Better Lucky Than Rich? Welfare Analysis of Automobile License Allocations in Beijing and Shanghai*

Shanjun Li†

January 2014

Abstract

Economists often favor market-based mechanisms (e.g., auction) over non-market based mechanisms (e.g., lottery) for allocating scarce public resources on grounds of economic efficiency and revenue generation. When the usage of the resources in question generates negative externalities and the amount of externalities is positively correlated with willingness-to-pay (WTP), the efficiency comparison can become ambiguous. Both types of mechanisms are being used in China's major cities to distribute limited vehicle licenses as a measure to combat worsening traffic congestion and urban pollution. While Beijing employs non-transferable lotteries, Shanghai uses an auction system. This study empirically quantifies the welfare consequences of the two systems by estimating a random coefficients discrete choice model of vehicle demand to recover consumers' WTP for a license. Rather than relying on the maintained exogeneity assumption on product attributes in the literature, we employ a novel strategy by taking advantage of a control group as well as information from household surveys to identify structural parameters. Our analysis finds that although the lottery system in Beijing has a large advantage in reducing externalities from automobile use than a uniform price auction, the advantage is offset by the significant welfare loss from misallocation. The lottery system foregone nearly 36 billion RMB (or \$6 billion) in social welfare in 2012 and the auction would have generated 21 billion RMB to Beijing municipal government, more than covering all the subsidies to the local public transit system.

Keywords: Auction, Lottery, Random Coefficients Utility Model, Resource Allocation

---

*I thank Antonio Bento, Sebastien Houde, Sarah Jacobson, Matthew Kahn, Neha Khanna, Joshua Linn, Yimin Liu, Stuart Rosenthal, Catherine Wolfram, Jun Yang and seminar participants at Binghamton University and 2013 Southern Economics Association Annual Meeting for comments and suggestions. Financial support from the Lehman fund at Cornell University is acknowledged. All errors are my own.

†Assistant Professor, Dyson School of Applied Economics and Management, Cornell University, 405 Warren Hall, Ithaca, NY 14853. Email: SL2448@cornell.edu.

1 Introduction

Market-based mechanisms (e.g., auction) have often been advocated for allocating scarce public resources on grounds of economic efficiency and revenue generation, as opposed to non-market based mechanisms (e.g., lottery). Both have been used widely in practice, often for different types of resources. For example, auctions are routinely used to sell mineral rights, timber, and radio spectrum while lotteries are employed in charter school admissions, graduate student offices, and jury duty.[1] Market-based mechanisms have the potential to achieve efficiency by using price signals to distribute the scarce resource to those with the highest willingness-to-pay (WTP), whereas lotteries with no-transferability can lead to mis-allocation and unrealized welfare gain. However, in the presence of negative externalities from the usage of the resources, the welfare comparison between the two mechanisms may become ambiguous when the amount of externalities is positively correlated with WTP. In this case, the social benefit (consumer welfare net of externalities), the basis for measuring social welfare, diverges from and may not be monotonically increasing with WTP, the basis for resource allocation under the market-based mechanisms.

The resource in question here is a license or permit to purchase a vehicle. Major cities in China have been experiencing the world's worst traffic congestion and air pollution as a result of rapid economic growth as well as vehicle ownership outpacing transport infrastructure and environmental regulation.[2] New automobile sales in China grew from 2.4 million units in 2001 to 19.30 million in 2012, surpassing the U.S. to become the largest market in 2009. Beijing, the second largest city in China and the poster child of traffic congestion and pollution, has the fastest growth in vehicle ownership, with the number of passenger vehicles increasing from one million in 2000 to nearly five million in 2012. As a result, Beijing has been routinely ranked as one of the cities with worst traffic conditions in the world, with the average traffic speed during peak hours below 15 miles per hour. The daily concentration of PM2.5, a key measure of air quality, frequently reaches over 10 or 20 times of the daily health limit of 20 micrograms per cube meter recommended by the World Health Organization (WHO). Automobiles are the largest source of PM2.5, accounting for 22 percent in the whole city

[1] The Federal Communications Commission used both administrative process and lotteries to allocate radio spectrum in the early years, both of which led to wide discontent due to long delays. The FCC finally adopted auctions in 1994 after many decades of persistent arguments by economists since Coase (1959). The first auction yielded ten times of revenue predicted by the Congressional Budget Office (Cramton 1995).

[2] Zheng and Kahn (2013) offer a comprehensive review on China's urban pollution challenges and government policies to deal with them.

and about one third in the urban core.[3]

To ease gridlock and improve air quality, Beijing municipal government introduced a quota system for new vehicle licenses in order to limit the growth of vehicle ownership. About 20,000 licenses are distributed each month through lotteries since 2011. The winners are determined from random drawings among the registered participants. The licenses are not transferable. A license is needed for first-time vehicle buyers, and those who purchase an old vehicle, accept a gifted vehicle, or transfer out-of-state registration to Beijing. Vehicle owners who scrap the used vehicle can transfer the old license to the new vehicle and do not need a new license. Winning the lottery has become more and more difficult over time: the odds decreased from 1:10 in January 2011 to 1:84 in August 2013.

Shanghai, the largest city in China, also has a quota system to restrict the number of vehicles but its license allocation is through an auction instead of lotteries. In fact, an auction system has been in place since 1986 but the goal of reducing vehicle ownership did not emerge until much recently. The current system started in 2008 and auctions are held monthly to distribute about 10,000 licenses. Each auction lasts 90 minutes in an online system and it is a multi-unit, discriminatory (pay as you bid) and dynamic auction where up to two bid revisions are allowed for each bidder during the last 30 minutes. The auction revenue is designated for improving public transit and transport infrastructure. In 2012, the auction system generated over 6.7 billion RMB to Shanghai municipal government. The average bid for a license reached over 92,000 RMB in March 2013, higher than the price of many entry-level vehicle models.

The main objective of this study is to empirically quantify the welfare consequences of the two allocation mechanisms in distributing vehicle licenses taking into account both allocation efficiency and externalities associated with vehicle usage. This is an important question for at least two reasons. First, traffic congestion and air pollution impose major costs to the society, especially in emerging markets (Parry et al. 2007). Creutzig and He (2008) estimate that the external costs from automobile usage in Beijing amount to over 7.5% of its GDP. Under the endorsement of China's Ministry of Environmental Protection, many large cities in China plan to adopt the license quota system and use one of the two allocation mechanisms. However, the impacts and welfare consequences of these policies are unknown.

[3]The second largest source of PM2.5 is coal burning (17 percent) followed by construction site dusts (16 percent) in 2012 according to Beijing Environmental Protection Bureau.

Second, there is theoretical ambiguity *a priori* in terms of the welfare comparison between the two mechanisms as we alluded to above because the usage of license (buying and ultimately driving a vehicle) is associated with negative externalities such as congestion and pollution. The negative externalities are very likely to be positively correlated with consumers' WTP for a license in that consumers with high WPT tend to have high income and are likely to use less fuel-efficient vehicles and drive more. Whether the lottery or the auction system is more efficient depends critically on the level of heterogeneity in consumers' WTP and its relationship with the externalities from vehicle usage. Therefore, the efficiency comparison between the two mechanisms and the magnitude of welfare impacts are ultimately empirical questions. Recovering consumer's WTP is the key to addressing these questions.

Since licenses are allocated through lotteries and trade is not allowed in Beijing, we do not observe consumers' WTP. In Shanghai, licenses are auctioned in a non-standard format and the bids may not reflect value as shown in an experimental study by Liao and Holt (2013). To recover consumers' WTP for a license, we estimate a random coefficients discrete choice model of vehicle demand that takes into account consumer preference heterogeneity and unobserved product attributes developed by Berry, Levinsohn, and Pakes (1995) (henceforth BLP). Similar to Petrin (2002), our estimation strategy employs both aggregate market-level data and information from household surveys to form moment conditions. From household surveys on new vehicle buyers, we obtain the share of buyers among different income groups. We used them to form (micro-) moment conditions and they are critical in identifying consumer preference heterogeneity. A key departure in identification strategy from this literature is that we do not rely on the maintained exogeneity assumption that unobserved product attributes are uncorrelated with observed product attributes. Rather, we employ the common trend assumption used in the difference-in-differences (DID) framework for impact evaluation. This is possible because our market-level data include data for four cities: Beijing, Nanjing, Shanghai, and Tianjin. Nanjing and Tianjin are two large cities next to Shanghai and Tianjin, respectively. They did not have the license quota system during the data period and as we show through graphs and regressions, the automobile markets in these two cities exhibit similar trends to Beijing and Shanghai in the absence of the policies.

Four key findings emerge from our analysis. First, both policies in Beijing and Shanghai significantly limited new vehicles sales: the lottery system in Beijing reduced over one million new vehicle sales from 2011 to 2012 while the Shanghai auction reduced about 1.4 million new vehicles from 2008 to 2012. These reductions are substantial and a reflection of the stringency of the quota system. Second, there is significant heterogeneity in WTP for licenses among consumers and across the two cities. Among those with positive WTP in Beijing in 2012, the average WTP is about 53,200 RMB with a standard deviation of about 83,500 RMB. The average WTP in Shanghai is lower than Beijing at 32,000 RMB with a standard deviation of 52,400 RMB. The difference could be partly attributed to the better public transit in Shanghai that directly benefited from auction revenue for many years. Third, the lottery system performs poorly relative to a uniform price auction even after taking externalities into consideration. While the lottery system in Beijing would lead to a non-trivial reduction in total external costs of 7 billion RMB relative to a uniform price auction in 2012, it resulted in nearly 43 billion RMB foregone in consumer surplus due to misallocation. These two results imply a welfare loss of nearly 36 billion RMB from the lottery system. In addition, the uniform price auction would have generated 21 billion revenue to Beijing municipal government in 2012, more than enough to cover all its subsidies to the local public transit system. Fourth, based on a range of plausible assumptions, the optimal level of quota in Beijing is lower than the existing level and further reducing the quota would increase net social welfare.

This study contributes to the literature in the following four dimensions. First, although theoretical literature on allocation mechanisms are abundant, there are very few empirical studies on quantifying and comparing welfare consequences from different allocation mechanisms. Glaeser and Luttmer (2003) study the rationing in the housing market under rent control in New York city and find significant misallocation of houses. Our study goes a step future to quantify the welfare loss from misallocation by exploring a rare opportunity where a market-based mechanism and a non-market mechanism are both used for the same type of resources.

Second, as we discussed above, in the presence of externalities, the efficiency comparison between the lottery and auction systems could be ambiguous. Our empirical analysis highlights this important point and showcases the advantage of the lottery system in reducing externalities. Nevertheless, the gain in consumer surplus from allocation through auctions dominates this (non-trivial) advantage of the lottery system in our context, mainly due to significant consumer heterogeneity in WTP.

Third, our analysis adds to the literature on environmental and transportation policies, especially the small literature on related policies in China. China is by far the largest emitter of greenhouse gases, accounting for nearly 30 percent of world emissions in 2012. It is the largest energy consumer and is also now the top importer of crude oil in the world. Therefore, these policies are not only important for China itself but also have significant implications for the world at large. However, our understanding of the impacts and effectiveness of these policies is very limited.[4]

Lastly, our study offers a novel identification strategy in structural demand estimation by employing the common trend assumption used in DID analysis. This alleviates the need to rely on the maintained assumption which is often deemed strong since vehicles attributes are likely to be jointly determined in the design process. Our empirical strategy to recover consumers' WTP for licenses combines both the structural demand model and the DID method. The structural demand model allows consumer heterogeneity that is critical for the welfare analysis while the DID framework provides us an alternative method to identify structural parameters while at the same time offering a benchmark for examining validity of the predictions from the structural model.

Before we proceed, it is perhaps helpful to further clarify the scope of our welfare analysis. The analysis focuses exclusively on allocation efficiency and external costs from automobile usage under the two allocation mechanisms. Automobiles are an important mode of transportation and the allocation of vehicle licenses are likely to have impacts on household location, job decision and schooling choices, all of which could have important welfare implications in the long term. Although part of these welfare impacts such as the desire to live in a large apartment outside of the city center and hence the need to have a vehicle are captured by the WTP that we estimate, we are not investigating the policy impacts in these dimensions. In addition, vehicle license quota systems and the choice of allocation mechanisms could affect housing prices, urban development, and carbon emissions (Glaeser and Kahn 2010). These are interesting and potentially important margins for analysis but would necessitate different data and methodology, hence we leave them for future research.

