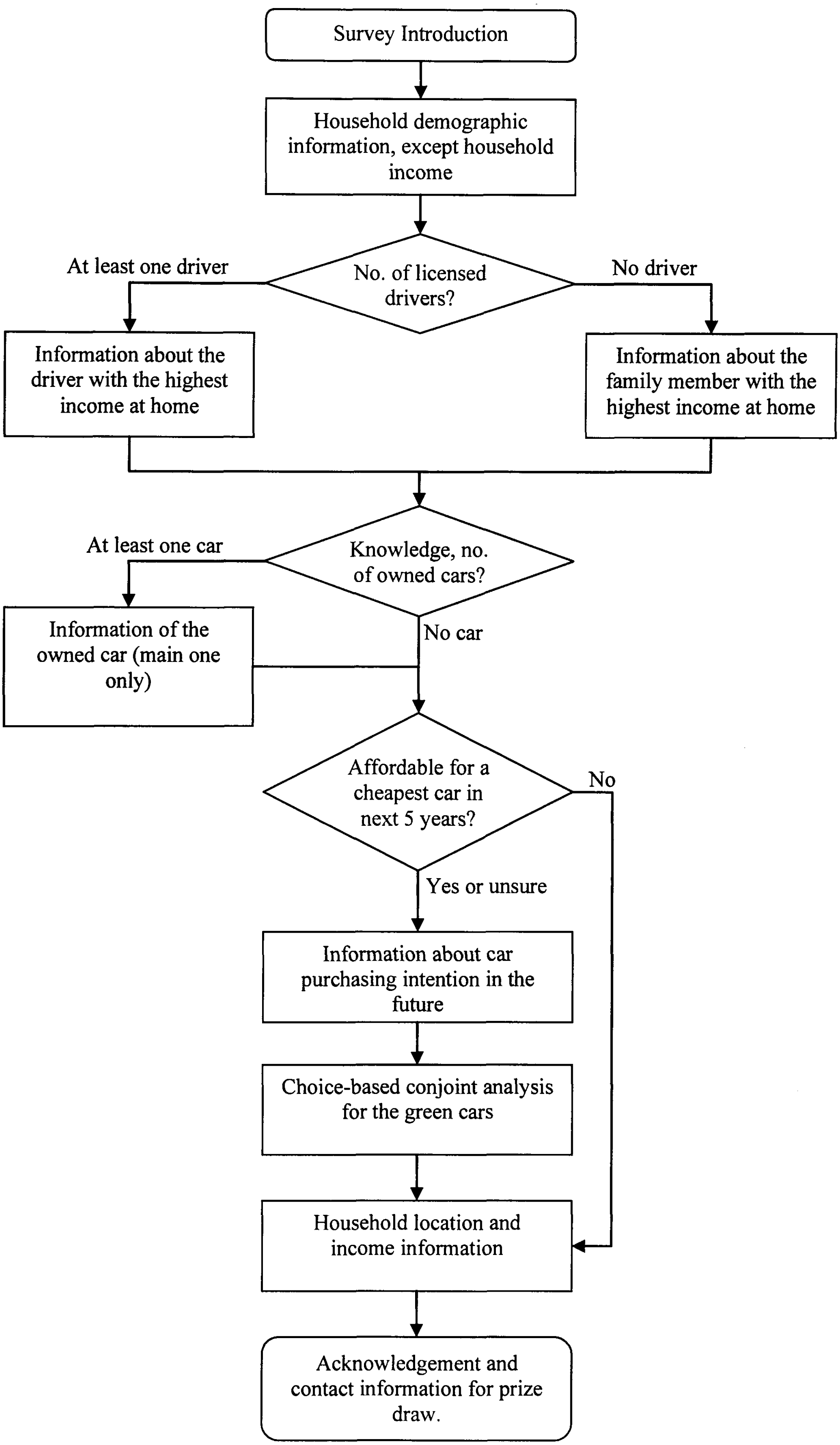
2.5.2 Primary data and questionnaire design

In addition to the secondary data, we conducted a household survey in China to collect data about the Chinese consumers' intentions and preferences to buy cars. In this survey, we are interested in their existing car ownership and choice preferences, their future purchase intentions for cars, and their stated preferences for a type of new product, green cars to be available in the market. We also collected some data about the demographic characteristics of the respondents and their families.

The design of the questionnaire is based on the extensive literature review as well as our knowledge of the local market situation in China. Bunch and Chen (2008), De Jong, et al. (2004) and Potoglou & Kanaroglou (2008a) provide comprehensive reviews on modelling car demand using the survey data at the disaggregate level. Besides, we review a number of individual studies of car ownership and car type choice models as summarised in Chapter 4, and many studies of modelling consumer choices for green cars as discussed in Chapter 5. Important information gathered from the literature review helped us develop the survey questions. We also take into account the local market situation when designing the questionnaire. For example, since the average car ownership level in China is low and it is fairly rare to have families owning multiple cars, we only ask the respondents (if their families own cars) to report on one owned car information in detail. If some households have more than one car, we ask the respondents to select the main one as the representative to fill in the questionnaire. Since we contend that the Chinese car market may differ significantly from the other more developed markets, we developed questions to understand the local market context. For example, we asked the respondents to evaluate their own knowledge level about cars and the car market, which we think is an important feature for that market, and we investigate in detail its significant effects on local consumer behaviour in Chapter 4.

Our questionnaire consists of following main sections. For the full questionnaire, please refer to the Appendix 1 of this thesis. The logical sequence of each questionnaire section is presented in Figure 2-1.

Figure 2-1: Flowchart of the questionnaire



1. Current car ownership, such as number of cars owned now, and detailed information of the currently owned car (for car owning households, and only asking the primary one if owning more than one car). In addition, we asked the respondents to evaluate their own knowledge level about cars no matter whether they were owning cars or not; they chose one from four given options: no knowledge, basic knowledge, familiar or very familiar with cars. Since most consumers in China do not have any experience of using cars, we collected subjective knowledge about the familiarity of cars instead of the more objective expertise about handling cars. More discussion about consumer knowledge is presented in Chapter 4.
2. Self-assessment of the affordability for the cheapest car in the next 5 years. If self-assessment shows the household cannot afford the cheapest car, the respondent will be directed to the section of household residential location and income, otherwise the survey will continue.
3. Car purchasing intention in the next 5 years. This includes the planned year of the next purchase, the preference for the country of origin, whether it is replacing the current car or buying one more, preference for a new or used car, preference for domestically made or imported cars, etc. In addition, we asked respondents to select and sort top 5 factors based on importance in his/her mind when purchasing a car. We also allowed respondents to customise their intended car specification in 5 categories and 32 attributes (purchasing price, vehicle size and general attributes, performance, safety/security and comfort/convenience equipments).
4. Choice-based conjoint analysis for electric, hybrid and petrol cars. We designed the conjoint analysis to involve product attributes including purchase price, running cost, three types of incentives for hybrid and electric cars, availability of

charging facility and vehicle range after full charging for the electric cars. Each attribute has three levels. We employ an orthogonal fractional factorial design, which was implemented through SPSS to generate 32 choice scenarios as shown in the Appendix 2. In addition, we adopt the hybrid conjoint technique discussed previously to customise the price of the petrol car using the respondent's preferred purchasing price provided earlier, so that the presented choice scenarios are more related to what respondents really consider. We randomly allocated 8 choice scenarios to every respondent and within each scenario, the respondent was asked to select one that his/her household would most likely purchase. More information on the design of the conjoint analysis for investigating consumer preferences for green cars is described in Chapter 5.

2.5.3 Survey data collection

Before formally starting the survey in China, we conducted a pilot survey with 12 participants to detect any mistakes or missing information in our questionnaire. Based on their feedback, we improved the wording and revised some questions in our survey.

Our survey was implemented online (www.surveymonkey.com). To mitigate the potential bias of the online survey towards the computer/internet users, we recruited students from two local universities in China (North China Electric Power University and China University of Mining and Technology) to help us collect data. During the Chinese New Year Holiday in 2010, these students went back to their home areas to collect data for us, which allows for getting data from different regions in China. For each student, we prepared an information card as shown in Figure 2-2 for them to easily approach the respondents. The card had a seasonal greeting for each survey participant and showed our online survey address. On the card, we also notified people that there would be a prize draw if they complete our survey. All students were able to provide their own computers for the respondents without access to the Internet to complete the survey. With the help of these students, we were able to access a wide range of households living in different areas of China and even those households who have no access to a computer.

Figure 2-2: information card used in the survey

In order to encourage the participation of the potential respondents, we designed three types of prizes for those who completed the survey successfully. The first prize was 500 RMB for one person, the second prize was 200 RMB each for two persons, and the third prize was 100 RMB each for ten persons. We also provided some allowances for the university students who helped us during the survey. For every completed questionnaire with a clear indication of who recommended the respondent to participate in the survey, we gave the named student 10 RMB allowance to acknowledge his/her help and cover the potential cost of using his/her computer or public internet café.

Our survey started on 20th January 2010 and ended on 30th April 2010. In total we have 760 respondents who completed the survey. It is worth noting that we did not force the respondents to answer every question in the survey, because we wanted to collect correct instead of potentially manipulated answers if the respondents had not wanted to disclose some information. Therefore, after removing some cases where some conflicting answers are detected or where information on some key variables (such as income, household head's age and sex) were missing, we have 563 cases available. We will present the data description for each chapter that uses the survey data for different purposes.

2.6 Summary

This chapter has presented the methods that are usually employed for new product forecasting and has focused on those which have been employed in the thesis. More attention has been devoted to diffusion models, conjoint analysis and discrete choice models. These new product forecasting methods rely on different types of data, either the secondary or the primary data which we have discussed in the second part of this chapter. Both these new product forecasting methods and data are then closely linked to the modelling methodologies that we use in the following contributing chapters.

The diffusion model is generally based on a sigmoid-shaped curve to model and forecast the growth pattern of innovations or new products at the macro level. Usually based on the secondary data, the diffusion model has the advantage of not requiring primary data (De Jong, et al., 2004). So, it is a typical approach used for modelling new products and in the thesis we test whether it can be used to model car demand in China. However, the literature has paid little attention to car sales forecasting based on the diffusion model and particularly its sales forecasting performances in comparison with other forecasting methods (Meade & Islam, 2006). Therefore, the next chapter presents a car sales forecasting approach based on the diffusion model that usually models car ownership levels. More specifically, we employ five different specifications of the diffusion model and compare their sales forecasting performances with those from three benchmark models (exponential smoothing, ARIMA and linear econometric models) that use sales data directly to forecast sales. The diffusion model and the benchmark models in the next chapter are estimated using the secondary data of car sales, GDP per capita and population.

For brand new products without any history, the conjoint analysis is widely used to investigate consumers' potential preferences for different features of the new products that may be still under design or development. There are different types of conjoint analysis. In particular, the choice-based conjoint analysis is useful when studying heterogeneous consumer preferences, as its corresponding modelling approach, the discrete choice models, can allow for different types of correlations between multiple choice alternatives. In most cases, researchers design the experiment and implement the conjoint analysis through surveys by themselves. Thus the data collected through the conjoint analysis is the primary data, except in rare cases researchers utilise the existing conjoint analysis data from other surveys. In Chapter 5 of this thesis, we conduct the choice-based conjoint analysis to investigate the consumer preferences for the green cars in China. Based on the primary data we collected through this conjoint analysis, we explore all possible correlations between conventional petrol cars and two types of the green cars (e.g. hybrid and electric cars) and finally identify the most appropriate choice structure of the Chinese consumers for the green cars.

In addition to using the conjoint analysis, the discrete choice models can also be employed to investigate the revealed preferences (RP) or purchase intentions for the existing products. In Chapter 4 of this thesis, we use the primary data collected through our survey and specify the discrete choice models to examine the key determinates of the car ownership and car type choices in China. We explicitly compare these key factors with those typically found in the developed car markets and highlight the differences and the corresponding insights. Importantly, we account for context-specific factors such as consumer knowledge in these models to demonstrate the importance of understanding specific local consumer behaviour when investing in new markets.

Finally, this chapter has also noted the recent empirical studies that combine the diffusion model with the discrete choice model to forecast new product demand. Such combination approach is reviewed in detail in Chapter 6 before we present the proposed dynamic segmentation approach to forecast the diffusion of green cars in China. Using both secondary and primary data, the proposed approach combines the segmental choice models at the individual level and the car ownership diffusion model at the aggregate level.

CHAPTER 3.

EMPLOYING DIFFUSION MODELS TO FORECAST SALES IN THE CHINESE CAR MARKET[9]

3.1 Introduction

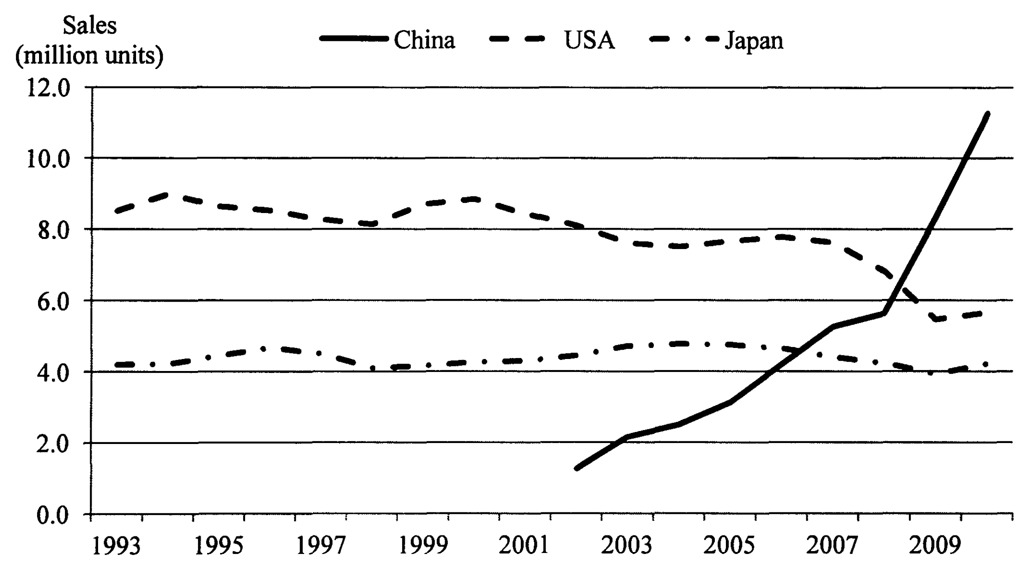
Forecasting aggregate car ownership level and how this will change over time is typically used to get an idea of the market potential. It is useful for governments and other policy makers when they have to make decisions related to infrastructure development or taxation policy (De Jong, et al., 2004). However, a car manufacturer is more interested in the forecast of car sales/demand because this will have an important bearing on some of the strategic decisions of that company such as whether to keep exporting or to produce their cars locally. In the EMs, car sales forecasting is even more critical because of both the short term revenue and long term market potentials.

Limited data can be a significant problem for marketers and forecasters. The short history of private car ownership in China means that there is limited time series data to predict future sales. In addition, preference survey data or consumer panels, which have been extensively utilised in more mature and developed markets, may also be

[9] This chapter is largely based on: Qian, L. and Soopramanien D., Employing diffusion models to forecast sales in the Chinese car market, to be submitted to *Journal of Business Research*.

limited or unavailable in such a market. China provides an interesting context for this research. The Chinese car market is a typical market in an emerging economy with a fast growth rate but with a short history. Starting from 2002, the Chinese car market has experienced massive growth and at the end of 2009 it overtook the United States to become the largest car market in the world in terms of annual sales. Historical annual new car sales in China, the United States and Japan are shown in Figure 3-1, which shows a non-linear growth pattern in China during the early stage of household vehicle adoption.

Figure 3-1: New car sales history in China, USA and Japan



Data Sources: China Associate of Automobile Manufacturers (CAAM)
US National Automobile Dealers Association (NADA)
Japan Automobile Manufacturers Association (JAMA)

However, the Chinese car market is far from reaching what is commonly termed the saturation level (where growth stagnates and sales/car ownership reaches a plateau). According to World Development Indicators published by the World Bank (see Table 3-1), the global average car ownership increased from 90 cars per 1000 people in early

1990s to 118 cars per 1000 people in 2005. The major car markets in developed economies, represented by the US, UK, Germany and Japan, have high car ownership rates with roughly one car owned by every two persons in 2005. China had a much lower car ownership level, where every 1000 people only owned 15 cars in 2005, although the figure had increased more than 10 times since 1990.

Table 3-1: Number of cars per 1000 people worldwide and across some countries

Year	World	China	India	Brazil	Germany	Japan	UK	US
1990	91.06	1.43	2.42	NA	386.30	283.32	340.58	573.28
1995	89.65	2.90	4.12	119.73[b]	494.90	356.19	352.02	484.84
2000	104.35	7.00	6.00	134.00	515.00	413.00	388.00	473.00
2005	117.91[a]	15.00	8.00[c]	136.00[a]	550.00	441.00[a]	457.00	461.00[d]

Note: [a]. data in 2004; [b]. data in 1996; [c]. data in 2003; [d]. Excludes personal passenger vans, passenger minivans, and utility-type vehicles, which are all treated as trucks

Data Source: 2006 World Development Indicators, World Bank

Researchers have contended that diffusion models are particularly useful when data is limited and more so in the EMs because they require limited information (Dargay & Gately, 1999; De Jong, et al., 2004). However, limited research has been conducted to forecast sales using diffusion models in the EMs. Furthermore, Meade and Islam (2006) stated that it is important to compare the different specifications of diffusion models and how they are used to forecast sales and that researchers should justify their choice of forecasting models by conducting effective comparison of models (Fildes, et al., 2008). Our objective is to propose an approach where we can forecast sales of cars in China by using the diffusion models that usually predict car ownership levels. Our research significantly differs from the existing body of work on car ownership modelling in that we extend the use of diffusion models to forecast sales. We consider five different specifications of diffusion models, including three time trend-only models (Gompertz, Logistic and Bass models) and two extended diffusion models with both time and GDP per capita as explanatory variables (extended Gompertz and Logistic models), and compare the sales forecasts from these models against those produced by three benchmark models that employ sales data directly: Exponential Smoothing, ARIMA and linear econometric models. In addition, in order to account for the potential impact of demand variations on the forecasting performance, we adopt the methodology of rolling forecasting to compare the performance of various models based on 5 different horizons. Intuitively, a model may perform well in terms of its forecasts over a fixed horizon, but if we change the length of that horizon, we may find that this is not the case (Fildes, 1992; Tashman, 2000). This is particularly important in the context of EMs where the market condition may be more unstable, so the rolling forecasting approach is employed to evaluate which model can provide the most robust forecasts. Specifically, we examine the mean and median values of various forecasting measures across 5 rolling horizons in order to select the best sales forecasting model.

Since three basic forms of diffusion models, which are Gompertz, Logistic and Bass models, have been reviewed in the methodology chapter previously, the next section of this chapter directly presents the proposed sales forecasting approach and the data that we use in this chapter. Section 3.3 discusses the model estimation results. The sales forecasting performance of the different models is then presented and compared in section 3.4. Finally, our conclusions are presented in the last section of the chapter.

3.2 Methodology and Data

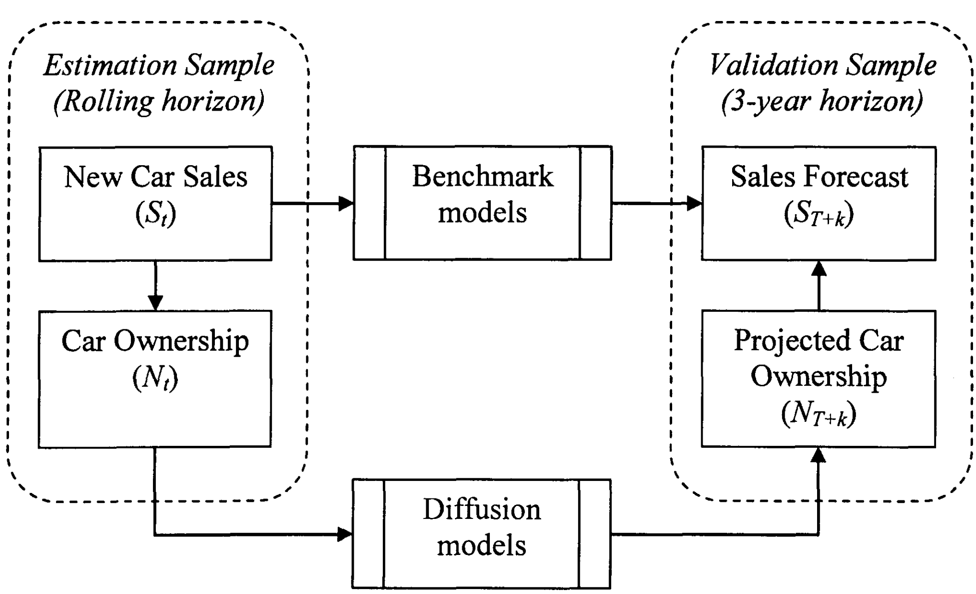
3.2.1 Conceptual model

Our approach that uses diffusion models to forecast car sales is depicted in Figure 3-2. Sales of new cars in the current period can be represented as the difference in stock of cars between the current and preceding period with some adjustments representing car scrapping. Thus, the relationship between new car sales S_t and car ownership level N_t can be derived as

$$S_t = N_t \cdot P_t - N_{t-1} \cdot P_{t-1} + A_t + e_t \tag{3-1}$$

where P_t is the population size so that $(N_t \cdot P_t)$ stands for total car stock at time t, A_t is the scrapping level at time t, and e_t is the error term. Because we are modelling the early stage of the emerging car market, where car scrapping volume is limited, the adjusting term A_t in equation (3-1) is approximated to zero.

Figure 3-2: Concept framework of forecasting sales using diffusion models



Therefore, car sales amount at time t is be re-expressed as

$$S_t = N_t \cdot P_t - N_{t-1} \cdot P_{t-1} + e_t \tag{3-2}$$

and thus car ownership level at time t can be expressed as

$$N_t = \frac{S_t + N_{t-1} \cdot P_{t-1} - e_t}{P_t} \tag{3-3}$$

When we have car sales data S_t, initial period car stock $(N_0 \cdot P_0)$ and population size P_t data, the historical car ownership levels can be derived through continuous iterations of equation (3-3). With the derived car ownership levels N_t, various diffusion models can be fitted in the T-period estimation horizon ($t = 1, 2, \ldots, T$), and then future car ownership can be extrapolated in a K-period validation horizon ($t = T+1, T+2, \ldots, T+K$). Finally, through iteratively applying equation (3-2) with the extrapolated car ownerships and population information, car sales forecasts can be derived over the validation horizon.

In this chapter, we use five diffusion model specifications, as summarised in Table 3-2, to model quarterly car ownership in China starting in 2002. The first three models are basic forms of Gompertz, Logistic and Bass models, and the other two are the extended specifications of the Gompertz and Logistic models that include both time and GDP variables. After parameter estimation, all these five models are used to predict car ownership levels in a 3-year horizon, which are further used to forecast car sales following the approach described above. It is worth mentioning that GDP information used in the validation sample is not the actual data but is also forecasted from the estimation sample by using the Holt's exponential smoothing method. This is because in a scenario where sales need to be forecasted, the future levels of GDP are unknown when one is using that variable as an input to the forecast model. Therefore, the car sales forecasts from the extended diffusion models are unconditional rather than conditional on future GDP level data.

Table 3-2: Summary of five diffusion models

Model	Definition of diffusion process
Gompertz	$N_t = M \cdot \exp\left(-\alpha \cdot \exp\left(-\beta \cdot t\right)\right)$
Logistic	$N_t = \dfrac{M}{1 + \alpha \cdot \exp(-\beta \cdot t)}$
Bass	$N_t = M \cdot \dfrac{1 - \exp\left(-(p+q)t\right)}{1 + (\frac{q}{p})\exp\left(-(p+q)t\right)}$
Extended Gompertz	$N_t = M \cdot \exp\left(-\alpha \cdot \exp\left(-\beta \cdot t + \gamma \cdot GDP_t\right)\right)$
Extended Logistic	$N_t = \dfrac{M}{1 + \alpha \cdot \exp(-\beta \cdot t + \gamma \cdot GDP_t)}$

In order to compare the forecasting performance of the diffusion-based models, three typical benchmark models are used to forecast car sales directly. Exponential smoothing and ARIMA models are estimated based on historical car sales. After calibrating all the available variables including seasonal effects, we specify following econometric model

$$S_t = \beta_0 + \beta_1 S_{t-1} + \beta_2 PRICE_t + \beta_3 GDP_t + e_t, \tag{3-4}$$

which includes one-period lagged sales, market average car price and GDP per capita as explanatory variables.

Three forecasting performance metrics are used to compare the forecasting accuracy of all competing models on the validation sample. These are *mean absolute deviation* (MAD), *mean absolute percentage error* (MAPE) and *root mean squared error*

(RMSE). Assuming A_t and F_t are actual and forecasted sales at time t respectively, three measures are defined as

$$MAD = \frac{1}{K}\sum_{t=T+1}^{T+K} |A_t - F_t|$$

$$MAPE = \frac{100}{K}\sum_{t=T+1}^{T+K} \left|\frac{A_t - F_t}{A_t}\right|$$

$$RMSE = \sqrt{\frac{1}{K}\sum_{t=T+1}^{T+K} (A_t - F_t)^2}$$

3.2.2 Data

We use secondary data that we have discussed in the previous chapter for estimating both diffusion models and the benchmark models in this chapter. Specifically, cars sales are regularly published since 2002 by the China Associate of Automobile Manufacturers (CAAM), from which we can easily obtain quarterly car stock in China by using year 2001 as the base year. Furthermore, we collect population data from China Statistical Yearbooks (2003-2009) published by National Bureau of Statistics of China (NBSC) to derive the car ownership level (N_t) that is measured as the number of cars owned by every 1000 people. Finally, quarterly data of GDP per capita is available on the website of NBSC, which is further deflated to be at constant 2001 price. It is worth noting that the quarterly GDP per capita demonstrates significant seasonality effects, so we extract the trend component of GDP per capita variable and use de-seasonalised GDP per capita in the extended diffusions as well as in the linear econometric models.

3.3 Model Estimation

3.3.1 Estimation of diffusion models

Nonlinear Least Squares (NLLS) instead of Ordinal Least Squares (OLS) method should be used to estimate the diffusion models to obtain the least biased parameters (Putsis & Srinivasan, 2000; Srinivasan & Mason, 1986). All 5 diffusion models are estimated in EViews 6. The parameters' starting values in the Bass model are set to be 0.03 for p and 0.42 for q, which are the average values of each parameter reported in Van den Bulte & Stremersch (2004) across several hundred of consumer durables. Furthermore, some attention is required here about our special treatment of the saturation level in the diffusion models. The requirement of stable stage data has been highlighted in the literature to estimate the saturation levels for every type of diffusion models (Dargay, et al., 2007; Kobos, et al., 2003; Srinivasan & Mason, 1986). An alternative way to address the problem of saturation levels in the developing markets is to assume a reasonable saturation level and then examine the resultant diffusion process (Button, et al., 1993; Chamon, et al., 2008; Wang, et al., 2006). Since we only have 8 years worth of car sales data from 2002 and car ownership in China is far from saturation, we are not able to estimate the saturation level and we instead compare the different models' performance based on different levels of assumed saturation threshold for car ownership. Furthermore, there is not much agreement regarding the saturation level of car ownership in China. Dargay & Gately (1999) assume an international saturation level of 620 cars per 1000 people for 26 countries, including China. Chamon et al. (2008) hold a higher saturation level assumption of 2 cars per 3 people to forecast car ownership of China and India in 2030. However, other studies report lower car ownership saturation level in China. Button et al. (1993) claim that a saturation range from 300 to 450 cars per 1000 people can reasonably be assumed for low-income countries due to high population density. In this research, we assume three different saturation scenarios, a high level of 600 cars per 1000 people in China, a middle level of 400 cars per 1000 people and a low level of 200 cars per 1000 people, and we further verify the sensitivity of the forecasts to these assumptions.

As we have already highlighted in the introduction section of this chapter, in order to compare the different models and how they fare in forecasting sales, we apply the rolling forecast methodology. In our study, the first estimation sample is from 2002Q1 to 2004Q4, and the following 3-year data is the corresponding validation sample. The rolling mechanism extends the estimation sample by every half year, followed by a 3-year forward validation sample. Thus, there are total 5 pairs of rolling horizons for estimation and validation purposes respectively, as shown in Table 3-3.

Table 3-3: Rolling horizons for model estimation and validation

Case	Estimation horizon	Validation horizon
1	2002Q1 to 2004Q4	2005Q1 to 2007Q4
2	2002Q1 to 2005Q2	2005Q3 to 2008Q2
3	2002Q1 to 2005Q4	2006Q1 to 2008Q4
4	2002Q1 to 2006Q2	2006Q3 to 2009Q2
5	2002Q1 to 2006Q4	2007Q1 to 2009Q4

Based on 5 different estimation samples and under 3 saturation scenarios, each diffusion model has 15 sets of parameter estimation results, as reported in Table 3-4.

Table 3-4: Parameter estimation result of all diffusion models

	Gompertz			Logistic			Bass		
Saturation level	α	β	R^2	α	β	R^2	p	q	R^2
Estimation sample (1): 2002Q1 to 2004Q4									
200	3.4695	0.0148	0.9958	30.7865	0.0485	0.9975	0.0011	0.0762	0.8042
400	4.1604	0.0121	0.9962	62.5207	0.0474	0.9976	0.0006	0.0748	0.8023
600	4.5650	0.0109	0.9964	94.2557	0.0471	0.9976	0.0004	0.0743	0.8016
Estimation sample (2): 2002Q1 to 2005Q2									
200	3.4722	0.0149	0.9973	30.7436	0.0483	0.9985	0.0012	0.0664	0.7714
400	4.1623	0.0122	0.9977	62.4117	0.0471	0.9985	0.0006	0.0649	0.7696
600	4.5664	0.0110	0.9978	94.0809	0.0468	0.9985	0.0004	0.0644	0.7691
Estimation sample (3): 2002Q1 to 2005Q4									
200	3.4812	0.0153	0.9971	30.8639	0.0488	0.9988	0.0012	0.0698	0.8522
400	4.1701	0.0125	0.9976	62.6215	0.0476	0.9989	0.0006	0.0681	0.8515
600	4.5737	0.0112	0.9978	94.3810	0.0472	0.9989	0.0004	0.0675	0.8513
Estimation sample (4): 2002Q1 to 2006Q2									
200	3.4956	0.0159	0.9959	31.1096	0.0499	0.9985	0.0012	0.0709	0.9053
400	4.1829	0.0129	0.9966	63.0707	0.0485	0.9987	0.0006	0.0690	0.9051
600	4.5859	0.0116	0.9969	95.0345	0.0481	0.9987	0.0004	0.0683	0.9050
Estimation sample (5): 2002Q1 to 2006Q4									
200	3.5128	0.0165	0.9947	31.3922	0.0509	0.9983	0.0012	0.0706	0.9292
400	4.1981	0.0133	0.9957	63.5790	0.0494	0.9985	0.0006	0.0684	0.9291
600	4.6002	0.0120	0.9961	95.7697	0.0490	0.9986	0.0004	0.0677	0.9290

To be continued on the next page

	Extended Gompertz				Extended Logistic			
Saturation level	α	β	γ	R^2	α	β	γ	R^2
Estimation sample (1): 2002Q1 to 2004Q4								
200	5.8848	0.0005	-1.9594	0.9985	93.3034	0.0184	-4.1148	0.9986
400	6.2082	0.0013	-1.4845	0.9985	175.3288	0.0195	-3.8270	0.9986
600	6.4724	0.0015	-1.2950	0.9985	257.8971	0.0198	-3.7358	0.9986
Estimation sample (2): 2002Q1 to 2005Q2								
200	5.1311	0.0040	-1.4544	0.9988	51.9992	0.0335	-1.9594	0.9987
400	5.5043	0.0044	-1.0410	0.9988	97.7047	0.0345	-1.6712	0.9987
600	5.7872	0.0044	-0.8827	0.9988	143.7394	0.0349	-1.5804	0.9987
Estimation sample (3): 2002Q1 to 2005Q4								
200	5.0190	0.0046	-1.3727	0.9992	51.1445	0.0339	-1.8988	0.9992
400	5.4102	0.0048	-0.9773	0.9992	96.5051	0.0348	-1.6262	0.9992
600	5.7012	0.0048	-0.8273	0.9992	142.1835	0.0351	-1.5409	0.9992
Estimation sample (4): 2002Q1 to 2006Q2								
200	5.1104	0.0041	-1.4400	0.9995	56.8127	0.0309	-2.2914	0.9995
400	5.4958	0.0044	-1.0359	0.9995	107.1780	0.0318	-2.0180	0.9995
600	5.7859	0.0044	-0.8823	0.9995	157.9104	0.0321	-1.9328	0.9995
Estimation sample (5): 2002Q1 to 2006Q4								
200	4.9929	0.0048	-1.3528	0.9997	54.7913	0.0320	-2.1554	0.9997
400	5.3967	0.0049	-0.9676	0.9997	103.6165	0.0328	-1.8912	0.9997
600	5.6943	0.0048	-0.8224	0.9997	152.7959	0.0331	-1.8092	0.9997

Note: all estimated parameters are significant at 1% level.

All estimated models have good model fits. Specifically, Gompertz and Logistic models (both simple and extended specifications) achieve R^2 values higher than 0.99 in every horizon. The goodness of fit of Bass model improves with more data points included in the model estimation. The estimated parameters of all diffusion models are statistically significant at 1% level with the expected signs across 5 rolling horizons. Regarding the stability of the estimates, we find that the different saturation levels affect model estimates for the same estimation sample, particularly α parameter in Gompertz and Logistic models and p in Bass model. When comparing the stability across rolling horizons, we find that the simple diffusion models tend to have more stable estimates than the extended ones, which suggests that the parameter variation increases with additional variables included in the model. Also, the stability of extended diffusion models is improved when more data points are used in the estimation. More importantly, since our main research purpose is to select the best sales forecasting model, the better estimation stability of simple diffusion models may not ensure their better forecasting capability. We are going to compare the forecasting performances of various models in the rolling validation samples in next section.

3.3.2 Estimation of benchmark models

To evaluate how useful these diffusion models are, we compare them against some benchmark models. These models directly forecast future sales in the different rolling horizons compared to the diffusion models where we use car ownership data and then obtain car sales forecasts. We employ three simple time series models estimated in a forecasting software package of ForecastPro. The three models are Exponential Smoothing with the linear trend and additive seasonality (ES-LA), Exponential

Smoothing with the linear trend and multiplicative seasonality (ES-LM), and an ARIMA model. The specifications of ARIMA model are SARIMA(0,1,0)*(0,1,0) for the first 3 horizons and SARIMA(0,1,0)*(1,0,0) and SARIMA(0,1,0)*(1,1,0) respectively for the last 2 rolling horizons. The last benchmark model, a linear econometric model in equation (3-4), is estimated using Ordinary Least Squares (OLS) method in EViews 6. Residual normality, correlation and heteroskedasticity tests were performed after OLS estimation in every rolling horizon case and no significant violations of OLS assumptions were detected.

3.4 Sales Forecasting

3.4.1 Quarterly sales forecasting

Based on the car ownership forecasts in the validation sample, the quarterly car sales forecasts from the diffusion models are generated using the equation (3-2). As we have discussed above, all benchmark models provide car sales forecasts in all 5 horizons. In order to evaluate the forecasting performance stability across different horizons, we report the mean and median values of three performance measures (MAD, MAPE and RMSE) in all rolling horizons for all models (Table 3-5). Therefore, the best forecasting model should produce the smallest average and median values of various performance measures across the validation samples in distinct horizons. It is worth noting that the median value here is more important in the dynamic market context, since the error measures may suffer high degree of variation and skewed distribution. When using the median of various measures across rolling horizons, we can achieve a clearer conclusion that the extended Logistic model is the best model in forecasting car sales in China. We present the comparison of the quarterly rolling forecasting results as follows.

Table 3-5: 3-year ahead quarterly rolling forecasting performance

		MAD		MAPE		RMSE	
Model	Saturation	Average	Median	Average	Median	Average	Median
Gompertz	200	4.18E+05	3.93E+05	30.48	30.54	4.62E+05	4.17E+05
	400	3.74E+05	3.53E+05	27.29	27.33	4.17E+05	3.80E+05
	600	3.54E+05	3.38E+05	25.88	25.89	3.96E+05	3.64E+05
Logistic	200	2.21E+05	2.10E+05	16.22	15.91	2.61E+05	2.45E+05
	400	1.97E+05	1.83E+05	14.54	14.16	2.31E+05	2.17E+05
	600	1.89E+05	1.76E+05	14.02	13.58	2.21E+05	2.07E+05
Bass	200	1.26E+05 [b]	1.22E+05 [b]	9.76 [b]	10.30	1.79E+05	2.00E+05
	400	1.45E+05	1.43E+05	11.10	11.58	2.03E+05	2.31E+05
	600	1.52E+05	1.50E+05	11.57	12.01	2.12E+05	2.41E+05
Extended Gompertz	200	2.20E+05	2.14E+05	16.12	15.32	2.75E+05	2.62E+05
	400	1.94E+05	1.94E+05	14.35	13.03	2.41E+05	2.33E+05
	600	1.84E+05	1.90E+05	13.69	12.13	2.27E+05	2.20E+05
Extended Logistic	200	1.39E+05	1.26E+05	10.53	10.11 [a]	1.68E+05 [a]	1.46E+05 [a]
	400	1.31E+05 [a]	1.21E+05 [a]	10.01 [a]	9.55 [a]	1.60E+05 [a]	1.53E+05 [a]
	600	1.28E+05 [a]	1.23E+05 [a]	9.87 [a]	9.25 [a]	1.58E+05 [a]	1.58E+05 [a]
ES_LA		1.67E+05	1.96E+05	12.18	10.35	2.10E+05	2.21E+05
ES_LM		1.70E+05	1.91E+05	12.37	11.08	2.12E+05	2.16E+05
ARIMA		2.17E+05	1.85E+05	16.64	12.62	2.67E+05	2.41E+05
Linear Regression		1.52E+05	1.67E+05	11.04	10.86	1.94E+05	2.02E+05

Note: [a] Extended Logistic model outperforms all other models with smaller mean and median values of measures across 5 rolling horizons. [b] Bass Model has smallest average and median of MAD as well as median of MAPE in low saturation scenario.

- Comparison of the basic diffusion models

For the three basic diffusion models, the Bass model is best at forecasting car sales, followed by the Logistic and the Gompertz models respectively. The average and median values of MAPE from Gompertz model range from 25% to 30% depending on the different saturation levels, while the MAPE of Bass model on average is only 10% to 12%. The similar performance differences between Gompertz and Bass models are identified when measured by MAD and RMSE. In addition, the Bass model also outperforms the Logistic model in most measurements, except when measured with the median of RMSE and in the scenarios of middle or high saturation levels.

Table 3-6: Quarterly forecasting improvement from basic to extended diffusions

	MAD reduction rate (%)			MAPE reduction rate (%)			RMSE reduction rate (%)		
Saturation	200	400	600	200	400	600	200	400	600
Gompertz vs. extended Gompertz	47.35	48.07	47.94	47.09	47.41	47.10	40.39	42.13	42.63
Logistic vs. extended Logistic	36.99	33.53	32.08	35.08	31.15	29.57	35.54	30.77	28.48

Note: The reduction rates are measured as the forecasting error decreasing percentage from the basic to the corresponding extended diffusion specifications

- Comparison of the different diffusion model specifications

When comparing the basic and extended diffusion models, the extended specifications are substantially better than the basic ones. Table 3-6 summarises the improvement ratios of average measures achieved in quarterly sales forecasting from the basic to the extended diffusions. The extended Gompertz model improves the car sales forecasting performance by 40%-50% compared with the basic specification. The extended Logistic model improves the forecasting performance by about 30% compared to its corresponding basic model. In addition, the extended Logistic model is better than the extended Gompertz model under all saturation scenarios. Furthermore, the extended Logistic model clearly outperforms the Bass model under both middle and high saturation level scenarios with smaller average and median values of all measures. For the low saturation level scenario, although the median of MAD is higher for the extended Logistic model, it achieves smaller median values in both MAPE and RMSE than the Bass model. Therefore, the best model amongst all the diffusion models to forecast car sales is the extended Logistic model with both time and GDP per capita variables.

- Comparison between diffusions and benchmark models

Amongst all the benchmark models, the ARIMA model is the worst model and the two exponential smoothing models have similar forecasting performances. Although the linear econometric model has a better forecasting performance than all the other benchmarks and some of the diffusion models, the extended Logistic model still outperforms it with both smaller mean and median values of three error measures. Therefore, the extended Logistic model is the best model to forecast car sales.

The importance of using rolling forecasting to evaluate which model to use is best demonstrated by illustrating what would happen if we were to use a fixed sample. As shown in Figure 3-3, the extended Logistic model forecasting performance is more stable than the linear econometric model. In the validation samples of 2005Q1-2007Q4 and 2007Q1-2009Q4, the linear econometric model is the worst model with significantly higher forecasting errors. However, if all models were just estimated based on data from 2002Q1 to 2005Q4 respectively, one would opt for the linear econometric model because it has the smallest MAD and MAPE compared to the extended specification of the logistic model and almost the same RMSE in the corresponding validation horizon (2006Q1-2008Q4). Therefore, in order to achieve a more robust conclusion on which model to use, it is necessary to examine different horizons for model estimation and validation respectively.

