Motorization, Vehicle Purchase and Use Behavior in China: A Shanghai Survey

By

NI, MENG-CHENG (JASON)

B.S. (National Taiwan University) 1998

M.S. (University of California, Berkeley) 2003

M.C.P. (University of California, Berkeley) 2003

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

TRANSPORTATION TECHNOLOGY AND POLICY

in the

OFFICE OF GRADUATE STUDIES

of the

UNICERSITY OF CALIFORNIA, DAVIS

Approved

Professor Daniel Sperling, Co-chair

Dr. Ken Kurani, Co-chair

Professor Patricia Mohktarian

Dr. Tom Turrentine

Committee in Charge

2008

中國之機動化暨交通工具購買及使用行為：
上海調研及案例分析

ABSTRACT

Motorization is the transition from non-motorized travel means (e.g., walk) to motorized travel means (e.g., car). China, as the most populous country in the world, has started the motorization process, and its results will have huge impact on the whole world in terms of transportation, energy, environment and automobile market.

At the national level, motorization is usually measured by the growth of auto ownership, and average income (GDP per capita) is usually considered the major driving force. However, this dissertation is focused on studying motorization at the individual or household. To do so, a pilot survey of 122 residents of Shanghai was conducted in late 2005, and a final survey of 1,037 people was conducted in mid-2006.

Research methodology, motorization pathway, and vehicle purchase (and use) behavior are three topics of this dissertation. Basically, this dissertation attempts to answer the following three questions: How to conduct survey research in China? What are motorization pathways in China? What is the vehicle purchase and use behavior in China? Regarding the first, trust from respondents is an important factor affecting people's motivation to participate in the study. A short, straight-forward questionnaire and a team speaking the local dialect help to facilitate survey research. In terms of motorization pathways, the survey shows that motorization pathways in Shanghai are diverse, complicated (multi-staged), and as one would expect at this point, do not include cars for many households and individuals. About half of the respondents don't simply follow my hypothetical motorization direction. In terms of the purchase and use behavior, variables

such as gender, perceptions of different aspects of the utility of different travel means, as well as personal or household income are significant. Purchasing a car may be considered a "family decision" as it is positively associated with household income; however, weekday car use seems to be a more personal choice as it is positively associated with personal income.

Last, although Shanghai itself can not represent the whole China, the results (motorization pathways; choice models) of the Shanghai study may be generalizable to other Chinese cities experiencing rapid economic growth and with various transportation alternatives.

(350 words)

KEY WORDS:

Motorization, China, Shanghai, Survey, Pathway, Choice Model

ACKNOWLEDGEMENTS

This is the most challenging “chapter” of the whole dissertation, because I have to thank so many people. First, I would like to thank Professor Dan Sperling (my dissertation co-chair) for giving me this opportunity to involve in such interesting and important topic. I would also like to express my gratitude to Dr. Ken Kurani, my dissertation co-chair, for working with me in the past several years. From Ken, I learned not only the “knowledge” but also “how to develop the knowledge”. Professor Patricia Mokhtarian is the third person I would like to specially thank to. Again, not only the knowledge learned from the class, but also her serious attitude on doing research impresses me very much. Also, I want to thank Dr. Tom Turrentine for sharing me so many great ideas in my proposal writing phase. There is an old Chinese saying: “一日為師，終身為父”. It means that a teacher sometimes can become almost like people’s father/mother, because he/she can teach people not only the book, but also the *life*. This Chinese saying is the best example for my five-year learning and working experiences in the Institute of Transportation Studies at UC Davis (ITS-Davis). Other than professors, many people at the ITS-Davis also play important role to make the China project happened, including Joe Krovoza, Ernie Hoftyzer, and especially, Mr. Yunshi Wang, Director of China Center of Energy and Transportation at ITS-Davis.

December 2004 was the start of my China *adventure*. From 2004 to 2006, I spent about one and half year in China, including Shanghai and Beijing. During my stay in China, I would like to thank Professor Chen XiaoHong (School of Transportation Engineering, Tongji University) and Professor Ma Jun (School of Automobile, Tongji University) for

their comments and student helpers (Joyce Fu, Frances Fan and Huang Rui, etc.) In addition to Tongji University, Ford China (Mr. Mei-Wei Chen, Mr. Keith Davey) also gave me lots of local help during my survey implementation. Also, I would like to thank Professor Jason Chang (National Taiwan University) for sharing me his China experiences and help me familiarized with Shanghai in a short time. The last, I need to say *thank you* to Dr. Jonathan Weinert, my classmate and the one (and only one) who went to China with me at the very beginning (December, 2004).

After returning from China to US, the data analysis and dissertation writing was a long process. I would like to thank many of my friends in China and US for the kindly advice and help, especially Ms. Raymos Au, who spend lots of time helping me in the data cleaning phase.

When I prepare this acknowledgement, a lot of good memories emerge. I will never forget how much help I received and will always be grateful. The last, I would like to dedicate this work to my mother, who supported and encouraged me from my first study abroad (from Taiwan to USA in 2000) to my second study abroad (from USA to China in 2004)!

博士論文的撰寫，對我而言，不但是知識求取的過程，更是自我的探索與實現。

本人在此，對所有曾幫助過我，或曾參與其中的人，再次表達由衷的感謝！

倪 孟 正

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER 1: INTRODUCTION

Motorization: Definition and World Studies

This dissertation starts by defining the term "motorization" as the changing of transportation from non-motorized means to motorized means. Others, such as Schipper et al. (2004) discuss "motorization" and Cherry et al. (2007) discuss "motorization pathway", but a discussion of the definition of motorization is usually missing or insufficient. The motorized means are transportation tools powered by fuel (gasoline, electricity, LPG, etc.) instead of human effort. At the individual level, motorization can be understood as a substitution from walk and bicycle to motorcycle and automobile, complicated by collective motorized modes such as taxi, bus, and transit. At the national level, motorization is usually reflected by the growth of the automobile fleet or per capita auto ownership. Instead of a single event, motorization usually takes place as a series of transitions over time, and these transitions at an individual, household, business enterprise, or other micro decision-making units are that unit's "motorization pathway." In this dissertation, I take the larger, and usually national, collection of motorization pathways to be motorization.

In human history, motorization plays an important role, as it is not merely about the substitution of automobiles and trucks for non-motorized travel modes; it further affects people's travel pattern, work and housing location choice at the individual level and brings in impacts on transportation, energy and environment at the world or national level.

Since the 1960s, many studies have been conducted to explain motorization in different countries around the world. *"At the national level, income alone typically explains more than 90 percent of the variation in motorization levels, and at the urban level more than 80 percent. The growth of national motor vehicle fleets parallels that of income: a 1 percent increase in income is associated with a 1 percent increase in motor vehicles, and this relationship has been relatively stable for the past 30 years"* (U.S. National Research Council and Chinese Academy of Engineering, 2003). Clearly, income is correlated with motorization at the aggregate national level. Figure 1 shows average income (measured as per capita GDP) versus motor vehicle ownership in 50 selected countries. Seven countries are labeled (including China) with the GDP per capita transformed to 1995 US dollars. For each country in Figure 1, a line segment connects the country's position in 1970 with its position in 1996. Clearly, for many countries, the slope (compare 1970 with 1996) between income and motor vehicle ownership (vehicle/1000 people) is close to 45 degrees. Focusing on similar measures in seven Asian cities in Figure 2, we see many cities are exceptions to the national scale generalization, for example, Bangkok and Jakarta. Therefore, I believe an "income-auto ownership" linear relationship will be further challenged at a more disaggregate level, e.g., individual, even if it holds at the aggregate level, e.g., nation. The main effort of this dissertation is to describe motorization pathways and vehicle purchase and use behavior at the individual level in Shanghai, China.

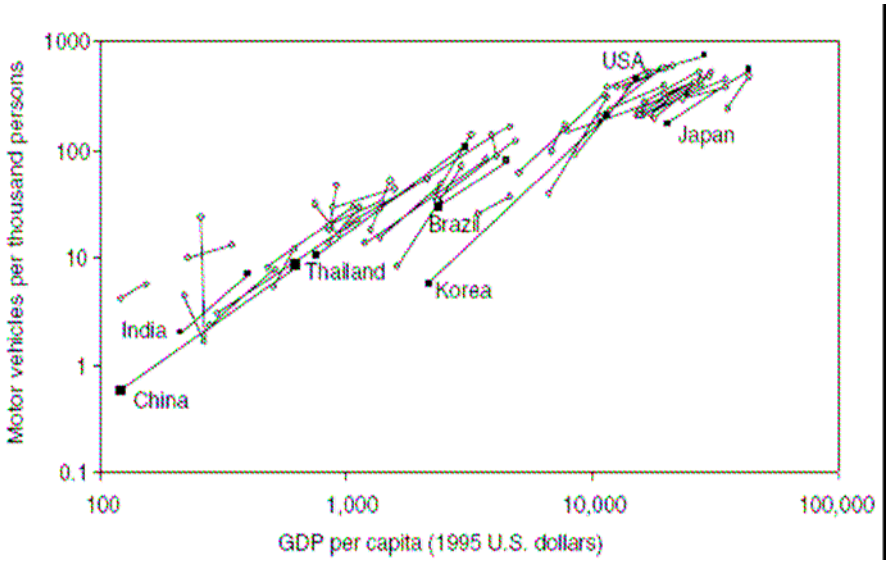


Figure 1: Income and Motor Vehicle Ownership (1970 vs. 1996)

Sources: U.S. National Research Council and Chinese Academy of Engineering, *Personal Cars and China* (Washington, DC: National Academies Press, 2003); Motorization data: International Road Federation (2001 and earlier); Other data: World Bank (2001 and earlier)

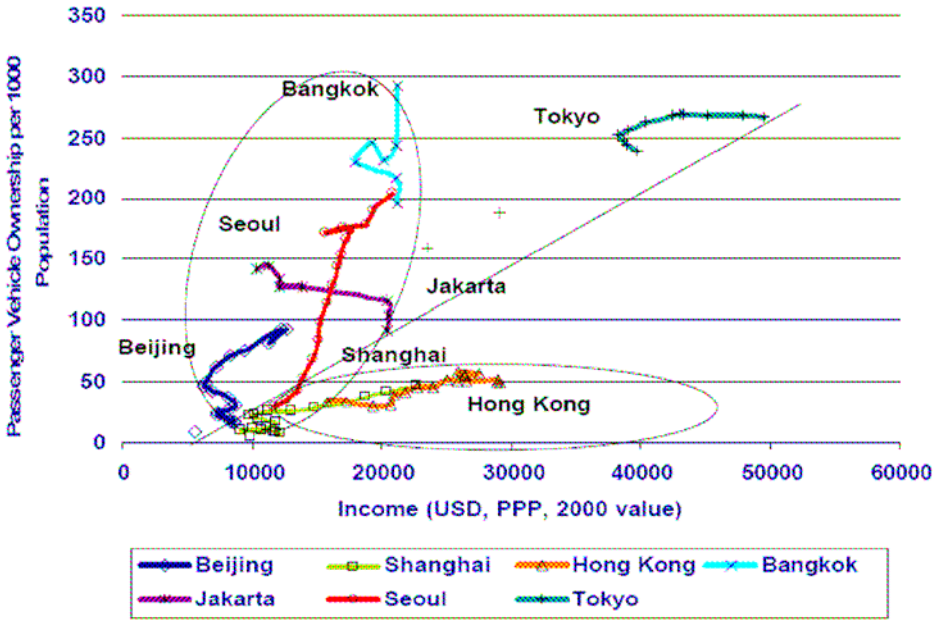


Figure 2: Income and Passenger Vehicle Ownership in Seven Asian Cities (1980 – 2002)

Source: Naoko Doi, APERC Database, 2005

Motorization in China

China, with a population of 1.3 billion people, is starting its motorization process. Although the level of motorization is still low compared to many other countries, China's growth of motorization is startlingly rapid – with local and regional implications for personal mobility, urban development, land use, and air quality and global implications for energy use and climate change. China had a 300 percent growth in auto ownership from 1999 to 2002 (National Bureau of Statistics of China: 1999, 2002). In 2001, the motor vehicle fleet size in China reached 18 million motor vehicles[1] (National Bureau of Statistics of China, 2001) versus 220 million in the US. However, from 1980 to 2002, the total number of motor vehicles in China grew ten-fold compared to a one-third increase in the US. To examine the "income vs. motorization" relationship in China, I notice that, from 1990 to 2000, the growth rate of automobiles (2.9 times) is slightly faster than the growth rate of GDP (2.6 times) as shown in Table 1 (Shen et al., 2002).

	1990	**2000**	**2000/1990**
GDP	100	262.3	2.6
Automobile (10,000 vehicles)	551.4	1,609	2.9
Bicycles/Others (10,000 vehicles)	421.3	3,772	9.0

Table 1: Vehicle (motorized and non-motorized) Growth in China (1990 – 2000)

Source: Shen et al., 2002

[1] Motor vehicles (or cars) include: private car, truck, and bus.

The fleet of non-motorized vehicles appears to have grown even faster than the fleet of motorized vehicles. We note in Table 1 that the number of "bicycles/others" grew 9-fold, from 4.21 million to 37.72 million over the 1990 to 2000. However, the true size of the bicycle population is estimated as over 50 million in 2000 (Figure 3) (Shen et al., 2002).

According to a recent study (Weinert, 2007), bicycles still remain the dominant two-wheeled vehicle in Chinese cities, mainly due to low income, high population density (and thus short trips), and extensive bicycle infrastructure (e.g., lanes, parking).

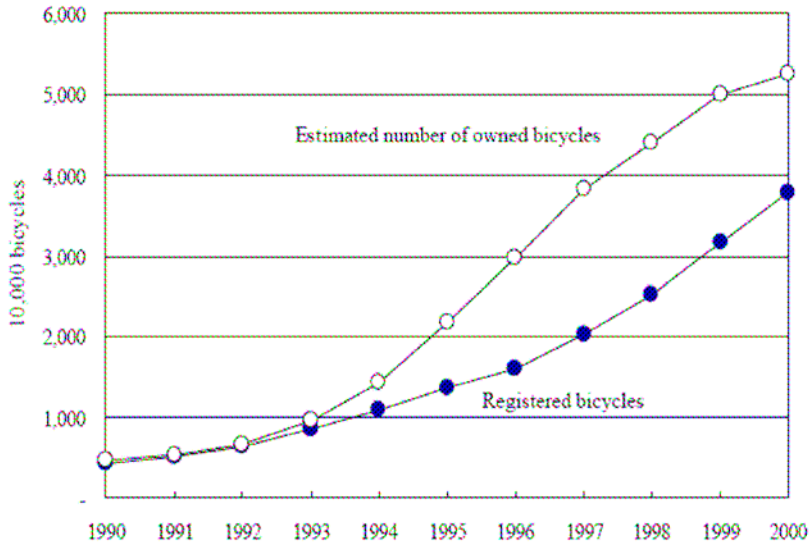


Figure 3: Estimated Bicycle Ownership in China

Source: Shen et al., 2002

In addition to the automobile and bicycle, the motorized two-wheeler (powered by electricity or LPG) and motorcycle are two types of personal mobility options for Chinese people. Take motorized two-wheeler for example, the shift from bicycles to electric bicycle (E-bike, a type of electric two-wheeler) also occurs at rapid pace

throughout China, especially in large cities. E-bike sales reached 10 million/year in 2005 as more bike and public transit users shifted to this mode (Weinert, 2007). Similar study indicated that the sales of the electric powered vehicle exceeded 16 million/year in 2006 (Cherry, 2007). According to one study (Weinert, 2007), over 30 million electric two-wheelers are now in regular use on Chinese streets, and the electric two-wheelers in China are called: "the world's most successful electric-drive vehicle".

Motorization also affects the auto industry of China. In 2002, China underwent a car-buying craze with more than 50 percent sales growth. In 2003, the Chinese automobile market grew by about 20 percent to reach a value of $23.9 billion and a volume of 1,528,200 units (Datamonitor, 2004). In 2004, data showed that the China automotive market still had the fastest growth rate in the world (PaoHua Economic Research Institute, 2005).

Several factors facilitated China's car buying craze. One major factor is the manufacturer's price-cut due to the cut-throat competition and the over-estimation of the China market. Another factor is that the Agriculture Bank of China offered 10 billion RMB (1.2 billion USD) in loan for car buyers in 2002. Although the car loan system is still at an initial phase, the availability of loans is expected to be a driving force for car purchases by many Chinese people with low but rising income. Moreover, the over 50 percent reduction of tariffs on imported cars after China's accession to WTO also makes the car purchase much easier than before. On January 1, 2006, the quota control of imported cars was abandoned, and the tariff on imported cars was further lowered to 25 percent on July 1, 2006. In addition, another important reason for the car buying craze is

the rapid increase of personal income (GDP/capita), which is usually considered a prerequisite of car purchase. As predicted in a study about the Chinese middle class (Farrell et al., 2007), the emergence of a Chinese "lower middle-class" (with annual household income of 25,001 – 40,000 RMB) will happen around 2010, and the "upper middle-class" (with annual household income of 40,001 – 100,000 RMB) will emerge around 2020. The increase of purchasing power suggests that more and more Chinese will be able to enjoy personal motorized transport and the "third consumer revolution" (cars) after bicycles and electronic goods is going to begin (*The Economist*, 2003).

To meet rising domestic demand, joint-ventures between foreign automobile makers and domestic automakers[2] dominate the Chinese auto industry, accounting for 97 percent of the entire China market (*China Automotive Industry Yearbook*, 2000). Foreign companies bring the car as not only a commercial good but also a "culture" into China. Thus, for many Chinese people, owning a car symbolizes a step toward a western modern life. "*Development has brought more interaction with the developed world and its culture, giving China's population greater exposure to Western values, ideas, and lifestyles*" (Gould, 2000).

Definition of a "Car" in China

Due to the reforms and "open-door policies" of 1978, China gradually made a transition from a state-planned economy to a market economy. According to the National Bureau of Statistics of China, from 1978 to 2004, the Chinese economy grew at an annual average

[2] Under current policies, foreign companies have to find local partners in order to enter the China market.

rate of 9.4 percent. China's spectacular economic growth over the past two decades, together with its population growth (especially the increase in the urban population), has resulted in high rates of industrialization, urban development, and most importantly – motorization (Shipper et al., 2004).

To describe the phenomenon, studies have been conducted to compare China's level of motorization with other countries in the world. For example: in 2002, the total number of motorized vehicles of all kinds was less than 100 per 1,000 people versus about 700 per 1,000 people in Europe and 900 per 1,000 people in the United States (Schipper et al., 2004). In addition, as indicated in the same study, China had approximately the same number of "cars" per capita in 2003 as the United States in 1910 (Shipper et al., 2004). However, such comparisons will only make sense if the definitions of "car" are the same in both Chinese and Western contexts.

The definition of a "car" (*qi che*) in official Chinese statistics is sometimes misleading, as it means "motor vehicles" in the English context. (Shipper et al., 2004) That is, the so-called "car" in China covers a wide variety of "motor vehicles" such as taxis, buses, vans, minibuses, trucks, as well as automobiles. As we can notice, the automobile, which is usually considered as a "car" in Western context, is merely a sub-category of the China-defined "car". Nevertheless, there are at least two major national definitions of a car in China, as in the following:

[Definition 1: GB/T3730.1-2001 Standard]

According to the Chinese national standard (GB/T3730.1-2001) of the General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, the definition of a car is: *"motor vehicle, non-railed, with 4 or more wheels."*; *"... used for carrying or trailing passengers and goods, or for other special purposes"*. It has two basic categories – passenger car and commercial vehicle. Table 2 describes in detail the "car" under GB/T3730.1-2001 Standard.

[Definition 2: China Statistical Yearbook]

GB/T3730.1-2001 Standard has two limitations for my research. First, it is a categorization based on vehicle function (not ownership). Second, this standard is so new few data have been collected or reported using this definition. Therefore, I refer to the *China Statistical Year book* which defines cars as "personal cars" in two basic categories, private and commercial. This is a categorization (Table 3) based on vehicle ownership, and it is more relevant to the purchase behavior theme of this study. In a nutshell, this dissertation will focus on the individual-level purchase behavior of the "private car" under the definition of the China Statistical Yearbook.

"CAR" in China (or motor vehicle)	
Passenger Car (less than 9 seating capacity)	**Commercial Vehicle**
1. Saloon (sedan) 2. Convertible saloon 3. Pullman saloon (pullman sedan、executive limousine) 4. Coupe 5. Convertible (open tourer、roadster、spider) 6. Hatchback (1~6 are usually called "Sedan" in China) 7. Station wagon 8. Multipurpose passenger car 9. Forward control passenger car 10. Off-road passenger car 11. Special purpose passenger car 12. Motor caravan 13. Armoured passenger car 14. Ambulance 15. Hearse	BUS（more than 9 seats） 1. minibus 2. City-bus 3. Interurban coach 4. Touring coach 5. Articulated bus 6. Trolley bus 7. Off-road bus 8. Special bus 9. Semi-Trailer Towing Vehicle GOODS VEHICLE 10. General purpose goods vehicle 11. Multipurpose goods vehicle 12. Trailer towing vehicle 13. Off-road goods vehicle 14. Special goods vehicle 15. Specialized goods vehicle

Table 2: China Car Category (GB/T3730.1-2001)

Source: GB/T3730.1-2001 Standard, General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China

"CAR" in China (or personal car)	
Private Car [definition of study subject in this dissertation]	**Commercial (non-private) Car**
1. Passenger Vehicle 2. Goods Vehicle	1. Passenger Vehicle 2. Goods Vehicle

Table 3: China Car Category (China Statistical Yearbook)

Source: *China Statistical Yearbook*

It seems that growth in the number of private cars is driving the growth in the number of "cars" in China. *"At one time, most of the vehicles on Chinese roads were commercial vehicles, accounting for 85% of the total. This composition changed little even in 1990, or ten years after reforms & opening-up began being implemented. It did, however, change drastically due to the emergence of private cars in the last decade, which caused the share of commercial vehicles to decline to 60 percent by 2000"* (Shen et al., 2002, p. 3) (Figure 4). By further examining GDP growth versus car penetration, I find that the 2.6-fold GDP growth (1990 – 2000) brought a 2.1-fold increase in commercial cars, while the private car grew 7.7-fold. Noticeably, private passenger car had an even higher growth – 15.2 times (Table 4).

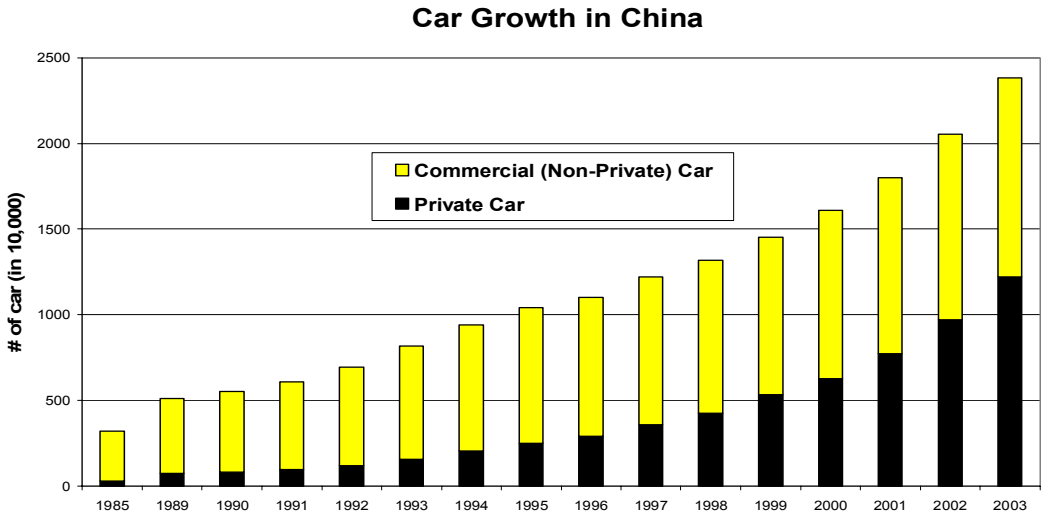


Figure 4: Private *vs.* Non-Private Car in China (1985 – 2003)

Source: *China Statistical Yearbook*, 1985 – 2003

	1990	2000	2000/1990	Annual Growth (%)
GDP (1990 = 100)	100	262.3	**2.6**	10.0
Total number of personal cars*	551.4	1,608.9	2.9	11.2
1. Commercial cars	469.8	983.6	**2.1**	7.7
2. Private cars	81.6	625.3	**7.7**	22.6
Private passenger cars	24.1	365.1	**15.2**	31.3

* In 10000 of cars

Table 4: Car Penetration and GDP growth in China (1990 – 2000)

Source: *China Statistical Yearbook*, 1990 and 2000

Shanghai: City of Wealth, Diversity and Car Debate

Established more than 700 years ago at the tip of the Changjiang River Delta on the East China Sea (Figure 5), Shanghai is the commercial hub and one of the most populous

cities in China. Growing from 5.7 million people in 1950[3], the number of official registered residents in Shanghai was 17.4 million in 2005. In addition to the registered residents, there are more than 5 million un-registered people as the "floating population" (Shanghai Bureau of Statistics, 2006).

Figure 5: Geographical Location of Shanghai in China
Source: Wikipedia

The Shanghai metropolitan area has a total area of 6,340.5 km^2, of which the city's suburbs cover 6,000 km^2. These suburbs are estimated to be the home of 6.8 million people. (Shanghai Bureau of Statistics, 2006) That is, 70 percent of the population is clustered in 5 percent of the urban area. There are currently 18 districts plus one county in Shanghai, as in Figure 6: the "suburbs" refer to Chongming county; eight districts (Nanhui, Fengxian, Baoshan, Minghang, Jiading, Jinshan, Songjiang, Qingpu) and the Pudong New Area. The "urban" area includes the remaining nine districts.

[3] Shanghai's rapid population growth in the 1950's was due to natural increase and unregulated in-migration from the outskirts of the city.

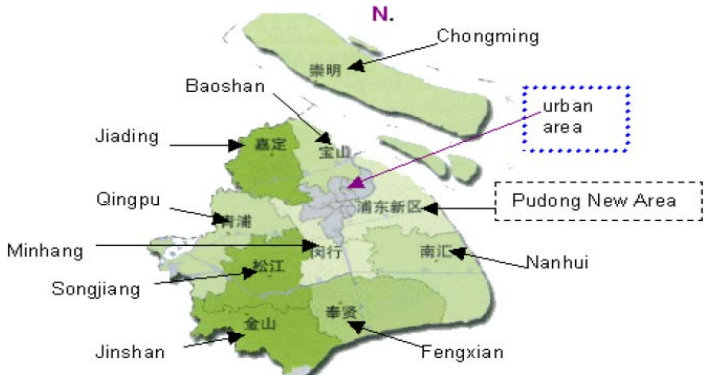


Figure 6: City Map of Shanghai (Urban vs. Suburb)

Source: Wikipedia

The average population density of Shanghai is 2,116 (people/km^2), however, the density distribution is uneven – from as high as 49,854 (people/km^2) in Huangpu district (urban area) to as low as 610 (people/km^2) in Chongming county (suburb area) (Shanghai Bureau of Statistics, 2004). In light of the high population density of the urban districts, current planning policy seeks to decentralize Shanghai by building seven satellite suburbs. As a result, massive new public transportation investments planned for the next two decades are aimed at lowering Shanghai's extremely high population density, supporting economic growth, and enhancing the quality of life (U.S. National Research Council and Chinese Academy of Engineering, 2003). In 2005, Shanghai had more than one thousand bus lines and the Shanghai Metro (subways and light rail) had four lines. According to the development plan of the city government, by 2010, another eight subway lines will be built.

In spite of the development of the public transportation system, from 1990 to 2000, the number of motor vehicles in Shanghai increased 2.5 times. At the end of 2006, the total number of automobiles in Shanghai reached 1.1 million plus 1.02 million motorcycles

and 0.27 million gas auxiliary-powered vehicles (Department of Traffic Police, City of Shanghai, 2006). As a consequence, the city, especially the urban area, has suffered from serious traffic congestion and air pollution problem.

Based on the above, an investigation of vehicle purchase behavior is timely and important for Shanghai. Generally speaking, the reasons we selected Shanghai for our case study are that it is experiencing high economic growth and rapid motorization, yet is attempting to shape the motorization process through local policies such as vehicle licensing and limits on the types of vehicles within the city. In the following, we discuss three important aspects motivating the Shanghai case study.

- **Wealth**

No single city or region can represent the whole of China due to significant regional differences in climate, level of urbanization, demographic composition, culture, languages[4], and most importantly, income.

Accurate comparisons of income are always challenging. One way is to use GDP/capita. We can compare the GDP/capita in "nominal" terms or in "purchase power parity (PPP)" terms. Usually, the PPP-adjusted GDP/capita makes more sense, since the same dollar has different purchasing powers in different countries/regions. Therefore, the nominal GDP/capita will be overestimated for a place whose currency is strong, but be underestimated for a place whose currency is weak. (Chi Hung Kwan, 2002) The following table shows the nominal versus PPP-adjusted GDP/capita in China and

[4] Although Mandarin is the official language in China, there are more than 50 local dialects.

Shanghai, 2000. For comparison, the PPP-adjusted GDP/capita in the US in 2000 was $36,200 (US CIA, *World Factbook*, 2001).

China	Shanghai
GDP/capita [nominal] = $840 (USD[5])	GDP/capita [nominal] = $1415 (USD)
GDP/capita [PPP, USA as benchmark] = $3940 (USD) (4.7 times the nominal)	GDP/capita [PPP, USA as benchmark] = $4245 (USD) (3 times the nominal)

Table 5: Nominal vs. PPP-adjusted GDP/capita in China and Shanghai (2000)

Source: World Bank, *World Development Report*, 2002

Figure 7 illustrates the big regional GDP/capita differences across China. East Coastal regions are the wealthiest with an average of almost $2,000 USD GDP/capita (PPP adjusted). Shanghai is at the top of the East Coastal region. If personal income is a prerequisite to personal vehicle purchase, Shanghai is a good place to study motorization pathways.

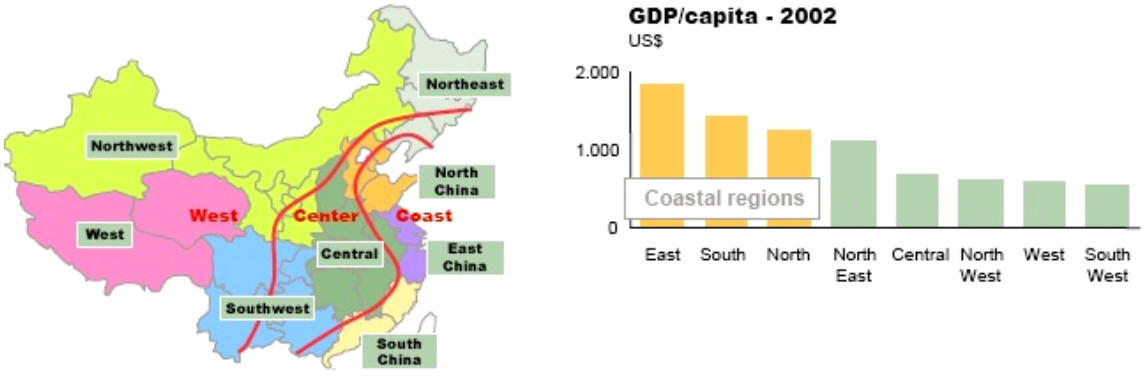


Figure 7: Regional Distribution of GDP/capita (PPP adjusted) in China

Source: Mercer Management Consulting, 2004. Data: *China Statistical Yearbook*

[5] 1 USD = 8.28 RMB (or Chinese Yuan), 2000 conversion rate

- **Diversity**

As mentioned, there are about five million floating (non-registered) people in Shanghai (Shanghai Bureau of Statistics, 2003) In terms of spatial distribution, there are one million floating people living in the urban area versus four million in the suburbs[6]. Among those floating people, 98.5 percent are actually Chinese from nearby (e.g., Jiangsu, Zhejiang) or remote rural (e.g., Anhui, Jiangxi) provinces. The remaining 1.5 percent is from other countries.

Japan	17,409
Taiwan	11,818
USA	8,248
S. Korea	7,135
Hong Kong	3,505
Singapore	3,263
Germany	2,541
Australia	2,499
Canada	2,352
Malaysia	1,955
U.K.	1,627
TOTAL Foreign	**72895 (people)**

Table 6: Foreign People in Shanghai

Source: *Shanghai Statistical Yearbook*, 2003

Historically, Shanghai has been China's commercial hub; it attracts people from different regions of China and different countries (Table 6). The former foreign-occupied[7] urban districts and newly developed suburban areas have already formed clusters of people

[6] The Pudong district itself attracted more than 1 million foreign residents.
[7] Many districts were occupied by foreign countries during WWII.

from different origins. Thus, Shanghai is very diverse, and this diversity can be realized by examining the vehicle purchase and use behavior (or motorization groups). In a nutshell, the diversity of Shanghai provides us a good opportunity to test hypotheses on explanatory variables of vehicle choice.

- **The "Car Debate"**

As mentioned, at aggregate national levels motorization is correlated to income. Past experiences of developed countries show that as consumers become wealthier, they tend to purchase automobiles. While this relationship generally holds throughout the world, there are significant variations among nations, and even among cities within the same nation. (U.S. National Research Council and Chinese Academy of Engineering et al., 2003) Today's China is a good example. As in Figure 8, we see a rather scattered distribution when plotting the GDP/capita versus vehicle ownership over different Chinese cities. Noticeably, Shanghai has a relatively low vehicle ownership given its high GDP/capita.

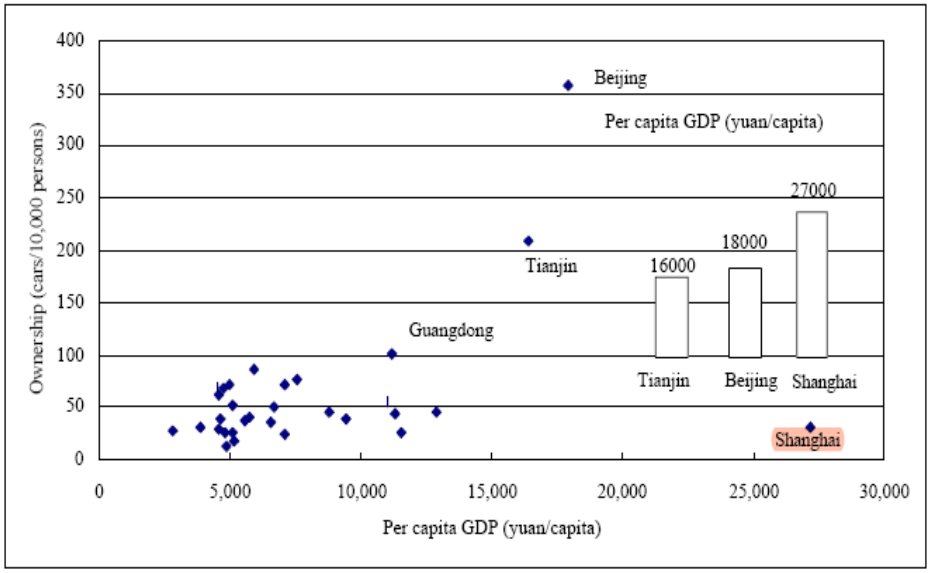


Figure 8: GDP/capita vs. Vehicle Ownership in China Cities (2002)

Source: Shen et al., 2002

Several reasons explain the relatively low car-ownership in Shanghai. One is government control: "*Shanghai is not a typical Asian city, given its surging economy and its world-class planning capabilities and strong government institutions"* (U.S. National Research Council and Chinese Academy of Engineering et al., 2003, Appendix B, p. 224). In the 1990s, Shanghai began to promote public transportation as one solution for decentralization. To match the "pro-public transportation" planning theme, motorization in Shanghai is regulated to some extent by car-restricting policies including: tight standards on fuel economy and emissions, as well as quota control of license plates[8].

However, Shanghai government's car-restricting policies are not consistent with the central (Beijing) government's "household car" idea. That is, in order to make the auto

[8] There is a "license auction" system held by the Shanghai Vehicle Management Bureau.

industry a pillar industry of China, Beijing expects that eventually every Chinese family will have one car. In fact, some scholars argue that the auto industry should also be the pillar industry of Shanghai (Wang et al., 2000). In addition to the policy conflicts, there are other factors challenging the car-restricting concepts of Shanghai. For example, local car companies usually lobby the city government to loosen controls on motor vehicle purchase to stimulate sales. As a consequence, government officials have begun to discuss the idea of separating the car use from purchase – for example, using congestion pricing to replace the existing license control. Besides, as many local people are employed in car companies (Shanghai VW, Shanghai GM); their car purchase is encouraged to some extent by the subsidy[9] from their employers. In fact, the so-called "transportation subsidy" (in terms of vehicle purchase, parking, fuel, etc.) is not unusual. In China, people sometimes don't bear all the costs of motor vehicle ownership, which causes difficulty to accurately estimate the effects of cost on people's vehicle purchase. Last, even within the Shanghai City government, there is a debate on car restricting policies. Instead of public transportation, some officials consider automobile-oriented development as a way to decentralize the city. A free license plate for people who purchase a house in the suburb has been considered (Wang et al., 2000).

Despite the "car debate", there was still a significant growth in the number of private cars in Shanghai over the past two decades. A study (Schipper et al., 2004) shows that, by comparing the trip mode shares for Shanghai between 1986 and 2001, Shanghai has decreasing shares of trips by walk and collective transit (bus and ferry), but rising shares for bike and private car.

[9] They can either get direct monetary subsidy or can purchase the car at a discounted price.

Focusing on the ownership of all kinds of private motor vehicles in Shanghai, comparing 2001 with 2003, I notice an increased share of private car and a decreased share of motorcycle, as indicated in Figure 9. I conclude from the aforementioned that although Shanghai's current private car ownership is relatively low compared to other mega-cities in China, it is growing at a rapid rate. How people make their vehicle purchase choices in a city with an on-going "car debate" is the last but not least reason motivating the Shanghai case study.

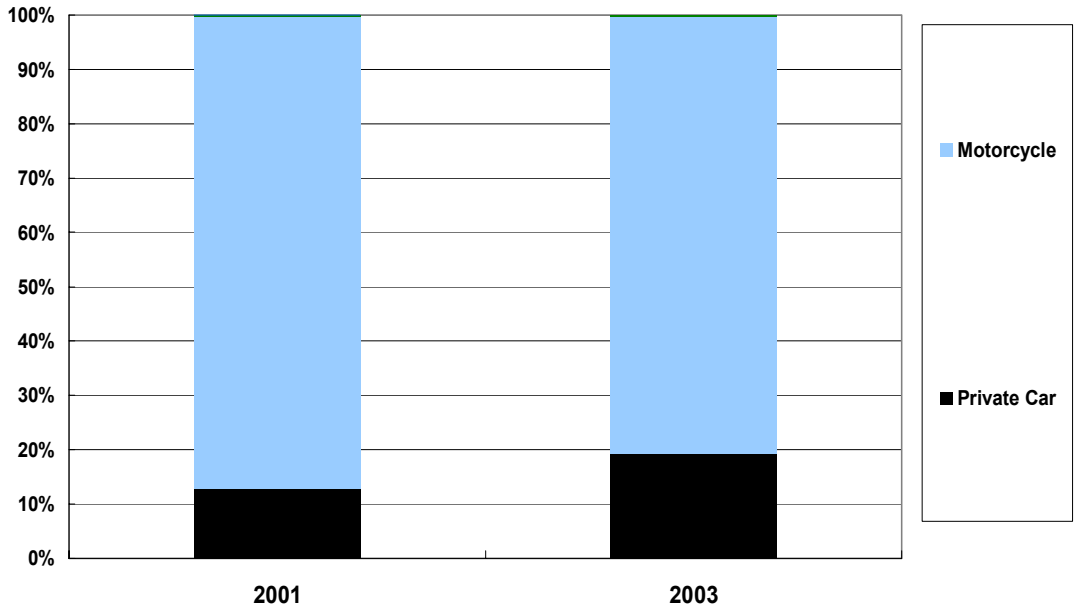


Figure 9: Private Motor Vehicle Ownership Share in Shanghai (2001 vs. 2003)

Source: *Shanghai Statistical Yearbook*, 2002 and 2004

Shanghai vs. Beijing

Shanghai and Beijing are two big cities in China; one is the commercial hub, the other is the nation's capital. In Figure 8, we see that Beijing has almost ten times the auto

ownership of Shanghai, but lower GDP/capita. Based on interviews[10], Table 7 summarizes and compares factors believed to affect differences in private car ownership between Shanghai and Beijing. (Signs indicate a negative or positive effect on car purchase, number of signs means the level of influence):

	Shanghai	**Beijing**
Car Restricting Policy	– –	–
Zoning Regulation	+	++
Public Transportation	– –	–
Population Growth	+	+
Personal Income (GDP/capita)	+ +	+

Table 7: Shanghai vs. Beijing – Assumed Factors Affecting Car Purchase

As discussed, Shanghai has many policies to restrict car purchases – the license quota-control is one famous example. People who wish to license an automobile must bid for a fixed number of license plates (usually 4,000 to 6,000) issued each month. In 2005, the auction price for a private car license was as high as $4,300 USD (with an additional $12 USD registration fee). However, there is no such restrictive car licensing policy in Beijing. On the contrary, in Beijing, bicycle lanes and sidewalks have been sacrificed in many places to allow more road space for autos. Moreover, as discussed, the central government in Beijing promotes the "household car" concept to demonstrate the government's resolution to make the auto industry a pillar industry in China.

[10] Several interviews have been done with local Chinese people including people from academia, car companies and others (e.g., taxi drivers).

In addition, some current zoning regulations in Beijing also make car purchases "necessary" – the Beijing city government doesn't allow new high-rise buildings in the central business district (CBD) to protect many historic sites in Beijing (Beijing is officially defined as the cultural and political center of China) and to save the skyline for the 2008 Olympics. A consequence of such zoning is urban sprawl and a car-oriented city. More and more people who cannot afford to live in the urban area choose to move to the suburbs and buy a car for the reason of convenience – based on personal interviews with staff of Energy Foundation of China and BMW China. Nevertheless, what is covered in Table 7 are merely factors mentioned by interviews in Shanghai and Beijing, there are presumably other factors affecting the vehicle purchase behavior in different regions of China. In the last part of this chapter, I discuss the goal of the Shanghai case study and outline of this dissertation.

Goal and Outline of the Dissertation

As discussed, the linear positive correlation at national levels between average income and motorization has been confirmed by past studies. However, I would like to explore and test if this relationship holds at the disaggregate level – that is, one major objective of this dissertation research is to ask what underlies the aggregate measures at the individual level of personal travel. To fulfill this objective, a questionnaire was designed and distributed to Shanghai residents in different phases of motorization, i.e., with different motorization pathways, including those who may not have yet acquired or used motorized modes. In addition to income, issues related to local policies, cultural beliefs, and socioeconomic context are discussed as they relate to Shanghai residents'

motorization pathways. All above aspects will be included as explanatory variables in the vehicle purchase and use models and tested statistically. The goal of this dissertation is to draw policy implications and sketch future perspectives for motorization in China from the case study of the wealthy and seemingly "under-motorized" City of Shanghai.[11]

There are six chapters of this dissertation. Chapter 1 is the introduction, and Chapter 2 is a literature review, which covers previous studies related to motorization, vehicle purchase behavior, and general survey research methodology in China. Hypotheses, survey techniques, and lessons learned from local survey implementation will be addressed as methodology in Chapter 3. Chapter 4 is a chapter about motorization pathway analyses, which explores the progression of vehicle use in Shanghai. Common pathway patterns and the direction of motorization will be discussed in that chapter. Chapter 5 first defines the dependent and independent variables of the discrete choice model. The attempt and results of developing discrete choice models of vehicle purchase (and use) will also be documented in the same chapter. Chapter 6 summarizes conclusions from my pilot and final surveys and aim to answer three important questions – How to conduct survey research in China? What is the motorization pathway in China? What is the vehicle purchase (and use) behavior in China? Policy implications and suggestions of future research directions will also be presented in Chapter 6.

[11] Clearly, calling Shanghai "under-motorized" may seem farcical to anyone who has traveled in that city. The description refers to the fact that Shanghai has far fewer automobiles than the simple aggregate motorization-income correlation would predict.

CHAPTER 2: LITERATURE AND KEY STATISTICS

Survey Research in China

As Chinese society underwent significant changes following the economic reforms of 1978, survey research became more common in China. In this chapter, I discuss eight prior transportation studies conducted in China based on survey research – the list is by no means complete. The eight are not all related to motorization in Shanghai, but were chosen for their discussion of survey techniques applied in Chinese contexts. I developed the survey methods for the pilot and final surveys by referring to these past studies. For example, by referring to previous case studies of Shanghai and Hong Kong (a city of similar motorization level as Shanghai), I developed some of my location-based sampling concepts. In Table 8, the eight studies are summarized and compared with my pilot and final surveys. Detailed methodology of the pilot and final surveys will be discussed in Chapter 3.

Survey Topic	Year	Location	Sampling Scheme	Distribution Method	Reward (promised in advance)
Vehicle Use Characteristic and Mode Choice Behavior	2006	Shanghai	Location-based convenience sampling	On-street	Yes
Vehicle Use Characteristic and Mode Choice Behavior	2006	Kunming	Location-based convenience sampling	On-street	Yes
Bicycle and Electric Two-wheeler User Survey	2006	Shijiazhuang	Location-based convenience sampling	On-street	No
Public Transportation Use and Transfer	2006	Shanghai	Convenience sampling	On-street	No
Shanghai Master Transportation Survey	2004	Shanghai	Random & Convenience sampling	Household, GPS, etc.	N/A
Vehicle Use Behavior	2002	Shanghai	Convenience sampling	On-street	No
Car Dependence	2001	Hong Kong	Location-based convenience sampling	On-street	N/A
Motorization and Obesity	1989 1991 1993 1997	Shangdong, Jiangsu, Hunan, Hubei, Henan, Guizhou, Guangxi, Liaoning	Multi-stage, random cluster sampling	Household	N/A
My Present Study: Vehicle Purchase and Use Behavior (Pilot Survey)	2005	Shanghai	Location-based and other convenience sampling	On-street	Yes
My Present Study: Vehicle Purchase and Use Behavior (Final Survey)	2006	Shanghai	Location-based and other convenience sampling	On-street, auto dealership, household, cell phone message	Yes

Table 8: Comparison of Survey Research Projects in China

- **Shanghai: Vehicle Use Characteristic and Mode Choice Behavior Survey**

In late May 2006, UC Berkeley researcher Chris Cherry conducted a survey on vehicle use characteristics and mode choice behavior in Shanghai. The survey targeted electric bike, bicycle and LPG scooter users. Travel diary, demographic information and attitudinal questions were included in the questionnaire. The survey was conducted during the periods of daily activity, from mid-morning to evening and during the middle of the week, from Tuesday to Friday, so that the previous day travel diary would represent a "typical" weekday (Monday to Thursday). Location-based convenience sampling was chosen:

> "Conducting a random household survey in China is logistically and institutionally difficult. As a result, targeted intercept surveys were conducted at locations that contain a representative sample of urban two-wheel vehicle users, specifically centralized parking facilities of major activity centers and trip generators throughout the urban area. These activity centers contain employment, social activities, and shopping that serve all demographic groups." (Cherry, 2007)

After the survey, the participants were offered rewards (parking fee payment, as promised in advance) as tokens of appreciation. A total of 696 responses were collected in Shanghai.

- **Kunming: Vehicle Use Characteristic and Mode Choice Behavior Survey**

Similarly, another vehicle use characteristic and mode choice behavior survey was

conducted in Kunming, a mid-size Chinese city, by Chris Cherry in early April 2006. The same questionnaire was used as in the Shanghai survey, and the sampling scheme was also a convenience sample based on locations – "These locations included major shopping centers that cater to all demographics of users as well as centralized bike parking facilities surrounding a large pedestrian mall in the center of the city that contains shopping, entertainment, and employment. Importantly, most of the survey sites were within the gas motorcycle restricted zone" (Cherry, 2007). However, the Kunming survey differs from the Shanghai survey by only targeting electric bicycle and bicycle users (without the LPG scooter users). Small gifts were promised in advance and offered at the end of the survey, and a total of 502 responses were collected.

- **Shijiazhuang: Bicycle and Electric Two-wheeler User Survey**

In June 2006, UC Davis researcher Jonathan Weinert conducted a survey targeting the bicycle and electric two-wheeler (E2W) users in Shijiazhuang, a mid-size city located in the south-central Hebei province of China. The sampling plan is location-based convenience – "Because of the institutional and logistical difficulty in conducting random household surveys in China… The survey was administered at bicycle and E2W parking lots along the main travel corridor (Zhongshan Lu) in Shijiazhuang in order to capture a diverse range of respondents from many different parts of the city" (Weinert, 2007). According to the author, the survey was implemented on both a workday and weekend day 7:30am – 11:30am and 3:00pm – 6:00pm to collect as broad a range of respondent types as possible. Besides, separate surveys were given to bicycle and E2W riders to identify differences between their travel behavior and attitudes. Finally, 751 responses were collected for bicycle users and 460 responses were collected for E2W users.

- **Shanghai: Public Transportation Use and Transfer Survey**

Tongji University implemented a survey of public transportation users in August 2006. The study sought to understand the trip purpose of subway/light rail riders and their attitudes toward making transfers to local bus systems. Convenience sampling was used: people waiting on eleven pre-selected subway/light rail platforms were interviewed. The survey was conducted during the morning and evening peak hours, i.e., 7:00am to 10:00am and 4:00pm to 7:00pm. A total of 7,816 surveys were collected with a response rate of 14.2 percent.

- **Shanghai: Master Transportation Survey**

The Shanghai city government and Shanghai City Comprehensive Transport Planning Institute (SCCTPI) implemented a large transportation survey in 2004. The survey covered the entire metropolitan area, divided into 309 transportation analysis zones (TAZ). It involved 40,000 surveyors; a total of 200,000 surveys were randomly (except for a convenience sample described later) distributed to 30,000 households, 20,000 organizations, and 70,000 car drivers. The population to be sampled consisted of multiple groups of people, thus, several sampling procedures and survey instruments were used. For example, Shanghai residents' travel behavior was measured through a household survey, while the origin-destination patterns of 2,300 taxis were gathered with GPS devices. A convenience sample was used, in part, as city employees were automatically selected for the survey regardless of their prevalence in the city's population.

- **Shanghai: Vehicle Use Behavior Survey**

Another survey in Shanghai was conducted by Tongji University in 2002 on vehicle use

behavior. The purpose of this study was to support the road planning of the central city of Shanghai. Convenience sampling was used to survey drivers (or car users). Questionnaires were distributed in five "road fee" collection areas in central Shanghai. 1,630 surveys were returned over a two-week period. The questionnaire included a simple vehicle use diary and questions about vehicle ownership.

- **Hong Kong: Car Dependence Survey**

In summer 2001, a survey on car dependence has been conducted in Hong Kong (Cullinane, 2003). 401 car-owning Hong Kong residents were interviewed. Their sampling scheme was a convenience sample based on locations. They used convenience sampling because "… it proved impossible to obtain the contact details of car owners in Hong Kong" (Cullinane, 2003). And they designed their location-based sampling in residential and commercial areas of Hong Kong to balance the probability of finding car owners with any "bias associated with approaching only car owners who were using their cars at the time."

- **China: Motorization and Obesity Survey**

Other researchers have asked whether motorization causes obesity. In 1989, Bell et al. (2002) conducted an initial survey in eight Chinese provinces (Shangdong, Jiangsu, Hunan, Hubei, Henan, Guizhou, Guangxi, and Liaoning) as well as follow-up surveys in 1991, 1993, and 1997. They used a multi-stage, random cluster sampling process in which four counties were selected in each province, within which neighborhoods were randomly selected from suburbs, townships, and villages. Finally, within in each of these neighborhoods, twenty households were selected at random.

Motorization in China

In Chapter 1, motorization is defined as the changing of transportation from non-motorized means to motorized means. At the national or province level, motorization is usually reflected by the growth of the automobile fleet or per capita auto ownership. Therefore, longitudinal data on vehicle ownership (or use) is necessary to describe motorization. In the following, the change of vehicle ownership in Shanghai, Beijing, and China over a 10-year period (1996 to 2005) will be presented. In addition, distance-based mode shares of Shanghai in 1986, 1995, 2000, and a forecast for 2020 will also be presented to provide another view of motorization.