The remainder of the paper proceeds as follows. Section 2 offers an intuitive discussion on how the existence of externalities affects the welfare comparison between the lottery and auction systems. Sector 3 describes the background of the study, policies and data used. Section 4 lays out the empirical model. The identification strategy and estimation method are presented in Section 5. In Section 6, we present estimation results. Section 7 conducts

[4]Various federal and local policies are being adopted in China to address urban congestion and air pollution including corporate fuel economy standards, emission standards, gasoline tax, driving restrictions and license quotas. Chen et al. (2011) study the environmental impacts of measures such as driving restriction and plant closure by the Beijing government to clear up the air before the 2008 Olympic Games. Xiao and Heng (2013) evaluate and compare the effects of favorable consumption tax treatment for small vehicles and gasoline tax.

welfare analysis and Section 8 concludes.

2 Allocation Mechanisms and Externalities

The purpose of this section is to offer an illustration on the welfare comparison between lottery and auction systems in the presence of negative externalities. Although we believe that the insight may be applicable in other contexts, we use a vehicle license quota system as the context of discussion.

The notion that the price mechanism can achieve efficient resource allocation is as old as modern economics. In allocating scarce public resources, governments or resource managers have relied on both market-based mechanisms and non-market based mechanisms often for different types of goods. Market-based mechanisms such as a well-designed auction can allocate the resource to those with the highest value and hence achieve efficiency while non-market based mechanisms such as an administration process or lottery do not. It is argued that non-market based mechanisms such as a lottery are chosen often out of concern of fairness or for political convenience (Taylor et al. 2003).

However, when the usage of resources generates market failures such as externalities, market-based mechanisms such as an auction may not yield efficient allocation because private value and social value diverge.[5] More importantly, within the context of allocating vehicle licenses, we illustrate below that lotteries need not be less efficient than auctions *a priori*. The usage of vehicle licenses and ultimately automobile use lead to a range of negative externalities, the most important of which are traffic congestion and air pollution in urban areas. In fact, the purpose of vehicle license quota systems used in Beijing and Shanghai is to deal with these two issues.[6]

Consider the following environment: (1) there are $\underline{Q}$ licenses to be allocated; (2) there

[5]The usage of licenses by firms may have implications on market power. This has been discussed in the context of spectrum auctions where different collection of winners may subsequently lead to different market structure (Cramton 2002).

[6]A vehicle quota system itself is not the first-best instrument for internalizing externalities from automobile use which leads to several externalities such as congestion, air pollution, traffic accidents and noise. In the first-best world, multiple instruments should be used to correct for these externalities. For example, real-time congestion pricing is the first-best instrument to deal with congestion externality while fuel taxes is the first-best instrument to internalize the externality from CO_2 emissions and other pollutants that are proportional to fuel use. A comparison between the quota system with other potentially more efficient policy instruments such as congestion pricing and fuel taxes is an interesting and important question. It will necessitate additional information and modeling efforts such as with respect to driving decisions and we leave it for future research.

are $N (N > \underline{Q})$ agents, each demanding at most one license; (3) each agent i has a private value (or WTP) V_i and the value is drawn from a known i.i.d. distribution with a support of $[0, V_0]$; (4) the usage of the license imposes an external cost of E_i, which is increasing in V_i. The fourth assumption is a key departure from a standard model of resource allocation. As we document below, consumers with high WTP for a license tend to have higher income. On average, they drive larger and less fuel-efficient vehicles and they drive more relative to those with a low WTP. So the usage of licenses by those a higher WTP is likely to generate larger external costs. Our estimates below show that neither the magnitude of external costs nor the difference in external costs across households is trivial. That is, the discussion below does have empirical relevance.

We compare two allocation mechanisms: a lottery where all agents can participate and have an equal chance of winning; and a uniform price auction where the $\underline{Q}$ highest bidders each gets one license and pays a price equal to the highest rejected bid. Harris and Raviv (1981) show that each agent bids her value in the uniform price auction and therefore the licenses will be allocated to the $\underline{Q}$ agents who have the highest value. We use the uniform price auction as a convenient way for exposition and the point is not lost with other types of auctions such as discriminatory auctions that can achieve the same allocation outcome.

In the top panel of Figure 1, the line V_0V depicts the (smoothed) WTP schedule and the line E_0E_1 depicts the external cost for the agents with the corresponding WTP. The external cost is an increasing function of the WTP and the social benefit of one unit of license is given by V_i-E_i. The quota is denoted by $\underline{Q}$. Under the lottery system, the licenses would go to a random set of $\underline{Q}$ agents and the total (expected) social benefit from the allocation is given by W_l=$(V_0E_0V^*\text{-}V^*E_1Q_1)\frac{O\underline{Q}}{OQ_1}$. Under the auction system, the licenses go to the agents with the WTP higher than V and the total social benefit is given by W_a=V_0E_0EV. Since V_i-E_i is increasing in V_i in this scenario, it is evident that the social benefit from the auction is higher: $W_a > (V_0E_0V^*)\frac{O\underline{Q}}{OQ^*} > (V_0E_0V^*)\frac{O\underline{Q}}{OQ_1} > W_l$.

The opposite is true in the bottom panel of Figure 1 where the external cost rises faster than the WTP so that the social benefit V_i-E_i is negative for those agents with the highest WTP to have the license. We make E_1 to be zero to simplify the exposition but it need not be the case. The total social benefit from the lottery is W_l=$(V\text{'}E_1Q_1\text{-}V_0E_0V\text{'})\frac{O\underline{Q}}{OQ_1}$ while that from the auction is W_a=$V\text{'}VE\text{-}V_0E_0V\text{'} < 0$. The lottery system leads to a higher social benefit than the auction in this case since W_l= $(V\text{'}VE\text{-}V_0E_0V\text{'}+ VEE_1Q_1)\frac{O\underline{Q}}{OQ_1} = W_a\frac{O\underline{Q}}{OQ_1} + VEE_1Q_1\frac{O\underline{Q}}{OQ_1} > W_a$. This is a scenario where agents with the highest private value generates the smallest social value, therefore, the auction will allocate the licenses to the wrong hands from the efficiency perspective.

Which of the two scenarios is playing out in reality depends on the WTP schedule as well as its relationship with the external cost curve. Although one can get a good sense on how household income relates to the fuel economy of the vehicles purchased, the driving pattern from household level data and hence the external costs, the WTP schedule for a license is not observed. The empirical goal is to develop a method to estimate the WTP schedule and conduct welfare comparison taking into account the external costs.

3 Policy and Data Description

In this section, we first describe the background for the quota system on new vehicles. We then discuss the details of lottery and auction policies in Beijing and Shanghai. Various data components for our analysis will be presented thereafter.

3.1 Background

During the past three decades, China has embarked on an extraordinary journey of economic growth with its GDP growing at about 10 percent a year: per capita GDP increased from less than $200 in 1980 to $6000 in 2012 in nominal terms according to the World Bank. As household income grows, luxury good consumption such as automobiles started to pick up dramatically around the turn of the century. The annual sales of new passenger vehicles increased from 2.4 million units in 2001 to over 19 million units in 2012 as depicted in Figure 2, surpassing the U.S. to become the largest market in 2009. Large cities in China are ahead of the curve relative to the nation in both economic growth and vehicle ownership. Beijing, the capital and the second largest city with about seven million households, has gone from a city on bikes to a city on cars during this period: the stock of passenger vehicles increased from 1.6 million units to nearly 5 million units as shown in Figure 2.[7]

The rapid growth in vehicle ownership leads to serious traffic congestion, despite significant efforts in expanding roads and public transit systems and a whole range of other traffic management policies such as driving restrictions and reducing public transit fares Yang et

[7]Among the 5 million vehicles, a little over 4 million are owned by households. The household vehicle ownership rate is 58 percent in Beijing, comparing to 46 percent in New York city based on 2010 U.S. Census.

al. (2013). While there were mostly bicycle traffic in 2001 in Beijing, the city is now often ranked routinely as one of the worst cities in traffic conditions.[8] The traffic speed on arterial roads within the 5th ring road during morning peak hours (7:00-9:00) on work days averaged 14.7 miles/hour (MPH) in 2011, reduced from 22.8 MPH in 2005. The average speed was 13.4 MPH during afternoon peak hours (17:00-19:00) in 2011, compared with 20.2MPH in 2005.[9] During the same period, air quality has worsened dramatically and the air quality index is frequently above the hazardous level defined by U.S. EPA which recommends that everyone should avoid all outdoor physical activities. The average daily concentration of PM2.5, a key measure of air quality frequently reaches over 250 micrograms per cube meter, more than 10 times the recommended daily level of 20 by WHO.[10]

Shanghai, the financial center and the largest city in China with 8.5 million households, has better traffic conditions and air quality. The traffic speed on arterial roads averages 21.2MPH and 22.3MPH during morning and evening peak, respectively. As shown in the bottom panel of Figure 2, the total number of vehicles in Shanghai is less than half of that in Beijing, despite having more households and higher average household income. This is largely due to the vehicle purchase restrictions that was put in place before vehicle ownership took off as we discuss below. Public transit plays a bigger role in Shanghai than in Beijing: in 2011, it accounted for 49% of all travels (not including walking) in Shanghai relative to 42% in Beijing. Partly due to fewer vehicles and better traffic conditions, the air quality in Shanghai is consistently better than Beijing. In a ranking of air quality among 28 major cities on the east coast, Beijng ranked at the bottom with an annual average PM10 concentration of 121 micrograms per cubic meter while Shanghai ranked at 8th with an average PM10 concentration of 79 in 2010, which is still significantly higher than the WHO recommended annual concentration limit of 20 micrograms per cubic meter.[11]

[8]IBMs Global Commuter Pain Survey in 2011 ranks Beijng as the second worst cities in commuting behind Mexico city: http://www.smartplanet.com/blog/transportation/10-worst-cities-for-commuting-2011-edition/892. This article based on TomTom's congestion index ranks Beijing the third worst traffic city in the world after Moscow and San Paulo: http://www.roadcrazed.com/top-5-worst-traffic-cities.

[9]Source: Beijing Annual Transportation Report 2006, 2012.

[10]The city may have seen the worst air pollution so far in January 2013. During the 24-hour period from 10am on January 12th, 18 of the 24 hourly PM2.5 reading in U.S. Embassy were beyond 500 micrograms per cube meter. The recorded was set at 866 micrograms per cube meter during the reading at 8pm on the 12th.

[11]Systematic monitoring of PM2.5 by the Chinese government only began in recent years. PM10 data are from 2010 Statistics of Ambient Air Quality of Environmental Protection Priority Cities, Ministry of Environmental Protection.

3.2 Policy Description

To address traffic congestion and air pollution, Beijing municipal government announced the policy of imposing quota on new vehicle licenses, the right to purchase a vehicle, on December 23, 2010. Monthly lotteries have been used to allocate about 20,000 licenses each month, starting from January of 2011. A license is needed for first-time buyers, and those who purchase an old vehicle, accept a gifted vehicle, transfer out-of-state registration to Beijing. Vehicle owners who scrap the used vehicle can transfer the old license to the new vehicle and do not need a new license. The eligible participants include Beijing residents and non-residents who have been paying income tax for at least five years in Beijing.[12] The licenses are assigned to winners through random drawings. The winners can then use the license to register their vehicles. Transferring a license from a winner to other people is prohibited: the winner and the vehicle owner must be the same person.[13]

Among all the licenses allocated, about 88% (or 17,600 each month) are for private vehicles and the rest are for institutions. The winners are determined in two different pools for these two categories. While the private licenses are allocated monthly, the institutional licenses are done bi-monthly. The first lottery was held on January 26th, 2011 and 17,600 private licenses were allocated among 187,420 participants, with a winning odds of roughly 1:10. In August, 2013, the winning odds reduced to 1:84.[14] The top panel of Figure 3 shows the monthly licenses allocated on the bottom and the monthly new vehicle sales on the top, the data source of which will be discussed in the next section. The dramatic decrease in vehicles sales since the start of the policy in January 2011 reflects the stringency of the policy relative to the demand for new vehicles. The difference between vehicle sales and the number of licenses allocated is due to the sales that do not need a license (vehicle replacement after scrappage). The winners have six months to purchase a new vehicle with the licenses before they become expired. Once expired, the license recycles back and is added to the system for

[12] A Beijing resident is someone with Beijing huko, something akin to a U.S. green card, and having a place to live in Beijing does not suffice.