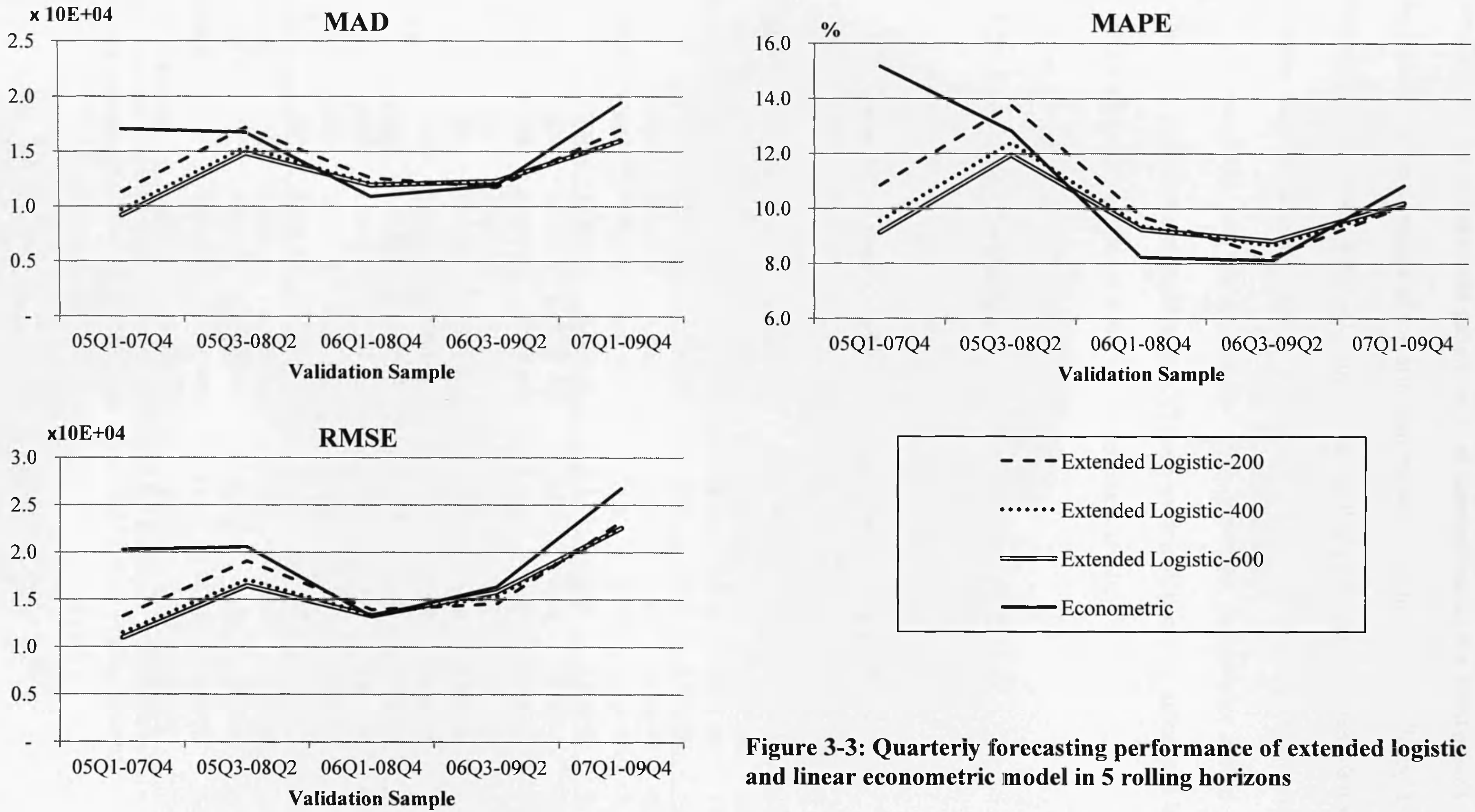


Figure 3-3: Quarterly forecasting performance of extended logistic and linear econometric model in 5 rolling horizons

We also note that the different saturation level assumptions do not substantially affect the forecasting performance of the diffusion models. As illustrated in Figure 3-3, the extended logistic model achieves a similar forecasting performance when we test it at the three different saturation levels and the results for the middle and high saturation levels respectively are not substantial. This indicates that 3-year ahead sales forecasting performance based on the diffusion models will not be influenced by the different assumptions that we make on the saturation levels.

3.4.2 Annual sales forecasting

Table 3-7: 3-year ahead annual rolling forecasting performance

		MAD		MAPE		RMSE	
Models	Saturation	Average	Median	Average	Median	Average	Median
Gompertz	200	1.66E+06	1.57E+06	30.71	31.29	1.78E+06	1.64E+06
	400	1.49E+06	1.41E+06	27.47	27.65	1.59E+06	1.49E+06
	600	1.41E+06	1.35E+06	26.03	26.22	1.51E+06	1.43E+06
Logistic	200	8.51E+05	8.14E+05	15.83	15.47	9.35E+05	8.24E+05
	400	7.21E+05	6.83E+05	13.52	12.58	8.03E+05	7.11E+05
	600	6.78E+05	6.44E+05	12.76	12.04	7.61E+05	6.76E+05
Bass	200	4.32E+05	4.24E+05	8.05[b]	8.99	5.84E+05	6.59E+05
	400	5.39E+05	5.16E+05	9.78	10.78	6.94E+05	7.88E+05
	600	5.75E+05	5.47E+05	10.36	11.25	7.35E+05	8.31E+05
Extended Gompertz	200	8.22E+05	8.08E+05	15.23	15.70	9.69E+05	9.90E+05
	400	7.03E+05	6.86E+05	13.12	13.40	8.15E+05	8.70E+05
	600	6.52E+05	6.34E+05	12.22	12.43	7.52E+05	8.17E+05
Extended Logistic	200	4.29E+05 [a]	3.71E+05 [a]	8.21	7.49 [a]	4.86E+05 [a]	4.46E+05 [a]
	400	3.91E+05 [a]	3.16E+05 [a]	7.58 [a]	6.89 [a]	4.62E+05 [a]	3.70E+05 [a]
	600	3.84E+05 [a]	3.16E+05 [a]	7.46 [a]	6.54 [a]	4.60E+05 [a]	3.98E+05 [a]
ES_LA		5.67E+05	6.92E+05	10.50	9.13	6.77E+05	8.35E+05
ES_LM		5.70E+05	6.94E+05	10.51	9.20	6.80E+05	8.20E+05
ARIMA		7.79E+05	7.20E+05	15.18	13.06	8.95E+05	7.58E+05
Linear Regression		4.66E+05	6.03E+05	8.45	8.41	5.77E+05	7.62E+05

Note: [a] Extended Logistic model outperforms all other models. [b] Bass model is the best in this case only (low saturation scenario and measured with average value of MAPE across 5 rolling horizons).

We use quarterly sales data but it is very likely that strategic decisions relating to car sales will be based on annual sales figures. Thus we further compare the annual sales forecasts based on a 3-year ahead rolling horizons, which are derived by combining every 4 quarters' prediction in a 3-year validation horizon. The corresponding forecasting performances of the different models are summarised in Table 3-7, where we also present both average and median values of three error measures. Following the same comparisons for the quarterly sales forecasts, we find that the extended Logistic model consistently achieves the smallest median values of forecasting errors in any saturation scenario. Again here we find that the Logistic model with the extended specification is the best model.

3.5 Conclusion

Aggregate car sales forecast is an important strategic market insight for car manufacturers when they have to assess the demand potential of any country's car market. Data on car sales may be limited or not widely available if the product is new in a market. In such a situation, we propose that market researchers consider the use of diffusion models to forecast car sales in an emerging market by using data on car ownership levels which may be more readily predictable whereas sales data is limited. We use the context of China where car adoption is still low compared to other developed more mature car markets. A more distinctive contribution of this chapter is that we have explored how diffusion models can be used to forecast car sales as opposed to simply forecasting car ownership levels.

Three well-known diffusion models are initially estimated (Gompertz, Logistic and Bass models) and then we extend their basic specifications by including a trend effect and the effect of GDP. The significant effect of income on the car diffusion and sales from our results corroborates with other findings in the literature (Dargay & Gately, 1999; Dargay, et al., 2007; Dargay, 2001). The superior forecasting performance of the extended Logistic model compared to the econometric model suggests that the diffusion models are better able to cope with the non-linearity characteristic of the market expansion in China. This is not surprising as previous research does stress on the importance of using models that suit the market dynamics (Fildes, et al., 1998; Fildes, et al., 2008).

The sales forecasting performances of the diffusion models are compared against the benchmark models (exponential smoothing, ARIMA and linear econometric model). We show that it is important to use a rolling forecast horizon approach instead of using a fixed validation sample to compare the models. We demonstrate that the rolling forecasting approach is particularly significant when it comes to choose a more robust model that can cope with the rapidly changing environment in the EMs such as China.

CHAPTER 4.

THE IMPORTANCE OF UNDERSTANDING LOCAL CONSUMER BEHAVIOUR: THE CHINESE CAR MARKET

4.1 Introduction

For any car manufacturer who wants to sell its products in China, it should have a clear understanding of the local consumer behaviour and particularly the differences in comparison to the established markets. Such a comparison provides car manufacturers with valuable insights to develop competitive products or strategies that meet local consumers' demand. It is particularly important for multinational car manufacturers who want to enter the Chinese car market or further benefit the growth potentials of this market. Their success or failure in this market largely depends on how they understand local consumer preferences for cars as well as the culture and then effectively deliver the "right" consumer products. In the car market, for example, the failure of Fiat in China from 2003 to 2007 is believed mainly due to its inability to respond to the Chinese consumer preferences[10]. Another case in the Chinese car market is Toyota in 2003, who clumsily advertised its SUV model (Prado) with two bowing stone lions (Doctoroff, 2005, p. 104), which were easily connected by the

[10] "Fiat 'must learn from past mistakes'", China Daily, 13/07/2009, http://www.chinadaily.com.cn/cndy/2009-07/13/content_8418360.htm

Chinese consumers with the similar stone lions on the Marco Polo bridge where the Japanese troops started their full-scale invasion in China in 1937. Therefore, the advertisement aroused immediate resentment from the Chinese consumers. Consequently Toyota had to formally apologise for that advertisement[11] and even changed the Chinese name of Prado one year later. When modelling car market demand, however, the existing literature extensively focuses on the developed markets and pays little attention to consumer behaviour in the EMs such as China. Thus, using the car market as an example, this chapter aims to bridge this gap and to shed some light on local consumer behaviour in China and its differences in comparison with in the more mature markets.

At aggregate level, researchers have conducted cross-country comparison of new product diffusion processes to differentiate key factors that determine new product diffusions in different countries (Meade & Islam, 2006; Yalcinkaya, 2008). For example, Gatignon, et al. (1989) find that the innovative effect (p) and imitative effect (q) of Bass model in a country could be affected by its cultural characteristics including the role of women in society (proportion of women in the workforce), cosmopolitanism (communication with foreign countries by post or by travel) and mobility level (car ownership). Takada & Jain (1991) show that a lagged introduction of new product in a country may lead to an accelerated adoption, while Tellefsen & Takada (1999) find that p and q of Bass diffusion can also be influenced by the levels of mass media in countries. Through comparing new product diffusions in developed and developing countries, Talukdar, et al. (2002) find that the average penetration potential of new products in developing countries is only one third of that in the

[11] "Toyota apologizes for 'humiliating' ads", 05/12/2003, China Daily, http://www.chinadaily.com.cn/en/doc/2003-12/05/content_287571.htm

developed ones and the adoption rate is slower in the developing countries. They also find that the stronger economic and population growth in the developing countries would indicate a significant rise of the penetration levels in the developing countries. More recently, Van den Bulte & Stremersch (2004) conduct a meta-analysis based on 52 different consumer durables across 28 countries. They show that the *q/p* ratio of Bass model is positively associated with income inequality in a country and the collectivistic cultures tend to have a higher *q/p* ratio than the individualistic ones.

In this chapter, we focus on two basic types of models at disaggregate level. The first one is the car ownership model, which mainly deals with the choice of different numbers of cars owned by each household. The second model concerns car type choice, which investigates how households select their specific types of cars out of all available vehicle types in the market. The more complex models, such as the vehicle transaction model and the joint model of vehicle holding and use, are more appropriate for the developed markets where there are longer histories of household car ownership and more individual level data. See Bunch & Chen (2008), De Jong, et al. (2004) and Potoglou & Kanaroglou (2008a) for the relevant reviews on these models.

Our research approach in this chapter consists in developing models at the disaggregate level to investigate consumer preferences in the Chinese car markets, and then comparing them with the typical effects found in more mature markets. Furthermore, it is also important for the car manufacturers to explore context specific features in a new market that they must take into account. Some preliminary analysis highlights that consumer knowledge about cars could be fairly low in China as the average car ownership level is much lower than the world average level and the car ownership history in China is short. Thus, using consumer knowledge as an example, we show the importance of accounting for the context-specific variables when understanding the local consumer behaviour. In addition, the importance of the consumer knowledge is further demontrated through investigating its effect on the consumers' future purchase intentions.

The remainder of this chapter is organised as follows. Section 4.2 reviews the literature relevant to car ownership and car type choices, and presents the key explanatory variables in our models and their typical effects in the developed markets. Section 4.3 discusses the literature of consumer knowledge in general followed with a brief introduction of the data used in this chapter. The modelling approaches and the estimation results are discussed in detail in section 4.5 and 4.6 for the two models respectively. Section 4.7 presents a segmentation analysis to highlight the different elasticity effects of consumers with different knowledge levels, which is followed by an additional section investigating the effect of consumer knowledge on purchase intentions. The last section of this chapter gives a summary of main findings and important implications.

4.2 Review of Literature and Key Explanatory Variables

In this section, we first review the modelling methodology for car ownership and car type choice models respectively. After the review for each model, we describe the explanatory variables in our model and their typical effects found in models in the developed car markets. This establishes our hypotheses about whether consumers in China behave differently from those in the developed markets and more specifically how they differ.

4.2.1 Car ownership model: methodology review

The car ownership model is mainly concerned with the households' decision about the number of cars to own, including whether to own a car or not. The car ownership model is the simplest model at the disaggregate level to investigate automobile demand (Bunch & Chen, 2008).

Table 4-1 summarises the recent empirical studies of car ownership models that we reviewed, where we compare their market context (data sources), sample sizes, dependent variables and model specifications. The key explanatory variables will be discussed in detail with their typical effects in the next subsection. It is not surprising that most studies, except Li, et al. (2010), are based on the developed markets. It is worth noting that Li, et al. (2010) only investigate the car ownership models in two big cities in China, Beijing and Chengdu, and these may not be representative of the whole Chinese market.

Table 4-1: A summary of recent car ownership models

No.	Study	Survey Area or Data Source (year of data collection)	Sample size	Dependent Variable	Model[a]
1	Bhat & Pulugurta (1998)	Boston, U.S. (1991), The Bay Area, U.S. (1990), Puget Sound, U.S. (1991), The Dutch national dataset (1987)	2,500 (Boston); 2,500 (Bay); 1,231 (Puget Sound); 1,307 (Dutch)	0, 1, 2, 3, 4 cars (U.S.) 0, 1, 2 cars (Dutch)	MNL* vs. ORL
2	Ryan & Han (1999)	Oahu, Honolulu (1995)	4,060	0, 1, 2, 3+ cars	MNL
3	Hess & Ong (2002)	Multnomah County, Oregon, U.S. (1992)	1,132	0, 1, 2, 3+ cars	ORL
4	Chu (2002)	New York (1997/1998)	3,397	0, 1, 2, 3+ cars	ORP
5	Nolan (2002)	Dublin, Ireland (1994/1995)	2,148	0, 1+ cars	Binary Probit
6	Kim & Kim (2004)	Nationwide Personal Transportation Survey, U.S. (1995)	42,033	0, 1, 2, 3+ cars	ORL
7	Giuliano & Dargay (2006)	The US National Personal Transportation Survey (1995); the British National Transport Survey (1995/1997)	10,000 each	0, 1, 2+ cars	ORP
8	Whelan (2007)	Family Expenditure survey, G.B. (1971-1996); National travel survey, G.B. (1991)	46,137	0, 1+ cars; 1, 2+ cars; 2, 3+ cars	Binary Dogit
9	Bhat & Guo (2007)	Alameda County, San Francisco Bay area, U.S. (2000)	2,954	No. of owned cars	Mixed ORL
10	Matas & Raymond (2008)	Spanish Household Surveys (1980, 1990, 2000)	23,696 (80s); 20,927 (90s); 28,963 (00s)	0, 1, 2, 3+ cars	ORP
11	Potoglou & Kanaroglou (2008b)	Hamilton, Canada (2005)	774	0, 1, 2, 3+ cars	MNL* vs. ORL
12	Potoglou & Susilo (2008)	Baltimore Metropolitan area, U.S. (2001); Dutch National Travel Survey (2005); Osaka Metropolitan area (2000)	3,496 (Baltimore), NA (Dutch and Osaka)	0, 1, 2, 3+ cars	MNL*, ORL & ORP
13	Li, et al. (2010)	Chengdu (2005) and Beijing (2006) in China	1,001 (Chengdu); 1,200 (Beijing)	0, 1+ cars	Binary Logit
14	Nolan (2010)	Living in Ireland Survey (1995 to 2001)	18,441	0, 1+ cars	Binary Probit

Note: [a] ORL is Ordered Logit model and ORP is Ordered Probit model;
* It indicates that the MNL model can perform better than Ordered models in the same research.

The dependent variable in the car ownership models is defined as the number cars owned. Due to the high car ownership levels in the developed markets, most studies define their dependent variable to be zero, one, two cars and even more. Consequently, the corresponding models employed in these studies are Multinomial Logit (MNL), Ordered Logit (ORL) or Ordered Probit (ORP) model. When some samples have a very small ratio of multiple car ownerships (Li, et al., 2010; Nolan, 2002) or the exact numbers of cars owned by the surveyed households are unavailable to the researchers (Nolan, 2010), a binary model (binary Logit or Probit) is usually employed with a binary dependent variable to indicate whether the household own at least one car or not. Whelan (2007) is an exception, because he designs three binary models instead of a multinomial model to study the multiple car ownerships in the UK. In addition, in the multinomial choice situations, some studies compare ordered (i.e. ORL or ORP) and non-ordered (i.e. MNL) specifications and they consistently find that the MNL model outperforms the ordered model with either better estimation convergence (Potoglou & Kanaroglou, 2008b; Potoglou & Susilo, 2008) or more accurate prediction in validation samples (Bhat & Pulugurta, 1998).

4.2.2 Car ownership model: key explanatory variables

Given our primary research interests in the cross-market comparison of consumer behaviour, we discuss the key explanatory variables included in our car ownership model and their typical effects in the developed markets. The selection of explanatory variables in our car ownership model is mainly based on the literature review. Specifically, we include three categories of explanatory variables usually used by empirical studies of the car ownership model: residential location of the household, alternative transport modes accessible to the household and demographic characteristics.

Table 4-2: A summary of explanatory variables of car ownership model and their effects in developed markets

Explanatory variable	Definition in our model	Typical effect in the developed markets
• Residential location	Category variable • Live in urban area of cities • Live in suburban area of cities • Live in cities without indicating whether urban or suburban areas • Living in towns or non-city areas (reference category)	Linear effect of decreasing car ownership level with the rising density of residential area
• Alternative transport modes	Three dummy variables • Public transport • Bike, E-bike or motorcycle • Company shuttle	–
• Consumer knowledge about cars and car market	Category variable: good knowledge, basic knowledge, no knowledge (reference category)	Not available
• Income	Household income in year 2009 (measured at 10,000 RMB), and its squared term	+ for the income, – for the income squared
• Family size	Number of family members	+
• Children	Number of children	+
• No. of working adults	No. of employed family members	+
• Drivers	No. of licensed drivers divided by the household size	+
• Age of head	Age of household head	+ for working ages, – for retired ages
• Gender of head	Dummy variable (male head = 1, otherwise = 0)	+ for the male head

Note: + stands for the positive effect on car ownership, – stands for negative effect on car ownership.

Furthermore, we include consumer knowledge as a context-specific variable in our car ownership level to explore the potential association between households' car ownership and their knowledge about cars or car market. Such association has not been investigated in the developed car markets, but could be important in the context of emerging market such as China, where local consumers have a shorter car adoption history and thus the whole market in the current period is still during the early stage of knowledge diffusion. Table 4-2 summarises all explanatory variables in our car ownership model and their typical effects in other car markets.

- *Residential location*

Household residential location is widely used in car ownership models to investigate the influence of urbanisation and household accessibility on the car ownership (Bhat & Pulugurta, 1998; Chu, 2002; Giuliano & Dargay, 2006; Kim & Kim, 2004; Li, et al., 2010; Nolan, 2010; Potoglou & Susilo, 2008; Whelan, 2007). The typical finding in the developed markets is that the denser area the households live in, the less likely they are to own cars. For example, by using the category of households living in rural areas as the reference, Bhat & Pulugurta (1998) find that households in urban and suburban areas have less propensity to own cars and the likelihood decreases along with the increase in residential density. The underlying reason is that the residential density is an important proxy of households' accessibility to public transport and local services. A similar relationship is identified by Whelan (2007), who defines five types of residential locations in the UK based on population density. In his study of car ownership, he finds that households in the Greater London and metropolitan districts have the lowest car ownership likelihood followed by those in other urban areas compared to those in rural areas.

In our survey, we collected information about which city the households are living in as well as the post code, which is then used to define their residential location. In our sample, some respondents only told us the city name without the post code, so we do not know whether they are living in the urban or suburban areas of the city and thus these households are classified into a separate group of households living in cities without knowing whether in urban or suburban areas. Therefore we finally define four categories of household residential locations: urban areas of cities, suburban areas of cities, cities but not knowing whether in urban or suburban areas, and other remote areas such as towns or non-city regions. One-way analysis of variance (ANOVA) is conducted to test the income differences between households in these categories. The P-value is 0.263 and thus we do not have sufficient evidence to show that the incomes differ across households in these 4 different residential locations.

- *Alternative transport modes*

The substitution effects are also explored in the car ownership model through incorporating the households' accessibility to alternative transport modes. Intuitively, if there is easy convenient access to other transport modes such as buses or the underground, households are less likely to own cars. Normally there are two different approaches to accommodate this factor in the empirical studies. Some studies ask for the households' own assessment on the convenience level or quality of public transport (Hess & Ong, 2002). It is also common that the availability of other transport modes is measured by the distance or time to get to the nearby public transport (Bhat & Guo, 2007; Kim & Kim, 2004; the Chengdu model in Li, et al., 2010) or number of public transport stops within a short distance (such as 500 meters) from home (Giuliano & Dargay, 2006; Hess & Ong, 2002; Potoglou & Kanaroglou, 2008b). In addition, Li, et al. (2010) also examine the substitution effect of households who own bikes or motorcycles on car ownerships. These studies indicate that the convenient access to the alternative transport modes does discourage households to own cars.

In our study, we explore whether the surveyed households have frequent access to three types of alternative transport modes typically available in China. The first one is public transports, including buses, undergrounds, light rails and trains. The second category includes bikes, electric bikes/mopeds or motor cycles, which are usually ridden by the individuals. The last one is company shuttle buses usually provided by employers to take employees to and from work. If there is a substitution effect as found in the literature, we would expect that the households with the frequent access to any of these alternative transport modes have a lower car ownership level in China.

- *Consumer knowledge*

In the context of China, due to the short history of the car market, cars generally represent a new product concept for most Chinese consumers and they are still learning about cars. Therefore, it is important to investigate whether the consumers' knowledge level is positively associated with their car adoption behaviour. If so, car manufacturers as well as governments might be able to influence market demand or consumer adoption behaviour through helping consumers know more and better about cars and the car market.

Given that most consumers have no previous experience (expertise) about how to use cars in China, we define the consumer knowledge mainly based on their familiarity about cars. In our survey, the respondents were asked to assess their own knowledge about cars and the car market. Four different levels of familiarity were available for them to select: unfamiliar, basic knowledge, familiar and very familiar. In our sample, most respondents think they are unfamiliar with cars (about 30% of respondents) or have a basic knowledge only (about 47% of respondents), which corroborates our expectation that the general consumer knowledge level in China is not high due to the short household car ownership history. Only 18% of our respondents think they are familiar with cars and the remaining 5% of respondents are very familiar with cars and car markets. So we combine those respondents who are familiar or very familiar with cars and get three categories of consumer knowledge levels: good, basic and no knowledge about cars. The last category is used as the reference category. Therefore, our hypothesis in the car ownership model is that whether consumers with the better knowledge are more likely to be car owners in China.

Table 4-3: A crosstab of consumer knowledge and annual income

	Knowledge levels			
Annual Income (RMB)	No	Basic	Good	Grand Total
<100k	104	131	52	287
100k-190k	29	61	39	129
190k-300k	15	37	20	72
>=300k	10	16	10	36
Grand Total	158	245	121	524

Table 4-3 presents a crosstab of consumer knowledge and annual household income. In general, there is a generally positive association between these two factors. On the one hand, in the low income group with annual incomes less than 100,000 RMB, most households have limited or even no knowledge about cars. On the other hand, in the mid to high income groups (from 100,000 to 300,000 RMB), there are more consumers with good knowledge than those without knowledge about cars at the same income level. In addition, we conduct an ANOVA test for the income across different knowledge levels. The P-value of 0.019 indicates an association existing between the income and the consumer knowledge. However, we also notice that all income groups are consistently dominated by consumers with the basic knowledge level. In addition, the Pearson correlation coefficient between the knowledge and income is only 0.123, if the knowledge variable is encoded as an ordinal variable from one to three, which indicates there is no strong multicollinearity between these two variables, so that we can include both of them in the model.

- *Demographic variables*

Demographic variables that are investigated in our car ownership model include household income and its squared term, number of working adults, household types, number of licensed drivers, and age and gender of the household head. We briefly review their typical effects in the literature.

- Household income

It is not surprising that income is one of the most important factors that influence household car adoption decisions. Household income is consistently found to have a significant positive effect on the household car ownership (Bhat & Guo, 2007; Bhat & Pulugurta, 1998; Chu, 2002; Giuliano & Dargay, 2006; Hess & Ong, 2002; Kim & Kim, 2004; Li, et al., 2010; Nolan, 2010; Potoglou & Kanaroglou, 2008b; Potoglou & Susilo, 2008; Ryan & Han, 1999; Whelan, 2007). Besides, by including the square of income additional to the income variable, Nolan (2010) shows a nonlinear relationship between household income and car ownership in Ireland, which implies that the Irish car market is approaching the saturation level[12]. Since the Chinese car market is expected to have huge potentials to grow, the income effect can also be investigated by including both income and income squared in the car ownership model.

- Number of working adults

[12] According to the World Development Indicators (WDI) of the World Bank, every 1000 Irish people own 451 cars.

The number of working adults in each household is another frequently employed demographic variable in the car ownership model. The typical finding is that the more employed adults, the higher probability for the households to own more cars, which implies that the working adults in the developed economies heavily rely on private cars to independently travel to their working places (Bhat & Guo, 2007; Bhat & Pulugurta, 1998; Kim & Kim, 2004; Matas & Raymond, 2008; Potoglou & Kanaroglou, 2008b; Potoglou & Susilo, 2008; Whelan, 2007).

- Number of drivers

The number of licensed drivers in the households has also been taken into account in empirical studies of car ownership models. The variable is usually defined as a ratio of the number of drivers to the household size and it is typically found to have a significant positive effect for households to own cars (Chu, 2002; Kim & Kim, 2004; Potoglou & Kanaroglou, 2008b; Whelan, 2007). In the developed markets, there might be a two-way effect between the number of drivers and household car ownership level, but in the context of Chinese car market, most consumers are first time buyers of cars and they usually learn driving before buying cars. So we include it as an explanatory variable of the car ownership model.

- Age and gender of the household head

Demographic information about the household heads mainly includes the age and gender. Generally, compared with the households with young heads in their 20s, the households with the mature heads are more likely to adopt cars (Matas & Raymond, 2008; Nolan, 2002, 2010). Other studies also find that the households with senior or retired heads are less likely to own cars than working households (Giuliano & Dargay, 2006; Matas & Raymond, 2008). In terms of the gender, it is typically found that in the developed markets either the households with male heads have the higher likelihood to adopt cars (Hess & Ong, 2002; Matas & Raymond, 2008), or the households with female heads are less likely to buy cars (Nolan, 2002).

- Household type

Household types reflect different home structures or household life stages, which are also believed to have important influence on household car ownership. Sometimes, the household types are jointly defined by number of children in the household with the household size (Kim & Kim, 2004; Potoglou & Kanaroglou, 2008b; Potoglou & Susilo, 2008; Whelan, 2007). It is also common that the number of children and family size are separately included in the car ownership model (Giuliano & Dargay, 2006; Hess & Ong, 2002; Kim & Kim, 2004; Nolan, 2002, 2010; Ryan & Han, 1999). In general, empirical studies find that the households with children and/or more members are more likely to own cars. We account for the number of children and family size separately in our car ownership model.

4.2.3 Car type choice model: methodology review

There are a number of empirical studies that examine the key factors affecting households' choices of different types of cars, as summarised in Table 4-4. We focus on recent empirical studies here and for more early research on car type choices, please refer to Table 1 in Choo & Mokhtarian (2004). It is worth noting that most car type choice studies are also conducted in the context of developed markets and particularly in the US. The definitions of dependent variable in the vehicle type choice models largely depend on classifications of the vehicles. The common approach classifies the different vehicles into a certain number of categories and investigates the households' choice among these categories (Cao, et al., 2006; Choo & Mokhtarian, 2004; Lave & Train, 1979). Alternatively, a more detailed approach directly employs the vehicle make and model as the dependent variable (Mannering & Mahmassani, 1985; McCarthy, 1996; Train & Winston, 2007). Also, some studies combine the vehicle classification (vehicle classes or vehicle make-model) with other factors, such as vehicle vintages (Berkovec & Rust, 1985; Mohammadian & Miller, 2003), fuel efficiency levels (McCarthy & Tay, 1998) or choices of vehicle acquisition (Mannering, et al., 2002). In addition, some researchers investigate the dependency of vehicle type choices on car ownership levels (Berkovec, 1985; Hensher, et al., 1989).

In terms of modelling approaches, the multinomial logit (MNL) model is usually used when the dependent variables are different vehicle classes or vehicle make-models. The only exception is Cao, et al. (2006), who use a partially degenerated nested logit (NL) model and find it is the best structure that groups two alternatives (car and Minivan) into a nest while leaving other alternatives (SUV and pickup truck) independent. When vehicle classes or make-models are combined with another factor such as vehicle age, the NL model is then commonly employed. One recent study employs the mixed logit model to investigate the underlying reasons of the declining market shares of the U.S. car manufacturers in their home market (Train & Winston, 2007).

Table 4-4: A list of vehicle type choice models

No.	Study	Survey Area (year)	Sample size	Dependent Variable/Alternative	Model
1	Lave & Train (1979)	Seven U.S. cities (1976)	541 new car buyers	10 vehicle classes: • sub-subcompact • sports cars • subcompact A & B • compact A & B, • intermediate • standard A & B • luxury	Multinomial Logit
2	Berkovec (1985)	U.S. (1978)	1048 households	Upper level: 4 vehicle ownership levels • 0 vehicle • 1 vehicle • 2 vehicles • 3 vehicles Lower level: 131 combinations of vehicle classes and vintages • 13 vehicle classes • 10 years (1969-1978) • "old car": all pre-1969 models	Nested Logit (sequential)
3	Berkovec & Rust (1985)	U.S. (1978)	237 single-vehicle households	Upper level: vehicle age groups • New cars (1977-1978) • Mid-age cars (1973-1976) • Old cars (1967-1972) Lower level: 5 vehicle classes • Subcompact • Compact • Intermediate • Standard • Luxury/Sport	Nested Logit
4	Mannering & Mahmassani (1985)	U.S. (1980)	220 new vehicle buyers	Vehicle make and model	Multinomial Logit
5	Hensher, et. al. (1989)	Sydney (1981-1982)	1172 households	Upper level: fleet size • 0 car • 1 car • 2 cars • 3 + cars Medium level: body type • Sedan • Station wagon • Other body types Lower level: vehicle models and vintages mix	3-level Nested Logit (sequential)
6	McCarthy (1996)	U.S. (1989)	1564 new vehicle buyers	Vehicle make and model	Multinomial Logit

To be continued on the next page

No.	Study	Survey Area (year)	Sample size	Dependent Variable/Alternative		Model
7	McCarthy & Tay (1998)	U.S. (1989)	1564 new vehicle buyers	Upper level: 3 fuel efficiency levels • High • Medium • Low	Lower level: 10 alternatives for each fuel efficiency (nest)	Nested Logit
8	Mannering et. al. (2002)	U.S. (1993-1995)	654 households acquiring new vehicles	Upper level: vehicle acquisition type • Cash • Non-cash (lease or finance)	Lower level: 10 alternatives (make-models)	Nested Logit
9	Mohammadian & Miller (2003)	Greater Toronto Area (1998)	597 new car purchases	Upper level: 6 vehicle classes • Subcompact • Compact • Midsize • Large • SPVs • Van	Lower level: 4 vehicle age groups • less than 1 year • 1 - 2 years • 3 - 7 years • 8 years or more	Nested Logit
10	Choo & Mokhtarian (2004)	San Francisco Bay Area (1998)	1571 households	9 vehicle classes: • Small, Compact, Mid-sized, Large, Luxury, Sports, Minivan/van, SUV, and Pickup (Reference)		Multinomial Logit
11	Cao et. al. (2006)	North California (2003)	1387 households	Upper level: • Car-van nest • SUV • Pickup	Lower level: • Car • Minivan	Partially degenerated NL
12	Train & Winston (2007)	U.S.(year NA)	458 consumers	Year 2000 new vehicle make-models		Mixed Logit

4.2.4 Car type choice model: key explanatory variables

The explanatory variables in the car type choice model typically include vehicle attributes, brand preference and the demographic characteristics of the household and the main driver (Choo & Mokhtarian, 2004). In addition, in order to account for the potential context-specific effects, we include the consumer knowledge and the primary use of owned cars as additional variables in our car type choice model. Table 4-5 summarises all explanatory variables in our car type choice model and their typical effects in the context of developed car markets.

- *Vehicle attributes*

In our car type choice model, we consider the following important vehicle attributes: the vehicle purchase price divided by household income, fuel cost, horsepower, turning radius, vehicle length, and the number of airbags. In terms of running cost, we calculate the expected annual running cost for each car-holding household by multiplying its annual mileage with its vehicle's fuel consumption rate. Here we briefly review their typical effects found in the other markets.

- Vehicle Price

Vehicle price is included in all studies we reviewed in Table 4-4 and it is usually assumed to interact with household income. We follow the common definition to divide the vehicle price by household income (Lave & Train, 1979; Mannering & Mahmassani, 1985; McCarthy, 1996; McCarthy & Tay, 1998). Furthermore, we include a squared term of the constructed price variable to explore the potential nonlinear effect between the vehicle price and its utility (Lave & Train, 1979), which implies that consumers may not prefer the lowest vehicle price. In addition, we use the manufacturer suggested retail price (MSRP) of each vehicle model instead of the purchase prices provided by each respondent, in order to avoid potential reporting errors (Mohammadian & Miller, 2003; Train & Winston, 2007). Not surprisingly, the price variable has a significantly negative effect across all studies, which implies that the higher priced vehicle is less likely to be chosen (Berkovec, 1985; Berkovec & Rust, 1985; Hensher, et al., 1989; Train & Winston, 2007).

Table 4-5: A summary of explanatory variables of car type choice model and their effects in developed markets

Explanatory Variables	Definition in our model	Typical effect in the developed markets
• Price/income	Vehicle purchase price divided by household income, and its squared	– for price/income, + for price/income squared
• Fuel cost	Annual fuel cost (1000 RMB)	– or insignificant
• Performance	Factor derived through principle component analysis with horsepower, turning radius and vehicle length	+
• Airbags	Number of airbags	+
• Brand's country of origin	Category variable: European brands, US brands, Japanese/Korean brands, Chinese brands (reference category)	+ for foreign brands, – for domestic US brands
• Age	Age of household head	+/–
• Gender	Gender of household head (male as reference category)	+ for small/economical cars (female head)
• Children	No. of children in the household	+ for large/spacious cars
• Residential location	Binary variable, whether the household lives in urban areas	+ for small cars
• No. of owned vehicles	No. of vehicles held by the household	+ for small cars
• Distance	Commute distance of household head	+ for large/spacious cars
• Consumer knowledge about cars and car market	Category variable: good knowledge, basic knowledge, no knowledge (reference category)	Not available
• Primary use of the car	Dummy variable, whether the car is primarily used for business purpose	Not available

Note: + stands for the positive effect on car type choice, – stands for negative effect on car type choice; +/– stands for the association without consistent indications of positive or negative effect.

- Fuel Cost

Another cost related variable is the vehicle fuel cost. Empirical studies define this variable to be either annual fuel cost (Hensher, et al., 1989; Mannering & Mahmassani, 1985; Mannering, et al., 2002), or fuel cost per mile travelled (Berkovec, 1985; Berkovec & Rust, 1985; Lave & Train, 1979; McCarthy, 1996; McCarthy & Tay, 1998; Train & Winston, 2007). In terms of its effect, there is no consistent conclusion yet. Some studies find that the fuel cost has a significantly negative effect on vehicle type choices (Berkovec & Rust, 1985; Hensher, et al., 1989; Mannering & Mahmassani, 1985; McCarthy, 1996; McCarthy & Tay, 1998), while other studies do not identify such statistical significance (Berkovec, 1985; Lave & Train, 1979; Mannering, et al., 2002). Such inconsistencies across different time periods are probably related to the fluctuation of the fuel price. When researchers collected data during the fuel price increasing periods, such as during the 1970s oil crisis or 1988 to 1990 (see U.S. oil price history in the Annual Energy Review[13]), households were found to be more sensitive to the fuel cost than in other periods when the fuel price was fairly stable.

- Vehicle Performance

After the cost-related attributes, the vehicle performance is the next important category of vehicle attributes included in the car type choice model. Among various vehicle performance measurements, horsepower and turning radius are the two most common attributes. Some studies divide the horsepower by the vehicle weight to measure vehicle acceleration ability (Berkovec & Rust, 1985; Lave & Train, 1979; Train & Winston, 2007), while others directly investigate the effect of the horsepower on vehicle type choices (Mannering & Mahmassani, 1985; Mannering, et al., 2002;

[13] See Figure 5.18 of Crude Oil Domestic First Purchase Prices (1949-2010) in the Annual Energy Review (U.S. Energy Information Administration, 2011)

McCarthy, 1996) or directly account for the acceleration time from 0 to 100 km/h (Hensher, et al., 1989). Generally, a vehicle with the larger horsepower or better acceleration performance is more attractive to consumers (Hensher, et al., 1989; McCarthy, 1996; Mohammadian & Miller, 2003; Train & Winston, 2007). In terms of the turning radius, some researchers find that vehicles with the smaller turning radius are more attractive, particularly in metropolitan areas (Berkovec & Rust, 1985; Mannering & Mahmassani, 1985), while other studies show that a vehicle with the greater turning radius is more likely to be chosen due to its correlation with a smoother vehicle ride and greater comfort (Mannering, et al., 2002). In addition, empirical studies in the developed markets consistently find the vehicle length has a significant positive effect on household car choice (Hensher, et al., 1989; McCarthy, 1996; McCarthy & Tay, 1998; Mohammadian & Miller, 2003; Train & Winston, 2007), which implies that consumers in those markets, typically US market, tend to prefer larger vehicles.

Table 4-6: Principle component analysis of vehicle performance factors

	Correlation Matrix			Factor Loading
Variable	Horsepower	Turning Radius	Vehicle Length	
Horsepower	1	0.834	0.727	0.371
Turning Radius		1	0.714	0.369
Vehicle Length			1	0.351
% Variance Explained				83.94%

Furthermore, we find that there is an issue of multicollinearity among all three factors about vehicle performance (see the correlation matrix in Table 4-6), which means it is inappropriate to employ them directly as explanatory variables in the car type choice model. Therefore, we follow Mohammadian and Miller (2003) to use principal component analysis (PCA) to create independent variables based on the information on existing variables. Since all three variables are highly correlated, the principal component analysis generates one factor related to vehicle performance. As the result, this factor can explain about 84% of the total variance of three factors in the sample.

- Airbags

Although airbags have become the standard equipment of cars nowadays, they were not widely adopted in the developed markets until early 1990s. That's why the inclusion of airbags as one of the vehicle attributes appears in more recent studies. More specifically, both McCarthy & Tay (1998) and all three models in Mannering, et al. (2002) show the significant effect of the airbags on increasing the choice probability of the corresponding vehicle. This is probably because consumers are usually risk-averse and the presence of airbags is thought to help improve the vehicle safety and reduce injury risks.

- *Brand preference*

The brand preference is a non-tangible vehicle attribute that measures consumers' attitude to different vehicle brands or more generally their countries of origin. It indicates how consumers perceive and compare different car manufacturers and products. To account for the potential brand preferences in China, we define four main categories based on the country of origin: European brand, American brand, Japanese/Korean brand, and local brand in China. The last one is used as the reference category in our model.

Most empirical studies reviewed in Table 4-4 investigate consumers' brand preferences when choosing different types of vehicles, except Lave & Train (1979) and Mohammadian & Miller (2003). Specifically, most studies compare the brand preferences between the US Big Three car manufacturers (General Motors, Ford and Chrysler) and different Japanese and European brands. For example, Mannering, et al. (2002) find in one of their models based on the vehicle leasing market, that the Chrysler vehicles are extremely disliked by consumers followed by GM, while there is no significant preference difference between the Ford and all foreign brands when consumers lease cars. However, consumer preferences change when they buy vehicles, as Mannering, et al. (2002) show that the US consumers demonstrate strong negative preferences against the domestic Big Three brands as well as Japanese brands when buying vehicles. Similar brand preferences are found in a recent study from Train & Winston (2007), who use the Japanese brands as a reference category. They find that the European brands are particularly welcomed followed by the Korean ones, while the preference difference between the Japanese and the US brands is insignificant. In summary, consumers in the US market prefer the foreign brands and particularly the European brands instead of their local brands, which may help explain the reason why the US car manufacturers have been suffering the market share decline in the past decades as explored in Train & Winston (2007).

- *Demographic variables*

We include the following demographic variables from the literature in our car type choice model: age and gender of household head, number of children, residential location (whether the household lives in urban areas or not), number of owned cars, commuting distance of the household head. It is worth noting that the income effect in the car type choice model is mainly explored through its interaction with the vehicle purchase price, as we have discussed previously.

- Age and gender

Empirical studies demonstrate that there is a relationship between the driver's age and gender and particular types of vehicles. Mohammadian and Miller (2003) show that households with lower average age of people are more likely to choose SUVs or pickups and are less likely to buy the second hand cars. At the same time, they also find that the more mature drivers are also more likely to buy new cars. More recently, research has shown that the older consumers are less likely to drive small cars or personalised vehicles, such as sports cars and SUVs (Choo & Mokhtarian, 2004), but more likely to own minivans (Cao, et al., 2006). In terms of the gender factor, the typical finding is that the female consumers prefer more economical vehicles. McCarthy and Tay (1998) find that women are more likely to buy fuel-saving vehicles. Moreover, Mohammadian & Miller (2003) find that the male drivers significantly prefer large cars, SUVs and pickups as well as minivans. Similarly, both Choo & Mokhtarian (2004) and Cao, et al.(2006) find that women significantly dislike pickup trucks.