- **China: Change of Vehicle Ownership (Urban vs. Rural, 1996 – 2005)**

The major source for vehicle ownership data is the China Statistical Yearbook published each year by the National Bureau of Statistics of China. In the China Statistical Yearbook, there are data about the "number of durable consumer goods owned per 100 urban/rural households". In Table 9, I present the ownership data for "urban households" in China from 1996 to 2005 for: bicycle, motorized two-wheeler, motorcycle and car, which are also the four major alternatives in my vehicle purchase choice model. Ownership of bicycle decreases over the 10-year period, whereas car ownership increases, especially from 2002 to 2005. There were no official data for the motorized two-wheeler until 2002; there was no data for car ownership until 1997.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Bicycle	193.23	179.10	182.05	183.03	162.72	165.42	142.71	143.55	140.21	120.04
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	2.72	4.25	6.50	9.54
Motorcycle	7.94	11.60	13.22	15.12	18.83	20.40	22.19	24.00	24.84	25.00
Car	N/A	0.19	0.25	0.34	0.51	0.60	0.88	1.36	2.18	3.37

Table 9: Numbers of Vehicles Owned per 100 Urban Households [China, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

Table 10 shows the ownership data for "rural households" in China from 1996 to 2005. The data for motorized two-wheeler and car are missing.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Bicycle	139.82	141.95	137.15	136.85	120.48	120.83	121.32	118.50	118.15	98.37
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Motorcycle	8.45	10.89	13.52	16.49	21.94	24.71	28.07	31.80	36.15	40.70
Car	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 10: Numbers of Vehicles Owned per 100 Rural Households [China, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

- **Shanghai: Change of Vehicle Ownership (Urban vs. Rural, 1996 – 2005)**

Table 11 presents the ownership data for "urban households" in Shanghai from 1996 to 2005 for bicycle, motorized two-wheeler, motorcycle and car. There were no official data for the motorized two-wheeler until 2002; and there was no data available for car ownership until 1998. Similar to the case of China, the ownership of bicycle decreases, but the ownership of motorized two-wheeler, motorcycle and car all increase in those ten

years. Besides, a big jump of car ownership happens in the end of 2002 reflecting the car buying craze in China (mentioned in Chapter 1).

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Bicycle	124.00	125.40	132.60	138.80	125.60	123.80	123.60	125.40	125.60	119.10
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	12.65	17.80	20.30	23.50
Motorcycle	1.00	1.20	1.40	2.00	1.20	1.40	2.88	3.80	3.00	2.70
Car	N/A	N/A	0.00	0.00	0.00	0.20	0.25	1.80	3.60	3.80

Table 11: Numbers of Vehicles Owned per 100 Urban Households [Shanghai, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

In terms of the vehicle ownership for rural households in Shanghai, a trend of decreasing numbers of bicycles and increasing numbers of motorcycles is shown in Table 12. This trend is similar to what I found for the rural households in the previous China case.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Bicycle	244.17	245.50	238.50	235.33	218.83	212.67	215.67	210.67	202.17	173.50
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Motorcycle	33.17	45.50	54.83	60.67	72.67	73.17	82.83	87.67	90.67	72.00
Car	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 12: Numbers of Vehicles Owned per 100 Rural Households [Shanghai, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

- **Beijing: Change of Vehicle Ownership (Urban vs. Rural, 1996 – 2005)**

The numbers of vehicles owned per 100 urban households (1996 – 2005) in Beijing are

shown in Table 13. Data for motorized two-wheelers before 2001 and data for car ownership before 1997 are not available. Similar to the case of China and Shanghai, the trend of motorization of Beijing is also the decrease of bicycle ownership and the increase of motorcycle; motorized two-wheeler, and car ownership. However, I found two aspects differentiating Beijing's motorization from previous cases. First, the growth rate of car ownership in Beijing is significantly faster than for Shanghai and China as a whole, especially after 2002. Second, although the growth of motorized vehicles (motorcycle, motorize two-wheeler, and car) is fast, bicycle is still the mainstream vehicle in urban Beijing. Moreover, the urban households of Beijing have the highest percentage of bicycle ownership as compared to urban households in Shanghai and China.

	1996	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**
Bicycle	249.00	209.40	221.00	220.10	230.70	230.60	201.15	202.06	191.83	193.71
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	3.19	4.61	3.99	5.08
Motorcycle	3.40	4.20	3.80	6.00	5.70	5.00	5.54	5.59	5.65	6.32
Car	N/A	0.80	1.00	2.50	2.50	2.60	4.05	6.60	12.64	14.06

Table 13: Numbers of Vehicles Owned per 100 Urban Households [Beijing, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

Table 14 present the vehicle ownership of rural households in Beijing (1996 – 2005). Data for motorized two-wheeler and car ownership are not available. A trend of decreasing ownership of bicycle and increasing ownership of motorcycle are identified as in previous cases of Shanghai and China, although the bicycle still remains the dominant

vehicle. Besides, compared to rural Shanghai and China, the rural households of Beijing generally have a higher percentage of bicycle ownership over the period of 1996 to 2005.

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Bicycle	260.40	261.60	258.40	248.67	219.20	216.80	208.13	212.80	208.67	196.00
Motorized Two-wheeler	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Motorcycle	17.47	23.33	27.60	36.67	36.67	36.00	41.33	41.87	42.93	45.33
Car	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 14: Numbers of Vehicles Owned per 100 Rural Households [Beijing, 1996 – 2005]

Source: *China Statistical Yearbook*, 1997 – 2006

To summarize the above tables, different motorization patterns can be indentified for Shanghai, Beijing and China as a whole. Further, the motorization patterns are also slightly different for urban and rural households within the same city or country. Nevertheless, one trend in common is the decreasing ownership of non-motorized vehicles, e.g., bicycle; and the growth of motorized vehicles, e.g., car – although bicycle is still the dominant vehicle type for (urban and rural) Shanghai, Beijing and China.

- **Shanghai: Change of Distance-based Mode Shares (1986, 1995, 2000, and 2020 forecast)**

In addition to vehicle ownership, motorization can also be studied by understanding the change of distance-based mode shares, which are discussed in terms of "passenger-km" in this section. The distance-based mode share data include not only the "purchasable" vehicles but also travel means, which cannot be purchased and owned by individuals, e.g.,

public transportation.

According to data from the Energy Foundation and the Shanghai City Comprehensive Transportation Planning Institute (SCCTPI) (Figure 10), in the 1980's, walk and public transportation were the two major transportation modes, followed by bicycle. However, because of decreasing of government subsidies[12] of public transportation, urban sprawl, change of commute pattern, etc.; the "private" modes (including bicycle, motorcycle and car) became the mainstream in the 1990's. In 2000, public transportation again took a large share due to the completion of several new networks. However, the motorcycle and, especially car, also grew hugely. As a result, the walk share shrank to only seven percent. Basically, it shows a "non-motorized (walk) to motorized (car)" motorization pattern. For the future (2020), as an effort to solve the urban congestion and pollution problem, the policy of the Shanghai government is to promote sustainable transportation – walk, bicycle, and public transportation are major options. However, the share of car is still expected to be 20 percent under the 2020 scenario of SCCTPI.

In addition to SCCTPI's forecast, in the book "*Personal Cars and China*" (U.S. National Research Council and Chinese Academy of Engineering, 2003), two future motorization scenarios of Shanghai are explored – a "high motorization" scenario is based on market forces playing a greater role in the economy, and government playing a lesser role. Therefore, Shanghai will follow the path of fast-growing cities in Asia with high car ownership such as Bangkok and Jakarta. A "low motorization" scenario assumes that the government plays an active role in restraining vehicle purchases and use. Thus, Shanghai

[12] One reason is to improve the competitiveness of transit operators.

will follow the path of cities such as Singapore, Tokyo, and Hong Kong. Both high and low motorization scenarios are shown in terms of distance-based mode shares (passenger-km) in Figure 10.

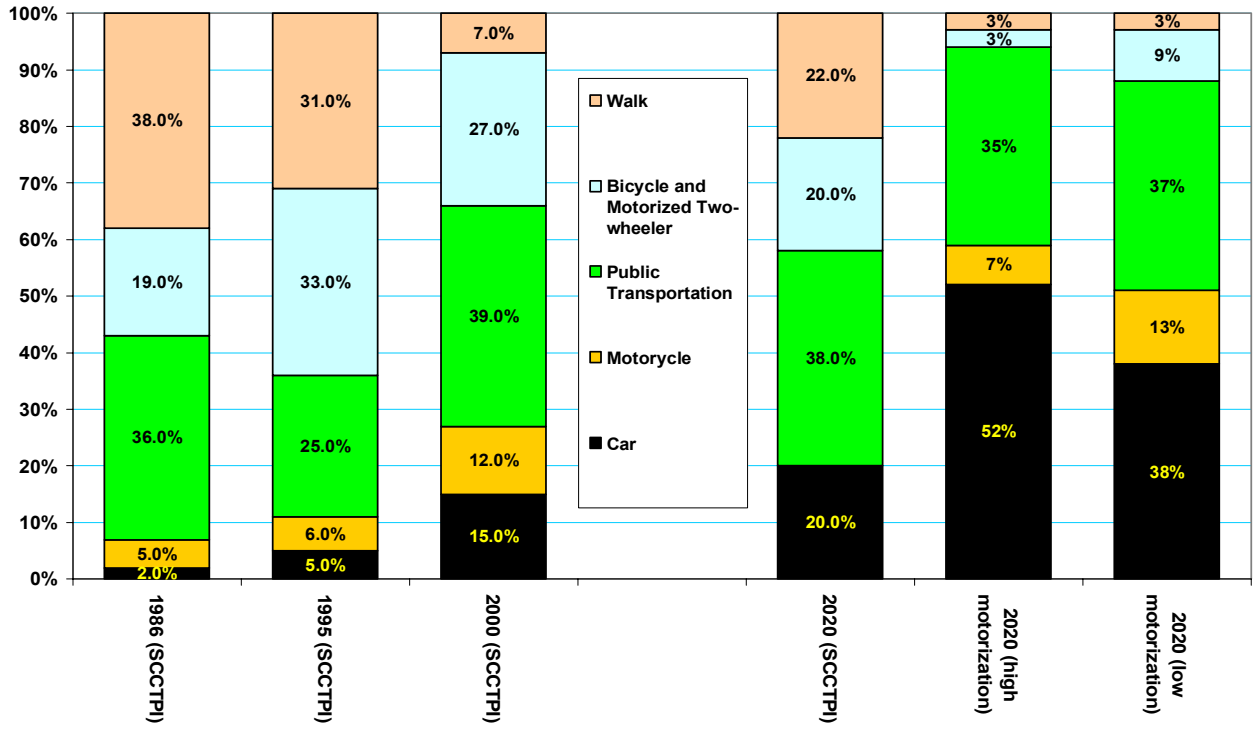


Figure 10: Distance-based Mode Shares of Shanghai (1986, 1995, 2000, 2020)

Source: Energy Foundation, SCCTPI, U.S. National Research Council and Chinese Academy of Engineering, *Personal Cars and China* (Washington, DC: National Academies Press, 2003)

The above discussion on vehicle ownership or distance-based mode share basically reflects the aggregate level of motorization at different time points. However, there are not many studies focusing on the "transition" or "motorization pathway" at the individual level. One attempt of my dissertation is to describe the motorization pathway at the individual level through a survey of Shanghai residents.

- **Concepts of the Discrete Choice Model**

A choice model is a quantitative method to estimate the level of influence and statistical significance of attributes affecting people's purchase behavior. Basically, the discrete choice model includes the decision-makers (individuals), alternatives, and variables describing the choice context.

Two common discrete choice model structures are the multinomial logit model (MNL), and the nested logit model (NL), which has MNL as a special case. Both models are utility-based, assuming $U_{in} = V_{in} + \varepsilon_{in}$, where U_{in} is the utility of individual n for alternative i, V_{in} is the deterministic part of U_{in} and can be expressed as a linear-in-parameters function of observed explanatory variables (X_{in}); ε_{in} is the error term. The logit model is derived from the assumption that the error terms are independent and identically Gumbel distributed. If the error terms are independent and identically Gumbel distributed (with location parameter 0 and scale parameter μ), the probability that a given individual chooses alternative *i* within choice set *C* is given by:

$$P_C(i) = \frac{e^{\mu V_i}}{\sum_{k \in C} e^{\mu V_k}}$$

An important property of the MNL model is the Independence of Irrelevant Alternatives (IIA), which states that the ratio of the probabilities (i.e., the relative *odds*) of any two alternatives is independent from the "choice set". That is, for any choice sets *S* and *T*, and $S \subseteq T \subseteq C$, for any alternative α_1 and α_2 in *S*, we can find:

$$\frac{P_S(a_1)}{P_S(a_2)} = \frac{P_T(a_1)}{P_T(a_2)}$$

Defined by Ben-Akiva and Lerman (1985) as: "The ratio of the choice probabilities of any two alternatives is entirely unaffected by the systematic utilities of any other alternatives", the IIA property sometimes becomes a limitation for practical application, as illustrated by the famous red bus/blue bus paradox (Ben-Akiva and Lerman, 1985).

The nested logit (NL) model provides a solution to partly overcome the limitation of the multinomial logit model (NL does not require IIA to hold). First derived by Ben-Akiva (1973), the NL model is an extension of the MNL model designed to capture the unmeasured correlations [13] among alternatives. The NL model is based on the partitioning of the choice set into several nests (*k=1, 2 ... n*):

$$C = \bigcup_{k=1}^{n} Ck$$

The utility function of an alternative is now composed of a term "specific to the alternative", and a term "associated with the nest". That is, for alternative $i \in C_k$:

$$U_i = V_i + \varepsilon_i + V_{C_k} + \varepsilon_{C_k}$$

The two error terms ε_i and ε_{Ck} are independent. For the multinomial logit model, error terms ε_i are supposed to be independent and identically Gumbel distributed, with scale parameter σ_k. The distribution of ε_{Ck} is such that the random variable $\max_{j \in C_k} U_j$ is Gumbel distributed with scale parameter μ.

[13] For example, shared unobserved attributes (error terms), which violate the independent and identically Gumbel distribution assumption.

The probability that a given individual chooses alternative *i* within choice set *C* is:

$$P_C(i) = P_C(C_k)P_{C_k}(i)$$

Obviously, this is a product of two probabilities. First, the probability of choosing C_k among *C* is denoted as:

$$P_C(C_k) = \frac{e^{\mu V'_{C_k}}}{\sum_{l=1}^{n} e^{\mu V'_{C_l}}}$$

The V'_{ck} is called the "composite utility[14]" of nest *k*. The μ is the scale parameter of $\varepsilon_{Ck,}$ which is Gumbel distributed. The composite utility for nest C_k is:

$$V'_{C_k} = V_{C_k} + \frac{1}{\sigma_k} \ln \sum_{j \in C_k} e^{\sigma_k V_j}$$

(V_{ck} is the component of the utility common to all alternatives in the nest $C_{k,}$)

Similarly, the probability of choosing *i* among C_k is denoted as:

$$P_{C_k}(i) = \frac{e^{\sigma_k V_i}}{\sum_{j \in C_k} e^{\sigma_k V_j}}$$

The σ_k is the scale parameter of $\varepsilon_{i,}$ which is also supposed to be independent and identically Gumbel distributed.

Above is a very simple introduction to two discrete choice models; details about the model development process will be presented in Chapter 5.

- **Studies of Vehicle Purchase Behavior in China**

In China, few studies have attempted to develop sophisticated choice models of vehicle

[14] It is also called "pseudo-utility", "expected maximum utility", "inclusive value".

purchase behavior. Nevertheless, the auto market survey of the Mercer Group (2004) provides one example of vehicle purchase behavior study in general[15]. Figure 11 presents the purchase criteria of Level 1 city car buyers[16]. Interestingly, instead of price, most people (96 percent) consider "safety" and "reliability" as the most important criteria for their car purchase.

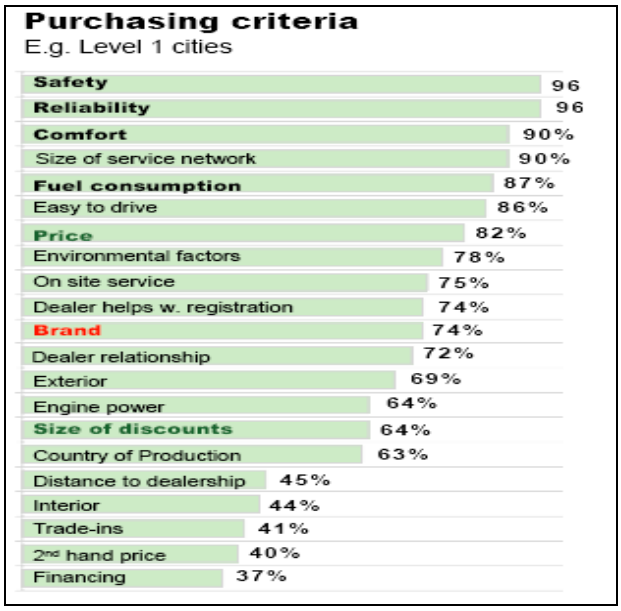


Figure 11: Car Purchase Criteria (Shanghai, Beijing, Guangzhou, Shenzhen)
Source: Mercer Consulting Group, 2004

However, another survey in Beijing (Energy Foundation, 2005) asked 150 potential car buyers to rank five most important considerations in their car purchase. As shown in Table 15, "price" is the top-ranked one for this sample.

[15] There were more than 2,000 participants in seven cities of China, including existing and potential car buyers.
[16] Shanghai, Beijing, Guangzhou, Shenzhen – as categorized by Mercer.

Importance Rank	Attribute
1st	Price
2nd	Brand
3rd	Fuel Economy
4th	Model/Style
5th	Power

Table 15: Five most Important Car Purchase Criteria (Beijing)

Source: Energy Foundation, 2005

Instead of analyzing vehicle purchase criteria as in the previous two cases, one survey in Hong Kong (Cullinane, 2003) focused on the deterrence factors of driving, which is related to the exogenous environment deterring vehicle purchase. Traffic congestion and parking availability at destination were the top two deterring of driving.

Extent of deterrence from driving in Hong Kong						
	Very much	Quite a lot	Neutral	Not very much	Not at all	Average (rank in brackets)
Traffic congestion	129 (32)	118 (29)	20 (5)	63 (16)	71 (18)	3.43 (1)
Parking costs at destination	60 (15)	138 (34)	19 (5)	93 (23)	91 (23)	2.96 (4)
Parking availability at destination	88 (22)	142 (35)	21 (5)	73 (18)	77 (19)	3.23 (2)
Unreliability of parking availability	54 (13)	146 (36)	32 (8)	85 (21)	84 (21)	3.00 (3)
Tunnel costs	21 (5)	75 (19)	37 (9)	119 (30)	148 (37)	2.26 (5)
Petrol costs	18 (5)	71 (18)	27 (7)	135 (34)	149 (37)	2.19 (6)
Route unfamiliarity	19 (5)	69 (17)	25 (6)	116 (29)	170 (43)	2.13 (7)
Stress of driving	2 (1)	24 (6)	14 (4)	125 (31)	5 (59)	1.01 (8)

Table 16: Deterrence Factors of Driving (Hong Kong)

Developing a robust model to represent Chinese car buyers is a challenging task due to huge regional variation. Nevertheless, "Shanghai Vehicle Purchase (and Use) Models" are proposed and developed in this dissertation. The models are supposed to address the influence and statistical significance of attributes of vehicle purchase and use behaviors in Shanghai.

CHAPTER 3: METHODOLOGY

Methodology Overview

In this chapter, I report on the design and implementation of the data collection to test to my hypothesis of motorization (pathway) in China. This chapter starts with the discussion of the hypothetical motorization pathway.

By hypothesis, motorization is a sequential transition from non-motorized to motorized travel modes at a national (or some other large aggregate) level. While the process can be thought of as adding to the modes people have available to them for daily travel, it also implies changes in the distribution of use of travel modes by individuals, e.g., if I get a bicycle, I walk less, and longer term changes in where people travel, e.g., if I get a car, I can eventually move to a suburb – which may not have good transit service. I characterize the personal transitions as a motorization pathway, and I hypothesize the direction of the motorization pathways is from non-motorized to motorized, from low cost to high cost, and from shared to private-owned (Figure 12).

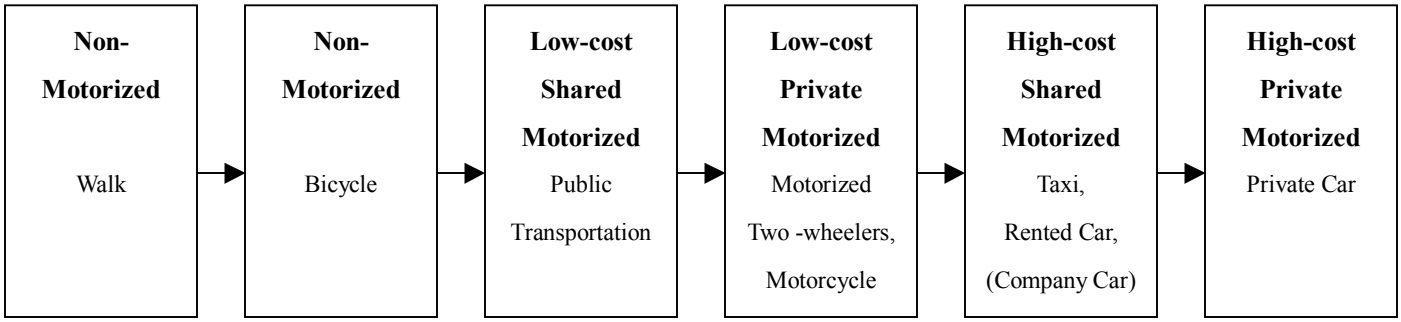


Figure 12: Hypothetical Motorization Direction

This survey research is designed to include people from all the steps of the hypothetical motorization pathway. Because we were unable to generate a single probability-based sampling frame for the residents of Shanghai, and because it is not necessary for the purposes of my research to accurately characterize Shanghai according to the distribution of its residents across the motorization pathway, I did not implement a probability-based sample. Rather, I hoped to insure that I had respondents in each step of the motorization pathway through the use of multiple convenience samples. I conducted the survey at several specific locations – a location-based variant of convenience sampling. For example, I conducted the surveys in subway stations for the subway riders. With respect to step in the motorization pathway, it was not exactly a choice-based sample, because any given subway rider might own a car, but still be using the subway. Still, by sampling at subway stations, I expected to interview some respondents who are no further along in their motorization pathways than low-cost, shared, motorized modes.

Two surveys – a pilot and a final – were designed and implemented. A pilot survey of 122 residents of Shanghai was conducted in late 2005, and a final survey of 1,037 people was conducted in mid-2006. The pilot survey was not merely a pre-test of the final survey, but included the additional goal of getting a basic sense of the geographical and socioeconomic context of Shanghai, which was important for determining sampling locations for the final survey. The final survey, containing substantially more questions than the pilot study, was used to test the idea of motorization pathways in Shanghai (Chapter 4) and provide the data for the vehicle choice model (Chapter 5). For instance, in the final survey, there are questions regarding the utility-based comparison of different travel modes. In addition, questions about personality, lifestyle and the exogenous

environment are included as they are considered to be variables affecting vehicle purchase behavior.

The primary goals of this chapter are to relate the survey design to my motorization pathway assumption, to document the research methods step by step, and to describe for others who may wish to conduct survey research in China what worked and what did not. In the following, there will be detailed discussion about both pilot and final survey. I will discuss sampling, survey instruments, and incentives. Implementation issues such as determining the survey time of day, and survey team selection and training are also discussed. The last part of this chapter will be focused on lessons learned from local implementation, and findings (about the research methodology) based on survey results.

Pilot Survey

- **Sampling**

As mentioned, one major objective of the pilot survey was to interview people at different steps of the hypothesized motorization pathway (Figure 12). To fulfill this objective, location-based sampling was selected – that is, I surveyed different locations related to travel modes (or purchased vehicle) as well as people's mobility characteristics including: income and transit access.

In the design of the sampling scheme, I first referred to the estimates of the true shares of vehicle purchase or use in Shanghai; I used data on distance-based mode share as a proxy (discussed in Chapter 2). The Shanghai City Comprehensive Transport Planning Institute

(SCCTPI, 2000) published distance-based mode share data for Shanghai. These data showed trip shares of 7 percent walk, 27 percent bicycle, 39 percent public transportation, 12 percent motorcycle, and 15 percent private car. Considering the shares of walk, motorcycle, and car trips are relatively small, I intentionally sampled from certain locations for the purpose of enriching those mode shares. That is, since I knew the total sample size of the pilot study would be small, I wanted to be sure that I had people from all steps of the motorization pathway in the sample. For instance, I surveyed onboard the ferry across the Yangtze River to sample people riding motorcycles or motorized two-wheelers. The ferry is their most common way to travel across the Yangtze River.[17] In the pilot sample, 75 percent of people surveyed in the ferry terminal or on the ferry were people using motorcycles or motorized two-wheelers (Figure 13).

Figure 13: Pilot Survey on Ferry

[17] People are not allowed to ride motorcycles and motorized two-wheelers on bridges or to bring those vehicles onto the subway.

People with different mobility characteristics, i.e., people at different steps of the motorization pathway, are conceived as different market segments. In order to cover the whole range of segments, I chose to sample at places where I expected to find people in certain segments. For example, I sampled at high-end shopping malls for high-income people who have the capability to purchase automobiles, whether they have done so or not; at Walmart stores for middle-income people or the so-called "salary class", whose income is more likely to afford them motorcycles, use of public transportation, or taxis; and in the old town Shanghai (Lao Ximen, Figure 14) for the lower-income people who are more likely to walk or ride bicycle. Further, where a respondent is in their mobility pathway may vary according to location and access to different travel modes. Such differences as between the central city and suburbs relate to people's living and working location and possible modes of travel. So we sampled at the Huang Du town outside the metropolitan area of Shanghai to compare to results from the other locations within the city. According to the above guidelines just discussed, ten locations were selected for the pilot survey (Figure 15).

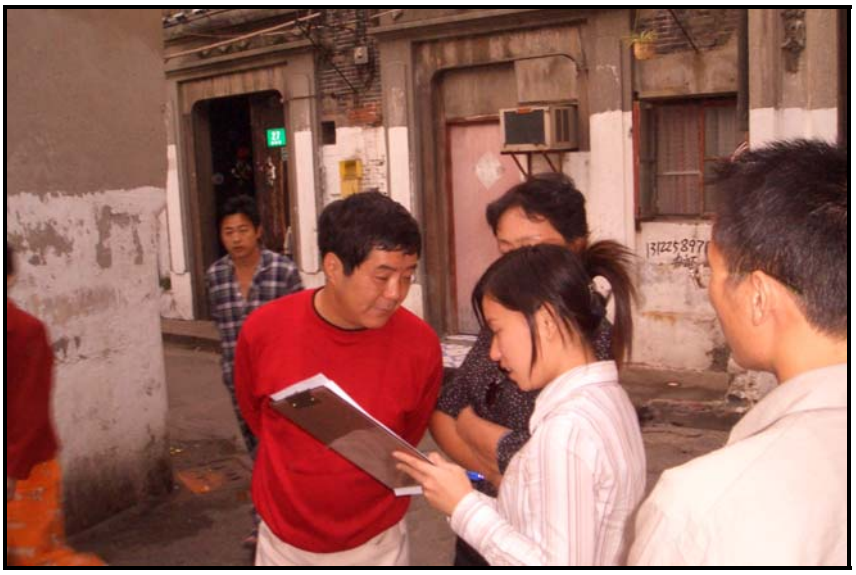


Figure 14: Pilot Survey at Old Town Shanghai (Lao Ximen)

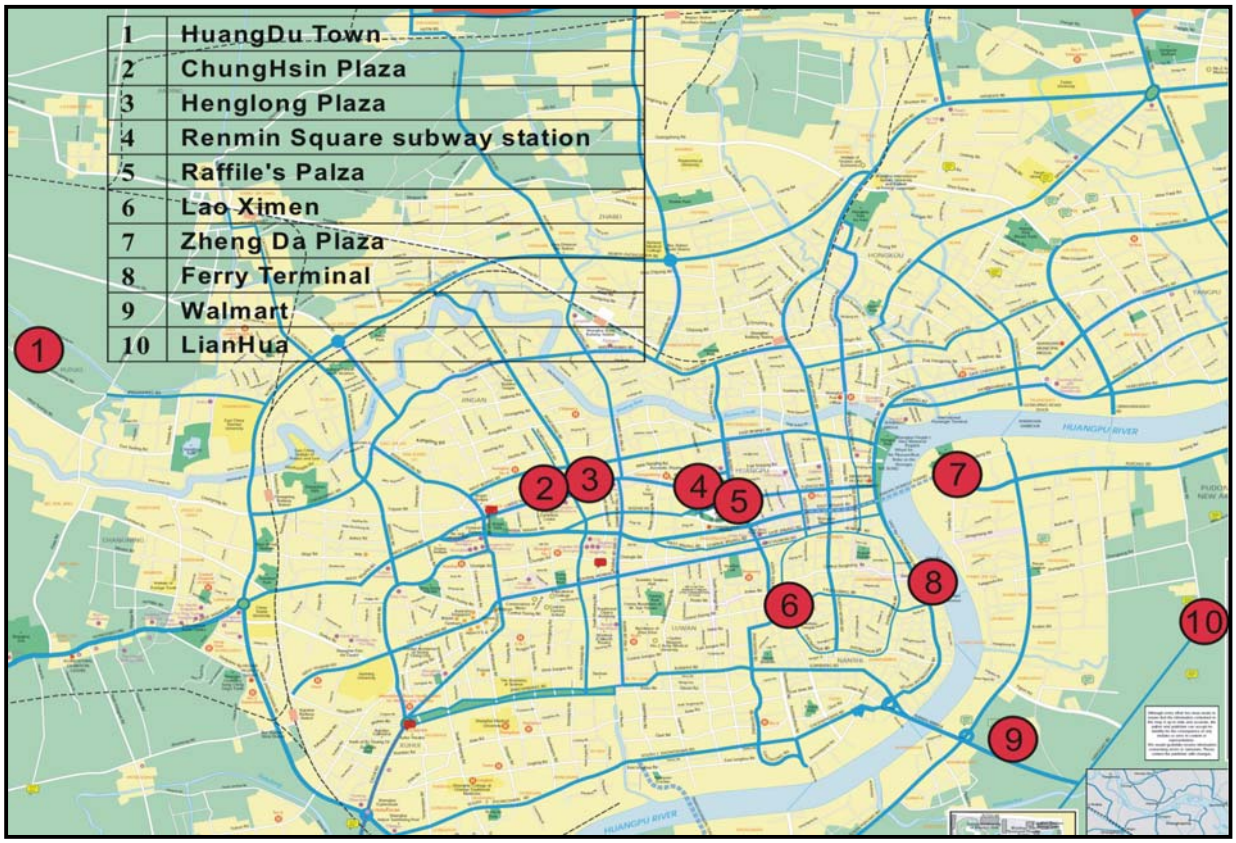


Figure 15: Survey Locations (Pilot Survey)

The pilot survey was conducted from October 14th to November 20th 2005. I distributed 164 questionnaires at the ten locations. In addition, 28 questionnaires were distributed through a (pure) convenience sample, i.e., not location-based, of colleagues and friends. A total of 122 questionnaires were returned for an overall response rate of 63.5 percent. Details about the survey locations and the response rates are provided in Table 17. High response rates were associated with locations where people were generally less busy or rushed. In the old town Shanghai (Lao Ximen), the many elderly people walking on the streets may have been less busy than people on their way home from work. The high response rates of the pure convenience sample may be due to their high level of trust for the survey-givers as this sample was made up of friends and colleagues.

ID	Survey Location	Survey Date (2005)	Distribution Method	# of Survey Distributed	# of Valid Response	Response Rate (percent)
1	Huang Du Town	11/20	On-street	13	8	61.5
2	Chung Hsin Plaza	10/21	On-street	9	4	44.4
3	Heng Long Plaza	10/21	On-street	10	5	50.0
4	Renmin Square subway station	10/28	On-street	27	11	40.7
5	Raffle's Plaza	10/16	On-street	18	10	55.6
6	Lao Ximen	10/16	On-street	5	5	100.0
7	Zheng Da Plaza	10/14	On-street	21	12	57.1
8	Ferry	10/28	On-street	17	16	94.1
9	Walmart	10/22	On-street	20	12	60.0
10	Lian Hua Supermarket	10/22	On-street	24	12	50.0
11	Peer Network	10/12 to 11/4	In person, e-mail	28	27	96.4

Table 17: Survey Locations and Response Rates (Pilot Survey)

- **Distribution Method**

In the pilot survey, I used on-street intercept methods except for some questionnaires sent by e-mail to known associates. There are two reasons for using on-street distribution. First, the on-street method can better facilitate the location-based sampling. Second, by actually visiting various places in Shanghai, I gained a better geographical and cultural understanding of this city. This was important since I planned to also use location-based sampling for the final survey. The pilot survey not only tested questions and sampling procedures, but was also an opportunity to comprehensively "picture" the metropolitan area of Shanghai.

- **Questionnaire Design**

Because the pilot study was mostly conducted as a street intercept survey, the questionnaire had to be simple and short to increase the probability that people would agree to take the questionnaire, and complete it once they agreed to take it. The questionnaire consisted of 20 questions, printed on three pages, double-sided. I estimated it could be finished in less than ten minutes. The questions asked about modes of travel owned and used, motivation for vehicle purchases, vehicle use patterns, and demographic data. I asked people to respond to questions about their past, present, and expectations for the future in order to observe transitions along the motorization pathway. I used both factual, e.g., "Do you currently own any vehicle?" and attitudinal, e.g., "Please rank three most important reasons for your vehicle purchase." questions (See Appendix for the questionnaire of the pilot survey).

- **Incentive**

Due to budget constraints for the pilot study, I initially did not plan to provide any incentives. However, in Chinese culture, it is better to offer something, even a token, when asking people for a favor – and I viewed respondents' participation as a favor they were doing for me. Therefore, I prepared caps with the Tongji University logo as incentives. (Figure 16) The cap is not expensive – $8 RMB or about $1 USD. Another reason for selecting this specific reward was to enhance a sense of trust that people were not being sold a product and that as the study was connected to an academic, not a commercial, study, their participation would have some public benefit. In practice, a few people declined the reward but still participated in the survey, especially respondents in the high-end shopping malls.

Figure 16: Survey Reward (Pilot Survey)

- **Local Implementation**

One major part of this research is the first-hand data collection in Shanghai. The "all-in-one" process of survey design and implementation conducted by myself (with the help of professors in Tongji University and UC Davis) provided the advantage of asking questions pertinent to what I am most interested in. In the following, three issues of local implementation will be discussed.

Finalizing the Survey Team and Training

The survey team consisted of three people – one from UC Davis (myself) and two graduate students from the Automobile Marketing and Management School, Tongji University. Each team member was required to be familiar with the topic. The team was chosen to provide a mix of males and females and the ability to communicate in both Mandarin and the local Shanghainese dialect. The language ability was crucial in certain circumstances. For example, some elderly people we surveyed in old town Shanghai only spoke Shanghainese. Further, even Mandarin-speaking residents of Shanghai appeared to trust the survey-givers more when they noticed the team also speaking Shanghainese.

Team members were provided with a standard introduction to inform people who we were, what this was survey about, how long it would take to finish, and about the incentive. Also, the surveyors needed to be able to provide adequate but not leading (or misleading) assistance when people had questions or had trouble understanding a question. Finally, to minimize selection bias due to an affinity for certain types of people, the surveyors were instructed to use an "every fifth-person" rule – that is, to approach every fifth person passing by.

Finalizing the Time of Day for Survey Distribution

In Shanghai, the typical working day is from 9:00am to 5:00pm. I used two time slots – from 12:00pm to 2:00pm and 5:00pm to 7:00pm – during weekdays to capture people coming out for lunch and leaving work. I also surveyed on weekend afternoons, especially at shopping locations. The idea was to be at our locations during times of higher foot traffic.

Finalizing the Survey Spot

The team determined the general survey locations shown in Figure 15. However, upon arriving at each general location, we first spent up to 30 minutes looking around the survey area, discussing strategies for how to approach people, and finalizing the exact survey spots. For example, after arriving at the Renmin Square subway station, the team decided to survey near the bookstores at the station because we noticed that people there appeared to not be in a hurry. Similarly, once at the shopping malls, we observed the upper level of the mall was a better place to intercept people as they generally appeared to be less hurried than those on lower levels. Although surveying at those "less hurried"

spots may under-sample people with a high value of time or under high time pressure, our goal is to maximize the possibility of people taking our survey.

Final Survey

- **Sampling**

The population for the final survey was people over 18 years old, living in Shanghai, distributed along the motorization pathway[18]. Similar to the pilot survey, location-based convenience sampling was used, with oversampling of people at certain steps of the pathway. In particular, since car ownership is still relatively rare in Shanghai, the survey team went to Ford dealerships, a car show place, and a driving school to sample more car owners than could be sampled from pure on-street intercepts. As shown in Figure 17, there were 24 survey locations (13 on-street locations, 3 automobile dealerships, and 8 communities) all over the seven districts in Shanghai, including Yangpu, Jing'an, Huangpu, and Xuhui in the center city, as well as Minhang, Jiading and Pudong in suburban areas.

[18] People need to be at least 18 years old to have driver's license in Shanghai.

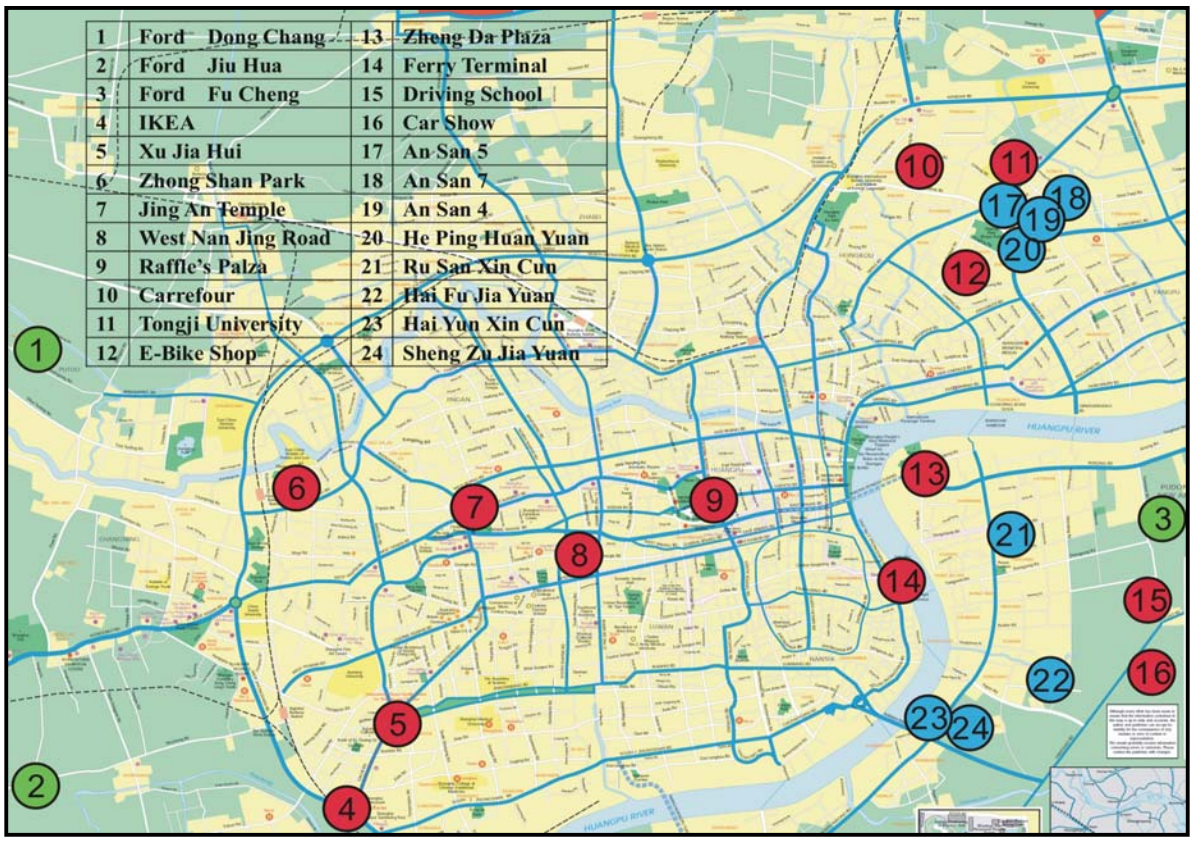


Figure 17: Survey Locations (Final Survey)

The final survey was administered between July 28th and October 27th, 2006. We distributed 1,569 questionnaires, and a total of 1,037 responses were collected including: 442 from on-street locations, 201 from the Ford dealerships, 316 from households, and 78 from a pure convenience sample of peer network. The overall response rate was 66 percent, which is slightly improved compared to the 63.5 percent response rate in the pilot survey. Details of locations, dates, and response rates by locations are shown in Table 18.

ID	Survey Location	Survey Date (2006)	Distribution Method	# of Surveys Distributed	# of Valid Responses	Response Rate (percent)
1	Ford Dong Chang	8/2 8/4 8/5	Dealership	53	40	75.5
2	Ford Jiu Hua	7/28 7/30 10/15 10/22	Dealership	285	133	47.0
3	Ford Fu Cheng	9/30	Dealership	59	28	47.5
4	IKEA	8/8	On-street	83	39	47.0
5	Xu Jia Hui	10/9	On-street	81	39	48.1
6	Zhong Shan Park Subway	10/22	On-street	82	40	48.8
7	Jing An Temple	9/23 9/24	On-street	71	40	56.3
8	West Nan Jing Rd.	8/5 8/12	On-street	69	38	55.0
9	Raffle's Plaza	9/24	On-street	50	35	70.0
10	Carrefour	8/13	On-street	66	41	62.1
11	Tongji University	10/13	On-street	68	40	58.8
12	E-bike Shop	10/26	On-street	15	8	53.3
13	Zheng Da Plaza	9/30	On-street	22	10	45.5
14	Ferry	10/24	On-street	69	35	50.7
15	Driving School	8/7	On-street	40	37	92.5

ID	Survey Location	Survey Date (2006)	Distribution Method	# of Survey Distributed	# of Valid Response	Response Rate (percent)
16	Car Show Place	8/20	On-street	58	40	69.0
17	An San 5	10/26	Household	40	40	100.0
18	An San 7	10/25	Household	40	40	100.0
19	An San 4	10/26	Household	40	40	100.0
20	He Ping Hua Yuan	10/25	Household	40	37	92.5
21	Ru San Xin Cun	9/12 9/22	Household	41	41	100.0
22	Hai Fu Jia Yuan	9/14 9/20	Household	41	41	100.0
23	Hai Yun Xin Cun	9/21	Household	37	37	100.0
24	Shen Zu Jia Yuan	9/21	Household	40	40	100.0
25	Peer Network	8/8 to 10/27	In person	79	78	98.7
26	All over Shanghai	9/1 to 10/27	Cell phone Message	41,754	78	0.18

Table 18: Survey Locations and Response Rate (Final Survey)

- **Distribution Method**

I used multiple methods to distribute the final questionnaire. I drew on lessons learned from the pilot survey to rewrite the questionnaire, to refine the on-street location selection, and to conduct additional experimentation with survey distribution techniques.

On-street: I re-visited some places from the pilot survey such as: Zheng Da Plaza and Raffle's Plaza, as these are near major subway lines and have high foot traffic. Also, I applied experiences from the pilot survey to determine additional locations that would improve the response rate. For example, Jing An temple was selected because there was a park nearby. Thus, it was considered a comfortable and convenient location and people in that area appeared to be less hurried.

Car dealership: One way I ensured the overall sample contained (enough) car owners was to distribute the final questionnaire at automotive sales dealerships. With cooperation from Ford China, the survey team was able to distribute questionnaires to car owners waiting while their cars were repaired or serviced. The waiting area was a comfortable and convenient place, and people tended to show trust toward us when they were informed that the survey was supported by Ford China. Note that other car-related sample locations were designed to expand the sample beyond owners of Ford vehicles, for example, the driving school and automotive show.

Household: Pre-arrangement to complete a questionnaire was important to ensure the success of the household surveys, since people need to have more trust to let survey-givers into their houses. Neighborhood Committees, which function like a rental or property management office, were enlisted to introduce us to households. The Neighborhood Committee is a semi-governmental organization and the smallest unit of the Communist Party in China. The professor I cooperated with has good connections with the Neighborhood Committee and helped to arrange the survey. With their aid, we surveyed households in eight different communities, attaining high response rates.

Peer network: Distributing the questionnaire in person to a network of peers of the survey-givers – a form of convenience sampling – was an efficient approach. We were able to solicit questionnaires from many people in a short time and the response rate was high (98.7 percent). However, to avoid possible selection bias in the overall sample, only a small portion of the surveys (n = 79) was distributed through the peer network.

Cell Phone Message/On-line: In addition to the above distribution methods, an on-line survey (www.china.v33.org) was implemented to expand the sample and test the current viability of a common survey research sampling medium in the U.S. A cell phone text message was the major tool to recruit people for the on-line survey. Cell phone users were contacted via random digit dialing covering all registered cell phone users in Shanghai metropolitan area[19]. However, it was not effective. The research team sent out 41,754 text messages during a 2-month time window, but only 78 people (in addition to the above 1,037 respondents) participated in the on-line survey. Considering the on-line survey as a fundamentally different survey type as well as its very low response rate, those 78 cases are excluded from analyses in this dissertation, except for the discussion of survey participation.

Nevertheless, the potential advantages of an on-line survey are that people may complete it in an anonymous situation and at a time of their choosing. The cell phone text message recruiting was originally considered efficient as it can reach a large group of people in a short time. To gain more trust from respondents, I posted the questionnaire with the official cooperation documents between UC Davis and Tongji University on the survey

[19] I purchased the machine to do the random digital dialing for about 2,000 RMB.

website (www.china.v33.org). However, the low response rate suggested several potential problems. The text message may have been lost in the large number of text messages Shanghai cell phone users may receive, and even if read may have been too impersonal and untrustworthy. The lack of personal access to computers in peoples' homes and work places could be another factor that limited people's willingness or ability to participate.

- **Questionnaire Design**

The questionnaire contained 73 questions organized into six parts (See Appendix for the questionnaire of the final survey):

Part 1 asked for background on vehicle use and purchase. Respondents were asked eight questions about what kind of travel tools they own, the most expensive currently-owned vehicle, the use conditions, purchase cost, and operation cost, etc. A question about their motorization pathway was also included:

> *Please sort the sequence of the travel means you have used from the very past (in your memory) till now. Please fill in the numbers inside the boxes.*
> *(Skip the box if you never use that travel mean)*
>
> **Ex:** [1] Bicycle, [2] Motorcycle, [3] Car = Bicycle → Motorcycle → Car
>
> ☐ Bicycle
> ☐ Walk
> ☐ Personal Car
> ☐ Public Transportation
> ☐ Auxiliary Power Vehicle[20], Motorcycle
> ☐ Taxi, Rented car

[20] In my survey, the original wording for motorized two-wheeler is "auxiliary power vehicle".

☐ Shared Company's Car

☐ NO, there is no such "pathway". (WHY? _____)

Basically, respondents sorted their personal history of travel means to trace their motorization pathway, or absence of one. More detailed discussion about the motorization pathway analysis will be presented in Chapter 4.

Part 2 asked respondents to rate six attributes (status, speed, availability, capacity, price, comfort) of nine different travel means, including: walk, bicycle, motorized two-wheeler, motorcycle, taxi, rented car, public transportation, company car, and private car.

Part 3 asked about exogenous environmental factors affecting peoples' decisions whether to use or buy a travel mode, i.e., to advance a step on the motorization pathway. This part included 18 questions covering issues about policy, regulations, social environment, etc.

Part 4 asked another 18 questions to triangulate six underlying lifestyles hypothesized to be related to vehicle purchase in Shanghai. The six lifestyles are: "status-seeking," "bandwagon," "happiness as the first priority," "freedom and control of life," "environmental concern," and "family-oriented."

Part 5 asked 12 questions regarding demographic and socio-economic information about the respondents, including: age, gender, education, income (personal and family), occupation, possession of driver's license, residence area and duration, and experience of living in foreign country.

Part 6 asked 11 questions about respondents' participation in the survey itself, including what aspects of the survey process affected their willingness to initiate and complete the questionnaire.

- **Incentive**

To attract people to our survey, I offered rewards that varied by the method of questionnaire distribution. In each case, I attempted to match the incentive to the people and the setting in which I surveyed them. For example, people who took the survey immediately (including on-street, at an automobile dealership, and in-home) were given taxi coupons worth 20 RMB (i.e., $2.5 USD) (Figure 18). People who finished the questionnaire on-line were eligible to enter a drawing for an i-Pod with value of 2,000 RMB (i.e., $250 USD) (Figure 19). Considering that people taking the on-line survey may be mostly students or young professionals, the chance to win an i-pod was considered an attractive reward for them. The total cost of incentives was about $2,000 USD – we gave out two i-pods, and about 60 percent of our respondents eventually took the taxi coupons as rewards.

Figure 18: Survey Reward (Taxi Coupon, Final Survey)

Figure 19: Survey Reward (i-Pod, Final Survey)

- **Local Implementation**

Finalizing the Survey Team and Training

The accuracy of the data is the goal. Especially in survey methods involving direct personal interaction with survey-givers, quality depends on whether survey-givers are responsible and professional in their approach to potential respondents and in their assistance to interviewees. Based on these requirements, I selected a group from transportation engineering students at Tongji University. These students are developing professional backgrounds in transportation. They understand the need for high quality data, and they are full of passion.

I distributed training materials to survey-givers and trained them as teams. Various scenarios were simulated in the training. They were told to be polite, patient, warmhearted and persevering. All the survey work should be finished in a harmonious and friendly manner. As in the pilot survey, the team members were trained to avoid

selection bias, for example by applying the "fifth-person" rule in approaching potential respondents.

Finalizing the Time of Day for Survey Distribution

Choosing a proper time to distribute questionnaires differed for the various means of distributing questionnaires. For example, as the on-street survey was conducted outdoors, successful implementation might be dependent on weather, especially in the summer time. So I chose weekend late afternoons, usually 5:00pm to 7:00pm, when it was not so hot in Shanghai for the on-street survey. As a trade-off, surveying during those specific time windows might introduce some potential sampling bias; however, our goal was to increase the response rate and facilitate the survey. As another example, during weekday work hours fewer people would be at home for the in-home household survey, therefore the team used weekdays from 7:00pm to 9:00pm for the household survey.

Finalizing the Survey Spot

As described above, on-street locations, car dealerships, and respondents' households were three major distribution methods in the final study. Due to the fact I had only one automotive sponsor for the study, I was able to distribute questionnaires only at the three Ford dealerships in Shanghai. Nevertheless, the following describes some guidelines about finalizing the survey spots for the on-street and household distribution.

On-street: Similar to the pilot survey, I first considered locations that would increase the probability I would be able to sample for motorization-segments. This meant locations that had both certain types of people and many of them who could be easily approached.

Therefore, I chose some shopping malls and supermarkets in Shanghai. Within this type of location, my selection procedure is illustrated by this example of how I chose a Carrefour store (the Quyang branch store). First, I reviewed all the addresses of Carrefour supermarkets in Shanghai. Carrefour supermarkets attract many people in their neighborhoods. In fall 2006, there were ten Carrefour markets in Shanghai: Lianyang, Wanli, Jinqiao, Qibao, Gubei, Wuning, Quyang, Nanfang, Baoshan and Zhongshan Park. The Lianyang, Gubei, Jinqiao and Wuning stores are located in neighborhoods where foreigners are the majority of residents, so these would not be good places to survey Chinese people. The Qibao and Zhongshan Park stores are close to Shanghai's most famous tour sites – high traffic, but many travelers from outside Shanghai, both Chinese and foreign. I judged the Nanfang, Baoshan and Jinqiao stores were not big enough to have enough foot traffic to make them viable locations for intercept surveys. The best option was the Quyang store because the residents are mostly local Shanghainese people and it was a large store with many customers. No government permission was needed for distributing surveys on-street.

Household: There are 17.4 million residents of Shanghai living in numerous residential districts spread over 6,341km^2 (Shanghai Bureau of Statistics, 2006). Sampling specific communities from within this large city was crucial to our research. I attempted to sample according to differences in household income cross-classified by access to public transportation. These were expected to affect the respondents' ability to buy automobiles (moving to the high cost, privatized end of the motorization pathway) and to use transit (which might facilitate some people moving from non-motorized modes to collective, motorized modes and to moderate the effect of higher income people moving to high-cost,

private modes). Locations chosen according to these criteria are listed in Table 19, and the identifying numbers correspond to those in Figure 17.

	Near Transit: less than 5 minutes walking distance to subway	**Far from Transit:** more than 5 minutes walking distance to subway
Low Income	Location 21	Locations 17, 18, 19, 23
Mid- to High Income	Location 24	Locations 20, 22

Table 19: Household Survey Locations (Final Survey)

Since it is not possible to know household income without first asking respondents for this information, I used publicly available housing prices as a proxy measure of wealth that would differentiate people according to their financial ability to buy an automobile. Suggested by the Tongji professors I worked with, people whose house price was below 8,000 RMB/m^2 were classified as low level of wealth; people whose house price was between 8,000 RMB/m^2 and 15,000 RMB/m^2 were classified in a middle level of wealth; and people whose house price was higher than 15,000 RMB/m^2 were classified in the high level of wealth. For transit access, I used proximity to subway as the criterion. I considered a place within a five-minute walking distance to the subway as "near" transit.

Lesson Learned from Local Implementation

The following describes several important lessons from the pilot and final surveys.

- **Lesson from Pilot Survey**

Lesson #1: "Dream life" is vague for most people

In pilot survey, questions about the future proved the most difficult for the respondents. I wanted people to imagine a "dream life" as a way for them to think about how they would want their life to be, including, how they would travel. A series of questions asked people to imagine a scenario of their dream life including income, vehicle, etc. 11 percent of respondents declined to answer at all; another 54 percent provided either vague or partial answers. Frequently, respondents needed further explanation. Even after the explanation, common comments were: "I have never thought about that," and "This does not relate to me, since I don't think I can change my life." The difficulty people had in answering these questions suggested this line of questioning would not be useful in the final survey. The comments suggest that many people in Shanghai may not be ready yet to imagine a future very different from their past and present, raising difficulties for studying many hypothetical topics.