[13] Although there are anecdotal evidence that some transferring occurred by having vehicle registered under the winner but paid and used by another person, this is unlikely to be widespread because the legal owner (the winner) not only has the liabilities in paying annual registration fee, traffic fines and emission inspections, but also is liable for damages and injuries in accidents.

[14] With the shrinking odds of winning the license, there is increasing discontent with the system. According to a survey of 800 residents in 2013 by Beijing Statistical Bureau, nearly 70 percent of the survey participants agree that the system needs to be improved. Among them, 42.2 percent would like to abolish the system and 41.5 percent would prefer a hybrid system of lottery and auction. 7.7 percent would like to switch to auctions.

distribution in future lotteries. The winners who allow their licenses to expire will not be allowed to participate in the lottery again within the next three years. Although there are some households who decide to register a vehicle outside of Beijing due to increased difficulty of winning the lottery, the households still have to get a permit to be able to drive these out-of-state vehicles in Beijing. More importantly, these vehicles are banned from entering the 5th ring road (within which the vast majority of business and population are located) during rush hours.

Shanghai is the first city to implement a vehicle license quota system in China and it auctioned its first license in 1986. Although the market for private vehicles was very small at that time, traffic congestion was a big problem due to insufficient road infrastructure. In fact, Shanghai experimented vehicle driving restrictions even before 1995. The auction system has evolved over time. Initially, it was a sealed-bid auction where reservation prices and quota levels varied across vehicles produced in Shanghai, non-Shanghai produced vehicles, and imports.[15] In 2003, a unified auction system without a reservation price was put in place for domestic vehicles and imports.

The current online system started in 2008 and it is one of the allocation mechanisms that our study focuses on. This auction format, henceforth Shanghai auction, can be characterized as a multi-unit, discriminatory (pay as you bid), and dynamic auction (Liao and Holt 2013). The auction is held monthly during a 90-minute period and bidders observe the current lowest accepted bid prior to submitting a bid. In the first hour, bidders can submit a single initial bid and in the last 30 minutes, each bidder can revised their bids up to two times. The revised bid however has to be within a window of 300 RMB below and above the current lowest accepted bid. The purpose of the bid revision period and the restriction on bid revisions is to reduce price volatility. The bottom panel in Figure 3 shows the average and lowest accepted bid in each month from 2008 to 2012. The average bid price increased from 23,370 to 69,346 RMB during this period. The plot shows that the average and the lowest winning bids are very close: the difference is usually less than 500 RMB or one to two percent of the average bid.[16] The winners are required to purchase a new vehicle within three

[15]In 1998, Shanghai government set the reservation price of 20,000 RMB for vehicles produced in Shanghai while the reservation price was set at 100,000 RMB for vehicles produced elsewhere. In protest, Hubei province which has a large automobile industry initiated an additional charge of 70,000 RMB for those who purchase brands produced in Shanghai. The trade war ended in 2000 when Shanghai removed the reservation price in auctions for domestically produced vehicles.

[16]During the first month of the new auction system in January 2008, the lowest price was 8,100 RMB compared with the average price of 23,370 RMB. This anomaly was due to a computer glitch. The abrupt change in price in December 2010 was due to the speculation that the government was about to phase out

months before the license gets expired. The vehicle and the license cannot be transferred within the first year of registration. Similar to Beijing, vehicle registered outside of Shanghai are not allowed to use the major roads during rush hours. Nevertheless, there is anecdotal evidence that some households choose to registered their vehicles in neighboring provinces due to high license prices.

3.3 Data Description

Our analysis focuses on policies in Beijing and Shanghai and we bring two nearby cities, Nanjing and Tianjin into analysis to serve as our control group. The assumption of common unobservables used in DID analysis will provide one set of moment conditions for our structural model as we discuss below. Nanjing is about 300km away from Shanghai and is the capital city of Jiangsu province that shares boarder with Shanghai. Tianjing is about 150km from Beijing. The characteristics of these four cities are shown in Table 1. Shanghai and Beijing are the two largest cities in China in population. Tianjin is the sixth and Nanjing is the 11th.[17] Among the four cities, Shanghai has the highest average household income while Tianjing has the lowest. Table 1 shows that Beijing has the smallest increase in average nominal income of 42% during the 5-year period while Shanghai has the largest increase of 57%, with the inflation being 13.7% during this period.

We rely on four main data sets together with a variety of auxiliary data for our analysis. The first main data set contains monthly vehicle sales by model (vintage-nameplate) in each of the four cities from 2008 to 2012. There are 21,228 observations with 1,769 distinct models. Figure 4 plots monthly sales (in log) in each city and displays two important features. First, sales in all four cities grew over time and tracked each other well before 2011 and the trend is more consistent across Beijing, Nanjing and Tianjin, reflecting the fact that Shanghai has an auction policy in place through the data period. Second, there was strong seasonality which is to a large extent due to holidays. For example, the sales in December 2010 went up dramatically in all cities but then dropped significantly in February 2011. This is due to the fact that Chinese New Year, the most important Chinese holiday, was in February 3rd in 2011.[18] Third, the sales increase in December 2011 appeared to be stronger in Beijing than

the policy.

[17]Beijing, Shanghai and Tianjin are three of the four municipalities or province-level cities. They are at the same level of administrative subdivision as provinces and are right below the central government.

[18]During the holiday season that starts from at least one week before the Chinese New Year and ends two weeks after, people are occupied with visiting families and friends. Big-item purchases such as buying a car

in other cities. This is due to the anticipation and more importantly the fact that the quota policy was announced in December 23, 2010. Very little if any discussion on the policy was made public before the announcement. However, once the policy was announced, many who planned to buy a vehicle in the next few months moved their purchase forward into the last week of December. In the main specification of our analysis, we remove the last two months in 2010 and the first two months in 2011 in Beijing to deal with the changes in purchase timing induced by the policy.

The second data set contains vehicle characteristics of each model in the sales data. These characteristics include price, fuel economy measured by liters per 100 miles, vehicle size (length by width), engine size, vehicle type (passenger car, SUV, or minivan), and vehicle segment (mini, small, medium, upper medium, large and luxury). The summary statistics are presented in Table 2. Vehicle prices are computed based on the Manufacturer Suggested Retail Prices (MSRPs) and the sales tax. MSRPs are set by manufacturers and are generally constant across locations and within a model year. There could be potential pitfalls in using MSRPs when they are different from the transaction prices due to promotions. Different from the promotion-heavy environment in the U.S., China's auto market has infrequent promotions from manufactures or dealers and retail prices are often very close to or the same as MSRPs. The sales tax is normally set at 10% but was reduced to 5 percent and 7.5 percent for vehicles with engine displacement no more than 1.6 liter in 2009 and 2010, respectively. The average price of a vehicle is over 300,000 RMB and the medium price is over 190,000, both significantly higher than the average household income in all four cities.[19]

The third data set is income distributions in each city in each year. China's National Bureau of Statistics conducts census every 10 years but the income distribution data at the household level are not publicly available. Instead, we obtain the average income by income quantiles in each city in each year from the statistical yearbooks of each city. We construct household income distribution based on these aggregate information together with Chinese Household Income Survey (2002), a national representative survey, conducted by researchers at the University of Michigan. The purpose of the survey was to measure and estimate income distribution in China. We adjust the income in the household income survey (14,971

tend to occur before the holiday season.

[19]MSRPs already include two types of taxes: consumption tax which ranges from one percent for vehicles with small engines to 40 percent for large engines, and value-added tax of 17 percent of before-tax price. These high taxes partly contribute to the high vehicle prices in China relative to those in the United States. Li et al. (2013) examine various factors in the price changes over time and also offer a detailed discussion on the industry.

observations) proportionally and separately for each of the quantiles so that the interpolated income distributions in a given year are consistent with the annual income statistics from the yearbooks.

The fourth set of data contains aggregate information from an annual national representative household survey among new vehicle buyers from Ford Motor Company. We were provided the shares of households by four income groups among new vehicles buyers in each of the four cities in each year. Table 3 presents these shares along with the shares by income group among all households, which are constructed based on the income distributions from each city in each year. The table shows that high income groups account for a disproportionally large share of vehicle buyers. For example, while the highest income group (annual household income over 144,000) accounts for 3.85% of all households in 2012 in Beijing, this group accounts for more than 20% among new vehicle buyers. These aggregate information will be used to form additional moment conditions that are crucial to identify consumer preference parameters especially those related to vehicle prices.

4 Empirical Model of Vehicle Demand

In order to compare the welfare outcomes of the two policies, we need to recover consumers' WTP for a license. We do not observe license prices in Beijing since the licenses are allocated through lottery for free. In Shanghai, although we observe average bids, they are unlikely to represent consumers' WTP. Liao and Holt (2013) use experiments to study the relationship between bid and WTP under the Shanghai auction and its welfare comparison with other formats such as a first-price auction. They show that initial bids in the first stage of the auction tend to be much lower than those in a first-price auction which themselves are lower than WTP. Bids are revised up in the second stage but they are still lower than WTP and the difference is larger among high WTP bidders.

Our strategy to recover consumers' WTP for a license is to estmate consumer surplus from buying a new vehicle. We set up and estimate a demand system that is obtained by aggregating over the discrete choices of individual buyers. In this section, we first specify the utility function, the basis of individual choices. We then discuss the aggregation process to obtain the market demand.

4.1 Utility Function Specification

Let $m = \{1, 2, 3, 4\}$ denote a market (i.e., Beijing, Nanjing, Shanghai, Tianjin) and a year-month by t from January 2008 to December 2012. Let i denote a household and $j \in \mathcal{J}$ denote a model or product (i.e., vintage-nameplate) where $\mathcal{J}$ is the choice set. Household i's utility from product j is a function of household demographics and product characteristics. A household chooses one product from a total of J models of new vehicles and an outside alternative in a given month. The outside alternative captures the decision of not purchasing any new vehicle in the current month. The indirect utility of household i from product j in market m at time t is defined as

$$u_{mtij} = \bar{u}(\mathrm{p}_j, \mathrm{b}_{mti}, \mathrm{X}_j, \xi_{mtj}, \mathrm{y}_{mti}, \mathrm{Z}_{mti}) + \epsilon_{mtij}, \tag{1}$$

where the first term on the right, $\bar{u}()$, denotes the deterministic component of the utility as a function of vehicle attributes and consumer characteristics. p_j is the price of product j and it does not vary across markets and months within a year as we discussed in the data section. b_{mti} is the price paid by consumer i (i.e., the bid) after winning the auction and this applies only in Shanghai (hence zero in other cities). As shown above, the winning bids in the discriminatory multi-unit auction have very small spread. Given that we do not observe the distribution of the bids, we use the average bid as the price paid by all the winners. The effect of this measurement error on our results should be small since the difference between the average winning bid and the lowest winning bid is generally less than two percent of the average bid.

X_j is a vector of observed product attributes (other than price) including a constant term, vehicle size, engine size and fuel cost per 100 kilometers.[20] ξ_{mtj} includes unobserved product attributes such as product quality or prestige and as well as unobserved demand shocks to be specified below. y_{mti} is the income of household i and Z_{mti} is a vector of (unobserved) household demographics. ϵ_{mtij} is an i.i.d. random taste shock and is assumed to follow a Type-I extreme value distribution. The utility from the outside good is defined as $\mathrm{u}_{mti0} = \epsilon_{mti0}$, where ϵ_{mti0} also follows a Type-I extreme value distribution.

[20] Unlike most commodities in China, Gasoline prices are still set by the National Development and Reform Commission (NDRC) and there is minimal price variation across regions.