- Number of children

The influence of children in the households on their vehicle type choices has been investigated in the recent studies (Cao, et al., 2006; Choo & Mokhtarian, 2004; Mohammadian & Miller, 2003; Train & Winston, 2007). These studies indicate that the households with children are likely to buy spacious vehicles, such as minivans, SUVs and station wagons.

- Residential location

It is interesting to see that the households' residential locations affect not only their choices of the number of cars to own, but also which types of cars to adopt. Its typical effect is that households living in metropolitan areas prefer small cars or cars with high fuel efficiency (Choo & Mokhtarian, 2004; Mannering, et al., 2002; McCarthy, 1996; McCarthy & Tay, 1998).

– Number of already owned vehicles.

A couple of empirical studies investigate the influence of household fleet size, i.e. number of owned vehicles, on their choices of different types of cars. An early study found that the households owning more than two vehicles have a significantly high probability of owning small-sized cars (Lave & Train, 1979). A recent study shows that the households with multiple vehicle ownerships are also more likely to own a pickup truck (Cao, et al., 2006). These results suggest that small cars or pickup trucks are usually not the consumers' first vehicle choice, but often the second or third vehicle owned to diversify the family fleet.

– Travel distance

The travel distance is also accounted for in the car type choice model. It is usually found that the spacious SUVs are more likely to be adopted by consumers who travel a lot (Cao, et al., 2006; Choo & Mokhtarian, 2004), and the compact or sports cars are more desired by people who perceive to mainly drive short distances (Choo & Mokhtarian, 2004).

- *Consumer knowledge and primary use of cars*

We account for two context specific factors, the consumer knowledge and the primary use of cars, in our car type choice model, so that car manufacturers that plan to enter this market could have greater insights about how to better position and promote different types of products. In the developed markets, cars are primarily used for daily commuting, while private cars may take additional roles in the EMs. For example, Vasconcellos (1997) shows that the middle class in the EMs tends to perceive the car as an essential tool for their "social reproduction" to ensure their living conditions. In our model, we explore whether consumers demonstrate different preferences of selecting cars if they use cars for different purposes. Specifically, we define a binary variable to indicate whether the owned cars are primarily used for business purposes. It is worth noting that company cars are not widely available in China except for senior staff, so that most people may have to rely on private cars or other transport modes for their business activities. Such investigations can help car manufacturers appropriately segment the market and define the target customers with different needs.

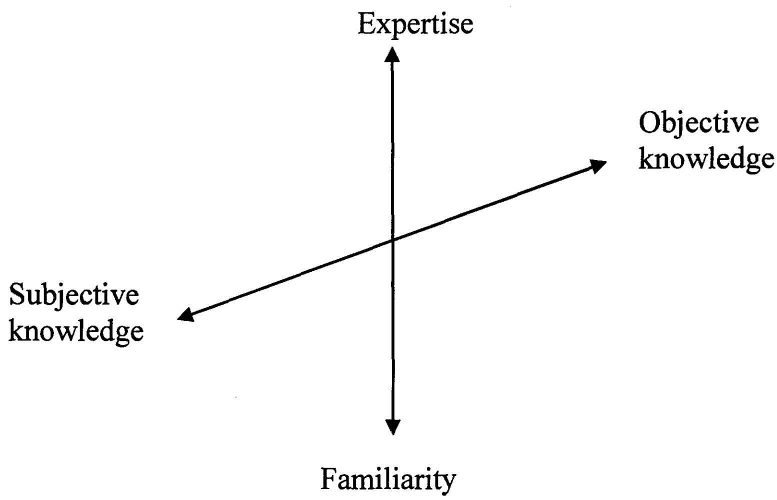
Regarding the definition of consumer knowledge, we use the same category variable as in the car ownership model and the lowest knowledge level is used as the reference category. By accounting for the consumer knowledge in the car type choice model, we explore whether consumers with different knowledge about cars demonstrate different preferences for different types of cars.

4.3 Consumer Knowledge

Consumer knowledge influences how consumers search for information (Bettman & Park, 1980; Brucks, 1985; Cowley & Mitchell, 2003; Rao & Sieben, 1992) and also it impacts on consumer choice behaviour (Maheswaran, et al., 1996; Mitchell & Dacin, 1996; Moorman, et al., 2004; Peracchio & Tybout, 1996; Rao & Monroe, 1988).

When attempting to measure consumer knowledge, Alba and Hutchinson (1987) suggest that consumer knowledge mainly consists of two components: familiarity and expertise. The former is the product experiences accumulated by the consumers and the latter is defined as the consumers' ability to successfully perform product-related tasks. Regarding the relationship of these two components, familiarity is thought to be necessary but insufficient for consumers to have enough expertise (Rao & Monroe, 1988). There is another measure of consumer knowledge, which differentiates two types of knowledge: objective knowledge (correct information actually stored in the memory) and subjective knowledge (self-assessed or perceived knowledge) (Brucks, 1985). When investigating the key determinants of these two different types of knowledge, Park, Mothersbaugh and Feick (1994) find that consumers' subjective knowledge is more associated with their product-related experience than with the stored product class information, while the objective knowledge is affected more by the stored product-class information than by the product experience. Therefore, the subjective knowledge is more related to the familiarity component of consumer knowledge, because both of them depend greatly on the product-related experience, which is mainly acquired through advertising exposure, information search, product usage and ownership. Similarly, the expertise component of knowledge is closer to the objective knowledge, which requires the product-class information, such as usage procedure, product features and brand information, to conduct product-related tasks. See Figure 4-1 for the illustration of two different measures of consumer knowledge.

Figure 4-1: Two different measures of consumer knowledge



When investigating the effect of knowledge on product choice or evaluation, researchers typically employ a segmentation approach that compares how consumers with different knowledge levels make different judgements or choices. For example, in the study that investigates the effect of consumer knowledge on the assessment of product quality using price information, Rao and Monroe (1988) find that consumers with low and high knowledge level tend to perceive a stronger positive association between product price and quality than those with the moderate knowledge, which implies the heterogeneous effects of consumer knowledge on their judgements. Similarly, Maheswaran, et al. (1996) examine how consumers' existing knowledge level affects their further learning and thereby product evaluations. Through comparing the different behaviour from two separate groups (novices versus experts), they find that the learning and product evaluations of the novices can be enhanced by better information organisation and message repetition, while the experts with better prior knowledge are more affected by the type of message delivered and the detailed content of the message.

It is not surprising that the studies about the effects of consumer knowledge are more relevant to new products in the market (Moreau, et al., 2001; Peracchio & Tybout, 1996; Wood & Lynch, 2002). This is because for new products, most consumers are learning about the products and furthermore different consumers may have significantly different knowledge levels about the new products. Thus, the accumulation of knowledge can effectively influence consumers' information search strategy and their choice behaviour thereafter. In comparison, mature products, such as cars in the developed economies, have existed for several decades so that consumers there commonly know how to use these products with generally high expertise. That can explain why the consumer knowledge factor has not been accounted for when modelling car market demand in the developed markets. Therefore, in our consumer survey, we asked each respondent to evaluate their own familiarity about cars and the car market (which is the subjected knowledge we collected), so that we can investigate how consumer knowledge levels are associated with their car adoption behaviour in China.

As we asked for the self-evaluated (subjective) knowledge in our survey, respondents were likely to interpret the question as their general experiences about both products and the market, which are highly influenced by car advertising exposure, car information search as well as car ownership and use. When answering this question, they may think of whether they frequently came across car advertisements, whether they knew different brands or models of cars, whether they actively learned information about cars (e.g. visiting automobile websites or reading car related magazines) and whether they owned and used cars etc. We acknowledge that this subjective knowledge is not very accurate to evaluate the actual knowledge level of respondents. Essentially, we can better design the question to be more specific with a finer scale of knowledge. It will also be desirable to acquire the objective knowledge of respondents, such as detail usage procedure and product features.

4.4 Data Description

Typically when exploring the effects of various factors on consumer choices, we need to use micro-level data. The car ownership and car type choice models in this chapter are based on the survey data we collected in China in early 2010 as we have discussed in Chapter 2. After removing cases with missing information on some key attributes, we are able to use 524 respondents for the car ownership model and 173 car owners for the car type choice model in this chapter.

Table 4-7 compares the major demographic characteristics of our sample that we used for the car ownership model against those in the 2009 national sample reported in the China Statistical Yearbook 2010 (National Bureau of Statistics of China, 2010). Our survey oversampled high-income groups and car owners in China, so that we can have enough observations for the car type choice model that is based on car owners. In our first model that measures the car ownership status of the whole population, we reweight our sample based on the household income and car ownership information from the national sample. In our second model of car type choices based on the owners only, there is no obvious indication of oversampling any specific group of car owners and we do not have national level information about all car owners in China, so we do not use the reweighting approach.

Table 4-7: Demographic distribution of car ownership model sample

Variables	Our Sample	National Sample[a]
Average family size	3.25	2.89
Average no. of children less than 18 years old	0.46	/
Average no. of working members	2.08	1.49
Average household disposable Income in 2009 (RMB)	125,391	49,635
No. of cars owned per 100 households	66.22	10.89
Average age of the household head	37.74	/
Proportion of male household head (%)	78.05	/
Average commuting distance of household head (km)	9.13	/
Sample size	524	65,506

[a] National Sample Data Source: China Statistical Yearbook 2010

In the car type choice model, we also collected secondary data on vehicle attributes from the magazine of *Orient Auto* in China and the automobile fuel consumption website developed by the Ministry of Industry and Information Technology of China[14]. See Chapter 2 for more discussions on the data and data collection.

4.5 Modelling Car Ownership in China

4.5.1 Modelling approach

We follow previous research and use discrete choice models to analyse households' car ownership decisions in China. More specifically, because owning multiple cars is fairly rare in China and our sample only has 9.5% of households owning two or more cars, we employ a binary choice model to investigate the households' choice between

[14] The website address is http://chinaafc.miit.gov.cn/index.html.

owning private cars or not. We specify the utility function of the household (i) owning cars depends on various explanatory variables (X_i), including demographic characteristics, residential location, alternative transport modes and consumer knowledge:

$$U_i = V_i + \varepsilon_i = \beta' X_i + \varepsilon_i \tag{4-1}$$

where V_i and ε_i are the deterministic and random portions of the utility respectively.

When the random utility is assumed to follow the logistic distribution, the model is a binary logit model, while it becomes a binary probit model with the random portion following the normal distribution (Franses & Paap, 2001; Greene, 2009). We use the binary logit model here[15], and the corresponding probability of owning at least a car is

$$P_i = \frac{\exp(\beta' X_i)}{1 + \exp(\beta' X_i)} \tag{4-2}$$

4.5.2 Estimation results and discussions

Because the effect of consumer knowledge has not been explored in other markets, we estimate two binary car ownership models, the first one without consumer knowledge and the second one with this variable, to demonstrate the value of the consumer knowledge in contributing to a better car ownership model in China. Both binary car ownership models are estimated using NLOGIT 4.0 (Greene, 2007)[16] and the estimated parameters of these two models are presented in Table 4-8.

---

[15] We also estimate a binary probit model, but it does not show the significant difference from the binary logit model.

[16] The NLOGIT code for the binary car ownership model is available in Appendix 3.1.

Table 4-8: Estimated parameter of the car ownership model

Explanatory variable	Model 1 Coeff.	Model 1 T-stat.	Model 2 Coeff.	Model 2 T-stat.
Constant of car owners	-6.98	-3.90***	-7.38	-4.05***
Demographic variable				
Household income	0.46	2.51**	0.43	2.29**
Income squared	-0.02	-1.75*	-0.02	-1.64
Number of working adults	-0.21	-0.54	-0.33	-0.80
No. of drivers	5.11	4.71***	5.02	4.49***
Gender of household head	-0.33	-0.63	-0.81	-1.38
Age of household head	0.05	2.18**	0.06	2.46**
No. of children	1.09	2.66***	0.91	2.12**
Family size	0.65	2.11**	0.74	2.33**
Residential location				
Urban area of cities	-1.87	-2.44**	-2.15	-2.60***
Suburban area of cities	-0.73	-0.80	-0.87	-0.89
Unknown whether urban or suburban areas	-1.16	-1.61	-1.45	-1.86*
Alternative transport modes				
Taking public transport	-3.88	-5.66***	-4.32	-5.75***
Riding bikes or motorcycles	-2.75	-5.28***	-2.87	-5.23***
Company shuttles	-3.06	-4.17***	-3.20	-4.16***
Consumer knowledge				
Good knowledge level			1.89	2.76***
Basic knowledge level			1.25	2.30**
No. of observations	524		524	
Log-likelihood at zero	-359.77		-359.77	
Log-likelihood at convergence	-84.39		-79.62	
Rho-square w. r. t. zero:	0.765		0.779	
Percentage of correct predictions (%)	74.42		77.86	
χ^2 against zero	550.76 (df = 15)[a]		560.31 (df = 17)[b]	
χ^2 against Model 1			9.55 (df = 2)[c]	

Note: [a] the critical value for 15 degrees of freedom is 25.00 at 5% significance level; [b] the critical value for 17 degrees of freedom is 27.59; [c] the critical value for 2 degrees of freedom is 5.99.
*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Overall, both models perform very well as indicated by their high values of the likelihood ratio index or Rho-square (0.765 and 0.779). Furthermore, our models achieve good predictive accuracy (74.42% and 77.86%). The likelihood ratio tests against the equal probability model are conducted and the high chi-square test values of 550.76 with 15 degrees of freedom and 560.31 with 17 degrees of freedom significantly reject the null hypothesis of no difference from the equal probability model. Finally and more importantly, the difference between the models with and without the knowledge variable is tested through another likelihood ratio test with 2 degrees of freedom. The resulted chi-square value of 9.55 is larger than the critical value of 5.99 at 5% significance level, which confirms that the second model including the consumer knowledge variable is significantly better than the first model without consumer knowledge in terms of explanatory power. Therefore, we discuss the estimated parameters based on the second model.

When discussing the estimated parameters, we are particularly interested in their differences in comparison to other markets, which are summarised in Table 4-9. Thus we discuss the estimated results along with the comparisons of their typical effects in the developed markets as we have reviewed in section 4.2.2.

Among all demographic variables, we find that two variables typically significant in the developed markets, the number of working adults and the gender of the household head, are insignificant in our car ownership model. The insignificance of the gender variable implies that the first household car purchase usually depends on the joint decision of the spouse no matter who is the head member. In addition, with the economic reform in China, Chinese women have become more independent and thus balance the influence with their partners on household purchase decisions (Doctoroff, 2005). In terms of the number of working adults, its significance in the literature is probably because most working adults rely on private cars for their independent commutes in the mature markets so that owning two or more cars is very popular. In China, however, most car-holding families own one car only, so that the dominating status of single car ownership cannot account for the different numbers of working adults in the Chinese households. In addition, we find that household income variable is significant but the income squared is insignificant in our model, which indicates the existence of a linear association between income and car ownership and further implies that the Chinese car market is far from the saturation level.

Table 4-9: Comparison of effects in car ownership model

Explanatory variable [a]	Typical effect in the developed markets	Finding in our model
• **Income**	+ for the income, – for the income squared	+ for income, insignificant for income squared
• **No. of working adults**	+	Insignificant
• Drivers	+	+
• Age of head	+ for working age, - for retired age	+
• **Gender of head**	+ for the male head	Insignificant
• Household type	+ for larger households and/or with children	+ for larger households + for households with children
• Residential location	Consumers in higher dense areas are less likely to own cars.	Urban households are less likely to own cars than households in remote areas.
• Availability of alternative transports	–	–
• **Consumer knowledge**	Not applicable & not tested	+

Note: + stands for the positive effect on car ownership, – stands for negative effect on car ownership; 'Insignificant' means the insignificant effect on car ownership;
[a] Explanatory variables in bold have different effects in China and the developed car markets.

The remaining three demographic variables are largely consistent with the findings in the developed markets. Specifically, the households with more licensed drivers, more family members or children are more likely to adopt cars in China. In addition, most family heads in our sample are at their working ages and the age variable demonstrates a positive effect on the car ownership in China, implying that the mature households are more likely to adopt cars than the young ones.

In terms of alternative transport modes, our models find the substitution effects in China as in other markets. Specifically, good accessibility to three alternative transport modes (public transport, bike/motorcycle and company shuttle) reduces the likelihood for consumers to adopt cars. It is worth noting that we only accounted for the substitution effect of alternative transport modes (by using dummy variables) on owning cars, but we are not able to address the effects of different quality levels of public transport and road system across different countries in this thesis.

With regard to the residential locations, the typical effect on car ownership in other markets is that the households living in the denser areas have the lower car ownership level. In our model, we find the similar effect of the residential locations on car ownership level in China. Specifically, by using the households in remote areas as the reference, the urban households show a significant negative propensity to own cars. The households in suburban areas have no significant difference from those in remote areas.

Finally, we find that there is a positive association between consumer's knowledge levels and their car ownerships in China. As cars are still considered as new products in China, this finding is consistent with the prior consumer research that consumers' subjective knowledge is positively associated with their product related experiences such as product ownership or use (Park, et al., 1994). On the one hand, car ownership experiences can enhance consumers' familiarity or knowledge about cars and car markets. On the other hand, better knowledge about cars can improve consumers' information search efficiency (Brucks, 1985), which may further help them make earlier adoptions. What is important here is that we explicitly show the significant association between a consumer's knowledge and their car ownership level. It provides important implications for car manufacturers that the market size can be expanded and consumers can be encouraged to make earlier purchases through helping them improve their knowledge about the products and the markets. For example, more trial driving activities can help consumer learn how to set realistic expectations about cars, which will positively influence market demand (Goering, 1985; Lakshmanan & Krishnan, 2011).

4.5.3 Model validation

We also examine the predictive capabilities of two car ownership models (Model 1 and Model 2). We first drew 10 random sub-samples from the whole dataset and each sub-sample consists of 90% observations. They were used as the estimation samples and the remaining 10% observations were kept as the validation samples. For every estimation sample, different car ownership model specifications were estimated separately and the estimated results were used to calculate the choice probabilities of each alternative in the validation sample. We will not report the estimated parameters of these 10 sub-samples here but show the predictive performances of different car ownership model specifications in the validation samples.

We follow Bhat & Pulugurta (1998) to compare forecasting performances of different discrete choice models using both aggregate and disaggregate measures. The aggregate measures are MAD, MAPE and RMSE that examine the accuracy of market share forecasts for different alternatives in the validation sample, and the disaggregate measure is the correctly predicted ratio, which is calculated through dividing the number of respondents who have been correctly predicted to make their choices, based on the estimated model, over the total number of respondents in the validation sample (Hensher, et al., 2005; Train, 2003). The prediction of respondent's choice is made for the alternative with the highest probability for each respondent.

Since we have 10 random sub-samples, we report both mean and median values of all forecasting performance measures across all validation samples in Table 4-10. It clearly shows that the Model 2 that accounts for consumer knowledge as a key explanatory variable in car ownership model can provide better forecasting performances than Model 1 without that variable, as indicated by the smaller error measures (both mean and median values of MAD, MAPE and RMSE) and higher percentage of correct predictions.

Table 4-10: Validation results of car ownership models

	Model 1	Model 2
Aggregate Measure		
Mean of MAD	21.28	19.01
Median of MAD	22.24	19.73
Mean of MAPE	45.09%	40.26%
Median of MAPE	48.39%	41.29%
Mean of RMSE	21.24	19.01
Median of RMSE	22.24	19.73
Disaggregate Measure		
Mean of correctly predicted ratio	74.24%	75.74%
Median of correctly predicted ratio	73.31%	76.11%

4.6 Modelling Car Type Choices in China

4.6.1 Modelling approach

The dependent variable of vehicle type choice model primarily depends on vehicle classification according to vehicle sizes as well as other special body types such as SUVs or MPVs (see Table 4-4). In addition to SUVs and MPVs, the passenger cars in China are usually classified into 6 categories based on their sizes: mini, small, compact, mid-sized, upper-mid and luxury classes[17]. Because our sample size of car owners is small, we do not have enough observations of SUV or MPV owners and we also need to combine 6 size classes of cars into three categories as follows. The mini class is merged into the small car category, the compact and mid-sized classes are combined to be a new mid-sized car category, and the upper-mid and luxury classes are combined together to be a large car category.

Along with the fast development of the car market in China, more than 75% of households in our sample who own cars made their adoptions from 2006 to early 2010 when we conducted the survey. In addition, we are more interested in the current or recent consumer behaviour in the market and exclude car type choices occurred at earlier time (such as 5 years ago), because in such a fast growing market, the earlier consumer preferences may differ significantly from now and thus they have limited values for studying the existing or future market demand. Therefore, our car type choice model is based on the new car buyers in China (from 2006 to early 2010). As a result, we have in total 173 cases left for the car type choice model.

[17] See detail vehicle classification schemes on two most popular auto websites in China: http://data.auto.sina.com.cn/ and http://db.auto.sohu.com/model-list.shtml

The choice set of each household is defined as follows. Each household has three different alternatives of small, mid-sized and large car types, from which consumers have made their decisions to buy which type of cars. First, we randomly selected a representative vehicle model in each class from all available models by applying a reweighting method. The weights are defined based on actual sales of different vehicle models from 2006 to 2009, which is based on an assumption that the vehicle models that had been sold more would have better market exposure and thus are more likely to be considered and compared by consumers. Second, since each respondent has reported one owned car, we can easily identify its class and then use it to replace the randomly selected one in the corresponding class. Thus, one actually owned car and two randomly selected cars form the different choice set for each respondent.

Given three alternatives in our choice set, we employ the multinomial logit (MNL) model (McFadden, 1974). The utility function of household (*i*) choosing car type (*n*) depends on car attributes (*X*), and choice-invariant factors (*Z*) including demographic variables, consumer knowledge and primary use of the car.

$$U_{in} = V_{in} + \varepsilon_{in} = \alpha' X_{in} + \beta_n' Z_i + \varepsilon_{in} \tag{4-3}$$

Thus choice probability of the MNL model that assumes that the random term (ε_{in}) follows the type I extreme value distribution is specified as

$$P_{in} = \frac{\exp(V_{in})}{\sum_{j=1}^{3} \exp(V_{jn})} \tag{4-4}$$

4.6.2 Estimation results and discussions

Depending on whether accounting for two context specific explanatory variables (consumer knowledge and primary use of owned cars), we have two different forms of car type choice model. Both models are also estimated using NLOGIT 4.0 (Greene, 2007)[18] and their estimation results are shown in Table 4-11.

After estimating the MNL model, we apply the Hausman test (Hausman & Mcfadden, 1984) to verify whether the independence from irrelevant alternatives (IIA) property is violated or not. The test result confirms that all the alternatives are independent and thus the MNL model is an appropriate model here. In general, our car type choice models achieve good performance with the likelihood ratio index (Rho-square) to be 0.242 and 0.300 respectively. Also, our models achieve the corrected predicted ratios to be 60.12% and 67.05%. In addition, the likelihood ratio test values of 91.87 and 114.01 indicate that both models are clearly superior to the equal probability model that has all parameters to be zero.

Furthermore, we conduct another likelihood ratio test to compare two different specifications of the car type choice model. The test statistic is 22.14 with 6 degrees of freedom, which is larger than the critical value of 12.59 at 5% significance level. This indicates that the Model 2 with two context specific variables of consumer knowledge and primary use of the car achieves higher explanatory power than the Model 1 without these two variables.

[18] The NLOGIT code for the car type choice model is available in Appendix 3.2.

Table 4-11: Estimated parameters of the car type choice model

	Model 1		Model 2	
Explanatory variable	Coeff.	T-stat.	Coeff.	T-stat.
Vehicle attributes				
Price/income	-0.59	-3.18 ***	-0.67	-3.31 ***
Price/income squared	0.01	2.05 **	0.02	2.39 **
Fuel cost	-0.41	-2.17 **	-0.48	-2.32 **
Performance	0.13	0.40	0.19	0.52
No. of airbags	0.36	2.96 ***	0.46	3.34 ***
European brands	1.00	2.54 **	1.07	2.62 ***
American brands	0.74	1.76 *	0.76	1.71 *
Japanese or Korean brands	0.03	0.09	0.10	0.28
Individual characteristics of mid-size car owners				
Mid-size car, constant	-0.64	-0.49	0.70	0.43
Age of household head	0.01	0.37	0.01	0.25
Gender of household head (male)	0.30	0.61	0.13	0.26
Living in urban areas	-0.85	-2.04 **	-1.04	-2.32 **
No. of owned cars	1.44	1.73 *	1.64	1.81 *
Commuting distance of household head	0.00	0.12	0.01	0.38
Basic knowledge			-1.45	-1.93 *
Good knowledge			-2.31	-2.92 ***
Use cars for business purposes			2.45	1.97 **
Individual characteristics of large car owners				
Large car, constant	-3.63	-2.09 **	-2.00	-0.97
Age of household head	0.00	0.17	-0.01	-0.22
Gender of household head (male)	1.39	2.05 **	1.41	1.97 **
Living in urban areas	-0.86	-1.65 *	-1.02	-1.82 *
No. of children	0.74	1.91 *	0.80	2.01 **
No. of owned cars	2.17	2.34 **	2.41	2.44 **
Commuting distance of household head	0.04	1.14	0.05	1.38
Basic knowledge			-2.44	-2.67 ***
Good knowledge			-2.08	-2.29 **
Use cars for business purposes			2.70	2.07 **
No. of observations	173		173	
Log-likelihood at zero	-190.06		-190.06	
Log-likelihood at convergence	-144.13		-133.06	
Rho-square w.r.t. zero	0.242		0.300	
Percentage of correct predictions	60.12		67.05	
χ^2 against zero	91.87 (df = 21)[a]		114.01 (df = 27)[b]	
χ^2 against Model 1			22.14 (df = 6)[c]	

Note: [a] the critical value for 21 degrees of freedom is 32.67 at 5% significance level; [b] the critical value for 27 degrees of freedom is 40.11; [c] the critical value for 6 degrees of freedom is 12.59.
*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

Table 4-12: Comparison of effects in car type choice model

Explanatory Variable [a]	Typical effects in the developed markets	Finding in our model
• Price/income	– for price/income, + for squared price/income	– for price/income, + for squared price/income
• Fuel cost	–	–
• **Performance**	+	Insignificant
• No. of airbags	+	+
• **Brand preference**	+ for foreign brands – for domestic US brands	+ for European brands followed by the US brands; Insignificant for Japanese/ Korean brands.
• **Age**	+/–	Insignificant
• Gender	+ for small cars (female)	+ for large cars (male)
• No. of children	+ for large/spacious cars	+ for large cars
• Residential location (urban)	+ for small cars	+ for small cars
• **No. of owned vehicles**	+ for small cars	+ for large cars
• **Commute distance**	+ for large/spacious cars	Insignificant
• **Knowledge level**	Not available	+ for small cars
• **Primary use of cars for business purposes**	Not available	– for small cars

Note: + stands for the positive effect, – stands for negative effect; +/– stands for the significant association without consistent indication of positive or negative effect. 'Insignificant' means insignificance for car ownership.
[a] Explanatory variables in bold have different effects in China and the developed markets.

We follow the same approach employed in the car ownership model to discuss the estimated parameters along with the comparison against their typical effects in the developed car markets (see Table 4-12 for the comparison summary). First of all, we find that the Chinese consumers place more emphasis on the monetary factors and safety than vehicle performance when selecting cars, which corroborates the findings observed by the local market experts (Garner, 2005, p. 78). Specifically, we find the Chinese consumers are fairly similar to those in the mature markets in being concerned about the vehicle price, running cost and the number of airbags. However, what is significantly different in China is that vehicle performance is insignificant.

This is probably because when consumers buy their first cars, they have no prior experiences about how different technical factors will affect their car use, but other tangible factors seem more sensible for them, such as how much money they need to pay and how safe the car can be. Of course, when consumers have more experiences about running cars, we expect that they will become more sensitive to the technical attributes.

With respect to the brand preference, using the local (Chinese) brands as the reference category, we find that the European brands are most preferred in China. This can be explained by the fact that both Germany and French car manufacturers were the earliest entrants into the Chinese car market, setting up their local joint venture firms in China before the Chinese government published the *Automotive Industry Plan* in 1994 to officially support the car market development. In particular, as a market leader in China, Volkswagen group has started to sell more cars in China, mainly under three main brands of Volkswagen, Audi and Skoda, than in its home market of Germany[19], and China has been considered as its "second home market"[20]. Its market advantage has been built through emphasising the reliable German engineering as well as providing dynamic and youthful products in China (Doctoroff, 2005, p. 134). After the European brands, the Chinese consumers also like the U.S. brands, as indicated by the positive effect significant at 10% level. It is interesting because these U.S. brands have dramatically lost market shares in their home market (Train & Winston, 2007). The reason is largely due to the success of General Motors (GM) in China since its first debut in 1998. Buick, a dead brand of GM in the United States, has revived and

[19] "VW China sales exceed home market", 27/06/2009, Leftlane, http://www.leftlanenews.com/vw-china-sales-exceed-home-market.html
[20] "VW sees leap in its China profits", 27/10/2010, Financial Times, http://www.ft.com/cms/s/0/a4eb38c8-e1b5-11df-b71e-00144feabdc0.html

becomes the category leader in China, so that it remains as one of four pillar brands after GM's recent bankruptcy crisis[21]. By the end of 2010, GM also achieved more annual sales in China than in the U.S.[22]. In addition, our model shows that the Asian brands from either Japan or Korea are not significantly preferred in China compared to the Chinese brands, probably due to their low quality images in the mind of Chinese consumers[23]. In particular, Japanese car makers need to be extremely careful about the sensitive nationalism in China against the Japanese brands. The improper advertisement case of Toyota that we have discussed previously is a typical example. This Japanese car manufacturer did not fully understand the potential influences of local culture, local consumers and historical events before conducting its marketing activity, which directly led to the negative effects on the company.

Among all demographic characteristics, the effects are generally similar as in other markets and the differences are only on a few variables. First, the age factor insignificantly influences consumers buying different types of cars in China. Second, the commute distance has limited influences on car type choices in China, while it is positively associated with the choice of spacious cars in the mature markets. In the developed markets, cars are mainly used for daily commuting. But it is different in the emerging markets when consumers own their first cars. Typically, the first cars are shared by the whole family and the commuting is one of the roles only. Other important roles of owning cars include conveniently sending children to school, shopping, weekend/holiday short trips, supporting business purpose or even showing

---

[21] "Buick Back From the Dead", 06/05/2010, The Detroit Bureau, http://www.thedetroitbureau.com/2010/05/buick-back-from-the-dead/

[22] "GM sells more vehicles in China than in the U.S.", 24/01/2011, USA Today, http://www.usatoday.com/money/autos/2011-01-24-gm-china-sales_N.htm

[23] "Breaking the Rules: Moments of Truth in the Car Buying Process", presentation from ACNielsen China 2007, http://jp.nielsen.com/industry/Breakingtherule_China.pdf.pdf

social status/wealth etc. Therefore, the insignificant effect of the commuting distance variable implies that other factors instead of commuting distance of household head are more important on influencing car type choice in China. Third, the households owning more than one car are more likely to adopt mid-sized or large cars, implying that the Chinese households usually choose small cars as their first cars, which is different in the U.S., where small sized cars are more preferred by the multiple car owners (Lave & Train, 1979). The strategic implication for the car manufacturers is that if they can develop small cars with satisfying characteristics such as reasonable price and high fuel efficiency, this segment could potentially attract a huge number of consumers as most consumers in China still have no car owning experiences

In order to explore the context specific effects when the consumers choose different types of cars, we account for consumer knowledge and primary use of owned cars in the car type choice model. The estimation results indicate that consumers with good knowledge about cars or the car market are less likely to own both midsized and large cars, and consumers with basic knowledge are more likely to choose small cars. This is probably because consumers with better knowledge have higher awareness about the high fuel efficiency and less pollution of smaller cars. In addition, as we have shown that the small cars are usually bought as the first cars in China, the owners of small cars might be preparing for re-purchasing or upgrading their cars in the future, so that they are more active in acquiring information about cars and thus they know more about the car market. When jointly considering the effects of consumer knowledge on car ownership and car type choice models, we find that in general car owners have higher knowledge than non-car owners, and within the segment of car owners, those who own small cars tend to know more than other car owners. Such heterogeneous knowledge levels across different car ownership status provides an important implication for the car manufacturers, suggesting that they could try to diversify their market strategies or products to meet the different requirements from consumers with different knowledge or experiences about cars. In addition, our finding also suggests that the government could promote the segment of small cars through enhancing consumers' knowledge or awareness about the fuel saving advantage of small cars in addition to the typical monetary incentives.

In terms of the primary use of cars, we find that if the cars are mainly used for the business purposes, the Chinese consumers are less likely to buy small cars, which implies the importance of "prestige face" in the Chinese culture (Ho, 1976; Tian & Dong, 2011). Driving small cars to do business with others might be perceived to be "face losing" and lack high social status, as social status is usually used as a "tool" to advertise expensive products such as cars or diamonds in China (Doctoroff, 2005, p. 29). This finding supports the current practice of car manufacturers that the large cars are developed and promoted with a good capability of meeting the business-related requirements, such as the larger space for rear seats and luxury interior decoration.

4.6.3 Model validation

We follow the same approach used for the car ownership model to compare the predictive performances of two car type choice models in the validation samples. Specifically, we examine the aggregate level market share forecasting errors (MAD, MAPE and RMSE) and disaggregate level correctly predicted ratio for all alternatives (Bhat & Pulugurta, 1998).

The forecasting performances of two different specifications of car type choice model are presented in Table 4-13. We compare both mean and median values of all measures, since we have 10 different validation sub-samples randomly selected from the whole sample. At the aggregate level, Model 2 consistently demonstrates superior predictive performance over Model 1, indicated by the smaller mean and median values of every error measure for the forecasted market shares. At the disaggregate level, Model 2 achieves higher ratio of correct predictions on average across all validation sample, although both models have the same medians of correctly predicted ratio.

Table 4-13: Validation results of car type choice models

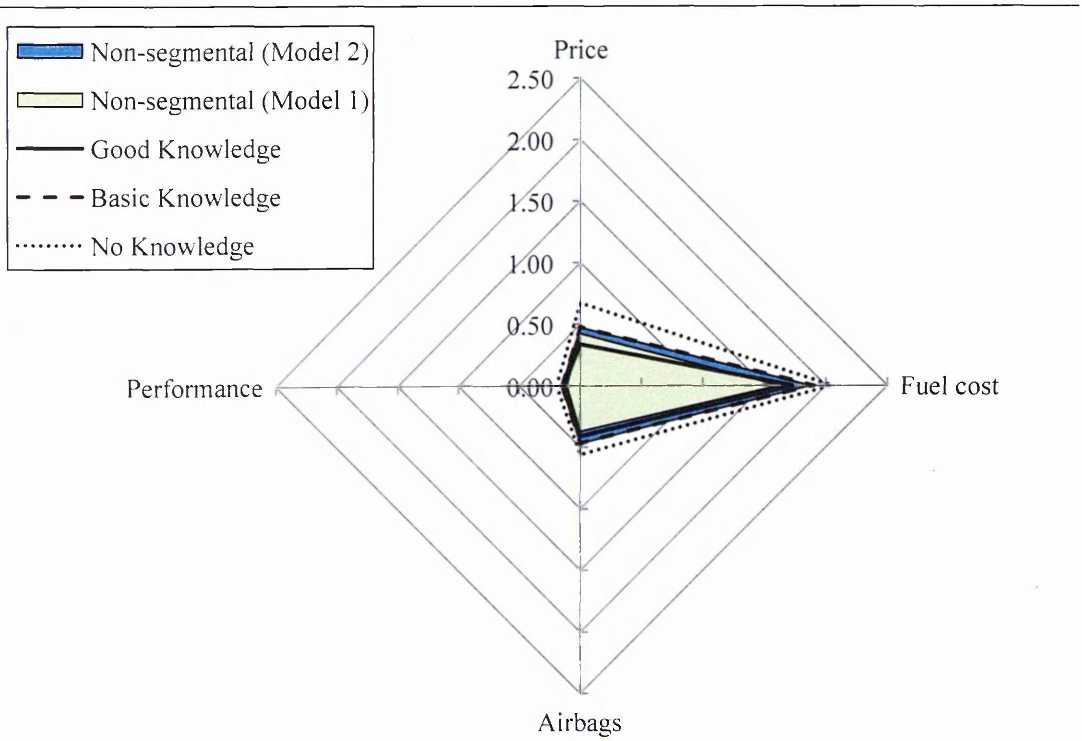
	Model 1	Model 2
Aggregate Measure		
Mean of MAD	17.69	13.87
Median of MAD	18.35	15.87
Mean of MAPE	64.02%	51.25%
Median of MAPE	55.40%	43.63%
Mean of RMSE	19.10	15.53
Median of RMSE	20.03	17.82
Disaggregate Measure		
Mean of correctly predicted ratio	54.23%	55.43%
Median of correctly predicted ratio	53.13%	53.13%

4.7 Segmentation Analysis based on Consumer Knowledge

As we have discussed previously, consumer knowledge is one of the important context specific characteristics, which should be accounted for in order to better understanding local consumer behaviour. In this section, we conduct a segmentation analysis to further examine the moderating effect of consumer knowledge on consumer preferences. Specifically, we define three consumer segments based on their knowledge levels and compare the elasticity effects of various vehicle attributes for the consumers in different segments. We also compare the segmental elasticity effects with the non-segmental elasticity effects based on both car type choice Model 1 and Model 2.

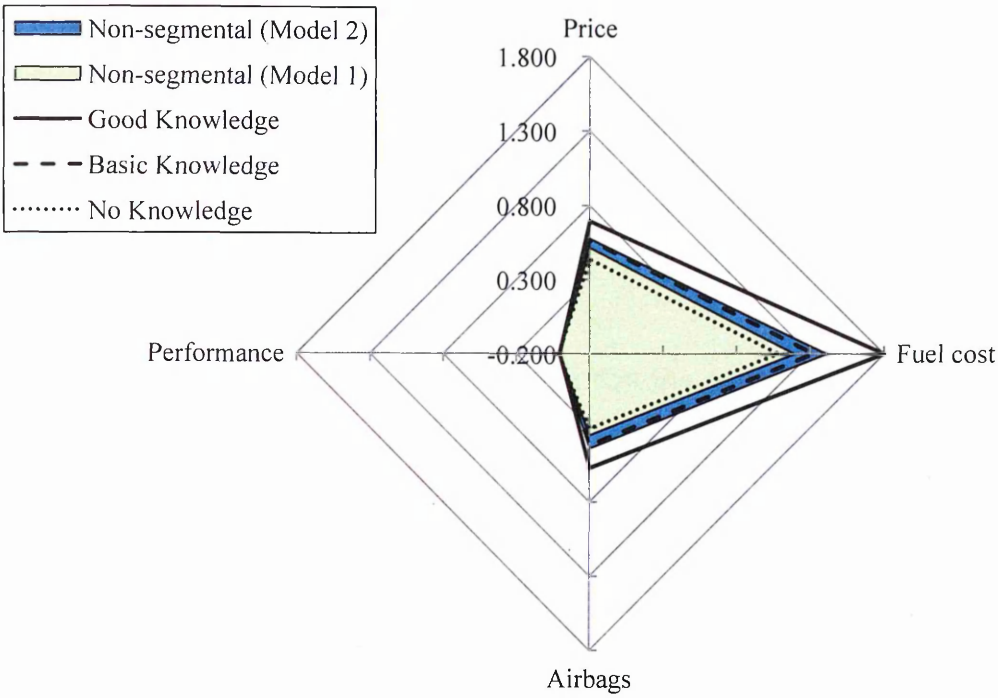
Comparisons of the direct-elasticity effects based on different consumer knowledge levels are shown in Figure 4-2, Figure 4-3 and Figure 4-4 respectively for three types of cars. It is worth noting that in spite of the negative elasticity effects of some variables (such as price and fuel cost), we show and compare the magnitudes of all elasticity effects in these figures.

Figure 4-2: Direct-elasticity effects based on knowledge levels (small cars)



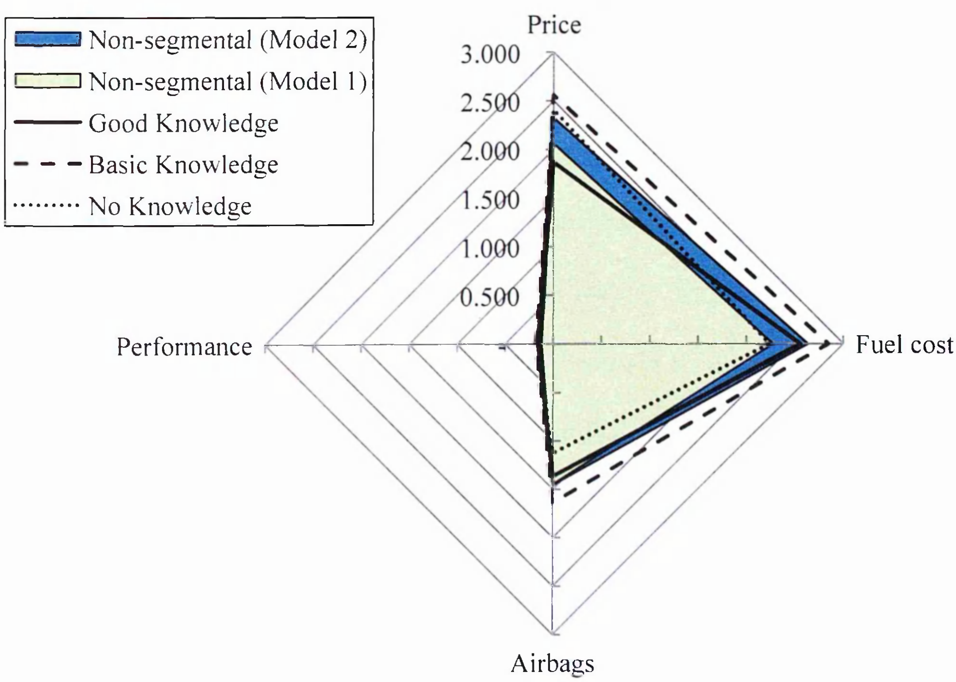
Note: p-values from ANOVA test for elasticity differences across segments are 0.001 (price), 0.531 (fuel cost), 0.122 (airbags), and 0.015 (performance) for small cars.