Lesson #2: Trust, convenience, and comfort are keys to success

Based on the pilot survey experience, trust is the number one factor affecting people's willingness to participate. That is why I got the lowest response rate (40.7 percent) in the Renmin Square subway station. It is not only because people are busy, but also people tended to confuse the survey team with street venders, or possibly scams. When potential respondents don't trust the surveyors, we usually cannot stop them to ask for their participation. To remedy this situation, surveyors began to show their student IDs when approaching people. Besides, a convenient and comfortable environment will yield a better response rate. A convenient environment means a place where people can quickly and easily complete their questionnaire. For example, a street corner with public chairs or tables is a more convenient location for people to take the survey. A comfortable

environment could be a place with air-conditioning in the hot summer time. However, it can also be a location where people feel mentally comfortable. Some of our respondents stated that a place with more privacy was a more comfortable place – for example, the little bookstore inside the subway station.

Lesson #3: "Getting to know the place" is also important

In addition to the experience gained directly from the design and implementation of the pilot survey, I learned a lot by traveling around Shanghai. As a researcher who conducts survey research in a new place for the first time, orientation to the setting is important to develop a basic sense of the people and the area. The implementation of on-street interviewing in the pilot phase helped me finalize the sample design and location selection for the final phase.

- **Lesson from Final Survey**

Lesson #1: A complicated and long survey is challenging

The most difficult and critical step was to design a questionnaire that I could reasonably expect people to complete in the generally short time available. However, as there are many aspects of a motorization pathway I want to research, the questionnaire necessarily requires some complexity. I attempted to balance time demands on respondents and questionnaire complexity so that people would agree to participate and not lose patience. However, many people refused to participate when they realized the survey was seven double-sided pages.

Moreover, some individual questions proved difficult. A few of these are classic problems

in questionnaire design, i.e., respondents confusing rating and ranking tasks. For example, in Part 2 of the questionnaire I asked people to rate their agreement or disagreement with how expensive they perceive different travel modes to be. A table listing nine travel modes, each of which I wanted them to rate followed this. Unfortunately, many people ranked the items instead.

Besides, I did not separate the category of public transportation into bus and metro (including subway and light rail). This may have led to some misunderstanding of this rating task, because in Shanghai the metro system and bus system are not organized and operated by a single centralized manager. The price of metro is generally much higher than bus. The ratings of travel speed (asked in another question) of transit may also be affected by the income of respondents, since different income groups may have been responding to different "public transportation" systems. Thus despite our exploration of Shanghai as part of the survey, I did not accurately characterize transit in the final survey.

Lesson #2: Anonymity/confidentiality, authorization, and study topic are the top three factors affecting people's motivation to participate

As mentioned, Part 6 of the survey is designed specifically to identify the factors affecting people's motivation to participate in the survey. The original Part 6 questions are shown in Table 20.

	How important to you is each of the following affecting your motivation of taking a survey?	**Not at all important**	**Slightly Important**	**Medium Important (Neutral)**	**Moderately Important**	**Extremely Important**
1.	How much time it will take	☐	☐	☐	☐	☐
2.	The topic (e.g., commercial *vs*. academic)	☐	☐	☐	☐	☐
3.	Authorization letter	☐	☐	☐	☐	☐
4.	My answer will be anonymous or kept confidential.	☐	☐	☐	☐	☐
5.	Guaranteed reward (non-cash)	☐	☐	☐	☐	☐
6.	Guaranteed reward (cash)	☐	☐	☐	☐	☐
7.	Drawing reward (non-cash)	☐	☐	☐	☐	☐
8.	Drawing reward (cash)	☐	☐	☐	☐	☐
9.	My friend refers me this survey.	☐	☐	☐	☐	☐
10.	Other, please specify_____					

Table 20: Original Questions about Factors of Survey Participation

According to the survey results of 1,037 respondents (those 78 on-line survey participants were excluded), anonymity/confidentiality, authorization, and study topic were the top three considerations in whether people initiated and completed a questionnaire. 55 percent of respondents considered that "my answer will be anonymous and confidential" to be moderately or extremely important. Similarly, 46 percent said that the "authorization letter" and 41 percent said "the survey topic" were moderately or extremely important.

Findings based on Survey Results

- **Factors of Survey Participation: Face-to-Face vs. On-line Survey**

Important and Not Important Factors

The final survey was conducted face-to-face interview. 1,037 people responded on-street, at the car dealership, household, or by peer network. A further 78 people were recruited by cell phone text messages to an on-line version of the questionnaire. The two surveys are compared in Table 21.

SURVERY TYPE	Face-to-Face	On-line
Sampling	Location-based and other convenience sampling	Random digital dialing
Survey Distribution	On-street, auto dealership, household, in person (peer network)	Cell phone text message
Survey Collection	● Same method as survey distribution. ● Distribution and collection is usually one-step process	● Internet ● Distribution and collection can be separate steps
In-survey Process	● Less privacy, less time flexibility ● More assistance and interaction	● More privacy, more time flexibility ● Less assistance and interaction

Table 21: Comparison between Face-to-Face and On-line Survey

Three important factors affecting the willingness to participate of 1,037 people taking part in the face-to-face interviews have been discussed briefly as the lesson of final survey implementation. However, I believe that people attracted by cell phone text messages to an on-line survey might view other factors as important to their participation. Table 22 and Table 23 show the importance distribution (on a five-point Likert scale) of nine factors affecting participation for face-to-face and on-line surveys. The values in the

cells are the percentage of sample (1,037 for face-to-face survey, 78 for on-line survey). The highest percentage for each factor is in bold-face. As I expected, "medium important/neutral" is the most frequent answer for many questions, especially for the people taking face-to-face survey. This response pattern, to certain extent, reflects the Chinese culture, which encourages people to be conservative, or, to be in the "middle". However, for the on-line survey respondents, more people express their attitude without placing themselves in the middle. "Medium important/neutral" is the most frequent answer to five questions for the on-line respondents compared to being most frequent for seven questions for people taking survey face-to-face. Besides, no one skipped a single question in the on-line survey, although the "no comment/skip" was an option provided equally for both face-to-face and on-line survey participants. This discussion suggests that the on-line survey respondents, compared to people taking the survey face-to-face, are more willing to express a "not-in-the-middle" attitude.

How important to you is each of the following affecting your motivation of taking a survey?	N/A (no comment)	Not at all important	Slightly Important	Medium Important (Neutral)	Moderately Important	Extremely Important
How much time it will take	2.16%	1.75%	16.87%	**43.59%**	26.59%	9.04%
The topic, e.g., commercial vs. academic study	2.70%	0.94%	15.92%	**39.54%**	34.55%	6.34%
Authorization letter	3.91%	1.08%	12.55%	36.57%	**37.92%**	7.96%
My answer will be anonymous and confidential.	2.16%	1.48%	12.42%	28.88%	**35.22%**	19.84%
Guaranteed reward (non-cash)	2.56%	7.15%	34.14%	**42.78%**	9.99%	3.37%

Guaranteed reward (cash)	2.29%	7.56%	34.01%	**41.70%**	11.47%	2.97%
Drawing reward (non-cash)	2.16%	7.56%	39.00%	**42.65%**	6.75%	1.89%
Drawing reward (cash)	2.83%	8.10%	37.92%	**41.70%**	7.69%	1.75%
My friend refers me this survey.	4.72%	6.34%	34.28%	**41.03%**	11.20%	2.43%

Table 22: Factors of Survey Participation (Face-to-Face Survey Type, 1037 Respondents, Five-Point Importance Scale)

How important to you is each of the following affecting your motivation of taking a survey?	N/A (no comment)	Not at all important	Slightly Important	Medium Important (Neutral)	Moderately Important	Extremely Important
How much time it will take	0.00%	3.85%	12.82%	29.49%	**32.05%**	21.79%
The topic, e.g., commercial vs. academic study	0.00%	2.56%	7.69%	33.33%	**35.90%**	20.51%
Authorization letter	0.00%	3.85%	6.41%	32.05%	**34.62%**	23.08%
My answer will be anonymous and confidential.	0.00%	1.28%	10.26%	28.21%	29.49%	**30.77%**
Guaranteed reward (non-cash)	0.00%	6.41%	11.54%	**48.72%**	20.51%	12.82%
Guaranteed reward (cash)	0.00%	7.69%	7.69%	**44.87%**	28.21%	11.54%
Drawing reward (non-cash)	0.00%	12.82%	7.69%	**55.13%**	16.67%	7.69%
Drawing reward (cash)	0.00%	12.82%	6.41%	**57.69%**	12.82%	10.26%
My friend refers me this survey.	0.00%	7.69%	11.54%	**39.74%**	26.92%	14.10%

Table 23: Factors of Survey Participation (On-line Survey Type, 78 Respondents, Five-Point Importance Scale)

To identify the important and not important factors of survey participation, I collapsed the five importance scales into three – "not-at-all to slightly important" (considered basically as not-important), "medium important/neutral", and "moderately to extremely important" (considered basically as important). As seen in Table 24, the top three important factors of participation (within red box) for face-to-face survey participants are: "confidentiality" (55.06 percent), "authorization letter" (45.88 percent), and "survey topic" (40.89 percent)[21]. Likely, the top three not-important factors of participation (within blue box) for the same group are: "drawing reward (non-cash)" (46.56 percent), "drawing reward (cash)" (46.02 percent), and "guaranteed reward (cash)" (41.57 percent)[22].

How important to you is each of the following affecting your motivation of taking a survey?	N/A (no comment)	Not at all important to Slightly Important (-)	Medium Important (Neutral)	Moderately Important to Extremely Important (+)	*Relative Importance Check*
How much time it will take	2.16%	18.62%	**43.59%**	35.63%	- < +
The topic, e.g., commercial vs. academic study	2.70%	16.86%	39.54%	**40.89%**	- < +
Authorization letter	3.91%	13.63%	36.57%	**45.88%**	- < +
My answer will be anonymous and confidential	2.16%	13.90%	28.88%	**55.06%**	- < +
Guaranteed reward (non-cash)	2.56%	41.29%	**42.78%**	13.36%	- > +
Guaranteed reward (cash)	2.29%	41.57%	**41.70%**	14.44%	- > +

[21] I determined the top "important factor" by comparing the combined percentage of "moderately to extremely important" across all nine factors listed.
[22] Top "not-important" factors were identified by comparing the combined percentage of "not-at-all to slightly important" across all nine factors listed.

Drawing reward (non-cash)	2.16%	**46.56%**	42.65%	8.64%	- > +
Drawing reward (cash)	2.83%	**46.02%**	41.70%	9.44%	- > +
My friend refers me this survey.	4.72%	40.62%	**41.03%**	13.63%	- > +

Table 24: Factors of Survey Participation (Face-to-Face Survey Type, 1037 Respondents, Three-Point Importance Scale)

How important to you is each of the following affecting your motivation of taking a survey?	N/A (no comment)	Not at all important to Slightly Important (-)	Medium Important (Neutral)	Moderately Important to Extremely Important (+)	*Relative Importance Check*
How much time it will take	0.00%	16.67%	29.49%	**53.84%**	- < +
The topic, e.g., commercial vs. academic study	0.00%	10.25%	33.33%	**56.41%**	- < +
Authorization letter	0.00%	10.26%	32.05%	**57.70%**	- < +
My answer will be anonymous and confidential	0.00%	11.54%	28.21%	**60.26%**	- < +
Guaranteed reward (non-cash)	0.00%	17.95%	**48.72%**	33.33%	- < +
Guaranteed reward (cash)	0.00%	15.38%	**44.87%**	39.75%	- < +
Drawing reward (non-cash)	0.00%	20.51%	**55.13%**	24.36%	- < +
Drawing reward (cash)	0.00%	19.23%	**57.69%**	23.08%	- < +

My friend refers me this survey.	0.00%	19.23%	**39.74%**	41.02%	- < +

Table 25: Factors of Survey Participation (On-line Survey Type, 78 Respondents, Three-Point Importance Scale)

In Table 25, the top three important factors of participation (within red box) for on-line survey respondents are: "confidentiality" (60.26 percent), "authorization letter" (57.7 percent), and "survey topic" (56.41 percent) – the same three factors as for face-to-face survey participants. The top three not-important factors of participation (within blue box) for on-line survey respondents are: "drawing reward (non-cash)" (20.51 percent), "drawing reward (cash)" (19.23 percent), and "friends' referral" (19.23 percent). In addition, in order to see the relative importance, I compared the combined percentage of "not-at-all to slightly important (negative)" with "moderately to extremely important (positive)" and found out that, for *each* question, more positive answers than negative ones were provided by people taking on-line survey.

To summarize, on-line respondents, being surveyed without personal contact but with privacy and freedom, are more willing to reveal "positive" attitudes. However, people involved in conventional face-to-face interviews, tend to skip questions or show "negative to middle" attitude. Besides, anonymity/confidentiality, authorization, and study topic were the top three important factors for both on-street and on-line respondents. Drawing rewards (cash and non-cash) were two not-important factors for both groups – no "guaranteed" reward could be one reason. However, it is also possible that people don't want to admit they took the survey for a reward (considered *not* socially-desirable in China). According to the initial findings, I assumed that the importance distribution

differs by the survey type, and chi-square analyses (for each single factor) were conducted to test my hypothesis, as in the following:

Time Cost vs. Survey Type

"Time cost" is the first factor of participation I test. "Not-at-all important" and "slightly important" are combined to reduce the number of cells with an expected count of less than five. However, the trade-off of this combination is the decrease of sensitivity of the original question at the negative end. The Pearson chi-squared test suggests that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). In Table 26, "neutral" is the most frequent answer (44.7 percent, highlighted in red) for people taking face-to-face survey. "Moderately important" (32.1 percent) is the most popular selection for on-line respondents, although the "neutral" also has similar frequency (29.5 percent) in that group.

			TIME COST				
			Not at all ~ Slightly Important	Neutral	Moderately Important	Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	211	**459**	269	88	1027
		% within SURVEY TYPE	20.5%	**44.7%**	26.2%	8.6%	100.0%
	On-line	Count	13	23	**25**	17	78
		% within SURVEY TYPE	16.7%	29.5%	**32.1%**	21.8%	100.0%
Total		Count	224	**482**	294	105	1105
		% within SURVEY TYPE	20.3%	**43.6%**	26.6%	9.5%	100.0%

Table 26: Cross-tabulation: Survey Type x Time Cost

Research Topic vs. Survey Type

Similarly, "not-at-all important" and "slightly important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test suggests that there are significant differences in the distributions of importance scale

across survey types (p-value = 0.000 < 0.05). Again, "neutral" (41.1 percent) is the most frequent answer for people taking face-to-face survey, and "moderately important" (35.9 percent) is the most frequent answer for on-line respondents (Table 27).

			RESEARCH TOPIC				
			Not at all ~ Slightly Important	Neutral	Moderately Important	Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	180	**419**	357	64	1020
		% within SURVEY TYPE	17.6%	**41.1%**	35.0%	6.3%	100.0%
	On-line	Count	8	26	**28**	16	78
		% within SURVEY TYPE	10.3%	33.3%	**35.9%**	20.5%	100.0%
Total		Count	188	**445**	385	80	1098
		% within SURVEY TYPE	17.1%	**40.5%**	35.1%	7.3%	100.0%

Table 27: Cross-tabulation: Survey Type x Research Topic

Authorization Letter vs. Survey Type

"Not-at-all important" and "slightly important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test indicates that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). In Table 28, "neutral" (39.1 percent) is the most frequent answer for face-to-face survey group, and "moderately important" (34.6 percent) is the top answer for on-line respondents.

			AUTHORIZATION				
			Not at all ~ Slightly Important	Neutral	Moderately Important	Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	158	**396**	381	79	1014
		% within SURVEY TYPE	15.6%	**39.1%**	37.6%	7.8%	100.0%
	On-line	Count	8	25	**27**	18	78
		% within SURVEY TYPE	10.3%	32.1%	**34.6%**	23.1%	100.0%
Total		Count	166	**421**	408	97	1092
		% within SURVEY TYPE	15.2%	**38.6%**	37.4%	8.9%	100.0%

Table 28: Cross-tabulation: Survey Type x Authorization

Confidentiality vs. Survey Type

Similarly, "not-at-all important" and "slightly important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test shows that there are significant differences in the distributions of importance scale across survey types (p-value = 0.021 < 0.05). In Table 29, "extremely important" is the top choice for on-line respondents; however, "moderately important" is the most frequent choice for people taking the survey face-to-face. In contrast to the previous three cases, most people, whether in the face-to-face or the on-line survey group, consider confidentiality as important "in general".

			CONFIDENTIALITY				
			Not at all ~ Slightly Important	Neutral	Moderately Important	Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	184	309	**356**	176	1025
		% within SURVEY TYPE	18.0%	30.1%	**34.7%**	17.2%	100.0%
	On-line	Count	9	22	23	**24**	78
		% within SURVEY TYPE	11.5%	28.2%	29.5%	**30.8%**	100.0%
Total		Count	193	331	**379**	200	1103
		% within SURVEY TYPE	17.5%	30.0%	**34.4%**	18.1%	100.0%

Table 29: Cross-tabulation: Survey Type x Confidentiality

Reward (Guaranteed, Non-cash) vs. Survey Type

In contrast to the previous cases, "moderately important" and "extremely important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test suggests that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). "Neutral" is the most frequent answer for both face-to-face (43.1 percent) and on-line survey groups (48.7 percent). However, the second most frequent answer is "slightly important" (37.7 percent) for the face-to-face group but "moderately to extremely important" (33.3 percent) for the

on-line survey group. (Table 30)

			REWARD (GUARANTEED, NON-CASH)				
			Not at all Important	Slightly Important	Neutral	Moderately ~ Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	74	386	**441**	122	1023
		% within SURVEY TYPE	7.2%	37.7%	**43.1%**	11.9%	100.0%
	On-line	Count	5	9	**38**	26	78
		% within SURVEY TYPE	6.4%	11.5%	**48.7%**	33.3%	100.0%
Total		Count	79	395	**479**	148	1101
		% within SURVEY TYPE	7.2%	35.9%	**43.5%**	13.4%	100.0%

Table 30: Cross-tabulation: Survey Type x Reward (Guaranteed, Non-Cash)

Reward (Guaranteed, Cash) vs. Survey Type

"Moderately important" and "extremely important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test shows that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). Similar to previous case, "neutral" is the most frequent answer for both face-to-face (40.3 percent) and on-line survey groups (44.9 percent). (Table 31)

			REWARD (GUARANTEED, CASH)				
			Not at all Important	Slightly Important	Neutral	Moderately ~ Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	78	389	**413**	146	1026
		% within SURVEY TYPE	7.6%	37.9%	**40.3%**	14.2%	100.0%
	On-line	Count	6	6	**35**	31	78
		% within SURVEY TYPE	7.7%	7.7%	**44.9%**	39.7%	100.0%
Total		Count	84	395	**448**	177	1104
		% within SURVEY TYPE	7.6%	35.8%	**40.6%**	16.0%	100.0%

Table 31: Cross-tabulation: Survey Type x Reward (Guaranteed, Cash)

Reward (Drawing, Non-cash) vs. Survey Type

"Moderately important" and "extremely important" are combined to reduce the number

of cells with an expected count of less than five. The Pearson chi-squared test indicates that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). "Neutral" is the most frequent answer for both face-to-face (41.6 percent) and on-line groups (55.1 percent) (Table 32). However, compared to people taking survey in person, more people in the on-line group consider the drawing non-cash reward as "not-at-all important" (12.8 percent).

			REWARD (DRAWING, NON-CASH)				
			Not at all Important	Slightly Important	Neutral	Moderately ~ Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	85	420	**427**	94	1026
		% within SURVEY TYPE	8.3%	40.9%	**41.6%**	9.2%	100.0%
	On-line	Count	10	6	**43**	19	78
		% within SURVEY TYPE	12.8%	7.7%	**55.1%**	24.4%	100.0%
Total		Count	95	426	**470**	113	1104
		% within SURVEY TYPE	8.6%	38.6%	**42.6%**	10.2%	100.0%

Table 32: Cross-tabulation: Survey Type x Reward (Drawing, Non-Cash)

Reward (Drawing, Cash) vs. Survey Type

Likely, "moderately important" and "extremely important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test shows that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). "Neutral" is the top selection for both face-to-face (42.1 percent) and on-line survey groups (57.7 percent) (Table 33). Besides, a noticeable 12.8 percent of on-line survey respondents consider the drawing cash reward as not-at-all important compared to 7.9 percent in the group taking face-to-face survey.

			REWARD (DRAWING, CASH)				
			Not at all Important	Slightly Important	Neutral	Moderately ~ Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	81	412	**430**	98	1021
		% within SURVEY TYPE	7.9%	40.4%	**42.1%**	9.6%	100.0%
	On-line	Count	10	5	**45**	18	78
		% within SURVEY TYPE	12.8%	6.4%	**57.7%**	23.1%	100.0%
Total		Count	91	417	**475**	116	1099
		% within SURVEY TYPE	8.3%	37.9%	**43.2%**	10.6%	100.0%

Table 33: Cross-tabulation: Survey Type x Reward (Drawing, Cash)

Friend's Referral vs. Survey Type

"Friend's Referral" is the last factor I test. "Moderately important" and "extremely important" are combined to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test suggests that there are significant differences in the distributions of importance scale across survey types (p-value = 0.000 < 0.05). "Neutral" is the most frequent answer for both face-to-face (41.6 percent) and on-line survey groups (39.7 percent). However, the second most frequent answer is "slightly important" (35.9 percent) for the face-to-face group but "moderately to extremely important" for the on-line survey group (41 percent) (Table 34).

			FRIEND'S REFERRAL				
			Not at all Important	Slightly Important	Neutral	Moderately ~ Extremely Important	Total
SURVEY TYPE	Face-to-Face	Count	70	362	**419**	157	1008
		% within SURVEY TYPE	6.9%	35.9%	**41.6%**	15.6%	100.0%
	On-line	Count	6	9	**31**	32	78
		% within SURVEY TYPE	7.7%	11.5%	**39.7%**	41.0%	100.0%
Total		Count	76	371	**450**	189	1086
		% within SURVEY TYPE	7.0%	34.2%	**41.4%**	17.4%	100.0%

Table 34: Cross-tabulation: Survey Type x Friend's Referral

- **Effectiveness of Sampling: Revisiting the Location-based Idea**

Reclassifying the Survey Location

As discussed, in the final survey, I did not implement a probability-based sample. Rather, in order to insure enough respondents in each step of the motorization pathway, I conducted the survey at several specific locations – a location-based variant of convenience sampling. That is, many of my survey locations are directly (or indirectly) related to vehicle purchase or use. For example, the auto dealership is for interviewing car owners, and the ferry place is for surveying the motorized two-wheeler or motorcycle riders. In Table 35, I reclassify the 26 final survey locations into six types as related to vehicle purchase or use.

LOCATION TYPE	Survey Locations
On-street (general)	IKEA, Xu Jia Hui, Zhong Shan Park Subway, Jing An Temple, West Nan Jing Rd., Raffle's Plaza, Carrefour, Zheng Da Plaza, Peer Network
Bicycle-related	Tongji University
Motorized Two-wheeler/ Motorcycle-related	Ferry, E-bike Shop
Car-related	Ford Dong Chang, Ford Jiu Hua, Ford Fu Cheng, Driving School, Car Show Place
Household	An San 4, An San 5, An San 7, He Ping Hua Yuan, Ru San Xin Cun, Hai Fu Jia Yuan, Hai Yun Xin Cun, Shen Zu Jia Yuan
On-line	All over Shanghai Metropolitan Area

Table 35: Reclassification of Location Type

In this research, the "effectiveness" of sampling is defined in terms of how well (or diversely) the sample is distributed across different motorization pathway steps.

Therefore, “How effective was my sampling idea? For example, did I really sample more car owners in auto dealerships?” are questions of interest. According to the results of the pilot survey, 75 percent of people surveyed in the ferry terminal or on the ferry were people using motorcycles or motorized two-wheelers. In the following, cross-tabulation and chi-square analyses are conducted to test the relationship between final survey location (six types) and: “most expensive vehicle owned”, “most frequently used travel means (weekday)”, and “most frequently used travel means (weekend).

Most Expensive Vehicle Owned vs. Location Type

The original choices for most expensive vehicle in the final survey are: no vehicle, bicycle, motorized two-wheeler, motorcycle and car. However, to reduce the number of cells with an expected count of less than five, “motorized two-wheeler” and “motorcycle” are combined. The Pearson chi-squared test suggests that there are significant differences in the distributions of most expensive vehicle owned across location types (p-value = $0.000 < 0.05$). Based on Table 36, not to my surprise, “bicycle (as the most expensive vehicle)” is over-represented (61.4 percent) in the sample of bicycle-related locations. However, “bicycle” is also the most frequent answer for people sampled from: general on-street locations (29.3 percent), household (42.3 percent), and on-line (38.7 percent). Besides, 92.7 percent of people sampled from motorized two-wheeler/motorcycle-related locations reported “motorized two-wheeler/motorcycle” as their most expensive vehicle owned. 61.8 percent of sample at car-related locations reported “car” as the most expensive vehicle owned. The above findings basically confirm the effectiveness of my location-based sampling scheme to interview people owning specific types of vehicles.

			MOST EXPENSIVE VEHICLE				
			No Vehicle	Bicycle	Motorized Two-wheeler /Motorcycle	Car	Total
LOCATION TYPE	On-street (general)	Count	84	**103**	89	76	352
		% within LOCATION TYP	23.9%	**29.3%**	25.3%	21.6%	100.0%
	Bicycle-related	Count	7	**27**	5	5	44
		% within LOCATION TYP	15.9%	**61.4%**	11.4%	11.4%	100.0%
	Motorized Two-wheeler/Motorcycle-relate	Count	0	2	**38**	1	41
		% within LOCATION TYP	.0%	4.9%	**92.7%**	2.4%	100.0%
	Car-related	Count	39	20	43	**165**	267
		% within LOCATION TYP	14.6%	7.5%	16.1%	**61.8%**	100.0%
	Household	Count	89	**129**	55	32	305
		% within LOCATION TYP	29.2%	**42.3%**	18.0%	10.5%	100.0%
	On-line	Count	26	**29**	5	15	75
		% within LOCATION TYP	34.7%	**38.7%**	6.7%	20.0%	100.0%
Total		Count	245	**310**	235	294	1084
		% within LOCATION TYP	22.6%	**28.6%**	21.7%	27.1%	100.0%

Table 36: Cross-tabulation: Location Type x Most Expensive Vehicle

Most Frequently Used Travel Means (weekday) vs. Location Type

Originally, there are a total of nine choices for the question of most frequently used travel means (weekday). To reduce the number of cells with an expected count of less than five, "walk" and "bicycle" are combined; "taxi", "rented car", "company car" are combined, and "motorized two-wheeler" and "motorcycle" are combined. The Pearson chi-squared test shows that there are significant differences in the distributions of most frequently used travel means (weekday) across location types (p-value = 0.000 < 0.05). As expected, the most frequently used travel means (weekday) is "walk/bicycle" for most people surveyed in bicycle-related locations (47.7 percent) (Table 37). "Motorized two-wheeler/motorcycle" is over-represented (86 percent) in the sample of motorized two-wheeler/motorcycle-related locations. Similarly, "car" is over-represented (50.9 percent) for the people sampled from car-related locations. Besides, "public transportation (as the most frequently used vehicle during weekday)" is the top answer

for people sampled from general on-street (33.4 percent) and household (42.5 percent) locations. In the on-line sample, "walk/bicycle" (46.2 percent) is considered as the most frequently used vehicle during weekday by most respondents. Again, the survey results suggest the effectiveness of location-based sampling at capturing specific groups of people based on their most frequently used travel means (weekday).

			MOST FREQUENT USED TRAVEL MEAN (weekday)					
			Walk/Bicycle	Public Transportation	Taxi/Rented Car/Company Car	Motorized Two-wheeler /Motorcycle	Car	Total
LOCATION TYPE	On-street (general)	Count	84	**116**	40	49	58	347
		% within LOCATION TY	24.2%	**33.4%**	11.5%	14.1%	16.7%	100.0%
	Bicycle-related	Count	**21**	15	0	3	5	44
		% within LOCATION TY	**47.7%**	34.1%	.0%	6.8%	11.4%	100.0%
	Motorized Two-wheeler/Motorcycle-relate	Count	4	2	0	**37**	0	43
		% within LOCATION TY	9.3%	4.7%	.0%	**86.0%**	.0%	100.0%
	Car-related	Count	21	46	41	26	**139**	273
		% within LOCATION TY	7.7%	16.8%	15.0%	9.5%	**50.9%**	100.0%
	Household	Count	107	**133**	33	19	21	313
		% within LOCATION TY	34.2%	**42.5%**	10.5%	6.1%	6.7%	100.0%
	On-line	Count	**36**	23	6	1	12	78
		% within LOCATION TY	**46.2%**	29.5%	7.7%	1.3%	15.4%	100.0%
Total		Count	273	**335**	120	135	235	1098
		% within LOCATION TY	24.9%	**30.5%**	10.9%	12.3%	21.4%	100.0%

Table 37: Cross-tabulation: Location Type x Most Frequently Used Travel Means (weekday)

Most Frequently Used Travel Means (weekend) vs. Location Type

As in the previous case, I combine original choice categories to reduce the number of cells with an expected count of less than five. The Pearson chi-squared test shows that there are significant differences in the distributions of most frequently used travel means (weekend) across location types (p-value = 0.000 < 0.05). Reasonably, "motorized two-wheeler/motorcycle" is over-represented (62.8 percent) in the sample of motorized two-wheeler/motorcycle-related locations, and "car" is over-represented (51.8 percent) in

the sample of car-related locations. "Public transportation (as the most frequently used vehicle during weekend)" is the top answer for respondents sampled from general on-street locations (40.8 percent), household (47.9 percent), and on-line (61.5 percent) (Table 38). However, instead of "walk/bicycle", the answer of "public transportation" stands out (53.5 percent) in the sample of bicycle-related locations. Therefore, the results suggest that the location-based sampling might not work perfectly to capture specific groups based on their most frequently used travel means on weekends.

			MOST FREQUENT USED TRAVEL MEAN (weekend)					
			Walk/Bicycle	Public Transportation	Taxi/Rented Car/Company Car	Motorized Two-wheeler /Motorcycle	Car	Total
LOCATION TYPE	On-street (general)	Count	55	**141**	61	30	59	346
		% within LOCATION TYP	15.9%	**40.8%**	17.6%	8.7%	17.1%	100.0%
	Bicycle-related	Count	7	**23**	7	1	5	43
		% within LOCATION TYP	16.3%	**53.5%**	16.3%	2.3%	11.6%	100.0%
	Motorized Two-wheeler/Motorcycle-relate	Count	4	10	1	**27**	1	43
		% within LOCATION TYP	9.3%	23.3%	2.3%	**62.8%**	2.3%	100.0%
	Car-related	Count	14	56	43	18	**141**	272
		% within LOCATION TYP	5.1%	20.6%	15.8%	6.6%	**51.8%**	100.0%
	Household	Count	52	**150**	70	9	32	313
		% within LOCATION TYP	16.6%	**47.9%**	22.4%	2.9%	10.2%	100.0%
	On-line	Count	11	**48**	6	2	11	78
		% within LOCATION TYP	14.1%	**61.5%**	7.7%	2.6%	14.1%	100.0%
Total		Count	143	**428**	188	87	249	1095
		% within LOCATION TYP	13.1%	**39.1%**	17.2%	7.9%	22.7%	100.0%

Table 38: Cross-tabulation: Location Type x Most Frequently Used Travel Means (weekend)

CHAPTER 4: MOTORIZATION PATHWAY

Question Design

An important part of my survey is to explore Chinese people's motorization pathway. In this research, motorization pathway is defined as the transition which individual people make among different travel means. As the name states, motorization is taken to mean a general likeliness to make a transition from non-motorized to motorized travel. I assume that Chinese people can accurately self-identify their motorization stages, from the past to current – for example, "I took bus before, but I am basically a bicycle person now" or "I started with walk, and now I use bus and bicycle interchangeably." Based on this assumption, I want to explore the motorization pathway in terms of patterns and direction of progression.

Sometimes, longitudinal studies are used to understand the progression of vehicle use (or purchase). However, as this survey was conducted only once, I used the following question to capture respondents' motorization pathways:

> *Please sort the sequence of the travel means you have used from the very past (in your memory) till now. Please fill in the numbers inside the boxes.*
> *(Skip the box if you never use that travel means)*
>
> **Ex:** [1] Bicycle, [2] Motorcycle, [3] Car = Bicycle → Motorcycle → Car
>
> ☐ Bicycle
> ☐ Walk
> ☐ Personal Car

☐ Public Transportation
☐ Auxiliary Power Vehicle[23], Motorcycle
☐ Taxi, Rented car
☐ Shared Company's Car

☐ NO, there is no such "pathway". (WHY? _____)

Basically, respondents sorted their personal history of travel means to trace their motorization pathway, or their lack of one. People respond to this question by assigning numbers (representing the sequence) to seven groups of travel means as shown above. Based on the experiences of a pilot survey, some travel means, e.g., motorized two-wheeler and motorcycle, or taxi and rented car, were combined to ensure a large enough sample (in the combined groups) for analysis. According to my definition of motorization pathway, people were encouraged to respond to this question as long as they have experience with "using" (not necessarily with purchasing) each travel means.

The goal of this analysis is two-fold. First, I want to see if a few common pathways are followed out of the many possible pathways. If so, what are those common patterns? Second, I would like to examine if the direction of motorization of Shanghai residents correspond to what I hypothesized. That is, is the daily experience of individual travelers one of increasing likeliness to rely on motorized means? The motorization pathway analysis will provide information about "What motorization stages are Shanghainese people currently (2006) in?" and "How do Shanghainese people get into their current motorization stages?" Thus, the pathway analysis is an introduction for further study on the vehicle purchase and use behavior of Shanghai (discussed in Chapter 5).

[23] In my survey, the original wording for motorized two-wheeler is "auxiliary power vehicle".

Data Cleaning

- **Modifying the Out-of-Range Responses**

In the survey, people can only assign one number (or leave it blank) to one travel means. That is, one travel means can only show up once in people's motorization pathway (but two different means can be assigned the same sequence number). The first reason is that, although there may be some people who have moved to a more motorized stage and then went back, I am more interested in understanding the motorization pathway before the "loop back" for a developing country like China, where I believe that most people are still *upgrading* their level of motorization. Secondly, by doing so, the complexity of possible answers can be greatly reduced – as a tradeoff of preventing this question from identifying a loop-back pattern, such as "walk → car → walk."

According to the question design, the biggest number associated with one travel means is 7, if people have experiences with using all seven listed travel means, and none of them is classified into the same motorization stage. However, in reviewing the survey results, I found three cases with out-of-range responses. Two people assigned "10" to private car, and one person assigned "8" to the group of motorized two-wheelers and motorcycles. I considered those values as the last travel means they used and changed them to the next highest values in the sequence of consecutive numbers given by the respondents.

- **Discarding and Recoding the Missing Data**

In addition to the out-of-range values, I noticed missing data, which would generate non-response errors. There are two types of missing data – people refused to comment on

all seven listed travel means (i.e., section non-response), or people refused to answer certain specific travel means (i.e., item non-response).

First, there are 45 respondents who didn't provide any answers to this question. That is, no sequence number was assigned to any of the seven travel means. I discarded these "section non-responses" from the analysis. Thus, the final sample for the pathway analysis consists of 992 cases (instead of the original 1,037).

Secondly, I recoded the "item non-responses" as "0." The reason I selected "0" to represent the item non-responses is that according to the original question, people *should* skip checking the box if they don't have experience of using that travel means. However, even with this instruction, we are still unable to distinguish the "should-be missing" (i.e., no use experience) from the "truly missing" (i.e., people overlooked or skipped the question for whatever reason). Eventually, I treated such item non-response as "should-be missing" in view of the survey instruction.

Data Validation

Based on the pre-test of similar questions in the pilot survey, the idea of a motorization pathway was considered an ambiguous topic for Chinese survey-takers. Therefore, a series of data validation processes have been conducted before the analysis to see if people really understand the questions and provide answers that are plausible and consistent throughout the survey.

- **Vehicle Use and Vehicle Purchase**

The first step of data validation is checking answers to the pathway questions against other questions about vehicle use or purchase. There are four questions related to people's vehicle purchase or use, as shown in Table 39.

	Vehicle Purchase (Ownership)		Vehicle Use	
	1. Vehicle Ownership (An Inventory)	**2. Most Expensive Vehicle Owned**	**3. Vehicle Use Experience (Motorization Pathway)**	**4. Most Frequently Used Travel Means (weekday/weekend)**
Question Type	Multiple Choice	Single Choice	Multiple Choice (ranking)	Single Choice
Choice Set	No Vehicle, Bicycle, Motorized Two-wheeler, Motorcycle, Car (5 choices)	No Vehicle, Bicycle, Motorized Two-wheeler, Motorcycle, Car (5 choices)	Walk, Bicycle, Public Transportation, Motorized Two-wheeler / Motorcycle, Taxi / Rented Car, (shared) Company Car, Car (7 choices)	Walk, Bicycle, Public Transportation, Motorized Two-wheeler, Motorcycle, Taxi, Rented Car, Company Car, Car (9 choices)

Table 39: Vehicle Purchase and Use Questions

Before validation, it is useful to explain the relationship between vehicle purchase and use in this survey; one typical case is illustrated in the following:

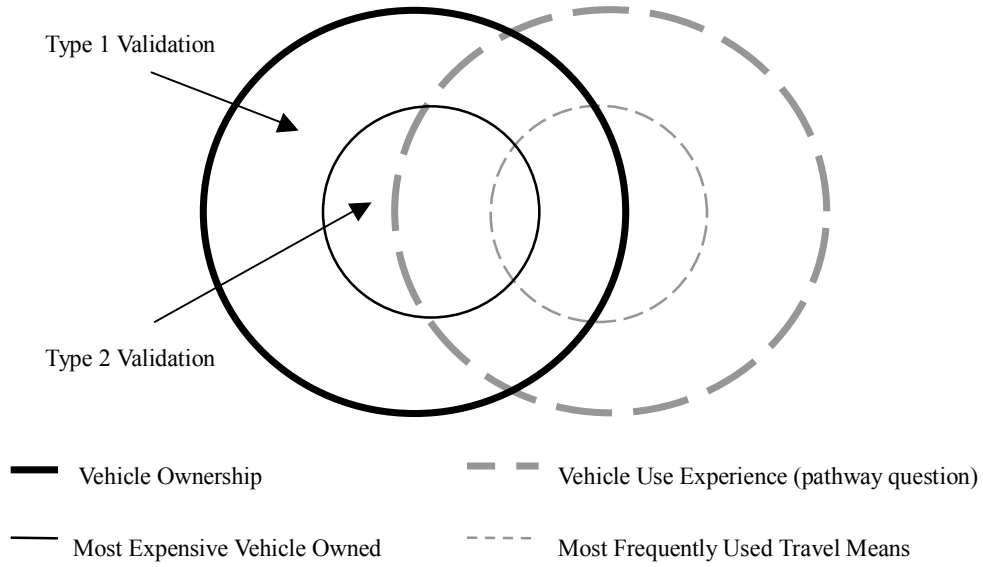


Figure 20: Type 1 and Type 2 Validation

The motorization question is shown as the big grey dashed circle in Figure 20. In Table 40, I describe my three types of validation.

Type 1 Validation	*How many people reported ownership of a vehicle but didn't report any experience using it?*
Type 2 Validation	*How many people reported ownership of a vehicle (as the most expensive one) but didn't report any experience using it?*
Type 3 Validation	*How many people reported the most frequently travel means but didn't report any experience using that travel means?*

Table 40: Three Types of Validation

My assumption for the Type 1 and Type 2 validation is there will be very few (or zero) cases falling into the areas pointed out in Figure 20. For example, if a person has no past experience using (defined as *driving* in the original survey) a car, it is unlikely that such a person owns a car or reports car as his/her most expensive vehicle owned. Besides,

according to the original survey questions (see Appendix), people responded to questions in Table 39 based on their *personal* experiences (without considering their family). Thus, in principle, people should not report owning a car if it is owned by spouse or parents but never driven by the respondent.

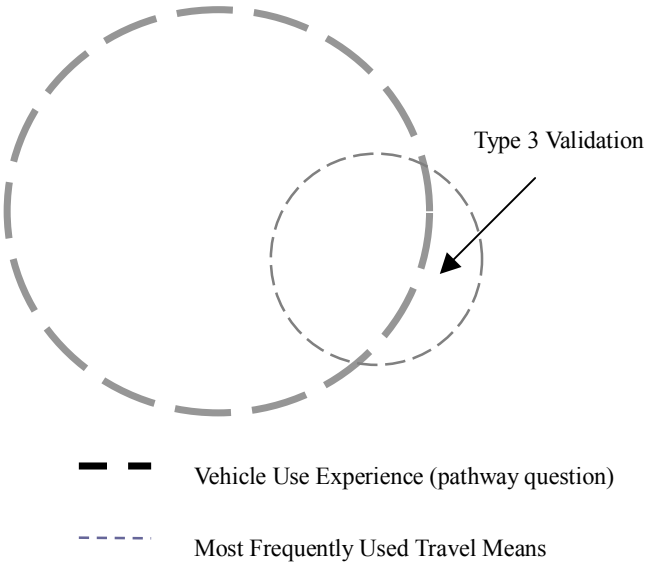


Figure 21: Type 3Validation

In addition, Figure 21 shows the Type 3 Validation. Similarly, I expect very few (or zero) cases to be in the indicated area. For example, if a person doesn't even have past experience of using a car, it is logically impossible for this person to report car as the most frequently used travel means during the weekday or weekend. Last, the selection of these three types of validation was arbitrary for a quick check on the internal consistency of data. Other validations such as: "how many people reported ownership of a vehicle (as the most expensive one) but didn't report the ownership of it?" can be conducted for a more thorough check.

- **Internal Consistency: Use Experiences**

Based on the above three types of validation, I first checked the internal consistency by cross-tabulating vehicle use experience by: vehicle ownership, most expensive vehicle owned, and most frequently used travel means during weekday and weekend.

Walk Experience

"Walk" is a tricky choice in our pathway question. Although it is not a "vehicle" which can be "purchased" per se, I still considered walk as a valid mode that can be identified into a stage of people's motorization pathway. Strictly speaking, everybody's first transportation mode should be walk, except for some special cases such as people with impaired personal mobility. However, the inclusion of walk as a choice was for the purpose of ensuring that everyone had an appropriate response to select (at a tradeoff of adding to the complexity). In this survey, interestingly, not everybody listed walk as the first motorization stage – there are 143 people (out of 992) who reported walk, but not as the first travel means; in addition, there are another 229 people (out of 992) who didn't rank walk at all. I didn't consider those abnormal answers on the walk mode as "implausible" because people who provided those answers might still tell true stories – based on their memory and judgment. For example, maybe they grew up in a household that had already moved beyond walking, so that even as children they had access to bike, bus, transit, or even car. Since "walk" cannot be owned, Table 41 and Table 42 show the results of the Type 3 Validation only. There were only six cases (in both Table 41 and Table 42) with a Type 3 violation (identified in red) indicating that most people provided consistent answers in this respect.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Walk	None	6	56	18	5	7	29	24	9	70	224
	YES	72	239	52	1	28	114	77	19	145	747
Total		78	295	70	6	35	143	101	28	215	971

Table 41: Walk Experience vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Walk	None	6	60	49	1	1	15	14	5	75	226
	YES	58	300	114	7	2	45	45	19	155	745
Total		64	360	163	8	3	60	59	24	230	971

Table 42: Walk Experience vs. Most Frequently Used Travel Means (weekend)

In the cross-tabulation tables, I put "none" instead of "NO" because "none" represents missing data, which can be no comment or NO. Nevertheless, for the pathway question specifically, I have assumed the missing data to be "no," i.e., 0, according to the original survey instruction.

Last, I didn't conduct another data cleaning for those "abnormal (red) cases" in Table 41 and Table 42 (and for all the following validation checks) because those cases are not *absolutely incorrect*. Taking "walk" as an example for Type 3 violation, people might self-identify their motorization stage as "beyond walk" in general; while still report walk as the most frequently used travel means during weekday. For Type 1 or 2 violation, some people might own luxury car for displaying only (as personal collection) without really using it.

Bicycle Use Experience

Table 43 and Table 44 show the Type 1 and Type 2 Validations for bicycle use and ownership. As we can see, not many cases fall into the disallowed categories.

Count

		Bicycle Ownership		
		None	YES	Total
Bicycle Use	None	181	17	198
	YES	376	418	794
Total		557	435	992

Table 43: Bicycle Use vs. Bicycle Ownership

Count

		Most Expensive Vehicle Owned					
		No Vehicle	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Bicycle Use	None	95	6	27	11	55	194
	YES	113	258	146	39	214	770
Total		208	264	173	50	269	964

Table 44: Bicycle Use vs. Most Expensive Vehicle Owned

Table 45 and Table 46 are the results of the Type 3 Validation; there are *no* inconsistent responses.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Bicycle Use	None	22	82	16	2	5	0	15	7	48	197
	YES	56	213	54	4	30	143	86	21	167	774
Total		78	295	70	6	35	143	101	28	215	971

Table 45: Bicycle Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Bicycle Use	None	18	78	32	3	2	0	12	5	46	196
	YES	46	282	131	5	1	60	47	19	184	775
Total		64	360	163	8	3	60	59	24	230	971

Table 46: Bicycle Use vs. Most Frequently Used Travel Means (weekend)

Public Transportation Use Experience

For public transportation I did not check the answers against vehicle ownership, since public transportation is defined as a travel means not owned by an individual. However, I found 39 cases (Type 3 violation) in Table 47; and 58 cases (Type 3 Violation) in Table 48. Those cases suggest that this specific group of people didn't provide consistent answers between the pathway and most frequently used travel means questions.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Public Transportation Us	None	16	39	20	2	5	42	40	16	85	265
	YES	62	256	50	4	30	101	61	12	130	706
Total		78	295	70	6	35	143	101	28	215	971

Table 47: Public Transportation Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Public Transportatic Use	None	13	58	47	4	1	20	26	13	87	269
	YES	51	302	116	4	2	40	33	11	143	702
Total		64	360	163	8	3	60	59	24	230	971

Table 48: Public Transportation Use vs. Most Frequently Used Travel Means (weekend)

Motorized Two-wheeler or Motorcycle Use Experience

This is a combined group with experiences using either motorized two-wheelers or motorcycles. I checked against the ownership data for the Type 1 and Type 2 Validations. Table 49, Table 50 and Table 51 show that the answers provided by the motorcycle owner group are more consistent than the motorized two-wheeler owner group, because there are fewer cases of violations (Type 1 and Type 2) in the motorcycle owner group.

Count

		Motorized Two-wheeler Ownership		
		None	YES	Total
Motorized Two-wheeler & Motorcycle Use	None	538	24	562
	YES	234	196	430
Total		772	220	992

Table 49: Motorized Two-wheeler / Motorcycle Use vs. Motorized Two-wheeler Ownership

Count

		Motorcycle Ownership		
		None	YES	Total
Motorized Two-wheeler & Motorcycle Use	None	561	1	562
	YES	361	69	430
Total		922	70	992

Table 50: Motorized Two-wheeler / Motorcycle Use vs. Motorcycle Ownership

Count

		Most Expensive Vehicle Owned					
		No Vehicle	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Motorized Two-wheeler & Motorcycle Use	None	161	213	19	1	154	548
	YES	47	51	154	49	115	416
Total		208	264	173	50	269	964

Table 51: Motorized Two-wheeler / Motorcycle Use vs. Most Expensive Vehicle Owned

Table 52 and Table 53 are results of Type 3 Validation. I was satisfied with the data quality since there are very few cases with the Type 3 problem.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Motorized Two-wheel & Motorcycle Use	None	55	197	48	2	19	107	4	0	124	556
	YES	23	98	22	4	16	36	97	28	91	415
Total		78	295	70	6	35	143	101	28	215	971

Table 52: Motorized Two-wheeler / Motorcycle Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Motorized Two-whe & Motorcycle Use	None	43	228	98	5	2	42	1	0	132	551
	YES	21	132	65	3	1	18	58	24	98	420
Total		64	360	163	8	3	60	59	24	230	971

Table 53: Motorized Two-wheeler / Motorcycle Use vs. Most Frequently Used Travel Means (weekend)

Taxi or Rented Car Use Experience

Similar to public transportation, taxi or rented car can not be "purchased or owned" by individual. Therefore, the following tables are focused on the consistency between vehicle use experiences and the most frequently used travel means. Based on Table 54 and Table 55, we see that people with taxi or rented car experience (in their motorization pathway) basically provide answers consistent with their most frequently used travel means.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Taxi & Rente Car Use	None	25	101	13	1	13	58	57	19	93	380
	YES	53	194	57	5	22	85	44	9	122	591
Total		78	295	70	6	35	143	101	28	215	971

Table 54: Taxi / Rented Car Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Taxi & Rente Car Use	None	22	156	33	4	2	25	36	14	87	379
	YES	42	204	130	4	1	35	23	10	143	592
Total		64	360	163	8	3	60	59	24	230	971

Table 55: Taxi / Rented Car Use vs. Most Frequently Used Travel Means (weekend)

Car Use Experience

Table 56 and Table 57 represent the results of Type 1 and Type 2 Validation against the ownership data. Table 58 and Table 59 are the Type 3 Validation against vehicle use frequency data. As I found, there were not many cases showing the violations. Therefore, I believed the basic data quality of this group of respondents.

Count

		Car Ownership		
		None	YES	Total
Car Use	None	615	16	631
	YES	97	264	361
Total		712	280	992

Table 56: Car Use vs. Car Ownership

Count

		Most Expensive Vehicle Owned					
		No Vehicle	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Car Use	None	173	225	155	46	16	615
	YES	35	39	18	4	253	349
Total		208	264	173	50	269	964

Table 57: Car Use vs. Most Expensive Vehicle Owned

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Car Use	None	60	239	45	4	19	121	90	27	10	615
	YES	18	56	25	2	16	22	11	1	205	356
Total		78	295	70	6	35	143	101	28	215	971

Table 58: Car Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Car Use	None	55	300	123	8	3	47	51	21	11	619
	YES	9	60	40	0	0	13	8	3	219	352
Total		64	360	163	8	3	60	59	24	230	971

Table 59: Car Use vs. Most Frequently Used Travel Means (weekend)

Company Car Use Experience

I checked the Type 3 problem for people who have ever used (shared) company cars. As shown in Table 60 and Table 61, the small number of cases with Type 3 violation indicated the consistency between questions.

Count

		Most Frequently Used Travel Means (weekday)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Company Car Use	None	61	229	44	5	5	116	88	26	162	736
	YES	17	66	26	1	30	27	13	2	53	235
Total		78	295	70	6	35	143	101	28	215	971

Table 60: Company Car Use vs. Most Frequently Used Travel Means (weekday)

Count

		Most Frequently Used Travel Means (weekend)									
		Walk	Public Transportation	Taxi	Rented Car	Company Car	Bicycle	Motorized Two-wheeler	Motorcycle	Car	Total
Company Car Use	None	53	279	116	7	0	46	49	20	166	736
	YES	11	81	47	1	3	14	10	4	64	235
Total		64	360	163	8	3	60	59	24	230	971

Table 61: Company Car Use vs. Most Frequently Used Travel Means (weekend)

- **Plausibility: Single Stage (with Single Mode) Pathway**

After the validation based on internal consistency, the next step of validation was to check the plausibility of certain answers. People reporting single stage (with single mode) pathway, e.g., the motorization pathway is walk only, were selected for plausibility check – *Did respondents who self-reported such pathways appear to provide plausible answers?* I first selected respondents with single stage (with single mode) pathways; then checked their top (most frequent) answers to three other questions – most expensive vehicle owned, license ownership, and occupation. Occupation was selected because it was considered as a demographic attribute related to motorization pathway. Besides, in my survey, there was less missing data on occupation than on other demographic attributes such as income.

	Single Stage (with Single Mode) Motorization Pathway Groups				
	Walk [7]*	**Bicycle** [25]	**Public Transportation** [5]	**Motorized Two-wheeler, or Motorcycle** [4]	**Car** [3]
Most Expensive Vehicle Owned?	No vehicle (70%)	Bicycle (84%)	No vehicle (60%)	Motorized Two-wheeler (75%)	Car (100%)
License Ownership?	No license (86%)	No license (92%)	No license (80%)	Both Motorcycle and Car (75%)	Car Only (67%)
Occupation?	Other (60%)	Factory Worker (30%)	N/A**	Private-owned Company Worker (50%)	N/A**

Table 62: Single Stage (with Single Mode) Pathway Groups and Top Answers to Three Validation Questions

Table 62 shows people with different types of single stage (with single mode) pathway and their top (most frequent) answers to three validation questions. One thing to be noted is that such motorization pathway is not applicable for every travel means in the choice set. For example, no one has "taxi or rented car only" and "company car only" pathway. The total cases of single stage (with single mode) pathway are only 44 (out of 992). Judging from people's most frequent answers to the three validation questions, I didn't find anything unreasonable or counter-intuitive – it suggested that the single stage (with single mode) pathway was a plausible, even if infrequent, answer. In conclusion, based

* Numbers in brackets indicate the number of cases.
** All answers are with equal frequency.

on the results of both internal consistency and plausibility checks, I am more confident about the fundamental validity of the motorization pathway data.