Following the literature, we specify the first part of u_{mtij}, the deterministic utility to be:

$$\bar{u}_{mtij} = \alpha_{mti} ln(\mathrm{p}_j + \mathrm{b}_{mti}) + \sum_{k=1}^{K} \mathrm{X}_{mtjk} \tilde{\beta}_{mtik} + \xi_j. \quad (2)$$

α_i measures consumer i's preference or distaste for price and it is defined as:

$$\alpha_{mti} = \alpha_0 + \alpha_1 ln\mathrm{y}_{mti} + \sigma \nu_{mti},$$

where α_i will be negatively related to income if α_1 is negative. One would expect high income households to be less sensitive to price due to diminishing marginal utility of income. α_i is also affected by unobserved household attributes captured by ν_{mti}. We assume that ν_{mti} has a standard normal distribution in the benchmark specification and σ is the standard deviation of a normal distribution.

A note on the functional form of consumer preference on price is in order. The literature on vehicle demand has used different specifications for the price and income interaction. For example, BLP and Petrin (2002) use ln(y_i-p_j) and the term has an natural explanation as the utility from the composite goods. As discussed above, the median price of vehicles in China are higher than the average income of most households, hence this specification does not work well in our context. One might argue that we should use the current payment on the vehicle rather than the price in the utility function. In China, most buyers make full cash payment on their purchases. Goldberg (1995) specifies the price and income interactions as α_i(y_i-p_j) where α_i varies across income categories. She argues that one can view the price and income variables to be proxies for vehicle capital cost and the lifetime wealth of the household, respectively. Berry et al. (1999) specify $\alpha_i p_j$ where α_i is inversely related to income. In our context, these specifications do not lead to the intuitive pattern of price elasticity where more expensive products have less elastic demand. In order to generate that pattern, the increase in household income among the buyers have to be faster than the price increase if we compare two products with different prices. We choose the current specification to allow income to affect consumer preference on price in a more flexible manner. The price variables should be viewed as relative to the outside good (which has a price of one). Therefore, our utility specification is homogenous of degree one in prices. In obtaining consumer surplus and welfare analysis, we rely on the price variable rather than the income variable because of the difficulty in directly interpreting the income variable in our context.[21]

[21]Household income can be a rather imprecise proxy of wealth in these cities where housing value can

X_{mtjk} is the kth attribute of product j. $\tilde{\beta}_{ik}$ is the random taste parameter of household i over product attribute k. It is a function of unobserved household demographics captured by ν_{ik}, which is assume to have a standard normal distribution.

$$\tilde{\beta}_{mtik} = \bar{\beta}_k + \sigma_k^u \nu_{mtik}. \tag{3}$$

The preference parameters defined above underscore consumer heterogeneity that our model tries to capture. The heterogeneity will translate into heterogeneity in consumers' WTP for a new vehicle, which is crucial for our welfare analysis. With all the components defined above, the utility function can be fully written out as the following:

$$u_{mtij} = (\alpha_0 + \alpha_1 ln \mathrm{y}_{mti} + \sigma \nu_{mti}) ln(\mathrm{p}_j + \mathrm{b}_{mti}) + \sum_{k=1}^{K} \mathrm{X}_{mtjk}(\bar{\beta}_k + \sigma_k^u \nu_{mtik}) + \xi_{mtj} + \epsilon_{mtij}. \tag{4}$$

4.2 Choice Probabilities and Aggregate Demand

Based on the i.i.d. Type-I extreme value distribution of ϵ_{mtij} and ϵ_{mti0}, the choice probability of household i for product j without any quantity constraint is

$$\mathrm{Pr}_{mtij}(\mathrm{p}_j, \mathrm{b}_{mti}, \mathrm{X}_j, \xi_{mtj}, \mathrm{y}_{mti}, \mathrm{Z}_{mti}) = \frac{\exp(\bar{u}_{mtij})}{1 + \sum_h [\exp(\bar{u}_{mtih})]}, \tag{5}$$

where b_{mti}=0. This equation can be used directly to generated aggregate sales in the market when there is no quantity constraint such as in Nanjing and Tianjin as well as in Beijing before the lottery was put in place in 2011. Denote N_{mt} as the number of households in the market and the market share of project j in market m in time t is

$$\mathrm{S}_{mtj} = \frac{1}{\mathrm{N}_{mt}} \sum_i \mathrm{Pr}_{mtij}. \tag{6}$$

In the case of quantity constraint, the aggregation needs to take into account the allocation mechanisms. There are two types of households: those who need to acquire a new license before purchasing a vehicle and those who do not. Under both lottery and auction policies,

account for the majority of the wealth. Housing values have increased several folds during the past 10 years in these cities. Many residents inherited housing from their parents and many others especially those who work at government agencies were given housing for a cost much lower than the market price. These people do not necessarily have as high income as those who purchase their houses from the market. One can treat the income variable as another variable (such as education) that does not have a monetary unit.

households who scrap a used vehicle can use the old license when buying a new vehicle. We do not have household level data and therefore do not observe the type of households in this regard. Instead, we explicitly model the type probabilities as a function of vehicle ownership rate in Beijing and Shanghai. Denote L_{mt} as the probability that a household in city m and time t would need to obtain a new license before buying a vehicle. We parameterize L_{mt} as a logistic function of vehicle ownership rate o_{mt}:

$$\mathrm{L}_{mt} = \frac{exp(\gamma_0 + \mathrm{o}_{mt} * \gamma_1)}{1 + exp(\gamma_0 + \mathrm{o}_{mt} * \gamma_1)}. \tag{7}$$

A positive coefficient γ_1 would imply that as vehicle ownership rate rises, the probably of a household needing a new license will decrease.

In Beijing, the households who need a new license must obtain the license through the lottery system. Denote the odds of obtaining a lottery in Beijing in a given month from January 2011 (t $>$ 36) as ρ. Denote c_{mti} as a random draw from a Bernoulli distribution with probability of L_{mt} being 1. The market share of product j in Beijing (m=1) is:

$$\mathrm{S}_{1tj[t>36]} = \frac{1}{\mathrm{N}_{mt}} \sum_i [\mathrm{Pr}_{1tij} * \mathbb{1}(\mathrm{c}_i = 1) * \rho + \mathrm{Pr}_{1tij} * \mathbb{1}(\mathrm{c}_i = 0)], \tag{8}$$

where $\mathbb{1}(.)$ is the indicator function. Pr_{1tij} is defined by equation (5) and b_{mti}=0. The market share of product j in Shanghai (m=3) is defined as:

$$\mathrm{S}_{3tj} = \frac{1}{\mathrm{N}_{mt}} \sum_i [\mathrm{Pr}_{3tij}(\mathrm{b}_{mti} > 0) * \mathbb{1}(\mathrm{c}_i = 1) + \mathrm{Pr}_{3tij}(\mathrm{b}_{mti} = 0) * \mathbb{1}(\mathrm{c}_i = 0)]. \tag{9}$$

5 Identification and Estimation

5.1 Constructing Moment Conditions

Our goal is to recover the preference parameters in equation (4) in order to estimate consumers' WTP for a license. The identification challenge comes from the fact that there are unobserved product attributes as well as demand shocks ξ_{mtj} in the utility function. The unobserved product attributes such as product quality are likely to be correlated with prices. Previous studies show that ignoring these unobserved product attributes biases the price coefficient toward zero and leads to wrong welfare calculations. This challenge in fact motivated the methodology in BLP. An additional issue in our context is that unobserved demand shocks in Shanghai are likely to be correlated with average bids and hence render them endogenous. Ignoring this can also bias the price coefficient toward zero.

To facilitate the discussion on identification and estimation below, notice that the utility function in equation (4) contains terms that vary by households as well as terms that do not. We separate these two categories and rewrite the utility function into the following:

$$u_{mtij} = \delta_{mtj} + \mu_{mtij} + \epsilon_{mtij}, \tag{10}$$

where δ_{mtj} is the household-invariant utility or the mean utility of product j in market m in time t. Based on equation (4), it is specified as follows

$$\begin{aligned} \delta_{mtj}(\theta_1) &= X_j\bar{\beta} + \xi_{mtj} \\ &= X_j\bar{\beta} + \xi_j + \eta_t + \mathbb{1}(m=3)\eta_t' + \zeta_{ms} + \kappa_m yr_t + e_{mtj} \\ &= \delta_j + \eta_t + \mathbb{1}(m=3)\eta_t' + \zeta_{ms} + \kappa_m yr_t + e_{mtj}, \end{aligned} \tag{11}$$

where we write ξ_{mtj} into several terms in the second line. ξ_j is unobserved product attributes. η_t captures time (year-month) fixed effects that control for common demand shocks and seasonalities across cities. $\mathbb{1}(m=3)\eta_t'$ captures Shanghai-specific time effects. ζ_{ms} is city-specific preferences for different vehicle segments where s is an index for segments. yr_t is year (1 to 5) and $\kappa_m yr_t$ captures city-specific time trend. e_{mtj} is time-varying and city-specific demand shocks. The last line combines $X_j\bar{\beta} + \xi_j$ into product dummies δ_j, absorbing the utility that is constant for all households across the markets. The parameters in the mean utility function is denoted as $\theta_1 = \{\delta_j, \eta_t, \eta_t', \zeta_{ms}, \kappa_m\}$.

The second part in equation (10), μ_{mtij}, is household-specific utility defined as:

$$\mu_{mtij}(\theta_2) = [\alpha_0 + \alpha_1 ln\mathrm{y}_{mti} + \sigma\nu_{mti}]ln(\mathrm{p}_{mtj} + \mathrm{b}_{mti}) + \sum_k x_{mjk}\nu_{mtik}\sigma_k^u. \tag{12}$$

The parameters in the household-specific utility are denoted as $\theta_2 = \{\alpha_0, \alpha_1, \sigma, \sigma^u\}$.

With this specification, we can rewrite the choice probabilities in equation (5) as following:

$$\mathrm{Pr}_{mtij}(\mathrm{p}_j, \mathrm{b}_{mti}, \mathrm{X}_j, \xi_{mtj}, \mathrm{y}_{mti}, \mathrm{Z}_{mti}) = \frac{\exp[\delta_{mtj}(\theta_1) + \mu_{mtij}(\theta_2)]}{1 + \sum_h \{\exp[\delta_{mth}(\theta_1) + \mu_{mtih}(\theta_2)]\}}. \tag{13}$$

The market shares can be written as $\mathrm{S}_{mtj}(\delta_{mtj}, \theta_2, \theta_3)$, where $\theta_3 = \{\gamma_0, \gamma_1, \rho\}$ which characterizes the license allocation processes described above.

In the choice probabilities, unobserved product attributes and demand shocks are absorbed in δ_{mtj} while the price and bid variables are in μ_{mtij}. If we could include market-time-product fixed effects subsuming δ_{mtj}, we can control for unobserved product attributes and demand shocks. However, this is impractical in this nonlinear model. BLP develop a methodology to back out δ_{mtj}. Under mild regularity conditions, for given vectors of θ_2 and θ_3, a unique vector of δ_{mt} for each market that equalizes the predicted market shares with observed market shares can be recovered through a contraction mapping algorithm:

$$\delta_{mt}^{n+1} = \delta_{mt}^{n} + ln(S_{mt}^{o}) - ln[\hat{S}(\delta_{mt}^{n}, \theta_2, \theta_3)], \tag{14}$$

where n is the number of iteration. S^o is a vector of observed market shares while $S()$ is predicted market shares. With the recovered δ_{mt} for given vectors of θ_2 and θ_3, θ_1 can be estimated using a linear framework following equation (11).

To estimate the model, we follow BLP by using a simulated GMM with the nested contraction mapping. The GMM is based on three sets of moment conditions. The first set is formed based on the city-year specific demand shocks in equation (11). The identification assumption is that these demand shocks are mean independent of city-year dummy variables, i.e., having zero mean at the city-year level:

$$\mathrm{E}\Big[e_{mtj}(\theta_2, \theta_3)|d_{mt}\Big] = 0, \tag{15}$$

where d_{mt} are city-year dummies.This assumption amounts to that time-varying demand shocks have a common trend across cities and the common trend is controlled by time fixed effects. What is left from the time trend e_{mtj} is not systemically different across cities. Note that we have also included city-segment fixed effects and these control for difference in levels in demand shocks. Since Beijing implemented the lottery in 2011 and 2012, this assumption implies that the lottery policy is exogenous to the time-varying demand-shocks in Beijng.