Figure 4-3: Direct-elasticity effects based on knowledge levels (mid-sized cars)



Note: p-values from ANOVA test for elasticity differences across segments are 0.269 (price), 0.037 (fuel cost), 0.000 (airbags), and 0.579 (performance) for mid-sized cars.

Figure 4-4: Direct-elasticity effects based on knowledge levels (large cars)



Note: p-values from ANOVA test for elasticity differences across segments are 0.370 (price), 0.256 (fuel cost), 0.001 (airbags), and 0.088 (performance) for large cars.

When comparing two types of non-segmental elasticity effects, we find that the car type choice Model 1 that does not account for the context-specific variables consistently underestimate the direct-elasticity effects of all vehicle attributes across all types of cars, as illustrated with smaller green areas for the Model 1 than the blue areas for the Model 2 in all these figures.

When looking at each type of cars, we find that the annual fuel cost is the most important factor for the owners of small and mid-sized cars. The large car buyers are more concerned about both fuel cost and purchase price, while vehicle performance is the least sensitive across all types of cars. More importantly, we note some differences on the elasticity effects for the consumers at different knowledge levels. For the small cars, consumers with no knowledge about cars demonstrate the highest direct-elasticity effects (see Figure 4-2), which is linked to our previous finding that the smaller cars are usually the preferred choice of the first time car buyers in China who have no prior car ownership experience. Regarding the mid-sized cars, we find that consumers with no knowledge have the lowest direct-elasticity effects (see Figure 4-3). This might be explained by the fact that the segment of mid-sized cars dominates the car market in China, so that consumers with better knowledge can conduct more information search and comparisons when selecting mid-sized cars. In terms of the large cars, we find that the consumers with the basic or moderate knowledge tend to have the highest direct-elasticity effects, while other consumers with either no knowledge or good knowledge have comparatively lower elasticity effects (see Figure 4-4). Such kind of "U-shaped" effect is also found by Rao & Monroe (1988) and they indicate that this effect occurs for products with a wide variation of product features. The segment of large cars in our study shares the similar features of wide variation, as the segment consists of upper-mid and luxury cars and their attribute values (such as price) can be extremely large for the luxury cars.

In addition, we conduct ANOVA tests to examine the significance of elasticity differences across segments (see reported *p*-values in the notes of above three figures). As we can see from the test result, consumers at different self-perceived knowledge levels tend to have significantly different elasticity effects on some attributes depending on what type of cars they own. For small car owners, the elasticity effects both purchase price and vehicle performance factor are significant different across three segments. For consumers who own midsize cars, their elasticity effects of fuel cost and airbags are significantly different across three knowledge levels. The airbag variable also demonstrates different elasticity effects for the large car owners across three segments.

In summary, by employing the consumer knowledge as an example, we have showed the importance of accounting for the context specific factors when investigating the local consumer behaviour, as shown with the different elasticity effects in Model 1 and Model 2. Moreover, through the segmentation analysis, we further demonstrate the importance of the context specific variables, as indicated by our findings that consumers at different knowledge levels have different elasticity effects on vehicle attributes and comparatively the non-segmental approach tends to provide biased estimations of the elasticity effects for consumers in different segments. Therefore, properly accounting for the context specific variables is particularly important for international companies to enter the emerging markets such as China, where there may exist significant different contextual characteristics from the companies' current markets (Burgess & Steenkamp, 2006; Johansson, 2009; Sheth, 2011).

4.8 The Effect of Consumer Knowledge on Purchase Intentions

We have discussed in detail the importance of accounting for the context specific variables, such as the consumer knowledge in our study, when understanding local consumer purchase behaviour that is observable in the market. In this section, we extend our study to examine the key determinants of future purchase intentions and particularly the potential effect of the context specific variables on the purchase intentions.

Consumer purchase intentions have been widely employed to forecast the actual purchase behaviour and product demand (Armstrong, et al., 2000; Bemmaor, 1995; Chandon, et al., 2005; Jamieson & Bass, 1989; Sun & Morwitz, 2010). Using intention information to predict sales is found to perform better for the durable goods that have already existed in the market (Morwitz, et al., 2007). Here, we are particularly interested in understanding the key factors influencing the consumer purchase intentions in the Chinese car market.

As we have discussed in Chapter 2, we collected data about the respondents' car adoption intentions. In the survey, we asked each respondent whether he/she has the intention or plan to buy a car within the next 5 years. With the respondents' choices out of three options (yes, not sure and no), we can develop a discrete choice model to investigate which explanatory variables significantly influence different levels of purchase intentions. We acknowledge here that it is one limitation of our survey that we only provided three levels of purchase intentions. It would be better to provide more detailed scale of purchase intentions or even ask the probability of purchase intentions (Morwitz, 2001; Van Ittersum & Feinberg, 2010). Given the three alternatives in the dependent variable, the developed model takes a form of the MNL model. We include following demographic variables in the model: household income, number of working adults, number of drivers, gender and age of the household head, family size, number of children, whether the household owns cars or not. All these demographic variables have the same definition as we have discussed in Section 4.2.

We follow the previous approach to develop two choice models to compare. The first model only accounts for the demographic variables, while the second one includes both the demographics and the consumer knowledge information[24]. The estimated parameters of both models are presented in Table 4-14.

In the first model, we find that most demographic variables, except household income and current car ownership, are insignificant for consumers' purchase intentions. Specifically, since most existing car owners adopted their cars a few years ago, they demonstrate significantly lower purchase intentions in the near future as indicated by the negative sign for the car owners with high purchase intentions. In addition, we also find that the individuals who are unsure of their car purchase intentions tend to have significantly lower household income than others. This suggests that those consumers without clear purchase intentions are mainly concerned about their affordability to the

[24] The NLOGIT code for the purchase intention model is available in Appendix 3.3.

car adoption and the daily use. Furthermore, we notice that for the car manufacturers that want to encourage consumers to buy more cars, it is difficult to influence either household income or existing car ownership in order to enhance consumer purchase intentions.

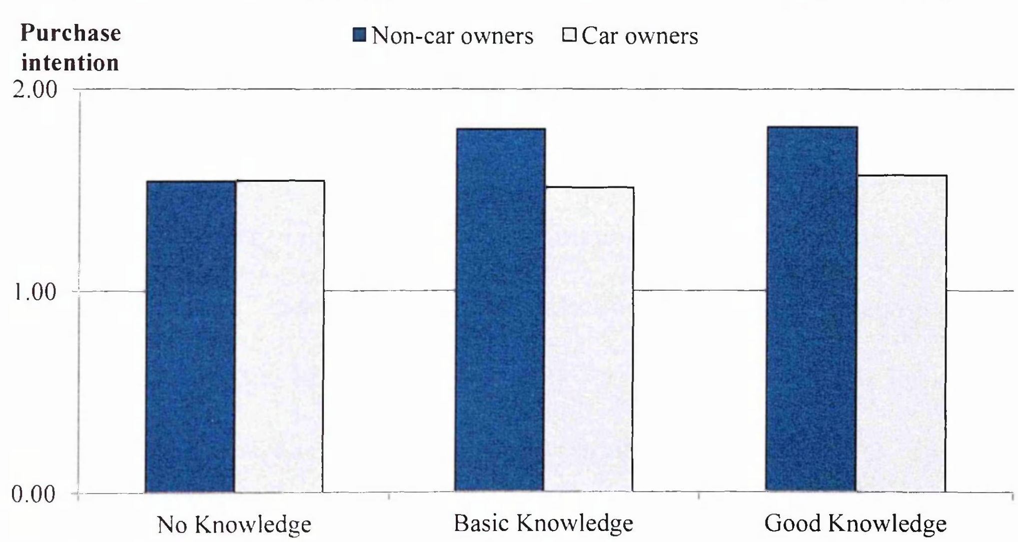
Table 4-14: Estimation parameters of the purchase intention model

	Model 1			Model 2		
Explanatory variable	Coeff.	T-stat.		Coeff.	T-stat.	
Individual characteristics for consumers with unsure purchase intensions						
Constant of unclear intention holders	1.79	1.52		1.62	1.39	
Household income	-0.18	-2.80	***	-0.18	-2.75	***
No. of working adults	0.52	1.61		0.45	1.38	
No. of drivers	0.27	0.33		-0.03	-0.04	
Gender of household head	-0.75	-1.40		-0.86	-1.60	
Age of household head	0.01	0.40		0.01	0.63	
Family size	-0.09	-0.34		-0.08	-0.30	
No. of children	0.34	0.75		0.20	0.42	
Existing car owners	-0.16	-0.24		-0.29	-0.44	
Basic knowledge about cars				1.10	1.96	*
Good knowledge about cars				0.62	0.81	
Individual characteristics for consumers with high purchase intensions						
Constant of high intention holders	2.22	2.02	**	1.62	1.48	
Household income	-0.06	-1.09		-0.07	-1.21	
No. of working adults	0.39	1.28		0.23	0.76	
No. of drivers	1.36	1.75	*	0.86	1.10	
Gender of household head	-0.23	-0.44		-0.55	-1.06	
Age of household head	-0.01	-0.57		0.00	0.07	
Family size	0.06	0.25		0.11	0.47	
No. of children	0.18	0.43		-0.02	-0.05	
Existing car owners	-1.23	-2.01	**	-1.53	-2.41	**
Basic knowledge about cars				1.91	3.56	***
Good knowledge about cars				1.56	2.19	**
No. of observations	524			524		
Log-likelihood at zero	-397.78			-397.78		
Log-likelihood at convergence	-375.87			-362.28		
Rho-square w. r. t. zero:	0.057			0.089		
χ^2 against zero	44.99	(df = 18)[a]		71.00	(df = 22)[b]	
χ^2 against Model 1				26.01	(df = 4)[c]	

Note: [a] the critical value for 18 degrees of freedom is 28.87 at 5% significance level; [b] the critical value for 22 degrees of freedom is 33.92; [c] the critical value for 4 degrees of freedom is 9.49.
*** significant at 1% level, ** significant at 5% level, * significant at 10% level.

As we have discussed previously, the context specific variables are important when understanding the existent local consumer behaviour. Using the consumer knowledge as an example again, we further demonstrate here that the context specific variables can also help us know more about consumer purchase intentions. In the second model as shown in Table 4-14, both income and existing car ownership level remain important, but more importantly we find that consumers with at least some basic knowledge about cars demonstrate significantly higher purchase intentions than those knowing nothing about cars. This can be explained by that fact that the car is a type of expensive durable goods, in particular in the EMs, so that consumers do not want to pay money for it if they totally have no knowledge about it.

Figure 4-5: Average purchase intentions across different knowledge levels



Note: the purchase intentions collected from the survey were coded into 0, 1 and 2 to stand for three levels of intentions.

Furthermore, we differentiate the purchase intentions between car owners and non-car owners, as most consumers in China do not own cars and some preliminary analyses indicate that the car owners may not have different intentions due to their different knowledge levels. Figure 4-5 compares the average purchase intention levels of car owners and non-car owners across different knowledge levels. Since we collected the self-evaluated knowledge from respondents, some car owners reported that they have no knowledge about cars. This is one limitation of using self-evaluated knowledge without references and it will be desirable in the future research to explore the objective knowledge. When comparing purchase intentions across different segments, it is clear that there is a significant gap of purchase intentions for non-car owners without knowledge versus with at least basic knowledge about cars, while the difference between any groups of car owners is insignificant. The Chi-square test in the segment of non-car owners has the P-value of 0.017, which indicates the significant association between non-car owners' knowledge levels and purchase intentions. The Chi-square test P-value for the car owners is 0.672, which suggests the limited influence of the knowledge on car owners' purchase intentions.

In summary, our analysis of purchase intentions further demonstrates the important effects of the context specific variables when investigating consumer purchase intentions. Our finding implies that in the Chinese car market where most consumers are non-car owners, consumer purchase intentions can be significantly enhanced through improving their knowledge level about cars. This is particularly important for the car manufacturers, as they can play an important role in the process of educating consumers through implementing many marketing activities, such as providing 'test driving' opportunities to the potential consumers (Goering, 1985; Lakshmanan & Krishnan, 2011).

4.9 Conclusions and Managerial Implications

With market globalisation, whether firms should adopt the standardised or localised marketing strategies across different markets has been a debate for decades (Albaum & Tse, 2001; Jain, 1989; Ryans, et al., 2003; Theodosiou & Leonidou, 2003; Zou, et al., 1997). In general, the standardisation strategy fits better for the markets that share similar characteristics in terms of economic development stage, government regulation, customs and traditions, consumer characteristics and so on (Katsikeas, et al., 2006; Szymanski, et al., 1993). However, for companies that want to extend the business into a completely different market context, the localisation or adaptation strategy is highly advocated for their success (Cavusgil & Zou, 1994; Johansson, 2009). This is particularly important for established companies in the more mature markets to move into the EMs, because the EMs demonstrate significantly different contextual characteristics from the mature markets that those companies are accustomed to (Alden, et al., 2006; Burgess & Steenkamp, 2006; Sheth, 2011).

In the context of car market, the existing body of empirical studies in the literature extensively focuses on the mature markets, with little research conducted to understand consumer behaviour in the Chinese car market and more importantly to investigate how the consumer behaviour that influences car demand is different in China compared to those mature markets. Furthermore, it is also important to account for context specific factors when understanding the local consumer behaviour, because the specific market context may present additional factors that do not exist in other markets but influence consumer choices in this market. Understanding the behavioural differences with the context specific factors is particularly valuable for multinational car manufacturers who want to invest in China and benefit the growth of the Chinese car market, so that they can effectively learn how the Chinese consumers behave in different ways and are influenced by some new factors that the car manufacturers have never experienced in other markets.

In this chapter, we model car ownership and car type choices. We summarise our main finding and the corresponding managerial implications as follows. In the car ownership model, we find that whether the household head is male or female is insignificant in the car ownership in China, which indicates the existence of a joint decision mechanism between husbands and wives on whether buying their first cars or not. The corresponding implication for the car makers is that they should balance their marketing focus between the male and the female consumers without ignoring the potential influence of the female on the first car purchase decisions. Also, we find that the number of working adults is insignificant in China, which is probably because the number of working adults is more associated with multiple car ownerships in the developed market context while the key decision for the Chinese consumers is whether to buy a car or not.

When investigating consumer preferences for different types of cars, we find more behavioural differences in China and the developed car markets than when looking at adoption. Specifically, our model shows that the Chinese consumers are more concerned about the "tangible" factors such as price/cost and safety instead of the vehicle performance. This finding supports our contention that the Chinese car market is not as mature as in those in the developed markets. In terms of brand preference, we find that the non-Asian brands, particularly those from the Europe, are more popular in China, which implies that the European car makers can emphasise their outstanding brand images when marketing their products, while the Asian car makers should strive to improve their brand values and emphasise other advantages (such as the high fuel efficiency of the Japanese cars) when promoting their products. After analysing consumer preferences, the consumer profiles of different types of cars are available. First time car buyers are more likely to adopt small cars. Small car owners are more likely to live in the central urban areas and to use their cars mainly for other purposes instead of business activities. When consumers live in suburban or remote areas or want to use cars for business purposes, they are more likely to buy mid-sized or large cars. In particular, households with children or owning multiple cars are more likely to own the large cars rather than the small or midsized cars in China. The implication here for car manufacturers is that they should diversify their marketing strategies and customise product positions for different consumers. For example, different promotion policies or pricing strategies might be applicable for the first time buyers versus the existing car owners, based on their heterogeneous preferences towards different types of cars.

In addition to comparing preference differences across different car markets, an important contribution of this chapter is that we account for the context specific variables when understanding local consumer behaviour. Using the consumer knowledge as an example of the context specific variables, we find there are context specific effects on car ownership, car type choice as well as purchase intention models, which indicate that it is critical to incorporate the context specific features in addition to the cross market comparison when analysing the local consumer behaviour. The localisation strategy through better understanding local consumer behaviour is crucial for the success of multinational car manufacturers in China, but its importance goes far beyond the car market. Both McDonald's and Nike has experiences of advertising improperly in China due to their low awareness of local culture (Chan, et al., 2007). The successful experiences of KFC and Amway in China, in comparison to the less successful McDonald's and the failing Best Buy[25], demonstrate that the importance of localised strategies applies to many other consumer markets when international companies well established in the developed markets extend their business into the EMs.

[25] "Consumers: How to find favour down at the shops", 26/10/2011, Financial Times, http://www.ft.com/cms/s/0/e4209f72-f5ac-11e0-824e-00144feab49a.html

CHAPTER 5.

MODELLING HETEROGENEOUS CONSUMER PREFERENCES FOR ALTERNATIVE FUEL CARS IN CHINA[26]

5.1 Introduction

Car use is a significant contributor of air pollution (Fenger, 1999; Potoglou & Kanaroglou, 2007). The rise in the adoption of cars in many developed and emerging markets also raises concerns about the stock of oil to satisfy the demand. According to the 2010 Key World Energy Statistics, world oil consumption ratio in the transport sector increased from 45.3% in 1973 to 61.40% in 2008 (International Energy Agency, 2010a). Given these two main environmental concerns, governments in many countries are implementing policies to encourage producers to manufacture and consumers to buy greener cars.

When the global financial crisis started to seriously affect the car industry in 2009, governments in major developed economies outlined a number of measures to support firms in the industry. Many of those policies focused on stimulating demand but this

[26] This chapter is largely based on: Qian, L. and Soopramanien D. (2011), Heterogeneous consumer preferences for alternative fuel cars in China, *Transportation Research Part D: Transport and Environment*, *16*(8), 607-613.

also represented an opportunity for governments to encourage consumers to buy "greener" cars. The Chinese central government was no exception and announced a set of stimulus policies to encourage consumers to buy greener cars. The largest and most important car show took place in China (Auto China) in 2010 with the theme of "For a Greener Tomorrow", which was another occasion for the country to show that it wants to lead the way in promoting the adoption and use of the greener cars.

In this chapter we empirically model consumers' potential preferences for buying alternative fuel cars (AFCs) in China. The market context of China is interesting because the majority of potential consumers have never owned a car and this can affect how these potential car buyers perceive the choice between conventional type of cars and the AFCs respectively. Previous research on modelling preference towards green cars tends to use one choice structure for multiple alternatives with different types of fuels in their empirical models (See examples in Bunch, et al., 1993; Caulfield, et al., 2010; Potoglou & Kanaroglou, 2007. We will discuss them in detail in the literature review section). In our study, we do not impose any prior assumption about how consumers perceive the different types of cars. The types of cars that are considered by consumers are likely to share some common attributes compared to those that are not considered. In the context of this chapter, this effectively means that if consumers ignore alternative fuel cars when they are deciding to buy a car, then any policy action to encourage them to buy green cars will not have the desired impact. So, it is important to take into consideration how consumers perceive the different types of cars when modelling consumer preferences for the green versus conventional cars.

In this chapter we also consider if different segments of the market behave differently when they have to choose between green and conventional types of cars. In the context of China, as most potential buyers of green cars will be non-car owners, we explore if these consumers are different to car owners in terms of their preferences towards green and conventional types of cars respectively. This comparison has not been given much attention in the literature as most studies in this area focus on car owners only and are conducted in more mature markets where car ownership levels are significantly higher than in China. Another contribution of this chapter is that we examine if consumers differentiate between different types of alternative fuel cars (electric cars and hybrid); other previous work tend to consider the market for green cars as a whole and do not take into account the possibility that some types of green cars may be perceived by consumers to be closer substitutes to conventional types of cars.

The remainder of this chapter is organised as follows. In the next section we briefly review the literature on modelling consumer car preferences and we also provide a brief introduction to the development of green cars in China as well as the corresponding governmental policies. The data and the models that we employ are described in section 5.3 followed by the empirical results and the detailed segmentation analysis. Finally, our conclusions and policy implications are presented in the last section of the chapter.

5.2 Literature Review and Market Background

5.2.1 Literature review

Most research on consumer preferences for alternative fuelled vehicles (AFVs) apply discrete choice models based on either the survey or conjoint experiment data. When some of the attributes of AFVs are either unavailable or limited in the market, choice based conjoint analysis is usually designed and implemented. Various hypothetical choice scenarios are typically presented and the preference for different offerings is explored without actual market data (Louviere, et al., 2000). With conjoint data, discrete choice modelling is used to analyse the relationship between consumers' choices and alternative attributes and consumer characteristics (Train, 2003).

Early research on the potential demand and consumer preference for AFVs were based on survey data collected in California, where air quality has been an important concern since the 1990s (Adler, et al., 2003; Brownstone, et al., 2000; Bunch, et al., 1993; Golob, et al., 1993). For example, a recent study from Adler et al. (2003) investigates conditions and incentives that would encourage California residents to adopt AFVs and these researchers employ a conjoint survey consisting of three hypothetical car choice: a gasoline, a hybrid-electric and a diesel vehicle. They use the NL model and find that that the high purchase cost of AFVs has a negative effect on adoption but that the potential fuel cost savings and incentives such as purchase tax reduction and free parking would encourage Californians to adopt AFVs. In Canada, Ewing and Sarigollu (2000) investigate consumer preferences for low-emission and zero emission vehicles and employ a conjoint survey in the suburban area of Montreal. Based on the MNL model, they find that the vehicle characteristics such as driving range, acceleration and refuelling time are critical factors affecting consumer choice for AFVs but also find that consumers are concerned about environmental issues. In addition, they also explore government regulations and indicate that the price subsidy would be an effective incentive but that increasing gasoline prices or providing faster lanes for AFVs are less important. See also Ewing and Sarigollu (1998), Potoglou and Kanaroglou (2007), and Mau, et al. (2008) for more studies of consumer preferences for AFVs in the context of Canada. Similar studies are also available in the European context (Caulfield, et al., 2010; Dagsvik, et al., 2002; Eggers & Eggers, 2011) and in South Korea (Ahn, et al., 2008; Kim, et al., 2007; Lee & Cho, 2009). A more recent literature review from Potoglou and Kanaroglou (2008a) also shows that the purchase price and maintenance and running costs are two critical factors influencing consumers choice for AFVs. In addition, other factors such as vehicle performance, fuel availability and driving range of AFVs are also found to affect demand. They also suggest that monetary incentives for AFVs buyers as well as potential green attributes of AFVs would be effective in encouraging adoption of such types of cars.

As far as we are aware, our review of the literature indicates that research on the potential demand and consumer preference for AFVs is mostly conducted in the context of developed car markets, which essentially treat all consumers/households as experienced car owners. The only exception is the research of Dagsvik and Liu (2009) which explores a rank-ordered choice model based on conjoint survey data collected in Shanghai in 2001. But their study provided limited insights about local consumer behaviour for the following reasons. Firstly, most urban households in China had no experience at all about conventional car ownership and use at the time of their survey, and thus their responses to alternative fuel vehicle choices are far from reliable. Secondly, consumer demographic variables were not included which means that the influence of different demographics on car preference was not studied. Thirdly, incentives or stimulus policies were not present in their conjoint experiment.

With regard to the modelling approach, different specifications of discrete choice models have been applied in the literature to model the demand for alternative fuel vehicles, such as the MNL model (Ewing & Sarigollu, 1998, 2000; Golob, et al., 1993; Mau, et al., 2008), the NL model (Adler, et al., 2003; Bunch, et al., 1993; Caulfield, et al., 2010; Potoglou & Kanaroglou, 2007), mixed logit model (Brownstone, et al., 2000; Kim, et al., 2007) and rank-ordered models (Dagsvik & Liu, 2009; Dagsvik, et al., 2002; Lee & Cho, 2009). As we have mentioned previously, the NL models have been explored to study how consumers perceive having an additional choice of alternative fuel vehicles in the market. Adler et al. (2003) and Potoglou and Kanaroglou (2007) mainly group alternatives based on the different vehicle types (e.g. car, Van, SUV and Pickup) and/or sizes (e.g. small, medium and large), instead of considering different fuel types. Bunch et al. (1993) only investigate one specific NL structure to group non-electric vehicles into one branch, based on the assumption that preferences for conventional petrol and other fuel types (e.g. methanol, ethanol, propane, or compressed natural gas) might be correlated whilst attitude towards electric vehicle is independent from both. Similarly, there is only one assumed NL structure used by Caulfield et al. (2010) which includes hybrid electric vehicle and alternative fuel vehicle in one nest and keeps conventional vehicle on its own, which assumes correlated consumer preferences between the hybrid and AFVs a priori. In order to fully explore consumer's perceptions about various AFVs, we propose in our study that it is important to test the extent to which AFVs and conventional vehicles are perceived to be substitutes without any priori assumptions. That is, for example, it may be interesting to investigate an alternative tree structure by grouping both electric and other alternative fuel vehicles into one branch because all of them are environmental friendly vehicles.

5.2.2 Background of the green cars in China

AFVs can generally be characterised as a new product concept and even more so in China. The Toyota Prius was first introduced to China at the end of 2005, and a local firm named BYD introduced the first plug-in hybrid car F3DM in 2008. However, neither car model was very successful in China: Cumulative sales of Prius from 2006 to 2009 were less than 3,800 units and total sales of F3DM in 2009 were only 48 units[27]. Moreover, similar to governments in developed car markets, the Chinese government also implemented stimulus policies in 2009 to support the market for alternative fuel vehicles. In January 2009, the "Ten cities and Thousand vehicles" program was initiated in China, which aims to select 10 cities each year and introduce 1000 hybrid, electric or hydrogen fuel cell vehicles in each of them (Huo, et al., 2010). This policy exclusively applies in the public sector to support the green vehicles used for taxis, bus and other public services. In June 2010, the Chinese government announced a subsidy policy for household green car buyers, which is only available for consumers in 5 selected cities (Shanghai, Changchun, Shenzhen, Hangzhou and Hefei). This policy is designed for the period between 2010 and 2012 and mainly supports plug-in hybrid and electric cars. The governmental subsidies for each hybrid and electric car can go up to RMB 50,000 and RMB 60,000 respectively, depending on the battery capacity. The short-term goal of the green car sector development in

[27] According to China Association of Automobile Manufacturers (http://www.caam.org.cn/)

China has been clearly stated in "Automobile Industry Adjustment and Revitalisation Plan" released in March 2009, which requires electric and hybrid cars to achieve 5% market share of the total passenger car sales by the end of 2011[28].

5.3 Methodology and Data

5.3.1 Modelling approach

In this chapter we also apply the discrete choice models where utility maximising behaviour is assumed. The utility function of consumer (i) choosing a car alternative (n) depends on car attributes (X), possible governmental policies (Y) and choice-invariant individual characteristics (Z):

$$U_{in} = V_{in} + \varepsilon_{in} = \alpha' X_{in} + \beta' Y_{in} + \gamma_n' Z_i + \varepsilon_{in} \tag{5-1}$$

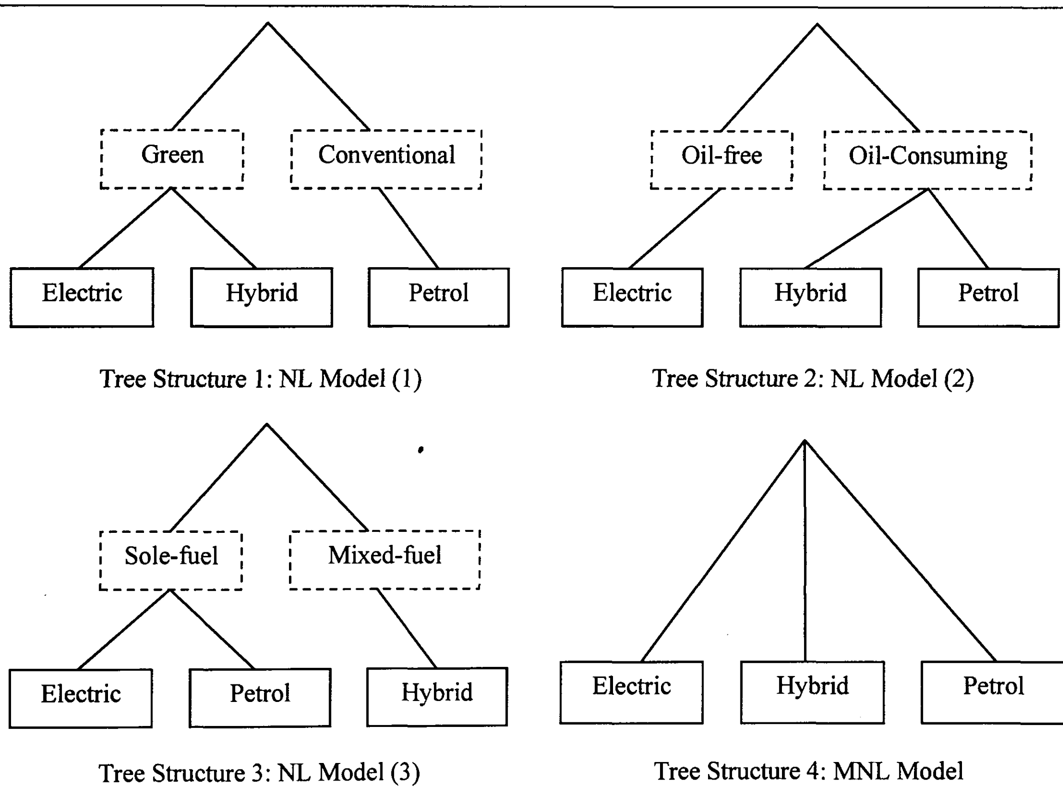
where utility (U_{in}) consists of the deterministic portion (V_{in}) and the error unexplained component is (ε_{in}).

The assumption on the error component (ε_{in}) dictates the specification of the discrete choice model. We first consider the MNL model (McFadden, 1974), which assumes that the error term follows the type I extreme value distribution or Gumbel distribution where the independence from irrelevant alternatives (IIA) is assumed. This effectively means that the MNL model assumes that the consumer perceives all the alternatives to be completely different from each other. If consumers do however perceive that some alternatives in the choice set are more similar than others, then the

[28] The State Council of China (http://www.gov.cn/zwgk/2009-03/20/content_1264324.htm)

IIA property does not hold and we have to consider models that allow for the possibility that consumer preferences for some alternatives are correlated. In this research we consider the NL model (Ben-Akiva & Lerman, 1985; Daly & Zachary, 1978; McFadden, 1978) and this assumes that the error term in equation (5-1) follows a type of generalised extreme value (GEV) distribution (Train, 2003). The NL model allows for the grouping of similar alternatives into a nest thereby relaxing the IIA assumption.

Figure 5-1: Tree structures of MNL model and three nested logit models



We propose different tree structures, as depicted in Figure 5-1, based on the different preference assumption towards alternative fuel cars and conventional types of cars. In the first structure, we assume that both hybrid and electric cars are perceived to be environmentally friendly cars and thus are grouped together to form a "Green" nest/group. In the second structure, preferences for hybrid and petrol cars are assumed to be positively correlated since both of types of cars consume oil (although at different efficiency levels), while the electric vehicle is completely oil-free. We also explore another structure as illustrated in tree structure 3 in Figure 5-1, which assumes that petrol and electric car alternatives are perceived to be similar given that both of them are solely fuelled. We have to note that the MNL model is in fact the restricted specification of the NL model. This means that we can obtain the MNL model from a NL model by simply making the assumption that all alternatives are independent as in tree structure 4.

5.3.2 Data

As we have discussed in Chapter 2, we conducted the choice based conjoint analysis experiment as part of our online survey in China. Students from two Chinese universities (North China Electric Power University and China University of Mining and Technology University) helped us to collect data during their winter holiday. When the students went back to their home cities, they visited local households with an invitation that described our research objectives and how to fill in the survey online, which helped us access households in different regions in China. We obviously, at the outset, recognised that an online survey could potentially bias the sample towards those who have computers and the Internet. So the students also collected data from those households who did not have access to a computer/the Internet. As part of this survey we also collected additional information such as demographic characteristics and car ownership status. After deleting some cases where information on key variables was missing, we are left with 527 usable cases for modelling consumer preferences towards alternative fuel cars in this chapter.

Table 5-1: Descriptive statistics for the survey sample

Variables	Our Sample	National Sample
Average family size	3.24	2.89
Average no. of children less than 18 years old	0.46	/
Average no. of working members	2.09	1.49
Average household disposable Income in 2009 (RMB)	124,900	49,635
No. of cars owned per 100 households	66.60	10.89
Average age of the household head (Years)	37.76	/
Proportion of male household head (%)	77.99	/
Average distance from home to work place for household head (km)	9.15	/
Sample size	527	65,506

Data Source of National Sample: China Statistical Yearbook 2010

The demographic characteristics of our sample are summarised in Table 5-1. When we compare our survey with the 2009 national urban and township household survey data reported in the China Statistical Yearbook 2010, there is no significant difference in family size but high-income groups and those who own cars are over represented. We therefore reweight our data based on household income and car ownership information from the national sample when we estimate the empirical models.

In the choice based conjoint analysis, respondents are presented scenarios of hypothetical cars with different fuel types: one conventional petrol car and two potential AFCs, hybrid and electric cars. We constructed these three choice alternatives of cars based on a list of attributes with three levels (Louviere, et al., 2000) (see Table 5-2).

Table 5-2: Attributes and levels in the choice-based conjoint analysis

Attributes	Vehicle 1	Vehicle 2	Vehicle 3
• Fuel Type	Petrol	Hybrid	Electric
• Purchase price (RMB)	Specified by the respondent	(1). 30% higher than the similar-sized petrol car (2). 50% higher than the similar-sized petrol car (3). 80% higher than the similar-sized petrol car	(1). 30% higher than the similar-sized petrol car (2). 50% higher than the similar-sized petrol car (3). 80% higher than the similar-sized petrol car
• Annual running cost (RMB)	Market average level based on vehicle price	(1). 20% less than the similar-sized petrol car (2). 40% less than the similar-sized petrol car (3). 60% less than the similar-sized petrol car	(1). 20% less than the similar-sized petrol car (2). 40% less than the similar-sized petrol car (3). 60% less than the similar-sized petrol car
• Availability of charging facility	NA	NA	(1). 10% of parking spaces (2). 40% of parking spaces (3). 70% of parking spaces
• Vehicle range with full charging	NA	NA	(1). 80km (2). 120km (3). 160km
• Incentives	NA	(1). 20,000 RMB allowance (2). Eligible for priority lane (3). Free Parking for 5 years	(1). 30,000 RMB allowance (2). Eligible for priority lane (3). Free Parking for 5 years

The selection of attributes and levels were largely based on the wide review of literature (Adler, et al., 2003; Bunch, et al., 1993; Ewing & Sarigollu, 1998, 2000; Potoglou & Kanaroglou, 2007, 2008a). The selected attributes were also adjusted based on our knowledge on local market. For example, as we found previously, the Chinese consumers are not highly concerned about the vehicle performance, so we did not include the performance attribute (e.g. accelerating time or speed) in the conjoint analysis. Also, based on the real price information we learned from the local market, the price of green cars was assumed to be up to 80% higher than that of the petrol cars in our study, while the literature studies usually allow for the price gap of 10% or 20% only. Three levels of attributes are usually used in the literature, which can accommodate the non-linear effects (Hensher, et al., 2005) but will not significantly increase the total number of choice profiles or combinations of conjoint analysis. The selected vehicle attributes included the purchase price, annual running cost (including fuel and maintenance costs), and charging, convenience and driving range for the electric vehicle. Furthermore, we designed three different incentives for potential buyers of hybrid or electric cars: one cash subsidy and two non-monetary incentives (fast lane or free parking). The cash subsidy for the electric car is slightly higher than that for the hybrid car because the electric car normally consumes no oil without any direct pollution. This difference in cash subsidy is similar to what the government is planning to offer. Priority lane and free parking policies are not yet available in the Chinese market but we investigate consumers' preferences towards such policies to test if and how consumers would respond to such incentives. Based on these attributes and levels, the complete experiment design has a total of 3^8=6561 combinations or scenarios, which are unrealistic to implement. We therefore derive the orthogonal fractional design with 32 scenarios using SPSS 17.0 (SPSS Inc.), which reduces the complexity of the tasks of the stated choice experiment for consumers.

We first introduce our experiment with detailed explanations about the three alternatives followed by their attributes and different incentives. Then 8 choice scenarios are randomly selected from a total of 32 scenarios and presented to each respondent, where the purchase price and running costs of the base alternative (petrol car) are customised based on the respondent's preference stated beforehand in the survey. The respondents are then asked to select one vehicle from each scenario. A sample choice scenario is shown in Figure 5-2.

Figure 5-2: A sample of choice scenarios

	Vehicle 1	Vehicle 2	Vehicle 3
Fuel Type	Petrol	Hybrid	Electric
Purchase price (RMB)	75,000	97,500	135,000
Annual running cost (RMB)	15,000	9,000	12,000
Incentives	Not Applicable	Priority Lane	30,000 RMB subsidy
Availability of charging facility	Not Applicable	Not Applicable	40% of parking slots
Vehicle range with full charging	Not Applicable	Not Applicable	80 km
Your Choice:	☐	☐	☐

5.4 Empirical Analysis

5.4.1 Estimation results

In our models, the alternative specific attributes include purchase price, annual running cost, charging facility and vehicle range for electric cars. The incentive variables consist of three types of incentives. We also account for the demographic characteristics of the households. We estimate the MNL model and the three different specifications of the NL models based on 4216 observations from 527 respondents using NLOGIT 4.0 (Greene, 2007)[29]. The results are presented in Table 5-3.

[29] The NLOGIT code for the MNL and NL models is available in Appendix 3.4.

Table 5-3: Parameter estimation of the MNL and NL models

Variable	MNL		NL (1)		NL (2)	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-test
Vehicle Attribute (α)						
Purchase price	-0.014	-16.018	-0.017	-13.253	-0.014	-15.433
Running Cost	-0.059	-7.263	-0.057	-6.835	-0.064	-7.596
Charging Facility	0.933	5.230	0.971	5.333	0.926	5.183
Vehicle Range	0.005	3.982	0.005	4.020	0.005	4.014
Incentive (β)						
Cash Subsidy	0.019	1.739	0.020	1.800	0.017	1.518
Free Parking	0.412	1.534	0.429	1.548	0.364	1.331
Priority lane	0.165	0.618	0.176	0.641	0.114	0.419
Individual characteristics for electric car (γ_1)						
Electric car, constant	-1.307	-3.113	-1.392	-3.203	-0.986	-2.33
No. of young children	-0.583	-5.464	-0.638	-5.638	-0.502	-4.682
No. of drivers	-0.434	-6.227	-0.474	-6.375	-0.364	-5.248
Family size	0.296	5.873	0.314	5.915	0.272	5.488
Household Income	0.004	2.142	0.004	2.027	0.004	2.649
No. of owned cars	0.296	1.869	0.294	1.776	0.349	2.238
Age of family head	-0.009	-1.759	-0.010	-1.746	-0.009	-1.821
Sex of family head	-0.821	-7.627	-0.855	-7.526	-0.821	-7.876
Working distance of family head	-0.032	-4.879	-0.034	-4.984	-0.029	-4.457
Individual characteristics for hybrid car (γ_2)						
Hybrid car, constant	0.866	2.771	0.827	2.525	0.869	2.752
No. of children	-0.500	-6.492	-0.559	-6.508	-0.516	-6.676
No. of drivers	-0.315	-6.158	-0.349	-6.166	-0.324	-6.302
Family size	0.049	1.284	0.064	1.546	0.055	1.436
Household Income	0.003	2.643	0.003	2.407	0.003	2.564
No. of owned cars	0.371	3.323	0.373	3.113	0.362	3.231
Age of family head	-0.007	-1.834	-0.007	-1.792	-0.007	-1.777
Sex of family head	-0.212	-2.473	-0.236	-2.547	-0.228	-2.648
Working distance of family head	-0.025	-5.528	-0.028	-5.506	-0.025	-5.464
Inclusive value (IV) parameters (λ)						
Green nest in NL(1)			0.920 [a]	32.196		
Oil-consuming nest in NL(2)					0.772 [b]	15.331
No. of observations	4216		4216		4216	
Log-likelihood at convergence	-3952.366		-3949.104		-3940.251	
Log-likelihood at zero	-4631.749		-4631.749		-4631.749	
Rho-square w. r. t. zero:	0.1467		0.1474		0.1493	
χ^2 against zero	1358.766 (df=25)		1365.290 (df=26)		1382.996 (df=26)	
χ^2 against MNL			6.524 (df = 1)		24.230 (df = 1)	

Note:

[a] St. Error of the IV parameter in NL(1) is 0.029, thus its Wald test against one is (0.920-1)/0.029 = -2.76

[b] St. Error of the IV parameter in NL(2) is 0.050, thus Wald test against one is (0.772-1)/0.050 = -4.56

5.4.1.1 MNL model

The log-likelihood ratio index (Rho-square) value of our MNL model is 0.1467 and falls within the range for models which have been reported in the literature: 0.129 in Ewing and Sarigollu (1998) and 0.1948 in Golob et al. (1993). The log-likelihood ratio test yields a chi-square value of 1358.77 with 25 degrees of freedom, which implies that this MNL model is significantly different from the model without any variables. We then apply the Hausman test (Hausman & Mcfadden, 1984) which tests the IIA assumption and at the 5% significance level we reject the hypothesis that consumers perceive the different car alternatives to be completely independent. This indicates that we need to explore models that account for correlated preferences of alternatives and for this reason we will not pay too much attention to the specific parameters of the MNL but will focus more on the NL models.

5.4.1.2 NL models

The NL models are estimated based on the three tree structures as shown in Figure 5-1. When estimating our NL models, it is worth noting that all these three structures have degenerate branches with only one alternative within each of them. Therefore, the corresponding two-level NL models must use the RU1 specification that normalises the lower level of the tree structure and imposes the IV parameter of the degenerate branch to be one (Hensher & Greene, 2002; Hensher, et al., 2005).

A key parameter of the NL model is the Inclusive Value (IV thereafter) parameter and this has to lie between 0 and 1 for the model to be consistent with the utility maximisation as described by the tree structures (Greene, 2009). The IV parameter is also an indicator of independence of all alternatives in the nest and $1-\lambda^2$ is the correlation index between pairs of alternatives within the same nest (Ben-Akiva & Lerman, 1985). Amongst the three choice structures for NL models, the third structure that groups electric and petrol cars together yields an IV parameter which is significantly greater than one, which indicates its inconsistency with random unity maximisation (RUM) criterion. Therefore, we only present the parameter estimation results for the NL models based on tree structures (1) and (2) in Figure 5-1. The IV parameters for these two structures are 0.920 for the "Green" nest and 0.772 for the "Oil-consuming" nest respectively. Two IV parameters are also significantly different from zero at 1% significance level. Moreover, the Wald-test statistics, calculated in the footnote section of Table 5-3, show that they are also significantly less than one. These results concur with the results of the Hausman test where we reject the IIA assumption.