Common Motorization Pathway

- **Frequency Distribution of Actual Patterns**

As mentioned, one goal of the motorization pathway analysis is to identify whether there are common pathway patterns out of the many possible pathways. The answer is no: a total of 331 motorization pathways were reported by 992 survey respondents (see Appendix for the complete distribution). Since people reported pathway patterns of different sequences and number of stages, generalizing common pathway patterns is challenging. In this research, I first try to identify the common motorization pathways according to the top (based on frequency) patterns.

In Table 63, 498 (more than 50% of the sample) respondents' motorization pathways are consolidated into the top 30 most frequent patterns. The codes W, B, P, M, T, S, and C stand for: Walk, Bicycle, Public Transportation, Motorized Two-wheeler or Motorcycle, Taxi or Rented Car, (share) Company Car, and (private) Car. The most frequent pathway, reported by 6.5 percent of sample, is: "Walk (W) → Bicycle (B) → Public Transportation (P) → Taxi or Rented Car (T)." Private car didn't appear until it appeared as the last mode in the 3rd most frequent pathway (W → B → P → T → C). Among the top 30 patterns, there are only eight patterns (covering 11% of the sample) that include car, which confirms car is not yet a widely used mobility option. In terms of the sequence, pathways all start with non-motorized means (walk, bicycle, etc.), and mostly end with

motorized means (taxi, car, etc.). In terms of number of stages, pathways generally have three or four stages (within the top 30 patterns). However, the "bicycle only" and "walk only" pathways ranked as the 5th and 26th most frequent.

Motorization Pathway (Top 30 Patterns, sorted by frequency)	**Frequency**	**Percentage of 992 cases**
WBPT	64	6.5%
WPT	33	3.3%
WBPTC	32	3.2%
WBP	30	3.0%
B	25	2.5%
WB	25	2.5%
BM	22	2.2%
WBM	20	2.0%
WBPTM	19	1.9%
WBPM	18	1.8%
BPT	17	1.7%
WP	17	1.7%
BC	15	1.5%
WBPMTSC	15	1.5%
WBPC	12	1.2%
WBPMT	12	1.2%
WBPTS	12	1.2%
WPBT	11	1.1%
WPTC	11	1.1%
BMC	10	1.0%
BWPT	10	1.0%
WBPMTC	10	1.0%
BWP	9	0.9%
WBMPT	8	0.8%
WPTS	8	0.8%
W	7	0.7%

WBMC	7	0.7%
WBMPTS	7	0.7%
BP	6	0.6%
WBMP	6	0.6%
TOTAL	**498**	**50.2%**

Table 63: Top 30 Motorization Pathway Patterns (50% of the cases)

- **Number of Motorization Stage**

In addition to the frequency distribution, reviewing the number of motorization stages is another way to define common motorization patterns. Table 64 classifies 992 respondents based on number of motorization stages. Pathways of three or four stages are the most common, accounting for more than half of the sample pathways (52%). In the remaining half, 18% people are simple pathways with one or two stages, and 30% are longer pathways with five, six or seven stages.

One thing to be noted is that, in addition to the previous 44 single stage (with single mode) cases, two cases with "single stage but multiple/concurrent modes[24]" – (BWCP) and (BWPTS) – were also considered in Table 64, thus making the total counts 46.

[24] Travel modes within the parentheses means those modes were reported in the same position of respondents' motorization pathway.

Number of Motorization Stages	Counts	Percentage of 992 cases
1 (with 7 patterns)	**46**	5%
2 (with 26 patterns)	130	13%
3 (with 77 patterns)	**255**	**26%**
4 (with 81 patterns)	**254**	**26%**
5 (with 58 patterns)	162	16%
6 (with 36 patterns)	68	7%
7 (with 46 patterns)	77	8%
TOTAL	**992**	**100%**

Table 64: Number of Motorization Stages (100% of the cases)

Concluding from above, a *common* motorization pathway can be pictured as a series of three (or more) transitions in the types of travel means that a Shanghainese person has used for their daily travel. This result is not surprising for Shanghai – a city with rising income and various transportation options available.

Direction of Motorization

- **Motorization Pathway: Looking Forward**

After the search for common motorization pathways, I wanted to understand if pathways are in the direction of motorization. Tree diagrams of the first two, three and four motorization stages were developed to examine the pathway in the "forward" direction (i.e., from past to current)[25]. In the tree diagram, pathway patterns were basically truncated and combined. For example, by considering the first two stages, patterns of "W → B → C" and "W → B → M → T" will be combined as two cases of "W → B". The

[25] Since about 2/3 of the motorization pathways ended within four stages in my survey, analyzing the first two to four stages is considered sufficient.

partial look at motorization patterns can greatly reduce the data complexity, as in Table 65.

Only first 2 stages	**Only first 3 stages**	**Only first 4 stages**	**Consider *all* stages**
60 Patterns (100% cases)	152 Patterns (100% cases)	236 Patterns (100% cases)	331 Patterns (100% cases)
take 2 patterns to reach 50%	take 5 patterns to reach 50%	take 14 patterns to reach 50%	take 30 patterns to reach 50%

Table 65: Partial and Complete Motorization Pathway (Looking Forward)

In the following, tree diagrams of seven stages are presented. The top (most frequent) branch and pattern are identified[26], and the progression of motorization is reviewed.

Motorization Pathway Starts with "Walk" (613 cases)

[26] In this research, branches are specifically defined as the consolidated patterns of the first (or last) two stages.

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
W	7	W	7	W	7
W(BP)	2	W(BP)(MT)	1	W(BP)(MT)C	1
W(BPS)	1	W(BP)C	1	W(BP)C	1
WB	**416**	W(BPS)(MT)	1	W(BPS)(MT)C	1
WC	7	**WB**	**25**	**WB**	**25**
WM	16	**WB(PM)**	**3**	**WB(PM)**	**1**
WP	146	**WB(PMT)**	**1**	**WB(PM)T**	**2**
WS	2	**WB(PT)**	**2**	**WB(PMT)(CS)**	**1**
WT	11	**WBC**	**9**	**WB(PT)S**	**2**
(WM)B	2	**WBM**	**82**	**WBC**	**4**
(WP)B	2	**WBP**	**276**	**WBCM**	**1**
(WP)T	1	**WBS**	**3**	**WBCP**	**2**
		WBT	**15**	**WBCS**	**1**
		WC	2	**WBCT**	**1**
		WCB	2	**WBM**	**20**
→ (consider stages forward)		WCP	3	**WBM(PT)**	**2**
		WM	6	**WBMC**	**7**
		WMB	2	**WBMP**	**34**
		WMC	2	**WBMS**	**2**
		WMP	3	**WBMT**	**17**
		WMT	3	**WBP**	**30**
		WP	17	**WBP(MT)**	**2**
		WPB	50	**WBPC**	**17**
		WPC	4	**WBPM**	**72**
		WPM	5	**WBPS**	**8**
		WPS	6	**WBPT**	**147**
		WPT	64	**WBSC**	**2**
		WS	1	**WBST**	**1**
		WSM	1	**WBT**	**2**
		WT	3	**WBTC**	**3**
		WTB	2	**WBTM**	**2**
		WTC	5	**WBTP**	**3**
		WTP	1	**WBTS**	**5**
		(WM)BP	2	WC	2
		(WP)BT	2	WCBP	2
		(WP)T	1	WCPB	2
				WCPT	1
				WM	6
				WMBP	2
				WMC	2
				WMP	1
				WMPB	2
				WMT	2
				WMTC	1
				WP	17
				WPB	6
				WPBC	1
				WPBM	19
				WPBS	3
				WPBT	21
				WPC	4
				WPM	2
				WPMT	3
				WPS	4
				WPST	2
				WPT	33
				WPTB	7
				WPTC	11
				WPTM	2
				WPTS	11
				WS	1
				WSM	1
				WT	3
				WTBC	1
				WTBM	1
				WTC	5
				WTPS	1
				(WM)BPT	2
				(WP)BT	1
				(WP)BTC	1
				(WP)T	1
N (2+ stages) = 606 N (1 stage) = 7		N (3+ stages) = 551 N (2 stages) = 55 N (1 stage) = 7		N (4+ stages) = 431 N (3 stages) = 120 N (2 stages) = 55 N (1 stage) = 7	

Table 66: Tree Diagram (Start with "Walk", Look Forward)

Not surprisingly, most of our respondents start their motorization pathways with "walk" (613 cases; 62% sample). In Table 66, the color-coding represents different branches. The most frequent branch is also shown under bold-faced font. "W → B → P → T" is the most frequent pattern (147 cases) within this group. In addition, "W → B → P → M" is also considered a frequent pattern (72 cases). The branch starting with "WB" is the most frequent branch (416 cases).

At the bottom of the table, we can see the *tree expansion* based on cases. Most people in this group have complicated patterns, i.e., four stages or more. Besides, the total number of patterns grows quickly from 12 (first 2 stages) to 71 (first 4 stages). Both findings above suggest that this is a diverse group.

Motorization Pathway Starts with "Bicycle" (246 cases)

People with "bicycle" as the initial motorization stage (246 cases; 25% sample) make up the second biggest group. The branch starting with "BM" is the most frequent (60 cases) within this group (Table 67). It is reasonable to see the "motorized two-wheeler or motorcycle" as the next immediate motorization stage of bicycle, because the "motorized two-wheeler or motorcycle" is similar to bicycle by basic features but more motorized. Single stage "B" is the top pattern (25 cases); however, "B → M" (22 cases), "B → W → P → T" (18 cases) and "B → P → T" (17 cases) are also with high frequency.

This is a diverse group, since the total numbers of patterns increase rapidly with the tree expansion. Besides, I noticed some singular patterns that cannot be classified into any branch. These are cases of unique motorization pathways.

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
(BC)T	1	(BC)TP	1	(BC)TP	1
(BM)T	1	(BM)T	1	(BM)T	1
(BW)(PM)	1	(BW)(PM)T	1	(BW)(PM)T	1
(BW)(PMT)	1	(BW)(PMT)(CS)	1	(BW)(PMT)(CS)	1
(BW)C	1	(BW)C(PT)	1	(BW)C(PT)	1
(BW)M	1	(BW)MS	1	(BW)MSC	1
(BW)P	2	(BW)PC	1	(BW)PCT	1
(BWCP)	1	(BW)PT	1	(BW)PTS	1
(BWPTS)	1	(BWCP)	1	(BWCP)	1
B	25	(BWPTS)	1	(BWPTS)	1
B(PT)	1	B	25	**B**	**25**
B(WC)	1	B(PT)C	1	B(PT)C	1
BC	23	B(WC)(PM)	1	B(WC)(PM)(TS)	1
BM	**60**	BC	15	BC	15
BP	59	BCM	5	BCM	5
BS	5	BCP	2	BCP	1
BT	15	BCW	1	BCPT	1
BW	47	**BM**	**22**	BCW	1
		BM(CS)	**2**	**BM**	**22**
		BMC	**12**	**BM(CS)**	**2**
→		**BMP**	**10**	**BMC**	**10**
(consider stages forward)		**BMS**	**5**	**BMCS**	**1**
		BMT	**6**	**BMCT**	**1**
		BMW	**3**	**BMP**	**2**
		BP	6	**BMPC**	**3**
		BP(CT)	1	**BMPT**	**4**
		BPC	4	**BMPW**	**1**
		BPM	12	**BMS**	**3**
		BPT	28	**BMSC**	**1**
		BPW	8	**BMST**	**1**
		BSC	2	**BMT**	**3**
		BST	2	**BMTC**	**1**
		BSW	1	**BMTP**	**1**
		BT	3	**BMTS**	**1**
		BTC	4	**BMWP**	**1**
		BTM	2	**BMWS**	**1**
		BTP	2	**BMWT**	**1**
		BTS	4	BP	6
		BW	4	BP(CT)	1
		BWC	3	BPC	4
		BWM	5	BPM	3
		BWP	32	BPMC	2
		BWT	3	BPMT	5
				BPMW	2
				BPT	17
				BPTC	4
				BPTM	1
				BPTS	4
				BPTW	2
				BPW	5
				BPWM	1
				BPWT	2
				BSC	2
				BST	2
				BSW	1
				BT	3
				BTC	4
				BTM	1
				BTMC	1
				BTP	1
				BTPC	1
				BTS	1
				BTSC	2
				BTSW	1
				BW	4
				BWC	2
				BWCP	1
				BWM	1
				BWMP	3
				BWMT	1
				BWP	9
				BWPC	1
				BWPM	4
				BWPT	18
				BWT	1
				BWTC	1
				BWTS	1
N (2+ stages) = 219 N (1 stage) = 27		N (3+ stages) = 168 N (2 stages) = 51 N (1 stage) = 27		N (4+ stages) = 81 N (3 stages) = 87 N (2 stages) = 51 N (1 stage) = 27	

Table 67: Tree Diagram (Start with "Bicycle", Look Forward)

Motorization Pathway Starts with "Public Transportation" (52 cases)

There are 52 respondents starting their motorization pathway with "public transportation" (6% sample). The most frequent branch within this group is "PT" (16 cases). Single stage "P" and "P → W → T" are both top patterns (5 cases); but "P → T → C" and "P → T → W" are also popular (4 cases). Compared to previous two groups, this is a less diverse group with less tree expansion. (Table 68)

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
P	5	P	5	**P**	**5**
PB	14	PB	2	PB	2
PC	3	PBC	2	PBC	1
PM	4	PBM	2	PBCT	1
PS	2	PBS	1	PBM	1
PT	**16**	PBT	2	PBMT	1
PW	8	PBW	5	PBS	1
		PC	1	PBT	2
		PCW	2	PBWC	1
→		PMB	1	PBWM	1
(consider stages forward)		PMT	2	PBWT	3
		PMW	1	PC	1
		PSC	1	PCWB	1
		PST	1	PCWT	1
		PT	**1**	PMBW	1
		PTB	**3**	PMT	2
		PTC	**4**	PMWT	1
		PTM	**1**	PSC	1
		PTS	**2**	PST	1
		PTW	**5**	**PT**	**1**
		PWB	1	**PTB**	**2**
		PWT	7	**PTBW**	**1**
				PTC	**4**
				PTMS	**1**
				PTS	**1**
				PTSC	**1**
				PTW	**4**
				PTWM	**1**
				PWB	1
				PWT	**5**
				PWTS	2
N (2+ stages) = 47 N (1 stage) = 5		N (3+ stages) = 43 N (2 stages) = 4 N (1 stage) = 5		N (4+ stages) = 17 N (3 stages) = 26 N (2 stages) = 4 N (1 stage) = 5	

Table 68: Tree Diagram (Start with "Public Transportation", Look Forward)

Motorization Pathway Starts with "Motorized Two-wheeler or Motorcycle" (36 cases)

36 respondents' starting stages of motorization pathway are "motorized two-wheeler or motorcycle" (4% sample). Branches starting with "MB" and "MW" are both listed as most frequent (8 cases). Likely, the top patterns within this group are "M → B" and "M" (4 cases). This is a less diverse group without rapid pattern growth. (Table 69)

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
M	4	M	4	**M**	**4**
MB	**8**	**MB**	**4**	**MB**	**4**
MC	1	**MBP**	**1**	**MBPW**	**1**
MP	6	**MBW**	**3**	**MBW**	**2**
MS	2	MC	1	**MBWP**	**1**
MT	7	MP	3	MC	1
MW	**8**	MPB	2	MP	3
		MPT	1	MPBT	1
		MS	1	MPBW	1
		MSW	1	MPTW	1
→		MT	1	MS	1
(consider stages forward)		MTB	2	MSW	1
		MTP	1	MT	1
		MTS	1	MTB	2
		MTW	2	MTP	1
		MW	**3**	MTSP	1
		MWB	**1**	MTW	1
		MWP	**3**	MTW(CS)	1
		MWT	**1**	**MW**	**3**
				MWBP	**1**
				MWP	**1**
				MWPB	**1**
				MWPC	**1**
				MWTP	**1**
N (2+ stages) = 32 N (1 stage) = 4		N (3+ stages) = 19 N (2 stages) = 13 N (1 stage) = 4		N (4+ stages) = 11 N (3 stages) = 8 N (2 stages) = 13 N (1 stage) = 4	

Table 69: Tree Diagram (Start with "Motorized Two-wheeler or Motorcycle", Look Forward)

Motorization Pathway Starts with "Taxi or Rented Car" (9 cases)

There are only nine respondents starting motorization pathway with "taxi or rented car" (1% sample). The branch of "TP" is the top one with three cases. However, in this small group, no top pattern can be properly identified when considering first four stages, because every pattern ends up with the same number of case. (Table 70)

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
TB	1	TBP	1	TBP	1
TC	2	TC	1	TC	1
TM	1	TCB	1	TCBW	1
TP	**3**	TMB	1	TMBW	1
TS	1	**TPB**	**1**	**TPB**	**1**
TW	1	**TPC**	**1**	**TPC**	**1**
		TPS	**1**	**TPS**	**1**
→ (consider stages forward)		TSW	1	TSW	1
		TWC	1	TWC	1
N (2+ stages) = 9 N (1 stage) = 0		N (3+ stages) = 8 N (2 stages) = 1 N (1 stage) = 0		N (4+ stages) = 2 N (3 stages) = 6 N (2 stages) = 1 N (1 stage) = 0	

Table 70: Tree Diagram (Start with "Taxi or Rented Car", Look Forward)

Motorization Pathway Starts with "Car" (32 cases)

This is a group with 32 cases (2% sample). The top branch is "CT" (11 cases), and the top pattern is exactly the two-stage "C → T" (4 cases). Thus, for people start motorization with car, taxi is their next immediate stage although most people eventually end the pathways with public transportation or even walk (Table 71). One possible explanation for this "backward motorization" is that people who grow up in an affluent (car) family choose to establish on their own later in their life. Another possible explanation is that people who start with the car phase choose to *downgrade* their motorization because they realize that those *low-motorized* travel modes (taxi, public transportation) are sometimes more convenience for them to get around Shanghai.

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
C	3	C	3	C	3
CB	1	CBW	1	CBW	1
CM	5	CM	1	CM	1
CP	6	CMP	1	CMPT	1
CS	1	CMT	1	CMTW	1
CT	**11**	CMW	2	CMW	2
CW	5	CPB	2	CPBW	2
		CPM	1	CPMW	1
		CPS	1	CPST	1
		CPT	1	CPT	1
→ (consider stages forward)		CPW	1	CPWT	1
		CSW	1	CSW	1
		CT	**4**	**CT**	**4**
		CTP	**2**	**CTPW**	**2**
		CTS	**2**	**CTSP**	**2**
		CTW	**3**	**CTW**	**2**
		CWB	1	**CTWP**	**1**
		CWP	2	CWBT	1
		CWT	2	CWPT	2
				CWT	2
N (2+ stages) = 29 N (1 stage) = 3		N (3+ stages) = 24 N (2 stages) = 5 N (1 stage) = 3		N (4+ stages) = 15 N (3 stages) = 9 N (2 stages) = 5 N (1 stage) = 3	

Table 71: Tree Diagram (Start with "Car", Look Forward)

Motorization Pathway Starts with "Company Car" (4 cases)

Very few people start motorization with "company car" (4 case, close to 0% sample). The *top* branch "ST" is relatively large with two cases. Similar to group of people starting with taxi or rented car, no top pattern can be identified when taking first four stages in to account. (Table 72)

Within First 2 Stages	Counts	Within First 3 Stages	Counts	Within First 4 Stages	Counts
SC	1	SCP	1	SCPT	1
ST	**2**	**STC**	**1**	**STCW**	**1**
SW	1	**STP**	**1**	**STPW**	**1**
		SW	1	SW	1
→ (consider stages forward)					
N (2+ stages) = 4 N (1 stage) = 0		N (3+ stages) = 3 N (2 stages) = 1 N (1 stage) = 0		N (4+ stages) = 3 N (3 stages) = 0 N (2 stages) = 1 N (1 stage) = 0	

Table 72: Tree Diagram (Start with "Company Car", Look Forward)

- **Motorization Pathway: Looking Backward**

Similar to examining the forward direction, tree diagrams of the last two, three and four motorization stages were developed to understand the motorization pathway in the "backward" direction (i.e., from now to past). Basically, I want to start with the ending stage of a motorization pathway, and then trace back where it comes from. This partial look of motorization patterns (backward) can reduce the data complexity, as in Table 73.

Only last 2 stages	Only last 3 stages	Only last 4 stages	Consider all stages
66 Patterns (100% cases)	174 Patterns (100% cases)	281 Patterns (100% cases)	331 Patterns (100% cases)
take 8 patterns to reach 50%	take 17 patterns to reach 50%	take 26 patterns to reach 50%	take 30 patterns to reach 50%

Table 73: Partial and Complete Motorization Pathway (Looking Backward)

In the following, tree diagrams of seven ending stages are presented. The most frequent branches and patterns are identified and the progression of motorization is reviewed.

Motorization Pathway currently Ends with "Walk" (60 cases)

60 respondents (6% sample) reported "walk" as their final motorization stage. If consider the last four stages, single-stage pattern "W" is the most frequent (7 cases) within this group. However, patterns such as "B → P → W" (5 cases), "B → W" (4 cases), "P → T → W" (4 cases) are also with relatively high frequency (Table 74).

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
MBW	**2**	**MBW**	**5**	**BW**	**14**
PMBW	**2**	**BW**	**4**	TW	13
TMBW	**1**	**PBW**	**2**	PW	10
BW	**4**	**CBW**	**2**	W	7
CPBW	**1**	**TBW**	**1**	MW	7
MPBW	**1**	PTW	9	SW	7
CBW	**1**	CTW	2	CW	2
TCBW	**1**	MTW	2		
PTBW	**1**	BPW	6		
PTW	4	TPW	2		
BPTW	2	MPW	1	←	
MPTW	2	SPW	1	(consider stages backward)	
CPTW	1	W	7		
CTW	2	MW	3		
CMTW	1	CMW	2		
MTW	1	PMW	2		
BPW	**5**	MSW	2		
MBPW	1	TSW	2		
CTPW	1	BSW	1		
STPW	1	CSW	1		
BMPW	1	SW	1		
TSPW	1	BCW	1		
W	**7**	TCW	1		
MW	3				
CMW	2				
BPMW	1				
CPMW	1				
MSW	1				
TMSW	1				
BTSW	1				
TSW	1				
BSW	1				
CSW	1				
SW	1				
BCW	1				
STCW	1				
N (4+ stages) = 23 N (3 stages) = 22 N (2 stages) = 8 N (1 stage) = 7		N (3+ stages) = 45 N (2 stages) = 8 N (1 stage) = 7		N (2+ stages) = 53 N (1 stage) = 7	

Table 74: Tree Diagram (End with "Walk", Look Backward)

Pathways ending with "BW" are the most frequent (14 cases). That is, for most people currently walking, their immediate previous stage was bicycle. This result is not surprising, since both bicycle and walk are non-motorized travel means. People who used to be in a biking stage can easily transfer to walking.

Motorization Pathway currently Ends with "Bicycle" (82 cases)

There are 82 respondents (8% sample) whose motorization pathway currently ends with bicycling. "WB" (27 cases) is the most frequent branch at the two-stage level. By expanding the tree diagram, "B" and "W → B" are both the top patterns (25 cases). Similar to previous group, this result suggests the transfer between two non-motorized travel means (W and B). Judging from the tree expansion, this group is less diverse compared to previous group (Table 75).

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WB	**25**	**WB**	**25**	**WB**	**27**
PWB	**1**	**PWB**	**2**	B	25
TPWB	**1**	B	25	PB	10
B	**25**	WPB	6	TB	9
WPB	6	PB	2	MB	7
PB	2	MPB	1	CB	2
WMPB	1	TPB	1	(CS)B	1
TPB	1	PTB	6	SB	1
WPTB	4	MTB	2		
PTB	2	WTB	1		
MTB	2	MB	4		
PWTB	1	SMB	2	←	
MB	4	CMB	1	(consider stages backward)	
TSMB	2	SCB	2		
SCMB	1	W(CS)B	1		
MSCB	1	PSB	1		
TSCB	1				
TW(CS)B	1				
TPSB	1				
N (4+ stages) = 13		N (3+ stages) = 26		N (2+ stages) = 57	
N (3 stages) = 12		N (2 stages) = 31		N (1 stage) = 25	
N (2 stages) = 31		N (1 stage) = 25			
N (1 stage) = 25					

Table 75: Tree Diagram (End with "Bicycle", Look backward)

Motorization Pathway currently Ends with "Public Transportation" (97 cases)

97 respondents end their motorization pathways with "public transportation" (9% sample). "W → B → P" (30 cases) is the most frequent pattern, and this pattern arises from the top branch "BP" (38 cases) at the two-stage level. (Table 76)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WBP	**30**	**WBP**	**30**	**BP**	**38**
BP	**6**	**BP**	**6**	WP	29
WMBP	**1**	**MBP**	**1**	MP	13
TBP	**1**	**TBP**	**1**	TP	6
WP	17	WP	17	P	5
BWP	9	BWP	10	SP	3
MBWP	1	MWP	1	CP	2
MWP	1	TWP	1	(BWCP)	1
CTWP	1	BMP	8		
WBMP	6	MP	3		
BMP	2	SMP	1		
MP	3	WMP	1	←	
TSMP	1	MTP	3	(consider stages backward)	
WMP	1	BTP	2		
BMTP	2	(BC)TP	1		
MTP	1	P	5		
BTP	1	TSP	3		
WBTP	1	BCP	1		
(BC)TP	1	WCP	1		
P	5	(BWCP)	1		
BTSP	1				
CTSP	1				
WTSP	1				
BCP	1				
BWCP	1				
(BWCP)	1				
N (4+ stages) = 18 N (3 stages) = 47 N (2 stages) = 26 N (1 stage) = 6		N (3+ stages) = 65 N (2 stages) = 26 N (1 stage) = 6		N (2+ stages) = 91 N (1 stage) = 6	

Table 76: Tree Diagram (End with "Public Transportation", Look Backward)

Motorization Pathway currently Ends with "Motorized Two-wheeler or Motorcycle" (130 cases)

There are 130 (13% sample) respondents having "motorized two-wheeler or motorcycle" as the final motorization stage. At the two-stage level, "BM" is the top branch (50 cases); likely, at the four-stage level, the "B → M" is the top pattern (22 cases). Besides, "W → B → M" (20 cases); "B → P → T → M" (19 cases) and "W → B → P → M" (18 cases) are nearly as frequent as the top pattern (Table 77).

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BM	**22**	**BM**	**22**	**BM**	**50**
WBM	**20**	**WBM**	**20**	PM	28
WPBM	**6**	**PBM**	**7**	TM	23
PBM	**1**	**TBM**	**1**	CM	8
PTBM	**1**	BPM	22	SM	8
WBPM	18	WPM	6	WM	8
BPM	3	PTM	21	M	4
CBPM	1	BTM	1	B(PM)	1
BWPM	4	CTM	1		
WPM	2	BCM	5		
BPTM	19	SCM	2		
WPTM	2	CM	1	←	
BTM	1	TSM	5	(consider stages backward)	
PCTM	1	CSM	2		
BCM	5	WSM	1		
TSCM	2	WM	6		
CM	1	BWM	2		
PTSM	3	M	4		
CTSM	1	WB(PM)	1		
WTSM	1				
TCSM	2				
WSM	1				
WM	6				
BWM	1				
PBWM	1				
M	4				
WB(PM)	1				
N (4+ stages) = 62 N (3 stages) = 35 N (2 stages) = 29 N (1 stage) = 4		N (3+ stages) = 97 N (2 stages) = 29 N (1 stage) = 4		N (2+ stages) = 126 N (1 stage) = 4	

Table 77: Tree Diagram (End with "Motorized Two-wheeler or Motorcycle", Look Backward)

Motorization Pathway currently Ends with "Taxi or Rented Car" (245 cases)

There are 245 respondents with final motorization stage as "taxi or rented car" (25% sample). Going back to include four stages, "W → B → P → T" (64 cases) is the most frequent pattern within this group. Moreover, the branch ending with "PT" is the most frequent branch if considering only the last two stages. This branch suggests that most people in "taxi or rented car" stage come from "public transportation". Compared to all groups discussed before, this is by far the most diverse one with many complicated patterns and rapid tree expansion (Table 78).

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WBPT	**64**	**BPT**	**83**	**PT**	**146**
BPT	**17**	**WPT**	**45**	MT	32
(WM)BPT	**1**	**MPT**	**12**	BT	21
MBPT	**1**	**CPT**	**3**	WT	17
WPT	**33**	**SPT**	**2**	ST	13
BWPT	**10**	**PT**	**1**	CT	9
CWPT	**1**	PMT	19	P(MT)	2
MWPT	**1**	BMT	10	(BM)T	1
BMPT	**11**	WMT	2	(PM)T	1
WMPT	**1**	MT	1	(WP)T	1
BCPT	**2**	PBT	14	C(PT)	1
CPT	**1**	BT	3	P(CT)	1
MSPT	**2**	WBT	3		
PT	**1**	(WP)BT	1		
BPMT	14	PWT	7		
PMT	2	BWT	3	←	
WPMT	2	WT	3	(consider stages backward)	
SPMT	1	CWT	2		
PBMT	4	MWT	2		
BMT	3	PST	7		
WBMT	3	BST	3		
WMT	2	MST	3		
MT	1	CT	4		
WPBT	12	PCT	4		
PBT	2	MCT	1		
BT	3	BP(MT)	2		
WBT	2	(BM)T	1		
CWBT	1	(BW)(PM)T	1		
(WP)BT	1	(WP)T	1		
PWT	5	(BW)C(PT)	1		
BPWT	2	BP(CT)	1		
PBWT	2				
BWT	1				
WT	3				
CWT	2				
BMWT	1				
PMWT	1				
BPST	3				
WPST	2				
MPST	1				
PST	1				
BST	2				
WBST	1				
BMST	2				
CMST	1				
CT	4				
BPCT	2				
(BW)PCT	1				
MPCT	1				
PMCT	1				
WBP(MT)	2				
(BM)T	1				
(BW)(PM)T	1				
(WP)T	1				
(BW)C(PT)	1				
BP(CT)	1				
N (4+ stages) = 155 N (3 stages) = 76 N (2 stages) = 14 N (1 stage) = 0		N (3+ stages) = 231 N (2 stages) = 14 N (1 stage) = 0		N (2+ stages) = 245 N (1 stage) = 0	

Table 78: Tree Diagram (End with "Taxi or Rented Car", Look Backward)

Motorization Pathway currently Ends with "Car" (259 cases)

The biggest group contains 259 respondents with final motorization stage as "car" (27% sample). The "B → P → T → C" (36 cases) is the most frequent pattern, and the branch ending with "TC" (94 cases) is the largest one within this group.

It is not surprising that most people who are currently in the "car" stage come from a "taxi or rented car" stage. As a matter of fact, taxi or rent car can be considered a type of car in terms of basic functionality. However, car is a private mode and can be owned. The idea of ownership causes a major difference between private car and the taxi or rented car. Thus, I assume the transfer (or upgrade) from taxi to private car may involve certain attributes such as vehicle status. Last, according to Table 79, this is a large and diverse group. Apparently, there are many complicated patterns and the tree expansion is also complicated.

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BPTC	**36**	**PTC**	**55**	**TC**	**94**
WPTC	**12**	**MTC**	**18**	SC	54
PTC	**4**	**BTC**	**7**	MC	39
MPTC	**3**	**STC**	**7**	PC	36
PMTC	**10**	**WTC**	**6**	BC	24
BMTC	**7**	**TC**	**1**	WC	5
WMTC	**1**	TSC	33	C	3
BTC	**4**	MSC	11	(PT)C	2
WBTC	**2**	BSC	5	(BP)C	1
PBTC	**1**	PSC	4	(MT)C	1
MSTC	**5**	(PT)SC	1		
PSTC	**2**	BMC	20		
WTC	**5**	PMC	9		
BWTC	**1**	TMC	7	←	
TC	**1**	WMC	2	(consider stages backward)	
MTSC	18	MC	1		
PTSC	10	BPC	16		
BTSC	3	MPC	7		
WTSC	2	TPC	6		
TMSC	7	WPC	5		
BMSC	3	PC	1		
(BW)MSC	1	SPC	1		
BSC	2	BC	15		
WBSC	2	WBC	4		
PBSC	1	PBC	2		
MPSC	2	TBC	2		
BPSC	1	SBC	1		
PSC	1	BWC	2		
B(PT)SC	1	WC	2		
BMC	10	TWC	1		
WBMC	7	C	3		
PBMC	2	B(PT)C	1		
SBMC	1	M(PT)C	1		
BPMC	8	W(BP)C	1		
SPMC	1	(BPS)(MT)C	1		
PTMC	5				
BTMC	1				
STMC	1				
WMC	2				
MC	1				
WBPC	12				
BPC	4				
BMPC	5				
SMPC	2				
BTPC	3				
MTPC	2				
TPC	1				
WPC	4				
BWPC	1				
PC	1				
TSPC	1				
BC	15				
WBC	4				
PBC	1				
WPBC	1				
PTBC	1				
WTBC	1				
TSBC	1				
BWC	2				
WC	2				
TWC	1				
C	3				
B(PT)C	1				
BM(PT)C	1				
W(BP)C	1				
W(BPS)(MT)C	1				
N (4+ stages) = 189 N (3 stages) = 47 N (2 stages) = 20 N (1 stage) = 3		N (3+ stages) = 236 N (2 stages) = 20 N (1 stage) = 3		N (2+ stages) = 256 N (1 stage) = 3	

Table 79: Tree Diagram (End with "Car", Look Backward)

Motorization Pathway currently Ends with "Company Car" (119 cases)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BPTS	**15**	**PTS**	**40**	**TS**	**59**
WPTS	**13**	**MTS**	**11**	CS	14
MPTS	**9**	**BTS**	**8**	MS	14
(BW)PTS	**1**	TCS	9	PS	12
CPTS	**1**	MCS	2	BS	4
PTS	**1**	PCS	2	WS	4
PMTS	**6**	(MT)CS	1	T(CS)	2
BMTS	**4**	TMS	7	(PT)S	2
WMTS	**1**	BMS	4	M(CS)	2
PBTS	**6**	CMS	1	T(CMS)	2
BTS	**1**	MS	1	(PMT)(CS)	2
WBTS	**1**	PMS	1	(BWPTS)	1
MTCS	5	TPS	4	(PM)(TS)	1
PTCS	3	WPS	4		
BTCS	1	BPS	2		
BMCS	1	MPS	2		
TMCS	1	PBS	3	←	
TPCS	2	TBS	1	(consider stages backward)	
(BP)(MT)CS	1	TWS	2		
PTMS	4	MWS	1		
BTMS	1	WS	1		
CTMS	1	(PM)T(CS)	2		
WTMS	1	B(PT)S	1		
BMS	3	M(PT)S	1		
WBMS	1	BM(CS)	2		
TCMS	1	PT(CMS)	2		
MS	1	(BW)(PMT)(CS)	1		
CPMS	1	B(PMT)(CS)	1		
MTPS	3	(BWPTS)	1		
TPS	1	(WC)(PM)(TS)	1		
WPS	4				
WBPS	2				
BMPS	1				
TMPS	1				
WPBS	2				
PBS	1				
CTBS	1				
BTWS	1				
CTWS	1				
BMWS	1				
WS	1				
B(PM)T(CS)	2				
WB(PT)S	1				
BM(PT)S	1				
BM(CS)	2				
WPT(CMS)	1				
BPT(CMS)	1				
(BW)(PMT)(CS)	1				
WB(PMT)(CS)	1				
(BWPTS)	1				
B(WC)(PM)(TS)	1				
N (4+ stages) = 102 N (3 stages) = 14 N (2 stages) = 2 N (1 stage) = 1		N (3+ stages) = 116 N (2 stages) = 2 N (1 stage) = 1		N (2+ stages) = 118 N (1 stage) = 1	

Table 80: Tree Diagram (End with "Company Car", Look Backward)

As from Table 80, 119 respondents reported their final motorization stage as "company car" (12% sample). The "B → P → T → S" (15 cases) and "W → P → T → S" (13 cases) are top two patterns within this group, and they are both associated with the top branch ending with "TS" (59 cases)

This is a diverse group. Various pathways are observed at the four stage level. Beside, most of the respondents (102 out of a total of 119 cases) have motorization pathways involve four stages or more, as shown at the bottom of Table 80.

- **Hypothetical Direction**

Looking Forward		Looking Backward	
Starting Stage	**Most Frequent Branch and Pattern within Group**	**Ending Stage**	**Most Frequent Branch and Pattern within Group**
Walk # [613 cases]	W → B * [416] W → B → P → T $ [147]	**Walk** [60 cases]	[14] B → W [7] W
Bicycle [246 cases]	B → M [60] B [25]	**Bicycle** [82 cases]	[27] W → B [25] B [25] W→ B
Public Transportation [52 cases]	P→ T [16] P [5] P → W → T [5]	**Public Transportation** [97 cases]	[38] B → P [30] W → B → P
Motorized Two-wheeler / Motorcycle [36 cases]	M → B [8] M → W [8] M [4] M → B [4]	**Motorized Two-wheeler / Motorcycle** [130 cases]	[50] B → M [22] B → M
Taxi / Rented Car [9 cases]	T → P [3] N/A (no top pattern)	**Taxi / Rented Car** [245 cases]	[146] * P → T [64] $ W→ B→ P → T

Car [32 cases]	C → T [11] C → T [4]	**Car** # [259 cases]	[94] T →C [36] B → P → T →C
Company Car [4 cases]	S → T [2] N/A (no top pattern)	**Company Car** [119 cases]	[59] T →S [15] B → P → T →S
# : Largest group (consider first 1 stage) across sample * : Largest branch (consider first 2 stages) across sample $: Largest pattern (consider first 4 stages) across sample [ ] Numbers in bracket indicate the number of cases		# : Largest group (consider last 1 stage) across sample * : Largest branch (consider last 2 stages) across sample $: Largest pattern (consider last 4 stages) across sample [ ] Numbers in bracket indicate the number of cases	

Table 81: Summary of Tree Diagrams

Table 81 summarizes the top patterns (within group and across sample) under different levels of data consolidation (forward and backward). The tradeoff of the data consolidation is that the more I combine the stages, the easier it is to identify the most frequent pathway; however, I will lose more detail of the actual motorization stages. As mentioned, the term "group" and "branch" stand for the patterns consolidated into the first or last one and two stages. The reason for consolidating data into the groups is to understand the sample distribution based on people's starting and ending stages. The reason for identifying the branches is to focus on the immediate next (or previous) step of people's starting and ending stages.

Based on the top branches in Table 81, I found that certain travel means are mutually interchangeable. For example, for people start motorization with walk, bicycle is their next stage. However, for people whose last stage in their pathway is walk, bicycle also shows up as their previous stage. Similar situation happens on that pairs of "BM", "TP", "TC", and "TS". Besides, the Table 81 revealed some hints on the direction of motorization. For example, patterns of "W → B" and "B → M" appeared not only when

looking forward, but also when looking backward.

In addition to the partial look at motorization direction of Table 81, in Chapter 3, I have proposed a hypothetical direction of motorization pathway (Figure 22).

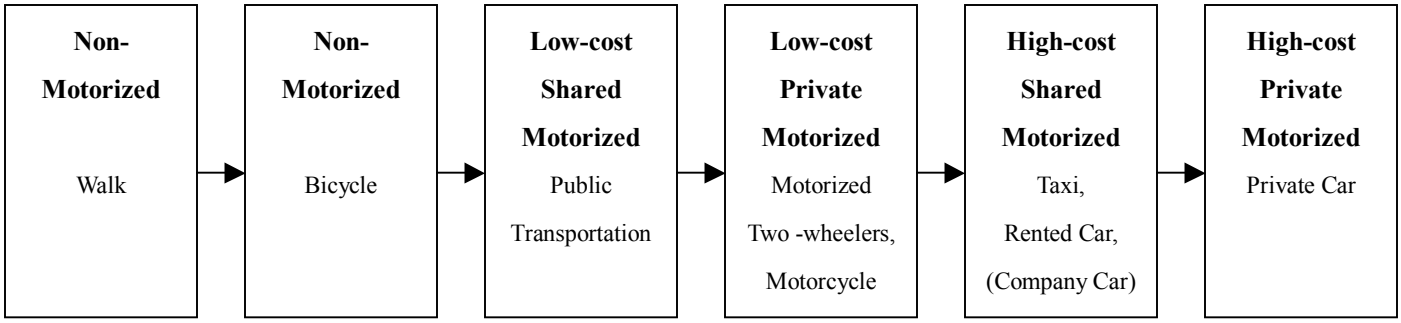


Figure 22: Hypothetical Motorization Direction

Based on my assumption, each travel mode of my motorizations pathway question is placed into a specific box in Figure 22, except for the company car. I assume company car can happen either after or at the same stage as the taxi or rented car, since company car basically serves as an intermediate stage between "taxi or rented car" and "car".

To understand the motorization direction, I checked the motorization pathway patterns of the entire sample (992 cases) to see if they match with the hypothetical direction of "W → B → P → M → T → S → C" or "W → B → P → M → T / S → C.[27]" Table 82 shows that basically more than half (53% if including single-stage pathway, 55% if not) of the sample follow my hypothetical motorization direction. However, there are also considerable people (43% if including single-stage pathway, 45% if not) not having the motorization pathway as I expected.

[27] No need to have the same number of stages; only need to match with the sequence.

Follow "Hypothetical Motorization Direction?"	Counts	Percentage of 992 cases	Percentage of 992 cases (without single-stage Pathway)
Single-stage and Single-mode Pathway (with 5 patterns)	44	4%	–
Yes (with 78 patterns)	523	53%	55%
No (with 248 patterns)	425	43%	45%

Table 82: Check of Hypothetical Motorization Direction

In order to understand more about the two groups (follow vs. not follow the hypothetical direction), I developed the pathway tree diagrams for each group. The looking-backward scenario was used to examine the pathways last two to four stages. Compared to looking-forward, the looking-backward scenario is favorable because it starts from what people currently do, then traces backward. The looking-backward scenario also focuses on the recent history of motorization in Shanghai, which I am more interested in.

Motorization Pathway currently Ends with "Walk"

I re-examined the group who end motorization with "walk". Since I hypothesized walk as the initial stage of pathway, all the cases in this groups (53 cases, excluding the single-stage pathway) are considered not following the expected direction. As in Table 83, "B → P → W" (5 cases) is the top pattern, and branch ending with "BW" (14 cases) is the most frequent at the two-stage level. Based on my hypothetical motorization direction, "walk" is the very basic, non-motorized mode and should be at the beginning of pathway. Therefore, any pattern ending with "walk" (if not single-stage) is considered "not following" the hypothetical direction. Different types of "not following direction" are summarized at the end of this chapter.

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
MBW	**2**	**MBW**	**5**	**BW**	**14**
PMBW	**2**	**BW**	**4**	TW	13
TMBW	**1**	**PBW**	**2**	PW	10
BW	**4**	**CBW**	**2**	MW	7
CPBW	**1**	**TBW**	**1**	SW	7
MPBW	**1**	PTW	9	CW	2
CBW	**1**	CTW	2		
TCBW	**1**	MTW	2		
PTBW	**1**	BPW	6		
PTW	4	TPW	2	←	
BPTW	2	MPW	1	(consider stages backward)	
MPTW	2	SPW	1		
CPTW	1	MW	3		
CTW	2	CMW	2		
CMTW	1	PMW	2		
MTW	1	MSW	2		
BPW	**5**	TSW	2		
MBPW	1	BSW	1		
CTPW	1	CSW	1		
STPW	1	SW	1		
BMPW	1	BCW	1		
TSPW	1	TCW	1		
MW	3				
CMW	2				
BPMW	1				
CPMW	1				
MSW	1				
TMSW	1				
BTSW	1				
TSW	1				
BSW	1				
CSW	1				
SW	1				
BCW	1				
STCW	1				
N (4+ stages) = 23 N (3 stages) = 22 N (2 stages) = 8		N (3+ stages) = 45 N (2 stages) = 8		N (2+ stages) = 53	

Table 83: Tree Diagram (End with "Walk", Look Backward, Don't Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Bicycle"

Table 84 and Table 85 show the patterns of people who have "bicycle" as the final motorization stage, but they are separated into groups of following and not-following the hypothetical motorization direction. Unlike previous tree diagrams, the total numbers of cases change under different levels of data consolidation. For example, the only pattern

(or branch) following the direction is "WB" at the two-stage level with 27 cases (Table 85). However, at the three-stage level, the total cases with the expected direction become 25 – two cases of pattern "P → W → B" are screened out. In Table 84, "PB" (10 cases) and "TB" (9 cases) are two big branches; and "W → P → B" (6 cases), "W → P → T → B" (4 cases), "M → B" (4 cases) are three frequent patterns.

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
PWB	1	PWB	2	**PB**	**10**
TPWB	1	**WPB**	**6**	TB	9
WPB	**6**	**PB**	**2**	MB	7
PB	**2**	**MPB**	**1**	CB	2
WMPB	**1**	**TPB**	**1**	(CS)B	1
TPB	**1**	PTB	6	SB	1
WPTB	4	MTB	2		
PTB	2	WTB	1		
MTB	2	MB	4		
PWTB	1	SMB	2	←	
MB	4	CMB	1	(consider stages backward)	
TSMB	2	SCB	2		
SCMB	1	W(CS)B	1		
MSCB	1	PSB	1		
TSCB	1				
TW(CS)B	1				
TPSB	1				
N (4+ stages) = 14 N (3 stages) = 12 N (2 stages) = 6		N (3+ stages) = 26 N (2 stages) = 6		N (2+ stages) = 30	

Table 84: Tree Diagram (End with "Bicycle", Look Backward, Don't Follow Hypothetical Direction)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WB	**25**	**WB**	**25**	**WB**	**27**
				← (consider stages backward)	
N (2 stages) = 25		N (2 stages) = 25		N (2+ stages) = 27	

Table 85: Tree Diagram (End with "Bicycle", Look Backward, Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Public Transportation"

In Table 86, I found the most frequent branch is "MP" (13 cases); however, the two popular pattern are "B → W → P" (9 cases) and "W → B → M → P" (6 cases). Likely, in Table 87, the most frequent branch is "BP" (38 cases) and the top pattern is "W → B → P."

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WMBP	1	MBP	1	**MP**	**13**
TBP	1	TBP	1	TP	6
BWP	**9**	BWP	10	SP	3
MBWP	1	MWP	1	CP	2
MWP	1	TWP	1		
CTWP	1	**BMP**	**8**		
WBMP	**6**	**MP**	**3**		
BMP	**2**	**SMP**	**1**	←	
MP	**3**	**WMP**	**1**	(consider stages backward)	
TSMP	**1**	MTP	3		
WMP	**1**	BTP	2		
BMTP	2	(BC)TP	1		
MTP	1	TSP	3		
BTP	1	BCP	1		
WBTP	1	WCP	1		
(BC)TP	1				
BTSP	1				
CTSP	1				
WTSP	1				
BCP	1				
BWCP	1				
N (4+ stages) = 18 N (3 stages) = 17 N (2 stages) = 3		N (3+ stages) = 35 N (2 stages) = 3		N (2+ stages) = 24	

Table 86: Tree Diagram (End with "Public Transportation", Look Backward, Don't Follow Hypothetical Direction)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WBP	**30**	**WBP**	**30**	**BP**	**38**
BP	**6**	**BP**	**6**	WP	29
WP	17	WP	17		
				←	
				(consider stages backward)	
N (3 stages) = 30 N (2 stages) = 23		N (3+ stages) = 30 N (2 stages) = 23		N (2+ stages) = 67	

Table 87: Tree Diagram (End with "Public Transportation", Look Backward, Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Motorized Two-wheeler or Motorcycle"

Considering people who don't follow expected direction, Table 88 shows a big branch of "TM" (23 cases) and top pattern "B → P → T → M" (19 cases). For people following the direction, Table 89 presents top branch of "BM" (50 cases) and the most frequent pattern "B → M" (22 cases).

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WPBM	6	PBM	7	**TM**	**23**
PBM	1	TBM	1	CM	8
PTBM	1	**PTM**	**21**	SM	8
CBPM	1	**BTM**	**1**	B(PM)	1
BWPM	4	**CTM**	**1**		
BPTM	**19**	BCM	5		
WPTM	**2**	SCM	2		
BTM	**1**	CM	1	←	
PCTM	**1**	TSM	5	(consider stages backward)	
BCM	5	CSM	2		
TSCM	2	WSM	1		
CM	1	BWM	2		
PTSM	3	WB(PM)	1		
CTSM	1				
WTSM	1				
TCSM	2				
WSM	1				
BWM	1				
PBWM	1				
WB(PM)	1				
N (4+ stages) = 39 N (3 stages) = 10 N (2 stages) = 1		N (3+ stages) = 49 N (2 stages) = 1		N (2+ stages) = 40	

Table 88: Tree Diagram (End with "Motorized Two-wheeler or Motorcycle", Look Backward, Don't Follow Hypothetical Direction)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BM	**22**	**BM**	**22**	**BM**	**50**
WBM	**20**	**WBM**	**20**	PM	28
WBPM	18	BPM	22	WM	8
BPM	3	WPM	6		
WPM	2	WM	6		
WM	6			←	
				(consider stages backward)	
N (4+ stages) = 18 N (3 stages) = 25 N (2 stages) = 28		N (3+ stages) = 48 N (2 stages) = 28		N (2+ stages) = 86	

Table 89: Tree Diagram (End with "Motorized Two-wheeler or Motorcycle", Look Backward, Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Taxi or Rented Car"

In Table 90, many patterns with small cases represent a diverse group. The top branch is "ST" (13 cases), and the "W → P → B → T" (12 cases), "B → M → P → T" (11 cases), "B → W → P → T" (10 cases) are all frequent patterns. For people who follow the hypothetical direction, it is a less diverse group. Branch "PT" dominates at the two-stage level with 146 cases. "W → B → P → T" is the most frequent pattern at the four-stage level with 64 cases. (Table 91)

(WM)BPT	**1**	**MPT**	**12**	ST	13
MBPT	**1**	**CPT**	**3**	CT	9
BWPT	**10**	**SPT**	**2**	P(MT)	2
CWPT	**1**	PBT	14	(BM)T	1
MWPT	**1**	(WP)BT	1	(PM)T	1
BMPT	**11**	PWT	7	(WP)T	1
WMPT	**1**	BWT	3	C(PT)	1
BCPT	**2**	CWT	2	P(CT)	1
CPT	**1**	MWT	2		
MSPT	**2**	PST	7		
SPMT	1	BST	3		
PBMT	4	MST	3	←	
WPBT	**12**	CT	4	(consider stages backward)	
PBT	2	PCT	4		
CWBT	1	MCT	1		
(WP)BT	1	BP(MT)	2		
PWT	5	(BM)T	1		
BPWT	2	(BW)(PM)T	1		
PBWT	2	(WP)T	1		
BWT	1	(BW)C(PT)	1		
CWT	2	BP(CT)	1		
BMWT	1				
PMWT	1				
BPST	3				
WPST	2				
MPST	1				
PST	1				
BST	2				
WBST	1				
BMST	2				
CMST	1				
CT	4				
BPCT	2				
(BW)PCT	1				
MPCT	1				
PMCT	1				
WBP(MT)	2				
(BM)T	1				
(BW)(PM)T	1				
(WP)T	1				
(BW)C(PT)	1				
BP(CT)	1				
N (4+ stages) = 69 N (3 stages) = 20 N (2 stages) = 6		N (3+ stages) = 69 N (2 stages) = 6		N (2+ stages) = 29	

Table 90: Tree Diagram (End with "Taxi or Rented Car", Look Backward, Don't follow Hypothetical Direction)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
WBPT	**64**	**BPT**	**83**	**PT**	**146**
BPT	**17**	**WPT**	**45**	MT	32
WPT	**33**	**PT**	**1**	BT	21
PT	**1**	PMT	19	WT	17
BPMT	14	BMT	10		
PMT	2	WMT	2		
WPMT	2	MT	1		
BMT	3	BT	3	← (consider stages backward)	
WBMT	3	WBT	3		
WMT	2	WT	3		
MT	1				
BT	3				
WBT	2				
WT	3				
N (4+ stages) =124 N (3 stages) = 20 N (2 stages) = 6		N (3+ stages) =162 N (2 stages) = 8		N (2+ stages) = 216	

Table 91: Tree Diagram (End with "Taxi or Rented Car", Look Backward, Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Car"

Two groups presented in the following have similar number of patterns (even at different levels of data consolidation). However, in terms of number of cases, people who follow the expected direction certainly form a larger group. Table 92 show the top branch as "(PT)C" (2 cases) and the most frequent patterns as "T → M → S → C" (7 case). Besides, "M → S → T → C", "P → T → M → C", and "B → M → P → C" are all popular patterns with 5 cases.