This common trend assumption (in the absence of the policy) is motivated by the graphical evidence in Figure 4 and it is a key assumption needed in the DID analysis for policy evaluation. Although one cannot test this assumption directly given that we do not observe the counterfactual of no policy for the treatment group (Beijing in our case), we have three years of data before the policy and we can examine if the pre-policy time trends are the same across the cities in a reduced-form framework. If they are the same before the policy, we would be more comfortable with the assumption (Heckman and Hotz 1989; Meyer 1995).

Since Shanghai implements an auction system throughout our data period, we do not have a pre-policy period for comparison. To allow for the possibility of different time trend between Shanghai and other cities, we include Shanghai-specific time effects in equation (11) in the benchmark specification as a conservative measure. Recall we have city-segment dummies (which swaps city fixed effects) and time fixed effects in the equation. This leaves us eight exclusion restrictions in the first set including city-year dummy variables for Beijing and Nanjing from 2002 to 2005 (Tianjin is the base group and year 2001 is the base year). In one of alternative specifications, we do not include Shanghai-specific time effects and assume common trend in all four cities (leaving us 12 exclusion restrictions) and we obtain very similar results.

The second set of moment conditions is constructed based on the aggregate information from the household survey presented in the right panel of Table 3. We match the predicted shares of households by income group by city among new vehicle buyers to those in the table. We use the fourth group as the base group and this gives us 12 moment conditions (four cities each with three income groups):

$$\mathrm{E}_t\left[\widetilde{\mathrm{S}}_{mgt|\mathrm{buyers}}(\theta_2,\theta_3) - \mathrm{S}_{mgt|\mathrm{buyers}}\right] = 0, \tag{16}$$

where g is a income group and $\widetilde{\mathrm{S}}_{mgt|\mathrm{buyers}}$ is the predicted share of income group g among vehicle buyers while $\mathrm{S}_{mgt|\mathrm{buyers}}$ is the observed counterpart. The former is calculated as:

$$\widetilde{\mathrm{S}}_{mgt|\mathrm{buyers}}(\theta_2,\theta_3) = \frac{\sum_{i=1}^{N_{mt}} d(\mathrm{y}_{mti} \in \mathrm{INC}_g)\sum_{j=1}^{J}\mathrm{Pr}_{mtij}}{\sum_{i=1}^{N_{mt}}\sum_{j=1}^{J}\mathrm{Pr}_{mtij}} \tag{17}$$

where d( ) is an indicator function being 1 for household i whose income (y_{mti}) falls into the income range of group g, INC_g.

The third set of moment conditions matches the predicted quantity of licenses to the observed quota in each month.

$$\mathrm{E}_t(\widetilde{\mathrm{Q}}_{mt}(\theta_2,\theta_3) - \mathrm{Q}_{mt}) = 0, \tag{18}$$

where $\widetilde{\mathrm{Q}}_{mt}$ is predicted quantity of licenses and it is calculated for Beijing (m=1 and t >36) and Shanghai (m=3) as the following

$$\begin{aligned} \widetilde{\mathrm{Q}}_{1t} &= \sum_i \sum_{j=1}^{J} [\mathrm{Pr}_{1tij} * \mathbb{1}(\mathrm{c}_i = 1) * \rho], \\ \widetilde{\mathrm{Q}}_{3t} &= \sum_i \sum_{j=1}^{J} [\mathrm{Pr}_{3tij}(\mathrm{b}_{mti} > 0) * \mathbb{1}(\mathrm{c}_i = 1)], \end{aligned} \tag{19}$$

where $\mathbb{1}(.)$ is the indicator function and the definitions of c_i are random draws from a Bernoulli distribution as discussed in Section 4.2. There could be a time gap between winning a license and purchasing a vehicle. In Beijing, winners have six months to purchase a new vehicle while in Shanghai, winners have three months before the license expires. Many consumers indeed take their time to purchase their vehicles. In the estimation, Q_{mt} is not the quota observed in that particular month; rather it is the average of the last six months and three months for Beijing and Shanghai, respectively.

We form the objective function by stacking these three sets of moment conditions. The procedure involves iteratively updating θ_2 and θ_3 and then δ_{mj} from the inner loop of contraction mapping to minimize the objective function. The estimation starts with an initial weighted matrix to obtain consistent initial estimates of the parameters and optimal weighting matrix. The model is then re-estimated using the new weighting matrix.

5.2 Further Discussions on Identification and Computation

Although our model follows closely with the BLP literature, our identification strategy represents an important departure. The maintained identification assumption is that unobserved product attributes are mean independent of those observed ones and the exclusion restrictions are given by the product attributes of other products within the firm and outside the firm. This assumption could be violated if firms choose product attributes (observed and unobserved) jointly (Klier and Linn 2012). We do not rely on this assumption. Instead, our first set of moment conditions (or macro-moments) are based on the assumption that unobserved demand shocks have a common trend across the cities, a critical assumption in DID analysis. We are able to offer some evidence to support this assumption.

It is worth mentioning that our identification strategy is also made possible by the fact that different households are paying different effective prices (price plus bid) for the same vehicle in Shanghai depending on whether they need a new license or not. This allows us to have all the price variables in the household-specific utility and be isolated from unobserved products and demand shocks. To understand this, imagine if we do not have data for Shanghai, we would have $\alpha_0 ln(p_j)$ entering the mean-utility term. We would then need to estimate α_0 for welfare analysis. Since the price variable and the unobserved product attributes would both appear in the mean utility, one would need to evoke some type of exogeneity assumption such as the one on unobserved product attributes maintained in the literature to deal with price endogeneity. Alternatively, one can assume away $\alpha_0 ln(p_j)$ from the utility specification so that the price variable is always interacted with household demographics such as income and hence appear in the household specific utility alone as in Berry et al. (1999) and Beresteanu and Li (2011). This could be a restrictive functional form and our estimation results do not support this form in our context.

Before concluding this section, we offer a few additional details for estimation. First, the estimation starts from generating a set of households in each year-month and in each market. Each of the households is defined by a vector of household demographics including income from the income distribution and unobserved household attributes from the standard normal. When generating the random draws, we use randomized Halton sequences to improve efficiency. Our results below are all based on 150 households in each year-month and market. For the benchmark specification, we tried 200 random draws but that made very little difference. Second, we define the market size to be half of the number of households in the city in a given year in the benchmark specification. To check the sensitivity of the result to this definition, we estimate a model where the market size is the total number of households as has been traditionally done in the literature for the U.S. market.[22] The results do not differ in any significant way as we will show below. Third, the convergence criterion for the simulated GMM (outer loop) is 10e-8 while that for the contraction mapping (inner loop) is set up to 10e-14. The convergence criterion for the contraction mapping starts from 10e-10 and increases as the search goes on in order to expedite the estimation. Last, we speed up the estimation process through a combination of two techniques. The first technique is to parallelize the computation across the four markets. The time savings from the parallel process more than offset the additional overhead time. The second technique is to modify equation (14) for the contraction mapping by employing the Newton-Rapson method where

[22]In the U.S., about 13 percent of household purchases a new vehicle each year before the economic downturn in 2008. In China, the number was about 5 percent in 2012.

the update is based on the derivate of the market share with respect to the mean utility δ_{mt}:

$$\delta_{mt}^{n+1} = \delta_{mt}^{n} + \left[\frac{\partial ln[\hat{S}(\delta_{mt}^{n}, \theta_2, \theta_3)]}{\partial \delta_{mt}^{n}}\right]^{-1} \left\{ln(S_{mt}^{o}) - ln[\hat{S}(\delta_{mt}^{n}, \theta_2, \theta_3)]\right\}. \tag{20}$$

Although additional time is needed to calculate the derivatives, we find there is still considerable time savings from fewer iterations due to the quadratic rate of convergence of the Newton-Rapson method.

6 Estimation Results

In this section, we first present evidence from the reduced-form regressions on the common trend assumption and the sales impact of the lottery policy in Beijing. Then we discuss the parameter estimates from the random coefficients discrete choice model.

6.1 Evidence from Reduce-form Regressions

To examine the validity of the common trend assumption across the cities, we estimate the following regression based on data from 2008 to 2010 (pre-policy period).

$$ln(S_{mtj}) = \delta_j + \lambda_{mt} + \eta_t + \mathbb{1}(m = 3)\eta_t' + \zeta_{ms} + e_{mtj},$$

where the dependent variable is the log market shares. δ_j is model (vintage-nameplate) dummies. λ_{mt} is city-year fixed effects to capture city-specific and time-varying demand shocks. The common trend assumption assumes that these shocks are the same across cities in a given year. The other terms are defined the same as in equation (11): we include time (year-month) fixed effects, Shanghai-specific time effects and city-segment fixed effects.

Table 4 presents the regression results for three specifications. The first two specifications use all observations from 2008 to 2010 while the third one drops the data in the last two months of 2010 in Beijing to remove the anticipation effect and more importantly the announcement effect in December 2010. In all specifications, the base group is Tianjin and the base year is 2008. The first specification does not include Shanghai-specific time fixed effects but include Shanghai-year fixed effects. The coefficient estimate on ln(price+bid) suggests a price elasticity of -4.846, which is a plausible magnitude. All the city-year interactions are small in magnitude and not statistically different from zero, suggesting a similar time trend across the four cities. The second specification include Shanghai-specific time fixed effects to control for monthly demand shocks in Shanghai that are different from the base group and could be correlated with the average bid. The price coefficient reduces to -5.089, consistent with the conjecture that unobserved demand shocks that are correlated with the average bid can bias the price coefficient toward zero. Nevertheless, the difference in the price coefficient estimates is quite small. The city-year interactions again have small and insignificant coefficient estimates. The third specification produces very similar results to the second specification, suggesting that the anticipation and announcement effects are not large enough to affect the results qualitatively. The evidence from Figure 4 and these results support the common trend assumption, the basis of our first set of moment conditions in the structural estimation.

We next use a DID framework to examine the impact of lottery policy on vehicle sales. These results will be compared with our estimates from the structural demand model. The equation for DID is very similar to equation except replacing city-year fixed effects with lottery policy dummies for Beijing in 2011 and 2012. The regression results are presented in Table 5. The first specification uses all observations from 2008 to 2012 while the other two drop observations in the last two months in 2010 and the first two months in 2011 in Beijing. While the first two specifications include city-specific time trend (up to second-order polynomials), the third one does not.

Using the full data set, the estimated impact of the lottery policy on sales is a reduction of 60.6% in 2011 and 50.7% in 2012. These imply that without the policy, the sales would have been 847,000 units in 2011 and 1.05 million units in 2012, compared with a pre-policy sales of 804,000 in 2010. The second specification produces slightly smaller sales impacts: 54.1% and 40.4% in 2011 and 2012, respectively. This is intuitive since we drop the last two months in 2010 where the increase in sales was partly due to the fact that people moved their purchase forward from the future. So without the policy, the sales would have been smaller in 2010. The sales impacts of the policy would have been smaller in 2011 and 2012 to be consistent with growth in other cities. These estimates imply that the sales would have been about 728,000 and 873,000 in the absence of the policy in 2011 and 2012. The third specification leads to slightly larger sales impacts than those from specification two. We will come back to these estimates for comparison once we obtain estimates from the structural model.

6.2 Parameter Estimates from the Demand System

Table 6 shows parameter estimates from the GMM estimation for three specifications. The first panel represent parameters in θ_2 which appear in the household-specific utility function in equation (12). The three parameters in the second panel are the auxiliary parameters θ_3 that are needed to incorporate the policies into the calculation of market shares as shown in equations (8) and (9). We do not present parameter estimates for θ_1 since they are not needed to perform our policy simulations and welfare analysis: θ_2, θ_3, and δ_{mtj}, the mean utilities from equation (20) suffice.