The respective performances of both NL models are further investigated by performing log-likelihood tests $\left(-2(\mathrm{LL}_{\mathrm{base}}-\mathrm{LL}_{\mathrm{NL}})\right)$. Compared with the MNL model, both NL models have much higher test values (6.524 and 24.230) than the critical $\chi^2_{(1)}$ value of 3.841 at 5% significance level. For two NL models with same number of parameters, we compare their log-likelihood values and find that the second tree structure, which groups hybrid and petrol cars in "oil-consuming" nest, is better, because its log-likelihood value (-3940.251) is higher than that under the first tree structure with a "green" nest (-3949.104). This would indicate that consumers perceive petrol cars and hybrid cars to be more similar to each other compared to electric cars.

In Table 5-3 the parameters of the vehicle attributes are significant and the effects are similar to those reported in the literature (Potoglou & Kanaroglou, 2008a). More specifically, the parameters of monetary attributes such as purchase price and running cost have negative signs, which imply consumers would prefer low-priced cars and low running costs. The positive parameters of charging facility and driving range indicate the importance of providing more charging facilities and designing electric cars with longer range batteries. With regard to the various incentives, the estimated parameters of all three incentives are insignificant which contrasts with previous research conducted in more mature markets. For example, both Ewing and Sarigollu (2000) and Potoglou and Kanaroglou (2007) find that a price subsidy such as waiving purchasing tax is a significant incentive to buy green cars in Canada. We contend that the insignificant effects of incentives have much to do with the fact that the Chinese car market is fairly "young" and that government policies on Green cars will probably take time to have an impact. Since cars are still a type of new products for most Chinese consumers, it will be a safer choice for them to buy a conventional petrol car as the first car given its mature technology and widely available service. Therefore, in addition to providing short term incentives such as cash subsidy, governments and car manufacturers should strive to decrease the price of green cars and also develop a user-friendly environment, so that green car buyers do not need to worry about other issues, such as how safe the battery is and where to charge battery.

With regard to the impact of the individual characteristics, households with young children and where there is more than one driver prefer conventional petrol cars. This may reflect consumers' concerns about the technology maturity (new concept) of

AFCs in the short term and that they believe that it would be safer to buy a conventional vehicle, particularly when they have young children and/or where there is a less experienced second driver at home. We also note that larger households tend to prefer alternative fuel cars, particularly electric cars. As expected, households with higher incomes are more willing to adopt alternative fuel cars as can be seen by the significantly positive effect of household income variable when interacted with either electric or hybrid cars in both NL models. We also tried to use dummy variables to account for the non-linear income effects, but their inclusion does not substantially change the main findings of consumer preferences for AFCs as well as the model performance.

In addition, households with a male head are less likely to adopt green cars and young household heads are found to have stronger preferences towards AFCs, although these effects are only significant at the 10% level. We also note that household heads who are long-distance commuters are less likely to choose AFCs, possibly due to their concerns about the lack of fuelling facilities and limited vehicle range, which is similar to other findings in the literature (Ewing & Sarigollu, 1998, 2000; Potoglou & Kanaroglou, 2007). The empirical results also show that car owners are more likely to switch to alternative fuel cars compared to non-car owners. This provides an interesting insight for the government and car manufacturers about the demand potential green cars in China.

To demonstrate why it is important to choose a model that appropriately fits the market context, we compare the elasticity effects between MNL and NL models respectively in Table 5-4. These elasticity effects are computed from the parameters so we do not need to go into the detail of the effects since we have already commented on the parameters of the model. Furthermore, since we have shown that the second tree structure is the best NL model, we simply compare the resulted elasticity effects with those from the MNL model. We note in the first instance that the second tree structure (see Figure 5-1) allows the cross-elasticities in the NL to vary for the hybrid or petrol cars across alternatives compared to the MNL. The magnitude of the elasticity effects further shows that hybrid and petrol cars are perceived by consumers to be similar. We note to that effect that the cross elasticity effects of the attributes for hybrid cars have a bigger impact on the probability of buying a petrol car than on the probability to buy an electric car.

Table 5-4: Cross-elasticity comparison between MNL and NL models

		Cross elasticity in MNL Model		Cross elasticity in NL Model (2)	
	Changing Attribute	Hybrid	Petrol	Hybrid	Petrol
Electric car	Purchase Price	0.415	0.415	0.368	0.368
	Running Cost	0.106	0.106	0.108	0.108
	Cash Subsidy	-0.044	-0.044	-0.039	-0.039
	Free parking	-0.016	-0.016	-0.014	-0.014
	Priority Lane	-0.005	-0.005	-0.004	-0.004
	Charging facility	-0.049	-0.049	-0.049	-0.049
	Vehicle Range	-0.083	-0.083	-0.083	-0.083
		Electric	Petrol	Electric	Petrol
Hybrid car	Purchase Price	1.206	1.206	0.92	1.232
	Running Cost	0.312	0.312	0.263	0.354
	Cash Subsidy	-0.078	-0.078	-0.053	-0.072
	Free parking	-0.046	-0.046	-0.031	-0.042
	Priority Lane	-0.016	-0.016	-0.009	-0.012
		Electric	Hybrid	Electric	Hybrid
Petrol car	Purchase Price	1.014	1.014	0.772	1.025
	Running Cost	0.600	0.600	0.509	0.677

5.4.2 Segmentation analysis

As we have discussed previously, we also explore if different types of consumers have heterogeneous preferences towards AFCs and conventional types of cars. We are interested in two segments in the case of the Chinese car market: car owners and non-car owners. Compared to other important car markets such as the USA, the low car ownership rate is a distinguishing feature of the Chinese car market. We estimate separate NL models for the two segments based on tree structures (1) and (2) in Figure 5-1. We do not present the full set of parameters but we compare the key IV parameters and the elasticity effects for these two segments.

Table 5-5: Comparison of IV parameters of segmental NL models by car ownership

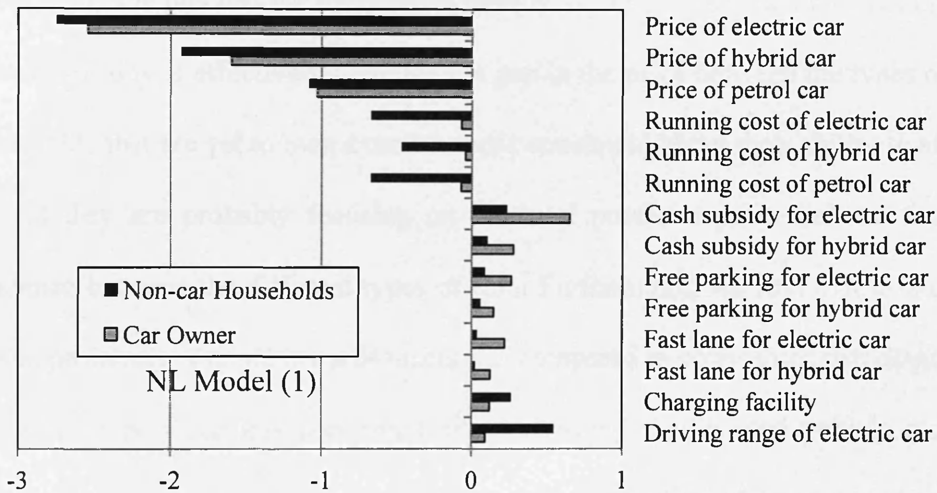
	NL Model (1)		NL Model (2)	
	IV Parameter (λ)	s.e. of (λ)	IV Parameter (λ)	s.e. of (λ)
Car Owners	0.578	0.058	0.985	0.185
Non-car households	0.948	0.037	0.725	0.063

For each segment, the estimated IV parameters for the two different tree structures are summarised in Table 5-5. For car owners, NL model (1) is better than the NL model (2) as the IV parameter in NL(1) is significantly different from both zero and one while the IV parameter in NL(2) is insignificantly different from one, which implies that car owners do consider hybrid and electric cars to be similar[30] and perceive petrol car to be a distinct alternative. However, this is not the case for those households who do not own cars. The IV parameter of NL(1) in this segment is insignificantly

[30] The correlation index is 0.666 ($1-0.578^2$).

different from one while the IV parameter of NL(2) for non-car owners' is significantly different from both one and zero, which implies a stronger preference correlation between hybrid and petrol cars for non-car owners. In other words, we find evidence that car owners and non-car owners do not think about the AFCs and conventional types of cars in the same way.

Figure 5-3: Direct-elasticities comparison based on car ownership status



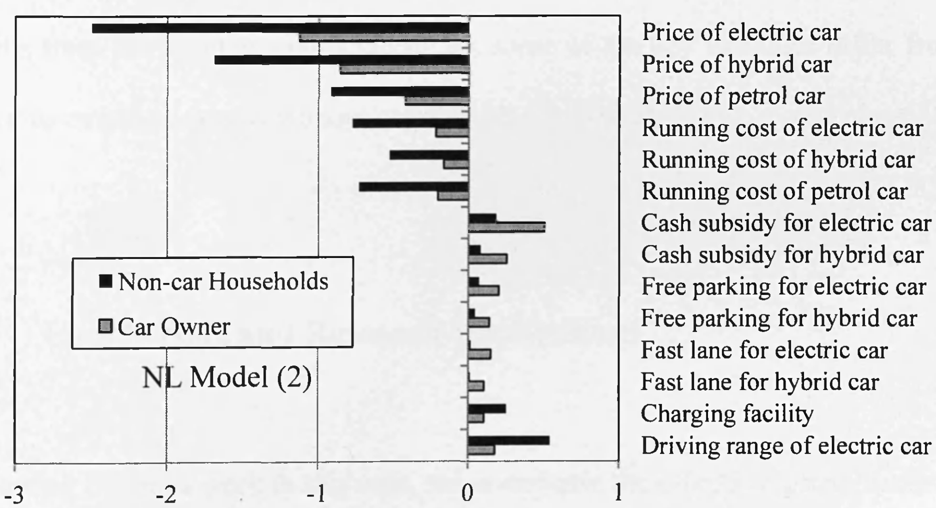


Figure 5-3 depicts and compares the elasticity effects for car owners and non-car owners for both NL models. We find that non -car owners are more sensitive than car owners to monetary attributes such as purchase price and annual running cost (compare the magnitudes of coefficients). This implies that existing car owners are less concerned by the high purchase price of AFCs compared to the non-car households. Our results suggest that car owners are less sensitive to the purchase price but more responsive to the cash subsidy than non-car households and this can be explained by the fact that car owners can already afford conventional types of cars and the cash subsidy is effectively reducing the gap in the price between the types of cars. Households that are yet to own a car are more concerned about their ability to afford a car and they are probably focusing on the total purchase price and not the price difference between the different types of cars. Furthermore, we find that existing car owners positively respond towards incentives compared to households that do not own cars. Car owners are less concerned about charging facility and vehicle range of electric cars can be explained by their better understanding of how to estimate the daily use of cars and assess whether the range is sufficient. In summary, the empirical results from the segmentation analysis for some of the key variables differ from the previous empirical results without segmenting the market.

5.5 Conclusions and Research Implications

Following previous work in this area, we investigate the effects of price, fuel/running cost and government incentives on consumer preferences to buy alternative fuel cars. Whilst the market context of the Chinese car market is interesting on its own right, we have aligned the modelling contributions of the chapter to some of the unique market features of that market. In that respect, this chapter pays particular attention to the comparison of the different empirical choice models and demonstrates the importance of choosing those models that fit a specific market context.

Unlike previous research, we do not assume a-priori how consumers perceive the choice between alternative fuel cars and conventional types of cars. Previous work in this area has also not explored whether consumers differentiate between the different types of alternative fuel cars. Our empirical results indicate that the IIA assumption of the multinomial logit model does not reflect how consumers in China think about the choice between alternative fuel cars and conventional type of cars. Chinese consumers do consider green and conventional types of cars when they are buying cars. However, they are more likely to consider hybrid cars and conventional types of cars but are less likely to consider electric cars. So if consumers are going to switch to greener types of cars, they are more likely to buy hybrid cars rather than electric cars.

Some of our empirical results are similar to those reported in previous research: Consumers are concerned about the relatively higher purchase price and running costs of alternative fuel cars and the lack of charging facilities and the range limitation of these types of cars are additional limiting factors. Monetary incentives generally have been found to have a significantly positive effect on consumer preferences for such types of cars in developed markets but we find that this is not the case in China. However, importantly, when we segment the market and consider car owners and non-car owners separately, the impact of the monetary incentives to buy alternative fuel cars differs between the two groups of consumers. Consumers who already own cars are more likely to react positively to these incentives to buy alternative fuel cars compared to non-car owners who are probably more likely to buy a new car in the short term. This raises an important policy implication for the Chinese government who needs to look more deeply into the reasons as to why those consumers who do not own cars are not motivated by the incentives to buy alternative fuel cars.

CHAPTER 6.

A DYNAMIC SEGMENTATION APPROACH TO FORECAST NEW PRODUCT DEMAND IN THE EMERGING MARKETS: THE GREEN CARS IN CHINA

6.1 Introduction

With the increasing importance of emerging economies such as the BRICs (Brazil, Russia, India and China), multinational corporations (MNCs) have launched new products almost simultaneously in the emerging markets (EMs) and in the developed markets. For instance, the time lag of Apple's iPhone3 sold in between China and U.S. was up to 854 days, but the gap has been significantly shortened to 167 days for iPad and 93 days for iPhone4 more recently[31]. Furthermore, governments and firms in the EMs also strive to develop innovative products by themselves in order to achieve their own advantages in the global competition.

Alternative fuel cars, or generally green cars, are such type of new products that are of great strategic interests to China. Petrol cars are currently dominating the market, but

[31] "Apple Gets Quicker In China", The Wall Street Journal, 20th September 2010, http://blogs.wsj.com/digits/2010/09/20/apple-gets-quicker-in-china/

the Chinese government is very keen to encourage consumers to adopt two types of alternative fuel cars, hybrid and electric cars. All car manufacturers, no matter whether they are multinational or domestic firms, have publicized their ambitions to develop or sell the greener cars in China in soon future. For example, Volkswagen has collaborated with its two Chinese partners to develop new brands specifically for electric cars[32]. General Motors also plans to develop its electric cars locally through a joint-venture company in China[33]. As a local pioneer, BYD Auto has already started to sell its full electric cars (E6) to the Chinese consumers[34].

It is a challenge in all markets to forecast demand for new products. When the markets are constantly and rapidly changing, it becomes more difficult to do so. In the context of Chinese car market, it is a big challenge for car manufacturers to forecast the demand potentials of these new products. It is because the existing body of marketing research has been largely based on the developed markets, while the EMs demonstrate significantly different institutional characteristics (Burgess & Steenkamp, 2006). More specifically, the EMs have higher level of within-market diversity or preference heterogeneity (Alden, et al., 2006; Batra, 1997; Burgess & Harris, 1999; Zhang, et al., 2008) and a segmentation approach is recommended to compare results across segments (Burgess & Steenkamp, 2006). Furthermore, market dynamics in the EMs should also be accounted for, which refer to their rapid socioeconomic changes during this period of economic growth (Batra, 1997; Burgess & Steenkamp, 2006). This also means that the segments that are constructed to address the preference heterogeneity

[32] "VW-SAIC venture to produce electric car", Automotive News China, 27th September 2011, http://www.autonewschina.com/en/article.asp?id=7621

[33] "G.M. Plans to Develop Electric Cars With China", The New York Times, 20th September 2011, http://www.nytimes.com/2011/09/21/business/global/gm-plans-to-develop-electric-cars-with-chinese-automaker.html

[34] "BYD Starts Sales of E6 Electric Car to Individuals in China", Bloomberg News, 26th October 2011, http://www.bloomberg.com/news/2011-10-26/byd-starts-sales-of-e6-electric-car-to-individuals-in-china.html

in the EMs may not be static over time. That is to say, the size of each pre-defined segment is more likely to change over time in the EMs, instead of the constant segment sizes usually assumed in the developed markets (see an example in Robertson, et al., 2007).

In order to forecast the demand for new and innovative products with limited history, some recent studies propose a market simulation approach based on conjoint analysis (Eggers & Eggers, 2011; Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006). With the conjoint experiment data, these studies normally employ one specification of discrete choice model for the whole sample and then conduct market simulation based on the changes of assumed attributes or demographic variables in the future. Although such non-segmental approach provides some valuable directions about how to forecast demand for new products, it cannot sufficiently account for the institutional characteristics of the EMs. In particular, such a non-segmental approach might not be able to sufficiently address the high level of preference heterogeneity typically found in the EMs, because it cannot account for different parameters for the same variable across different groups of people. Therefore the segmental approach, which can accommodate different parameters across several segments, is proposed to address the heterogeneous preferences in the EMs (Burgess & Steenkamp, 2006).

In this chapter, we address and demonstrate the challenge and the value of accounting for the preference heterogeneity and market dynamics in forecasting demand for new products in the context of EMs. We propose a dynamic segmentation approach that can account for not only the heterogeneity across segments, but also within-segment heterogeneity through allowing for different choice specification for each segment.

Furthermore, our approach also addresses the market dynamics, because we consider the dynamic change of segment sizes (i.e. dynamic segmentation) in addition to the market simulation based on the changes of product attributes and demographics over time. When we apply this approach to forecast the demand for alternative fuel cars (hybrid and electric cars) in the Chinese car market, we define two segments, car owners and non-car owners, because consumers in these two different segments are likely to demonstrate different preferences towards the alternative fuel cars. At the same time, we account for the fact that the segment sizes are dynamic over time, as an increasing number of Chinese consumers are switching from non-car owners to car owners, which can be modelled using a diffusion process. We then combine the segment-specific choice models with the segmentation diffusion model so that we can forecast market shares of alternative fuel cars in China.

To demonstrate the benefit of our approach we define a benchmark model. The benchmark model in our research is based on non-segmental approach from the existing literature (Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006). Due to its non-segmental structure, the benchmark model cannot address heterogeneity across segments. Moreover, it fails to allow for different choice structures for different segments, so that the heterogeneity of consumer behaviour within each segment cannot be appropriately accommodated in the benchmark model. In comparison, our proposed approach can allow for flexible choice structures in different segments, which can more appropriately account for consumer behaviour. In addition, our study designs and tests different scenarios of market change in the future, including the changes of vehicle price and fuel price, to demonstrate how the diffusion of green cars will be affected by different factors and also allow car manufacturers to effectively use the approach to understand the developing market.

The remainder of this chapter is organised as follows. In the next section we briefly review the literature on modelling and forecasting demand for new products. Section 6.3 presents the proposed modelling methodology followed by the specification of an empirical application in Section 6.4. The empirical forecasting results, scenario analysis and model validation are presented in Section 6.5 with the comparison to those from the benchmark model. The final section ends the chapter with our conclusions and research insights.

6.2 Literature Review

When modelling the demand and adoption of new products, diffusion models have been developed since 1960s at the aggregate level (Bass, 1969; Gregg, et al., 1964; Mahajan, et al., 2000; Meade & Islam, 2006). Diffusion models are recommended when data is limited and have been widely applied in many areas, such as automobile (Bouachera & Mazraati, 2008; Dargay & Gately, 1999; Dargay, et al., 2007; Kobos, et al., 2003), telecommunication (Robertson, et al., 2007; Sundqvist, et al., 2005; Wu & Chu, 2010), and other durable goods (Bass, 1969; Bottomley & Fildes, 1998; Tsai, et al., 2010). However, diffusion models have some limitations (Lee, et al., 2006; Urban, et al., 1990). Firstly, diffusion models are usually used to forecast demand at aggregate level so that they cannot directly accommodate competitions among different products or between new and mature products. Secondly, all diffusion models need some historical data for parameter estimation so that forecasting demand for brand new products remains a challenge using such data. In addition, multi-generation diffusion method or the analogy approach extends the simple diffusion models based on the assumption of similar diffusion patterns across different generations of products (Islam & Meade, 1997; Jun & Park, 1999; Kim, et al., 2005; Mahajan & Muller, 1996; Norton & Bass, 1987). However, the validity of this assumption could be a limitation in EMs where there is a high degree of market dynamics, so that the diffusion pattern of new products may differ significantly from that of the mature products. More importantly, Goodwin, Dyussekeneva & Meeran (forthcoming) find that the use of such analogy approach may lead to high forecasting errors particularly in highly dynamic markets, which constrain its capability to forecast demand of new products in the EMs

At a disaggregated level, discrete choice models based on survey data and conjoint analysis are widely used to investigate consumers' stated preference (SP) and model the potential demand for new products, such as alternative fuel vehicles (AFVs) (Dagsvik & Liu, 2009; Dagsvik, et al., 2002; Eggers & Eggers, 2011; Ewing & Sarigollu, 2000; Lee & Cho, 2009; Potoglou & Kanaroglou, 2007; Qian & Soopramanien, 2011). However, the conjoint analysis also has its limitations in terms of its forecasting insights that it yields (Wittink & Bergestuen, 2001). Firstly, researchers only take a snapshot of consumer behaviour in each conjoint experiment, so that changing patterns of utility variables over time are not captured. Although dynamic discrete choice models can potentially capture these changes over time (Aguirregabiria & Mira, 2010), they require panel data of revealed preference (RP), which are typically unavailable for new products. Secondly, all conjoint experiments are commonly based on the assumption that all alternatives are fully available in the market, which is far from the market reality and this gap also impacts the forecasting performance of conjoint analysis. In our study, we propose to use the diffusion models to accommodate different supplies of the products in the market.

Some recent studies have combined a discrete choice model with a diffusion model to forecast the potential of new products (Jun & Park, 1999; Kumar, et al., 2002; Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006), or have proposed a scenario-based conjoint adoption model to forecast the adoption of green cars (Eggers & Eggers, 2011). We adopt the same approach in our study but we address its following limitations to fit better with the requirements of the EMs. First, these studies all employ a non-segmentation approach by estimating a discrete choice model for the whole sample, so that the consumer preference heterogeneity across segments cannot be sufficiently accounted for. Secondly, the discrete choice models used in these studies do not verify whether the choice process matches the specification of the model. Both Lee, et al. (2006) and Lee, et al. (2008) use rank ordered logit model for consumer choices for television sets and home networking products respectively. But there is no obvious sequence or order between different types of TV sets (e.g. CRT, projection and LCD TVs) or home networking solutions (e.g. Ethernet, PLC and Wireless LAN). In addition, Eggers & Eggers (2011), Jun & Park (1999) and Kumar, et al. (2002) commonly employ a multinomial logit (MNL) model specification to model the choices from several different alternatives without testing the independence from irrelevant alternatives (IIA) property or exploring the potential correlations between some alternatives. In summary, as we demonstrate in this chapter, it is important to know if the empirical model that we use reflects the way in which consumers make decisions; otherwise we will obtain biased forecasts.

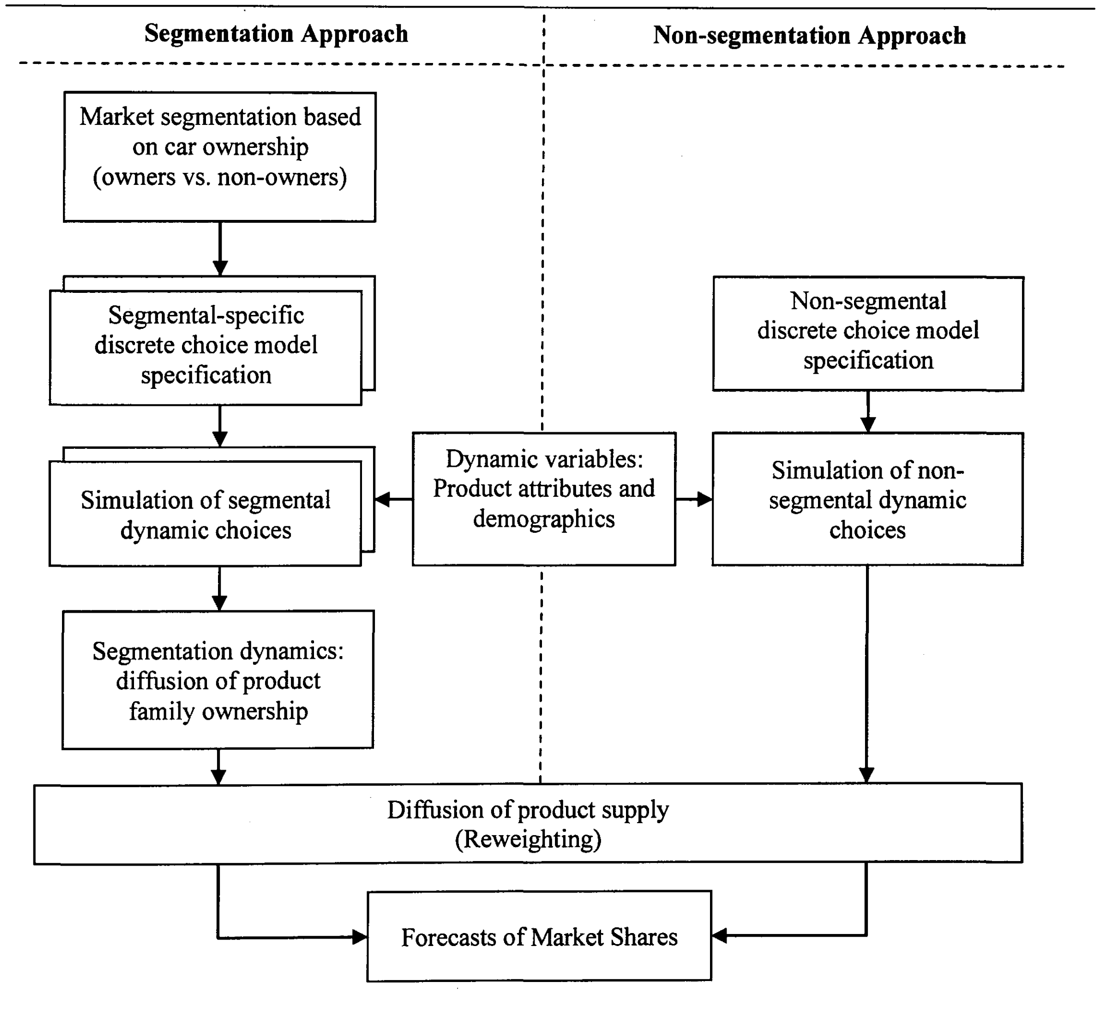
6.3 Modelling Methodology

After dividing the whole market into the two segments of car owners and non-car owners, our proposed approach consists of the following four-step procedure to forecast market shares for new products (see the left part of Figure 6-1).

1. We compare and specify the most appropriate discrete choice model for each segment based on conjoint experiment data so that the within-segment preference heterogeneity can be properly accounted for.
2. We conduct market simulation through accommodating different change scenarios of several important explanatory variables (called dynamic variables hereafter), which generates the dynamic market shares of different products within each segment at different periods.
3. We then model the segmentation dynamics by estimating a diffusion process, which provides us the dynamic sizes of the car owners' segment over time. These dynamic segment sizes are further used as the weights of averaging market shares of different products across two segments.
4. Finally, we employ another diffusion process to reweight the market share forecasts to address different supplies of new and mature products in the real market context.

With regard to the benchmark model, the non-segmental approach estimates the choice model based on the whole sample data and then applies a simulation approach with the same dynamic variables (see right hand side of Figure 6-1). There is no additional step for the non-segmental approach to account for segmentation dynamics, but it goes through the same product supply reweighting process.

Figure 6-1: the Segmentation vs. Non-segmentation approaches



6.3.1 Segment-specific discrete choice models

We follow the same modelling approach used in last chapter to identify the appropriate specification of a discrete choice model for different segments of consumers. Given multiple different alternatives, utility function of consumer i in segment k (k = car owners or non-car owners) choosing alternative n depends on product attributes, marketing or governmental policies X and choice-invariant socio-demographics Z:

$$U_{in}^{k} = V_{in}^{k} + \varepsilon_{in}^{k} = \alpha_k' X_{in} + \gamma_{nk}' Z_i + \varepsilon_{in}^{k} \tag{6-1}$$

where the utility U_{in}^{k} consists of a deterministic portion V_{in}^{k} and an error term ε_{in}^{k}. The coefficients of α_k' and γ_{nk}' are segment-dependent, which demonstrate that consumers have segment-specific preferences. If the error term is assumed to follow i.i.d type I extreme value distribution, we can use the multinomial logit (MNL) model (McFadden, 1974) and its choice probability is denoted as $P(MNL)_{in}^{k}$. The well known independence from irrelevant alternatives (IIA) property of the MNL model effectively means that consumers are assumed to perceive all alternatives completely independent from each other. If the error tem follows the generalised extreme value (GEV) distribution, it becomes the nested logit (NL) model (Ben-Akiva & Lerman, 1985; Daly & Zachary, 1978; McFadden, 1978; Williams, 1977), whose choice probability is denoted as $P(NL)_{in}^{k}$.

As we have discusses in last chapter, the NL model is the generalisation of the MNL model, because the IIA property is not held in the NL model so that it allows for different correlations and hence substitution effects among alternatives (Train, 2003). In our study, we are able to specify different types of the NL model to account for different choice structures across segments. Since the NL model reduces to the MNL model when the inclusive value (IV) parameter equals one, we use the NL model to present our approach hereafter.

6.3.2 Simulation of segmental dynamic choices

The discrete choice models specified in the preceding step are usually static, so the simulation approach is used to account for the potential dynamics of market demand through feeding some dynamic variables into the (static) segmental-specific discrete choice models. Thus, the new utility function with the dynamic (or time-dependent) effects is

$$U_{int}^{k} = V_{int}^{k} + \varepsilon_{int}^{k} = \alpha_k' X_{int} + \gamma_{nk}' Z_{it} + \varepsilon_{int}^{k} \tag{6-2}$$

where X_{int} and Z_{it} are dynamic variables, such as vehicle purchase price, fuel price and household income that are used in the market simulation of our empirical study. Also, we define different scenarios for dynamic variables to investigate the sensitivity of new product diffusion to these dynamic variables (Eggers & Eggers, 2011). We follow the assumption that derived coefficients in the utility function are constant over time (Lee, et al., 2008; Lee, et al., 2006).

After the market simulation, consumers at different time periods will have different choice probabilities. The time-dependent choice probability is denoted as $P(NL)_{int}^{k}$. Thus the market share of alternative n in segment k at time t is the average choice probabilities of all consumers for this alternative in the same segment:

$$S_{nt}^{k} = \frac{\sum_{i=1}^{I_k} P(NL)_{int}^{k}}{I_k} \tag{6-3}$$

where I_k is the segment size, i.e. the number of respondents in segment k.

6.3.3 Segmentation dynamics diffusion

Given two segments of car owners and non-car owners in the market, it is reasonable to expect that the car owners may have significant different preferences for new products of alternative fuel cars from the non-car owners, who have had little experience about cars. The additional benefit of such segmentation mechanism is that the dynamics of car ownership can be modelled using a diffusion process. For example, the Gompertz model is a popular diffusion specification for the car ownership (Dargay & Gately, 1999; Dargay, et al., 2007). Using this model, the percentage of the car owners' segment is defined as:

$$Q_t = e^{-ae^{-bt}} \tag{6-4}$$

where the saturation level is assumed to be 100% because all households can potentially be car owners, and a and b are two positive parameters that define the displacement and growth rate of the diffusion curve. Conversely, $1 - Q_t$ is percentage of non-car owners. The size of the segments, car owners and non-car owners is effectively defined by equation (6-4). It follows that the overall market share of alternative n at time t is the weighted average of the segmental market shares in two segments:

$$S_{nt} = Q_t S_{nt}^{k=\text{Car}} + (1 - Q_t) S_{nt}^{k=\text{No\_Car}} \tag{6-5}$$

where $S_{nt}^{k=\text{Car}}$ is the market share of alternative n in the car owners' segment at time t, which is weighted by the size of this segment Q_t. Similarly, $S_{nt}^{k=\text{No\_Car}}$ is the market share of alternative n in the non-car owners' segment.

6.3.4 Product supply diffusion

Although S_{nt} derived in equation (6-5) is the overall market share of alternative n at time t, it is still conditional on the assumption that all alternatives are equally available in the market, as usually assumed in the conjoint analysis (Wittink & Bergestuen, 2001). The product demand problem we mainly address in this chapter will be affected by the different products supplied to the market and even more so for new products with fewer product offerings than the mature ones. In our study, we account for the supplies of different products over time through an exogenous reweighting process. Intuitively, at the start of the commercialisation of a new technology, only a limited number of new products are launched by the pioneering firms. If the new products are successful, more brands or types of the same product will be offered by the followers (Lilien & Yoon, 1990). We can therefore assume that such a sigmoid growth pattern can also be modelled using a Gompertz diffusion process as follows

$$N_{nt} = \bar{N} \cdot e^{-\alpha e^{-\beta t_n}} \tag{6-6}$$

Where $\bar{N}$ is the saturation level and α and β are structural parameters to be estimated. In the context of EMs without stable stage data of the diffusion, we are not able to estimate the saturation level (Dargay, et al., 2007; Kobos, et al., 2003). Instead, we can adopt the scenario analysis approach by assuming different saturation levels (Button, et al., 1993; Chamon, et al., 2008). Furthermore, given that the new products (such as the electric or hybrid cars) are a new generation of the mature products (such as the petrol cars) within the same product family, we assume that the product supply of new products will approximately follow a similar but lagged path of the existing ones. Thus diffusion parameters in equation (6-6), i.e. α and β, are same for all alternatives. Thus, we can construct the product supply likelihood as

$$A_{nt} = \frac{N_{nt}}{\sum_{j=1}^{J} N_{jt}} \tag{6-7}$$

Thus the forecasted market share of product n at time t can be derived through a reweighting process as

$$MS_{nt} = \frac{S_{nt}A_{nt}}{\sum_{j=1}^{J} S_{jt}A_{jt}} \tag{6-8}$$

It is worth noting that this reweighting process in equation (6-8) is intended to provide more realistic forecasts, but will not change the forecasting comparison conclusion between our proposed approach and the benchmark approach, because both go through the same reweighting process (see Figure 6-1).

6.4 Empirical Application

When we apply the dynamic segmentation approach to forecast the demand or diffusion of the green cars in China, we essentially use both primary and secondary data that we have discussed in Chapter 2.

6.4.1 Estimation of segmental discrete choice models

The estimation of segment-specific discrete choice models is based on the choice-based conjoint analysis used in the last chapter, which is employing the primary data we collected through our survey. Specifically, the alternative-specific variables (*X*) consists of vehicle purchase price, annual running cost, charging facility and vehicle range for electric cars, and three types of incentives (cash subsidy, free parking and priority lane). The alternative-constant vector (*Z*) includes a list of demographic variables, such as number of young children, number of licensed drivers, family size, household income, household car ownership, and age, sex and working distance of household head.

In the last chapter, we show that for the whole sample data, the NL model that accounts for preference heterogeneity is more appropriate than the MNL model. More specifically, we find there are two potential tree choice structures (called NL1 and NL2 models, see Figure 6-2) that can achieve significantly better estimation performance than the MNL model. In order to appropriately account for heterogeneous consumer behaviour across different segments, we further explore whether different groups of consumers make their decisions differently, which is reflected through their different choice structures. Specifically, we find that the first tree structure in Figure 6-2 fits better for car owners and the second tree structure is more appropriate for non-car owners in China.

Figure 6-2: Two tree structures for the Nested Logit models

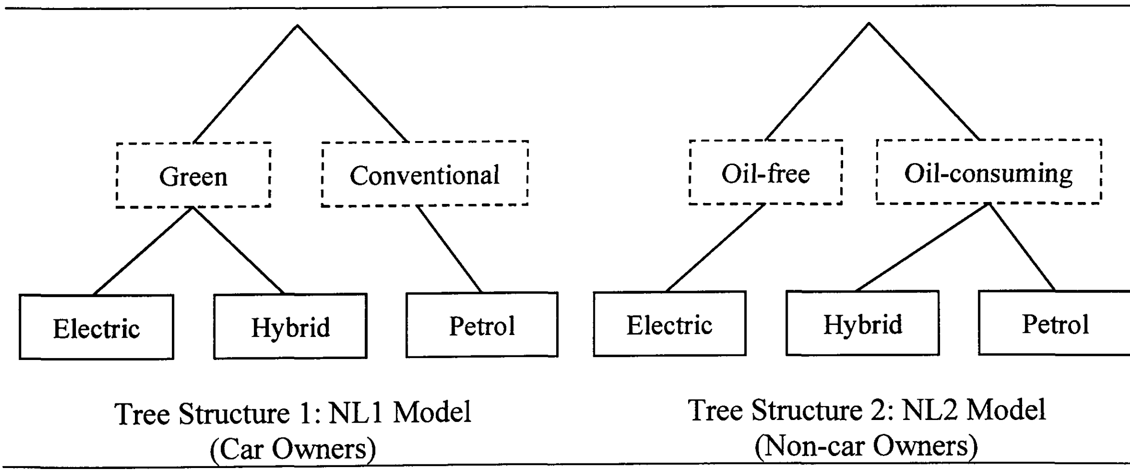


Table 6-1: Parameter estimation results of segmental NL models

Variable	Segment of Car owners				Segment of Non-car owners			
	NL(1) Model		NL(2) Model		NL(1) Model		NL(2) Model	
	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
Vehicle attributes or incentives (α)								
Purchase price	-0.012 ***	-7.161	-0.005 ***	-4.344	-0.018 ***	-10.347	-0.016 ***	-12.460
Running Cost	-0.006	-0.418	-0.017	-1.462	-0.065 ***	-5.796	-0.073 ***	-6.376
Charging Facility	0.463	1.467	0.388	1.324	0.992 ***	4.191	0.956 ***	4.084
Vehicle Range	0.001	0.439	0.002	0.865	0.006 ***	3.367	0.006 ***	3.336
Cash Subsidy	0.060 ***	2.928	0.043 **	2.421	0.020	1.320	0.015	1.027
Free Parking	1.266 **	2.482	0.937 **	2.217	0.415	1.150	0.326	0.899
Priority lane	0.931 *	1.833	0.620	1.468	0.164	0.458	0.078	0.217
Individual characteristics for electric car (γ_1)								
Electric car, constant	-3.845 ***	-3.908	-2.987 ***	-3.826	-1.138 **	-2.020	-0.619	-1.111
No. of young children	0.702 ***	3.256	0.541 ***	3.538	-0.768 ***	-5.083	-0.627 ***	-4.329
No. of drivers	-0.888 ***	-4.415	-0.427 ***	-2.992	-0.496 ***	-5.181	-0.373 ***	-4.141
Family size	0.217 *	1.763	0.058	0.653	0.345 ***	5.010	0.307 ***	4.720
Household Income	0.001	0.479	0.004 *	1.703	0.003	1.107	0.004	1.453
Age of family head	0.040 ***	3.23	0.025 ***	2.735	-0.015 **	-2.061	-0.014 **	-2.073
Sex of family head	-0.295	-1.03	-0.148	-0.700	-0.854 ***	-5.796	-0.846 ***	-6.189
Working distance of family head	-0.067 ***	-4.111	-0.041 ***	-3.502	-0.038 ***	-4.272	-0.032 ***	-3.761
Individual characteristics for hybrid car (γ_2)								
Hybrid car, constant	-1.090	-1.426	-0.380	-0.724	1.081 **	2.534	1.121 ***	2.670
No. of children	0.443 **	2.416	0.282 ***	2.639	-0.657 ***	-5.763	-0.637 ***	-6.080
No. of drivers	-0.969 ***	-5.436	-0.514 ***	-6.135	-0.362 ***	-5.008	-0.354 ***	-5.223
Family size	0.257 **	2.467	0.109 *	1.887	0.068	1.274	0.067	1.315
Household Income	-0.001	-0.337	0.001	0.696	0.004 *	1.897	0.004 *	1.958
Age of family head	0.029 ***	2.771	0.014 **	2.424	-0.011 **	-2.008	-0.010 **	-1.961
Sex of family head	-0.598 **	-2.536	-0.391 ***	-2.833	-0.240 **	-2.009	-0.252 **	-2.203
Working distance of family head	-0.054 ***	-3.957	-0.029 ***	-3.928	-0.030 ***	-4.660	-0.028 ***	-4.632
Inclusive value (IV) parameters (λ)								
Green nest in NL(1)[a]	0.578 ***	9.934			0.948 ***	25.466		
Oil-consuming nest in NL(2)[b]			0.985 ***	5.311			0.725 ***	11.517
Rho-square w.r.t. zero	0.1586		0.1492		0.1549		0.1590	
χ^2 against MNL[c]	36.346 (df = 1)		0.022 (df = 1)		1.740 (df = 1)		23.428 (df = 1)	

Note: *** significant at 1% level; ** significant at 5% level; * significant at 10% level.

[a] Wald statistics against the unity are -7.251 and -1.398 for car owners and non-car owners segments respectively. [b] Wald statistics against unity are -0.081 and -4.372 for two segments respectively. [c] Critical value of χ^2 is 3.84 at 5% significant level for one degree of freedom.

The model estimation results are presented in Table 6-1. Main findings of the segmental discrete choice models have been discussed in the segmentation analysis section of last chapter. Essentially, we find that car owners tend to perceive that hybrid car is more correlated with electric car than with petrol car. In addition, we find that existing car owners in China are not concerned with product attributes except purchase price but sensitive to governmental incentives such as cash subsidy and free parking. Regarding the segment of non-car owners, the better performance of NL2 model suggests that the non-car owners are more likely to think hybrid car to be more related to conventional petrol cars than to the electric cars. The estimated coefficients in this subsample are quite different from the car owners' sub-sample. Importantly, all vehicle attributes are significant at 1% level with the correct signs, but no incentives are significant for non-car owners in China.

6.4.2 Selection of dynamic variables

In the EMs, the market dynamics can occur because of the potential changes of both demographic characteristics and product attributes. We employ household income, vehicle purchase price and fuel price as examples of key dynamic variables and apply the simulation technique to account for the dynamic effects of these factors on market demand in the future.