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
MPTC	**3**	**STC**	**7**	(PT)C	2
PBTC	**1**	(PT)SC	1	(BP)C	1
MSTC	**5**	TMC	7	(MT)C	1
PSTC	**2**	MPC	7		
BWTC	**1**	TPC	6		
TMSC	7	SPC	1		
(BW)MSC	1	PBC	2	←	
PBSC	1	TBC	2	(consider stages backward)	
MPSC	2	SBC	1		
B(PT)SC	1	BWC	2		
PBMC	2	TWC	1		
SBMC	1	B(PT)C	1		
SPMC	1	M(PT)C	1		
PTMC	5	W(BP)C	1		
BTMC	1	(BPS)(MT)C	1		
STMC	1				
BMPC	5				
SMPC	2				
BTPC	3				
MTPC	2				
TPC	1				
BWPC	1				
TSPC	1				
PBC	1				
WPBC	1				
PTBC	1				
WTBC	1				
TSBC	1				
BWC	2				
TWC	1				
B(PT)C	1				
BM(PT)C	1				
W(BP)C	1				
W(BPS)(MT)C	1				
N (4+ stages) = 55 N (3 stages) = 7 N (2 stages) = 0		N (3+ stages) = 41 N (2 stages) = 0		N (2+ stages) = 4	

Table 92: Tree Diagram (End with "Car", Look Backward, Don't follow Hypothetical Direction)

For people following the expected direction, Table 93 indicates the top branch of "TC" (94 cases), which expands into the top pattern "B → P → T → C" (36 cases)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BPTC	**36**	**PTC**	**55**	**TC**	**94**
WPTC	**12**	**MTC**	**18**	SC	54
PTC	**4**	**BTC**	**7**	MC	39
PMTC	**10**	**WTC**	**6**	PC	36
BMTC	**7**	**TC**	**1**	BC	24
WMTC	**1**	TSC	33	WC	5
BTC	**4**	MSC	11		
WBTC	**2**	BSC	5		
WTC	**5**	PSC	4		
TC	**1**	BMC	20	←	
MTSC	18	PMC	9	(consider stages backward)	
PTSC	10	WMC	2		
BTSC	3	MC	1		
WTSC	2	BPC	16		
BMSC	3	WPC	5		
BSC	2	PC	1		
WBSC	2	BC	15		
BPSC	1	WBC	4		
PSC	1	WC	2		
BMC	10				
WBMC	7				
BPMC	8				
WMC	2				
MC	1				
WBPC	12				
BPC	4				
WPC	4				
PC	1				
BC	15				
WBC	4				
WC	2				
N (4+ stages) = 154 N (3 stages) = 35 N (2 stages) = 5		N (3+ stages) = 210 N (2 stages) = 5		N (2+ stages) = 252	

Table 93: Tree Diagram (End with "Car", Look Backward, Follow Hypothetical Direction)

Motorization Pathway currently Ends with "Company Car"

Table 94 shows a diverse group of people who are: currently at the "company car" stage but without having an expected motorization pathway. "CS" is the top branch (14 cases), indicating the interchangeability between company car and private car. "M → P → T → S" (9 cases) is the top pattern.

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
MPTS	**9**	TCS	9	CS	14
(BW)PTS	**1**	MCS	2	T(CS)	2
CPTS	**1**	PCS	2	(PT)S	2
PBTS	**6**	(MT)CS	1	M(CS)	2
MTCS	5	TMS	7	T(CMS)	2
PTCS	3	CMS	1	(PMT)(CS)	2
BTCS	1	TPS	4	(BWPTS)	1
BMCS	1	MPS	2	(PM)(TS)	1
TMCS	1	PBS	3		
TPCS	2	TBS	1		
(BP)(MT)CS	1	TWS	2		
PTMS	4	MWS	1	←	
BTMS	1	(PM)T(CS)	2	(consider stages backward)	
CTMS	1	B(PT)S	1		
WTMS	1	M(PT)S	1		
TCMS	1	BM(CS)	2		
CPMS	1	PT(CMS)	2		
MTPS	3	(BW)(PMT)(CS)	1		
TPS	1	B(PMT)(CS)	1		
BMPS	1	(BWPTS)	1		
TMPS	1	(WC)(PM)(TS)	1		
WPBS	2				
PBS	1				
CTBS	1				
BTWS	1				
CTWS	1				
BMWS	1				
B(PM)T(CS)	2				
WB(PT)S	1				
BM(PT)S	1				
BM(CS)	2				
WPT(CMS)	1				
BPT(CMS)	1				
(BW)(PMT)(CS)	1				
WB(PMT)(CS)	1				
(BWPTS)	1				
B(WC)(PM)(TS)	1				
N (4+ stages) = 60 N (3 stages) = 4 N (2 stages) = 0 N (1 stage) = 1		N (3+ stages) = 46 N (2 stages) = 0 N (1 stage) = 1		N (2+ stages) = 25 N (1 stage) = 1	

Table 94: Tree Diagram (End with "Company Car", Look Backward, Don't follow Hypothetical Direction)

Within Last 4 Stages	Counts	Within Last 3 Stages	Counts	Within Last 2 Stages	Counts
BPTS	**15**	**PTS**	**40**	**TS**	**59**
WPTS	**13**	**MTS**	**11**	MS	14
PTS	**1**	**BTS**	**8**	PS	12
PMTS	**6**	BMS	4	BS	4
BMTS	**4**	MS	1	WS	4
WMTS	**1**	PMS	1		
BTS	**1**	WPS	4		
WBTS	**1**	BPS	2		
BMS	3	WS	1	←	
WBMS	1			(consider stages backward)	
MS	1				
WPS	4				
WBPS	2				
WS	1				
N (4+ stages) = 43 N (3 stages) = 9 N (2 stages) = 2		N (3+ stages) = 70 N (2 stages) = 2		N (2+ stages) = 93	

Table 95: Tree Diagram (End with "Company Car", Look Backward, Follow Hypothetical Direction)

Table 95 presents a bigger (in terms of cases) but less diverse group of people who follow hypothetical direction. "TS" is the biggest branch (59 cases); the "B → P → T → S" (15 cases) and "W → P → T → S" (13 cases) are both frequent patterns.

Looking Backward **Follow Hypothetical Direction**		**Looking Backward** ***Don't* Follow Hypothetical Direction**	
Ending Stage	**Most Frequent Branch and Pattern within Group**	**Ending Stage**	**Most Frequent Branch and Pattern within Group**
Walk	–	**Walk**	[14] B → W [5] B → P → W
Bicycle	[27] W → B [25] B	**Bicycle**	[10] P → B [25] $ W → P → B
Public Transportation	[38] B → P [30] W → B → P	**Public Transportation**	[13] M → P [9] B → W → P
Motorized Two-wheeler / Motorcycle	[50] B → M [22] B → M	**Motorized Two-wheeler / Motorcycle**	[23]* T → M [19] B → P → T→ M
Taxi / Rented Car	[146] * P → T [64] $ W→ B→ P → T	**Taxi / Rented Car**	[13] S → T [12] W→ P→ B → T
Car	[94] T →C [36] B → P → T →C	**Car**	[2] (PT) →C [7] T → M → S →C
Company Car	[59] T →S [15] B → P → T →S	**Company Car**	[14] C →S [9] M → P → T →S
*: Largest branch (consider first 2 stages) across sample $: Largest pattern (consider first 4 stages) across sample [ ] Numbers in bracket indicate the number of cases		*: Largest branch (consider last 2 stages) across sample $: Largest pattern (consider last 4 stages) across sample [ ] Numbers in bracket indicate the number of cases	

Table 96: Summary of Tree Diagrams (Follow vs. Don't Follow Hypothetical Motorization Direction)

Table 96 summaries the results of tree diagrams under the looking-backward scenario. The entire sample were classified by ending motorization stage and "follow vs. don't

follow the hypothetical motorization direction". Number of cases, most frequent patterns and branches (within the group and across the sample) were presented.

As shown in Table 96, the most frequent patterns and branches for people who follow the expected motorization direction are basically the same as what I identified for the pooled data in Table 81. Reasonably, for people who do not follow the expected direction, their motorization patterns are more complicated compared to people who follow the expected direction. This situation matches with the finding in Table 82, that is, given the similar number of cases (follow vs. don't follow the direction), people who don't follow the motorization direction have three times more patterns than people who follow. Both findings above indicate that people who do not follow the hypothetical motorization direction is a more diverse group.

Table 97 shows the check of top patterns and branches (for people who do not follow the hypothetical motorization direction) against my three assumed rules of transition in Figure 22 – from non-motorized to motorized, from low cost to high cost, and from shared to private-owned. For example, bicycle is considered "higher cost" compared to walk, although both are non-motorized travel means. Thus, the branch "B → W" violates the rule of low cost to high cost. Apparently, based on the following table, most people's *unexpected* motorization pathways involve the violation of low cost to high cost. There are two possible explanations: first, the definition of so-called high-cost means and low-cost means may not be as I assumed. That is, walk may be perceived as a higher *cost* travel means by respondents than bicycle since walking requires more time (assuming that people have a non-zero value of time) and effort involved. Second, even if the cost

definition of travel means is what I assumed; the direction of transition may not be as hypothesized. Taking the "S → T" as an example, a company car may be purchased (and shared) for status purpose, i.e., to represent the company, before the company staff go for the taxi – a lower (purchase, operating) cost travel means. In a developing city like Shanghai, I assume that the status-seeking is a factor affecting the motorization pathway as well as the vehicle purchase/use behavior.

Look Backward ***Don't*** **Follow Hypothetical Direction**		**Against Three Hypothetical Rules of Motorization Direction?**		
Ending Stage	**Most Frequent Branch and Pattern within Group**	**"Non-motorized"→ "Motorized"**	**"Low-cost" → "High-cost"**	**"Shared" → "Private-owned"**
Walk	[14] B → W		x	
	[5] B → P → W	x	x	
Bicycle	[10] P → B	x		
	[25] W → P → B	x		
Public Transportation	[13] M → P			x
	[9] B → W → P		x	
Motorized Two-wheeler / Motorcycle	[23] T → M		x	
	[19] B → P → T→ M		x	
Taxi / Rented Car	[13] S → T		x	
	[12] W→ P→ B → T	x		
Car	[2] (PT) → C		x	
	[7] T → M → S →C		x	
Company Car	[14] C → S			x
	[9] M → P → T →S			x

Table 97: Patterns against Three Hypothetical Rules of Motorization Direction

The last analysis of this chapter is to understand about – Who are those people following (and not following) the hypothetical motorization direction?

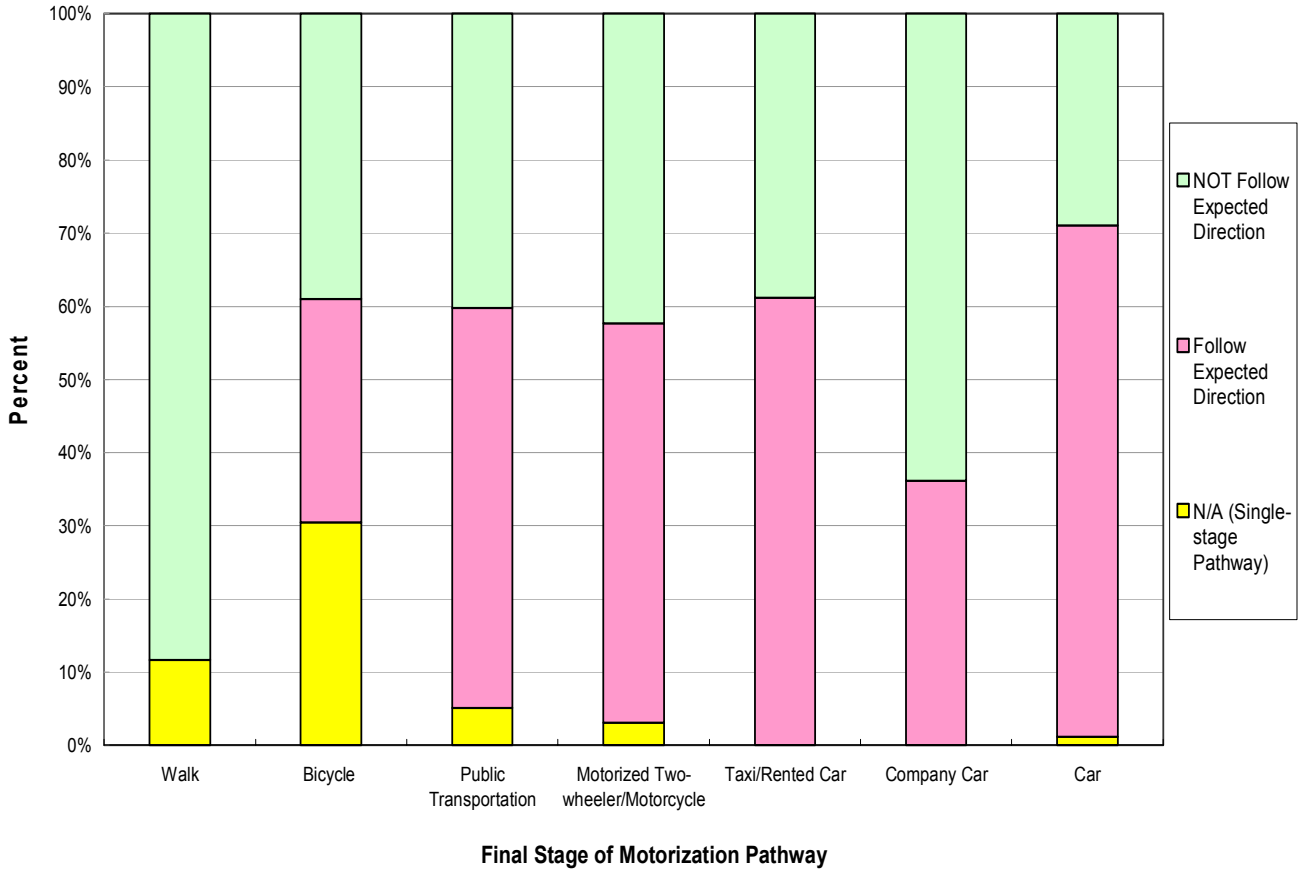


Figure 23: Final Stage of Motorization vs. Hypothetical Motorization Direction

Count

		Final Stage of Motorization							
		Walk	Bicycle	Public Transportation	Motorized Two-wheeler & Motorcycle	Taxi & Rented Car	Company Car	Car	Total
Follow Motorization Direction	Yes	0	25	53	71	150	43	181	523
	No	53	32	39	55	95	76	75	425
	N/A	7	25	5	4	0	0	3	44
Total		60	82	97	130	245	119	259	992

Table 98: Cross-tabulation: Final Stage of Motorization vs. Hypothetical Motorization Direction

Figure 23 and Table 98 show the bar chart and cross-tabulation of the "Final Stage of Motorization *vs.* Hypothetical Motorization Direction." A Pearson chi-squared test indicates that there are significant differences in the distributions of "following the hypothetical direction or not" across groups with different final motorization stages (p-value = 0.000 < 0.05). The group ending motorization with car has the highest percentage of people following the hypothetical motorization direction. However, for people currently in the walk or bicycle stage, most of them (in terms of percentage) did not experience a motorization as I expected. Basically, most of the highly-motorized people, i.e., people who are currently in advanced motorized means such as car, come from a typical (or expected) pathway[28], except for people currently using company car. Although company cars sometimes are considered functionally-similar to private-owned cars, people currently using company cars are more likely to have experienced different (or more diverse) pathways from people driving their own car.

[28] On the other hand, people who did not experience a typical motorization pathway may also have had car as one stage in their pathway (just not the final stage).

CHAPTER 5: VEHICLE PURCHASE AND USE MODELS

In this chapter I develop the vehicle purchase and use choice models based on the survey data. Factor analysis (as a pre-step of the model development), data preparation and model estimation are documented, and the implication of model results is discussed.

Factor Analysis

The main purpose of factor analysis is to reduce the number of variables to be used in the purchase and use models. Based on the pilot survey and my understanding of Chinese culture, Chinese people tend to avoid responding (or provide vague answers) to lifestyle-related questions as such questions are usually considered too personal, especially some sensitive topics, such as: status seeking lifestyle. To better measure lifestyle-related concepts, several questions were asked in the survey. I conducted factor analysis on these several measures in order to reduce the number of variables. By triangulating lifestyle and exogenous environment concepts, I attempt to measure them from several points of view in an effort to get a better overall assessment of these potentially sensitive ideas.

- **Questions about Exogenous Environment**

In PART III, 18 statements were presented to the respondents about the exogenous environment as related to vehicle purchase. People were asked to rate each exogenous item on a five-point (Likert) scale from "strongly disagree (1)" to "strongly agree (5)." In the questionnaire design phase, response to a pilot survey question *"Why did you*

purchase this vehicle?" was modified into 18 statement ratings in PART III. Some of these statements are related to similar underlying ideas expected to form potential factors resulting from a factor analysis. The statements and their expected factors are color-coded in (Table 99). Numbers in the first column of Table 99, Table 111, Table 112, and Table 124 represent the variable IDs.

PART III (original designed)	**Statement**	**Possible Factors**
125	The subsidy is important for my vehicle purchase	Cost-related Factor
126	The loan is important for my vehicle purchase	
131	The fuel price is what I care for my vehicle purchase	
127	The government fees affect my vehicle purchase	Government Policy
128	The new transportation law is what I need to consider	
140	Policy restrictions (e.g. ban of van using expressway) affect my vehicle purchase	
130	Owning a car is a pre-requisite for marriage	Culture Belief
132	Buying a house should be prior to vehicle purchase.	
135	Saving is a virtue, and we should always not spend too much	
133	Transportation environment in Shanghai is dangerous for "pedestrians"	Dangerous Transportation Environment
134	Transportation environment in Shanghai is dangerous for "drivers"	
123	The F-1 game and promotion in Shanghai affect my vehicle purchase decision	
124	I consider buying a vehicle for infrequent need.	
129	Buying a (domestic) car shows a direct support to China's auto industry.	
136	A vehicle which makes me look "better" will also bring me better career and social life	
137	Vehicle is just a business tool	
138	Vehicle (especially car) can depreciate very quickly	
139	In peak hours or rainy day, I wish I have personal vehicle	

Table 99: Original Design of Survey PART III

Determine the Number of Factors

Factor analysis includes creating and interpreting factors. Table 100 shows the descriptive statistics; the working sample (N) is 969 (instead of the original 1,037) due to the selection of "Exclude Cases Listwise." In addition, the Sig. (p-value) of the "Bartlett's

Test of Sphericity" is 0.00, and the "Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy" is 0.718[29] (Table 101). These imply that there exists underlying factors (Bartlett's Test of Sphericity) in these input scales and that the sample size of this input is acceptable (KMO Measure of Sampling Adequacy).

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
III_V123	2.28	.872	**969**
V124	2.33	.916	**969**
V125	3.16	1.066	**969**
V126	2.88	1.031	**969**
V127	3.47	1.052	**969**
V128	3.19	1.000	**969**
V129	3.19	1.067	**969**
V130	2.45	.989	**969**
V131	3.54	.960	**969**
V132	3.58	1.112	**969**
V133	3.19	1.013	**969**
V134	3.14	1.019	**969**
V135	3.69	.987	**969**
V136	3.31	.986	**969**
V137	2.73	.967	**969**
V138	3.22	.945	**969**
V139	3.97	.855	**969**
V140_III	3.36	.944	**969**

Table 100

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		**.718**
Bartlett's Test of Sphericity	Approx. Chi-Square	2662.800
	df	153
	Sig.	**.000**

Table 101

Principal axis factoring (common factor analysis) was conducted to extract factors with

[29] "KMO Test of Sampling Adequacy": 0.8 = meritorious; 0.7 = middling

maximum iterations for convergence = 100. I first tried the "Eigenvalues > 1" as the criterion of factor extraction. The Total Variance Explained (Table 102) shows that there are six "initial Eigenvalues" greater than 1, while there are only two "final Eigenvalues" (referred by SPSS as "Extraction Sums of Squared Loadings") greater than 1.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	**3.174**	17.635	17.635	**2.552**	14.181	14.181
2	**1.912**	10.624	28.259	**1.448**	8.047	22.227
3	**1.519**	8.439	36.698	.900	4.998	27.226
4	**1.196**	6.642	43.340	.583	3.238	30.463
5	**1.158**	6.433	49.772	.423	2.352	32.815
6	**1.050**	5.831	55.603	.339	1.881	34.696
7	.958	5.320	60.923			
8	.911	5.062	65.985			
9	.839	4.660	70.646			
10	.754	4.187	74.833			
11	.737	4.094	78.927			
12	.694	3.855	82.782			
13	.640	3.556	86.338			
14	.589	3.273	89.611			
15	.556	3.089	92.701			
16	.502	2.790	95.491			
17	.491	2.730	98.221			
18	.320	1.779	100.000			

Extraction Method: Principal Axis Factoring.

Table 102

The initial (suggested six factors) and final (suggested two factors) Eigenvalues over 1 are criteria purely based on the statistical consideration. However, in order to check the possible solutions more thoroughly, factor analysis was re-run by pre-specifying the number of factors to be, respectively, two through six; and each solution was examined individually for both statistical consideration and conceptual strength.

Table 103 through Table 107 present the factor loading matrices with Varimax rotation[30]. In the following Factor Matrices, high loading values were highlighted; however, I suppressed absolute loadings less than 0.25 (as a rule of thumb). If a given variable has no loadings greater than 0.25, the variable is not contributing much to common factor space and could be discarded from the factor analysis, even if retained for other purposes such as inclusion as an explanatory variable in its own right. Similarly, if a factor has no loadings greater than 0.25, it is a rather weak factor, and we might look for a solution containing fewer factors.

Rotated Factor Matrix[a]

	Factor	
	1	2
III_V123	.347	
V124	.301	
V125	**.590**	
V126	**.555**	
V127	**.541**	
V128	**.525**	
V129		
V130	.271	
V131	**.500**	
V132		
V133		**.714**
V134		**.810**
V135		.350
V136	.280	
V137		
V138		.262
V139		.253
V140_III	.356	

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

Table 103

Rotated Factor Matrix [a]

	Factor		
	1	2	3
III_V123			**.620**
V124			**.605**
V125	**.536**		
V126	**.483**		.259
V127	**.629**		
V128	**.548**		
V129			
V130			.357
V131	**.557**		
V132			
V133		**.711**	
V134		**.833**	
V135		.331	
V136			
V137			
V138		.251	
V139	.273		
V140_III	.384		

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Table 104

[30] Compared to the un-rotated solution, we always prefer the rotated solution because it provides more interpretable structure (the dimensions of factor space are more closely aligned with clusters of factors, resulting in more loadings that tend to be either high in magnitude, or close to zero).

Rotated Factor Matrix[a]

	Factor			
	1	2	3	4
III_V123			**.601**	
V124			**.596**	
V125	**.541**			
V126	**.470**			
V127	**.679**			
V128	**.556**			
V129				.328
V130			.367	.258
V131	**.556**			
V132				
V133		**.707**		
V134		**.889**		
V135				.325
V136				**.419**
V137				
V138				.262
V139				**.425**
V140_III	.348			

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Table 105

Rotated Factor Matrix[a]

	Factor				
	1	2	3	4	5
III_V123			**.590**		
V124			**.620**		
V125	**.540**				
V126	**.472**			.295	
V127	**.674**				
V128	**.548**				
V129				.437	
V130			.307	.352	
V131	**.569**				
V132					
V133		**.669**			
V134		**.939**			
V135				.314	
V136				.376	
V137					.285
V138					.519
V139					.344
V140_III	.349				

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table 106

Rotated Factor Matrix[a]

	Factor					
	1	2	3	4	5	6
III_V123			**.603**			
V124			**.634**			
V125	**.495**		.266			
V126	**.467**			.306		
V127	**.686**					
V128	**.577**					
V129				**.454**		
V130			.283	**.352**		
V131	**.527**				.263	
V132					.272	
V133		**.710**				
V134		**.886**				
V135				.292		
V136				**.352**		
V137						.457
V138						.426
V139					.613	
V140_III	.367					

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.

Table 107

As mentioned, statistical consideration and conceptual strength were main criteria for determination of number of factors. Based on the two criteria, the "4-factor solution" (Table 105) seems to be the most preferred because the loadings are generally high enough and distributed evenly across factors. And, more importantly, the "4-factor solution" is conceptually interpretable (will be discussed later).

After deciding on the "4-factor solution", I found out a number of the variables actually have low communalities, e.g. the V132, V137, and V138 (Table 108). Those variables have little in common with the rest of the variables, and they did not load heavily on any factor in the preferred 4-factor solution. Therefore, I removed them and re-did the factor analysis. Table 109 shows the communalities after removing V132, V137, V138, and Table 110 presents the final rotated factor matrix.

Communalities

	Initial	Extraction
III_V123	.235	.385
V124	.226	.369
V125	.301	.353
V126	.273	.298
V127	.312	.465
V128	.261	.326
V129	.144	.137
V130	.170	.206
V131	.263	.320
V132	.074	**.061**
V133	.460	.533
V134	.474	.823
V135	.164	.176
V136	.151	.232
V137	.108	**.095**
V138	.114	**.096**
V139	.164	.240
V140_III	.178	.193

Extraction Method: Principal Axis Factoring.

Table 108: Original Communalities (4-factor solution, PART III)

Communalities

	Initial	Extraction
III_V123	.224	.365
V124	.220	.406
V125	.292	.347
V126	.268	.307
V127	.308	.495
V128	.262	.332
V129	.139	.191
V130	.169	.240
V131	.249	.305
V133	.458	.543
V134	.466	.821
V135	.158	.205
V136	.142	.211
V139	.130	.165
V140_III	.163	.182

Extraction Method: Principal Axis Factoring.

Table 109: Communalities after Removing V132, V137, V138 (4-factor solution, PART III)

Rotated Factor Matrix [a]

	Factor			
	1	2	3	4
III_V123			**.584**	
V124			**.629**	
V125	**.522**			
V126	.446			
V127	**.701**			
V128	**.558**			
V129				**.414**
V130			.356	.333
V131	**.544**			
V133		**.723**		
V134		**.897**		
V135				.337
V136				**.410**
V139				.329
V140_III	.355			

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Table 110: Final 4-Factor Solution of PART III (Rotated Factor Matrix)

Name the Factors

Table 111 shows the final 4-factor solution with factor loadings (in descending order within the factor), and the factors are presented in descending order by final Eigenvalues. The factor names (capturing the common meanings of the variables that load on each factor) are – "importance of cost-related policy/regulation", "dangerous transportation environment", "attraction of infrequent need/impromptu purchase", and "pro vehicle purchase cultural/social belief".

The factor solution in Table 111 can be compared to what I expected from the original questionnaire design (Table 99). Similarity and differences are discussed below. In order to better interpret future results of the choice model[31], those four factors were named with signs embedded. For example, instead of being named as "cultural belief", the fourth

[31] The factor scores will be used as explanatory variables in the vehicle purchase choice model.

factor was named as "*pro* vehicle purchase cultural/social belief". Detailed explanation of the factor names follows.

PART III (factor analyzed)	Statement	Factor Loading	Factor Name
127	The government fees affect my vehicle purchase	0.701	Importance of Cost-related Policy/Regulation [Final Eigenvalue = 1.801]
128	The new transportation law is what I need to consider	0.558	
131	The fuel price is what I care for my vehicle purchase	0.544	
125	The subsidy is important for my vehicle purchase	0.522	
126	The loan is important for my vehicle purchase	0.446	
140	Policy restrictions (e.g. ban of van using expressway) affect my vehicle purchase	0.355	
134	Transportation environment in Shanghai is dangerous for "drivers"	0.897	Dangerous Transportation Environment [Final Eigenvalue = 1.414]
133	Transportation environment in Shanghai is dangerous for "pedestrians"	0.723	
124	I consider buying a vehicle for infrequent need.	0.629	Attraction of Infrequent Need/ Impromptu Purchase [Final Eigenvalue = 1.056]
123	The F-1 game and promotion in Shanghai affect my vehicle purchase decision	0.584	
130	Owning a car is a pre-requisite for marriage	0.356	
129	Buying a (domestic) car shows a direct support to China's auto industry.	0.414	Pro Vehicle Purchase Social/Culture Factor [Final Eigenvalue = 0.845]
136	A vehicle which makes me look "better" will also bring me better career and social life	0.410	
135	Saving is a virtue, and we should always not spend too much	0.337	
130	Owning a car is a pre-requisite for marriage	0.333	
139	In peak hours or rainy day, I wish I have personal vehicle	0.329	

Table 111: Final 4-Factor Solution of PART III (Factor Name)

Factor I: Importance of Cost-related Policy/Regulation

The first factor involves six variables related to: government fee, transportation law, fuel price, subsidy, loan, and policy restriction (e.g., vehicle license control). The government fee, with the highest loading (0.701), covered the concepts of all fees imposed by Shanghai government, especially the high vehicle license fee through the auction system. Explanations about the definition of the government fee were provided for respondents during the survey.

This factor was named "importance of cost-related policy/regulation" and it combined most of the variables that I originally expected to form two factors: one for cost and one for policy. People who score highly on this factor will tend to think exogenous cost-related policies/regulations are important. The "cost-related" policies/regulations factor, similar to the (monetary) operation cost, is assumed to have negative effect on vehicle purchase. The presence of this factor is not surprising as prices and costs are usually considered as a factor in conventional vehicle purchase model, especially in the developing countries. The "new transportation law" statement specifically means a new law in effect in 2006 saying that drivers have full financial responsibility whenever he/she is involved in an accident with a pedestrian. During the survey completion process, people received explanations if they were not clear about the definition of transportation law.

Factor II: Dangerous Transportation Environment

As I expected, this factor includes two variables: dangerous transportation environment for driver and for pedestrian; both with high factor loadings (0.897 for driver-related variable, 0.723 for pedestrian-related variable). The extraction of this variable is statistically and conceptually reasonable. People who score highly on this factor will tend to consider the transportation environment in Shanghai as dangerous for drivers and/or pedestrians. Thus, this factor is assumed to have negative effect on vehicle purchase.

Factor III: Attraction of Infrequent Need/ Impromptu Purchase

The factor of "attraction of infrequent need/impromptu purchase" was not expected during the initial survey design phase; however, this result is still conceptually

interpretable. This factor combines variables about "car for infrequent need", "F1 (Formula One racing) promotion" and "car as prerequisite of marriage". The "F1 promotion" is certainly an attraction for impromptu vehicle purchase because it stimulates people's desire beyond the basic mobility needs. The marketing of F1 is successful (especially in young generation) in Shanghai in creating not only a short-term craze but a long-term "lifestyle" (Figure 24). The variable "car as prerequisite of marriage" was loaded in this factor, although it was also loaded in the following cultural-related factor. According to an article from Singtao Daily (Chinese newspaper published in North America), the "prerequisite" of marriage for Shanghai people has changed from previous "big three items" (TV, washing machine and refrigerator) to "house, bank saving, and car". In general, people who score highly on this factor will tend to consider those attractions of infrequent need/impromptu purchase as important, and this factor is assumed to have positive effect on vehicle purchase.

Figure 24: The F1 Life-style Store in Shanghai

Factor IV: Pro Vehicle Purchase Cultural/Social Belief

This is the last and the weakest factor according to its low Eigenvalue. This factor includes several social/cultural variables which are positive toward vehicle purchase including: "car brings in a better social life," "car as prerequisite to marriage," etc. Basically, people who score highly on this factor tend to have "pro-auto" social/cultural beliefs identified by each variable that loads on this factor, and this factor is assumed to have positive effect on vehicle purchase. However, there were still exceptions. For example, *"saving is a virtue, and we should always not spend too much"* (V135) was grouped into this *pro* vehicle purchase factor; nevertheless, its loading (0.337) is not very high.

- **Questions about Lifestyle**

The survey PART IV contained another 18 statements designed to identify six vehicle purchase-related lifestyle factors (Table 112). The idea of 18 statements were developed based on the responses to a pilot survey question *"Why did you purchase this vehicle [Personal Attribute]?"* , local auto advertising, and social studies about today's China. For instance, the statement of *"I am the king in my own territory"* (V148) was a popular advertising slogan, and the concept of "Pursue Freedom and Control of Life" was usually promoted by local automobile companies. Another example is, the "enjoy now" is an emerging lifestyle for the young generation in China, as reported in several local newspapers, such as Shanghai Daily. Besides, some statements (denoted by asterisks in Table 112) were (partially or completely) borrowed from (or inspired by) UC Davis professor Patricia Mokhtarian's class handouts (TTP 200, *Transportation Survey Methods*) and sample surveys (Telecommuting Survey, 1992; Mobility Survey, 1998; pre-test of

E-shopping Survey, 2006) distributed in class.

PART IV (original designed)	Statement	Possible Factors
141	Traveling with family is a happy thing*	Family-oriented
156	Family is the most important in my daily life*	
158	Going back home is such a joy after working a whole day*	
142	The vehicle that I own needs to be well-known. (brand, model, etc.)*	Status-seeking
152	I enjoy catching everybody's eyes*	
154	For me, a lot of the fun of having something nice is showing it off*	
143	Life is so short; we should enjoy life as soon as we can	Enjoy "NOW"
147	I think about getting a loan for my vehicle purchase	
149	Saving money is hard for me	
144	When many of my friends own certain type of vehicle, I will consider buying one*	Bandwagon Effect
146	In the past (or now), I am crazy about some idol	
151	I pay attention on fashion	
145	Eventually, everybody will pay the price for the pollution on earth*	Environmental-concerned
153	I am aware of the information about new fuel/fuel efficient vehicles*	
157	For environmental protection, I am willing to use walk/bike to replace the motor vehicle use	
148	"*I am the king in my own territory*" is a good concept.	Pursue Freedom & Control of Life
150	Freedom and the control of life is what I pursue*	
155	I don't like to share things with others*	

Table 112: Original Design of Survey PART IV

The survey asked respondents to rate each statement on a five-point (Likert) scale from "strongly disagree (1)" to "strongly agree (5)." Our respondents found it difficult to comment on certain statements because those statements have negative implications or were too personal. For example, people tend to not reveal opinions toward the "status-seeking" and "enjoy now" lifestyles, because they are sometimes related to negative images of "being greedy" and "living without a goal". To remedy this problem, three statements (instead of one) were used to triangulate one factor.

Determine the Number of Factors

Similar to PART III, the working sample (N) is 968 (instead of the original 1,037) due to the selection of "Exclude Cases Listwise" (Table 113). The Sig. (p-value) of the "Bartlett's Test of Sphericity" is 0.00, indicating that there *are* underlying factors. The "Kaiser-Meyer-Olkin Measure of Sampling Adequacy" is 0.790, showing that the sample size of this input is acceptable (Table 114).

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
VI_V141	4.26	.743	**968**
V142	3.26	.984	**968**
V143	3.21	1.022	**968**
V144	3.14	.892	**968**
V145	3.92	.890	**968**
V146	2.93	1.017	**968**
V147	3.02	1.042	**968**
V148	3.10	1.002	**968**
V149	2.83	.941	**968**
V150	3.80	.789	**968**
V151	3.50	.883	**968**
V152	3.04	.881	**968**
V153	3.27	.900	**968**
V154	2.56	.947	**968**
V155	2.95	.977	**968**
V156	4.02	.765	**968**
V157	3.45	1.003	**968**
V158_VI	4.20	.738	**968**

Table 113

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		**.790**
Bartlett's Test of Sphericity	Approx. Chi-Square	2530.511
	df	153
	Sig.	**.000**

Table 114

To extract factors, the principal axis factoring (common factor analysis) was performed with maximum iterations for convergence = 100. The "Eigenvalues over 1" was initially

selected as the criterion of factor extraction. Table 115 indicates that there are five "initial Eigenvalues" over 1, while there are two "final Eigenvalues" (Extraction Sums of Squared Loadings) over 1.

Total Variance Explained

Factor	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	**3.308**	18.378	18.378	**2.671**	14.841	14.841
2	**2.195**	12.194	30.572	**1.572**	8.734	23.575
3	**1.217**	6.762	37.334	.503	2.795	26.370
4	**1.168**	6.491	43.825	.483	2.682	29.053
5	**1.077**	5.985	49.811	.387	2.150	31.203
6	.996	5.531	55.342			
7	.879	4.883	60.225			
8	.842	4.678	64.903			
9	.795	4.417	69.320			
10	.786	4.366	73.686			
11	.697	3.874	77.559			
12	.690	3.834	81.393			
13	.651	3.617	85.011			
14	.629	3.497	88.508			
15	.556	3.088	91.596			
16	.539	2.992	94.588			
17	.502	2.788	97.375			
18	.472	2.625	100.000			

Extraction Method: Principal Axis Factoring.

Table 115

In order to review more possible solutions, I re-ran the factor analysis by pre-specifying the number of factors to be two through six. Table 116 through Table 120 present the factor loading matrices after Varimax rotation. Like PART III, high loading values were highlighted in red; and the absolute loadings less than 0.25 were suppressed in the following factor matrices.

Rotated Factor Matrix[a]

	Factor	
	1	2
VI_V141		**.600**
V142	.412	
V143	.390	
V144		
V145		**.458**
V146	.375	
V147	**.457**	
V148	**.459**	
V149	.313	
V150	.348	.393
V151	**.484**	
V152	**.663**	
V153	.286	
V154	**.518**	
V155	.257	
V156		**.565**
V157		.282
V158_VI		**.710**

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

Table 116

Rotated Factor Matrix[a]

	Factor		
	1	2	3
VI_V141	**.594**		
V142		.296	.292
V143		**.479**	
V144		.305	
V145	.461		
V146		.305	
V147		.343	.293
V148		**.405**	
V149		**.435**	
V150	.381	.262	
V151			**.643**
V152		.361	**.602**
V153			.371
V154		**.408**	.304
V155			
V156	**.565**		
V157	.287		
V158_VI	**.705**		

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 5 iterations.

Table 117

Rotated Factor Matrix[a]

	Factor			
	1	2	3	4
VI_V141	**.525**			.269
V142		.253		.286
V143			.392	.353
V144			.361	
V145	.431			
V146			.302	
V147		.280	.309	
V148			.304	.387
V149			**.407**	
V150				**.615**
V151		**.596**		.308
V152		**.609**	.342	
V153		.405		
V154		.338	**.441**	
V155			.256	
V156	**.630**			
V157	.352			
V158_VI	**.678**			

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Table 118

Rotated Factor Matrix[a]

	Factor				
	1	2	3	4	5
VI_V141	**.512**				.286
V142					**.547**
V143			**.401**	.295	
V144			.357		
V145	.465				
V146			.298		
V147		.305	.297		
V148			.294	.382	
V149			**.404**		
V150				**.624**	
V151		**.594**		.288	
V152		**.621**	.318		
V153		.399			
V154		.341	**.435**		
V155			.253		
V156	**.615**				
V157	.380				
V158_VI	**.666**				

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 6 iterations.

Table 119

Rotated Factor Matrix[a]

	Factor					
	1	2	3	4	5	6
VI_V141	**.516**				.329	
V142					**.616**	
V143			.372	.303		
V144						.442
V145	.448					
V146		.274				
V147		.348				
V148			.433			
V149						
V150	.298		**.622**			
V151		**.592**				
V152		**.638**				
V153		.401				
V154		.349		.425		
V155				.411		
V156	**.650**					
V157	.339					
V158_VI	**.685**					

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 11 iterations.

Table 120

Taking into account both statistical and conceptual strength, the "4-factor solution" seems to be the most preferred solution (Table 118). The "2-factor solution" and "3-factor solution" both have factors containing too many variables, thus making them hard to interpret. However, the "5-factor solution" and "6-factor solution" both have "weak" factors with too few variables.

Given the "4-factor solution," V155 (with a low communality of 0.082, Table 121) has little in common with the rest of the variables, and it did not load heavily on any factor in the preferred 4-factor solution. Therefore, I removed V155 and re-did the factor analysis. Table 122 shows the communalities after removing V155, and Table 123 presents the final rotated factor matrix.

Communalities

	Initial	Extraction
VI_V141	.281	.364
V142	.233	.224
V143	.204	.293
V144	.117	.172
V145	.189	.224
V146	.132	.148
V147	.179	.202
V148	.211	.273
V149	.115	.185
V150	.281	.461
V151	.295	.464
V152	.346	.511
V153	.119	.178
V154	.241	.329
V155	.089	**.082**
V156	.272	.398
V157	.106	.154
V158_VI	.347	.499

Extraction Method: Principal Axis Factoring.

Table 121: Original Communalities (4-factor solution, PART IV)

Communalities

	Initial	Extraction
VI_V141	.271	.355
V142	.228	.223
V143	.198	.275
V144	.115	.201
V145	.189	.229
V146	.129	.150
V147	.180	.213
V148	.211	.280
V149	.113	.181
V150	.278	.463
V151	.296	.464
V152	.345	.516
V153	.118	.175
V154	.226	.297
V156	.267	.382
V157	.108	.170
V158_VI	.347	.502

Extraction Method: Principal Axis Factoring.

Table 122: Communalities after Removing V155 (4-factor solution, PART IV)

Rotated Factor Matrix [a]

	Factor			
	1	2	3	4
VI_V141	**.529**		.251	
V142		.273	.302	
V143			**.369**	.350
V144				**.391**
V145	.440			
V146		.252		.282
V147		.309		.297
V148			**.389**	.298
V149				**.392**
V150			**.619**	
V151		**.594**	.311	
V152		**.642**		.274
V153		**.406**		
V154		.375		**.361**
V156	**.617**			
V157	.360			
V158_VI	**.682**			

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 9 iterations.

Table 123: Final 4-Factor Solution of PART IV (Rotated Factor Matrix)

Name the Factor

To appropriately name the underlying factors, I examine the factor loadings, which indicate the correlation between the underlying factors and the original variables/statements. Table 124 shows the final 4-factor solution with factor loadings sorted in descending order within the factor, and the factors are presented in descending order by final Eigenvalues. Four lifestyle factors were extracted and named based on the loading patterns: "family-oriented and *green*," "status-seeking," "pursue freedom and control of life," and "enjoy now."

PART IV (factor analyzed)	Statement	Factor Loading	Factor Name
158	Going back home is such a joy after working a whole day	0.682	Family-oriented and "Green" [Final Eigenvalue = 1.636]
156	Family is the most important in my daily life	0.617	
141	Traveling with family is a happy thing	0.529	
145	Eventually, everybody will pay the price for the pollution on earth	0.440	
157	For environmental protection, I am willing to use walk/bike to replace the motor vehicle use	0.360	
152	I enjoy catching everybody's eyes.	0.642	Status-seeking [Final Eigenvalue = 1.383]
151	I pay attention on fashion	0.594	
153	I am aware of the information about new fuel/fuel efficient vehicles	0.406	
154	For me, a lot of the fun of having something nice is showing it off	0.375	
147	I think about getting a loan for my vehicle purchase	0.309	
142	The vehicle that I own needs to be well-known. (brand, model, etc.)	0.273	
146	In the past (or now), I am crazy about some idol	0.252	
150	Freedom and the control of life is what I pursue	0.619	Pursue Freedom and Control of Life [Final Eigenvalue = 1.072]
148	"*I am the king in my own territory*" is a good concept.	0.389	
143	Life is so short; we should enjoy life as soon as we can	0.369	
151	I pay attention on fashion	0.311	
142	The vehicle that I own needs to be well-known. (brand, model, etc.)	0.302	
141	Traveling with family is a happy thing	0.251	
149	Saving money is hard for me	0.392	Enjoy "NOW" [Final Eigenvalue = 0.986]
144	When many of my friends own certain type of vehicle, I will consider buying one	0.391	
154	For me, a lot of the fun of having something nice is showing it off	0.361	
143	Life is so short; we should enjoy life as soon as we can	0.350	
148	"*I am the king in my own territory*" is a good concept.	0.298	
147	I think about getting a loan for my vehicle purchase	0.297	
146	In the past (or now), I am crazy about some idol	0.282	
152	I enjoy catching everybody's eyes.	0.274	

Table 124: Final 4-Factor Solution of PART IV (Factor Name)

Basically, the factor (names) in Table 124 differed from what I expected in constructing the questionnaire design (Table 112). Only four factors, instead of the originally expected six factors, were identified. A distinct "bandwagon effect" was not indentified: most of the variables expected to be related to a bandwagon effect instead loaded onto the factor of status-seeking. Also, the "environmental-concerned" and "family-oriented" lifestyles were combined as the "family-oriented and *green.*" Details about the factor interpretation and naming process follow.

Factor I: Family-oriented and Green

According to the factor loadings, the top three variables in this factor were the variables originally designed to triangulate the family-oriented lifestyle. The statement (variable): *"Going back home is such a joy after working whole day"* had the highest loading (0.682). The remaining two variables in this factor were related to the awareness of environmental protection, thus making the factor name "family-oriented and *green*".

The combination of the concepts of family-oriented and "green" is conceptually interpretable. In fact, as one example in Figure 25, Ford China attempted to link the two concepts – family-oriented and "love green" – and to market it in an internet commercial of SUV (an event of "driving your family to a nice country-side"). Basically, people who score highly on this factor are assumed to have a "compound" characteristic of being family-oriented and environmental-concerned.

Figure 25: Internet Commercial of Family-oriented and Green Life

Source: www.ford.com.cn

Factor II: Status-seeking

"The idea that individuals are motivated by status considerations is a very old one in economics and can be traced back to thinkers such as David Hume (1978) and Thorstein Veblen (1899)" (Fisher, 2001). Basically, by consuming (or owning) luxurious goods or services, people show off their wealth, power over others, or confirm certain social membership (e.g., the uniform cars for Chinese government officials) in order to obtain/secure more resources.

In my research, the status-seeking purchase was defined as the "consumption of luxurious goods for displaying" and modeled as lifestyle-related. As the result of factor analysis, three variables (V142, V152 and V154) originally designed for triangulating the

status-seeking lifestyle were all included in this factor, and the statement (V152) *"I enjoy catching everybody's eyes"* was assigned the highest loading (0.642). However, two variables (V146, V151) for testing the "bandwagon effect" were also included in this factor. The bandwagon effect indicated that people often do (or believe) things when they see *many* other people do (or believe) the same thing. One possible explanation for the bandwagon effect is that in a society in which culture and values are being redefined, people tend to "follow the crowd" before they finally come to a clear realization of their own (new) values. To certain extent, the bandwagon effect is similar to the status-seeking; for example, "following the crowd" can be considered similar to "confirming social membership". Besides, some variables originally designed for testing other factors were also combined into this factor, such as: V153 was for the environmental-concerned and V147 for the "enjoy now" lifestyle. Instead of naming one factor with compound concepts like in Factor I, the name "status-seeking" was eventually used because of the high loadings of (all three) status-seeking variables

The finding about the status-seeking lifestyle is not surprising, both Figure 26 and Figure 27 imply a status-seeking social environment. Figure 26 is the internet commercial of Lexus in China. The image and the Chinese slogan convey not only the functionality of the car, i.e., being sporty and powerful, but also the vehicle status, i.e., being on the top of peak. Figure 27 shows the official cars used for important government meeting in China. It is common to see the use of official (or uniform) cars to represent certain status, e.g., the Communalist Party, or to confirm social membership. Although the official government cars may not be purchased by common people, owning (or being assigned) such car can be a status symbol and a way to get special attention from the society.

Figure 26: Internet Commercial about Vehicle Status in China
Source: www.lexus.com.cn

Figure 27: Official Car Used in Important Government Meeting in China
Source: autos.cn.yahoo.com

<u>Factor III</u>: Pursue Freedom and Control of Life

This factor contains two variables/statements (V148, V150) originally designed for the

lifestyle of pursuing freedom[32]. The variables with the highest loading is *"Freedom and the control of life is what I pursue"* (V150, 0.619) and the variable with the second highest loading is *"I am the king in my own territory is a good concept"* (V148, 0.389). Although this factor also included other variables with lower loadings, I named this factor as "Pursue Freedom and Control of Life" based on statistical and conceptual considerations. This variable is assumed to have positive effect in the vehicle purchase.

Factor IV: Enjoy Now

The "Enjoy Now" is a recent, complicated but influential lifestyle in China. According to information from local media (TV, newspapers) during my stay in Shanghai, this type of lifestyle exists more among the young generation and has become a social phenomenon. One usually-cited example is certain young people's craze (or abuse) of the on-line games. Figure 28 is a picture of an internet café in Shanghai. Most such internet cafés are open 24 hours a day; and people (especially students) usually play on-line games there day and night. One explanation is that the young generation has difficulty finding a long-term life goal in a society undergoing rapid change. Therefore, the instant happiness; or even "virtual" instant happiness, became important in their lives. People with such an attitude or lifestyle tend to spend most of their money to just "enjoy now" without following the traditional Chinese virtue to save money. Sometimes, people with this type of lifestyle even borrow money (loan, credit card) to make their dreams, e.g., a car, happen faster.

[32] The third variable (V155) for the freedom-pursuing lifestyle was discarded due to the low communality.

Figure 28: Internet Café in Shanghai
Source: cn.yahoo.com

The factor analysis results show all three variables (V143, V147, and V149) originally designed for the "enjoy now" lifestyle were assigned to this factor. The variable with the highest loading is *"Saving money is hard for me"* (V149, 0.392). Many variables were not taken into account in the naming process due to low factor loadings. Basically, people who score highly on this factor will tend to have a characteristic of "enjoy now" or a "loose" spending habit.

Dependent Variables

In this chapter, the major work is to develop choice models for vehicle use and purchase. There is big overlap between vehicle use and purchase; however, some exceptions exist. For example, wealthy people may purchase a "status" car for displaying to friends but seldom drive it. Or, some people just want to use a vehicle without purchasing it, e.g.,

taking taxi. There were three questions in my survey about vehicle purchase and use as in the following.

- **Most Expensive Vehicle Owned**

The original question was *"Among all vehicle(s) you own, which one is the MOST EXPENSIVE? (check one)"*. The frequency distribution of responses was: bicycle (27.1%), motorized two-wheeler (17.26%), motorcycle (4.92%), car (26.9%), no vehicle owned (21.12%) and missing data (2.7%). Although the motorcycle was considered a small group (less than 5%), I still included it into the choice set based on some of its distinct mobility characteristics. Five choices (including purchase nothing) were the dependent variables in the model of most expensive vehicle owned.

- **Most Frequently Used Travel Means (weekday)**

In addition to vehicle purchase, vehicle use was modeled by considering both weekday and weekend travel patterns. The original question of weekday travel was *"Which type of transportation means below do you use mostly during week days? (check one)"*, and the frequency distribution of answers was: walk (8.58%), public transportation (29.89%), taxi (6.85%), rented car (0.68%), shared company car (3.47%), bicycle (14.37%), motorized two-wheeler (10.03%), motorcycle (2.89%), car (21.22%), other (0.39%) and missing data (1.64%). In the model estimation, choices with less than 5% cases were combined. Thus, the choice of rented car, taxi, and shared company car were combined, and motorcycle was combined with motorized two-wheeler. This "reasonable" combination was based on the motorization pathway hypothesized in Chapter 4 – from non-motorized to motorized, from low cost to high cost, and from shared to private-owned. Eventually,

the weekday vehicle use was modeled under six dependent variables.

- **Most Frequently Used Travel Means (weekend)**

The original question of weekend travel was *"Which type of transportation means below do you use mostly during weekend? (check one)"*, and the frequency distribution of answers was: walk (6.85%), public transportation (36.74%), taxi (16.30%), rented car (0.77%), shared company car (0.48%), bicycle (5.88%), motorized two-wheeler (5.69%), motorcycle (2.51%), car (22.76%), other (0.10%) and missing data (1.92%). The weekend vehicle use has a different pattern from the weekday; for instance, the less use of bicycle and the more use of taxi. Similar scheme was used to combine alternatives with less than 5% cases. As a result, the same six alternatives (as in the weekday case) were modeled for the weekend vehicle use.

Explanatory Variables

This section describes the explanatory variables used in the vehicle choice model. A total of 36 explanatory variables were used, including variables drawn directly from the survey and variables defined based on the post-processing of the survey data, e.g., factors. In general, the 36 explanatory variables were categorized into the following five groups.

- **General Vehicle Use/Purchase Background (4 variables)**

This group contains four variables from the survey PART I. The original questions are: *"Do you own any parking space AT WORK? (Y/N)"*, *"Do you own any parking space AT HOME? (Y/N)"*, *"How far is your WORKING PLACE to the closest subway line? (1=*

Within 500 meters, 2 = 500 ~ 1000 meters, 3 = More than 1001 meters, 4 = N/A, I don't work, I don't know, etc)", and *"How far is your HOME to the closest subway line? (1= Within 500 meters, 2 = 500 ~ 1000 meters, 3 = More than 1001 meters, 4 = N/A, I don't know)"*. The YES and NO were coded as 1 and 0 for the first two questions (and also throughout the whole survey). For the last two questions, the answers were treated as ordinal and coded as it were in the survey. Answers of N/A were replaced by the geographical mean (detail will be discussed in the data imputation section).

- **Perceived Utility of Travel Means (6 variables)**

In PART II of the survey, a series of statements were designed to understand respondents' perceived utility of various travel means. One sample statement is: *"I think the following travel means is – a 'SYMBOL OF SUCCESS' (1 = Strongly Disagree; 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree)"*. In PART II, people rated (in Likert scale) six types of utility (symbol of success/status, speed, availability, capacity, price, and comfort) for nine travel means (walk, public transportation, taxi, rented car, shared company car, bicycle, motorized two-wheeler, motorcycle, and car). Each type of utility rating was considered a *generic* variable, that is, one whose values could differ by alternative, but whose coefficient in the utility function could be equal across alternatives. In the event of data consolidation, certain utility ratings were combined using the "mean utility" in order to match with the dependent variables actually used in the model[33].

- **Exogenous Environment Factor (4 variables)**

[33] The modeling program (Limdep) requires the explanatory variables to be entered into the dataset for each alternative. Therefore, the numbers of sets of utility ratings and dependent variables have to match each other.

As discussed in the factor analysis, 18 variables (survey PART III) were factor-analyzed to generate four factors – "importance of cost-related policy/regulation", "dangerous transportation environment", "attraction of impromptu/infrequent purchase", and "pro vehicle purchase cultural/social belief". The factor scores were used as explanatory variables.

- **Lifestyle Factor (4 variables)**

Similarly, four types of lifestyle were extracted from 18 original variables (survey PART IV) – "family-oriented and green", "status-seeking", "pursue freedom and control of life", and "enjoy now". The factor scores were used as explanatory variables.