The first specification is the benchmark model and our preferred specification. Below we discuss the coefficient estimates and compare results across different specifications. We note however, that the magnitude of the preference parameters by themselves are hard to interpret and we defer much of the discussion on the comparison across specifications in the next two sections where we simulate sales and conduct welfare analysis using these parameters. In the benchmark specification, the coefficient estimate on Ln(price+bid) is negative while that on the interaction between this price variable and ln(income) is positive. This suggests that households with a higher income are less price sensitive. Ln(income) ranges from 0.55 to 6.72, implying a negative coefficient on the price variable for all households. The variable ln(income) in the specification is to capture that the utility difference between a new vehicle and the outside good varies by income. For most of the vehicle models (about 95%), the partial effect of ln(income) is positive, implying that the utility difference increases with income.

The next five coefficients are random coefficients, representing the standard deviation estimates of the normal distribution for preferences on each vehicle characteristics. The random coefficient on constant captures the variation (due to unobserved household demographics) in the utility difference between a new vehicle and the outside good. Three out of five random coefficients are statistically significant, adding consumer heterogeneity to what is implied by income heterogeneity.

To get a sense of the magnitude of coefficient estimates on price variables, we calculate price elasticities based on these coefficient estimates and the estimated mean-utilities. The average own price elasticity is -10.51 with a range of -8.70 to -15.97. Models with a higher price tend to have a smaller elasticity in magnitude, consistent with the intuition. The average elasticity is somewhat larger in magnitude than the estimates obtained for the U.S. market which range from -3 to -8.4 (BLP, Goldberg 1995, Petrin 2002, and Beresteanu and

Li 2011).[23]

However, we believe our estimates are reasonable. In addition to the fact that we have a different identification strategy as discussed in Section 5 , the difference could be attributed to at least the following two reasons. First, the income level in these four cities in China is less than one half of the U.S. income level during the data period of 1981 to 1993 used in Petrin (2002). To the extent that higher income would reduce price sensitivity, the differences in income could lead to the differences in price elasticities. Second, vehicle prices in our data are much higher than MSPRs in the U.S. for the same brand.[24] For example, a Hyundai Sonata GLS Sedan with 2.4 Liter engine with base options had a MSRP of $19,695 in the U.S., and a similar model produced in China had an MSPR of 178,800 RMB (over $28,000). That is, one need to adjust our price elasticities downward (in magnitude) in order to compare them with the elasticities in the U.S. market. High vehicle prices in China are in part due to the fact that they include three types of taxes on top of the prices that dealers get: value-added tax, consumption tax, and sales tax. For an average vehicle, these three amount to about one third of the vehicle price.[25]

The first auxiliary parameter ρ is the ratio of total license allocated over the number of would-be vehicle buyers (without the quota constraint) that would need a new license (e.g., fist-time buyers) under the quota system. It measures the stringency of the quota system: the smaller it is, the more stringent the quota is. It is a very important parameter in estimating the counterfactual sales under the policy. The parameter is estimated to be 0.202 in the benchmark specification, implying that only one out of five would-be buyers that need a license is able to obtain a license through the lottery. This should not be compared with the observed odds in the lottery because the pool of participants includes not only those who enter the market for a new vehicle in the current month but also unmet demand in the past months and future buyers. However, our empirical model abstracts away from these strategic behaviors on the part of consumers in defining the market size. Nevertheless, as

[23]Petrin (2002) based on data from 1981 to 1993 in the U.S. market and Beresteanu and Li (2011) based on data from 1999 to 2006 both use micro-moments for estimation, yielding an average price elasticity of -6 and -8.4, respectively.

[24]Imports account for less than 3% of the auto market in China. Most brands sold in U.S. are available in China but they are produced there by joint ventures between foreign and domestic auto makers. Please see Li, Xiao and Liu (2013) for a discussion on China'a auto industry.

[25]Value-add tax is 17%. Consumption tax varies greatly with engine size from 1% for vehicles with a engine size less than one liter to 40% for those with a engine size larger than 4 liter. An average vehicle with a engine size of 1.5 to 2.0 liter has a tax of 5%. Sales tax is 10% except when 5% and 7.5% were levied on vehicles with small engines during 2009 and 2010, respectively.

we will show below, the estimate of 0.202 (together with other parameters) leads to a very reasonable counterfactual sales without the policy.

The second and third auxiliary parameters define the probability of a buyer needing a new license as given by equation (7). The positive coefficient γ_1 suggests that as vehicle ownership goes up, the share of potential buyers who need a new license decreases. This is intuitive given that as vehicle ownership increases, more and more households would need to replace their old vehicles (by scrappage) with new vehicles and hence do not need new licenses. These two parameter estimates imply that about 72% percent of potential buyers in Shanghai and 69% in Beijing in 2012 would need a license should they decide to buy a new vehicle.

To examine the importance of the first set of moment conditions that are based on the common trend assumption, we estimate the model without these moment conditions and the results are reported under alternative one. The coefficient estimates on ln(price+bid) and its interaction with income are both larger in magnitude although they have the same signs as those from the benchmark specification. The average own price elasticity is -13.55 with a range from -20.48 to -11.07. These are about 30 percent larger than those from the benchmark specification in magnitude. Another important difference is that the estimate of ρ is 0.395, almost twice as large as that from the benchmark model. Compared with the estimate from the benchmark model, this estimate implies that the quota is much less stringent and should therefore predict a smaller sales under the counterfactual scenario of no quota. Further comparison between these specifications will be made in the following sections.

Alternative specification two in Table 6 does not utilize the second set of moment conditions (i.e., micro-moments) that are based on shares of new vehicle purchases by income group. These moment conditions are important in recovering the heterogeneity in WTP for a new car and price sensitivity across income groups. Without these micro-moment conditions, the parameter estimate on the price and income interaction term has a negative sign. This implies that high income groups are more price sensitive, running against our intuition as well as the results from the benchmark model. The estimates of price elasticities has a mean of -13.56 and a range from -21.13 to -11.49. These are similar to those from alternative one and are larger in magnitude to those from the benchmark model. A key difference is that the elasticities are larger in magnitude for more expensive vehicles, opposite to the results from the first two specifications. It is interesting to note that the auxiliary parameters are very close to those from the benchmark model, suggesting that these parameters are largely identified through the first and third sets of moment conditions.

Table 7 presents results from three additional specifications as further robustness checks. Alternative four examines the sensitivity of the results to the definition of market size. In the benchmark model, we assume half of the households in a city participate in the market for new vehicles, implying the market size of a month is the number of households divided by 24. Alternative four assumes the market size to be the number of all households as is often done in the studies on the U.S. market. The coefficients by and large are very similar to those in the benchmark specification. This implies that the mean utilities must be smaller in this specification in order to generate the same number of new vehicle sales as in the benchmark specification.

Alternative five assumes that 85% of the lottery winners use the license to purchase new vehicles instead of used vehicles in Beijing and the number in Shanghai is 95%. As specified by the policy, buyers of used vehicles need a license if they do not already have one (i.e., from scrapping an old vehicle). We do not have detailed statistics on the ultimate usage of the licenses and we use these two numbers as upper bound of the respective shares. We choose a lower share for Shanghai since the buyers there tend to have higher income than in Beijing.[26]

In all the first five specifications, random draws for random coefficients are drawn from standard normal distributions which are unbounded. To remove the impacts of extreme values, alternative six uses random draws from truncated normal distributions by removing the draws below 2.5 percentile and above 97.5 percentile. The coefficient estimates are very close to those from the benchmark specification and so are the simulation results and welfare analysis shown below.

6.3 Impacts on Vehicle Sales

Table 8 presents the simulated sales under the counterfactual of no policy for various specifications. Under the benchmark specification, the counterfactual sales in Beijing are 779,272 and 1,142,064 in 2011 and 2012, relative to observed sales of 334, 308 and 520,442 under the

[26]In May 2011, about 1600 licenses from the lottery were used for used vehicle purchases in Beijing according to information from used vehicle dealers: http: //auto.sina.com.cn/news/2011-06-21/0752790152.shtml. In June and July 2011, the ratio was 11.7% and 14%, respectively: http: //auto.sohu.com/20110816/n316431063.shtml.

policy. This suggests that the lottery policy reduced sales by 57 and 54 percent, respectively in 2011 and 2012. The estimated sales impact in 2011 is very close to those presented in the DID analysis in Table 5, although the sales impact in 2012 is larger than those in Table 5, especially those from specifications two and three. Moreover, the 46.5% sales growth under the no-policy scenario from 2011 to 2012 would be quite a bit larger than what was observed in Tianjin and Nanjing (13% and 5%). This larger uptake in sales, relative to other cities, could be in part due to the combination of the following two reasons. First, the number of households in Beijing increased by 2.8% from 2011 to 2012, compared to 0.6% in Nanjing, 1.4% in Shanghai, and 4.3% in Tianjin. Second, the average income of the top quantile increased from 158,233 to 189,213 from 2011 to 2012 in Beijing, a 19.6% growth. This is much larger than the other three cities: 6.5%, 10.9% and 7.9% in Nanjing, Shanghai and Tianjin, respectively. Given that the high-income group contributes disproportionally more to new vehicle purchases as shown in Table 3, the larger income growth in this group could lead to a larger increase in vehicle sales.

Under alternative one without the first set of moment conditions, the counterfactual sales in Beijing without the policy are 591,789 and 884,541 respectively. The 25% drop in sales from 2010 to 2011 without the policy is hard to explain given that the sales decrease during the same period in Nanjing and Tianjin was 6.6% and 0.02%, respectively. Since the estimation does not use the common trend assumption, the large sales drop from 2010 to 2011 in our data due to the implementation of the lottery policy in 2011 is (wrongly) attributed to a negative demand shock in 2011 in Beijing relative to other cities. This then leads to a larger estimate on ρ that implies less stringent quota and a large drop in sales without the policy, as shown in the last column of Table 6 compared to that in the benchmark specification.

It is interesting to note that although alternative specification two does not provide sensible pattern of price sensitivities, the simulation results on sales impacts are close to those from the benchmark model. This is because the policy impacts are mainly identified through the first set of moment conditions that utilize common trend assumption, just as in a DID analysis. The micro-moment conditions, on the other hand, are very important for identifying consumer WTP and subsequent welfare comparisons.

The bottom panel of Table 8 shows the counterfactual sales without the auction policy from 2008 to 2012 in Shanghai. The results from the benchmark specification suggest that the policy reduced the sales by about 52% during this period. Interestingly, all three specifications including alternative one produce very similar results. This is because the identification of the policy impacts in Shanghai is through the changes in treatment intensity (i.e., average bids) over time. That is, the estimation of the policy impacts in Shanghai is largely based on the parameter estimates on the price variables.

Table 9 presents the sales impacts under three other alternative specifications based on different assumptions on the market size, license usage for used cars, and distribution of random draws. As we have discussed in the above section, the parameter estimates from these specifications are very similar to those in the benchmark specification. Consistent with that, this table shows that the sales impacts are by and large similar to the estimates from the benchmark specification in Table 8.

7 Welfare Analysis

The purpose of this section is to compare welfare consequences under the lottery and auction systems and we focus on 2012 for illustration. The comparison is performed on two dimensions: welfare gain through allocation of the licenses (allocative efficiency), and the external costs associated with automobile usage.

7.1 WTP Schedules

We first derive consumer WTP schedule for a license in Beijing and Shanghai in 2012 based on parameter estimates obtained above. The WTP for a license by household i is given by the expected consumer surplus (CS) from the most preferred vehicle:

$$\mathrm{E}(\mathrm{CS}_i) = \mathrm{E}_\epsilon\Big[\max_{j=1,\dots,J} \frac{\bar{u}_{ij} + \epsilon_{ij}}{\mathrm{MU}_{ij}}\Big] = \frac{1}{\mathrm{MU}}\ln\Big[\sum_{j=1,\dots,J} exp(\bar{u_{ij}})\Big] \text{ if } \mathrm{MU}_{ij} \text{ is constant,} \tag{21}$$

where MU_{ij} is marginal utility of money. The last equality holds only if the marginal utility of money is constant. However, because our utility specification is nonlinear in price, we cannot use the closed-form expression in the last equality to estimate consumer surplus. Instead, we follow the simulation method developed by Herriges and Kling (1999) to estimate the expected CS from the most preferred vehicle.[27]

[27] Independent of the estimation, we separately generate 1000 households for each month in 2012 for each market and each household is characterized by a vector of income and random draws of unobserved household attributes. For each household, we draw 200 vectors of ϵ_i to increase precision.