- Household income

Amongst all the demographic variables, we select household income as a representative variable in the market dynamics simulation, because it is the factor that we can surely expect a clear changing pattern (i.e. growth) in the future[35]. In our simulation, we assume the household income level will follow the economic growth in China, which is predicted to maintain a fairly high speed in near future and then gradually slow down. For example, Goldman Sachs predicts that the annual GDP growth rates of China will be 7.9% in 2011-2020 and drop to 5.7% in 2021-2030 (O'Neill & Stupnytska, 2009). For the illustrative purpose, we assume that house hold income in our simulation will increase annually by 8.50% in next 5 years (2011-2014), followed with 7.50% from 2015 to 2019, 6.50% from 2020 to 2024 and 5.50% from 2025 to 2029 respectively. We apply these income assumptions to all households because we conduct hypothesis tests that confirm that there is no significant difference in terms of income growth rates for households at different income segments in China.

- Vehicle purchase price

The discrete choice models developed previously indicate that vehicle purchase price is an important factor that significantly affects consumer choices for both segments of the market. Empirically, the prices for durable goods are normally assumed to follow an exponentially declining trend (Bayus, 1992), which has been applied by Lee et al. (2006) to estimate the price change of different TVs:

$$\text{Price}_t = m \cdot e^{\delta t} \tag{6-9}$$

where m is the initial price and δ is the parameter of price decline when $\delta < 0$. However, it is impossible to directly estimate the price function for new products such as hybrid or electric cars, because there is no historical price information. Instead, given the current higher prices of the hybrid or electric cars, we assume that

[35] We also note that the income is insignificant for each segment, which suggests the income differences across two segments are important. Thus to include or exclude it from the market simulation will not significantly change the simulation result.

their prices will decrease gradually and approach that of the conventional petrol cars in the long term, because these new products may benefit more from technological innovations and economies of scale in the future. Given that the long term price of the petrol cars can be easily extrapolated using its historical information, we can derive the price change parameter δ of the hybrid or electric cars using the transformed equation

$$\delta = \frac{\ln(\text{Price}_{\text{T}}/m)}{\text{T}} \tag{6-10}$$

where the long term price (Price_T) of either hybrid or electric cars is assumed to be same as the petrol cars' price after a T periods, and the initial price level m of the hybrid or electric cars might be available in the market before their product launch. With the price change parameter δ, we are able to extrapolate the prices for both the hybrid and the electric cars through applying the parameter back into equation (6-9).

In our study, we collect the secondary data of the petrol cars' price information from the magazine of *Orient Auto* in China, which regularly publishes detail specifications of all available passenger cars in the market. We calculate quarterly market average price of the petrol cars from Q4/2002 to Q4/2007, and then estimate the price function (6-9) of the petrol cars. The quarterly price change parameter δ of the petrol cars is -0.0089, which is significantly different from zero. It implies that on average the petrol cars in China have an annual price decreasing rate of 3.495% ($=1 - e^{4\delta}$).

In order to estimate the price changing parameters of the hybrid or electric cars, we derive the initial prices m of the hybrid and electric cars based on two representative models in the market. They are BYD F3DM and Chery RIICH M1-EV for the hybrid and electric cars respectively. Specifically, the F3DM is the only hybrid car recognised in China's first batch of green vehicle recommendation catalogue[36], which qualifies the green cars for government incentives. The price of F3DM is twice of its petrol-fuelled version (F3). The first electric car available in China is RIICH M1-EV, which was launched by another local firm Chery in November 2010[37]. Its price ranges from 149.8k RMB to 229.8k RMB, which is on average 3.90 times of the price of the similar model fuelled by the petrol (RIICH M1). In terms of the length of T periods, we do not have any priori but set up three scenarios, i.e. fast ($T = 20$ years), moderate ($T = 30$ years) and slow (T = 40 years) cases. Based on the initial price and different length of T periods, Table 6-2 presents the estimated price changing parameters as well as the equivalent annual price decrease rates for both hybrid and electric cars. With the price changing parameters, we then apply equation (6-9) to extrapolate the periodical prices of two types of green cars respectively in the forecasting horizon.

Table 6-2: Scenario-based parameters of the price functions for the hybrid and electric cars

Scenario	Fast (T=20)		Moderate (T=30)		Slow (T=40)	
	Parameter (δ)	Annual price decrease rate	Parameter (δ)	Annual price decrease rate	Parameter (δ)	Annual price decrease rate
Hybrid	-0.0583	5.67%	-0.0508	4.95%	-0.0470	4.59%
Electric	-0.0873	8.36%	-0.0701	6.77%	-0.0614	5.96%

Note: The calculation of annual price discount rate follows the formula of $(1 - e^{\delta})$.

- Fuel price

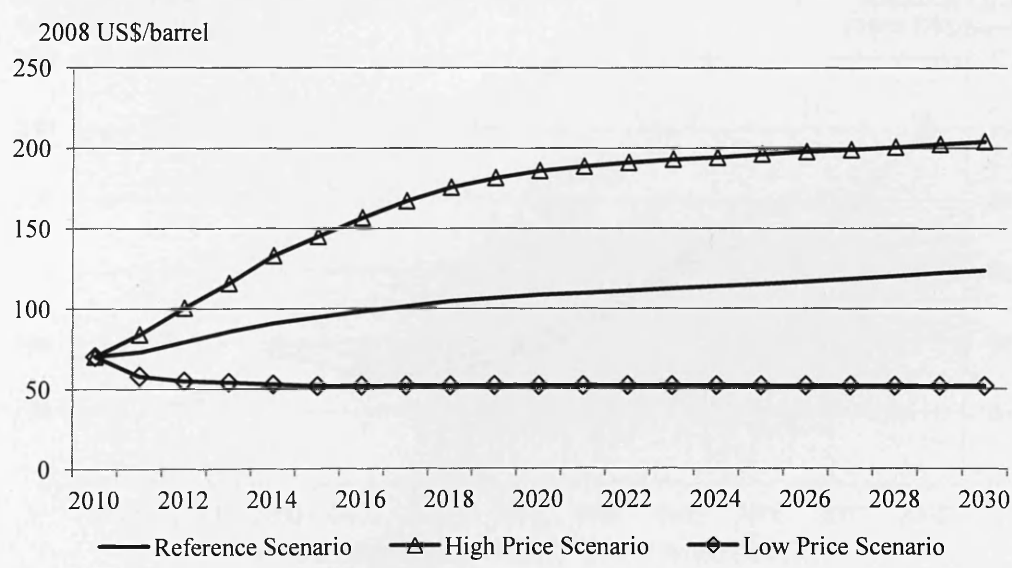
In addition to the purchase price, we include the fuel price as another dynamic variable in the market simulation exercise. It is because we find that the running cost,

[36] It was published by the Ministry of Industry and Information Technology of China, 11 August 2009 http://www.miit.gov.cn/n11293472/n11505629/n11506277/n11984220/n11984250/12504301.html

[37] "Chery Riich M1 EV hits the market", China Automotive Review, 24 November 2010 http://www.chinaautoreview.com/pub/CARArticle.aspx?ID=5019

whose variation is mainly from fuel cost, is the second most important car attribute (after the car purchase price) that significantly affects consumer choices in the non-car owners' segment as well as the total population. Since it is not our focus here to forecast fuel price, we adopt the secondary data of the world oil price projections from a professional agency within the U.S. Department of Energy, which provides annual world average oil price forecasts in three scenarios with the horizon from 2010 to 2030 (U.S. Energy Information Administration, 2010). As shown in Figure 6-3, the high fuel price scenario projects that the world oil price will exceed $200 per barrel by 2030 (at 2008 price level, same in follows), while the low price scenario estimates that the oil price will stably keep at $50 per barrel in next 2 decades. Between them, the reference scenario forecasts that the oil price will slowly increase to be about $123 per barrel in 2030.

Figure 6-3: Three scenarios of annual world oil price (2010-2030)

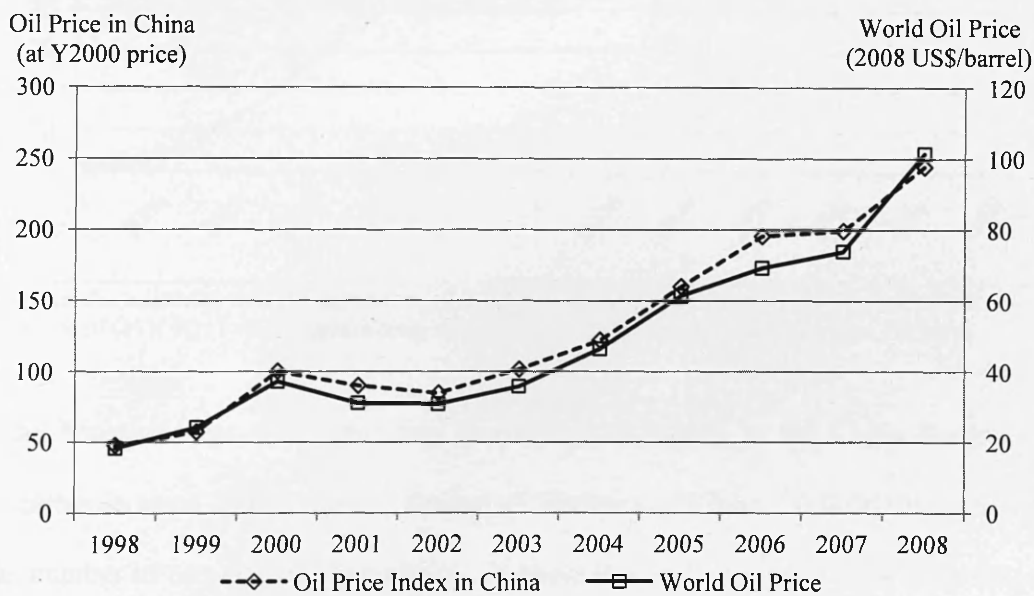


Data Source: U.S. Energy Information Administration

Besides, we compare the oil price history in China with that of the world to assess if we can use the projected world oil price as a proxy of the future oil price in China.

More specifically, we collect and compare the annual world oil price published in the U.S. Annual Energy Outlook 2010 (U.S. Energy Information Administration, 2010) and the Chinese annual oil price index from the Development Research Centre of the State Council of China[38] from 1998 to 2008. As shown in Figure 6-4, we observe that the historical oil price in China did closely follow the world price and their correlation index is 0.991. A linear regression function is further estimated with the world oil price as explanatory variable and the oil price index in China as dependent variable. The regression achieves a high R-square value of 0.982 and the estimated intercept is insignificant from zero (P-value = 0.223), which confirm that it is appropriate for us to use the projections of world oil price in the market dynamics simulation in China.

Figure 6-4: Comparison of oil price history in China and the world level (1998-2008)



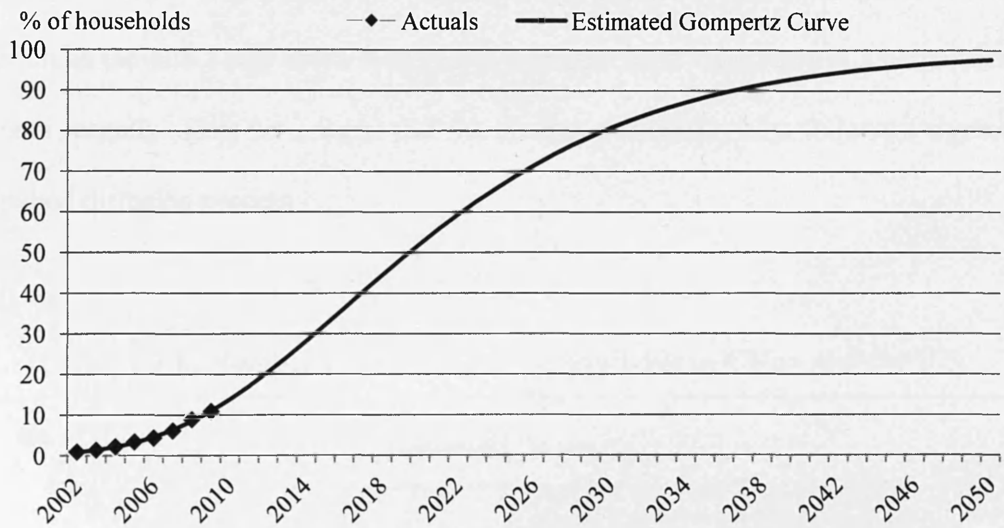
Data Source: The Development Research Centre of the State Council of China
The U.S. Energy Information Administration

[38] The website address is http://www.drcnet.com.cn.

6.4.3 Estimation of the segmentation diffusion model

Through modelling the segment-specific discrete choice models, we have specified the appropriate choice structure for each segment and also estimated the heterogeneous preferences. Furthermore, we expect the segment size of the car owning households in China will continue to grow following a diffusion process and we assume in long term that all households can be potential car owners in the market.

Figure 6-5: Diffusion of the percentage of car-owning households in China



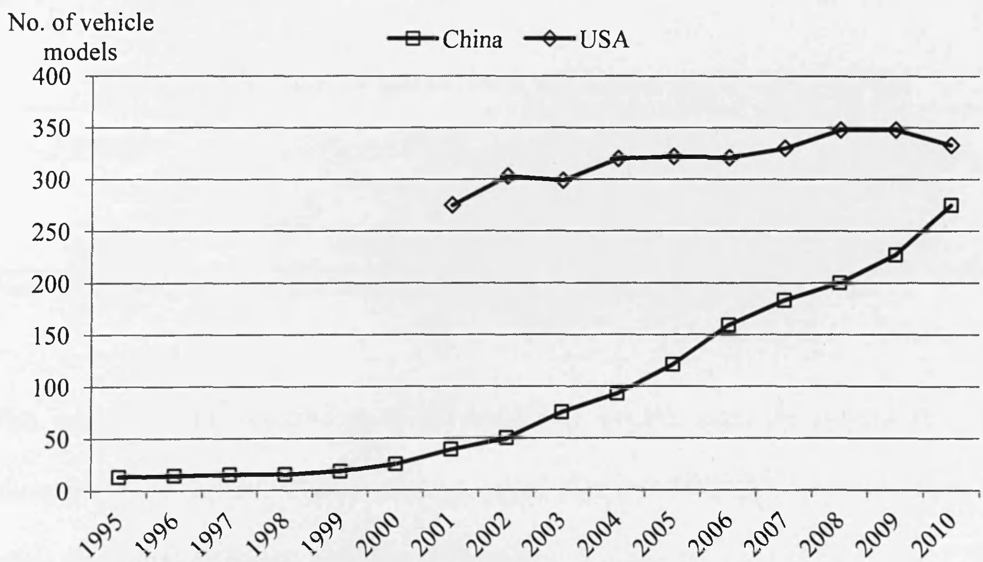
Note: the estimated value of parameter (*a*) is 5.378 with standard error of 0.121; the estimated value of (*b*) is 0.111 with standard error of 0.003. Both parameters are significant at 1% level.

The household car ownership data in China is available in the China Statistical Yearbooks since 2002 (National Bureau of Statistics of China, 2003-2010), reported as number of cars per 100 households. Because this market is still at the early stage of car adoption and most car-owning households only own one car, the percentage of the car-owning households can be approximated with the number of cars per 100 households. We use this secondary data of the car ownership data from 2002 to 2009 to estimate the Gompertz model specified in equation (6-4). We follow the same estimation method described in Chapter 3 to estimate the Gompertz model here. Based on estimated parameters, we extrapolate the annual proportions of car-owning households in China in following decades (see Figure 6-5).

6.4.4 Estimation of product supply diffusion model

When comparing the car model supply between in China and in U.S. (see Figure 6-6), we find that the number of available new car models is rapidly increasing when the market is expanding (such as in China), while the growth in more mature market (such as the U.S.) was much slower and even has been stagnating at a certain level more recently. Thus we assume that the product availability also follows a sigmoid-shaped diffusion process.

Figure 6-6: Number of new car models available in China and the U.S.



Data source: China Associate of Automobile Manufacturers (CAAM)
J.D. Power New Car Database (http://www.jdpower.com/autos/new-cars/)

In the diffusion model specification of equation (6-6), we argue that the saturation level in China might be higher than the level in the U.S. because by the end of 2010 there were more than 270 new car models available in China and its growth trend is still continuing. More importantly, market concentration level in China is much lower than in the U.S. According to annual sales in 2009, market shares achieved by top four manufacturers (CR4 ratio) was only about 37% in China[39], compared to 64% in the U.S.[40]. So we assume a reasonably higher saturation level of 500 new car models in China.

The product availability diffusion model of equation (6-6) is estimated using the secondary data of the number of new car models available in China from 1995 to 2010. The sample includes all petrol cars available after the Chinese government published *Automobile Industry Policy* in 1994, which is thought as the watershed of the Chinese automotive industry (Chin, 2010). Table 6-3 presents the estimation results, based on which we can project the availability of the petrol cars in the future.

Table 6-3: Estimated parameters of product availability model

Parameter	Estimated value	Standard error	t-statistics
α	7.526	0.503	14.974
β	0.153	0.005	28.256

Note: estimation sample is no. of available car models in China (1995-2010).

With respect to the availability of the hybrid or electric cars, we assume they will follow a similar diffusion process as the petrol cars but with later starting points. We define the initial diffusion periods of the hybrid and electric cars as the year of 2009

[39] According to sales statistics from China Associate of Automobile Manufacturers (CAAM)
[40] According to NADA 2010 Data Report from National Automobile Dealers Association, http://www.nada.org/Publications/NADADATA/2010/default.htm

and 2011 respectively, because the F3DM started its large-scale sales in 2009 and the RIICH M1-EV plans to start its sales in 2011. Thus, we derive the different availability diffusion processes for the hybrid and electric cars and then apply equation (6-7) to calculate the availability likelihoods for three different types of cars at different time periods, which predict that there will be 60% of new car models to be either the hybrid or the electric cars by 2030.

6.4.5 Specification of the benchmark model

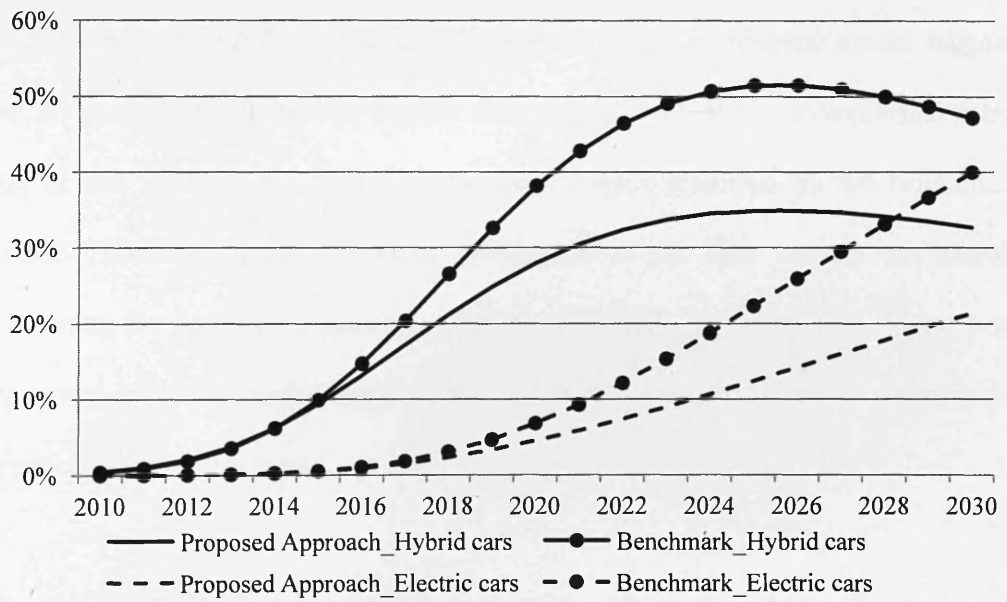
The benchmark model in this study is based on the non-segmental approach. We estimate the benchmark model based on the NL specification using the whole sample. That is the benchmark model does not account for the possibility that different groups of consumers may make their choices in different ways. Specifically, we use the second tree structure with "Oil-consuming" nest (see Figure 6-2) for the benchmark NL model, because it is a better model. Based on the estimation NL model for the whole sample, we conduct market simulation using the same dynamic variables of income, vehicle purchase price and fuel price. Finally, the product availability likelihoods calculated in section 6.4.4 are applied to generate the benchmark market share forecasts.

6.5 Forecasting Result and Scenario Analysis

6.5.1 Market share forecasts for green cars

In this subsection we present the market share forecasting results generated by the proposed dynamic segmentation approach and compare them with those from the benchmark model. Here we only present the forecasts based on a base scenario and more scenario analysis will be discussed in next subsection of scenario analysis. The base scenario is specified with the moderate vehicle price decreasing speed for the green cars (i.e. the case of T=30) and the reference case for the fuel price changes in future. We define our forecasting horizon to be next 2 decades from 2010 to 2030, because currently both car manufacturers and governments are more interested in exploring the long term demand potential of the green cars sector in order to better design its development strategies and industrial policies.

Figure 6-7: Market share forecasting for hybrid and electric cars in base scenario (2010-2030)



Note: the base scenario is specified with the moderate price decreasing for green cars as well as the reference fuel price assumption.

The market share forecasts in the base scenario from both hybrid and electric cars are presented in Figure 6-7, which includes the results from our proposed dynamic segmentation approach as well as the benchmark model. Within the first 5 years, there is no significant difference between the forecasts from both models, which predict that the hybrid car will start the relatively fast growth and will achieve 10% market shares by 2015, while the market expansion of the electric car will be slower. More significant distinctions between different approaches appear in the long term, where our approach provides more conservative predictions than the benchmark model for both types of the green cars. More specifically, our approach forecasts that the hybrid car will not reach a market share of 30% until after 2020 and will achieve the highest level close to 35% in 2025. More optimistically, the benchmark model predicts that the hybrid cars can achieve more than 50% market shares. Regarding the electric car, we forecast that its market share will grow slowly with market share slightly exceeding 20% by 2030, but the benchmark model predicts a much faster growing trend with about 40% shares from the electric cars by 2030. By aggregating the forecasts for both the hybrid and the electric cars, our proposed model suggests that the green cars will have to equally share the market with the conventional petrol cars in late 2020s, instead of the dominating status predicted by the benchmark model. Therefore, these significant differences imply that without sufficiently accounting for preference heterogeneity in the rapidly growing market, the non-segmental approach has the tendency to over-predict the market growth for both the hybrid and the electric cars.

We can follow the same approach to forecast the market share of petrol cars. Actually the market share of petrol cars equals one minus the green cars' market share, because fundamentally we use the discrete choice model to estimate the choice probabilities of three alternatives and the following reweighting procedures also normalise the total market share to be one.

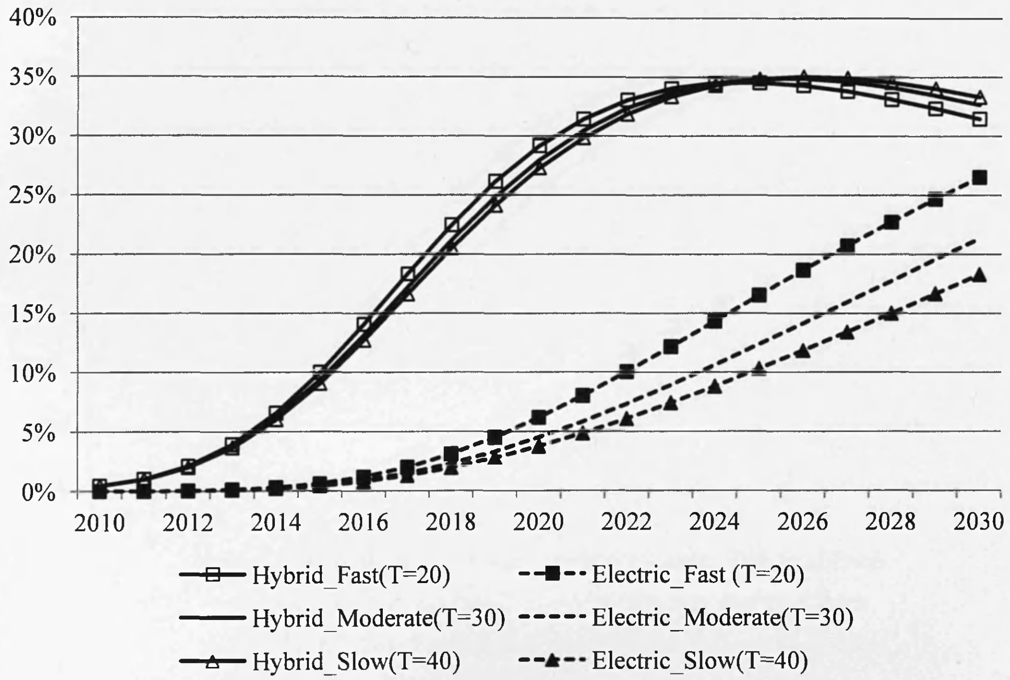
6.5.2 Scenario analysis

In addition to the base scenario, we conduct scenario analysis to investigate the sensitivity of market demand to different assumptions of two dynamic variables: the vehicle purchase price and the fuel price. Each factor has 3 different levels so that we have 9 available scenarios. Due to the space constraints, we first present the scenario analysis based on each factor individually without further comparing with the benchmark model, and then show the extreme case scenario that takes both factors into consideration (i.e. the most optimistic and the most pessimistic cases).

Depending on the different rates of change speeds for the green cars' purchase prices (see Table 6-2), three scenarios of market share forecasting are presented for both the hybrid and the electric cars in Figure 6-8, where the moderate one is the base scenario discussed in section 6.5.1. The underlying assumption about vehicle price is that the cost of green cars will decrease faster than that of the petrol cars, because the petrol cars are a mature product while the green cars are new to the market. The current high price of green car is largely due to its battery cost. We expect the future innovation in the battery technology will significantly reduce the price of green cars, but we are not sure when and how fast such price reduction will happen, so we make three different scenarios about the price of green cars.

In these vehicle price scenarios, the market share projections for the hybrid cars are quite close to each other, which implies the limited sensitivity of the hybrid cars' diffusion to different price changing speeds. On the contrary, different vehicle price changing scenarios are found to significantly affect the market share forecasts of the electric cars, ranging from 18% to 26% by 2030 depending on different price decreasing speeds.

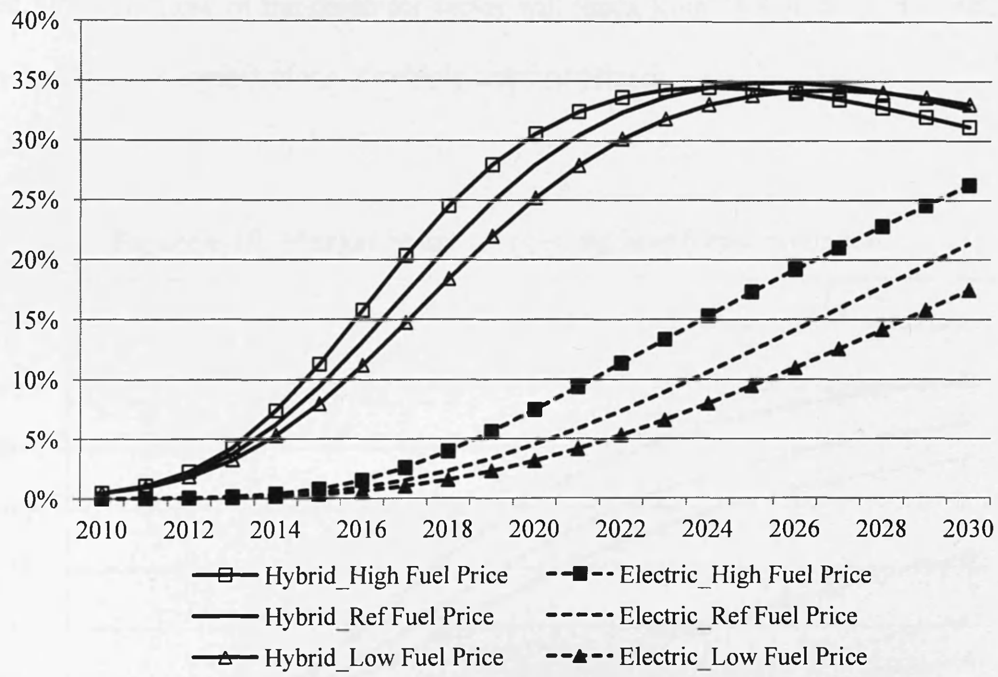
Figure 6-8: Scenario analysis based on different vehicle price change speeds



Note: Green cars' price change scenarios are defined in Table 6-2.
In this figure, fuel price follow the reference scenario.

The scenario analysis based on different fuel price levels is presented in Figure 6-9. If other factors remain the same as in the base scenario, the electric cars will increasingly benefit from the higher fuel prices, whose market share will grow to above 26% by 2030 in the scenario of high fuel price compared to 17% only in the low fuel price scenario. With regard to the hybrid cars, we find that they will only benefit from the high fuel price scenario in the first decade of our forecasting horizon until 2020. After that, the differences of the hybrid cars' market shares due to different fuel prices will gradually diminish. After 2026, the hybrid cars in the high fuel price scenario will even have less market shares than in the low fuel price scenario, but such difference can be completely offset by the market share increments of the electric cars in the high fuel price scenario.

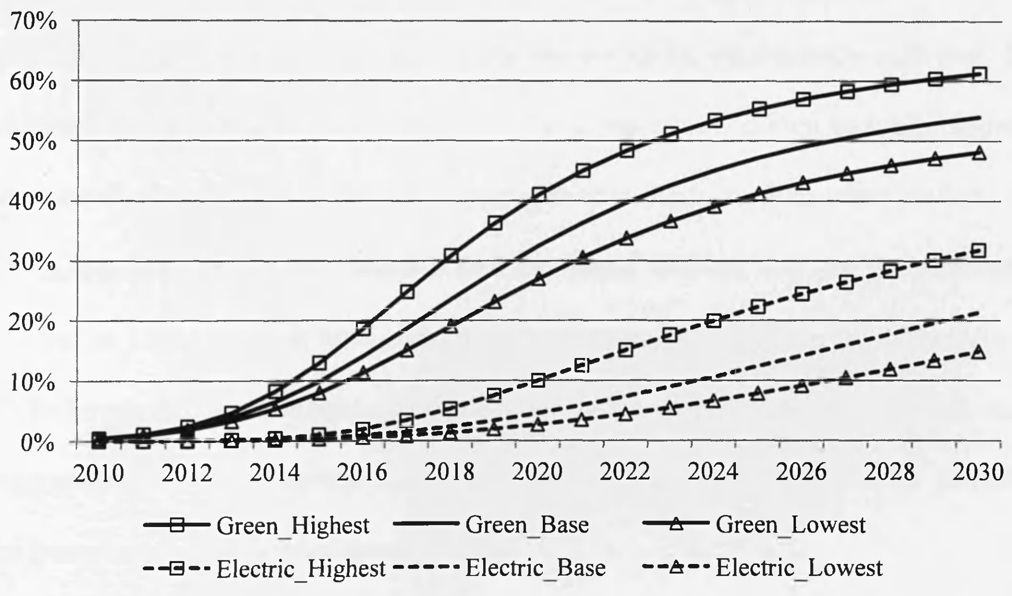
Figure 6-9: Scenario analysis based on different fuel price assumptions



Note: Fuel price scenarios are shown in Figure 6-3.
In this figure, the vehicle prices follow the moderate price changing scenario (T=30).

In summary, the electric cars will be more affected by the different scenarios of purchase price and fuel cost than the hybrid cars, as the electric cars will achieve higher market shares if their purchase prices decrease faster or the fuel price remains higher. Therefore, we present two extreme cases of market share forecasting for both the electric cars and the whole green car sector (see Figure 6-10). In the most optimistic case, which is specified with the fast decreasing speed for green cars' price (T = 20) and the high fuel price scenario, the market shares of both the electric car and the whole green car sector will grow fastest and expect to reach about 31% and 61% shares respectively by 2030. In contrast, the most pessimistic case is defined with the slow price decreasing speed (T=40) and the low fuel price scenario, where the market shares of the electric cars and the green car sector will have the slowest growths and ultimately reach about 15% and 48% respectively in 2030. Therefore, the forecasted market shares of the electric cars will vary between 15% and 31% and those of the green car sector will range from 48% to 61%, depending on the different combinations of vehicle and fuel prices.

Figure 6-10: Market Share forecasting in extreme scenarios



Note: The highest scenario is specified with fast decreasing trend for green cars' price (T=20) and high fuel price. The lowest scenario is specified with slow decreasing trend for green cars' price (T=40) and low fuel price. The base scenario follows the same definition as Figure 6-7.

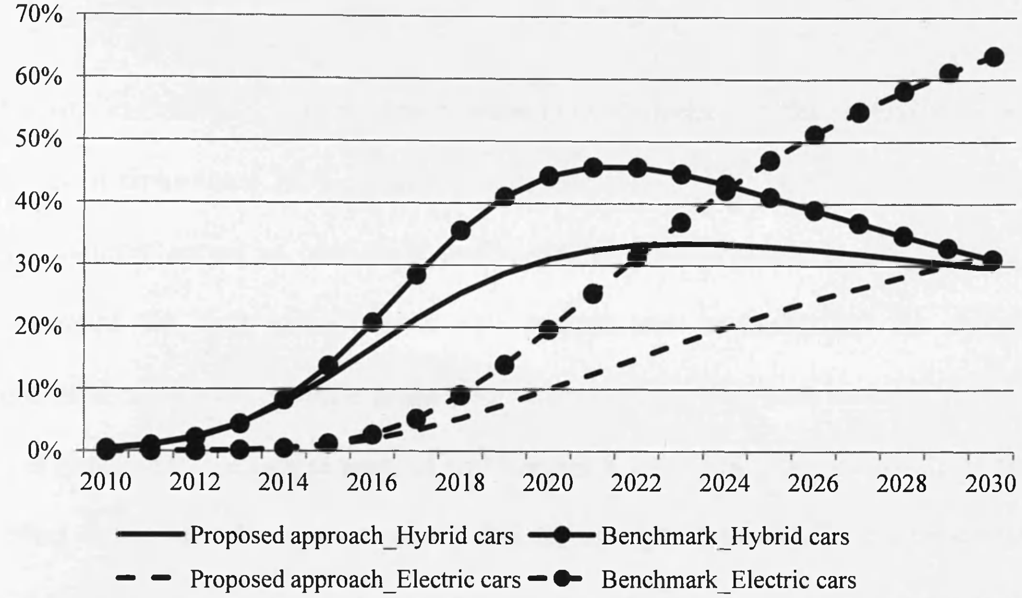
6.5.3 Model validation

We are not able to conduct the typical model validation exercise based on the forecasting performance in a validation sample, because green cars (particularly electric cars) are not widely available in the market and we do not have enough data to verify the forecasts in a 20-year horizon as used in this study. In a similar case of predicting diffusion of digital televisions without sales history, one validation approach used by Gupta, et al.(1999) is to compare their forecasts with the adoption pattern of older generation of products (i.e. colour television in their case). However, this approach will be problematic in our case due to potential issues in both demand and supply aspects. First of all, the early stage market demand of conventional petrol cars in China mainly came from the public sectors in the last century, while household buyers dominate the car market in China now. We expect that the adoption behaviour of public and private buyers to be significantly different. In addition, the development of petrol cars in China was mainly driven by multinational car manufacturers to introduce their existing mature products from other markets to China, but currently all car manufacturers, no matter whether they are multinational or local in China, are still exploring the key technologies for green cars. Therefore, the differences in both demand and supply aspects between petrol cars and green cars suggest that it is inappropriate to use the adoption history of petrol cars to validate the green car market development in China.

We have shown in the base scenario that our proposed approach will provide smaller market share forecasts for green cars than the benchmark approach (see Figure 6-7). Since we are not able to directly validate the forecasts using sales data, we investigate the model robustness and specifically whether the proposed approach can consistently produce more conservative forecasts for green cars' demand in comparison to the benchmark approach in two extreme scenarios. The highest scenario is specified with high fuel price and fast vehicle price decreasing trend, while the lowest scenario has low fuel price and slow vehicle price decreasing trend.

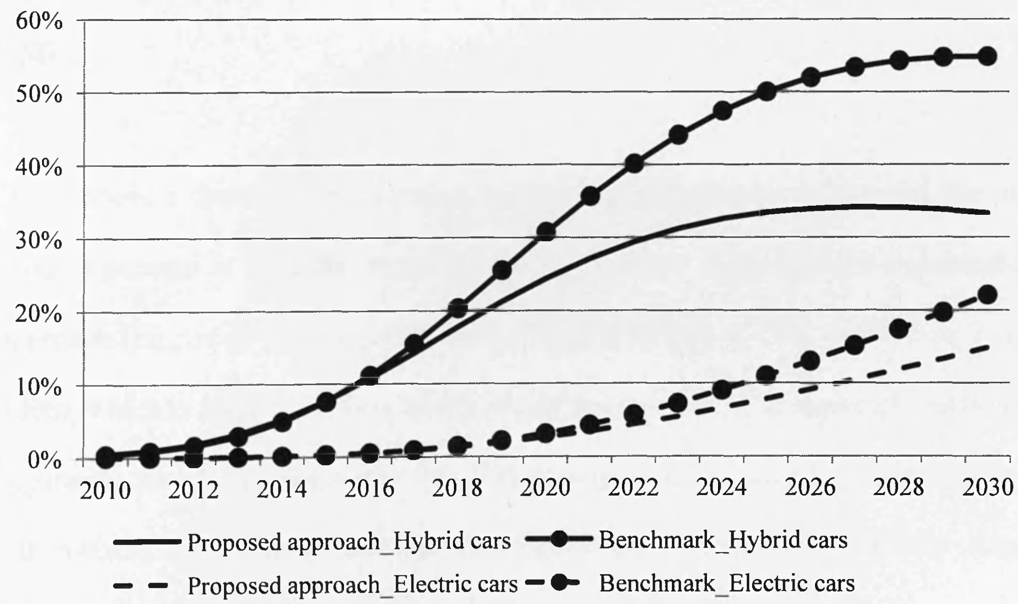
Figure 6-11 and Figure 6-12 present market share forecasts for hybrid and electric cars in two extreme scenarios respectively. We can see in both scenarios that the benchmark approach forecasts higher market potentials of hybrid and electric cars than the proposed approach. Specifically, in the highest scenario (Figure 6-11), the benchmark model predicts that electric cars can achieve more than 60% of market shares by 2030, which doubles the forecasts from the proposed approach. Also, the predicted market share of hybrid cars can have a peak point of 45% in 2021 based on the benchmark model, but the proposed model only predicts the highest market share of only 33% for hybrid cars in this scenario. In the lowest scenario (Figure 6-12), the benchmark model predicts that the hybrid cars' market share will increase to 55% in 2030, which is much greater than our forecasts (34%). Similarly, the benchmark model in the lowest scenario forecasts the faster growth of electric cars' market share than our proposed approach (22% versus 14% market share by 2030). In summary, the robust analysis in two extreme scenarios demonstrates that the proposed approach that can better account for preference heterogeneity will consistently provide less extreme predictions about the demand of green cars than the benchmark approach.

Figure 6-11: Market share forecasting for hybrid and electric cars in the highest scenario (2010-2030)



Note: The highest scenario is specified with fast decreasing trend for green cars' price (T=20) and high fuel price.

Figure 6-12: Market share forecasting for hybrid and electric cars in the lowest scenario (2010-2030)



Note: The lowest scenario is specified with slow decreasing trend for green cars' price (T=40) and low fuel price.

6.6 Conclusion

Growth in Emerging markets has become more important in the context of global financial crisis since 2008. Unsurprisingly, consumers now in the EMs have easier and quicker access to new technologies and innovative products, and thus it is important for both manufacturers and governments to understand the demand potential of the new products in the EMs, especially for long term business strategy. It is important to recognise some of key features of these markets and how these may affect the demand. In terms of models that can be applied, this effectively means that approaches used to forecast demand in such markets need to account for the differences between emerging markets and other more mature markets. The non-segmental approach for choice modelling which is typically used in the developed markets may be inappropriate to forecast demand for new products in the EMs, because it does not account for the preference heterogeneity typically found in the EMs.

We propose a dynamic segmentation approach in this chapter to forecast the new product demand in the EMs. Previous research tends to apply the non-segmentation approach (Eggers & Eggers, 2011; Lee, et al., 2008; Lee & Cho, 2009; Lee, et al., 2006), which is used as the benchmark model in our study. Our approach starts with separating the whole market into two distinct segments, such as car owners and non-car owners, which helps address the significant preference differences across segments in the EMs (Burgess & Steenkamp, 2006). For each segment, we then specify the most appropriate choice structures given different alternatives, so that we can further accommodate the within-segment preference heterogeneity by identifying different forms of NL model for each segment. The non-segmentation benchmark model, however, can only specify a discrete choice model for the whole sample, and thus it is not good at sufficiently accounting for preference heterogeneity both across and within segments. In a fast growing market context, we expect the consumer choices in the future will also be influenced by the dynamic changes of product attributes. Therefore we define different scenarios of product attribute changes (e.g. vehicle price and fuel price) to address their influences and the sensitivity on the future market shares. In the next step, we consider a market trend that the Chinese consumers are transiting from non-car owners to car owners, which will deeply affect the market share forecasts given the preference heterogeneity between these two segments. Thus we specify a diffusion model to account for the continuous growth of car owners' segment, which is then used as the weight to combine the predicted market shares from each segment for different alternatives. In contrast, the benchmark model cannot address such transition between segments given its non-segmentation characteristics. The last step of our approach is to address the influence of product supply on the market share forecasting, as new products will not equally available as the mature ones in the market and their supplies will change over time.

By applying the proposed approach to forecasting market shares of the alternative fuel cars in China, our approach forecasts that the aggregate market share of hybrid and electric cars will grow up to 54% by 2030 in the base scenario, while the benchmark approach provides a more optimistic forecast with more than 87% market shares held by the green cars in 2030. This implies that when the non-segmental (benchmark) approach fails to account for enough preference heterogeneity, it will overestimate the demand potential of the green cars in China. In addition, we use several scenario analyses to investigate the sensitivity of the green cars' market demand to different assumptions on vehicle price decreasing speed and fuel price. Compared with the hybrid cars, we find that the electric cars will be significantly affected by the different scenarios of both purchase price and fuel cost and the effects will gradually increase over time. We also validate the model through investigating the robustness of our approach in two extreme scenarios. The validation analyses in both the highest and lowest scenarios show that our approach consistently provides less extreme forecasts than the benchmark models for the green cars' growth in China.