- **Demographics (18 variables)**

Demographic variables were drawn from the original 12 questions in the survey PART V. However, dummy variables were created for certain variables with nominal values. For example, "factory worker (dummy)" and "boss (dummy)" were derived from the nominal variable of "employment". Eventually, 18 demographical variables were used in the choice model – age, motorcycle license (dummy), car license (dummy), motorcycle and car license (dummy), male, education, not working/retired (dummy), student (dummy), factory worker (dummy), office worker (dummy), boss (dummy), personal income, household income, household size, urban[34] (dummy), Shanghai resident, year of living in Shanghai, and experience of living abroad. Table 125 presents the categorization of 36 explanatory variables used in the modeling process.

[34] This dummy variable (living in urban Shanghai) was created from the variable of living locations (by district).

General Vehicle Use/Purchase Background	parking at work, parking at home, distance from work to closest subway, distance from home to closest subway **(4 variables)**
Perceived Utility	perceived utility in terms of: symbol of success, speed, availability, capacity, price, comfort **(6 variables for each of 9 travel means)**
Exogenous Environment Factor	importance of cost-related policy/regulation, dangerous transportation environment, attraction of impromptu/infrequent purchase, pro vehicle purchase cultural/social belief **(4 variables)**
Lifestyle Factor	family-oriented and *green*, status-seeking, pursue freedom and control of life, enjoy now **(4 variables)**
Demographics	age, motorcycle license*, car license*, motorcycle and car license*, male, education, not working/retired*, student*, factory worker*, office worker*, boss*, personal income, family income, household size, urban*, Shanghai resident, year of living in Shanghai, experience of living abroad **(* indicates dummy variables) (18 variables)**

Table 125: Categorization of Explanatory Variables

Data Preparation

- **Distribution of Missing Data**

Before running the model, data cleaning, imputation, and consolidation were required. Understanding the distribution of missing data is the first step. Table 126, Table 127 , and Table 128 show the distribution of missing data for three datasets (one for each dependent variable). The first two columns (in Table 126, Table 127, and Table 128) show the "missing-valid pattern". A total of 79 responses will be provided (in each case) if no data is missing. Columns 3 to 6 show the (cumulative) frequency and (cumulative) percent associated with a specific missing-valid pattern. Columns 7 and 8 show the cumulative frequency and percent for the "remaining" cases; that is, the sum of frequency and "remaining" frequency should be 1,037 (total number of cases).

For data cleaning purposes, a cut-off threshold (missing more than eight responses) was selected by considering two criteria: the percentage of data left after excluding cases beyond the threshold (the "cumulative percent") and the percentage of imputed data (% of data imputed). The percentage of imputed data was calculated by dividing "cumulative frequency of total missing (column 10)" by "total number of cells" (i.e., cumulative frequency x 79).

Missing	Valid	Frequency	Cumulative Frequency	Percent	Cumulative Percent	*(Remaining)* Cumulative Frequency	*(Remaining)* Cumulative Percent	TOTAL MISSING (frequency)	TOTAL MISSING (cumulative frequency)	% of Data Imputed
0	79	481	481	46.4%	46.4%	556	53.6%	0	0	0.00%
1	78	124	605	12.0%	58.3%	432	41.7%	124	124	0.26%
2	77	185	790	17.8%	76.2%	247	23.8%	370	494	0.79%
3	76	69	859	6.7%	82.8%	178	17.2%	207	701	1.03%
4	75	47	906	4.5%	87.4%	131	12.6%	188	889	1.24%
5	74	35	941	3.4%	90.7%	96	9.3%	175	1064	1.43%
6	73	18	959	1.7%	92.5%	78	7.5%	108	1172	1.55%
7	72	14	973	1.4%	93.8%	64	6.2%	98	1270	1.65%
8	71	9	**982**	0.9%	**94.7%**	**55**	**5.3%**	72	1342	**1.73%**
9	70	11	993	1.1%	95.8%	44	4.2%	99	1441	1.84%
10	69	6	999	0.6%	96.3%	38	3.7%	60	1501	1.90%
11	68	3	1002	0.3%	96.6%	35	3.4%	33	1534	1.94%
12	67	2	1004	0.2%	96.8%	33	3.2%	24	1558	1.96%
13	66	2	1006	0.2%	97.0%	31	3.0%	26	1584	1.99%
14	65	2	1008	0.2%	97.2%	29	2.8%	28	1612	2.02%
15	64	5	1013	0.5%	97.7%	24	2.3%	75	1687	2.11%
17	62	1	1014	0.1%	97.8%	23	2.2%	17	1704	2.13%
18	61	2	1016	0.2%	98.0%	21	2.0%	36	1740	2.17%
19	60	1	1017	0.1%	98.1%	20	1.9%	19	1759	2.19%
20	59	2	1019	0.2%	98.3%	18	1.7%	40	1799	2.23%
22	57	1	1020	0.1%	98.4%	17	1.6%	22	1821	2.26%
29	50	1	1021	0.1%	98.5%	16	1.5%	29	1850	2.29%
43	36	1	1022	0.1%	98.6%	15	1.4%	43	1893	2.34%
44	35	1	1023	0.1%	98.6%	14	1.4%	44	1937	2.40%
45	34	1	1024	0.1%	98.7%	13	1.3%	45	1982	2.45%
47	32	2	1026	0.2%	98.9%	11	1.1%	94	2076	2.56%
48	31	1	1027	0.1%	99.0%	10	1.0%	48	2124	2.62%
49	30	2	1029	0.2%	99.2%	8	0.8%	98	2222	2.73%
50	29	1	1030	0.1%	99.3%	7	0.7%	50	2272	2.79%
51	28	2	1032	0.2%	99.5%	5	0.5%	102	2374	2.91%
52	27	1	1033	0.1%	99.6%	4	0.4%	52	2426	2.97%
54	25	1	1034	0.1%	99.7%	3	0.3%	54	2480	3.04%
58	21	1	1035	0.1%	99.8%	2	0.2%	58	2538	3.10%
59	20	1	1036	0.1%	99.9%	1	0.1%	59	2597	3.17%
63	16	1	1037	0.1%	100.0%	0	0.0%	63	2660	3.25%

Table 126: Distribution of Missing Data (Dependent Variable = Most Expensive Vehicle Owned)

Missing	Valid	Frequency	Cumulative Frequency	Percent	Cumulative Percent	*(Remaining)* Cumulative Frequency	*(Remaining)* Cumulative Percent	TOTAL MISSING (frequency)	TOTAL MISSING (cumulative frequency)	% of Data Imputed
0	79	484	484	46.7%	46.7%	553	53.3%	0	0	0.00%
1	78	121	605	11.7%	58.3%	432	41.7%	121	121	0.25%
2	77	186	791	17.9%	76.3%	246	23.7%	372	493	0.79%
3	76	68	859	6.6%	82.8%	178	17.2%	204	697	1.03%
4	75	50	909	4.8%	87.7%	128	12.3%	200	897	1.25%
5	74	33	942	3.2%	90.8%	95	9.2%	165	1062	1.43%
6	73	18	960	1.7%	92.6%	77	7.4%	108	1170	1.54%
7	72	13	973	1.3%	93.8%	64	6.2%	91	1261	1.64%
8	71	10	**983**	1.0%	**94.8%**	**54**	**5.2%**	80	1341	**1.73%**
9	70	11	994	1.1%	95.9%	43	4.1%	99	1440	1.83%
10	69	5	999	0.5%	96.3%	38	3.7%	50	1490	1.89%
11	68	3	1002	0.3%	96.6%	35	3.4%	33	1523	1.92%
12	67	2	1004	0.2%	96.8%	33	3.2%	24	1547	1.95%
13	66	2	1006	0.2%	97.0%	31	3.0%	26	1573	1.98%
14	65	2	1008	0.2%	97.2%	29	2.8%	28	1601	2.01%
15	64	5	1013	0.5%	97.7%	24	2.3%	75	1676	2.09%
17	62	1	1014	0.1%	97.8%	23	2.2%	17	1693	2.11%
18	61	2	1016	0.2%	98.0%	21	2.0%	36	1729	2.15%
19	60	1	1017	0.1%	98.1%	20	1.9%	19	1748	2.18%
20	59	2	1019	0.2%	98.3%	18	1.7%	40	1788	2.22%
22	57	1	1020	0.1%	98.4%	17	1.6%	22	1810	2.25%
29	50	1	1021	0.1%	98.5%	16	1.5%	29	1839	2.28%
43	36	1	1022	0.1%	98.6%	15	1.4%	43	1882	2.33%
44	35	1	1023	0.1%	98.6%	14	1.4%	44	1926	2.38%
45	34	1	1024	0.1%	98.7%	13	1.3%	45	1971	2.44%
47	32	2	1026	0.2%	98.9%	11	1.1%	94	2065	2.55%
48	31	1	1027	0.1%	99.0%	10	1.0%	48	2113	2.60%
49	30	2	1029	0.2%	99.2%	8	0.8%	98	2211	2.72%
50	29	1	1030	0.1%	99.3%	7	0.7%	50	2261	2.78%
51	28	2	1032	0.2%	99.5%	5	0.5%	102	2363	2.90%
52	27	1	1033	0.1%	99.6%	4	0.4%	52	2415	2.96%
54	25	1	1034	0.1%	99.7%	3	0.3%	54	2469	3.02%
58	21	1	1035	0.1%	99.8%	2	0.2%	58	2527	3.09%
59	20	1	1036	0.1%	99.9%	1	0.1%	59	2586	3.16%
63	16	1	1037	0.1%	100.0%	0	0.0%	63	2649	3.23%

Table 127: Distribution of Missing Data (Dependent Variable = Most Frequently Used Travel Means on Weekday)

Missing	Valid	Frequency	Cumulative Frequency	Percent	Cumulative Percent	*(Remaining)* Cumulative Frequency	*(Remaining)* Cumulative Percent	TOTAL MISSING (frequency)	TOTAL MISSING (cumulative frequency)	% of Data Imputed
0	79	485	485	46.8%	46.8%	552	53.2%	0	0	0.00%
1	78	121	606	11.7%	58.4%	431	41.6%	121	121	0.25%
2	77	185	791	17.8%	76.3%	246	23.7%	370	491	0.79%
3	76	68	859	6.6%	82.8%	178	17.2%	204	695	1.02%
4	75	49	908	4.7%	87.6%	129	12.4%	196	891	1.24%
5	74	34	942	3.3%	90.8%	95	9.2%	170	1061	1.43%
6	73	17	959	1.6%	92.5%	78	7.5%	102	1163	1.54%
7	72	14	973	1.4%	93.8%	64	6.2%	98	1261	1.64%
8	71	10	**983**	1.0%	**94.8%**	**54**	**5.2%**	80	1341	**1.73%**
9	70	11	994	1.1%	95.9%	43	4.1%	99	1440	1.83%
10	69	4	998	0.4%	96.2%	39	3.8%	40	1480	1.88%
11	68	4	1002	0.4%	96.6%	35	3.4%	44	1524	1.93%
12	67	2	1004	0.2%	96.8%	33	3.2%	24	1548	1.95%
13	66	2	1006	0.2%	97.0%	31	3.0%	26	1574	1.98%
14	65	2	1008	0.2%	97.2%	29	2.8%	28	1602	2.01%
15	64	3	1011	0.3%	97.5%	26	2.5%	45	1647	2.06%
16	63	2	1013	0.2%	97.7%	24	2.3%	32	1679	2.10%
17	62	1	1014	0.1%	97.8%	23	2.2%	17	1696	2.12%
18	61	2	1016	0.2%	98.0%	21	2.0%	36	1732	2.16%
19	60	1	1017	0.1%	98.1%	20	1.9%	19	1751	2.18%
20	59	2	1019	0.2%	98.3%	18	1.7%	40	1791	2.22%
22	57	1	1020	0.1%	98.4%	17	1.6%	22	1813	2.25%
29	50	1	1021	0.1%	98.5%	16	1.5%	29	1842	2.28%
43	36	1	1022	0.1%	98.6%	15	1.4%	43	1885	2.33%
44	35	1	1023	0.1%	98.6%	14	1.4%	44	1929	2.39%
45	34	1	1024	0.1%	98.7%	13	1.3%	45	1974	2.44%
47	32	2	1026	0.2%	98.9%	11	1.1%	94	2068	2.55%
48	31	1	1027	0.1%	99.0%	10	1.0%	48	2116	2.61%
49	30	2	1029	0.2%	99.2%	8	0.8%	98	2214	2.72%
50	29	1	1030	0.1%	99.3%	7	0.7%	50	2264	2.78%
51	28	2	1032	0.2%	99.5%	5	0.5%	102	2366	2.90%
52	27	1	1033	0.1%	99.6%	4	0.4%	52	2418	2.96%
54	25	1	1034	0.1%	99.7%	3	0.3%	54	2472	3.03%
58	21	1	1035	0.1%	99.8%	2	0.2%	58	2530	3.09%
59	20	1	1036	0.1%	99.9%	1	0.1%	59	2589	3.16%
63	16	1	1037	0.1%	100.0%	0	0.0%	63	2652	3.24%

Table 128: Distribution of Missing Data (Dependent Variable = Most Frequently Used Travel Means on Weekend)

- **Data Cleaning**

As mentioned, the cases with more than eight missing responses were discarded. By using this cut-off, for all three datasets, I was able to keep about 95% of the original cases (after the data cleaning) and impute less than 2% of the responses for those cases. Figure 29 shows the flow chart of the data preparation; it also indicates the sample size after the data cleaning (982, 983, and 983 for three dependent variables). However, in the model estimation, same datasets (982 cases for three dependent variables) were eventually used for the benefit of comparing the vehicle purchase and use behavior of the same group of respondents.

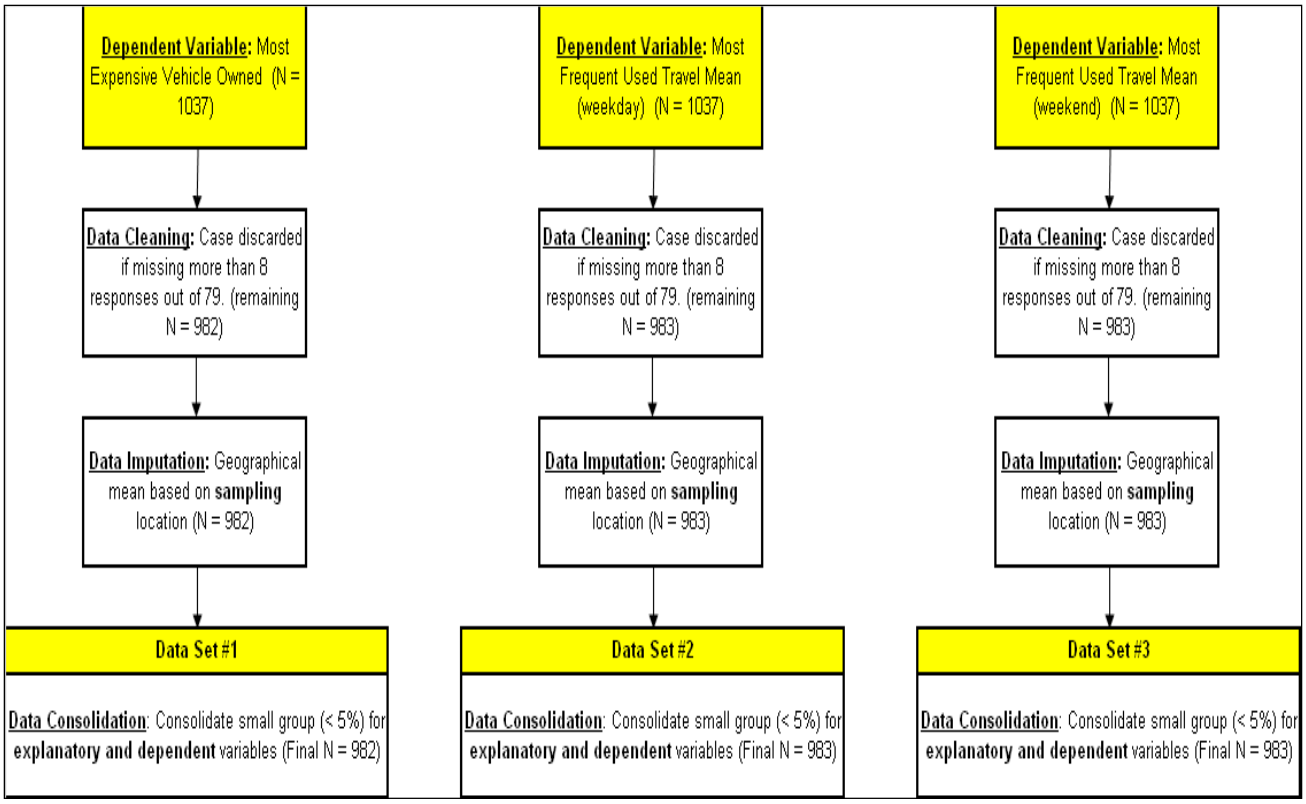


Figure 29: Flow Charts of Data Preparation

- **Data Imputation**

After discarding cases with too many missing data, the remaining missing data were

filled in by geographically-segmented means. Basically, the data imputation was used for the "real" missing (the blanks) and the N/A category. The imputation for N/A category was applied to only two questions of the survey, one is about the distance from home to the closest subway line, the other is about the distance from workplace to the closest subway line (as in the following):

How far is your working place to the "closest subway line"?

- ☐ Within 500 meters
- ☐ 501 ~ 1000 meters
- ☐ More than 1001 meters
- ☐ **N/A** (I don't work, I don't know, etc)

Originally, the N/A category was created to ensure that everyone has an appropriate response to select (e.g., try to differentiate the answer of "I don't work" from real missing data). However, since the answers of this question were treated as ordinal (1= Within 500 meters; 2= 501 ~ 1000 meters; 3= More than 1001 meters); the N/A was imputed (into ordinal measurement) based on geographically-segmented mean. For the "distance from workplace to the closest subway line", 67 N/A cells have been imputed; for the "distance from home to the closest subway line", 23 N/A cells have been imputed.

Instead of the housing location, the geographical means were actually calculated based on the sampling location for the following three reasons. First, I assume people within same sampling location are more homogeneous (in terms of variables related to vehicle purchase and use) than people within same housing location. This is actually the

fundamental assumption when I designed the location-based sampling (see Chapter 3). In Table 129, we can see the sampling locations as related to people's vehicle purchase and mobility characteristics, represented by location types. Second, the "sampling locations" already include the idea of "housing locations" – the household survey, as one example. The sampling location stratification can actually be considered as a hybrid stratification of "on-street" and "housing" locations. Third, the sampling location variable itself contains no missing data because that information was recorded by surveyors. However, there are missing data in the housing location variable, which was provided by respondents.

Location Type	25 Data Imputation Groups: Sampling Locations
On-street (general)	IKEA, Xu Jia Hui, Zhong Shan Park Subway, Jing An Temple, West Nan Jing Rd., Raffle's Plaza, Carrefour, Zheng Da Plaza, Peer Network
Bicycle-related	Tongji University
Motorized Two-wheeler/ Motorcycle-related	Ferry, E-bike Shop
Car-related	Ford Dong Chang, Ford Jiu Hua, Ford Fu Cheng, Driving School, Car Show Place
Household	An San 4, An San 5, An San 7, He Ping Hua Yuan, Ru San Xin Cun, Hai Fu Jia Yuan, Hai Yun Xin Cun, Shen Zu Jia Yuan

Table 129: 25 Data Imputation Groups

In addition to the original data from the survey, the imputation was also conducted on data derived from the original data, like the factor scores. When missing data would have precluded the direct estimation of a factor score, I imputed the original variables and re-ran the factor analysis, instead of directly imputing the factor scores. Similar results of the factor analysis were obtained (number of factors, factor names, etc.) using the

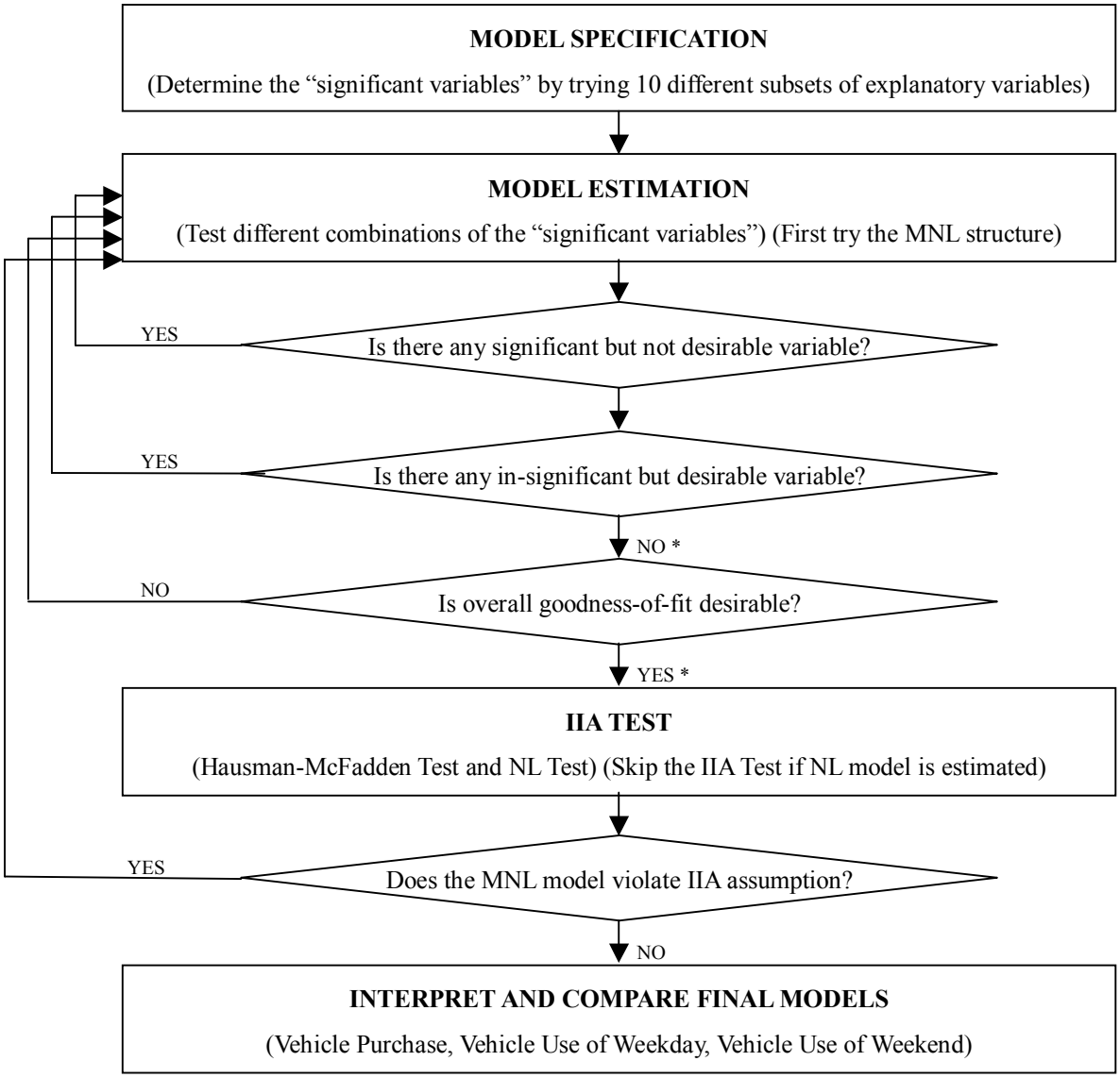
imputed variables. The new factor scores were hence used as explanatory variables for model estimation.

- **Data Consolidation**

The last step of the data preparation is the data consolidation. Theoretically, I consolidated (without deleting) both explanatory and dependent variables containing small group (< 5% of the sample). However, in practice, I did not consolidate certain variables with small group which I was interested in studying, for example, the dependent variable of "motorcycle" (4.92% of the sample).

Model Specification

After the data preparation, a 4-step procedure has been conducted to develop the vehicle choice models. The first two steps involved model specification and estimation. In this phase, multinomial logit (MNL) model was developed to estimate the probability of purchasing and using different vehicles based on the collective effect of personal, vehicle, and exogenous environment variables. Conceptual and statistical robustness were considered to determine the final model specification. The third step was Independence from Irrelevant Alternative (IIA) Test. The MNL model has to fulfill the IIA assumption in order to be valid. If the MNL violates the IIA assumption, a nested logit (NL) model structure will be tested and considered as a superior (final) model. The last step was to interpret the results in terms of the final model specification and structures. Vehicle purchase and use behaviors were also compared, since three models were created using the same dataset. Figure 30 presents the flow chart of the model development procedures.



*: Arbitrary stopping rules were applied to avoid infinite loop.

Figure 30: Model Development Procedures

- **MODEL 1: Most Expensive Vehicle Owned**

Among all 36 explanatory variables, except the perceived utility (six variables), the values of remaining 30 variables doesn't change by alternative. For those 30 variables, each variable must be assigned a different weight for at least one subset of the

alternatives. Otherwise, those variables will not be able to distinguish the choices if they were entered into the model directly with the same coefficients across all alternatives. Intuitively, "no purchase" was selected as the base alternative with zero coefficients for all its variables.

Initially, I allowed the coefficients of each variable to vary by non-base alternatives (bicycle, motorized two-wheeler, motorcycle, and car). That is, each explanatory variable was entered into the model as alternative-specific variables (ASVs). Although parsimony (or model simplicity) could be one concern, the choice of creating all ASVs instead of variables with constant coefficients across alternatives was for the sake of flexibility – for example, family income was assumed to have different influences between "car purchase" and "bicycle purchase". However, if the model results eventually indicated that a certain variable has similar coefficients for some of its ASVs, those coefficients were constrained to be equal.

If all explanatory variables are modeled as ASVs, the model will contain 36 x 4 = 144 ASVs, plus another 4 alternative-specific constants (ASCs). However, Limdep 7 (the modeling software) has a limitation of estimating models containing a maximum of 90 variables at the same time. Therefore, I tested ten different subsets of variables (Table 130) to identify the "significant variables" (Table 131), which are the variables with p-value < 0.05, then pooled them across subsets as the initial model specification (53 candidate ASVs, Table 132).

PART I only	Model contains only variables (ASVs) from survey PART I
PART II only	Model contains only variables (ASVs) from survey PART II
PART III only	Model contains only variables (ASVs) from survey PART III
PART IV only	Model contains only variables (ASVs) from survey PART IV
PART V only	Model contains only variables (ASVs) from survey PART V
PART V + I	Model contains variables (ASVs) from survey PART V and PART I
PART V + II	Model contains variables (ASVs) from survey PART V and PART II
PART V + III	Model contains variables (ASVs) from survey PART V and PART III
PART V + IV	Model contains variables (ASVs) from survey PART V and PART IV
PART I ~IV + Sig. V	Model contains variables (ASVs) from survey PART I ~ IV and significant variables (ASVs) identified in the "PART V only" scenario.

Table 130: 10 Subsets of Explanatory Variables to Determine the Significant Variables

Tested Scenario	PART I only	PART II only	PART III only	PART IV only	PART V only	PART V+I	PART V+II	PART V+III	PART V+IV	PART I~IV+Sig. V
"Significant" Variables with P < 0.05	N/A	USTA2	COST5	FAMILYG5	AGE2	WPARK2	USTA2	AGE2	AGE5	WPARK2
		USTA3	INFREQ4	FREEDOM2	AGE5	WPARK3	USTA3	AGE5	LICENM3	WPARK3
		UFAST2		FREEDOM3	LICENM3	HPARK2	UFAST2	LICENM3	LICENM4	WPARK4
		UAVAIL2			LICENM4	HPARK3	UAVAIL2	LICENM4	LICENC3	WPARK5
		UAVAIL3			LICENC5	WSUB2	UAVAIL3	LICENC5	LICENC5	WSUB2
		UAVAIL5			LICENMC3	WSUB3	UAVAIL5	LICENMC3	LICENMC3	WSUB3
		UCAP2			LICENMC4	HSUB4	AGE5	LICENMC4	LICENMC4	WSUB5
		UCOMF2			LICENMC5	AGE5	LICENM3	LICENMC5	LICENMC5	HSUB4
		UCOMF3			MALE2	LICENM3	LICENM4	MALE2	MALE2	HSUB5
					MALE3	LICENM4	LICENC5	MALE3	MALE3	USTA3
					MALE4	LICENC5	LICENMC3	MALE4	MALE4	UFAST2
					MALE5	LICENMC4	LICENMC4	MALE5	MALE5	UAVAIL2
					EMPSTU2	LICENMC5	LICENMC5	EMPSTU2	EMPSTU2	UAVAIL3
					EMPSTU5	MALE2	MALE2	EMPSTU5	EMPSTU5	UAVAIL5
					EMPFAC2	MALE3	MALE3	EMPFAC2	EMPFAC2	UCAP3
					EMPBOSS2	MALE4	MALE4	EMPBOSS2	EMPBOSS2	UCOMF2
					EMPBOSS3	MALE5	MALE5	EMPBOSS3	EMPBOSS3	INFREQ4
					EMPBOSS5	EMPSTU2	EMPSTU2	EMPBOSS5	EMPBOSS5	FAMILYG5
					INCFAM5	EMPSTU5	EMPSTU3	INCFAM5	INCFAM5	MALE2
					URBAN5	EMPFAC2	EMPSTU5	URBAN5	HHSIZE2	MALE3
					SHRESI5	EMPBOSS2	EMPFAC2	SHRESI3	URBAN5	MALE4
						EMPBOSS3	EMPBOSS2	SHRESI5	SHRESI3	MALE5
						EMPBOSS5	EMPBOSS3	COST5	SHRESI5	EMPSTU2
						INCFAM5	EMPBOSS5	INFREQ4	FAMILYG5	EMPFAC2
						URBAN5	INCFAM5	SOCIAL2	FREEDOM2	EMPBOSS2
						SHRESI5	URBAN5		FREEDOM3	EMPBOSS3
							SHRESI5		FREEDOM5	EMPBOSS5
										URBAN5
										SHRESI5

Table 131: Significant Variables of 10 Subsets of Variables [Dependent Variable = Most Expensive Vehicle Owned]

53 Candidate ASVs	Description
WPARK2	Parking Space at Work (for bicycle)
WPARK3	Parking Space at Work (for motorized two-wheeler)
WPARK4	Parking Space at Work (for motorcycle)
WPARK5	Parking Space at Work (for car)
HPARK2	Parking Space at Home (for bicycle)
HPARK3	Parking Space at Home (for motorized two-wheeler)
WSUB2	Distance from Work to Subway (for bicycle)
WSUB3	Distance from Work to Subway (for motorized two-wheeler)
WSUB5	Distance from Work to Subway (for car)
HSUB4	Distance from Home to Subway (for motorcycle)
HSUB5	Distance from Home to Subway (for car)
USTA2	Perceived Uitility of Status (for bicycle)
USTA3	Perceived Uitility of Status (for motorized two-wheeler)
UFAST2	Perceived Uitility of Speed (for bicycle)
UAVAIL2	Perceived Uitility of Availability (for bicycle)
UAVAIL3	Perceived Uitility of Availability (for motorized two-wheeler)
UAVAIL5	Perceived Uitility of Availability (for car)
UCAP2	Perceived Uitility of Carrying Capacity (for bicycle)
UCAP3	Perceived Uitility of Carrying Capacity (for motorized two-wheeler)
UCOMF2	Perceived Uitility of Comfort (for bicycle)
UCOMF3	Perceived Uitility of Comfort (for motorized two-wheeler)
COST5	[Factor] Cost-related Policy/Regulation (for car)
INFREQ4	[Factor] Attraction for Infrequent Use (for motorcycle)
SOCIAL2	[Factor] Pro-vehicle Purchase Cultural/Social Belief (for bicycle)
FAMILYG5	[Factor] Family-oriented and Green (for car)
FREEDOM2	[Factor] Pursue Freedom and Control of Life (for bicycle)
FREEDOM3	[Factor] Pursue Freedom and Control of Life (for motorized two-wheeler)
FREEDOM5	[Factor] Pursue Freedom and Control of Life (for car)
AGE2	Age (for bicycle)
AGE5	Age (for car)
LICENM3	Own Motorized two-wheeler/Motorcycle License (for motorized two-wheeler)
LICENM4	Own Motorized two-wheeler/Motorcycle License (for motorcycle)
LICENC3	Own Car License (for motorized two-wheeler)
LICENC5	Own Car License (for car)
LICENMC3	Own Car + Motorized two-wheeler/Motorcycle Licenses (for motorized two-wheeler)
LICENMC4	Own Car + Motorized two-wheeler/Motorcycle Licenses (for motorcycle)
LICENMC5	Own Car + Motorized two-wheeler/Motorcycle Licenses (for car)
MALE2	Male (for bicycle)
MALE3	Male (for motorized two-wheeler)
MALE4	Male (for motorcycle)
MALE5	Male (for car)
EMPSTU2	Student (for bicycle)
EMPSTU3	Student (for motorized two-wheeler)
EMPSTU5	Student (for car)
EMPFAC2	Factory Worker (for bicycle)
EMPBOSS2	Business Owner (for bicycle)
EMPBOSS3	Business Owner (for motorized two-wheeler)
EMPBOSS5	Business Owner (for car)
INCFAM5	Family Income (for car)
URBAN5	Live in Urban Area (for car)
SHRESI3	Registered Shanghai Resident (for motorized two-wheeler)
SHRESI5	Registered Shanghai Resident (for car)
HHSIZE2	Household Size (for bicycle)

Table 132: Initial Model Specification [Dependent Variable = Most Expensive Vehicle Owned]

- **MODEL 2: Most Frequently Used Travel Means (weekday)**

With similar concept to previous model, "walk" was chosen as the base alternative with all variable coefficients equal to zero. Table 133 shows those significant variables identified from 10 subsets of explanatory variables. Significant variables were color-coded based on five base subsets/scenarios – "PART I only" to "PART V only". Table 134 indicates the initial model specification of weekday vehicle use model. As we can observe, all variables were modeled as ASVs, and 35 candidate ASVs were identified as the initial model specification.

Tested Scenario	PART I only	PART II only	PART III only	PART IV only	PART V only	PART V+I	PART V+II	PART V+III	PART V+IV	PART I~IV+Sig. V
"Significant" Variables with P < 0.05	WPARK2	USTA2	DANGER2	FAMILYG3	AGE2	AGE2	AGE2	AGE2	AGE2	AGE2
	WPARK4	USTA4	COST2	FAMILYG5	AGE3	AGE3	AGE3	AGE3	LICENC3	LICENC3
	WPARK5	UAVAIL2		FAMILYG6	LICENC3	LICENC3	LICENC3	LICENC3	LICENC6	LICENC6
	WPARK6	UAVAIL4		FREEDOM3	LICENC6	LICENC6	LICENC6	LICENC6	LICENM5	LICENM5
	HPARK6	UAVAIL5			LICENM5	LICENM5	LICENM5	LICENM5	LICENMC3	LICENMC3
	WSUB6	UAVAIL6			LICENMC3	LICENMC3	LICENMC3	LICENMC3	LICENMC6	LICENMC6
	HSUB6				LICENMC6	LICENMC6	LICENMC6	LICENMC6	MALE4	MALE4
					MALE4	MALE4	MALE4	MALE4	MALE5	MALE5
					MALE5	MALE5	MALE5	MALE5	EMPFAC5	EMPBOSS6
					EMPNONE3	EMPFAC5	EMPNONE3	EMPNONE3	EMPBOSS4	INCPER6
					EMPFAC5	EMPBOSS4	EMPFAC5	EMPFAC5	EMPBOSS6	WPARK2
					EMPBOSS4	EMPBOSS6	EMPBOSS4	EMPBOSS4	INCPER6	WPARK4
					EMPBOSS6	INCFAM6	EMPBOSS6	EMPBOSS6	INCFAM6	WPARK5
					INCPER6	WPARK2	INCPER6	INCPER6	FAMILYG6	WPARK6
					INCFAM6	WPARK4	USTA2	INCFAM6		USTA2
						WPARK5	USTA4			USTA4
						WPARK6	UAVAIL2			UAVAIL2
							UAVAIL4			UAVAIL4
							UAVAIL5			UAVAIL5
							UAVAIL6			UAVAIL6
										DANGER2
										INFREQ6
										FAMILYG3
										FAMILYG6

Table 133: Significant Variables of 10 Subsets of Variables [Dependent Variable = Most Frequently Used Travel Means (weekday)]

35 Candidate ASVs	Description
WPARK2	Parking Space at Work (for public transportation)
WPARK4	Parking Space at Work (for bicycle)
WPARK5	Parking Space at Work (for motorized two-wheeler, motorcycle)
WPARK6	Parking Space at Work (for car)
HPARK6	Parking Space at Home (for car)
WSUB6	Distance from Work to Subway (for car)
HSUB6	Distance from Home to Subway (for car)
USTA2	Perceived Uitility of Status (for public transportation)
USTA4	Perceived Uitility of Status (for bicycle)
UAVAIL2	Perceived Uitility of Availability (for public transportation)
UAVAIL4	Perceived Uitility of Availability (for bicycle)
UAVAIL5	Perceived Uitility of Availability (for motorized two-wheeler, motorcycle)
UAVAIL6	Perceived Uitility of Availability (for car)
DANGER2	[Factor] Dangerous Transportation Environment (for public transportation)
COST2	[Factor] Cost-related Policy/Regulation (for public transportation)
INFREQ6	[Factor] Attraction for Infrequent Use (for car)
FAMILYG3	[Factor] Family-oriented and Green (for taxi, rented car, share company car)
FAMILYG5	[Factor] Family-oriented and Green (for motorized two-wheeler, motorcycle)
FAMILYG6	[Factor] Family-oriented and Green (for car)
FREEDOM3	[Factor] Pursue Freedom and Control of Life (for taxi, rented car, share company car)
AGE2	Age (for public transportation)
AGE3	Age (for taxi, rented car, share company car)
LICENC3	Own Car License (for taxi, rented car, share company car)
LICENC6	Own Car License (for car)
LICENM5	Own Motorized two-wheeler/Motorcycle License (for motorized two-wheeler, motorcycle)
LICENMC3	Own Car + Motorized two-wheeler/Motorcycle Licenses (for taxi, rented car, share company car)
LICENMC6	Own Car + Motorized two-wheeler/Motorcycle Licenses (for car)
MALE4	Male (for bicycle)
MALE5	Male (for motorized two-wheeler, motorcycle)
EMPNONE3	No-work or Retired (for taxi, rented car, share company car)
EMPFAC5	Factory Worker (for motorized two-wheeler, motorcycle)
EMPBOSS4	Business Owner (for bicycle)
EMPBOSS6	Business Owner (for car)
INCPER6	Personal Income (for car)
INCFAM6	Family Income (for car)

Table 134: Initial Model Specification [Dependent Variable = Most Frequently Used Travel Means (weekday)]

- **MODEL 3: Most Frequently Used Travel Means (weekend)**

In this model, "walk" was also set as the base alternative with all variable coefficients equal to zero. Table 135 shows those significant variables identified from 10 scenarios, and the same color-coding scheme was applied.

Tested Scenario	PART I only	PART II only	PART III only	PART IV only	PART V only	PART V+I	PART V+II	PART V+III	PART V+IV	PART I~IV+Sig. V
"Significant" Variables with P < 0.05	HSUB5	USTA2	COST6	FAMILYG6	AGE5	AGE5	LICENC5	AGE5	AGE5	LICENC5
	HSUB6	USTA4		STASEEK2	LICENC5	LICENC6	LICENC6	LICENC5	LICENC5	LICENC6
		UFAST4		STASEEK3	LICENC6	LICENM5	LICENM5	LICENC6	LICENC6	LICENM5
		UAVAIL5		STASEEK6	LICENM5	LICENMC6	LICENMC6	LICENM5	LICENM5	LICENMC6
		UAVAIL6		ENJOY6	LICENMC6	MALE4	MALE4	LICENMC6	LICENMC5	MALE4
		UCAP3			MALE4	EMPOFF4	EMPNONE3	MALE4	LICENMC6	INCPER6
		UCOMF6			EMPOFF4	INCPER3	EMPOFF4	EMPOFF4	MALE4	INCFAM3
					INCPER6	INCPER6	INCPER6	INCPER6	EMPNONE3	INCFAM6
					INCFAM3	INCFAM6	INCFAM3	INCFAM3	EMPOFF4	USTA2
					INCFAM6		INCFAM6	INCFAM6	INCPER6	USTA4
							USTA2	COST6	INCFAM3	UFAST4
							USTA4		INCFAM6	UFAST5
							UFAST4			UAVAIL6
							UAVAIL6			UCAP3
										UCAP5
										COST6
										SOCIAL5
										STASEEK2
										STASEEK3

Table 135: Significant Variables of 10 Subsets of Variables [Dependent Variable = Most Frequently Used Travel Means (weekend)]

Table 136 indicates the initial model specification of weekend vehicle use model. Thirty-one candidate ASVs (parentheses were used to specify the alternatives which ASVs were associated with) were identified as the union of those significant variables in Table 135.

31 Candidate ASVs	Description
HSUB5	Distance from Home to Subway (for motorized two-wheeler, motorcycle)
HSUB6	Distance from Home to Subway (for car)
USTA2	Perceived Utility of Status (for public transportation)
USTA4	Perceived Utility of Status (for bicycle)
UFAST4	Perceived Utility of Speed (for bicycle)
UFAST5	Perceived Utility of Speed (for motorized two-wheeler, motorcycle)
UAVAIL5	Perceived Utility of Availability (for motorized two-wheeler, motorcycle)
UAVAIL6	Perceived Utility of Availability (for car)
UCAP3	Perceived Utility of Carrying Capacity (for taxi, rented car, share company car)
UCAP5	Perceived Utility of Carrying Capacity (for motorized two-wheeler, motorcycle)
UCOMF6	Perceived Utility of Comfort (for car)
COST6	[Factor] Cost-related Policy/Regulation (for car)
SOCIAL5	[Factor] Pro-vehicle Purchase Cultural/Social Belief (for motorized two-wheeler, motorcycle)
FAMILYG6	[Factor] Family-oriented and Green (for car)
STASEEK2	[Factor] Status-seeking (for public transportation)
STASEEK3	[Factor] Status-seeking (for taxi, rented car, share company car)
STASEEK6	[Factor] Status-seeking (for car)
ENJOY6	[Factor] Enjoy Now (for car)
AGE5	Age (for motorized two-wheeler, motorcycle)
LICENC5	Own Car License (for motorized two-wheeler, motorcycle)
LICENC6	Own Car License (for car)
LICENM5	Own Motorized two-wheeler/Motorcycle License (for motorized two-wheeler, motorcycle)
LICENMC5	Own Car + Motorized two-wheeler/Motorcycle Licenses (for motorized two-wheeler, motorcycle)
LICENMC6	Own Car + Motorized two-wheeler/Motorcycle Licenses (for car)
MALE4	Male (for bicycle)
EMPNONE3	No-work or Retired (for taxi, rented car, share company car)
EMPOFF4	Office Worker (for bicycle)
INCPER3	Personal Income (for taxi, rented car, share company car)
INCPER6	Personal Income (for car)
INCFAM3	Family Income (for taxi, rented car, share company car)
INCFAM6	Family Income (for car)

Table 136: Initial Model Specification [Dependent Variable = Most Frequently Used Travel Means (weekend)]

Model Estimation

- **MODEL 1: Most Expensive Vehicle Owned**

In this phase of model estimation, variables in the initial specification were modeled under the MNL structure. However, MNL was only considered as the intermediate model, and the IIA test was conducted to tell if the NL is a superior model structure than MNL. Based on the IIA test, the NL might be selected as the best (final) model structure. By adding and subtracting variables, the final model was expected to have a slightly different

specification (from the initial one) but to achieve better statistical and conceptual robustness. Table 137 shows the estimated MNL model with four ASCs and 38 ASVs, representing 23 different types of explanatory variables. All the variables were conceptually interpretable and statistically significant with p-value less than 0.05 – except for the MALE5, which is considered as a "desirable" variable and is just slightly beyond the cut-off (at p-value = 0.504). As mentioned, the models (for three dependent variables) in this phase are considered intermediate. More sophisticated tables will be provided for the model results after the IIA tests (at the end of this chapter).

Table 138 presents the goodness-of-fit statistics. The ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.409, indicating that the estimated (full) model explains 40.9% (equally-likely model as the base) of the information in the survey data about Shanghai people's most expensive vehicle purchase. The ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.370, which means the estimated model can still explain 37% of the information in the data under a more strict (market-share model) base.

In addition, the χ_c^2 value of 1096.1 means there is a significant difference between the estimated model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1294.2 shows that the estimated model also significantly differs from the equally-likely (EL) model (all coefficients are equal to zero) at $\alpha << 0.005$.

#	Variable	Coefficient	P-value
1	ASC2	-2.519	0.0006
2	ASC3	-3.214	0.0000
3	ASC4	-3.234	0.0001
4	ASC5	-8.575	0.0000
5	WPARK2	-1.169	0.0003
6	WPARK3	-1.428	0.0000
7	WPARK4	-1.663	0.0002
8	HPARK2	-1.008	0.0010
9	WSUB2	0.362	0.0031
10	WSUB3	0.401	0.0011
11	HSUB4	0.535	0.0304
12	USTA2	0.275	0.0047
13	USTA3	0.448	0.0000
14	UFAST2	0.359	0.0026
15	UAVAIL2	0.587	0.0000
16	UAVAIL3	0.468	0.0000
17	UAVAIL5	0.380	0.0041
18	AGE5	0.048	0.0000
19	LICENM3	2.669	0.0000
20	LICENM4	4.828	0.0000
21	LICENC5	3.800	0.0000
22	LICENMC3	1.482	0.0044
23	LICENMC4	3.609	0.0000
24	LICENMC5	4.193	0.0000
25	MALE3	0.558	0.0044
26	MALE4	1.164	0.0127
27	MALE5	0.507	**0.0504**
28	EMPSTU2	1.708	0.0000
29	EMPSTU5	1.540	0.0049
30	EMPFAC2	0.814	0.0009
31	EMPBOSS2	1.517	0.0100
32	EMPBOSS3	1.330	0.0155
33	EMPBOSS5	2.628	0.0000
34	INCFAM5	0.440	0.0000
35	HHSIZE2	-0.215	0.0121
36	URBAN5	-1.086	0.0024
37	SHRESI5	0.940	0.0024
38	COST5	-0.424	0.0048
39	INFREQ4	0.530	0.0189
40	FAMILYG5	-0.412	0.0050
41	FREEDOM2	-0.336	0.0195
42	FREEDOM3	-0.319	0.0217

Table 137: Estimation of MNL Model [Dependent Variable = Most Expensive Vehicle Owned]

Number of observations **(purchase nothing = 211, bicycle = 257, motorized two-wheeler = 195, motorcycle = 50, car = 269)**	982
Final log-likelihood, L (β)	-933.378
Log-likelihood for market-share model, L (MS)	-1481.411
Log-likelihood for equally-likely (EL) model, L (0)	-1580.468
Total Number of Estimated Parameters = ASVs+ ASCs = 38 + 4	42
ρ_o^2 = 1– [L (β) / L (0)]	0.409
Adjusted ρ_o^2 = 1– {[L (β) – Total # of Estimated Parameters] / L (0)}	0.383
ρ_c^2 = 1– [L (β) / L (MS)]	0.370
Adjusted ρ_c^2 = 1– {[L (β) – Total # of Estimated ASVs] / L (MS)}	0.344
χ_o^2 (between the final model and the EL model)	1294.2
χ_c^2 (between the final model and the MS model)	1096.1

Table 138: Goodness-of-fit Statistics for Estimated MNL Model [Dependent Variable = Most Expensive Vehicle Owned]

- **MODEL 2: Most Frequently Used Travel Means (weekday)**

Table 139 presents the estimated MNL model with five ASCs and 21 ASVs, representing 12 different types of explanatory variables. All the variables were conceptually interpretable and statistically significant with p-values less than 0.05.

Table 140 indicates the goodness-of-fit statistics; the ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.299, indicating that the estimated (full) model explains 29.9% (equally-likely model as the base) of the information in the survey data. The ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.231, which means the estimated model can still explain 23.1% of the information in the data under a more strict (market-share model) base.

The χ_c^2 value of 773.8 means there is a significant difference between the estimated model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1100.1 shows that the estimated model also significantly differs from the equally-likely (EL) model at $\alpha << 0.005$.

#	Variable	Coefficient	P-value
1	ASC2	1.163	0.0055
2	ASC3	-0.117	0.4974
3	ASC4	-1.896	0.0023
4	ASC5	-1.638	0.0090
5	ASC6	-4.492	0.0000
6	WPARK2	-0.803	0.0016
7	WPARK4	-1.095	0.0001
8	WPARK5	-0.867	0.0060
9	USTA2	0.221	0.0072
10	USTA4	0.355	0.0002
11	UAVAIL2	0.267	0.0012
12	UAVAIL4	0.525	0.0000
13	UAVAIL5	0.501	0.0002
14	UAVAIL6	0.404	0.0011
15	FAMILYG6	-0.258	0.0471
16	FREEDOM3	0.456	0.0017
17	AGE2	-0.018	0.0014
18	LICENC3	1.461	0.0000
19	LICENC6	3.734	0.0000
20	LICENM5	2.257	0.0000
21	LICENMC3	1.813	0.0000
22	LICENMC6	3.416	0.0000
23	MALE4	0.659	0.0009
24	MALE5	1.299	0.0000
25	EMPBOSS6	1.303	0.0000
26	INCPER6	0.294	0.0000

Table 139: Estimation of MNL Model [Dependent Variable = Most Frequently Used Travel Means (weekday)]

Number of observations **(walk = 80; public transportation = 291; taxi, rented car, shared company car = 118; bicycle = 142; motorized two-wheeler, motorcycle = 137; car = 214)**	982
Final log-likelihood, L (β)	-1288.2
Log-likelihood for market-share model, L (MS)	-1675.1
Log-likelihood for equally-likely (EL) model, L (0)	-1838.2
Total Number of Estimated Parameters = ASVs+ ASCs = 21 + 5	26
$\rho_o^2 = 1 - [L(\beta) / L(0)]$	0.299
Adjusted $\rho_o^2 = 1 - \{[L(\beta) - $ Total # of Estimated Parameters$] / L(0)\}$	0.285
$\rho_c^2 = 1 - [L(\beta) / L(MS)]$	0.231
Adjusted $\rho_c^2 = 1 - \{[L(\beta) - $ Total # of Estimated ASVs$] / L(MS)\}$	0.218
χ_o^2 (between the final model and the EL model)	1100.1
χ_c^2 (between the final model and the MS model)	773.8

Table 140: Goodness-of-fit Statistics for Estimated MNL Model [Dependent Variable = Most Frequently Used Travel Means (weekday)]

- **MODEL 3: Most Frequently Used Travel Means (weekend)**

Similarly, Table 141 presents the estimated MNL model with five ASCs and 17 ASVs, representing 13 different types of explanatory variables. All the variables were conceptually interpretable and statistically significant with p-values less than 0.05.

Table 142 indicates the goodness-of-fit statistics; the ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.299, indicating that the estimated (full) model explains 29.9% (equally-likely model as the base) of the information in the survey data. The ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.209, which means the estimated model can still explain 20.9% of the information in the data under a more strict (market-share model) base.

The χ_c^2 value of 652.0 means there is a significant difference between the estimated model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1050.7 means that the estimated model significantly differs from the equally-likely (EL) model at $\alpha << 0.005$.

#	Variable	Coefficient	P-value
1	ASC2	1.194	0.0000
2	ASC3	-2.165	0.0004
3	ASC4	-3.106	0.0000
4	ASC5	-1.321	0.0034
5	ASC6	-5.210	0.0000
6	HSUB5	0.442	0.0083
7	USTA2	0.165	0.0238
8	USTA4	0.345	0.0169
9	UFAST4	0.480	0.0068
10	UAVAIL6	0.301	0.0086
11	UCAP3	0.290	0.0190
12	COST6	-0.479	0.0002
13	LICENC6	2.693	0.0000
14	LICENM5	2.828	0.0000
15	LICENMC5	1.623	0.0000
16	LICENMC6	2.832	0.0000
17	MALE4	0.848	0.0029
18	EMPNONE3	0.885	0.0097
19	INCPER3	0.271	0.0000
20	INCPER6	0.384	0.0000
21	INCFAM3	0.157	0.0204
22	INCFAM6	0.263	0.0075

Table 141: Estimation of MNL Model [Dependent Variable = Most Frequently Used Travel Means (weekend)]

Number of observations **(walk = 67; public transportation = 358; taxi, rented car, shared company car = 183; bicycle = 59; motorized two-wheeler, motorcycle = 89; car = 226)**	982
Final log-likelihood, L (β)	-1234.2
Log-likelihood for market-share model, L (MS)	-1560.2
Log-likelihood for equally-likely (EL) model, L (0)	-1759.5
Total Number of Estimated Parameters = ASVs+ ASCs = 17 + 5	22
ρ_o^2 = 1– [L (β) / L (0)]	0.299
Adjusted ρ_o^2 = 1– {[L (β) – Total # of Estimated Parameters] / L (0)}	0.286
ρ_c^2 = 1– [L (β) / L (MS)]	0.209
Adjusted ρ_c^2 = 1– {[L (β) – Total # of Estimated ASVs] / L (MS)}	0.198
χ_o^2 (between the final model and the EL model)	1050.7
χ_c^2 (between the final model and the MS model)	652.0

Table 142: Goodness-of-fit Statistics for Estimated MNL Model [Dependent Variable = Most Frequently Used Travel Means (weekend)]

Independence from Irrelevant Alternatives Test (IIA)

- **MODEL 1: Most Expensive Vehicle Owned**

Hausman-McFadden test

The Independence from Irrelevant Alternatives (IIA) assumption states that "the relative odds of choosing one alternative over another should not differ with the presence or absence of other alternatives in the choice set."