Figure 5 depicts the WTP schedule or the demand curve for licenses in Beijing (top panel) and Shanghai (bottom panel) in 2012. These schedules are simulated based on the benchmark specification in Table 6. The WTP schedules can be viewed as populated by households with different WTP since each household demands no more than one vehicle in a period. The red vertical line denotes the level of quota: 259,800 in Beijing and 108,100 in Shanghai. The WTP schedules exhibit two salient features. First, there is considerable heterogeneity with some households willing to pay more than 200,000 RMB for a license in both markets. Our random coefficients utility framework is well suited to capture consumer heterogeneity, which is critical for welfare comparison under different allocation mechanisms as we will show below.

The second feature is that the WTP schedule in Beijing is higher than that in Shanghai, implying that households in Beijing on average are willing to pay a higher price for a license. This is a reflection that consumers in Shanghai have better access to the public transit and hence have less strong demand for a new vehicle. This is also consistent with the sales figures in Table 8: without the lottery and auction policies, the sales in Beijing and Shanghai in 2012 would have been 1,142,064 and 712, 215, respectively. Although Shanghai has more households and a higher average income than Beijing, the sales are lower in Shanghai. As we discussed in Section 2, Shanghai has been carrying out the auctions to distribute licenses for nearly three decades and the auction revenue has been used to build road infrastructure and improve public transit.[28]

The market clearing prices under a uniform price auction, the highest rejected bid in such a auction, are given by the interaction between the vertical lines and the WTP schedules. Based on the simulations, the clearing price would have been 82,210 RMB in Beijing and 53,020 RMB in Shanghai. The simulated clearing price in Shanghai is slightly higher than the observed lowest bid of 52,800 from the Shanghai auction in 2012. This comparison is consistent with the experimental results in Liao and Holt (2013) that the lowest bid from the Shanghai auction is slightly smaller than the WTP. It is important to note that our estimation does not have any explicit constraints on WTP and clearing prices. Therefore, the fact that the simulated clearing price based on model estimates is close to the observed lowest bid in the data provides another validity check for our empirical method. This adds confidence for us to take the estimated model to perform the welfare analysis below.

Under alternative specification one presented in Table 6, the clearing prices in a uniform

[28]In the long run, the WTP is affected by the auction system and hence endogenous. Our analysis focus on the short-run.

price auction are 91,067 RMB and 90,973 RMB for Beijing and Shanghai, respectively. Although the clearing price in Beijing is close to what is estimated based on the benchmark specification, the clearing price in Shanghai is 80% higher than the lowest bids observed in 2012. This finding together with the poor sales prediction by this specification as discussed in the previous section underscore the importance of the first set of moment conditions in our estimation. Based on the parameter estimates from alternative specification two in Table 6, the clearing prices are 99,744 RMB in Beijing and 67,898 RMB in Shanghai, both higher than those from the benchmark specification.

7.2 Consumer Surplus and Externalities

We now compare consumer surplus from license allocations and externalities associated with automobile usage under the two mechanisms. As discussed in Section 2, although auctions can achieve more efficient allocations than lottery in general, this may not be true when the resources being allocated generate negative externalities and the externalities increase with WTP. Automobile usage generates multiple externalities including congestion, air pollution, and accident externalities (Parry et al. 2007). These externalities could impose a huge burden on social welfare especially in rapidly developing countries where they are not adequately controlled for.

There are good reasons to believe that the externalities are positively correlated with consumer WTP for a vehicle license. First, households with a high WTP for a license tend to be those with a higher household income. They tend to drive larger and more luxurious vehicles that have lower fuel economy and hence burn more gasoline for the same travel distance.[29] Second, households with a higher income tend to drive more. Small and Van Dender (2007) estimate the elasticity of annual vehicle miles traveled (VMT) with respect to income to be 0.11 using the U.S. state-level data. The relationship holds up among households in Beijing according to 2010 Beijing Household Travel Survey, a representative household survey in Beijing conducted by Beijing Transportation Research Center. Among the seven income groups in the survey, the average annual household VMT ranges from 15,300 to 25,100 km from the lowest income group (less than 50,000 RMB in annual income)

[29]Larger and heavier vehicles also impose larger accident externalities as shown in (Anderson and Auffhammer 2013; Jacobsen 2013). In addition, truck-based SUVs also impose larger externalities than passenger cars due to vehicle design (White 2004; Anderson 2008; Li 2012). However, this difference and accident externalities in general are found to be small relative to congestion and air pollution in the slow-moving urban traffic.

to the highest income group (more than 300,000 RMB). The average VMT is 16,100 km among all 133,14 households in the data.

We first examine consumer surplus in Beijing from lottery and auction mechanisms under the observed quota level of 259,800 units in 2012. We use a uniform price auction rather than the Shanghai auction format as the basis for comparison. As shown in Liao and Holt (2013), the Shanghai auction is non-standard and may not yield the efficient outcome. Table 10 presents the simulation results for two different scenarios regarding the time-span for the external cost analysis to be discussed below, bearing in mind that consumer surplus does not vary between the two scenarios.

Under a uniform price auction, the clearing price would have been 82,210 RMB in Beijing in 2012 as depicted in the top panel of Figure 5. The total consumer surplus given by the area to the left of the vertical line in Figure 5 would have been 48.35 billion, but only 11 percent of it (or 5.39 billion) is realized under the lottery system, resulting in a welfare loss of nearly 43 billon. This striking result is due to: (1) there is a great deal of heterogeneity in WTP for a license as shown in the figure, and (2) the lottery allocates the license randomly. The significant welfare loss is robust to the three assumptions (market size, license usage for used vehicles, and extreme random draws) examined under the three alternative specifications presented in Table 7: it is 41.1, 38.4 and 39.8 billion RMB, respectively.

We now turn to differences in total external costs under the two mechanisms. The average fuel economy of new vehicles sold under the lottery is 8.92 liters/100km compared to 9.19 under the auction. This is consistent with our discussion above: the auction system will lead to a market with more high income households than the lottery system. In addition, the average annual VMT of 17,760km under the lottery is also smaller than that from the auction, 19,800km.[30] To quantify the external costs associated with the usage of license, we need to make two critical assumptions.

The first is with regard to the time horizon over which the external costs accrue. Although a license can theoretically stay with the households over multiple vehicle lifetimes, it is probably not a very meaningful exercise to take such a long-term view in calculating external costs. Policies in China tend to change frequently, often in response to the fast-changing environment. In fact, the Beijing municipal government has been considering proposals to adopt higher gasoline taxes (termed pollution charges) and congesting prices that could more

[30] We assign an annual VMT to each household according to the seven income groups in 2010 Beijing Household Travel survey.

effectively deal with driving decisions. How long the quota system will last and how binding it will be is uncertain. We carry out our analysis based on two scenarios: 15 years and 10 years, which bound the average lifetime of a vehicle. We believe that these could be plausible time horizons to consider from the policy perspective. In addition, our demand analysis assumes that a consumer's WTP for a license is the consumer surplus for the most preferred vehicle during the vehicle lifetime. The WTP estimates obtained under this assumption are quite plausible as we discussed above. Lastly, we note that different time horizons would affect our welfare estimates but would not affect the comparison of the two mechanisms qualitatively.

The second assumption is on the external costs per gallon of gasoline consumed.[31] We take the results from Creutzig and He (2008) that estimate the externalities from automobile usage in Beijing. The most plausible estimate from their study is 0.85 RMB/km (in 2012 terms) and congestion and air pollution account for about 80%. This translates to 9.7 RMB per liter of gasoline consumed at the average level of fuel economy of 8.74 liters/100km. This implies external costs of \$6.02 from one gallon of gasoline used by automobiles. Parry and Timilsina (2009) estimate the external costs from automobile usage to be about \$6.05 (in 2005 \$) per gallon of gasoline in Mexico city.[32] Mexico city probably offers a reasonable comparison to Beijing in that both have very severe congestion and air pollution. In addition, the average household income is roughly the same, about one fourth of that in the US. For final calculation of external costs, we recognize that China currently imposes a gasoline tax of 1 RMB per liter to deal with the externalities from automobile usage. So we take 8.7 RMB/km as the (un-internalized) external costs for our welfare calculations, keeping in mind that although variations on the unit external cost affect the magnitude of welfare estimates, they do not qualitatively affect the efficiency comparison between the two mechanisms.

Figure 6 plots the external cost curve (the dashed downward sloping curves) for the 15-year time horizon in the top panel and 10-year time-horizon in the bottom panel. The external cost curves depict the total external costs from the automobile usage during a given time horizon for the corresponding households on the WTP curve. The estimates are obtained based on a five percent annual discount rate. Since the external costs increase with

[31] Congestion is directly related to travel distance while CO_2 emissions is proportional to gasoline consumption. Air pollution is more complicated as it relates to engine running time even for the same distance. In addition, it also has to do with vehicle age as old vehicles pollute disproportionally more (Kahn 1996). The literature has generally translated externalities on a per-unit of gasoline basis in the discussion on the optimal gasoline tax as a (second-best) policy instrument.

[32] Parry and Small (2005) estimate the external costs of one gallon of gasoline consumption to be around \$0.83 and \$1.23 (in 2000 \$) in the US and UK, respectively.

WTP, the auction system leads to larger externalities than the lottery system. As shown in Table 10, the total external costs during a 15-year time horizon is estimated to be 33.02 billion RMB under the lottery system and 40.21 under the auction system, implying a non-trivial difference of nearly 7 billion RMB. With a 10-year time horizon, the estimates are 24.57 billion RMB and 29.92 billion RMB, respectively.

7.3 Social Welfare and Comparison

Does the difference in external costs under the two mechanisms change their efficiency ranking? Although the difference shrinks the gap in social welfare between the two, it does not change their order in our context. The net social welfare (total surplus minus total external costs) is estimated to be -27.63 billion under the lottery, compared with 8.13 billion under the auction. The net social welfare from the lottery system being negative is due to: (1) the external costs are larger than consumer surplus for a larger number of households with relatively low WTP as shown in Figure 6; and (2) the level of quota is set too high relative to the optimal level to be discussed below.

The difference in net social welfare from the two mechanisms is 35.76 billion RMB. It is smaller than the difference in consumer surplus of 42.96 billion. Nevertheless, it implies that by using the lottery system instead of a uniform price auction or other efficient auction designs, Beijing municipal government left nearly 36 billion RMB on the table. If we use a 10-year time horizon to calculate total external costs, the difference in net social welfare is about 38 billion. The findings imply that under our model estimates and a set of reasonable assumptions, the lottery system performs poorly relative to a uniform price auction in terms of efficiency. In terms of government revenue, a uniform price auction would have generated over 21 billion RMB for the municipal government, more than covering the total subsidy of 17 billion to the public transit system in 2012. Alternatively, if the revenue were to be use for additional support to the public transit as Shanghai does, it could more than double the existing subsidy level.

The last exercise we conduct is to examine the optimal level of quota in Beijing. Our analysis is not meant to devise the optimal policy for dealing with congestion and air pollution. In fact, real-time congestion pricing and gasoline taxes should be more effective policies than the quota system since they directly target the source of externalities: driving and gasoline consumption. Nevertheless, we investigate another margin of improvement within the framework of the quota system. The optimal level of quota is determined by the intersection of the WTP and external cost curves as shown in Figure 6. Under the 15-year time horizon, the optimal level of quota is 125,612, about half of the quota used in 2012. The clearing price from a uniform price auction would be 157,500 RMB, nearly twice as high as it would be with the existing quota. As shown in the second panel of Table 10, consumer surplus would be 34.87 billion, less than 48.35 billion under the existing quota. However, the total external costs would be 20.87 billion with the smaller quota, about half of those from the existing quota. As a result, the net social welfare would be 14.12 billion, compared with 8.13 billion under the existing quota. Government revenue would be 19.78 billion, only slightly smaller. If we use a 10-year time horizon, the optimal quota would be higher since the unit external costs would be smaller. The net social welfare would increase to 20.79 billion RMB.