Through this empirical application, we essentially demonstrate the importance of appropriately accounting for the key features of the studies market when developing models to forecast demand of new products in the EMs. As we show through the comparison against the benchmark model, if the market features such as preference heterogeneity and dynamic dichotomy of car ownership status are not well accounted for, the derived market forecasts are likely to suffer significant bias, which will then mislead the strategic investment decisions of car manufacturers as well as the industry or incentive policies designed by governments.

CHAPTER 7.

SUMMARY OF RESEARCH AND RESEARCH IMPLICATIONS

This chapter provides a summary of the research work that has been conducted in this thesis and highlights the key research contributions. In section 7.1 we summarise the main background and research proposition of this thesis. In section 7.2 we highlight the specific contributions in each contributing chapter and discuss the key points when extending our research into other EMs. We also suggest some avenues for further research in this chapter.

7.1 Summary of the Main Research Proposition

The emerging markets (EMs) have been increasingly important in the world economy, particularly in the context of the global financial crisis since 2008. When most developed car markets suffered weak demand or serious fall in demand in 2008, only four major developing markets, Brazil, Russia, India and China (BRICs) within the top 10 car markets successfully maintained positive market growth (The International Organization of Motor Vehicle Manufacturers, 2009). The emerging markets usually demonstrate different institutional characteristics from the developed markets, such as the high level of market heterogeneity (Alden, et al., 2006; Burgess

& Steenkamp, 2006; Sheth, 2011). This thesis considers the case of the car market and explores the marketing modelling challenges that relate to demand forecasting in such a market context. In China, car sales in 2002 were only 1.25 million units, but the figure sharply increased to 11.27 million by 2010. The urban household car ownership level in China was only 0.88 cars per 100 households in 2002, and the rate rose to 10.89 cars in 2009. Significant market growth in China enabled it to overtake the U.S. to be the largest automobile market in the world in terms of annual sales in 2009.

Research on modelling car market demand is well established in the developed economies but little attention has been paid to the emerging car markets and the corresponding challenges that researchers face when they have to predict the demand or preferences for cars in the EMs. When modelling the car market demand in the EMs, researchers usually encounter several challenges such as limited sales data for demand forecasting and "not well" established consumer preferences which are mainly related to the fact that this is a new product concept in that market. The scarcity of thorough research in that context has formed and shaped the main research contention and proposition of the thesis.

In the context of EMs such as China, cars are a type of new products for most consumers. In general, there are two main categories of methodologies for new product forecasting: the diffusion model at the aggregate level and the discrete choice model/conjoint analysis at the disaggregate level (Wind, 1981). The diffusion model has the advantage of not requiring primary data, and thus it is thought to be more feasible to model car market demand in the EMs (De Jong, et al., 2004). We have reviewed different specifications of the diffusion model and their general applications in Chapter 2. At the disaggregate level, researchers are more interested in the underlying consumer preferences towards the adoption and choice decisions for the new products. The conjoint analysis is a widely used approach to investigate consumers' preferences for different features of the new products and we have discussed the key elements when conducting conjoint analysis and in particularly the choice-based conjoint analysis. The useful modelling approach with the choice based conjoint analysis data is the discrete choice models. Based on different assumptions about the correlations between choice alternatives, the discrete choice models have the modelling flexibility to account for different forms of preference heterogeneity.

In this thesis, we apply these well developed techniques that have been applied in other markets and assess how they perform in modelling demand and preferences for cars in a new market context. More importantly, we also take into consideration the different characteristics in the market context of China when using these approaches which means that we have to modify the existing approaches that have been widely used in more mature markets. In this thesis, we also propose novel modelling approaches that are inspired by the specific problems in the Chinese car market, but we think these approaches and their modelling ideas can also be replicated and tested to predict demand for other new products in other emerging economies.

7.2 Contributing Chapters of the Thesis

The first market problem that we have addressed in this thesis is how to better forecast market sales at the aggregate level with limited sales data in history. The interesting aspect of this work is to show how we can obtain demand forecasts in a situation where data on demand and sales may be unavailable or limited which is typically the case in emerging markets such as China. We propose in this thesis that we can use the car ownership data that can be extrapolated with the diffusion model to forecast car sales in a market context with short sales history such as in China. As Meade & Islam (2006) point out, little attention in the literature has been paid to how to use the diffusion model to forecast sales and compare its sales forecasting performance against other models. When forecasting quarterly and annual car sales in China, we empirically compare three basic specifications of diffusion model (Gompertz, Logistic and Bass models), two extended forms of diffusion model (Gompertz and Logistic models with both time and GDP per capita as independent variables) and three types of the benchmark models (exponential smoothing, ARIMA and linear econometric models) that directly forecast sales. By using the rolling forecasting approach to compare the forecasting results, we find that the extended Logistic model outperforms all benchmark models as well as all other diffusion model specification and provides the best sales forecasts. The superior performance of the extended Logistic model in forecasting car sales in China indicates that, first, we should select the sales forecasting method that can better cope with the market characteristics (Fildes, et al., 1998; Fildes, et al., 2008), such as the non-linear growth in the Chinese car market in our case, and second, the income effect is also significant on the car diffusion and sales in China (Dargay & Gately, 1999; Dargay, et al., 2007; Dargay, 2001). We also explicitly show that it is important to use a rolling forecast horizon approach instead of using a fixed validation sample to compare the models (Fildes, 1992; Tashman, 2000). We demonstrate that the rolling forecasting approach is particularly significant when it comes to choosing a more robust model that can cope with the rapidly changing environment in the EMs such as China.

The second research question that we tackle in the thesis is how to better understand local consumer behaviour for cars at the disaggregate level in China. Due to the short development history of the Chinese car market, local consumer behaviour in that market has not been thoroughly investigated in the literature yet. But it is also important that we highlight the value of this exercise for multinational companies that need to understand whether and how Chinese consumers differ from those in other developed markets. In this thesis, we compare the effect of variables that affect choice and adoption decisions in China and in other markets. We show that it is important for researchers to acknowledge that they need to explicitly account for variables that are specific to a market context and how these variables influence consumers' decisions. We illustrate this by looking at how consumer knowledge about cars affects their decisions to buy cars and the type of cars that they wish to buy. Different specifications of discrete choice models have been developed to explore the consumer preferences in terms of different car ownership, car type choices and future purchase intentions. The corresponding managerial implications have been put forward for car manufacturers and/or governments to effectively influence car market demand. We particularly show the significant effects of the context specific variables on local consumer preferences for cars. The preference heterogeneity across different knowledge levels has also been highlighted through a segmentation analysis in this thesis. We demonstrate through model validation that the model that accounts for the context specific variables can provide better forecasts of consumer choices than those without these variables. We also note that the importance of accounting for local market characteristics when developing marketing strategies can be applicable in the other market sectors such as retailing or fast food when the international companies established in the developed markets try to enter the EMs.

The above two research challenges that we have examined are more about the existing market demand or the conventional petrol cars existing in the market. More recently, the new concept of alternative fuel cars or green cars has attracted extensive market attention as well as government support and industry investments, which seems to be important not only in the developed car markets, but also in the EMs such as China. Therefore this thesis has further explored local consumer preferences about the new product of the green cars in this fast growing car market. We collected primary data based on the conjoint based conjoint analysis conducted in China, which includes two types of green cars (hybrid and electric cars) and one conventional petrol car in the choice set. We have investigated the effects of price, fuel/running cost and government incentives on consumer preferences to buy alternative fuel cars. More importantly, our approach differs from previous research on investigating consumer preferences for green cars that consistently holds prior assumptions about how consumers perceive the choice between different green cars and the conventional type of cars. Instead, we have compared all possible choice structures and have demonstrated the importance of choosing those models that fit a specific market context. Empirically, we find that the Chinese consumers tend to think that the hybrid cars are more similar to the petrol cars than to the electric cars, which essentially implies that if consumers are going to switch to the green cars, they are more likely to buy the hybrid cars rather than the electric cars. Given the low car ownership level in China, we have also compared the preferences of car owners versus non-car owners towards the green cars. The segmentation analysis shows that these two segments of consumers think about the green cars and conventional type of petrol cars in different ways. Car owners consider hybrid and electric cars to be similar and perceive petrol cars to be independent from them, while non-car owners perceive a stronger preference correlation between hybrid and petrol cars. In addition, the segmentation analysis also shows that car owners tend to be more sensitive to the incentive policies such as the cash subsidy than the non car owners.

After investigating static consumer preferences towards the green cars, the next research question we address in this thesis is how to effectively utilise the preference information to better forecast the new product demand in the EMs. We have tackled this research question through proposing a dynamic segmentation approach that can better account for the typical context characteristics in the EMs, such as the high level of market heterogeneity and market dynamics (Alden, et al., 2006; Burgess & Steenkamp, 2006; Sheth, 2011). Specifically, we consider that markets are segmented and that in each segment consumers can perceive alternatives and make decisions differently. The main advantage of the proposed approach over the non-segmental approach that is typically used in the literature is that our approach is able to account for not only the preferences across segments, but also within-segment heterogeneity through specifying the segment-specific choice structures. By adopting the segmentation method, our approach can also accommodate the dynamic changes of the market structure (e.g. car ownership) in the EMs. We apply this approach to forecast the long term diffusion of the green cars in China and we empirically demonstrate that if the market features such as preference heterogeneity are not well accounted for in the forecasting model, the derived demand forecasts are more likely to suffer significant bias. In addition, we apply simulation to analyse the different impacts of important factors, such as vehicle purchase price and fuel price, on the diffusion of the green cars. Such analysis can provide decision makers with a useful tool to project the new product demand based on different scenarios of the key factors in the future. As we are not able to directly validate the model, we further examine the robustness of our approach in two extreme cases. By comparing the forecasts from our approach and the benchmark model in both highest and lowest scenarios, we present a consistent conclusion that the propose model will produce less extreme forecasts for the demand of green cars.

We use the Chinese car market as an example of an emerging market context for the empirical analysis in this thesis. Importantly, when it comes to the generalisation of the findings of the research in the thesis, the proposed modelling approaches rather than the specific findings can be applied and tested for other products in other emerging markets. However, it is likely that different emerging markets have completely different context features that should be accounted for. For example, the slower car ownership growth in India (see Table 3-1) may suggest that the Gompertz model instead of the Logistic model will work better to predict car sales in India. In addition, the context specific variables, such as consumer knowledge we have selected in Chapter 4 may not be applicable in other EMs. Also, regarding the preferences towards the alternative fuel cars, taking Brazil as an example, it has developed more flexible-fuel vehicles, another type of alternative fuel vehicles that run on the mixture of petrol and ethanol fuel[41], than any other developed markets. Thus, the Brazilian consumers' preferences for the alternative fuel cars might be different from in China. Therefore, what is more valuable from this thesis is that researchers can follow our general modelling approach that we always understand local market characteristics before developing various models that can account for these context specific characteristics to investigate the local market demand.

7.3 Limitations and Directions for Future Research

This thesis has some limitations which represent avenues for further research. We discuss them following their sequence in respective contributing chapters.

- It would be interesting to see how our proposed diffusion models fare in forecasting car demand in other emerging markets such as in India and Brazil. For instance, we may be interested in whether different forms of the diffusion model will contribute to the better sales forecasting performance in different emerging market contexts. We can also develop cross-cultural diffusion models to explore the context specific factors or the different levels of the factor that differentiate the new product diffusion across several markets (Yalcinkaya, 2008).

---

[41] " Flexible-fuel vehicles in Brazil", Wikipedia, http://en.wikipedia.org/wiki/Flexible-fuel_vehicles_in_Brazil

- When using diffusion models to forecast car sales, we did not extend the Bass model with GDP per capita as an additional explanatory variable. The existing extensions of Bass model mainly focus on the inclusion of decision variables such as price and/or advertising (see the extensive review in Bass, et al., 2000), while there is little attention paid to the effect of macroeconomic factors such as GDP in the Bass model. So in the future research, we can explore how to properly develop an extended specification of Bass model to account for the GDP effect and then compare its sales forecasting performance with other extended diffusion models.

- We used the convenience sampling in our survey. We acknowledge that this sampling approach may have some limitations, such as poor representativeness for the whole population. Our solution to mitigate this limitation is that we use the national survey data to reweight our sample when estimating various models. Other sampling approaches, such as stratified sampling, can be explored if there are enough resources and a good sampling framework when conducting the survey.

- We develop car ownership and car type choice models respectively in the thesis to explore cross-market comparison of consumer behaviour for cars. However, there could be a joint decision making mechanism that consumers' decisions of adopting cars or not might also be influenced by their comparison or information searching for different types of cars. Therefore, we might be able to develop a joint model that can simultaneously describe car ownership and car type choice. In addition, consumer knowledge here is a subjective concept and we have not been able to differentiate between knowledge about cars and knowledge about the car market. For future research, it may be useful to look at how these two types of knowledge independently can affect car choice. For example, it may be interesting to look at how knowledge of cars affects the brands that consumers buy especially in such a market where consumers are less familiar with such a product. Furthermore, it would be desirable to collect data on consumer knowledge based on objective measures instead of self-evaluation.

- In this thesis, we explore consumer intentions to purchase cars in the future. There are some limitations about measuring intentions. We only provided three levels of intentions for respondents to select. We have to acknowledge that it may be desirable to capture better information about purchase intention. For example we can ask respondents to tell us their probability of purchase, which has been found to provide better predictions (Morwitz, 2001). It may also be interesting to look at what the consumers intend to purchase during cumulative time intervals, such as next year, next 2 years and next 3 years (Van Ittersum & Feinberg, 2010).

- We also acknowledge some limitations on using the intention data to forecast new product adoption behaviour. First of all, Morwitz, et al.(2007) shows the intentions are more accurate to predict actual purchases for existing products than for new ones. As cars are new to most consumers in China, respondents may have reporting bias when stating their purchase intentions. Secondly, the length of time horizon provided to the respondents (e.g. we asked the purchase intentions in next 5 years) also affects the prediction accuracy and the short time horizon is preferred (Morwitz, et al., 2007). Also, the actual purchase behaviour might change due to the unexpected change of explanatory variables (e.g. income shift or promotion) in the future (Sun and Morwitz, 2010). Furthermore, researchers find that the participation of intention survey also influences the respondents' purchase behaviour later (Morwitz, 2001). Regarding how to adjust the intention data, Morwitz (2001) suggests that the current intention can be adjusted using the historical bias in intention measures for the same type of behaviour but it requires the panel survey data. Van Ittersum & Feinberg (2010) summarise two main approaches to address the discrepancies between intentions and purchase behaviour. The first approach is to develop models that account for these discrepancies. For example, Sun and Morwitz (2010) propose a unified model that combines the intention data with actual purchase behaviour to address three types of discrepancies between intentions and actual purchase. The second approach to address the bias of intentions is to design a better scale of intentions data to be collected through the survey, such as asking the purchase probability (Morwitz, 2001; Van Ittersum & Feinberg, 2010).

- The attributes that we have considered in our choice based conjoint experiment for the green cars can be extended. It might be beneficial to further explore consumer's preferences for more attributes such as less emission levels or environmental impacts of various green cars. Our conjoint analysis follows the orthogonal design, but recently more design strategies have been developed, such as designing the stated preference experiment based on the revealed preference information, called "SP-off-RP" questions (Train & Wilson, 2008). Thus different experimental design strategies can be explored in the future and the derived results can be further compared to show whether the different experimental design contributes to our better understanding about the consumer preferences.

- Regarding the effects of incentive policies on the green cars, we have noticed that different organisations in China, including central government, local government and car manufacturers, have different incentive policies for the green car buyers. At one level it may be interesting to see which incentive policy is more effective. It may be important to investigate how these market players can coordinate to design the best incentive policies to support the development and adoptions of the green cars.

- In Chapter 6 when forecasting the demand for the green cars, we follow what is commonly assumed in the literature regarding constant preferences and unchanged choice structures for each segment over time. If panel data is available or becomes available, researchers can develop the dynamic choice models, which can account for the dynamic preferences and potential changes of the choice structures. Based on the dynamic preferences in the history, we might be able to further predict consumers' future preferences, which may help the forecasting of future market demand that we are interested in. In addition, we use an exogenous diffusion model to address the influence of product supply on the demand. It would be better to simultaneously model the decision making from both manufacturers and consumers to endogenously derive product supply versus demand (Orbach & Fruchter, 2011). Another limitation of this chapter is that we cannot directly validate our forecasting performance based on a validation sample. In addition, we have assumed that the long term penetration level is 100% in this research and it would be worth examining how the forecasts change if we assume other penetration levels.

- In the future research, we can explore a type of probability flow models to describe multiple stages of consumer adoption process. For example, Ozan, et al., (2007) propose an adoption model that consists of 4 stages' process (product awareness, need recognition, alternative evaluation and purchase decision) and there are different transition probabilities from the former stage to the later one. In addition, we also need to consider the speed of transition, which measures the time that consumers will take to move from one state to another. By linking the transition probabilities with the transition speed across all stages, we might be able to generate time series forecasts of product adoption.

7.4 Concluding Remarks

This thesis has made a significant contribution to the literature on forecasting and modelling demand in the fast growing EMs. This thesis also proposes some managerial insights for car manufacturers that are operating in such a market context. In the context of the Chinese car market, this thesis addresses four different aspects of the local market challenges: aggregate level demand forecasting, consumer preferences for the existing products, consumer preferences towards the new products of green cars, and the dynamic diffusion forecasting of the green cars. In the thesis, we empirically draw a clearer picture of the car market growth and more importantly the underlying consumer behaviour for cars in China, which we believe can better support decision making of car manufacturers as well as the policy design of local governments.

Regarding the methodological contributions, this thesis not only explores the validity of the existing approaches in this emerging market, but also proposes new models that can account for the specific characteristics of this market when forecasting new product demand. More generally, we demonstrate the importance of taking into consideration the local market characteristics when addressing the marketing challenges or problems in a different market context. We further contend that this research approach should be followed not only when moving from the more matured markets to the EMs, but also across different EMs.

APPENDICES

Appendix 1: Questionnaire for Chinese Household Vehicle Adoption Survey

Survey Introduction

First of all, we appreciate your time to attend our online survey. This survey is as part of my research for a PhD at Lancaster University in the UK. I am a Chinese PhD student working on better understanding the car market in China.

We seek to generate better forecast of the Chinese car market and we need your help in obtaining data on preferences for car purchase decisions. **Please note that even if you do not currently have a car we want to know your views too, because they are important as well.**

Depending on different family, our survey may take you 20 minutes to finish, which is completely anonymous and will not include your names, address and other characteristics that can identify you. We can assure you that all information collected in the survey will be strictly kept confidential and will be used for academic research only.

There are two main sections in the survey. The first section will ask you questions about your household status and your current car ownership. The second section will be about your future car purchasing intention and preferences.

This online survey will be live for about 3 months, until mid of April, 2010. After the survey is finished, all survey respondents are automatically entered in a prize draw as an incentive. If you are a winner you will be contacted. So we only need your contact details if you want to enter the prize draw. There will be following prizes:

- One first prize with RMB 500
- Two second prizes with RMB 200 each
- Ten third prizes with RMB 100 each

Good Luck!

Only a few questions in the survey are mandatory to answer, which are marked with an asterisk (*). For other questions, if you feel difficult to answer or cannot answer, you are free to skip some of them.

If you have any question about this survey, please feel free to contact us. (Contact Person: Lixian Qian; Email: l.qian@lancaster.ac.uk; MSN: lixian_qian@hotmail.com; telephone: +44 1524 594471)

Sincerely thank you for your cooperation!

Lancaster University Management School
Lancaster China Management Centre

Current Household Information

Firstly, we want to know about your household.

1. **Including yourself, how many members does your household have currently?**
   - 1
   - 2
   - 3
   - 4
   - 5
   - >=6
2. **How many children (under 18) live in your household?**
   - 1
   - 2
   - >=3
3. **How many adults in your household are employed?**
   - 1
   - 2
   - 3
   - >=4
4. **Do you own or rent your current residence?**
   - Own
   - Rent from the government
   - Rent from the private
   - Other(Please specify)________
5. **Do you have remaining mortgage to pay?**
   - Yes
   - No
6. **How big is your current home, measured by square meters of the building area?**
   - Smaller than 60 square meters
   - 60-90 square meters
   - 90-120 square meters
   - 120-140 square meters
   - Larger than 140 square meters

Licensed Vehicle Drivers

In this section, we want to know about the driving licence holders in your household.

1. *** In your household, how many driving licence holders? (including type A, type B, and C1 or C2 licence in China)?**
    - 0
    - 1
    - 2
    - 3 or more than 3

[Routing: If "0", the respondent will be directed to Section of "About the highest income member in Your Family", otherwise next page]

Licensed Vehicle Drivers

In this section, we want to know more about the driving licence holders in your household.

If there is more than one driving licence holder in your households, please select the one with the highest income to finish following questions.

1. **What is the relation of this driving licence holder to you?**
   - Myself
   - Husband/Wife/Unmarried Partner
   - Father/Mother/In Law
   - Grandfather/Grandmother
   - Brother/Sister
   - Son/Daughter
   - Other

2. **What is the sex of this driving licence holder?**
   - Male
   - Female

3. **What is the age group of this driving licence holder?**
   - 18-24
   - 25-34
   - 35-44
   - 45-54
   - 55-64
   - 65-69

4. **Which of the following best describes the current employment status of this driving licence holder?**
   - Full-time employed
   - Part-time employed
   - Self-employed/Freelanced
   - Unemployed, but look for a job now
   - Homemaker

- Retired
- Student
- Others

5. If employed, how far is your home from the working place of this driving licence holder? (Skip this question if not employed)

- Work at home.
- Shorter than 1 km
- 1-3 km
- 4-6 km
- 7-10 km
- 11-25 km
- Longer than 25 km

6. Is there free parking space at working place of this driving licence holder? (Skip this question if not employed)

- Yes
- No

7. What is the most frequently used transportation of this driving licence holder to his/her working place?

- Walking, Bicycle
- Moped, Motorcycle
- Public Transport (Bus, Underground or Train)
- Employer's Shuttle Bus
- Taxis
- Car (Drive alone)
- Carpool (including sharing a car with your family member)
- Other (Please specify) __________

[Skip to Section of "Current Car Ownership"]

About the highest income member in Your Family

Since there is no driving licence holder in your household, we want to know more information about the family member with the highest income.

1. **What is the relation of this family member to you?**
    - Myself
    - Husband/Wife/Unmarried Partner
    - Father/Mother/In Law
    - Grandfather/Grandmother
    - Brother/Sister
    - Son/Daughter
    - Other

2. **What is the sex of this family member?**
    - Male
    - Female

3. **What is the age group of this family member?**
    - 18-24
    - 25-34
    - 35-44
    - 45-54
    - 55-64
    - 65-69

4. **Which of the following best describes the current employment status of this family member?**
    - Full-time employed
    - Part-time employed
    - Self-employed/Freelanced
    - Unemployed, but look for a job now
    - Homemaker
    - Retired
    - Student
    - Others

5. **If employed, how far is your home from the working place of this family member? (Skip this question if not employed)**
   - Work at home.
   - Shorter than 1 km
   - 1-3 km
   - 4-6 km
   - 7-10 km
   - 11-25 km
   - Longer than 25 km

6. **Is there free parking space in the working place of this family member? (Skip this question if not employed)**
   - Yes
   - No

7. **What is the most frequently used transportation for this family member to his/her working place? (If not employed, select the most frequently used transportation during of daily travelling)**
   - Walking, Bicycle
   - Moped, Motorcycle
   - Public Transport (Bus, Underground or Train)
   - Employer's Shuttle Bus
   - Taxis
   - Car (Drive alone)
   - Carpool (including sharing a car with your family member)
   - Other (Please specify)__________

Current Car Ownership

In this section, we want to know your household's current car ownership.

1. *** How many passenger vehicles do your household members use, including purchased, long-term rented and company cars? (Passenger vehicles include sedan, hatchback, SUV, MPV, Sports car/Coupe and Station Wagons)**
    - 0
    - 1
    - 2
    - >= 3

[Routing: if the answer is "0", the respondent will be directed to Section of "future car adoption intention"; otherwise, the respondent will continue with next page.]

2. **Please rate your knowledge level for automobile and automobile market.**
    - Very familiar
    - Familiar
    - Basic knowledge
    - unfamiliar

Current Car Ownership

We want to know more about your owned vehicles.

If your household have more than one vehicle, please select the **main** one as the representative.

1. **What is the make or brand of this vehicle? (e.g. Volkswagen, Toyota, Buick)**

   [Drop down list here including both domestic made and imported makes.]

2. **What is the model of this vehicle? (e.g. Passat, Corolla, Excelle)**

3. **How did your household get this vehicle?**
   - Purchased Please go to Q4
   - Long-term rented Please go to Q5
   - Company Car Please go to Q5
   - Other (Please specify) Please go to Q5

     ____

4. **Was this vehicle bought by loan or did you pay it by instalment?**
   - Yes
   - No

5. **Was this vehicle a new or used one when purchased?**
   - New car
   - Second hand car

6. **Was this vehicle imported or domestically made?**
   - Imported
   - Domestically made (including made by joint-ventures)

7. **What is the body type of this vehicle?**
   - Hatchback/2-box car
   - Sedan/3-box car
   - Sports Utility Vehicle (SUV, including off-road vehicle and jeep)

- Multi-purpose Vehicle (MPV, such as Buick GL8, Honda Odyssey or JAC Refine)
- Sports Car/Coupe (Such as MG TF, Geely China Dragon or BYD S8)
- Station Wagon (Such as Buick Excelle SW or Brilliance Splendour Wagon)
- Other

8. When did you obtain or start to use this vehicle?

_________Year ____________Month

9. If you purchased the car in 2009, please answer this question, otherwise please skip to question 10.

How would you make your decision if there were no purchase incentive policy in 2009, i.e. halved purchasing tax for small engine cars?

- Would have no car purchased.
- Would have bought a similar car in terms of price and engine size
- Would have bought a different car in terms of price and engine size

10. Excluding purchasing tax and other expenses, what was the purchasing price of this vehicle, measured by 10,000 RMB?

____________________ (10,000 RMB)

11. What is the engine size of this vehicle?

- Less than 1.0L
- 1.0L - 1.3L
- 1.4L – 1.6L
- 1.7L – 2.0L
- 2.1L – 2.4L
- 2.5L – 2.9L
- Equal or larger than 3.0L

12. What is the transmission type of this vehicle?

- Manual (MT)
- Automatic (AT)
- Automated Mechanical Transmission/Manumatic (AMT)
- Continuously Variable Transmission (CVT)

- Direct-Shift Gearbox (DSG)
- Other

13. What is the fuel type of this vehicle?

- Petrol
- Diesel
- Liquefied petroleum gas (LPG)
- Hybrid electric
- Electric
- Other_______

14. Who is the main user/driver of this vehicle in your family?

- Yourself
- Your spouse/husband/wife/unmarried partner
- Your father/mother/in-law
- Your son/daughter
- Other __________________

15. What is the primary purpose of this vehicle?

- Driving between home and working place
- Escort child(ren) to/from school or other family members to/from working places
- Business requirement during the work
- Trips or Shopping in holiday and weekends
- Visiting relatives/friends
- Other (Please specify)__________________

16. Please select the other purposes of this vehicle? (you can choose more than one purposes here)

- o Driving between home and working place
- o Escort child(ren) to/from school or other family members to/from working places
- o Business requirement during the work
- o Trips or Shopping in holiday and weekends
- o Visiting relatives/friends

- Other (Please specify)________________

17. What is the annual mileage of this vehicle? If purchased within 1 year, please calculate proportionately.

- Less than 4000km
- 4001-8000km
- 8001-12000km
- 12001-16000km
- 16001-20000km
- 20001-24000km
- 24001-28000km
- 28001-32000km
- 32001-36000km
- 36001-40000km
- More than 40000km

18. Please tell us the approximate running costs of this vehicle in last year.

- Monthly Fuel cost ____________RMB/Month
- Annual Insurance cost ____________RMB/Year
- Annual Maintenance/Repair cost ____________RMB/Year
- Annual Tolls/Parking cost ____________RMB/Year
- Other expenses/fees ____________RMB/Year

  (Such as road tax, vehicle & vessel tax, MOT)

19. Does your household receive some subsidy from the household members' employers for your car running cost? How much is it per month in RMB?

- No Subsidy
- Less than 100 RMB/Month
- 101-300 RMB/Month
- 301-500 RMB/Month
- 501-1000 RMB/Month
- 1001-1500 RMB/Month
- 1501-2000 RMB/Month
- 2001-2500 RMB/Month

- More than 2500 RMB/Month

20. Before you decided to buy this vehicle, did you consider other made/model?

- Yes →go to Q21
- No →go to Q22

21. Please tell us your second choice besides this vehicle?

Make: [select from the drop-down list]

Model: ________________________

22. Did you get rid of other vehicle within 6 months before or after you got this vehicle?

- Yes →go to Q23 and Q24
- No →go to next page

23. What was the make (brand) of that vehicle you got rid of?

____

24. How did you get rid of it?

- Sold to dealer
- Sold to private
- Left with household members/relatives/friends
- Scrapped
- Stolen
- Other__________________

Future Car Purchasing Intention

In this section, we would like to know about your car purchasing intention and plan in next 5 years (2010-2015)

1. *** Based on your household income growth expectation, can your household afford a lowest priced car (new or second hand) in next 5 years? Please note that the lowest priced new car is currently around RMB 30,000.**
    - Yes
    - No
    - Not sure

[Routing: if the answer is "No", the respondent will skip the Future Car Purchasing Intention section and be directed to the last section of "Household Location & Income"; otherwise, the respondent will continue with next page.]

Future Car Purchasing Intention

We want to know more about your car purchasing intention and plan in next 5 years (2010-2015)

1. **If you can afford, do you intend or plan to buy a passenger vehicle in next 5 years, including replacing currently owned vehicle?**
    - Yes →Please go to Q3
    - Not sure →Please go to Q3
    - No →Please go to Q2

2. **Please select your reason not to buy a vehicle in the next 5 years.**
    - Cannot afford the car price with limited income
    - Although my household can afford the car price, we cannot afford the running cost every month.
    - We will have other big expense in next five years, such as housing, education, etc.
    - Although we can afford a car, we feel it is unnecessary or not economic for us to own a car.
    - We have had a car in our household, but we don't plan to replace it in next 5 years.
    - Other (Please specify)______________________

3. **If you intend to buy a car, when do you plan to buy it?**
    - 2010
    - 2011
    - 2012
    - 2013
    - 2014
    - 2015
    - No detail plan now.

4. **If you intend to buy a car, will you buy a new or a used car?**
    - New car
    - Second hand car

5. **If you intend to buy a car, will you buy an imported or a domestically made car?**
    - Imported
    - Domestically made (including made by joint-ventures)

6. **Which origin of brand do you prefer for your future car?**
    - Chinese local brands (e.g. Chery, Geely, BYD)
    - Brands from the US (e.g. Buick, Chevrolet, Ford)
    - Brands from Germany (e.g. VW, Audi, Benz, BMW)
    - Brands from Japan (e.g. Toyota, Honda, Nissan, Mazda)
    - Brands from South Korea (e.g. Hyundai, Kia)
    - Brands from France (e.g. Citron, Peugeot)
    - Doesn't matter for me
    - Others (Please Specify) ______________

7. **If purchasing a car in future, will you replace the existing one or add one more (including buying the first car)?**
    - Replace the existing one.
    - Add another car (including buying your first car).

8. **Will you purchase a vehicle by loan or by instalment?**
    - Yes
    - Will compare and then decide
    - No

9. **From the available models in current market, have you found one vehicle that best matches your intention?**
    - Yes →Please go to Q9
    - No →Please go to next page
    - Don't know. →Please go to next page

10. **Please tell us the make and model which is best matched with your intention.**

    Make

    [select from drop-down list]

    Model:______________________

Future Car Purchasing Intention

We also want to know five most important factors for your household's car purchasing intention.

1. **According to the order of importance in your opinion, please select and rank top 5 factors you will consider when buying a vehicle for your household.**
    - Safety, which is the vehicle ability to prevent the accidents
    - Reliability, which is the vehicle ability without running faults and be able to work continuously.
    - Performance, including engine size, power, speed, etc.
    - Handling, including the driving-related characteristics, such as brakes, power steering and gear shifting etc.
    - Space & Capacity, including passenger and luggage space
    - Comfort, including leg room, seat comfort, etc
    - Brand
    - Style, including appearance, colour, looks etc
    - Purchasing Price
    - Running Economy, including fuel consumption, maintenance & repairing costs, etc.
    - After sale Service
    - Depreciation
    - Fuel type
    - Emission/pollution level
    - Age (if considering used cars, including age, mileage, vehicle condition)
    - Other factors

Future Car Purchasing Intention: Your Own Design

Based on your economic and household conditions, please select attributes in following 5 aspects to design a car matching your future car purchasing intention.

Please note except that the first one is mandatory to answer, you are free to select all or part of attributes in other 4 aspects to design your intended car.

1. *** If you have car purchasing intention, what will be your most possible car price range?**
    - Less than 50,000 RMB
    - 50,000-100,000 RMB
    - 100,000-150,000 RMB
    - 150,000-200,000 RMB
    - 200,000-300,000 RMB
    - Higher than 300,000 RMB

    [The intended price range will be used as the base price of petrol car in stated choice experiments]

2. **Size and General Attributes**

	Body Type	**No. of doors**	**Car size (Only for Sedan & hatchback)**	**No. of seats**
Your Design	o Hatchback o Sedan o SUV o MPV o Coupe o Estate o Doesn't matter	o 2 doors o 4 doors o Doesn't matter o Doesn't matter	o Mini (e.g. Chery QQ) o Small (e.g. Honda Fit) o Midsize (e.g. Buick Excelle) o Upper-mid (e.g. VW Passat) o Luxury (e.g. Toyota Crown) o Doesn't matter	o 2 o 4 o 5 o 6 o 7 o Doesn't matter
	Boot size	**Fuel capacity**	**Fuel type**	**Gearbox**
Your Design	o Small (<400L) o Midsize (400-500L) o Large (>500L) o Doesn't matter	o Small (<=50L) o Midsize (50-65L) o Large (>65L) o Doesn't matter	o Gasoline o Diesel o CNG o Hybrid o EV o HFCV o Doesn't matter	o MT o AT o AMT o CVT o DSG o Doesn't matter

3. Performance

	Engine size	Top Speed (km/hr)	Urban Fuel consumption (L/100km)	0-100 km/h acceleration (s)
Your Design	o <1.0L o 1.0-1.3L o 1.4-1.6L o 1.7-2.0L o 2.1-2.4L o 2.5-2.9L o >=3.0L o Doesn't matter o Don't know	o <150 o 150-180 o 180-210 o 210-230 o >230 o Doesn't matter o Don't know	o <8L o 8L-10L o 10L-12L o 12L-15L o >15L o Doesn't matter o Don't know	o <9s o 9-11s o 11-13s o 13-15s o >15s o Doesn't matter o Don't know

4. Safety & Security

	Airbags	Alarm	Door Locking	Immobiliser	Power steering
Your Design	o 0 o 1 o 2 o 4 o >=6 o Doesn't matter o Don't know	o Need o No need o Doesn't matter o Don't know	o Central controlled o Remote & Central controlled o Keyless Entry System o Doesn't matter o Don't know	o Need o No need o Doesn't matter o Don't know	o Need o No need o Doesn't matter o Don't know
	Parking Distance Control	**Seatbelt alarm**	**ABS**	**ESP**	
Your Design	o Need o No need o Doesn't matter o Don't know	o Need o No needs o Doesn't matter o Don't know	o Need o No Need o Doesn't matter o Don't know	o Need o No need o Doesn't matter o Don't know	

5. Comfort & convenience

	Audio system	Steering Adjustment	Sunroof	Door Mirror Adjustment	Power Window
Your Design	o Radio only o CD o DVD o Doesn't matter o Don't know	o No o 2D (height) o 4D (height & depth) o Doesn't matter o Don't know	o No need o Manual o Electric o Doesn't matter o Don't know	o Manual o Electric o Electric & heated o Doesn't matter o Don't know	o No o Front only o Both front and rear windows o Doesn't matter o Don't know
	Seat Trim	**Front Seat Adjustment**	**Rear seat folding**	**Sat Nav (GPS)**	
Your Design	o Cloth o Leather o Doesn't matter o Don't know	o Manual o Electric o Electric & heated o Doesn't matter o Don't know	o No o Fold o Split & fold o Doesn't matter o Don't know	o Need o No need o Doesn't matter o Don't know	

Hypothetical Choices

In the following second portion, we will conduct hypothetical choice experiment for green car demand.

You will be provided with 8 choice exercises, and each of them consists of three vehicles with different fuel types:

- Petrol car: overwhelming majority of existing cars in the market is consuming petrol.
- Hybrid Electric car: this type of vehicles is equipped with an electric motor as the supplement to conventional petrol engine to drive the vehicle. Hybrid car here is specified as those without the requirement of charging.
- Electric car: they are purely based on battery instead of conventional petrol to propel the vehicle. Due to the capacity limitation of the battery, electric car has the range limitation after one charging and has to be recharged frequently.

In the following exercises, we set up several features for each vehicle:

- Price: the purchasing price of each vehicle.
- Running cost: the annual total expense to run a car, including fuel or charging cost, maintenance and repairing costs, insurance, and toll and parking costs, etc.
- Incentive: in order to encourage the usage of green car, there might be some different incentives, such as cash allowance, free parking or using priority lane.
- Charging facility: the availability of charging facilities for electric cars, measured by percentage of parking spaces.
- Range: the length of road the electric car can run after full charging.

In each exercise, please read and compare each hypothetical vehicle and its features, and then select one that your household would **most likely** purchase. Please repeat these steps for 8 exercises.

1. **Eight exercises will be randomly selected for you based on your birthday. Please select the quarter your birthday belongs to.**
    - First Quarter (Jan. – Mar.) *[route to 1st group of exercise]*
    - Second Quarter (Apr. – Jun.) *[route to 2nd group of exercise]*
    - Third Quarter (Jul. – Sept.) *[route to 3rd group of exercise]*
    - Fourth Quarter (Oct. – Dec.) *[route to 4th group of exercise]*

[Two examples for respondents with intended purchasing price of less than 50,000 RMB. Please refer to Appendix 2: for the 32 experimental scenarios created using orthogonal design]

All hypothetical choices are mandatory to answer.

1. **Suppose there are following 3 cars in the market, please select one your household would most likely purchase.**

	Vehicle 1	Vehicle 2	Vehicle 3
Fuel Type	Petrol	Hybrid	Electric
Price (RMB)	40,000	72,000	52,000
Annual Running cost (RMB)	12,000	4,800	7,200
Incentive	No	Cash Allowance 20,000 RMB	Using priority lane
Charging facility	Not Applicable	Not Applicable	40% of parking spaces
Range	Not Applicable	Not Applicable	80km
Your choice:	☐	☐	☐

2. **Suppose there are following 3 cars in the market, please select one your household would most likely purchase.**

	Vehicle 1	Vehicle 2	Vehicle 3
Fuel Type	Petrol	Hybrid	Electric
Price (RMB)	40,000	60,000	52,000
Annual Running cost (RMB)	12,000	9,600	4,800
Incentive	No	Cash allowance 20,000 RMB	Cash allowance 30,000 RMB
Charging Facility	Not Applicable	Not Applicable	70% of parking spaces
Range	Not Applicable	Not Applicable	80km
Your choice:	☐	☐	☐

Household Location & Income

Finally, we want to know your residence location and Income information.

We can assure you that all information collected will be strictly kept confidential. But if you are not able to answer some questions below, you are free to skip them.

1. **Which province, autonomous region or municipal city of China Mainland is your household living in?**
   [Drop-down list here]

2. **Which city is your household living in?**
   ____

3. **Please provide the postal code of your residence place.**
   ____

4. **What was your household total annual income in 2009, after deducting income tax and various social security schemes paid by yourselves?**
   - Less than 20,000 RMB
   - 20,000 – 39,999 RMB
   - 40,000 – 59,999 RMB
   - 60,000 – 79,999 RMB
   - 80,000 – 99,999 RMB
   - 100,000 – 129,999 RMB
   - 130,000 – 159,999 RMB
   - 160,000 – 189,999 RMB
   - 190,000 – 219,999 RMB
   - 220,000 – 259,999 RMB
   - 260,000 – 299,999 RMB
   - 300,000 – 349,999 RMB
   - 350,000 – 399,999 RMB
   - 400,000 – 449,999 RMB
   - 450,000 – 499,999 RMB
   - 500,000 RMB or more

Thanks for your Cooperation

Thank you for completing the survey.