The IIA assumption is a central condition for the MNL model structure to be valid. IIA will be violated when observed explanatory variables are correlated with unobserved ones, or when the unobserved variables for one alternative are correlated with those of another alternative. In my three models, certain alternatives could be considered similar

(e.g., motorized two-wheeler and motorcycle); therefore, it is very likely that IIA will be violated. However, if the IIA assumption is violated, MNL is not the appropriate model structure, and another model structure (such as NL) or specification must be applied.

In this research, I tested for IIA violations using two approaches. First, the Hausman-McFadden test (Hausman and McFadden, 1984) was conducted by comparing the coefficients of the model estimated on the full choice set with those of a model estimated on a subset of alternatives. Second, MNL model results were compared with several more general nested logit (NL) model formulations, which have the MNL model as a special case.

In terms of the Hausman-McFadden test, if the IIA property holds, the parameters of a model involving the full choice set should be the same as those involving only a subset of the full choice set. This test can be presented as in the following:

H_o (Null hypothesis): $\beta^R = \beta^U$, where β^R is the vector of (true) parameters for the model involving the restricted choice set, and β^U is the vector of (true) parameters for the model involving the unrestricted or full choice set.

H_a (Alternative hypothesis): $\beta^R \neq \beta^U$.

Test-statistic: $[\hat{\beta}^R - \hat{\beta}^U]'[V^R - V^U]^{-1}[\hat{\beta}^R - \hat{\beta}^U]$, where $\hat{\beta}^U$ and V^U are, respectively, the vector of coefficient estimates and the estimated variance-covariance matrix of $\hat{\beta}$ for the "unrestricted model"; and $\hat{\beta}^R$ and V^R are

the vector of estimated coefficients and variance-covariance matrix of $\hat{\beta}$ for the "restricted model". This statistic is asymptotically chi-squared distributed with the degrees of freedom equal to the number of identifiable parameters in $\hat{\beta}^R$.

The Hausman-McFadden test has been conducted on eight reduced choice sets of the MNL model of vehicle purchase, namely dropping alternatives: 2, 3, 4, 5, (2 & 3), (2 & 5), (3 & 4), and (4 & 5) respectively. In every case, the test statistic could not be computed. However, the inability to carry out the Hausman-McFadden test is quite common, since this test *"requires inversion of the difference between two closely related matrices, which may be non-positive-definite or nearly singular and thus cause computational and inference problems"* (Small and Hsaio, 1985). If IIA holds, by definition $\beta^R = \beta^U$, and therefore the variance-covariance matrices of the two vectors of parameter estimators, V^R and V^U, are also likely to be similar. If that is true, then their difference will be a matrix of relatively small numbers, and inverting such a matrix to compute the "test statistic" will be similar to division by zero. Nevertheless, the computational failures are only suggestive (but not conclusive) that IIA holds. The NL test needs to be conducted to further confirm.

NL Test

As mentioned, the nested logit (NL) model is a generalized format of MNL in which alternatives that are suspected of sharing unobserved characteristics are grouped together into nests. Within the nests, IIA is assumed to hold, but it is not necessary for IIA to hold between alternatives in different nests. When the NL model is estimated, the

decision-making involves both the choices of the nests and the alternatives within the nests.

The NL test of IIA is basically to see if the NL and MNL models are significantly different. First, to make the NL model theoretically consistent, the "inclusive value (IV) parameter" of each nest needs to be between 0 and 1. If any of the IV parameters are significantly less than one, the NL model is significantly different (and better) than the MNL model, and the NL should be used to fix the IIA violation of MNL. On the other hand, if the IV parameter of each nest is equal to 1, the NL model is equivalent to the MNL model as a special case of NL. Statistically, the NL test can be presented as:

H_o (Null hypothesis): IV parameter $\theta = 1$.

H_a (Alternative hypothesis): IV parameter $\theta \neq 1$.

Test-statistic: $\frac{\hat{\theta} - 1}{s.e.(\hat{\theta})}$, where $\hat{\theta}$ is the estimated IV parameter and $s.e.(\hat{\theta})$ is the (estimated) standard error of the estimated IV parameter. This statistic asymptotically follows the t-distribution with degrees of freedom equal to the number of observations minus the number of estimated parameters in the model.

Table 143 shows ten NL model structures for the test. The assumption of those NL structures was based on simple conceptual grounds – for example, the combination of motorcycle and motorized two-wheeler (purchase), and the separation of car (purchase) from other alternatives.

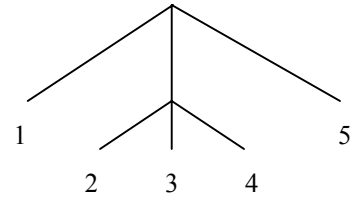
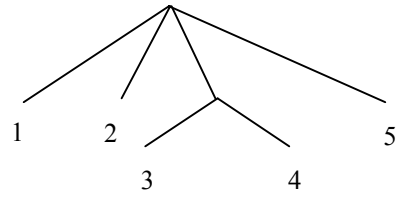
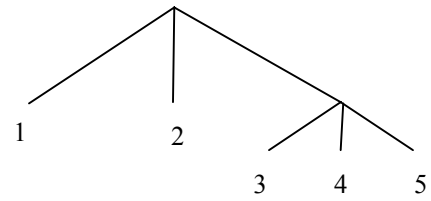
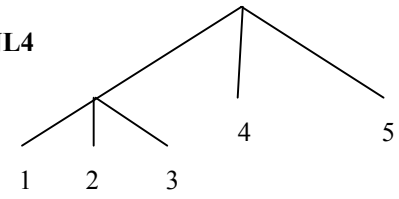
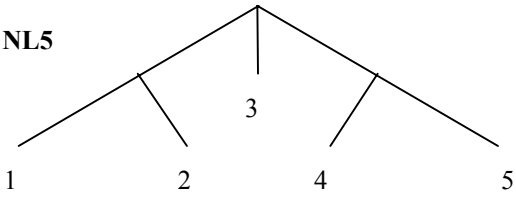
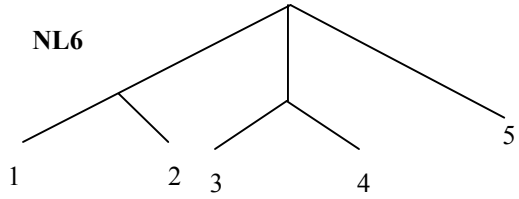
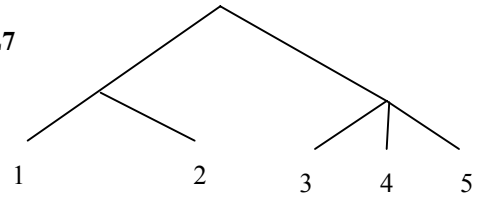
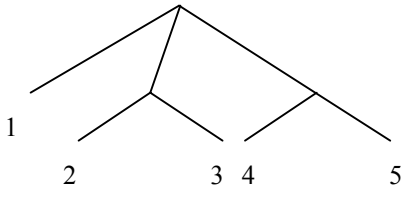
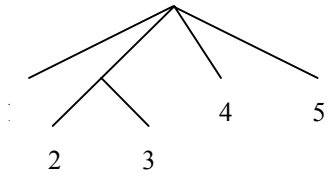
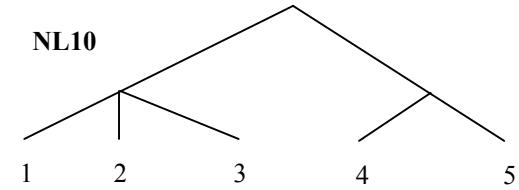
1= no purchase, 2= bicycle, 3= motorized two-wheeler, 4= motorcycle, 5= car	
NL1 	**NL2** 
NL3 	**NL4** 
NL5 	**NL6** 
NL7 	**NL8** 
NL9 	**NL10** 

Table 143: 10 Nested Logit Model Structures Tested [Dependent Variable = Most Expensive Vehicle Owned]

Table 144 summarizes the NL test results using the "initial model specification" (NL1 to NL10). Three models failed to reject the null hypothesis that the IV parameter was equal to one, indicating that NL is equivalent to the MNL model. There were seven NL models able to reject the null hypothesis; however, two of them were with the estimated IV parameter greater than one (highlighted in red), which is theoretically impermissible.

Only five NL models (highlighted in blue) had reasonable estimated IV parameter (less than one), and thus I am able to reject the null hypothesis, suggesting that they were superior models than the original MNL. By comparing the ρ_o^2, NL4 was selected as the best model structure. However, different specifications were tested by adding (insignificant but desirable) and subtracting (significant but not desirable) variables. NL4-1 (shown in the last column of the table) was the final vehicle purchase model, with $\rho_o^2 = 0.466$.

	NL1	NL2	NL3	NL4	NL5	NL6	NL7	NL8	NL9	NL10	NL4-1 (Final)
1= no purchase 2= bicycle 3= motorized two-wheeler 4= motorcycle 5= car	tree = a(1), b(2,3,4), c(5)	tree =a(1), b(2), c(3,4), d(5)	tree = a(1), b(2), c(3, 4, 5)	tree = a(1,2,3) b(4), c(5)	tree = a(1,2), b(3), c(4,5)	tree = a(1,2), b(3,4), c(5)	tree = a(1,2), b(3,4,5)	tree = a(1), b(2,3), c(4,5)	tree = a(1), b(2,3), c(4), d(5)	tree = a(1,2,3), b(4,5)	tree = a(1,2,3) b(4), c(5)
ρ^2	0.429	0.391	0.438	0.489	0.426	0.407	0.409	0.423	0.444	0.431	0.466
IV parameter estimate	0.735	1.185	1.705	0.243	0.832	0.878	1.535	0.755	0.638	0.597	0.203
Standard Error of IV (tau)	0.116	0.347	0.207	0.116	0.132	0.183	0.180	0.104	0.122	0.096	0.095
Test statistic	-2.284	0.535	3.404	-6.546	-1.273	-0.665	2.979	-2.360	-2.978	-4.217	-8.353
Number of observations			4910								4910
Estimated parameters			42								31
Degrees of freedom			4868								4879
95% critical value for t-distribution (two-tailed)			1.96								1.96
Conclusion	Reject Ho	Fail to reject Ho	Reject Ho	Reject Ho	Fail to reject Ho	Fail to reject Ho	Reject Ho	Reject Ho	Reject Ho	Reject Ho	Reject Ho

Table 144: Summary of Nest Logit Model Test [Dependent Variable = Most Expensive Vehicle Owned]

- **MODEL 2: Most Frequently Used Travel Means (weekday)**

Hausman-McFadden test

The Hausman-McFadden test has been conducted on eight reduced choice sets, namely dropping alternatives: 2, 3, 4, 5, 6, (2 & 3), (3 & 4), and (4 & 6) respectively. In every case, the test statistic could not be computed, which suggests (but is not conclusive) that

the IIA holds for the MNL model of weekday vehicle use.

NL Test

For the weekday and weekend vehicle use, I tried the same ten NL models (Table 145).

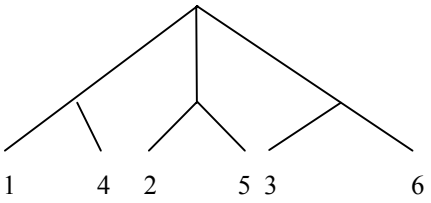
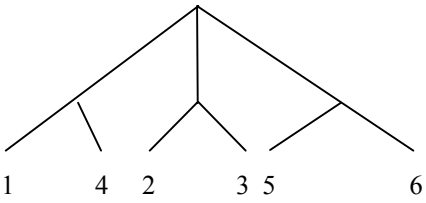
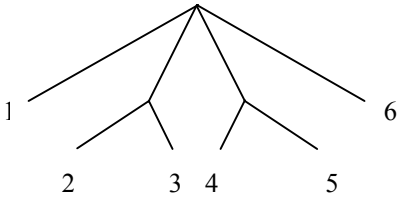
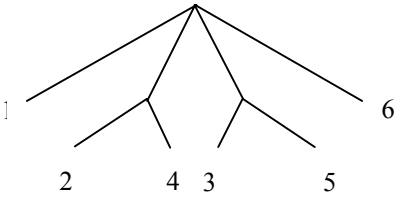
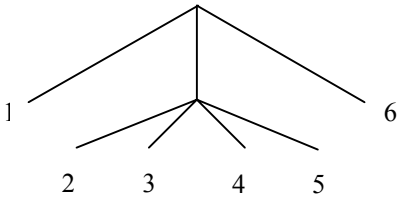
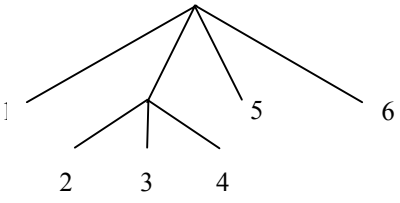
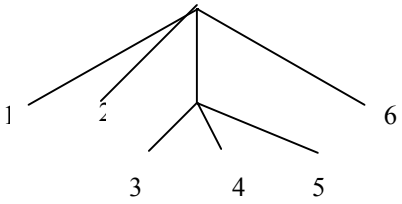
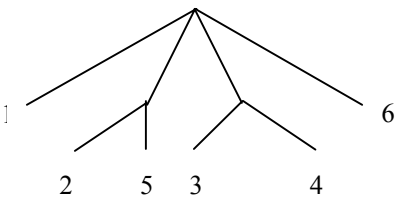
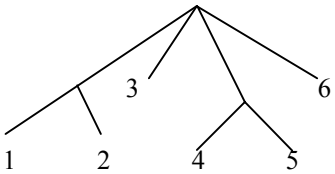
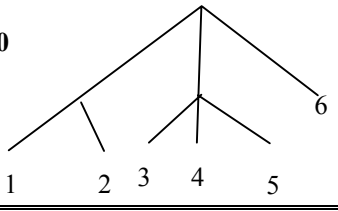
1= walk, 2= public transportation, 3= taxi, rented car, shared company car, 4= bicycle, 5= motorized two-wheeler, motorcycle, 6= car	
NL1 	**NL2** 
NL3 	**NL4** 
NL5 	**NL6** 
NL7 	**NL8** 
NL9 	**NL10** 

Table 145: 10 Nested Logit Model Structures Tested [Dependent Variable = Most Frequently Used Travel Means (weekday) (weekend)]

The assumption of those NL structures was based on the motorization pathway idea mentioned in Chapter 4 – for example, private type vehicles were combined into one nest, while the shared type vehicles were combined into the other.

Table 146 summarizes the ten tested NL models, two models failed to reject the null hypothesis (IV parameter = 1), indicating that NL is equivalent to the MNL model. Eight NL models rejected the null hypothesis, but three of them were with the estimated IV parameter greater than one (highlighted in red), which is theoretically impermissible. The remaining five NL models (highlighted in blue) were with estimated IV parameter less than one and able to reject the null hypothesis, suggesting that they were superior models than the original MNL. Similar to the MODEL 1, different model specifications were tested by adding (insignificant but desirable) and subtracting (significant but not desirable) variables. Eventually, the NL3 (circled) was selected as the final model of weekday vehicle use behavior. The ρ_o^2 of NL3 ($\rho_o^2 = 0.299$) is essentially equivalent to the highest one, of NL8 ($\rho_o^2 = 0.300$). I considered NL3 the better model because those two ρ_o^2 values (NL3, NL8) are close, but NL3 had a more interpretable model specification and NL structure.

	NL1	NL2	NL3 (Final)	NL4	NL5	NL6	NL7	NL8	NL9	NL10
1= walk 2= public transportation 3= taxi, rented car, shared company car 4= bicycle 5= motorized two-wheeler, motorcycle 6= car	tree = a(1,4), b(2,5), c(3,6)	tree = a(1,4), b(2,3), c(5,6)	tree = a(1), b(2,3), c(4, 5), d(6)	tree = a(1), b(2,4), c(3,5), d(6)	tree = a(1), b(2,3,4,5), c(6)	tree = a(1), b(2,3,4), c(5), d(6)	tree = a(1), b(2), c(3,4,5), d(6)	tree = a(1), b(2,5), c(3,4), d(6)	tree = a(1,2), b(3), c(4,5), d(6)	tree = a(1,2), b(3,4,5), c(6)
ρ^2	0.270	0.270	0.299	0.301	0.365	0.344	0.284	0.300	0.291	0.273
IV parameter estimate	1.772	1.531	0.675	1.710	0.996	1.299	0.667	0.575	0.507	0.709
Standard Error of IV (tau)	0.284	0.180	0.137	0.281	0.253	0.304	0.117	0.118	0.107	0.113
Test statistic	2.722	2.954	-2.363	2.524	-0.015	0.981	-2.839	-3.587	-4.612	-2.567
Number of observations			5892							
Estimated parameters			26							
Degrees of freedom			5866							
95% critical value for t-distribution (two-tailed)			1.96							
Conclusion	Reject Ho	Reject Ho	**Reject Ho**	Reject Ho	Fail to reject Ho	Fail to reject Ho	**Reject Ho**	**Reject Ho**	**Reject Ho**	**Reject Ho**

Table 146: Summary of Nest Logit Model Test [Dependent Variable = Most Frequently Used Travel Means (weekday)]

- **MODEL 3: Most Frequently Used Travel Means (weekend)**

Hausman-McFadden test

Similar to Model 2, the Hausman-McFadden test has been conducted on eight reduced choice sets dropping alternatives: 2, 3, 4, 5, 6, (2 & 3), (3 & 4), and (4 & 6) respectively. In every case, the test statistic could not be computed, suggesting that the IIA holds for the MNL model of weekend vehicle use.

NL Test

Table 147 presented the results of ten NL models (with initial model specification). Unlike previous two tests, most (eight out of ten) NL models failed to reject the null hypothesis (IV parameter = 1), which means NL was not a superior structure to MNL. Two NL models rejected the null hypothesis. However, they have the estimated IV parameter greater than one (highlighted in red), which is theoretically impermissible.

Thus, the IIA test results for the NL models strongly suggest that no NL models are superior to previous MNL model. That is, the IIA property of the MNL model holds, and the MNL model previously developed will be considered the final model of weekend vehicle use. More detail discussion will follow in the model interpretation section of this chapter.

	NL1	NL2	NL3	NL4	NL5	NL6	NL7	NL8	NL9	NL10
1= walk 2= public transportation 3= taxi, rented car, shared company car 4= bicycle 5= motorized two-wheeler, motorcycle 6= car	tree = a(1,4), b(2,5), c(3,6)	tree = a(1,4), b(2,3), c(5,6)	tree = a(1), b(2,3), c(4, 5), d(6)	tree = a(1), b(2,4), c(3,5), d(6)	tree = a(1), b(2,3,4,5), c(6)	tree = a(1), b(2,3,4), c(5), d(6)	tree = a(1), b(2), c(3,4,5), d(6)	tree = a(1), b(2,5), c(3,4), d(6)	tree = a(1,2), b(3), c(4,5), d(6)	tree = a(1,2), b(3,4,5), c(6)
ρ^2	0.299	0.304	0.330	0.329	0.394	0.391	0.285	0.330	0.298	0.290
IV parameter estimate	1.046	1.817	1.475	1.135	1.720	2.079	0.971	0.682	0.893	1.032
Standard Error of IV (tau)	0.172	0.253	0.344	0.239	0.537	0.548	0.187	0.172	0.231	0.190
Test statistic	0.267	3.235	1.382	0.564	1.342	1.970	-0.155	-1.850	-0.464	0.171
Number of observations			5892							
Estimated parameters			22							
Degrees of freedom			5870							
95% critical value for t-distribution (two-tailed)			1.96							
Conclusion	Fail to reject Ho	Reject Ho	Fail to reject Ho	Fail to reject Ho	Fail to reject Ho	Reject Ho	Fail to reject Ho	Fail to reject Ho	Fail to reject Ho	Fail to reject Ho

Table 147: Summary of Nested Logit Model Test [Dependent Variable = Most Frequently Used Travel Means (weekend)]

Interpretation of Model Results

- **MODEL 1: Most Expensive Vehicle Owned**

As shown in Table 148, the NL model of most expensive vehicle owned was created with four ASCs and 27 ASVs, representing 16 different explanatory variables. All explanatory variables were statistically significant.

In term of goodness-of-fit statistics, the ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.47, indicating that the estimated (full) model explains 47% (equally-likely model as the base) of the information in the survey data. The ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.35, which means the estimated model can still explain 35% of the information in the data under a more strict (market-share model) base.

The χ_c^2 value of 1031.7 means there is a significant difference between the estimated model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1683.3 means that the estimated model significantly differs from the equally-likely (EL) model at $\alpha << 0.005$.

To explain the model, in the following, I will first describe the model results by explanatory variables (by row), which are categorized based on different parts of the original survey. Thus, we can see how the variables' influences vary across different alternatives. Then, the model results will be interpreted by alternative (by column) so as to develop the profile of buyers/users of different type of vehicles.

Explanatory Variables	*Alternatives (Base Alternative = Purchase Nothing)*			
	Bicycle	**Motorized Two-wheeler**	**Motorcycle**	**Car**
ASC	-1.15 (0.2937)	-1.83 (0.0794)	-12.98 (0.097)	-28.25 (0.0482)
[PART I] Vehicle Purchase and Use Background				
Parking Space at Work	-3.1 (0.0003)	-3.47 (0.0001)	-5.9 (0.0283)	
Parking Space at Home	-1.08 (0.0018)			
Distance from Work to Subway	0.48 (0.0005)	0.56 (0.0001)		
[PART II] Perceived Utility				
Perceived Uitility of Status	0.3 (0.0051)	0.48 (0.0001)		
Perceived Uitility of Speed	0.41 (0.0011)			
Perceived Uitility of Availability	0.63 (0)	0.51 (0)		
[PART IV] Lifestyle Factor				
Pursue Freedom and Control of Life	-0.48 (0.0035)	-0.48 (0.0039)		
[PART V] Demographics				
Own Motorized two-wheeler/Motorcycle License		2.79 (0)	19.85 (0.0236)	
Own Car License				16.45 (0.0353)
Own Car + Motorized two-wheeler/Motorcycle Licenses		1.8 (0.0016)	17.71 (0.0336)	18.41 (0.0291)
Male		0.56 (0.005)		
Student	1.7 (0)			
Factory Worker	0.69 (0.0065)			
Business Owner	1.7 (0.0184)	1.59 (0.0241)		8.2 (0.0243)
Family Income				2.26 (0.05)
Household Size	-0.21 (0.021)			
Goodness-of-fit Statistic				
Number of observations (purchase nothing = 211, bicycle = 257, motorized two-wheeler = 195, motorcycle = 50, car = 269)	982			
Final log-likelihood, L (β)	-965.6			
Log-likelihood for market share model, L (MS)	-1481.4			
Log-likelihood for equally-likely (EL) model, L (0)	-1807.2			
Total Number of Estimated Parameters = ASVs+ ASCs = 27 + 4	31			
ρo^2 = 1- [L (β) / L (0)]	0.47			
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.45			
ρc^2 = 1- [L (β) / L (MS)]	0.35			
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.33			
Xo^2 (between the final model and the EL model)	1683.3			
Xc^2 (between the final model and the MS model)	1031.7			

Table 148: Nested Logit Model of Most Expensive Vehicle Owned

[Explanatory Variable] General Vehicle Use/Purchase Background

"(Own) Parking Space at Work", "(Own) Parking Space at Home", and "Distance from Work to Subway" are three variables identified as significant. Both variables related to parking show negative influences on non-car alternatives. One possible explanation is, although the "parking space" is originally defined to cover all type of parking (bicycle, motorized two-wheeler, motorcycle, and car), our respondents might tend to consider parking as "car parking only". For most of the non-car parking, there is no designated space in Shanghai, and people generally don't need to pay for using/owning it. Therefore,

people might not perceive the non-car parking as the "real" parking space. Finally, the model shows that people who own (car) parking space at work are less likely to purchase bicycle, motorized two-wheeler and motorcycle (as the most expensive vehicles); people who own (car) parking space at home are less likely to purchase bicycle (as the most expensive vehicle). The expected positive influence of having parking space at work/home for car purchase was not significant. Perhaps having (or not having) a work/home parking space will *not* affect the purchase decision for people who really want (and can afford) a car. The third variable indicates that the further people's workplaces are away from the main subway line, the more likely they will purchase bicycle and motorized two-wheeler (as the most expensive vehicle).

[Explanatory Variable]Perceived Utility of Travel Means

Three (out of six) types of perceived utility are identified as significant – status, speed, and availability; each type has positive sign as expected. The utility of status is significant for bicycle and motorized two-wheeler, indicating that the stronger perception of status people have for those two types of vehicle, the more likely they will purchase them as the most expensive vehicles. Similarly, the perception of utility of speed positively affects the purchase of bicycle (as the most expensive vehicle); and the perception of utility of availability positively affects the purchase of bicycle and motorized two-wheeler (as the most expensive vehicle). Surprisingly, none of the perceived utility, including status, was significant for car purchase. Perhaps the car purchase in Shanghai was less determined by people's utility perception but more by the "actual" constraints, such as income, license control. To further investigate (taking status perception as one example); the cross-tabulation between "car as the most expensive

vehicle owned" and the "car is a symbol of success" has been conducted (Table 149). The Pearson chi-squared test suggests that there are differences in the distributions of status perception between people who purchased cars as the most expensive vehicle and people who don't (p-value = 0.052). However, the statistical significance is borderline, and the fact that this variable is not significant in the model indicates that any explanatory power it might have has been adequately captured by other variables in the model.

			Car is a symbol of success (1=strongly disagree; 5=strongly agree)					
			1	2	3	4	5	Total
Car as the Most Expensive Vehicle Owned	NO	Count	11	57	137	291	217	713
		% within Car as the Most Expensive Vehicle Owned	1.5%	8.0%	19.2%	40.8%	30.4%	100.0%
	YES	Count	6	8	56	106	93	269
		% within Car as the Most Expensive Vehicle Owned	2.2%	3.0%	20.8%	39.4%	34.6%	100.0%
Total		Count	17	65	193	397	310	982
		% within Car as the Most Expensive Vehicle Owned	1.7%	6.6%	19.7%	40.4%	31.6%	100.0%

Table 149: Cross-tabulation: "Car as the Most Expensive Vehicle Owned" x "*Car is a symbol of success*" (Utility Perception)

[Explanatory Variable]Lifestyle Factor

"Pursue Freedom and Control of Life" is the only significant variable for bicycle and motorized two-wheeler identified from the lifestyle factors. Its negative sign indicates that people with freedom-pursuing type of lifestyle are less likely to buy bicycle and motorized two-wheeler.

[Explanatory Variable]Demographics

Nine demographic variables turn out to be significant, and three variables are related to

the ownership of drivers' licenses. As expected, owning motorized two-wheeler or motorcycle license is positively associated with the purchase of motorized two-wheeler and motorcycle. Similarly, the ownership of car drivers' license has positive sign and high magnitude (16.45) for car purchase. Gender (male) is positively associated with the ownership of motorized two-wheeler. Based on my experience conducting the pilot and final surveys in Shanghai, people riding motorized two-wheelers are mostly male and many of them use the vehicles to operate delivery services. However, safety can be another reason, compared to male, female might use public transportation more (if available) for short-distance travel. Three occupation variables are estimated to be correlated with the most expensive mode owned. Students and factory workers are more likely to own bicycles as their most expensive mode; business owners are more likely to own bicycles, motorized two-wheelers, and especially cars (coefficient = 8.2). Family income is significant and positively associated with the car purchase. To test various possible income effects, measures of personal and family income were tested; only family income is significant, and then only for car ownership. This result implies that car purchase, considered as a big expense for Chinese people, is not solely determined by personal income. Household size is negatively associated with the bicycle ownership, but not any other mode. Reasonably, people with big family might be more likely to own more expensive modes with greater passenger capacity and range constraint for their household travel needs (pick up kids, etc.).

The inclusion of ownership of drivers' licenses as explanatory variables may be confounding, because the ownership of drivers' licenses can be considered a pre-requisite of purchasing or using a vehicle. There will be little or no variation on mode-specific

license variables within each group of people who own that vehicle, and incorporating those variables in the model will not provide much behavioral insight. Therefore, I re-run the model with the same specification (as in Table 148) excluding those drivers' license variables (Table 150). As expected, Table 150 presents lower ρ_o^2 (0.34) than previous ρ_o^2 (0.47) in Table 148. However, the new ρ_o^2 value is still within the acceptable range. Besides, the coefficients and the p-values don't change for most the variables, except for "family income (for car)" (the new p-value is higher than 0.05, but it is still included as a borderline insignificant but desirable variable).

Explanatory Variables	*Alternatives (Base Alternative = Purchase Nothing)*			
	Bicycle	**Motorized Two-wheeler**	**Motorcycle**	**Car**
ASC	-1.42 (0.1954)	-1.51 (0.1394)	-7.82 (0.1961)	-33.75 (0.0786)
[PART I] Vehicle Purchase and Use Background				
Parking Space at Work	-3.04 (0.0006)	-3.38 (0.0001)	-7.33 (0.0365)	
Parking Space at Home	-1.06 (0.0014)			
Distance from Work to Subway	0.46 (0.001)	0.47 (0.0009)		
[PART II] Perceived Utility				
Perceived Utility of Status	0.29 (0.0055)	0.47 (0)		
Perceived Utility of Speed	0.46 (0.0002)			
Perceived Utility of Availability	0.66 (0)	0.48 (0)		
[PART IV] Lifestyle Factor				
Pursue Freedom and Control of Life	-0.48 (0.0031)	-0.44 (0.0056)		
[PART V] Demographics				
Male		0.72 (0.0002)		
Student	1.87 (0)			
Factory Worker	0.73 (0.0038)			
Business Owner	2.02 (0.0064)	2.01 (0.0044)		10.58 (0.0386)
Family Income				**4.29 (0.07)**
Household Size	-0.23 (0.0133)			
Goodness-of-fit Statistic				
Number of observations (purchase nothing = 211, bicycle = 257, motorized two-wheeler = 195, motorcycle = 50, car = 269)	982			
Final log-likelihood, L (β)	-1186.1			
Log-likelihood for market share model, L (MS)	-1481.4			
Log-likelihood for equally-likely (EL) model, L (0)	-1807.2			
Total Number of Estimated Parameters = ASVs+ ASCs = 21 + 4	25			
ρo^2 = 1- [L (β) / L (0)]	**0.34**			
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.33			
ρc^2 = 1- [L (β) / L (MS)]	**0.20**			
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.19			
Xo^2 (between the final model and the EL model)	1242.3			
Xc^2 (between the final model and the MS model)	590.9			

Table 150: Nested Logit Model of Most Expensive Vehicle Owned (without "license" variables)

[Vehicle Alternative] "Who are the bicycle buyers?"

Looking at previous tables by column, 14 variables are related to whether bicycles are the most expensive vehicle owned by a respondent. People whose most expensive mode they own is a bicycle tend not to own parking space at work or home and their workplaces tend to be further from the main subway line. In addition, they perceive the utility of status, speed and the availability (convenience) of bicycle; however, they are less likely to have the freedom-pursuing lifestyle per se. Occupation cannot differentiate bicycle buyers very much; they can be factory workers, students or business owners. The last, people with big family are less likely to buy bicycle (as the most expensive vehicle).

[Vehicle Alternative] "Who are the motorized two-wheeler buyers?"

The buyers of motorized two-wheeler are somewhat similar to bicycle buyers. Their workplaces tend to be further away from major subway and without reserved parking spaces. They realize the utility of status and availability of the motorized two-wheeler. The last, they are more likely to be male, business owner and have motorized two-wheeler's licenses.

[Vehicle Alternative] "Who are the motorcycle buyers?"

Only three variables (as ASCs) are associated with the motorcycle purchase. People owning motorcycle tend to have the motorcycle license but don't have (car) parking at work.

[Vehicle Alternative] "Who are the car buyers?"

Four variables (as ASCs) influence the car purchase in this model. The profile for people

buying car (as the most expensive vehicle) in Shanghai is – having car driver's license, business owner and with high family income.

- **MODEL 2: Most Frequently Used Travel Means (weekday)**

In terms of the weekday vehicle use, the final (NL) model was developed with five ASCs and 21 ASVs, representing 12 different variables, as shown in Table 151. All explanatory variables were statistically significant.

Explanatory Variables	*Alternatives (Base Alternative = Walk)*				
	Public Transportation	Taxi, Rented Car, Shared Company Car	Bicycle	Motorized Two-wheeler, Motorcycle	Car
ASC	2.16 (0.0072)	0.61 (0.2368)	-2.21 (0.0049)	-1.74 (0.0233)	-6.64 (0.0001)
[PART I] Vehicle Purchase and Use Background					
Parking Space at Work	-1.04 (0.0017)		-1.48 (0.0007)	-1.23 (0.0066)	
[PART II] Perceived Utility					
Perceived Uitility of Status	0.25 (0.0129)		0.43 (0.0003)		
Perceived Uitility of Availability	0.27 (0.0062)		0.72 (0.0001)	0.69 (0.0002)	0.6 (0.0063)
[PART IV] Lifestyle Factor					
Family-oriented and Green					-0.4 (0.0595)
Pursue Freedom and Control of Life		0.46 (0.0042)			
[PART V] Demographics					
Age	-0.02 (0.0023)				
Own Motorized two-wheeler/Motorcycle License				2.86 (0)	
Own Car License		1.58 (0)			5.4 (0)
Own Car + Motorized two-wheeler/Motorcycle Licenses		1.99 (0)			4.9 (0)
Male			1.08 (0.0038)	1.72 (0)	
Business Owner					1.95 (0.0017)
Personal Income					0.44 (0.0009)
Goodness-of-fit Statistic					
Number of observations (walk = 80; public transportation = 291; taxi, rented car, shared company car = 118; bicycle = 142; motorized two-wheeler, motorcycle = 137; car = 214)	982				
Final log-likelihood, L (β)	-1288.2				
Log-likelihood for market share model, L (MS)	-1675.1				
Log-likelihood for equally-likely (EL) model, L (0)	-1838.2				
Total Number of Estimated Parameters = ASVs+ ASCs = 21 + 5	26				
ρo^2 = 1- [L (β) / L (0)]	0.30				
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.29				
ρc^2 = 1- [L (β) / L (MS)]	0.23				
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.22				
Xo^2 (between the final model and the EL model)	1100.1				
Xc^2 (between the final model and the MS model)	773.8				

Table 151: Nested Logit Model of Most Frequently Used Travel Means (weekday)

In term of goodness-of-fit statistics, the ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.30, indicating that the estimated (full) model explains 30% (equally-likely model as the base) of the information in the survey data. The ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.23, which means the estimated model can still explain 23% of the information in the data under a more strict (market-share model) base.

The χ_c^2 value of 773.8 means there is a significant difference between the final model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1100.1 means that the estimated model significantly differs from the equally-likely (EL) model at $\alpha << 0.005$.

[Explanatory Variable] General Vehicle Use/Purchase Background

Like previous vehicle purchase model, owning (car) parking at work negatively influence the weekday use of public transportation, bicycle, motorized two-wheeler and motorcycle with similar magnitude.

[Explanatory Variable]Perceived Utility of Travel Means

Though the perceived status of motorized two wheelers and motorcycles, taxis, rented cars, shared company cars, and privately owned cars, is not significantly related to the use of any of these modes, the status of using public transit and bicycle is positively associated, respectively, with the use of public transportation and bicycle. Availability is positively associated with use of public transportation, bicycle, motorized two-wheeler, motorcycle and car, but not taxis, rented cars, and shared company cars.

[Explanatory Variable]Lifestyle Factor

The lifestyle factor I called "family oriented and green" could have different effects on the same mode for different reasons; therefore I did not have a simple hypothesis regarding this factor. For example, a family-oriented person might be more motivated to use a car to fulfill family needs that cannot be well served by other types of vehicles, e.g., driving the whole family out for a picnic. On the other hand, if a person is family-oriented and environmentally-concerned (green); he/she might choose to stay home more, intentionally reduce the car use, or use other modes more aligned with this factor. Based on the model results, the factor "family-oriented and green" turns out to be negatively associated with car use on weekdays. The freedom-pursuing lifestyle is positively associated with the use of taxi, rented car and shared company car. A freedom-pursuing person might be more likely to use some "ad hoc" type of vehicle such as taxi.

[Explanatory Variable]Demographics

According to the model results, age is slightly negatively associated with the use of public transportation (-0.02). Similar to previous vehicle ownership model, possession of drivers' licenses also positively affect vehicle use. Owning car drivers' license is positively related to not only the car use, but also the use of taxi, rented car and shared company car. Men are more likely to use bicycles, motorized two-wheelers and motorcycles during weekday. Business owners and people with high personal incomes are more likely to be weekday car users.

As mentioned in previous model, the inclusion of those "license" variables is debatable and does not provide much behavioral insight. I re-run the model with the same specification (as in Table 151) excluding those license variables. As expected, Table 152 shows lower ρ_o^2 (0.22) than previous ρ_o^2 (0.30) in Table 151. But the new ρ_o^2 value is within my acceptable range. The coefficients and the p-values don't change for most of the variables, and the variable "family-oriented and green" is still included as a borderline insignificant but desirable variable (p-value = 0.0832).

Explanatory Variables	*Alternatives (Base Alternative = Walk)*				
	Public Transportation	**Taxi, Rented Car, Shared Company Car**	**Bicycle**	**Motorized Two-wheeler, Motorcycle**	**Car**
ASC	2.61 (0.0293)	2.05 (0.0424)	-2.44 (0.0067)	-1.73 (0.048)	-9.51 (0.0015)
[PART I] Vehicle Purchase and Use Background					
Parking Space at Work	-0.93 (0.0069)		-1.62 (0.005)	-1.44 (0.0131)	
[PART II] Perceived Utility					
Perceived Utility of Status	0.3 (0.0043)		0.46 (0.0004)		
Perceived Utility of Availability	0.3 (0.0052)		0.93 (0)	0.89 (0.0001)	1.13 (0.0045)
[PART IV] Lifestyle Factor					
Family-oriented and Green					**-0.48 (0.0832)**
Pursue Freedom and Control of Life		0.5 (0.0023)			
[PART V] Demographics					
Age	-0.02 (0.0294)				
Male			1.35 (0.0281)	2.33 (0.0003)	
Business Owner					2.27 (0.0092)
Personal Income					1.31 (0.0011)
Goodness-of-fit Statistic					
Number of observations (walk = 80; public transportation = 291; taxi, rented car, shared company car = 118; bicycle = 142; motorized two-wheeler, motorcycle = 137; car = 214)	982				
Final log-likelihood, L (β)	-1433.9				
Log-likelihood for market share model, L (MS)	-1675.1				
Log-likelihood for equally-likely (EL) model, L (0)	-1838.2				
Total Number of Estimated Parameters = ASVs+ ASCs = 16 + 5	21				
ρo^2 = 1- [L (β) / L (0)]	**0.22**				
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.21				
ρc^2 = 1- [L (β) / L (MS)]	**0.14**				
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.13				
Xo^2 (between the final model and the EL model)	808.6				
Xc^2 (between the final model and the MS model)	482.3				

Table 152: Nested Logit Model of Most Frequently Used Travel Means (weekday) (without "license" variables)

[Vehicle Alternative] "Who are the public transportation riders?"

Public transportation riders in Shanghai don't own (car) parking space at work, but they perceive the utility of status and availability of using public transportation. Increasing age is associated with lower public transit use. It is possible that when people getting older, they are able to (possibly due to higher income) go for something "nicer." Further, the public transportation service in Shanghai (during weekday) may not be safe or comfortable enough for them (Figure 31).

Figure 31: Subway of Shanghai (weekday)
Source: cn.yahoo.com

[Vehicle Alternative] "Who are the taxi, rented car, and shared company car users?"

People using taxi, rented car and shared company car during weekday are more likely to have freedom-pursuing lifestyle. Owning car, motorized two-wheeler, or motorcycle drivers' license all positively affect the weekday use of these modes.

[Vehicle Alternative] "Who are the bicycle riders?"

Weekday cyclists in Shanghai don't own (car) parking space at work but perceive the

utility of status and availability of riding bicycle. They are more likely to be male.

[Vehicle Alternative] "Who are the motorized two-wheeler and motorcycle riders?"

People using these two types of two-wheeled vehicles during weekday don't own (car) parking space at work, perceive the utility of the availability of motorized two-wheelers and motorcycles, are more likely to be male, and have the appropriate vehicle drivers' license (with the highest coefficient = 2.86).

[Vehicle Alternative] "Who are the car users?"

People who drive a car on weekdays in Shanghai value its availability, i.e., using their car anytime, anywhere they want. In terms of lifestyle, they are less likely to be family-oriented or environmentally-concerned. They are more likely to be business owners, to have car drivers' license and high personal incomes. In contrast to purchasing car as the most expensive vehicle (which is positively associated with family income), weekday car use seems to be a more "personal" decision, associated with personal income.

- **MODEL 3: Most Frequently Used Travel Means (weekend)**

The final MNL model of weekend vehicle use was achieved with five ASCs and 17 ASVs. All explanatory variables in the model were statistically significant (Table 153).

Explanatory Variables	*Alternatives (Base Alternative = Walk)*				
	Public Transportation	Taxi, Rented Car, Shared Company Car	Bicycle	Motorized Two-wheeler, Motorcycle	Car
ASC	1.19 (0)	-2.16 (0.0004)	-3.11 (0)	-1.32 (0.0034)	-5.21 (0)
[PART I] Vehicle Purchase and Use Background					
Distance from Home to Subway				0.44 (0.0083)	
[PART II] Perceived Utility					
Perceived Uitility of Status	0.16 (0.0238)		0.35 (0.0169)		
Perceived Uitility of Speed			0.48 (0.0068)		
Perceived Uitility of Availability					0.3 (0.0086)
Perceived Uitility of Carrying Capacity		0.29 (0.019)			
[PART III] Exogenous Environment Factor					
Cost-related Policy/Regulation					-0.48 (0.0002)
[PART V] Demographics					
Own Motorized two-wheeler/Motorcycle License				2.83 (0)	
Own Car License					2.69 (0)
Own Car + Motorized two-wheeler/Motorcycle Licenses				1.62 (0)	2.83 (0)
Male			0.85 (0.0029)		
No-work or Retired		0.88 (0.0097)			
Personal Income		0.27 (0)			0.38 (0)
Family Income		0.16 (0.0204)			0.26 (0.0075)
Goodness-of-fit Statistic					
Number of observations (walk = 67; public transportation = 358; taxi, rented car, shared company car = 183; bicycle = 59; motorized two-wheeler, motorcycle = 89; car = 226)	982				
Final log-likelihood, L (β)	-1234.2				
Log-likelihood for market share model, L (MS)	-1560.2				
Log-likelihood for equally-likely (EL) model, L (0)	-1759.5				
Total Number of Estimated Parameters = ASVs+ ASCs = 17 + 5	22				
ρo^2 = 1- [L (β) / L (0)]	0.30				
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.29				
ρc^2 = 1- [L (β) / L (MS)]	0.21				
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.20				
Xo^2 (between the final model and the EL model)	1050.7				
Xc^2 (between the final model and the MS model)	652.0				

Table 153: Multinomial Logit Model of Most Frequently Used Travel Means (weekend)

In term of goodness-of-fit statistics, the ρ_o^2 (the Rho square value of estimated model with equally-likely model as the base) is 0.30, indicating that the estimated (full) model explains 30% (equally-likely model as the base) of the information in the survey data. However, the ρ_c^2 (the Rho square value of estimated model with market-share model as the base) is 0.21, which means the estimated model can still explain 21% of the

information in the data under a more strict (market-share model) base.

The χ_c^2 value of 652.0 means there is a significant difference between the final model and the MS model at $\alpha << 0.005$. Similarly, the χ_o^2 value of 1050.7 means that the estimated model significantly differs from the equally-likely (EL) model at $\alpha << 0.005$.

[Explanatory Variable] General Vehicle Use/Purchase Background

"Distance from Home to Subway" is the only significant variable associated with using motorized two-wheelers and motorcycles on weekends. Its positive sign suggests that the further people's home tend to be from the main subway line the more likely they are to use motorized two-wheeler or motorcycle on weekend.

[Explanatory Variable]Perceived Utility of Travel Means

The perception of the status of using these modes is positively associated with the use of public transportation and bicycle. In addition, the perception of speed, availability have positive effect on the use of bicycle, car. The carrying capacity, not being identified in previous models, is positively associated with the weekend use of taxi, rented car and shared company car. Surprisingly, the carrying capacity is not significant to car use. Perhaps a private car full of personal belongings is less considered "with capacity" than an always-empty taxi. Alternatively, perhaps car users and non-users alike have a similar perception of the carrying capacity of the car, in which case it cannot help distinguish users from non-users. To confirm, a cross-tabulation between "car as the most frequently used means" and the "car has carrying capacity" has been conducted (Table 154). The Pearson chi-squared test indicates that there is *no* difference in the distributions

of capacity perception between people who use car most frequently during weekend and people who don't (p-value = 0.263).

			Car has carrying capacity (1=strongly disagree; 5=strongly agree)					
			1	2	3	4	5	Total
Car as the Most Frequently Used Means (weekend)	NO	Count	5	31	96	363	261	756
		% within Car as Most Frequently Used Means (weekend)	.7%	4.1%	12.7%	48.0%	34.5%	100.0%
	YES	Count	1	4	21	115	85	226
		% within Car as Most Frequently Used Means (weekend)	.4%	1.8%	9.3%	50.9%	37.6%	100.0%
Total		Count	6	35	117	478	346	982
		% within Car as Most Frequently Used Means (weekend)	.6%	3.6%	11.9%	48.7%	35.2%	100.0%

Table 154: Cross-tabulation: "Car as the Most Frequently Used Means" x *"Car has Carrying Capacity"* (Utility Perception)

[Explanatory Variable]Exogenous Environment Factor

The "Cost-related Policy/Regulation" is negatively associated with the car use. As expected, this result reflects some policy concepts in Shanghai, for example, the car "ownership" license control and the accompanying high cost of license plates.

[Explanatory Variable]Demographics

Similar to weekday vehicle use, the ownership of drivers' licenses positively affect the weekend vehicle use, and male is positively associated with the use of bicycle. People without job (or retired) are more likely to take taxi, rented car or shared company car during weekend. Personal and family income are both identified as significant and positively associated with the use of car-related travel means (taxi, rented car, shared company car, and private car).

Similar to the previous two cases, an alternative model with identical specification but no "license variables" is developed. As expected, Table 155 shows lower ρ_o^2 (0.23) than previous ρ_o^2 (0.30) of Table 153. However, the new ρ_o^2 value is within my acceptable range. Basically, the coefficients (signs and magnitudes) and the p-values don't change for most of the variables from the previous model to this alternative model.

Explanatory Variables	*Alternatives (Base Alternative = Walk)*				
	Public Transportation	Taxi, Rented Car, Shared Company Car	Bicycle	Motorized Two-wheeler, Motorcycle	Car
ASC	1.07 (0)	-2.15 (0.0003)	-3.16 (0)	-0.71 (0.0798)	-5.85 (0)
[PART I] Vehicle Purchase and Use Background					
Distance from Home to Subway				0.42 (0.0051)	
[PART II] Perceived Utility					
Perceived Utility of Status	0.21 (0.0033)		0.35 (0.0153)		
Perceived Utility of Speed			0.52 (0.0036)		
Perceived Utility of Availability					0.45 (0)
Perceived Utility of Carrying Capacity		0.29 (0.019)			
[PART III] Exogenous Environment Factor					
Cost-related Policy/Regulation					-0.37 (0.0008)
[PART V] Demographics					
Male			0.71 (0.0129)		
No-work or Retired		0.9 (0.0034)			
Personal Income		0.26 (0)			0.63 (0)
Family Income		0.16 (0.0175)			0.32 (0.0005)
Goodness-of-fit Statistic					
Number of observations (walk = 67; public transportation = 358; taxi, rented car, shared company car = 183; bicycle = 59; motorized two-wheeler, motorcycle = 89; car = 226)	982				
Final log-likelihood, L (β)	-1349.1				
Log-likelihood for market share model, L (MS)	-1560.2				
Log-likelihood for equally-likely (EL) model, L (0)	-1759.5				
Total Number of Estimated Parameters = ASVs+ ASCs = 13 + 5	18				
ρo^2 = 1- [L (β) / L (0)]	**0.23**				
Adjusted ρo^2 = 1- {[L (β) - Total # of Estimated Parameters] / L (0)}	0.22				
ρc^2 = 1- [L (β) / L (MS)]	**0.14**				
Adjusted ρc^2 = 1- {[L (β) - Total # of Estimated ASVs] / L (MS)}	0.13				
Xo^2 (between the final model and the EL model)	820.8				
Xc^2 (between the final model and the MS model)	422.2				

Table 155: Multinomial Logit Model of Most Frequently Used Travel Means (weekend) (without "license" variables)

In fact, according to Table 156, the personal and family incomes are highly correlated (0.705). Such collinearity could be a concern if both variables appear in a model. There are common two ways to detect this problem – one variable may have the

counterintuitive sign, and/or one or both variables may have high standard errors (suggesting insignificant). However, in this model, the two highly-correlated variables are both significant and with the expected sign. Therefore, I keep both variables in, since the data are sensitive enough to successfully distinguish separate effects of personal and family income.

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Interval by Interval	Pearson's R	.694	.007	74.052	.000[c]
Ordinal by Ordinal	Spearman Correlation	**.705**	.008	76.344	.000[c]
N of Valid Cases		5892			

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.
c. Based on normal approximation.

Table 156: Correlation between Personal and Family Income

[Vehicle Alternative] "Who are the public transportation riders?"

In terms of weekend vehicle use, the perception of its status is the only significant (positive) variable for public transportation.

[Vehicle Alternative] "Who are the taxi, rented car, and shared company car users?"

People using those three types of vehicle realize the utility of carry capacity. In terms of demographic characteristics, they are retired (or without job) and with high personal or family income.

[Vehicle Alternative] "Who are the bicycle riders?"

The weekend bicycle riders are more likely to be male. Besides, the utility of bicycle status and speed are two variables positively affecting their weekend use behavior.

[Vehicle Alternative] "Who are the motorized two-wheeler and motorcycle riders?"

The people riding motorized two-wheeler or motorcycle live away from major subway lines and having the motorized two-wheeler or motorcycle drivers' licenses.

[Vehicle Alternative] "Who are the car users?"

Weekend car user is positively affected by the utility of availability and negatively affected by the cost-related policy and regulation. Owning the car drivers' license and high personal/family income are another two significant (and positive) variables toward the weekend car use.

CHAPTER 6: CONCLUSION

The last dissertation chapter recaps the answers to three key questions: How to conduct survey research in China? What is the motorization pathway in China? What is the vehicle purchase (and use) behavior in China? In addition, implications and recommendations are provided and future research directions are discussed.

Q1: How to conduct survey research in China?

Several lessons can be drawn from the pilot and final surveys about how to conduct survey research in China.

Lessons from Pilot Survey

Lesson #1: The concept of "dream life" is vague for most people

In pilot survey, there were a series of questions asking about a "dream life" as a way for respondents to think about how they would want their life to be, including, how they would travel. However, such questions about the future proved the most difficult for the respondents. 11 percent of respondents declined to answer at all; another 54 percent provided either vague or partial answers. Many people commented on this question with "I have never thought about that," and "This does not relate to me, since I don't think I can change my life." The comments suggest that some people (especially people surveyed in the lower-income neighborhood, like the old town Shanghai) may not be ready yet to imagine a future very different from their past and present, raising difficulties for studying hypothetical topics.

Lesson #2: Trust, convenience, and comfort are keys to success

Considering the pilot survey was mostly conducted on street, trust was the number one factor affecting people's willingness to participate. It is not only that people we approached were busy, but they also tended to confuse the survey team with street venders, or possibly scams. To remedy this situation, our team began to show student IDs when approaching people. Besides, we find out that a convenient and comfortable environment will increase people's motivation of participation. A convenient environment means a place where people can quickly and easily complete their questionnaire. A comfortable environment could be a place with air-conditioning in the hot summer time or a location where people feel mentally comfortable.

Lesson #3: Getting to know the place is also important

As a researcher who conducts survey research in a new place for the first time, orientation to the setting is important to develop a basic sense of the people and the area. The implementation of on-street interviewing in the pilot phase helped me finalize the sample design and location selection for the final phase.

Lessons from Final Survey

Lesson #1: A complicated and long survey is challenging

Due to the motorization pathway, vehicle purchase, and vehicle use I want to research, the final survey was long and complex. I attempted to balance time demands on respondents and questionnaire complexity so that people would agree to participate and not lose patience. However, many people refused to participate when

they realized the survey was seven double-sided pages. In terms of individual questions, most people complain about the PART II (utility rating) because it is vague, e.g., the term "symbol of success," and repetitive – requiring respondents to rate six types utility on nine different travel means.

Lesson #2: Anonymity/confidentiality, authorization, and study topic are top three factors affecting people's motivation to participate

According to the survey results of 1,037 respondents, anonymity/confidentiality, authorization, and study topic were the top three considerations in whether people initiated and completed a questionnaire. 55 percent of respondents considered that "my answer will be anonymous and confidential" to be moderately or extremely important. Similarly, 46 percent said that the "authorization letter" and 41 percent said "the survey topic" were moderately or extremely important.

- **Implications and/or Recommendations**

In the following, recommendations are provided about "how to conduct survey research in China?" in terms of survey type, sampling, questionnaire design, and local implementation.

Survey Type: Single survey type is recommended

Face-to-face and on-line are two survey types used in the final survey – 1,037 people responded to face-to-face interviews on-street, at car dealerships, in their households, or through the author's peer network. A further 78 people were recruited by cell phone text messages to an on-line version of the questionnaire. According to the interview

experiences and the analysis of factors affecting the willingness of participation, face-to-face respondents have different answers than the on-line respondents. The on-line respondents, surveyed without in-person contact but with better privacy and freedom, are more likely to reveal positive attitudes. People involved in a face-to-face interview tend to skip questions, e.g., sensitive or too-personal questions, or show negative-to-middle attitudes. It is recommended that single survey type should be used to ensure people respond to the same questionnaire under similar (if not standard) environment.

Sampling: Use location-based sampling strategically but carefully

Location-based, convenience sampling was used to overcome the inability to collect a random sample. According to the cross-tabulations and Chi-square tests (conducted in Chapter 3), the location-based sampling scheme used in this study was effective at capturing groups owning specific types of vehicles and using specific travel means during weekdays. However, based on the same tests, the location-based sampling might not work perfectly to capture specific groups based on their most frequent used travel mean on weekends. Understanding or estimating the true share of population is essential for the selection of sampling locations. Understanding the relationship between locations and sampling, for example, the relationship between car dealerships and car buyers is also important.

Questionnaire Design: Be short, straightforward, and clear. Avoid vague or sensitive questions

Although there is no specific definition of a long questionnaire, the final survey experiences suggest that my seven-page, double-sided questionnaire was too long for

many respondents. Based on my discussion with professors in the Department of Transportation Engineering of Tongji University, the survey they usually conduct is one-page, single-side containing less than ten questions, and requires less than ten minutes to finish (compared to my survey, which required at least 30 minutes). Moreover, in my survey, there are some vague or sensitive questions such as utility perception or income. Those questions require more effort and more trust from respondents and affect their motivation of participating or completing the survey. I recommend that, if the questionnaire cannot be shortened, the questions should be straightforward and clear, and the overall questionnaire structure should be easy to follow.

Local Implementation: Trust is the top factor

Gaining trust from people is the top factor of the local survey implementation. Without basic trust from people, a face-to-face interview cannot even start. Therefore, the cooperation with local, authorized organizations, e.g., university and government agency, is recommended to increase trust from potential respondents. In addition, finding a convenient and comfortable location, a team speaking local Shanghainese in this case, will all help to facilitate implementation.

Q2: What is the motorization pathway in China?

As defined in Chapter 4, a motorization pathway is the transition which individual people make among different travel means. In this dissertation, three aspects of motorization pathways were analyzed: common pathway patterns, total number of stages (without looping back through the same modes), and the sequence (motorization direction).

Motorization pathways are diverse, complicated (multi-staged) and mostly without car

In the motorization pathway analysis, based on the location-based sample, the relative frequency of motorization pathways between people at different motorization stages is reviewed. Overall, motorization pathways in my sample of Shanghai residents are diverse – the 992 respondents reported 331 distinct pathways – and complicated – more than 50% of people have pathway involving more than three stages. Among the top 30 patterns, there are only eight patterns (covering 11% of the sample) that include car, which suggests car is not *yet* a widely used mobility option. Concluding from above, a common motorization pathway within my sample is a series of three (or more) transitions in the types of travel modes that a Shanghainese person has used for their daily travel.

Hypothetical motorization direction is challenged

A hypothetical motorization direction from non-motorized to motorized, from low cost to high cost, and from shared to private-owned was proposed. Only about half of respondents come to their current motorization stage by following the hypothetical motorization direction. There are two possible explanations: first, the definition of so-called high-cost and low-cost means may not be as I assumed. That is, walk (hypothesized as low-cost) may be perceived as a higher cost travel means by respondents than bicycle (hypothesized as high cost) due to potentially longer travel times (assuming that people has non-zero value of time) and more effort involved. Second, even if the cost definition of travel means is what I assumed; the direction of transition may not be as hypothesized. In conclusion, motorization pathways in Shanghai are diverse and complex.

- **Implications and/or Recommendations**

Examining all patterns provides more insights

People experience motorization pathways with different sequences and number of stages. Therefore, listing out all the patterns (with the corresponding number of cases) provides more insights than simply observing or talking with people in Shanghai. For example, by examining the motorization in the backward direction (which cannot be developed unless we list all patterns), we understand that most people who are currently in the car stage come from a taxi or rented car stage. This result is interesting but not surprising. As a matter of fact, taxi or rent car can be considered a type of car in terms of basic functionality. However, car is a private mode and can be owned. The idea of ownership is a major difference between private car and the taxi or rented car. I assume the "upgrade" from taxi to private car may also involve certain attributes such as vehicle status.

Some policy can be considered based on the insights of motorization pathway analysis. For instance, a better taxi (or rental car) service *might* slow the growth of private car ownership (i.e., keep people stay in the taxi stage and from move to the car stage), since many people in the private car stage of motorization came from the taxi or rented car. Besides, Shanghai government should realize the diversity and complexity of motorization pathways happening in the city and have policies to guide those pathways toward a more sustainable future – innovative transportation, land use, or even energy policies should be considered as options.

Modified method to study the motorization pathway is recommended

Future research could build on this analysis by using a stratified sample based on people's

age, ownership of driver license or living location to create a more robust accounting of motorization pathways. It would also be useful to distinguish motorization stages that occurred prior to the respondent moving to Shanghai if they are immigrants from somewhere else. Finally, it would be good to extend the sample to residents of other regions besides Shanghai because their motorization pathways may be completely different from Shanghai residents due to the huge regional variation in China.

Q3: What is the vehicle purchase and use behavior in China?

Comparison of three models

In Table 157, I compare the results of three vehicle choice models (the versions without the debatable license-holding variables mentioned in Chapter 5) – most expensive vehicle owned, most frequently used travel means (weekday), and most frequently used travel means (weekend).

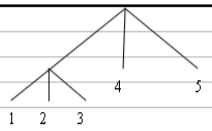
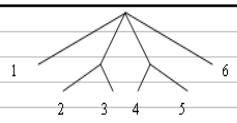
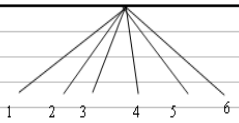
	purchase nothing = 1, bicycle = 2, motorized two-wheeler = 3, motorcycle = 4, car = 5	walk = 1; public transportation = 2; taxi, rented car, shared company car = 3; bicycle = 4; motorized two-wheeler, motorcycle = 5; car = 6	walk = 1; public transportation = 2; taxi, rented car, shared company car = 3; bicycle = 4; motorized two-wheeler, motorcycle = 5; car = 6
			
	Most Expensive Vehicle Owned 13 Explanatory Variables	**Most Frequently Used Travel Means (weekday)** 9 Explanatory Variables	**Most Frequently Used Travel Means (weekend)** 10 Explanatory Variables
PART I	Parking Space at Work Parking Space at Home Distance from Work to Subway	Parking Space at Work	Distance from Home to Subway
PART II	**Perceived Utility of Status** Perceived Utility of Speed **Perceived Utility of Availability**	**Perceived Utility of Status** **Perceived Utility of Availability**	**Perceived Utility of Status** Perceived Utility of Speed **Perceived Utility of Availability** Perceived Utility of Carrying Capacity
PART III			Cost-related Policy/Regulation
PART IV	Pursue Freedom and Control of Life	Family-oriented and Green Pursue Freedom and Control of Life	
PART V	**Male** Student Factory Worker Business Owner **Family Income** Household Size	Age **Male** Business Owner **Personal Income**	**Male** No-work or Retired **Personal Income** **Family Income**

Table 157: Comparison of Three Models

Different variable specifications and even model structures (MNL and NL) were identified. However, some common variables (highlighted in blue) such as the perceived utility of status and income (personal or family) exist in all three models. By looking at those three models side-by-side, I conclude that they are not quite similar, although I originally expected three almost identical models for the vehicle purchase and use.

Income – revisiting at the individual level

A positive correlation between income and vehicle purchase has been confirmed at the aggregate (national, inter-national) level. In this dissertation, individual and household incomes have been analyzed as explanatory variables in modeling the most expensive vehicle owned and the most frequently used travel means for weekday and weekend.

Vehicle purchase and use behaviors are more complicated when considering the disaggregate level. For example, at national level, the cross-nation comparisons of GDP/capita show an almost linear increasing relationship with aggregate auto ownership. However, more variation are found when researching into disaggregate level – based on the model results, several variables, i.e., gender and utility perception of different travel modes, other than *just* income are significant. Moreover, different type (personal and family) of incomes are associated with different models, and the relationship between income and vehicle purchase/use would be different from the one identified at the national level.

The status of "bicycle" – utility of low/non-motorized travel means is recognized

One interesting finding about the utility perception is that most of the utility perceptions, e.g., status and availability, are identified as significant to low/non-motorized travel means, e.g., bicycle, walk, instead of the highly-motorized travel means, e.g., car. In their questionnaires, some people mentioned the status of riding a bicycle or walking as, "I have enough money and am retired, so I no more need to drive to work or commute everyday." Interestingly, we find out similar case as the "motorization backward" in the pathway analyses. That is, some people's most frequently used travel means change from car to bicycle or walk.

- **Implications and/or Recommendations**

Income is not the sole driving force and its effect can be further differentiated

According the conclusion of Chapter 5, income is not the sole driving force for vehicle use or purchase. Variables such as gender, utility perception are also important. Besides,

the income effect can be further differentiated, for example, purchasing car is (so far) considered a "family decision" (positively associated with family income); whereas, the weekday car use seems to be a more "personal choice" (positively associated with personal income.)

Using vehicle choice model and motorization pathway to better forecast the future

Ideally, the results of vehicle choice model and motorization pathway should inform each other, and we did find such case as the "backward motorization vs. utility of low-motorized travel means." I suggest, for better describe current motorization and forecast the future, the pathway analysis and choice model should be conducted and analyzed at the same time. After all, the choice model can only present a "snap shot" of status quo and should be complemented with pathway analysis, which involves deeper understanding of the transition.

Future Research

Combining the results of motorization pathway and vehicle choice model analysis in Shanghai, an important and intriguing message from this dissertation is that the process of motorization is complicated; there are multiple pathways influenced by not only income but also utility perceptions of different modes, lifestyle orientations, and exogenous environment attributes.

Although Shanghai itself can not represent the whole China (actually, no single city or region can represent whole China), the Shanghai study is representative of a city

experiencing rapid economic growth and with various transportation alternatives. Thus, the results of the motorization pathways and choice models of the Shanghai study may be generalizable to certain cities which are expected to experience similar growth and diversity.

Using the Shanghai study to reflect on China, I speculate that motorization in China is an interaction between its internal development and the globalization. The term "half-globalization" was used by Dr. Y. T. Lee (1986, Nobel Laureate) in a speech on November, 2007. That is, when China leverages global resources to achieve its economic growth (and hence the increase of auto ownership), many issues (such as greenhouse gas emissions) are generated but not globally addressed. Because of the rapid change in developing countries, properly addressing global issues such as the environment, economy, and auto markets is necessary but challenging. Nevertheless, a general recommendation for future research direction is that a global perspective is needed, even for people conducting research locally in China. For example, in addition to the inventory check of vehicle ownership, issues about oil price, global warming (to understand how those issues affect the vehicle purchase) should be included in future motorization studies. Thus, the research is expected to provide not only a solution/recommendation to China, but also a "global solution." After all, the contribution of the China research should not only come from the how well it addresses the local issues, but also from how well it address the linkage between China-specific issues with the world.

REFERENCES

- **Chapter 1**

[1] Schipper, Lee and Ng, Wei-Shiuen (2004) Rapid Motorization in China: Environmental and Social Challenges, EMBARQ, World Resource Institute, Washington D.C.

[2] Christopher, Cherry R. (2007) Electric Two-Wheelers in China: Analysis of Environmental, Safety, and Mobility Impacts, Doctoral Dissertation, University of California, Berkeley.

[3] Doi, Naoko (2005) APERC Database, Asia Pacific Energy Research Centre, The Institute of Energy Economics, Japan.

[4] National Bureau of Statistics of China (1985 – 2006) *China Statistical Yearbook.*

[5] Shen, Zhongyuan, Kokichi, Ito, and Li Zhidong (2002) Outlook for China's Motorization and Energy Consumption, Institute for Energy Economics Japan (IEEJ), p.3.

[6] Weinert, Jonathan (2007) The Rise of Electric Two-wheelers in China: Factors for their Success and Implications for the Future, Doctoral Dissertation, University of California, Davis.

[7] Datamonitor (2004) Automobile Manufacturers in China: Industry Profile.

[8] PaoHua Economic Research Institute (2005) China Automotive Market Report.

[9] *The Economist* (2003)

[10] Farrell, Diana, Gersch, Ulrich A. and Stephenson, Elizabeth (2007) The value of China's emerging middle class, *The McKinsey Quarterly* (2006 special Edition).

[11] China Automotive Industry Yearbook House (2000) *China Automotive Industry Yearbook* (*中國汽車工業年鑒*).

[12] Gould, Stephen and Wong, Nancy Y. C. (2000) The Intertextual Construction of Emerging Consumer Culture in China as Observed in the Movie Ermo: A Postmodern, Sinicization Reading.

[13] Chinese national standard (GB/T3730.1-2001) of the General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China.

[14] Shanghai Bureau of Statistics (2002, 2003, 2004, 2006) *Shanghai Statistical Yearbook*

[15] www.wikipedia.org (Accessed 2005)

[16] Committee on the Future of Personal Transport Vehicles in China, National Research Council, National Academy of Engineering, Chinese Academy of Engineering (2003) *Personal Cars and China*, National Academy Press, p. 224. (Appendix B).

[17] Department of Traffic Police, City of Shanghai (2006) Motor Vehicle Database.

[18] Chi Hung Kwan (2002) How Far is Coastal China behind the Industrialized Countries? – An Analysis Based on Purchasing Power Parity, *China In Transition*.

[19] US CIA (2001), *World Factbook*.

[20] World Bank (2002), *World Development Report*.

[21] Mercer Management Consulting (2004), Chinese Automotive Market 2010: Facts, Trends and Strategic Challenges for Profitable Growth.

[22] Wang Xia and Qi Fang (2000) The Coordinated Development of Housing and Transport and The Restructuring of Shanghai's Spatial Layout, *City Planning Review* Vol . 24, No. 3 (published in China).

- **Chapter 2**

[23] Shanghai City Comprehensive Transport Planning Institute (2004) Master Transportation Survey

[24] Cullinan, Sharon and Cullinane, Kevin (2003) Car Dependence in a Public Transport Dominated City: Evidence from Hong Kong, *Transportation Research* Part D8, pp. 129-138.

[25] Bell, A. Colin, Ge, Keyou and Popkin, Barry M. (2002) The road to obesity or the path to prevention: motorized transportation and obesity in China. *Obes Res*. 2002; 10: p.p. 277–283.

[26] Shanghai City Comprehensive Transport Planning Institute (2000) Distance-based Mode Share Data

[27] Ben-Akiva, Moshe and Lerman, Steven R. (1985) Discrete Choice Analysis: Theory and Application to Travel Demand, *MIT Press*.

[28] Energy Foundation (2005) Car Purchase Survey, Beijing

- **Chapter 5**

[29] Mokhtarian, Patricia (2006) Class Handouts of TTP 200 (Transportation Survey Methods).

[30] Fisher, Walter H. (2001) Status Preference, Wealth, and Dynamics in the Open Economy. *Economics Series*, Institute for Advanced Studies, Vienna, p. 1.

[31] www.ford.com.cn (Accessed 2007).

[32] www.lexus.com.cn (Accessed 2007).

[33] autos.cn.yahoo.com (Accessed 2007).

[34] cn.yahoo.com (Accessed 2007).

[35] Small, Kenneth A. and Hsiao, Cheng (1985) Multinomial logit specification tests, *International Economic Review* 26(3), 619-627.

APPENDICES

- **Pilot Survey Questionnaire**

Survey Date: ____/____/____ Survey Time: _____AM/PM

Survey Location: _____ Surveyor: _____

PART 1: About Current

1. Do you have any vehicle now? Please describe the vehicle you have **purchased**[35] **most recently.** If you have one, **what** is that? If NOT, to **Q7**

- ☐ Bicycle
- ☐ Auxiliary Power Vehicle[36] [APV] (_____ Powered by?)
- ☐ Motorcycle[37] (_____ cc.?)
- ☐ Car (_____ Year; _____ Make; _____ Model)
- ☐ Other type of vehicle (_____ please specify)

2. If not you, who **owns** (registered under) this vehicle? (_____)

3. If not you, who **uses** this vehicle most often? (_____)

4. "**When and where** did you buy it?"

[35] Excluding: Getting vehicle as a gift from others.
[36] 2- or 3-wheeler powered by electricity or LPG. The "APV" is called as "motorized two-wheeler" in my dissertation.
[37] 2- or 3-wheeler powered by gasoline

5. "**Why** did you purchase this vehicle?" (Please designate 1st, 2nd and 3rd reasons in EACH attribute category)

   [Personal Attribute]

   - ☐ This vehicle makes me look "better"! And I think it will be a good tool for my career (finding job, get promotion) and social life.
   - ☐ This is a cool stuff. I feel satisfied owning this vehicle.
   - ☐ Many of my friends, colleagues have this vehicle.
   - ☐ I have used this type of vehicle before and was satisfied.
   - ☐ Freedom. For example, going out of the city to enjoy life, visiting family home. Basically, I can go anywhere anytime I want!
   - ☐ I enjoy the time traveling with my vehicle
   - ☐ I am saving money and I don't want to spend too much on transportation.
   - ☐ Other (_____ please specify)

   [Vehicle Attribute]

   - ☐ This vehicle helps me to carry things. (e.g., goods, kids, friends)
   - ☐ This vehicle is what I can afford right now. (_____ purchase price? _____registration fee? _____license cost? _____ operating cost/month?)
   - ☐ This vehicle saves me time.
   - ☐ This vehicle gives me safety
   - ☐ This vehicle gives me comfort
   - ☐ This vehicle gives me privacy
   - ☐ This vehicle is green and clean

- ☐ This vehicle is used for business purposes
- ☐ Other (_____ please specify)

[Exogenous Environment Attribute]

- ☐ Public transportation is not available where I work or live.
- ☐ Public transportation is not convenient (too slow, too many transfers) for my week day life (e.g., school, work)
- ☐ Public transportation is not comfortable (too crowded, too bumpy) for my week day life (e.g., school, work)
- ☐ Public transportation is not convenient (too slow, too many transfers) for my weekend life (e.g., shopping, going to movie, going outside Shanghai)
- ☐ Public transportation is not comfortable (too crowded, too bumpy) for my weekend life (e.g., shopping, going to movie, going outside Shanghai)
- ☐ There is no parking space near my house.
- ☐ Parking cost is too high for me.
- ☐ Gas cost is too high for me.
- ☐ Traffic congestion is too bad in Shanghai.
- ☐ Other (_____ please specify)

6. Are you **happy with** the vehicle(s) you have now?
   - ☐ YES
   - ☐ NO
7. Please describe your **current life**?

   _____ Personal Income (RMB/month)

_____ Household Income (RMB/month)

_____ Job

_____ Type of house

_____ Living location (address)

_____ Leisure: What you usually do when you have free time?

PART 2: About Past

8. Did you have any vehicle? Please describe **last** (one) vehicle you have **purchased.** If you have one, **what** is that? If NOT, to **Q11.**

   ☐ Bicycle

   ☐ Auxiliary Power Vehicle (_____ Powered by?)

   ☐ Motorcycle (_____ cc.?)

   ☐ Car (_____ Year; _____ Make; _____ Model)

   ☐ Other type of vehicle (_____ please specify)

9. Please describe your **past life** during that period?

   _____ Personal Income (RMB/month)

   _____ Household Income (RMB/month)

   _____ Job

   _____ Living location

   _____ Leisure: What you usually do when you have free time?

10. To your best memory, please tell us **all vehicles you have purchased in sequence** by assigning numbers under the following boxes. (e.g., "1" = most past; "6" = most recent)

No vehicle But I take taxi, or rent vehicles	**Bicycle**	**Private Car**	**No vehicle** But I take public transportation
( )	( )	( )	( )
Motorcycle	**Auxiliary Power Vehicle**	**No vehicle** But I walk	**No vehicle** But I use company vehicle
( )	( )	( )	( )

<u>PART 3:</u> About Future

11. Do you have a ***dream life***? How is would it be different from you current life?

_____ Personal Income (RMB/month)

_____ Household Income (RMB/month)

_____ Job

_____ Type of house

_____ Living location

_____ Leisure: What you usually do when you have free time?

_____ Type of Vehicle

PART 4: About Yourself

12. What is your **age**? (_____)

13. What is your **gender**?

- ☐ Male
- ☐ Female

14. What is your **education level**?

- ☐ None
- ☐ Elementary school
- ☐ High school
- ☐ College
- ☐ Graduate School
- ☐ Other (_____ please specify)

15. Are you born in Shanghai?

- ☐ YES
- ☐ NO (from _____ province)

16. Do you have experience **living in other country**? If YES, _____ which country?

How long_____?

17. Please describe your **typical "week day" trip**:

From_____ to _____

Rang: _____ (km)

Trip Purpose: _____

18. **How** do you get there? (check **one** you use **most "frequently"**)

- ☐ Walk
- ☐ Public transportation (bus, Shanghai Metro, light rail, etc.)
- ☐ Taxi or rented vehicle
- ☐ I use company car or government vehicle
- ☐ I use my own (or family) car
- ☐ I use my own (or family) motorcycle
- ☐ I use my own (or family) auxiliary power vehicle
- ☐ I use my own (or family) bicycle
- ☐ Others (_____ please specify)

19. Please describe your **typical "weekend" trip.**

From_____ to _____

Rang: _____ (km)

Trip Purpose: _____

20. **How** do you get there? (check **one** you use **most "frequently"**)

- ☐ Walk
- ☐ Public transportation (bus, Shanghai Metro, light rail, etc.)
- ☐ Taxi or rented vehicle
- ☐ I use company car or government vehicle
- ☐ I use my own (or family) car
- ☐ I use my own (or family) motorcycle
- ☐ I use my own (or family) auxiliary power vehicle
- ☐ I use my own (or family) bicycle
- ☐ Others (_____ please specify)

END, thank you for taking the survey!!

- **Final Survey Questionnaire**

Dear Shanghai Resident,

Did you have experience in purchasing any vehicle? The Institute of Transportation Studies at the University of California, Davis (USA) working with Tongji University is conducting a survey on vehicle purchase behavior. Basically, we want to understand what factors affect your purchase decision on different type of vehicles.

The term "vehicle purchase" in this survey is specifically defined as: *vehicle purchased, owned/registered, and used by an individual*. In this study, there are four categories of "vehicle" – **private car, motorcycle, auxiliary power vehicle** (e.g., electric bicycle) and **bicycle** – together with a "**purchase nothing**" option, including: walking, taking public transportation or taxi, renting cars, and sharing company-owned vehicle.

You are eligible to participate this survey as long as you are **over 18** and have your **daily activities in Shanghai** metropolitan area. You are randomly chosen and there are two ways for you to participate:

I. Fill out the questionnaire on site	**$20 RMB or gift with equal value** (guaranteed)
II. Internet (**www.china.v33.org**)	**$2000 RMB** (draw) + Special prize for people refer 5+ friends.

If you choose to take the survey on-site, you can get **$20 RMB** right after the survey. You can also log on: **www.china.v33.org** - once we receive your completed survey, you will be automatically enrolled a draw for cash prize of **$2000 RMB**. We also draw people who refer 5+ friends for a special prize!

We will have the drawing every two weeks, your chances of wining can be high! To ensure your inclusion in the drawing, please complete the survey by **September 30th, 2006.**

Six parts are included in this survey: 1) your vehicle use/purchase background, 2) your attitude toward various means of travel, 3) your opinions about vehicle purchase, 4) your lifestyle, 5) information about yourself, and 6) your opinions about survey research. This survey should take **less than 20 minutes**.

All information you provide will be kept in secure database to ensure your privacy and confidentiality. You will not be identified in any reports or data bases. This research will be used **only for academic purpose**.

Thank you in advance for participating in this valuable study, your answers are very important for our research. You opinion will also affect future policy-making in Shanghai. Results from the survey will be put on the survey website (www.china.v33.org) after May 2007.

If you have any question, feel free to contact me (mni@ucdavis.edu); you can also contact Professor Ma Jun (majun@gmx.net) or Professor Chen XiaoHong (chenxh@mail.tongji.edu.cn) of Tongji University.

Sincerely,

JASON NI [倪孟正]
Ph.D. candidate,
University of California, Davis

Survey Date: ____/____/____ Survey Time: _____AM/PM

Survey Location: _____ Surveyor: _____ I.D.: _ _ _ _

PART 1: Vehicle Use/Purchase Background

Please check ONE answer for each question. (Unless it specifies "check all that apply")

1.

a.	Please give us the list of **ALL VEHICLES you currently own**. (check all that apply)
	☐ I DON'T OWN ANY VEHICLE (if check this box, Go to **Q2.**) ☐ Bicycle ☐ Auxiliary Power Vehicle (____electricity/_____LPG) ☐ Motorcycle ☐ Car (_____ when you bought it? _____ Make; ____New/____Used _____Domestic/_____Imported) ☐ Other (_____ please specify)
b.	Among all vehicle(s) you own, which one is the **MOST EXPENSIVE? (check one)**
	☐ Bicycle ☐ Auxiliary Power Vehicle ☐ Motorcycle ☐ Car ☐ Other (_____ please specify)
FOLLOWING QUESTIONS (c. ~ h.) ARE ABOUT THE "MOST EXPENSIVE" VEHICLE CHECKED IN b.	

c.	What is the purchase price of that vehicle?
	_____ RMB (vehicle) _____ RMB (license fee)
d.	Do you get any "subsidy" on the purchase/use of your vehicle? (check all that apply)
	☐ No, I don't get any subsidy ☐ Purchase ☐ Maintenance ☐ Fuel ☐ Parking ☐ Government Fee (e.g., toll) ☐ Other (_____ please specify)
e.	How often do you use it? _____ days in a week.
f.	For what purpose do you purchase it? (check all that apply)
	☐ School ☐ Work ☐ Entertainment, Recreation (e.g., movie, travel) ☐ Shopping ☐ Personal Business (e.g., see doctor, visit family members) ☐ Taking others (family/friends) where they need to go ☐ Other (_____ please specify)
g.	Do you own any **parking space AT WORK**?
	☐ YES ☐ NO (Normally it takes how long to find parking? _____hr; _____min)
h.	Do you own any **parking space AT HOME**?
	☐ YES ☐ NO (Normally it takes how long to find parking? _____hr; _____min)
IF YOU OWN "PERSONAL CAR", PLEASE ANSWER (i ~ l)	
i.	How many personal car(s) do you have? _____
j.	How far did you travel "yesterday"? Total _____ km

k.	Your monthly operating cost: (RMB/month)
	Fuel: _____ Maintenance and car fix: _____ Parking Fee (at home): _____ Parking Fee (not at home, e.g., office, shopping): _____ Road Fee (e.g., toll, road maintenance fee): _____ Insurance: _____ Other (_____ please specify)
l.	Your total mileage: _____ (10000 km) _____ km/year; _____ km/month

2. How far is your **WORKING PLACE** to the **"closest subway line"**?
   - ☐ Within 500 meters
   - ☐ 501 ~ 1000 meters
   - ☐ More than 1001 meters
   - ☐ N/A (I don't work, I don't know, etc)

3. How far is your **HOME** to the **"closest subway line"**?
   - ☐ Within 500 meters
   - ☐ 501 ~ 1000 meters
   - ☐ More than 1001 meters
   - ☐ N/A (I don't know, etc)

4. Which type of "**transportation means**" below do you **use *mostly*** during **week days**? (check one)
   - ☐ Walk
   - ☐ Public Transportation
   - ☐ Taxi
   - ☐ Rented car
   - ☐ Shared company car
   - ☐ Bicycle
   - ☐ Auxiliary Power Vehicle
   - ☐ Motorcycle
   - ☐ Car

☐ Other (_____ please specify)

5. What is your **purpose** for ***most*** *of your trips* on **week days**? (please skip if you seldom have trip during week days) (check one)
   - ☐ School
   - ☐ Work
   - ☐ Entertainment, Recreation (e.g., movie, travel)
   - ☐ Shopping
   - ☐ Personal Business (e.g., see doctor, visit family members)
   - ☐ Taking others (family/friends) where they need to go
   - ☐ Other (_____ please specify)

6. Which type of "**transportation means**" below do you **use *mostly*** during **weekend**? (check one)
   - ☐ Walk
   - ☐ Public Transportation
   - ☐ Taxi
   - ☐ Rented car
   - ☐ Shared company car
   - ☐ Bicycle
   - ☐ Auxiliary Power Vehicle
   - ☐ Motorcycle
   - ☐ Car
   - ☐ Other (_____ please specify)

7. What is your **purpose** for ***most*** *of your trips* on **weekend**? (please skip if you seldom have trip during week days)
   - ☐ School
   - ☐ Work
   - ☐ Entertainment, Recreation (e.g., movie, travel)
   - ☐ Shopping
   - ☐ Personal Business (e.g., see doctor, visit family members)
   - ☐ Taking others (family/friends) where they need to go
   - ☐ Other (_____ please specify)

8. Please **sort the sequence** of the travel means you have used from the very past (in

your memory) till now. Please fill in the numbers inside the boxes. (Skip the box if you never use that travel mean)

Ex: [1] Bicycle, [2] Motorcycle, [3] Car = Bicycle → Motorcycle → Car

☐ Bicycle
☐ Walk
☐ Personal Car
☐ Public Transportation
☐ Auxiliary Power Vehicle, Motorcycle
☐ Taxi, Rented car
☐ Shared Company's Car

☐ NO, there is no such "pathway". (WHY?_____)

PART 2: What do you think about various means of travel?

Please indicate *how well* each of the *following characteristics* describes various means of travel by checking the boxes. (From Strongly Disagree to Strongly Agree). Please do your best to respond to all nine means of travel, even if you are less familiar with some of them than others

I think the following travel mean is – a "SYMBOL OF SUCCESS"

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐

Taxi	☐	☐	☐	☐	☐
Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

I think the following travel mean is – "FAST"

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐
Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

I think the following travel mean is – "AVAIABLE WHEN NEEDED"

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐
Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

I think the following travel mean has – "MORE CARRYING CAPACITY"

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐

Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

I think the following travel mean is – "PRICY"

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐
Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

I think the following travel mean is – "COMFORTABLE"

	Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**

Walk	☐	☐	☐	☐	☐
Bicycle	☐	☐	☐	☐	☐
Auxiliary Power Vehicle	☐	☐	☐	☐	☐
Motorcycle	☐	☐	☐	☐	☐
Taxi	☐	☐	☐	☐	☐
Rented Car	☐	☐	☐	☐	☐
Public Transportation	☐	☐	☐	☐	☐
Company Car	☐	☐	☐	☐	☐
Private Car	☐	☐	☐	☐	☐

PART 3: Your Opinion about Vehicle Purchase

For each of the following statements, please check the response that best expresses your opinion.

		Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1.	The F-1 game and promotion in Shanghai affect my vehicle purchase decision.	☐	☐	☐	☐	☐
2	I consider buying a vehicle for infrequent need. (ex: visit rural family home, the travel once a year or airport pickup)	☐	☐	☐	☐	☐

3.	The subsidy is important for my vehicle purchase.	☐	☐	☐	☐	☐
4.	The loan is important for my vehicle purchase.	☐	☐	☐	☐	☐
5.	The government fees (e.g., license fee) affect my vehicle purchase decision	☐	☐	☐	☐	☐
6.	The new transportation law is what I need to consider before purchase vehicle.	☐	☐	☐	☐	☐
7.	Buying a (domestic-made) car shows a direct support to China's auto industry.	☐	☐	☐	☐	☐
8.	Owning a car is a pre-requisite for marriage.	☐	☐	☐	☐	☐
9.	The fuel price is what I care for my vehicle purchase.	☐	☐	☐	☐	☐
10.	Buying a house should be prior to vehicle purchase.	☐	☐	☐	☐	☐
11.	Transportation environment in Shanghai is dangerous for "pedestrians".	☐	☐	☐	☐	☐
12.	Transportation environment in Shanghai is dangerous for "drivers".	☐	☐	☐	☐	☐
13.	Saving is a virtue, and we should always not spend too much.	☐	☐	☐	☐	☐
14.	A vehicle which makes me look "better" will also bring me better career and social life.	☐	☐	☐	☐	☐
15.	Vehicle is just a business tool.	☐	☐	☐	☐	☐
16.	Vehicle (especially car) can depreciate very quickly.	☐	☐	☐	☐	☐

17.	In peak hours or rainy day, I wish I have personal vehicle.	☐	☐	☐	☐	☐
18.	Policy restrictions (e.g., ban of motorcycle entering CBD, ban of van using expressway) affect my vehicle purchase decision,	☐	☐	☐	☐	☐

PART 4: Your lifestyle as it related to vehicle purchase

In the following, there are some questions regarding your lifestyle as it related to vehicle purchase behavior. Please answer to your best.

		Strongly Disagree	**Disagree**	**Neutral**	**Agree**	**Strongly Agree**
1.	For me, a lot of the fun of having something nice is showing it off.	☐	☐	☐	☐	☐
2	The vehicle that I own needs to be well-known. (brand, model, etc.)	☐	☐	☐	☐	☐
3.	I enjoy catching everybody's eyes.	☐	☐	☐	☐	☐
4.	When many of my friends/colleagues own a certain type of vehicle, I will consider purchasing one of it.	☐	☐	☐	☐	☐
5.	I pay attention on fashion.	☐	☐	☐	☐	☐
6.	In the past (or now), I am crazy about some idol.	☐	☐	☐	☐	☐
7.	Life is so short; we should enjoy life as soon as we can.	☐	☐	☐	☐	☐
8.	I think about getting a loan for my vehicle purchase.	☐	☐	☐	☐	☐
9.	Saving money is hard for me.	☐	☐	☐	☐	☐

10.	Freedom and the control of life is what I pursue.	☐	☐	☐	☐	☐
11.	*"I am the king in my own territory"* is a good concept.	☐	☐	☐	☐	☐
12.	I don't like to share things with others.	☐	☐	☐	☐	☐
13.	I am aware of the information about new fuel or fuel efficient vehicles.	☐	☐	☐	☐	☐
14.	For environmental protection sake, I am willing to use walk/bike to take place of the motor vehicle use.	☐	☐	☐	☐	☐
15.	Eventually, everybody will pay the price for the pollution on earth.	☐	☐	☐	☐	☐
16.	Family is the most important in my daily life.	☐	☐	☐	☐	☐
17.	Going back home is such a joy after working a whole day.	☐	☐	☐	☐	☐
18.	Traveling with family is a happy thing,	☐	☐	☐	☐	☐

PART 5: About Yourself

1. What is your age?

   _____

2. Do you have driver's/rider's license?
   - ☐ No
   - ☐ Yes, Motorcycle only
   - ☐ Yes, Car only
   - ☐ Yes, both motorcycle and car

3. What is your gender?
   - ☐ Male
   - ☐ Female

4. What is your education level?
   - ☐ None
   - ☐ Elementary school
   - ☐ Junior High school
   - ☐ Senior High school
   - ☐ Some College (including community college)
   - ☐ College graduate
   - ☐ Graduate school
   - ☐ Other (_____ please specify)

5. What is your occupation?
   - ☐ None
   - ☐ Student
   - ☐ _____Teacher, or _____Professor
   - ☐ Factory Worker
   - ☐ Office Worker (_____state-owned enterprise, or _____ private-owned company, or _____foreign company]
   - ☐ Business Owner
   - ☐ Government
   - ☐ Doctor/Medical care personals
   - ☐ Show/Entertainment business
   - ☐ Transportation/Delivery/Driver
   - ☐ Other (_____ please specify)

6. How much is your **personal income** range (RMB per month, AFTER TAX)?
   - ☐ Below 500
   - ☐ 501~1000
   - ☐ 1001 ~ 2000
   - ☐ 2001 ~ 3000
   - ☐ 3001 ~ 4000
   - ☐ 4001 ~ 5000
   - ☐ 5000 ~ 10000

☐ Above 10001

7. How much is your **family income** range (RMB per month, AFTER TAX)?
   - ☐ Below 500
   - ☐ 501~1000
   - ☐ 1001~2000
   - ☐ 2001~3000
   - ☐ 3001~4000
   - ☐ 4001~5000
   - ☐ 5001~10000
   - ☐ 10001~20000
   - ☐ Above 20001

8. How many members are in your household? (live together, including you)

   _____

9. Which district in Shanghai do you currently live in?

□ Huangpu	□ Luwan	□ Xuhui	□ Changning
□ Jing'an	□ Putuo	□ Zhabei	□ Hongkou
□ Yangpu	□ Baoshan	□ Jiading	□ Qingpu
□ Songjiang	□ Minhang	□ Jinshan	□ Pudong
□ Nanhui	□ Fengxian	□ Chonming	

10. Are you a registered Shanghai resident?
    - ☐ YES
    - ☐ NO

11. How long do you live in Shanghai?

    _____ Years; _____ Months

12. Have you ever lived abroad?
    - ☐ YES, _____which country? _____for how long?
    - ☐ NO

PART 6: You opinion about our survey

This is the last part; we would like to thank you for your patience. However, we certainly want to understand your opinion toward participating survey research. Your comment will be valuable for us as well as for future researchers conducting survey in China.

	How important to you is each of the following affecting your motivation of taking a survey?	**Not at all important**	**Slightly Important**	**Neutral**	**Moderately Important**	**Extremely Important**
1.	How much time it will take	☐	☐	☐	☐	☐
2.	The topic (e.g., commercial *vs*. academic)	☐	☐	☐	☐	☐
3.	Authorization letter	☐	☐	☐	☐	☐
4.	My answer will be anonymous or kept confidential.	☐	☐	☐	☐	☐
5.	Guaranteed reward (non-cash)	☐	☐	☐	☐	☐
6.	Guaranteed reward (cash)	☐	☐	☐	☐	☐
7.	Drawing reward (non-cash)	☐	☐	☐	☐	☐
8.	Drawing reward (cash)	☐	☐	☐	☐	☐
9.	My friend refers me this survey.	☐	☐	☐	☐	☐

10.	Other, please specify_____	☐	☐	☐	☐	☐

11. How do you know about our survey? (please check all that apply)

- ☐ On-street
- ☐ Dealership
- ☐ Neighborhood Committee/Housing Property Management
- ☐ Cell phone message
- ☐ Email
- ☐ Colleagues/Friends/Classmates/Family members
- ☐ Other, please specify_____

END, thank you for taking the survey!!

SPECIAL PRIZE!!

- Please refer at least one person you know to participate this interesting survey! If you refer 5 or more people (including 5, by providing their cell phone number OR Email address), you will automatically join a drawing for our special prize!

☐ Yes, I will like to refer my friends:

Person 1:_____

Person 2:_____

Person 3:_____

Person 4:_____

Person 5:_____

(Please add your friends' cell phone numbers in the following blank, if you want to

refer more than 5 people)

☐ No, thanks

- **Distribution of Motorization Pathway Patterns**

Motorization Pathway (331 patterns, sorted by counts)	**Counts**	**Percentage of 992 cases** (The shaded cells cover more than 50% sample)	**Follow Hypothetical Motorization Direction?**	**Number of Motorization Stages**
WBPT	64	6.45%	Yes	4
WPT	33	3.33%	Yes	3
WBPTC	32	3.23%	Yes	5
WBP	30	3.02%	Yes	3
B	25	2.52%	N/A	1
WB	25	2.52%	Yes	2
BM	22	2.22%	Yes	2
WBM	20	2.02%	Yes	3
WBPM	18	1.81%	Yes	4
WBPTM	19	1.92%	No	5
BPT	17	1.71%	Yes	3
WP	17	1.71%	Yes	2
BC	15	1.51%	Yes	2
WBPMTSC*	**15**	**1.51%**	**Yes**	**7**
WBPC	12	1.21%	Yes	4
WBPMT	12	1.21%	Yes	5
WBPTS	12	1.21%	Yes	5
WPBT	11	1.11%	No	4
WPTC	11	1.11%	Yes	4
BMC	10	1.01%	Yes	3
BWPT	10	1.01%	No	4
WBPMTC	10	1.01%	Yes	6
BWP	9	0.91%	No	3
WBMPT	8	0.81%	No	5

* This is exactly the hypothetical motorization pathway pattern.

WPTS	8	0.81%	Yes	4
W	7	0.71%	N/A	1
WBMC	7	0.71%	Yes	4
WBMPTS	7	0.71%	No	6
BP	6	0.60%	Yes	2
WBMP	6	0.60%	No	4
WBPMC	6	0.60%	Yes	5
WBPTMSC	6	0.60%	No	7
WM	6	0.60%	Yes	2
WPB	6	0.60%	No	3
WPBM	6	0.60%	No	4
BCM	5	0.50%	No	3
BPW	5	0.50%	No	3
BWPTS	5	0.50%	No	5
P	5	0.50%	N/A	1
PWT	5	0.50%	No	3
WBMPTSC	5	0.50%	No	7
WBPTMC	5	0.50%	No	6
WPBTS	5	0.50%	No	5
WTC	5	0.50%	Yes	3
BPC	4	0.40%	Yes	3
BTC	4	0.40%	Yes	3
BW	4	0.40%	No	2
BWPM	4	0.40%	No	4
CT	4	0.40%	No	2
M	4	0.40%	N/A	1
MB	4	0.40%	No	2
PTC	4	0.40%	Yes	3
PTW	4	0.40%	No	3
WBC	4	0.40%	Yes	3
WBMTC	4	0.40%	Yes	5
WBPMTS	4	0.40%	Yes	6
WPBMT	4	0.40%	No	5
WPC	4	0.40%	Yes	3

WPS	4	0.40%	Yes	3
WPTB	4	0.40%	No	4
BMPC	3	0.30%	No	4
BMPT	3	0.30%	No	4
BMS	3	0.30%	Yes	3
BMT	3	0.30%	Yes	3
BPM	3	0.30%	Yes	3
BPTC	3	0.30%	Yes	4
BPTS	3	0.30%	Yes	4
BT	3	0.30%	Yes	2
C	3	0.30%	N/A	1
MP	3	0.30%	No	2
MW	3	0.30%	No	2
WBMPTC	3	0.30%	No	6
WBMT	3	0.30%	Yes	4
WBMTPS	3	0.30%	No	6
WBPMSTC	3	0.30%	No	7
WBPMTCS	3	0.30%	No	7
WBPST	3	0.30%	No	5
WBPTMS	3	0.30%	No	6
WPBMTS	3	0.30%	No	6
WT	3	0.30%	Yes	2
BM(CS)	2	0.20%	No	3
BMP	2	0.20%	No	3
BPMC	2	0.20%	Yes	4
BPMT	2	0.20%	Yes	4
BPMTSC	2	0.20%	Yes	6
BPTW	2	0.20%	No	4
BPWT	2	0.20%	No	4
BSC	2	0.20%	Yes	3
BST	2	0.20%	No	3
BTSC	2	0.20%	Yes	4
BWC	2	0.20%	No	3
CMW	2	0.20%	No	3

CTW	2	0.20%	No	3
CWT	2	0.20%	No	3
MBW	2	0.20%	No	3
MTB	2	0.20%	No	3
PB	2	0.20%	No	2
PBT	2	0.20%	No	3
PBWT	2	0.20%	No	4
PMT	2	0.20%	Yes	3
PTB	2	0.20%	No	3
WB(PM)T(CS)	2	0.20%	No	5
WBMTP	2	0.20%	No	5
WBMTPCS	2	0.20%	No	7
WBP(MT)	2	0.20%	No	4
WBPCT	2	0.20%	No	5
WBPS	2	0.20%	Yes	4
WBPTSC	2	0.20%	Yes	6
WBSC	2	0.20%	Yes	4
WBT	2	0.20%	Yes	3
WBTC	2	0.20%	Yes	4
WBTPC	2	0.20%	No	5
WBTSMPC	2	0.20%	No	7
WC	2	0.20%	Yes	2
WMC	2	0.20%	Yes	3
WMT	2	0.20%	Yes	3
WPBMC	2	0.20%	No	5
WPBMSTC	2	0.20%	No	7
WPBS	2	0.20%	No	4
WPBTCSM	2	0.20%	No	7
WPM	2	0.20%	Yes	3
WPMT	2	0.20%	Yes	4
WPST	2	0.20%	No	4
WPTM	2	0.20%	No	4
(BC)TP	1	0.10%	No	3
(BM)T	1	0.10%	No	2

(BW)(PM)T	1	0.10%	No	3
(BW)(PMT)(CS)	1	0.10%	No	3
(BW)C(PT)	1	0.10%	No	3
(BW)MSC	1	0.10%	No	4
(BW)PCT	1	0.10%	No	4
(BW)PTS	1	0.10%	No	4
(BWCP)	1	0.10%	No	1
(BWPTS)	1	0.10%	No	1
(WM)BPT	1	0.10%	No	4
(WM)BPTC	1	0.10%	No	5
(WP)BT	1	0.10%	No	3
(WP)BTCMS	1	0.10%	No	6
(WP)T	1	0.10%	No	2
B(PT)C	1	0.10%	No	3
B(WC)(PM)(TS)	1	0.10%	No	4
BCP	1	0.10%	No	3
BCPT	1	0.10%	No	4
BCW	1	0.10%	No	3
BMCS	1	0.10%	No	4
BMCTPW	1	0.10%	No	6
BMPTS	1	0.10%	No	5
BMPW	1	0.10%	No	4
BMSC	1	0.10%	Yes	4
BMST	1	0.10%	No	4
BMTC	1	0.10%	Yes	4
BMTPC	1	0.10%	No	5
BMTS	1	0.10%	Yes	4
BMWPT	1	0.10%	No	5
BMWS	1	0.10%	No	4
BMWT	1	0.10%	No	4
BP(CT)	1	0.10%	No	3
BPMTS	1	0.10%	Yes	5
BPMW	1	0.10%	No	4
BPMWTSC	1	0.10%	No	7

BPTCS	1	0.10%	No	5
BPTMS	1	0.10%	No	5
BPTSC	1	0.10%	Yes	5
BPWMTCS	1	0.10%	No	7
BSW	1	0.10%	No	3
BTM	1	0.10%	No	3
BTMC	1	0.10%	No	4
BTP	1	0.10%	No	3
BTPC	1	0.10%	No	4
BTS	1	0.10%	Yes	3
BTSW	1	0.10%	No	4
BWCPMTS	1	0.10%	No	7
BWM	1	0.10%	No	3
BWMPCT	1	0.10%	No	6
BWMPT	1	0.10%	No	5
BWMPTS	1	0.10%	No	6
BWMTS	1	0.10%	No	5
BWPC	1	0.10%	No	4
BWPT(CMS)	1	0.10%	No	5
BWPTC	1	0.10%	No	5
BWPTSM	1	0.10%	No	6
BWT	1	0.10%	No	3
BWTC	1	0.10%	No	4
BWTSP	1	0.10%	No	5
CBW	1	0.10%	No	3
CM	1	0.10%	No	2
CMPTW	1	0.10%	No	5
CMTW	1	0.10%	No	4
CPBW	1	0.10%	No	4
CPBWTMS	1	0.10%	No	7
CPMW	1	0.10%	No	4
CPSTMBW	1	0.10%	No	7
CPT	1	0.10%	No	3
CPWTB	1	0.10%	No	5

CSW	1	0.10%	No	3
CTPWB	1	0.10%	No	5
CTPWBMS	1	0.10%	No	7
CTSP	1	0.10%	No	4
CTSPMBW	1	0.10%	No	7
CTWP	1	0.10%	No	4
CWBT	1	0.10%	No	4
CWPT	1	0.10%	No	4
CWPTSMB	1	0.10%	No	7
MBPW	1	0.10%	No	4
MBWP	1	0.10%	No	4
MC	1	0.10%	Yes	2
MPBTWS	1	0.10%	No	6
MPBW	1	0.10%	No	4
MPTW	1	0.10%	No	4
MS	1	0.10%	Yes	2
MSW	1	0.10%	No	3
MT	1	0.10%	Yes	2
MTP	1	0.10%	No	3
MTSPW	1	0.10%	No	5
MTW	1	0.10%	No	3
MTW(CS)B	1	0.10%	No	5
MWBPSTC	1	0.10%	No	7
MWP	1	0.10%	No	3
MWPBT	1	0.10%	No	5
MWPCTBS	1	0.10%	No	7
MWTPSB	1	0.10%	No	6
PBC	1	0.10%	No	3
PBCTWS	1	0.10%	No	6
PBM	1	0.10%	No	3
PBMTC	1	0.10%	No	5
PBS	1	0.10%	No	3
PBWCMST	1	0.10%	No	7
PBWM	1	0.10%	No	4

PBWTSM	1	0.10%	No	6
PC	1	0.10%	Yes	2
PCWBMST	1	0.10%	No	7
PCWTSCM	1	0.10%	No	7
PMBW	1	0.10%	No	4
PMWT	1	0.10%	No	4
PSC	1	0.10%	Yes	3
PST	1	0.10%	No	3
PT	1	0.10%	Yes	2
PTBW	1	0.10%	No	4
PTMSW	1	0.10%	No	5
PTS	1	0.10%	Yes	3
PTSC	1	0.10%	Yes	4
PTWMSCB	1	0.10%	No	7
PWB	1	0.10%	No	3
PWTSC	1	0.10%	No	5
PWTSCB	1	0.10%	No	6
SCPTW	1	0.10%	No	5
STCW	1	0.10%	No	4
STPW	1	0.10%	No	4
SW	1	0.10%	No	2
TBP	1	0.10%	No	3
TC	1	0.10%	Yes	2
TCBW	1	0.10%	No	4
TMBWCP	1	0.10%	No	6
TPB	1	0.10%	No	3
TPC	1	0.10%	No	3
TPS	1	0.10%	No	3
TSW	1	0.10%	No	3
TWC	1	0.10%	No	3
W(BP)(MT)CS	1	0.10%	No	5
W(BP)C	1	0.10%	No	3
W(BPS)(MT)C	1	0.10%	No	4
WB(PM)	1	0.10%	No	3

WB(PMT)(CS)	1	0.10%	No	4
WB(PT)S	1	0.10%	No	4
WB(PT)SC	1	0.10%	No	5
WBCMSPT	1	0.10%	No	7
WBCPT	1	0.10%	No	5
WBCPTSM	1	0.10%	No	7
WBCSPMT	1	0.10%	No	7
WBCTSMP	1	0.10%	No	7
WBM(PT)C	1	0.10%	No	5
WBM(PT)S	1	0.10%	No	5
WBMPC	1	0.10%	No	5
WBMPS	1	0.10%	No	5
WBMPSC	1	0.10%	No	6
WBMPST	1	0.10%	No	6
WBMPTCS	1	0.10%	No	7
WBMSC	1	0.10%	Yes	5
WBMSPT	1	0.10%	No	6
WBMTPC	1	0.10%	No	6
WBMTSC	1	0.10%	Yes	6
WBMTSPC	1	0.10%	No	7
WBPCTM	1	0.10%	No	6
WBPCTMS	1	0.10%	No	7
WBPCTSM	1	0.10%	No	7
WBPMCT	1	0.10%	No	6
WBPSC	1	0.10%	Yes	5
WBPSTC	1	0.10%	No	6
WBPSTMC	1	0.10%	No	7
WBPT(CMS)	1	0.10%	No	5
WBPTCS	1	0.10%	No	6
WBPTMCS	1	0.10%	No	7
WBPTSCM	1	0.10%	No	7
WBST	1	0.10%	No	4
WBTCPMS	1	0.10%	No	7
WBTMPS	1	0.10%	No	6

WBTMPSC	1	0.10%	No	7
WBTP	1	0.10%	No	4
WBTS	1	0.10%	Yes	4
WBTSP	1	0.10%	No	5
WBTSPMC	1	0.10%	No	7
WCBPM	1	0.10%	No	5
WCBPTSM	1	0.10%	No	7
WCPBTMS	1	0.10%	No	7
WCPBTS	1	0.10%	No	6
WCPTS	1	0.10%	No	5
WMBP	1	0.10%	No	4
WMBPT	1	0.10%	No	5
WMP	1	0.10%	No	3
WMPB	1	0.10%	No	4
WMPBTSC	1	0.10%	No	7
WMTC	1	0.10%	Yes	4
WPBC	1	0.10%	No	4
WPBMTC	1	0.10%	No	6
WPBMTCS	1	0.10%	No	7
WPBSC	1	0.10%	No	5
WPBTC	1	0.10%	No	5
WPBTCS	1	0.10%	No	6
WPBTMSC	1	0.10%	No	7
WPMTSBC	1	0.10%	No	7
WPTBC	1	0.10%	No	5
WPTBM	1	0.10%	No	5
WPTBMSC	1	0.10%	No	7
WPTSBMC	1	0.10%	No	7
WPTSC	1	0.10%	Yes	5
WPTSMB	1	0.10%	No	6
WS	1	0.10%	Yes	2
WSM	1	0.10%	No	3
WTBC	1	0.10%	No	4
WTBMPC	1	0.10%	No	6

WTPSCMB	1	0.10%	No	7