1. **If you are selected to attend this survey by a university student, please specify university and his/her name.**
   Student's name ______________________

2. **If you want to enter the prize draw activity to win a prize, please provide your email address or phone number so that we can contact you for the prize result.**
   Email: ____________________________________
   Or Telephone no. (including area code)
   ____

Appendix 2: 32 Choice Cards from the Orthogonal Experiment Design

Card ID	Price of HEV	Price of EV	Running cost of HEV	Running cost of EV	Incentive of HEV	Incentive of EV	Charging Facility	Range of EV
1	80% higher	30% higher	60% less	40% less	Cash Allowance	Priority Lane	40% of parking slots	80km
2	50% higher	30% higher	20% less	60% less	Cash Allowance	Cash Allowance	70% of parking slots	80km
3	80% higher	50% higher	60% less	60% less	Priority Lane	Cash Allowance	10% of parking slots	80km
4	50% higher	50% higher	20% less	60% less	Priority Lane	Priority Lane	10% of parking slots	120km
5	80% higher	80% higher	60% less	40% less	Priority Lane	Cash Allowance	70% of parking slots	80km
6	30% higher	50% higher	20% less	40% less	Free Parking	Free Parking	10% of parking slots	80km
7	50% higher	80% higher	40% less	20% less	Cash Allowance	Priority Lane	10% of parking slots	80km
8	80% higher	80% higher	20% less	20% less	Cash Allowance	Cash Allowance	10% of parking slots	80km
9	80% higher	50% higher	40% less	40% less	Free Parking	Cash Allowance	10% of parking slots	120km
10	30% higher	30% higher	40% less	20% less	Priority Lane	Cash Allowance	10% of parking slots	80km
11	50% higher	80% higher	20% less	40% less	Cash Allowance	Cash Allowance	10% of parking slots	80km
12	30% higher	30% higher	20% less	40% less	Cash Allowance	Cash Allowance	70% of parking slots	120km
13	80% higher	80% higher	60% less	60% less	Cash Allowance	Free Parking	10% of parking slots	80km
14	80% higher	30% higher	20% less	20% less	Priority Lane	Free Parking	10% of parking slots	160km
15	50% higher	80% higher	60% less	20% less	Free Parking	Cash Allowance	40% of parking slots	120km
16	80% higher	50% higher	20% less	20% less	Cash Allowance	Cash Allowance	40% of parking slots	160km

To be continued on the next page

Card ID	Price of HEV	Price of EV	Running cost of HEV	Running cost of EV	Incentive of HEV	Incentive of EV	Charging Facility	Range of EV
17	30% higher	50% higher	60% less	20% less	Cash Allowance	Priority Lane	70% of parking slots	160km
18	80% higher	80% higher	20% less	20% less	Free Parking	Free Parking	70% of parking slots	80km
19	30% higher	80% higher	40% less	20% less	Priority Lane	Cash Allowance	40% of parking slots	80km
20	30% higher	80% higher	60% less	20% less	Cash Allowance	Free Parking	10% of parking slots	120km
21	80% higher	50% higher	20% less	20% less	Cash Allowance	Cash Allowance	40% of parking slots	80km
22	80% higher	80% higher	40% less	40% less	Cash Allowance	Priority Lane	10% of parking slots	160km
23	80% higher	80% higher	40% less	60% less	Free Parking	Cash Allowance	70% of parking slots	160km
24	80% higher	80% higher	20% less	20% less	Cash Allowance	Cash Allowance	10% of parking slots	120km
25	30% higher	80% higher	20% less	60% less	Cash Allowance	Cash Allowance	10% of parking slots	160km
26	30% higher	80% higher	20% less	60% less	Free Parking	Priority Lane	40% of parking slots	80km
27	50% higher	50% higher	40% less	20% less	Cash Allowance	Free Parking	70% of parking slots	80km
28	80% higher	30% higher	20% less	20% less	Free Parking	Priority Lane	10% of parking slots	80km
29	80% higher	80% higher	20% less	20% less	Priority Lane	Priority Lane	70% of parking slots	120km
30	50% higher	30% higher	60% less	20% less	Free Parking	Cash Allowance	10% of parking slots	160km
31	80% higher	30% higher	40% less	60% less	Cash Allowance	Free Parking	40% of parking slots	120km
32	50% higher	80% higher	20% less	40% less	Priority Lane	Free Parking	40% of parking slots	160km

Appendix 3: NLOGIT Code for Discrete Choice Models

A3.1 Binary logit model for car ownership

```
Sample  ; all$
NLOGIT; LHS= chosen,cset,alt_ij
        ; choices = NoCar, Car
        ; wts=wght
        ; model:

        U(NoCar)= 0 /

        U(Car)=asc_car                  ? Constant of car owners
        +b_inc * inc_new                ? Household income
        +b_incsq * inc_sq               ? Square of income
        +b_wadult * wadult              ? Working adults
        +b_driver * dri_pro             ? Drivers
        +b_ownhouse * ownhouse          ? House ownership
        +b_male * hsex                  ? Gender of household head
        +b_age * age                    ? Age of household head
        +b_chd * child                  ? Number of children
        +b_fsize * fam_size             ? Family size
        +b_urban * urban                ? Urban areas of cities
        +b_suburb * suburb              ? Suburban area of cities
        +b_NA_Urb * NA_urban            ? Unsure of urban or suburban area
        +midknow * midknow              ? Basic knowledge level
        +highknow * hiknow              ? Good knowledge level

        ;crosstab$
```

A3.2 Multinomial logit model for car type choices

```
Sample  ; all$
Nlogit  ; Lhs=chosen, cset, index
        ; Choices=Small, Midsize, Large
        ; maxit=200
        ; Model:

        ? Attributes of small cars ?
        U(Small) = Price * Pri_inc       ? Price/Income
        +price2 * pri_inc2               ? Square of (price/income)
        +fuelcost * YRFLCOST             ? Annual fuel cost
        +Perform * Factor                ? Performance factor
        +Airbag * airbags                ? No. of airbags
        +Eur_COO * Europe                ? European brands
        +JPKR_COO * JPKR                 ? Japanese/Korean brands
        +USA_COO * American/             ? American brands

        ? Attributes of mid-sized cars & Individual demographics?
        U(Midsize) = Price * Pri_inc     ? Price/Income
        +price2 * pri_inc2               ? Square of (price/income)
        +fuelcost * YRFLCOST             ? Annual fuel cost
        +Perform * Factor                ? Performance factor
        +Airbag * airbags                ? No. of airbags
        +Eur_COO * Europe                ? European brands
        +JPKR_COO * JPKR                 ? Japanese/Korean brands
        +USA_COO * American              ? American brands
        +M_age * age                     ? Age of household head
        +M_male * hsex                   ? Gender of household head
        +m_urban * urban                 ? Living in urban area
        +M_cars * cars                   ? No. of owned cars
        +M_dist * distance               ? Commuting distance
        +M_WkUse * Work                  ? Primary use of cars (work)
        +M_midknow * Midknow             ? Basic knowledge level
        +M_Hiknow * Highknow             ? Good knowledge level
        +M_Dummy /                       ? Constant of mid-sized cars

        ? Attributes of large cars & Individual demographics?
        U(Large) = Price * Pri_inc       ? Price/income
        +price2 * pri_inc2               ? Square of (price/income)
        +fuelcost * YRFLCOST             ? Annual fuel cost
        +Perform * Factor                ? Performance factor
        +Airbag * airbags                ? No. of airbags
        +Eur_COO * Europe                ? European brands
        +JPKR_COO * JPKR                 ? Japanese/Korean brands
        +USA_COO * American              ? American brands
        +L_age * age                     ? Age of household head
        +L_male * hsex                   ? Gender of household head
        +L_urban * urban                 ? Living in urban area
        +L_child * Child                 ? No. of children
        +L_cars * Cars                   ? No. of owned cars
        +L_wkuse * work                  ? Primary use of cars (work)
        +L_dist * distance               ? Commuting distance
        +L_midknow * Midknow             ? Basic knowledge level
        +L_Hiknow * Highknow             ? Good knowledge level
        +L_Dummy                         ? Constant of large cars

        ; crosstable$
```

A3.3 Multinomial logit model for purchase intentions

```
Sample; all$
NLOGIT; LHS= chosen,cset,index
        ; choices = NoInt, Unsure, HighInt
        ; wts=wght
        ; model:

        U(NoInt)= 0 /

        ? Demographics of individuals unsure of intentions ?
        U(Unsure)= asc_uns              ? Constant for consumers unsure of intention
        +uns_inc * inc_new              ? Household income
        +uns_wdl * wadult               ? No. of working adults
        +uns_dri * dri_pro              ? Drivers
        +uns_sex * Hsex                 ? Gender of household head
        +uns_age * age                  ? Age of household head
        +uns_fmsz * famsize             ? Family size
        +uns_chd * child                ? No. of children
        +uns_car * cars_bi              ? Owning cars already or not
        +uns_mknw * midknow             ? Basic knowledge level
        +uns_hknw * hiknow/             ? Good knowledge level

        U(HighInt)= asc_high            ? Constant for consumers with high intentions
        +hig_inc * inc_new              ? Household income
        +hig_wdl * wadult               ? No. of working adults
        +hig_dri * dri_pro              ? Drivers
        +hig_sex * Hsex                 ? Gender of household head
        +hig_age * age                  ? Age of household head
        +hig_fmsz * famsize             ? Family size
        +hig_chd * child                ? No. of children
        +hig_car * cars_bi              ? Owning cars already or not
        +hig_mknw * midknow             ? Basic knowledge level
        +hig_hknw * hiknow              ? Good knowledge level

        ;crosstab$
```

A3.4 Nested logit model for the choices with alternative fuel cars

```
Sample ; all$
NLOGIT; Lhs=CARFUEL
        ; Choices=Electric,Hybrid,Petrol

        ? Following defines 3 NL tree structures. Only one structure is used in each model. ?

        ? Tree structure 1 in Figure 5-1?
        ; Tree=GREEN(Electric,Hybrid),Conv(Petrol)
        ; IVSET: (Conv)=[1.0]

        ? Tree structure 2 in Figure 5-1?
        ; Tree=No_Oil(Electric),Oil(Hybrid,Petrol)
        ; IVSET: (No_Oil)=[1.0]

        ? Tree structure 3 in Figure 5-1?
        ; Tree=One_Fuel(Electric,Petrol),Mixed(Hybrid)
        ; IVSET: (Mixed)=[1.0]

        ; Wts=WGHT
        ; Pds=8                        ? 8 choice scenarios for each respondents
        ; maxit=150
        ; start=logit                  ? Starting from the MNL model

        ; Model:

        ? Attributes and incentives of electric cars, as well as demographic characteristics ?
        U(Electric)=bprice * Price     ? Price of electric cars
        +bcost * RCost                 ? Running cost
        +bCash * Cash                  ? Cash subsidy
        +bpark * FreePark              ? Free parking
        +bPLane * Prilane              ? Fast lane
        +bcharge * chargef             ? Charging facility
        +bRange * Range                ? Range
        +e_Asc                         ? Constant of electric cars
        +be_hhs * hhsize               ? Household size
        +be_child * Child              ? No. of children
        +be_drive * Driver             ? No. of drivers
        +be_inc * income               ? Household income
        +be_age * age                  ? Age of household head
        +be_sex * sex                  ? Gender of household head
        +be_dis * Distance             ? Commuting distance
        +be_car * Cars/                ? No. of owned cars

        ? Attributes and incentives of hybrid cars, as well as demographic characteristics ?
        U(Hybrid)=bprice * Price ?Price of hybrid cars
        +bcost * RCost                 ? Running cost
        +bCash * Cash                  ? Cash subsidy
        +bpark * FreePark              ? Free parking
        +bPLane * Prilane              ? Fast lane
        +h_Asc                         ? Constant of hybrid cars
        +bh_hhs * hhsize               ? Household size
        +bh_child * Child              ? No. of children
        +bh_drive * Driver             ? No. of drivers
        +bh_inc * income               ? Household income
        +bh_age * age                  ? Age of household head
        +bh_sex * sex                  ? Gender of household head
        +bh_dis * Distance             ? Commuting distance
        +bh_car * Cars/                ? No. of owned cars
```

```
? Attributes of petrol cars ?
U(Petrol)=bprice * Price          ? Price of petrol cars
+bcost * RCost                    ? Running cost

; Crosstab$
```

BIBLIOGRAPHY

Adler, T., Wargelin, L., & Kostyniuk, L. (2003). *Incentives for Alternate Fuel Vehicles: A Large-Scale Stated Preference Experiment.* Paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne, Switzerland.

Aguirregabiria, V., & Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics, 156*(1), 38-67.

Ahn, J., Jeong, G., & Kim, Y. (2008). A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Economics, 30*(5), 2091-2104.

Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research, 13*(4), 411-454.

Albaum, G., & Tse, D. K. (2001). Adaptation of international marketing strategy components, competitive advantage, and firm performance: A study of Hong Kong exporters. *Journal of International Marketing, 9*(4), 59-81.

Alden, D. L., Steenkamp, J. B. E. M., & Batra, R. (2006). Consumer attitudes toward marketplace globalization: Structure, antecedents and consequences. *International Journal of Research in Marketing, 23*(3), 227-239.

Armstrong, J. S., Morwitz, V. G., & Kumar, V. (2000). Sales forecasts for existing consumer produces and services: Do purchase intentions contribute to accuracy? *International Journal of Forecasting, 16*(3), 383-397.

Bass, F. M. (1969). New Product Growth Model for Consumer Durables. *Management Science, 15*(5), 215-227.

Bass, F. M., Jain, D. C., & Krishnan, T. V. (2000). Modeling the Marketing-Mix Influence in New-Product Diffusion. In V. Mahajan, E. Muller & Y. Wind (Eds.), *New-Product Diffusion Models* (pp. 99-122). Dordrecht: Kluwer Academic Publishers.

Bass, F. M., Krishnan, T. V., & Jain, D. C. (1994). Why the Bass Model Fits without Decision Variables. *Marketing Science, 13*(3), 203-223.

Batra, R. (1997). Executive insights: Marketing issues and challenges in transitional economies. *Journal of International Marketing, 5*(4), 95-114.

Bayus, B. L. (1992). The Dynamic Pricing of Next Generation Consumer Durables. *Marketing Science, 11*(3), 251-265.

Bemmaor, A. C. (1995). Predicting Behavior from Intention-to-Buy Measures - the Parametric Case. *Journal of Marketing Research, 32*(2), 176-191.

Ben-Akiva, M., & Lerman, S. R. (1985). *Discrete choice analysis : theory and application to travel demand.* Cambridge, MA.: MIT Press.

Berkovec, J. (1985). New Car Sales and Used Car Stocks - a Model of the Automobile Market. *Rand Journal of Economics, 16*(2), 195-214.

Berkovec, J., & Rust, J. (1985). A Nested Logit Model of Automobile Holdings for One Vehicle Households. *Transportation Research Part B-Methodological, 19*(4), 275-285.

Bettman, J. R., & Park, C. W. (1980). Effects of Prior Knowledge and Experience and Phase of the Choice Process on Consumer Decision-Processes - a Protocol Analysis. *Journal of Consumer Research, 7*(3), 234-248.

Bhat, C. R., & Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B-Methodological, 41*(5), 506-526.

Bhat, C. R., & Pulugurta, V. (1998). A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. *Transportation Research Part B-Methodological, 32*(1), 61-75.

Bottomley, P. A., & Fildes, R. (1998). The role of prices in models of innovation diffusion. *Journal of Forecasting, 17*(7), 539-555.

Bouachera, T., & Mazraati, M. (2008). Fuel demand and car ownership modelling in India. *OPEC Review, 31*(1), 27-51.

Brownstone, D., Bunch, D. S., & Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B-Methodological, 34*(5), 315-338.

Brucks, M. (1985). The Effects of Product Class Knowledge on Information Search Behavior. *Journal of Consumer Research, 12*(1), 1-16.

Bunch, D. S., Bradley, M., Golob, T. F., Kitamura, R., & Occhiuzzo, G. P. (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.

Bunch, D. S., & Chen, B. (2008). Automobile demand and type choice. In D. A. Hensher & K. J. Button (Eds.), *Handbook of Transport Modelling* (2nd ed., pp. 541-557). Amsterdam ; London: Elsevier.

Burgess, S. M., & Harris, M. (1999). Social Identity in an Emerging Consumer Market: How You Do the Wash May Say a Lot About Who You Think You Are. *Advances in Consumer Research, 26*(1), 170-175.

Burgess, S. M., & Steenkamp, J. B. E. M. (2006). Marketing renaissance: How research in emerging markets advances marketing science and practice. *International Journal of Research in Marketing, 23*(4), 337-356.

Burgstaller, S., Rothery, T., Archambault, P., & Yuzawa, K. (2009). Identifying global long-term winners. global automotive research team report on November 27, 2009. Goldman Sachs.

Button, K., Ngoe, N., & Hine, J. (1993). Modeling Vehicle Ownership and Use in Low-Income Countries. *Journal of Transport Economics and Policy, 27*(1), 51-67.

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2006). Neighborhood design and vehicle type choice: Evidence from Northern California. *Transportation Research Part D-Transport and Environment, 11*(2), 133-145.

Cattin, P., & Wittink, D. R. (1982). Commercial Use of Conjoint-Analysis - a Survey. *Journal of Marketing, 46*(3), 44-53.

Caulfield, B., Farrell, S., & McMahon, B. (2010). Examining individuals preferences for hybrid electric and alternatively fuelled vehicles. *Transport Policy, 17*(6), 381-387.

Cavusgil, S. T., & Zou, S. (1994). Marketing Strategy-Performance Relationship: An Investigation of the Empirical Link in Export Market Ventures. *The Journal of Marketing, 58*(1), 1-21.

Chamon, M., Mauro, P., & Okawa, Y. (2008). Mass car ownership in the emerging market giants. *Economic Policy, 23*(54), 243-296.

Chan, K., Li, L., Diehl, S., & Terlutter, R. (2007). Consumers' response to offensive advertising: a cross cultural study. *International Marketing Review, 24*(5), 606-628.

Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2005). Do intentions really predict behavior? Self-generated validity effects in survey research. *Journal of Marketing, 69*(2), 1-14.

Chin, G. T. (2010). *China's automotive modernization : the party-state and multinational corporations*. Basingstoke: Palgrave Macmillan.

Choo, S., & Mokhtarian, P. L. (2004). What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice. *Transportation Research Part A-Policy and Practice, 38*(3), 201-222.

Chu, Y.-L. (2002). Automobile Ownership Analysis using Ordered Probit Models. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 60-67.

Cowley, E., & Mitchell, A. A. (2003). The moderating effect of product knowledge on the learning and organization of product information. *Journal of Consumer Research, 30*(3), 443-454.

Daganzo, C. (1979). *Multinomial probit : the theory and its application to demand forecasting*. New York ; London: Academic Press.

Dagsvik, J. K., & Liu, G. (2009). A framework for analyzing rank-ordered data with application to automobile demand. *Transportation Research Part A-Policy and Practice, 43*(1), 1-12.

Dagsvik, J. K., Wennemo, T., Wetterwald, D. G., & Aaberge, R. (2002). Potential demand for alternative fuel vehicles. *Transportation Research Part B-Methodological, 36*(4), 361-384.

Daly, A., & Zachary, S. (1978). Improved multiple choice models. In D. A. Hensher & M. Q. Dalvi (Eds.), *Determinants of Mode Choice*. Sussex: Saxon House.

Dargay, J., & Gately, D. (1999). Income's effect on car and vehicle ownership, worldwide: 1960-2015. *Transportation Research Part A-Policy and Practice, 33*(2), 101-138.

Dargay, J., Gately, D., & Sommer, M. (2007). Vehicle ownership and income growth, worldwide: 1960-2030. *Energy Journal, 28*(4), 143-170.

Dargay, J. M. (2001). The effect of income on car ownership: evidence of asymmetry. *Transportation Research Part A-Policy and Practice, 35*(9), 807-821.

De Jong, G., Fox, J., Daly, A., Pieters, M., & Smit, R. (2004). Comparison of car ownership models. *Transport Reviews, 24*(4), 379-408.

Doctoroff, T. (2005). *Billions : selling to the new Chinese consumer* (1st ed.). Basingstoke: Palgrave Macmillan.

Eggers, F., & Eggers, F. (2011). Where have all the flowers gone? Forecasting green trends in the automobile industry with a choice-based conjoint adoption model. *Technological Forecasting and Social Change, 78*(1), 51-62.

Elrod, T., Louviere, J. J., & Davey, K. S. (1992). An Empirical-Comparison of Ratings-Based and Choice-Based Conjoint Models. *Journal of Marketing Research, 29*(3), 368-377.

Ewing, G., & Sarigollu, E. (1998). Car fuel-type choice under travel demand management and economic incentives. *Transportation Research Part D-Transport and Environment, 3*(6), 429-444.

Ewing, G., & Sarigollu, E. (2000). Assessing consumer preferences for clean-fuel vehicles: A discrete choice experiment. *Journal of Public Policy & Marketing, 19*(1), 106-118.

Fenger, J. (1999). Urban air quality. *Atmospheric Environment, 33*(29), 4877-4900.

Fildes, R. (1992). The Evaluation of Extrapolative Forecasting Methods. *International Journal of Forecasting, 8*(1), 81-98.

Fildes, R., Hibon, M., Makridakis, S., & Meade, N. (1998). Generalising about univariate forecasting methods: further empirical evidence. *International Journal of Forecasting, 14*(3), 339-358.

Fildes, R., Nikolopoulos, K., Crone, S. F., & Syntetos, A. A. (2008). Forecasting and operational research: a review. *Journal of the Operational Research Society, 59*(9), 1150-1172.

Franses, P. H., & Paap, R. (2001). *Quantitative models in marketing research.* Cambridge: Cambridge University Press.

Garner, J. (2005). *The rise of the Chinese consumer : theory and evidence*. Chichester: John Wiley & Sons.

Gatignon, H., Eliashberg, J., & Robertson, T. S. (1989). Modeling Multinational Diffusion Patterns - an Efficient Methodology. *Marketing Science, 8*(3), 231-247.

Giuliano, G., & Dargay, J. (2006). Car ownership, travel and land use: a comparison of the US and Great Britain. *Transportation Research Part A-Policy and Practice, 40*(2), 106-124.

Goering, P. A. (1985). Effects of Product Trial on Consumer Expectations, Demand, and Prices. *Journal of Consumer Research, 12*(1), 74-82.

Golob, T. F., Kitamura, R., Bradley, M., & Bunch, D. S. (1993). Predicting the Market Penetration of Electric and Clean-Fuel Vehicles. *The Science of The Total Environment, 134*(1-3), 371-381.

Goodwin, P., Dyussekeneva, K., & Meeran, S. (forthcoming). The use of analogies in forecasting the annual sales of new electronics products. *IMA Journal of Management Mathematics*.

Green, P. E., & Krieger, A. M. (1988). Choice rules and sensitivity analysis in conjoint simulators. *Journal of the Academy of Marketing Science, 16*(1), 114-127.

Green, P. E., & Krieger, A. M. (1996). Individualized hybrid models for conjoint analysis. *Management Science, 42*(6), 850-867.

Green, P. E., Krieger, A. M., & Wind, Y. (2001). Thirty years of conjoint analysis: Reflections and prospects. *Interfaces, 31*(3), S56-S73.

Green, P. E., & Rao, V. R. (1971). Conjoint Measurement for Quantifying Judgmental Data. *Journal of Marketing Research, 8*(3), 355-363.

Green, P. E., & Wind, Y. (1975). New Way to Measure Consumers Judgments. *Harvard Business Review, 53*(4), 107-117.

Greene, W. H. (2007). *NLOGIT 4.0*. New York: Econometric Software Inc.

Greene, W. H. (2009). Discrete Choice Modelling. In T. Mills & K. Patterson (Eds.), *Palgrave Handbook of Econometrics: Volume 2: Applied Econometrics* (pp. 473-556). London: Palgrave.

Gregg, J. V., Hossell, C. H., & Richardson, J. T. (1964). *Mathematical trend curves: an aid to forecasting*. Edinburgh: Oliver & Boyd.

Gupta, S., Jain, D. C., & Sawhney, M. S. (1999). Modeling the evolution of markets with indirect network externalities: An application to digital television. *Marketing Science, 18*(3), 396-416.

Gustafsson, A., Herrmann, A., & Huber, F. (2007). *Conjoint measurement : methods and applications* (4th ed. ed.). Berlin: Springer.

Guyon, H., & Petiot, J.-F. (2011). Market share predictions : A new model with rating-based conjoint analysis. *International Journal of Market Research, 53*(6), 831-857.

Hauser, J. R., & Rao, V. R. (2004). Conjoint Analysis, Related Modeling, and Applications. In Y. Wind & P. E. Green (Eds.), *Marketing Research and Modeling: Progress and Prospects: A Tribute to Paul E. Green* (pp. 141-168). Boston, London: Kluwer Academic Publishers.

Hausman, J., & Mcfadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica, 52*(5), 1219-1240.

Hausman, J. A., & Wise, D. A. (1978). Conditional Probit Model for Qualitative Choice - Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences. *Econometrica, 46*(2), 403-426.

Hensher, D. A., Barnard, P. O., Smith, N. C., & Milthorpe, F. W. (1989). An Empirical-Model of Household Automobile Holdings. *Applied Economics, 21*(1), 35-57.

Hensher, D. A., & Greene, W. H. (2002). Specification and estimation of the nested logit model: alternative normalisations. *Transportation Research Part B-Methodological, 36*(1), 1-17.

Hensher, D. A., & Greene, W. H. (2003). The Mixed Logit model: The state of practice. *Transportation, 30*(2), 133-176.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2005). *Applied choice analysis: a primer*. Cambridge; New York: Cambridge University Press.

Hess, D. B., & Ong, P. M. (2002). Traditional neighborhoods and auto ownership. *Transportation Research Record: Journal of the Transportation Research Board, 1805*, 35-44.

Hess, S. (2005). *Advanced discrete choice models with applications to transport demand.* Unpublished PhD thesis, Imperial College London, London.

Ho, D. Y.-f. (1976). On the Concept of Face. *American Journal of Sociology, 81*(4), 867-884.

Holweg, M., Luo, J., & Oliver, N. (2009). The past, present and future of China's automotive industry: a value chain perspective. *International Journal of Technological Learning, Innovation and Development, 2*(1-2), 76-118.

Huo, H., Zhang, Q. A., Wang, M. Q., Streets, D. G., & He, K. B. (2010). Environmental Implication of Electric Vehicles in China. *Environmental Science & Technology, 44*(13), 4856-4861.

International Energy Agency. (2009). *World Energy Outlook 2009.* Paris: International Energy Agency.

International Energy Agency. (2010a). *Key World Energy Statistics 2010.* Paris: International Energy Agency.

International Energy Agency. (2010b). *World Energy Outlook 2010.* Paris: International Energy Agency.

Islam, T., & Meade, N. (1997). The diffusion of successive generations of a technology: A more general model. *Technological Forecasting and Social Change, 56*(1), 49-60.

Jain, S. C. (1989). Standardization of International Marketing Strategy - Some Research Hypotheses. *Journal of Marketing, 53*(1), 70-79.

Jamieson, L. F., & Bass, F. M. (1989). Adjusting Stated Intention Measures to Predict Trial Purchase of New Products - a Comparison of Models and Methods. *Journal of Marketing Research, 26*(3), 336-345.

Johansson, J. K. (2009). *Global marketing : foreign entry, local marketing, & global management* (5th ed.). Boston: McGraw-Hill Irwin.

Jun, D. B., & Park, Y. S. (1999). A Choice-Based Diffusion Model for Multiple Generations of Products. *Technological Forecasting and Social Change, 61*(1), 45-58.

Katsikeas, C. S., Samiee, S., & Theodosiou, M. (2006). Strategy fit and performance consequences of international marketing standardization. *Strategic Management Journal, 27*(9), 867-890.

Kim, H. S., & Kim, E. (2004). Effects of Public Transit on Automobile Ownership and Use in Households of the USA. *Review of Urban & Regional Development Studies, 16*(3), 245-262.

Kim, W. J., Lee, J. D., & Kim, T. Y. (2005). Demand forecasting for multigenerational products combining discrete choice and dynamics of diffusion under technological trajectories. *Technological Forecasting and Social Change, 72*(7), 825-849.

Kim, Y., Jeong, G., Ahn, J., & Lee, J. (2007). Consumer preferences for alternative fuel vehicles in South Korea. *International Journal of Automotive Technology and Management, 7*(4), 327-342.

Kobos, P. H., Erickson, J. D., & Drennen, T. E. (2003). Scenario analysis of Chinese passenger vehicle growth. *Contemporary Economic Policy, 21*(2), 200-217.

Kumar, V., Nagpal, A., & Venkatesan, R. (2002). Forecasting category sales and market share for wireless telephone subscribers: a combined approach. *International Journal of Forecasting, 18*(4), 583-603.

Lakshmanan, A., & Krishnan, H. S. (2011). The Aha! Experience: Insight and Discontinuous Learning in Product Usage. *Journal of Marketing, 75*(6), 105-123.

Lave, C. A., & Train, K. (1979). A Disaggregate Model of Auto-Type Choice. *Transportation Research Part A-Policy and Practice, 13*(1), 1-9.

Lee, C. Y., Lee, J. D., & Kim, Y. B. (2008). Demand forecasting for new technology with a short history in a competitive environment: the case of the home networking market in South Korea. *Technological Forecasting and Social Change, 75*(1), 91-106.

Lee, J., & Cho, Y. (2009). Demand forecasting of diesel passenger car considering consumer preference and government regulation in South Korea. *Transportation Research Part A-Policy and Practice, 43*(4), 420-429.

Lee, J. S., Cho, Y. S., Lee, J. D., & Lee, C. Y. (2006). Forecasting future demand for large-screen television sets using conjoint analysis with diffusion model. *Technological Forecasting and Social Change, 73*(4), 362-376.

Li, J., Walker, J. L., Srinivasan, S., & Anderson, W. P. (2010). Modeling Private Car Ownership in China: Investigation of Urban Form Impact Across Megacities. *Transportation Research Record: Journal of the Transportation Research Board, 2193*, 76-84.

Lilien, G. L., & Yoon, E. S. (1990). The Timing of Competitive Market Entry - an Exploratory-Study of New Industrial-Products. *Management Science, 36*(5), 568-585.

Louviere, J. J. (1988). *Analyzing decision making : metric conjoint analysis*: Sage.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: analysis and application*. Cambridge, MA: Cambridge University Press.

Mahajan, V., & Muller, E. (1979). Innovation Diffusion and New Product Growth-Models in Marketing. *Journal of Marketing, 43*(4), 55-68.

Mahajan, V., & Muller, E. (1996). Timing, diffusion, and substitution of successive generations of technological innovations: The IBM mainframe case. *Technological Forecasting and Social Change, 51*(2), 109-132.

Mahajan, V., Muller, E., & Wind, Y. (2000). *New product diffusion models*. Boston, Mass. ; London: Kluwer Academic.

Maheswaran, D., Sternthal, B., & Gürhan, Z. (1996). Acquisition and Impact of Consumer Expertise. *Journal of Consumer Psychology, 5*(2), 115-133.

Mannering, F., & Mahmassani, H. (1985). Consumer Valuation of Foreign and Domestic Vehicle Attributes - Econometric-Analysis and Implications for Auto Demand. *Transportation Research Part A-Policy and Practice, 19*(3), 243-251.

Mannering, F., Winston, C., & Starkey, W. (2002). An exploratory analysis of automobile leasing by US households. *Journal of Urban Economics, 52*(1), 154-176.

Matas, A., & Raymond, J. L. (2008). Changes in the structure of car ownership in Spain. *Transportation Research Part A-Policy and Practice, 42*(1), 187-202.

Mau, P., Eyzaguirre, J., Jaccard, M., Collins-Dodd, C., & Tiedemann, K. (2008). The 'neighbor effect': Simulating dynamics in consumer preferences for new vehicle technologies. *Ecological Economics, 68*(1-2), 504-516.

McCarthy, P. S. (1996). Market price and income elasticities of new vehicle demands. *Review of Economics and Statistics, 78*(3), 543-547.

McCarthy, P. S., & Tay, R. S. (1998). New vehicle consumption and fuel efficiency: A nested logit approach. *Transportation Research Part E-Logistics and Transportation Review, 34*(1), 39-51.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarmebka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.

McFadden, D. (1978). Modeling the choice of residential location. In A. Karlquist, L. Lundqvist, F. Snickars & J. Weibull (Eds.), *Spatial Interaction Theory and Planning Models* (pp. 75-96). Amsterdam: North-Holland.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics, 15*(5), 447-470.

Meade, N., & Islam, T. (1995). Forecasting with growth curves: An empirical comparison. *International Journal of Forecasting, 11*(2), 199-215.

Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation - A 25-year review. *International Journal of Forecasting, 22*(3), 519-545.

Mitchell, A. A., & Dacin, P. A. (1996). The assessment of alternative measures of consumer expertise. *Journal of Consumer Research, 23*(3), 219-239.

Mohammadian, A., & Miller, E. J. (2003). Empirical Investigation of Household Vehicle Type Choice Decisions. *Transportation Research Record: Journal of the Transportation Research Board, 1854*, 99-106.

Moorman, C., Diehl, K., Brinberg, D., & Kidwell, B. (2004). Subjective knowledge, search locations, and consumer choice. *Journal of Consumer Research, 31*(3), 673-680.

Moreau, C. P., Lehmann, D. R., & Markman, A. B. (2001). Entrenched Knowledge Structures and Consumer Response to New Products. *Journal of Marketing Research, 38*(1), 14-29.

Morwitz, V. G. (2001). Methods for forecasting from intentions data. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners* (pp. 33-56). Boston: Kluwer.

Morwitz, V. G., Steckel, J. H., & Gupta, A. (2007). When do purchase intentions predict sales? *International Journal of Forecasting, 23*(3), 347-364.

National Bureau of Statistics of China. (2003-2010). *China Statistical Yearbooks (various issues)*. Beijing: China Statistics Press.

National Bureau of Statistics of China. (2010). *China Statistical Yearbook 2010.* Beijing: China Statistics Press.

Nolan, A. (2002). *The Determinants of Urban Households' Transport Decisions: A Microeconometric Study using Irish Data*. Paper presented at the Royal Economic Society Annual Conference 2002, Warwick University, UK.

Nolan, A. (2010). A Dynamic Analysis of Household Car Ownership. *Transportation Research Part A-Policy and Practice, 2010*(44), 446-455.

Norton, J. A., & Bass, F. M. (1987). A Diffusion-Theory Model of Adoption and Substitution for Successive Generations of High-Technology Products. *Management Science, 33*(9), 1069-1086.

O'Neill, J., & Stupnytska, A. (2009). The Long-Term Outlook for the BRICs and N-11 Post Crisis. Global Economics Paper No: 192. New York: Goldman Sachs.

Orbach, Y., & Fruchter, G. E. (2011). Forecasting sales and product evolution: The case of the hybrid/electric car. *Technological Forecasting and Social Change, 78*(7), 1210-1226.

Ozan, E., Sireli, Y., & Kauffmann, P. (2007). A new market adoption model for the information systems industry. *Engineering Management Journal, 19*(1), 13-21.

Park, C. W., Mothersbaugh, D. L., & Feick, L. (1994). Consumer Knowledge Assessment. *Journal of Consumer Research, 21*(1), 71-82.

Peracchio, L. A., & Tybout, A. M. (1996). The moderating role of prior knowledge in schema-based product evaluation. *Journal of Consumer Research, 23*(3), 177-192.

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

Potoglou, D., & Kanaroglou, P. S. (2008a). Disaggregate Demand Analyses for Conventional and Alternative Fueled Automobiles: A Review. *International Journal of Sustainable Transportation, 2*(4), 234-259.

Potoglou, D., & Kanaroglou, P. S. (2008b). Modelling car ownership in urban areas: a case study of Hamilton, Canada. *Journal of Transport Geography, 16*(1), 42-54.

Potoglou, D., & Susilo, Y. O. (2008). Comparison of Vehicle-Ownership Models. *Transportation Research Record: Journal of the Transportation Research Board, 2076*, 97-105.

Putsis, W. P., & Srinivasan, V. (2000). Estimation Techniques for Macro Diffusion Models. In V. Mahajan, E. Muller & Y. Wind (Eds.), *New-Product Diffusion Models* (pp. 263-291). Dordrecht: Kluwer Academic Publishers.

Qian, L., & Soopramanien, D. (2011). Heterogeneous consumer preferences for alternative fuel cars in China. *Transportation Research Part D: Transport and Environment, 16*(8), 607-613.

Raghavarao, D., Wiley, J. B., & Chitturi, P. (2011). *Choice-based conjoint analysis : models and designs*. Boca Raton ; London: CRC Press.

Rao, A. R., & Monroe, K. B. (1988). The Moderating Effect of Prior Knowledge on Cue Utilization in Product Evaluations. *Journal of Consumer Research, 15*(2), 253-264.

Rao, A. R., & Sieben, W. A. (1992). The Effect of Prior Knowledge on Price Acceptability and the Type of Information Examined. *Journal of Consumer Research, 19*(2), 256-270.

Robertson, A., Soopramanien, D., & Fildes, R. (2007). Segmental new-product diffusion of residential broadband services. *Telecommunications Policy, 31*(5), 265-275.

Rogers, G., & Soopramanien, D. (2009). The truth is out there! How external validity can lead to better marketing decisions. *International Journal of Market Research, 51*(2), 163-180.

Ryan, J. M., & Han, G. (1999). Vehicle-ownership model using family structure and accessibility application to Honolulu, Hawaii. *Transportation Research Record: Journal of the Transportation Research Board, 1676*, 1-10.

Ryans, J. K., Griffith, D. A., & White, D. S. (2003). Standardization/adaptation of international marketing strategy - Necessary conditions for the advancement of knowledge. *International Marketing Review, 20*(6), 588-603.

Schmittlein, D. C., & Mahajan, V. (1982). Maximum Likelihood Estimation for an Innovation Diffusion Model of New Product Acceptance. *Marketing Science, 1*(1), 57-78.

Sheth, J. N. (2011). Impact of Emerging Markets on Marketing: Rethinking Existing Perspectives and Practices. *Journal of Marketing, 75*(4), 166-182.

Srinivasan, V., & Mason, C. H. (1986). Nonlinear Least Squares Estimation of New Product Diffusion Models. *Marketing Science, 5*(2), 169-178.

Sun, B. H., & Morwitz, V. G. (2010). Stated intentions and purchase behavior: A unified model. *International Journal of Research in Marketing, 27*(4), 356-366.

Sundqvist, S., Frank, L., & Puumalainen, K. (2005). The effects of country characteristics, cultural similarity and adoption timing on the diffusion of wireless communications. *Journal of Business Research, 58*(1), 107-110.

Szymanski, D. M., Bharadwaj, S. G., & Varadarajan, P. R. (1993). Standardization Versus Adaptation of International Marketing Strategy - an Empirical-Investigation. *Journal of Marketing, 57*(4), 1-17.

Takada, H., & Jain, D. (1991). Cross-National Analysis of Diffusion of Consumer Durable Goods in Pacific Rim Countries. *Journal of Marketing, 55*(2), 48-54.

Talukdar, D., Sudhir, K., & Ainslie, A. (2002). Investigating new product diffusion across products and countries. *Marketing Science, 21*(1), 97-114.

Tanner, J. C. (1958). An Analysis of Increases in Motor Vehicles in Great Britain and the United States. Research Note RN/3341. Harmondsworth: Road Research Laboratory.

Tanner, J. C. (1975). Forecasts of vehicles and traffic in Great Britain: 1974 revision. Department of Environment, TRRL Report LR650. Crowthorne: Transport and Road Research Laboratory.

Tashman, J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting, 16*(4), 437-450.

Tellefsen, T., & Takada, H. (1999). The relationship between mass media availability and the multicountry diffusion of consumer products. *Journal of International Marketing, 7*(1), 77-96.

The International Organization of Motor Vehicle Manufacturers. (2009). *World Motor Vehicle Production by Country and Type: 2007-2008 Passenger Cars*. Paris: The International Organization of Motor Vehicle Manufacturers (OICA).

Theodosiou, M., & Leonidou, L. C. (2003). Standardization versus adaptation of international marketing strategy: an integrative assessment of the empirical research. *International Business Review, 12*(2), 141-171.

Tian, K., & Dong, L. (2011). *Consumer-citizens of China : the role of foreign brands in the imagined future China*. London ; New York: Routledge.

Train, K., & Wilson, W. W. (2008). Estimation on stated-preference experiments constructed from revealed-preference choices. *Transportation Research Part B-Methodological, 42*(3), 191-203.

Train, K. E. (2003). *Discrete choice methods with simulation.* Cambridge, MA.: Cambridge University Press.

Train, K. E., & Winston, C. (2007). Vehicle choice behavior and the declining market share of US automakers. *International Economic Review, 48*(4), 1469-1496.

Tsai, B. H., Li, Y. M., & Lee, G. H. (2010). Forecasting global adoption of crystal display televisions with modified product diffusion model. *Computers & Industrial Engineering, 58*(4), 553-562.

U.S. Energy Information Administration. (2010). *Annual Energy Outlook 2010.* Washington D.C.: U.S. Department of Energy.

U.S. Energy Information Administration. (2011). *Annual Energy Review 2010.* Washtington D.C.: U.S. Department of Energy.

Urban, G. L., Hauser, J. R., & Roberts, J. H. (1990). Prelaunch Forecasting of New Automobiles. *Management Science, 36*(4), 401-421.

Van den Bulte, C., & Lilien, G. L. (1997). Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science, 16*(4), 338-353.

Van den Bulte, C., & Stremersch, S. (2004). Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Science, 23*(4), 530-544.

Van Ittersum, K., & Feinberg, F. M. (2010). Cumulative Timed Intent: A New Predictive Tool for Technology Adoption. *Journal of Marketing Research, 47*(5), 808-822.

Vasconcellos, E. A. (1997). The demand for cars in developing countries. *Transportation Research Part a-Policy and Practice, 31*(3), 245-258.

Wang, M., Huo, H., Johnson, L., & He, D. (2006). Projection of Chinese Motor Vehicle Growth, Oil Demand, and CO_2 Emissions through 2050. Research Paper ANL/EDS/06-6. Chicago: Argonne National Laboratory.

Whelan, G. (2007). Modelling car ownership in Great Britain. *Transportation Research Part A-Policy and Practice, 41*(3), 205-219.

Williams, H. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A, 9*(3), 285-344.

Wind, Y. (1981). A Framework for Classifying New-Product Forecasting Methods. In Y. Wind, V. Mahajan & R. N. Cardozo (Eds.), *New-Product Forecasting: Models and Applications* (pp. 3-42). Lexington: LexingtonBooks.

Wind, Y., & Green, P. E. (2004). *Marketing research and modeling : progress and prospects : a tribute to Paul E. Green*. Boston ; London: Kluwer Academic Publishers.

Wind, Y., Mahajan, V., & Cardozo, R. N. (1981). *New-product forecasting : models and applications*. Lexington: Lexington Books.

Wittink, D. R., & Bergestuen, T. (2001). Forecasting with conjoint analysis. In J. S. Armstrong (Ed.), *Principles of Forecasting: A Handbook of Researchers and Practitioners* (pp. 147-167). New York: Kluwer Academic Publishers.

Wittink, D. R., & Cattin, P. (1989). Commercial Use of Conjoint-Analysis - an Update. *Journal of Marketing, 53*(3), 91-96.

Wood, S. L., & Lynch, J. G. (2002). Prior knowledge and complacency in new product learning. *Journal of Consumer Research, 29*(3), 416-426.

Wu, F. S., & Chu, W. L. (2010). Diffusion models of mobile telephony. *Journal of Business Research, 63*(5), 497-501.

Yalcinkaya, G. (2008). A culture-based approach to understanding the adoption and diffusion of new products across countries. *International Marketing Review, 25*(2), 202-214.

Zhang, X. A., Grigoriou, N., & Li, L. (2008). The myth of China as a single market - The influence of personal value differences on buying decisions. *International Journal of Market Research, 50*(3), 377-402.

Zou, S., Andrus, D. M., & Norvell, D. W. (1997). Standardization of international marketing strategy by firms from a developing country. *International Marketing Review, 14*(2), 107-123.