8 Conclusion

Governments and resource managers have relied on both market-based mechanisms and non-market based mechanisms to allocate scarce public resources. Transit authorities in many large cities in China are charged with implementing a license quota system to limit new vehicle purchases in an effort to combat worsening traffic congestion and air pollution. How should they distribute the limited number of vehicle licenses? One might be quick to point out that market-based mechanisms that distribute the resources to those with highest value would be more efficient. However, if the usage of the resources (the license to purchase a vehicle and ultimately vehicle usage in our case) generates negative externalities that are positively correlated with willingness-to-pay, the efficiency advantage of market-based mechanisms may disappear because the usage of the resources by those with highest private benefit could lead to smaller social benefit.

Although there is a large and distinguished theoretical literature on resource allocation mechanisms since Coase (1959) and Vickrey (1961), with the exception of the auction literature, empirical studies that quantify the misallocation of resources and its welfare consequences are sparse. This study offers the first empirical investigation on the welfare consequences of the lottery and auction systems used in Beijing and Shanghai, two largest cities in China, to distribute vehicle licenses. The adoption of the two types of allocation mechanisms for the same type of resources offers a unique opportunity. The analysis builds upon the random coefficients discrete choice model and develop a novel empirical strategy that combines the common trend assumption between treatment and control groups in impact evaluation literature with the micro-moment conditions as in Petrin (2002) to identify structural parameters, without relying on the exogeneity assumption between observed and unobserved product attributers commonly used in this literature.

Our analysis shows that the lottery system offers a large advantage over the auction system in reducing automobile externalities within a quota system. Nevertheless, this advantage is offset by the efficiency loss from misallocation under the lottery system. The estimated welfare loss from the lottery system in Beijing is 36 billion RMB (or $6 billion) in 2012. This underscores the importance of allocation efficiency in the presence of large consumer heterogeneity. This large loss is especially concerning given that Beijing is often a leader in setting state policies. Guiyang, the capital city of Guizhou province, and Guangzhou, the capital city of Guangdong province started license lotteries in July 2011 and August 2012, respectively. The China's Ministry of Environmental Protection has endorsed the vehicle quota system and it is projected that about 25 major cities in China will adopt a quota system by 2015. Although the rational for choosing lottery over auction has not been made public, a reasonable conjecture is that the decision reflects concerns for equality. Our analysis shows that a uniform price auction could have generated 21 billon RMB revenue to Beijing municipal government, which could then be used to double its subsidies to local public transit system or for other measures to address the distributional concern.

Air pollution and traffic congestion are arguably two of the most pressing issues for China's urban residents, whose share increased from about a quarter to more than 50 percent of the total population during the last two decades and is still growing. Caught off-guard by these urban challenges, policy makers in China are scrambling for solutions and are adopting a range of policies that typically are of the nature of command and control to deal with externalities from automobile usage. The vehicle license quota system being adopted by urban transit authorities in China is a blunt instrument that does not directly attack the sources of externalities. In contrast, market-based mechanisms to correct for externalities have long been advocated by economists since at least Pigou. Among real-world implementations, congestion pricing schemes outlined in Vickrey (1963) have been adopted in London, Singapore and Stockholm and have been found effective in reducing congestion (Anas and Lindsey 2011). Fuel taxes offer a second-best and administratively simple policy to deal with multiple externalities from automobile use (Parry et al. 2007). In this paper, we focus on the welfare consequences of the two allocation mechanisms within the confines of the license quota system. The comparison between the quota system itself with road pricing and gasoline taxes is worthy of future research.

References

Anas, Alex and Robin Lindsey, "Reducing Urban Road Transportation Externalities: Road Pricing in Theory and In Practice," *Review of Environmental Economics and Policy*, 2011, *5* (1), 66–88.

Anderson, Michael L., "Safety for Whom? The Effect of Light Trucks on Traffic Fatalities," *Journal of Health Economics*, 2008, *27*, 973–989.

_ **and Maximilian Auffhammer**, "Pounds That Kill: The External Costs of Vehicle Weight," *Review of Economic Studies*, 2013.

Beresteanu, Arie and Shanjun Li, "Gasoline Prices, Government Support, and the Demand for Hybrid Vehicles," *International Economic Review*, 2011, *52* (1), 161–182.

Berry, Steven, James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

_ **,** _ **, and** _ , "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy," *American Economic Review*, 1999, *89* (3), 400–430.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi, "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality," 2011. Working Paper.

Coase, Ronald, "The Federal Communications Commission," *Journal of Law and Economics*, 1959, *1*, 1–40.

Cramton, Peter, "Money Out of Thin Air: The Nationwide Narrowband PCS Auction," *Journal of Economics and Management Strategy*, 1995, *4*, 267–343.

_ , "Spectrum Auctions," in Martin Cave, Sumit Majumdar, and Ingo Vogelsang, eds., *Handbook of Telecommunications Economics*, Amsterdam: Elsevier Science, 2002, pp. 605–639.

Creutzig, Felix and Dongquan He, "Climate Change Mitigation and Co-Benefits of Feasible Transport Demand Policies in Beijing," *Transportation Research Part D*, 2008, *14* (2), 120–131.

Glaeser, Edward L. and Erzo Luttmer, "The Misallocation of Housing Under Rent Control," *American Economic Review*, 2003, *93*, 1027–1046.

_ **and Matthew E. Kahn**, "The Greenness of Cities: Carbon Dioxide Emissions and Urban Development," *Journal of Urban Economics*, 2010, *67* (3), 404–418.

Goldberg, Pinelope, "Product Differentiation and Oligopoly in International Markets: The Case of the US Automobile Industry," *Econometrica*, 1995, *63* (4), 891–951.

Harris, Milton and Artur Raviv, "Allocation Mechanisms and the Design of Auctions," *Econometrica*, 1981, (6), 1477–1499.

Heckman, James and V. Joseph Hotz, "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Pro-grams: The Case of Manpower Training," *Journal of the American Statistical Association*, 1989, *84*, 862–880.

Herriges, Joseph and Catherine Kling, "Nonlinear Income Effects in Random Utility Models," *Reivew of Economics and Statistics*, 1999, *81* (1), 62–72.

Jacobsen, Mark, "Fuel Economy and Safety: The Influences of Vehicle Class and Driver Behavior," *American Economic Journal: Applied Economics*, 2013, *5* (3), 1–26.

Kahn, Matthew, "New Evidence on Trends in Vehicle Emissions," *Rand Journal of Economics*, 1996, *27* (1), 183–196.

Klier, Thomas and Joshua Linn, "New-vehicle characteristics and the cost of the Corporate Average Fuel Economy standard," *Rand Jounal of Economics*, 2012, *43* (1), 186–213.

Li, Shanjun, "Traffic Safety and Vehicle Choice: Quantifying the Arms Race on American Roads," *Journal of Applied Econometrics*, 2012, *27* (1), 34–62.

_ **, Junji Xiao, and Yiming Liu**, "The Price Evolution in Chinas Automobile Market," 2013. Working paper.

Liao, Evan Zuofu and Charles Holt, "The Pursuit of Revenue Reduction: An Experimental Analysis of the Shanghai License Plate Auction," 2013. working paper.

Meyer, Bruce, "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, 1995, *13*, 151–161.

Parry, Ian and Govinda Timilsina, "Pricing Externalities from Passenger Transportation in Mexico City," 2009. World Bank Policy Research Working Paper.

_ **and Kenneth Small**, "Does Britain or the United States Have the Right Gasoline Tax?," *American Economic Review*, 2005, *95* (4), 1276–1289.

_ **, Winston Harrington, and Margaret Walls**, "Automobile Externalities and Policies," *Journal of Economic Literature*, 2007, *XLV*, 374–400.

Petrin, Amil, "Quantifying the benefit of new products: the case of minivan," *Journal of Political Economy*, 2002, *110* (4), 705–729.

Small, Kenneth and Kurt Van Dender, "Fuel Efficiency and Motor Vehicle Travel: The Declining Rebound Effect," *Energy Journal*, 2007, *28* (1), 25–51.

Taylor, Grant, Kevin K. Tsuib, and Lijing Zhu, "Lottery or Waiting-line Auction?," *Journal of Public Economics*, 2003, (5-6).

Vickrey, William, "Counterspeculation and Competitive Sealed Tenders," *Journal of Finance*, 1961.

_ , "Pricing in Urban and Suburban Transport," *American Economic Review*, 1963, (2).

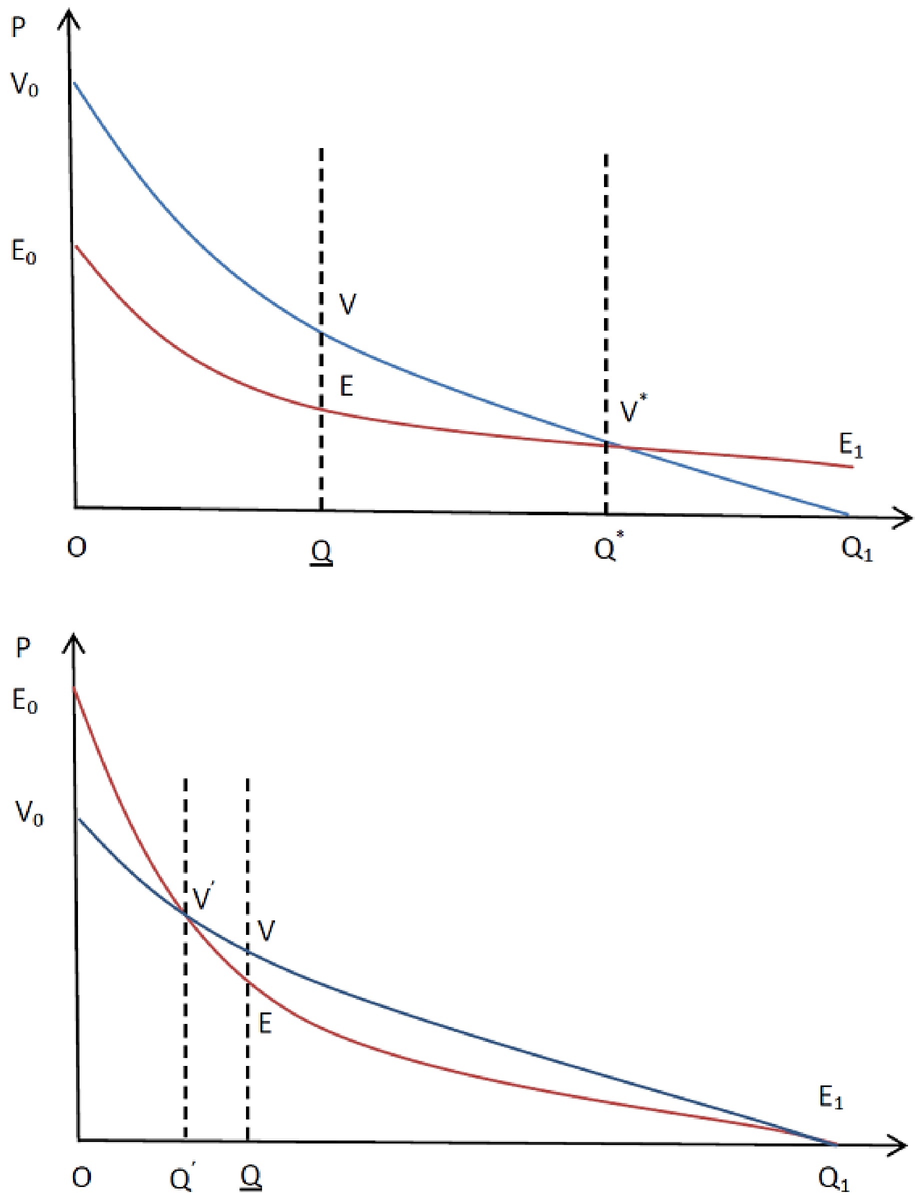
White, Michelle, "The 'arms race' on American Roads: The Effect of SUV's and Pickup Trucks on Traffic Safety," *Journal of Law and Economics*, 2004, *XLVII* (2), 333–356.

Xiao, Junji and Ju Heng, "The impacts of air-pollution motivated automobile consumption tax adjustments of China," *Review of Economics and Statistics*, 2013.

Yang, Jun, Ping Qin, Ying Liu, and Antung Liu, "An Overview of Beijing's Car Lottery Policy," 2013. working paper.

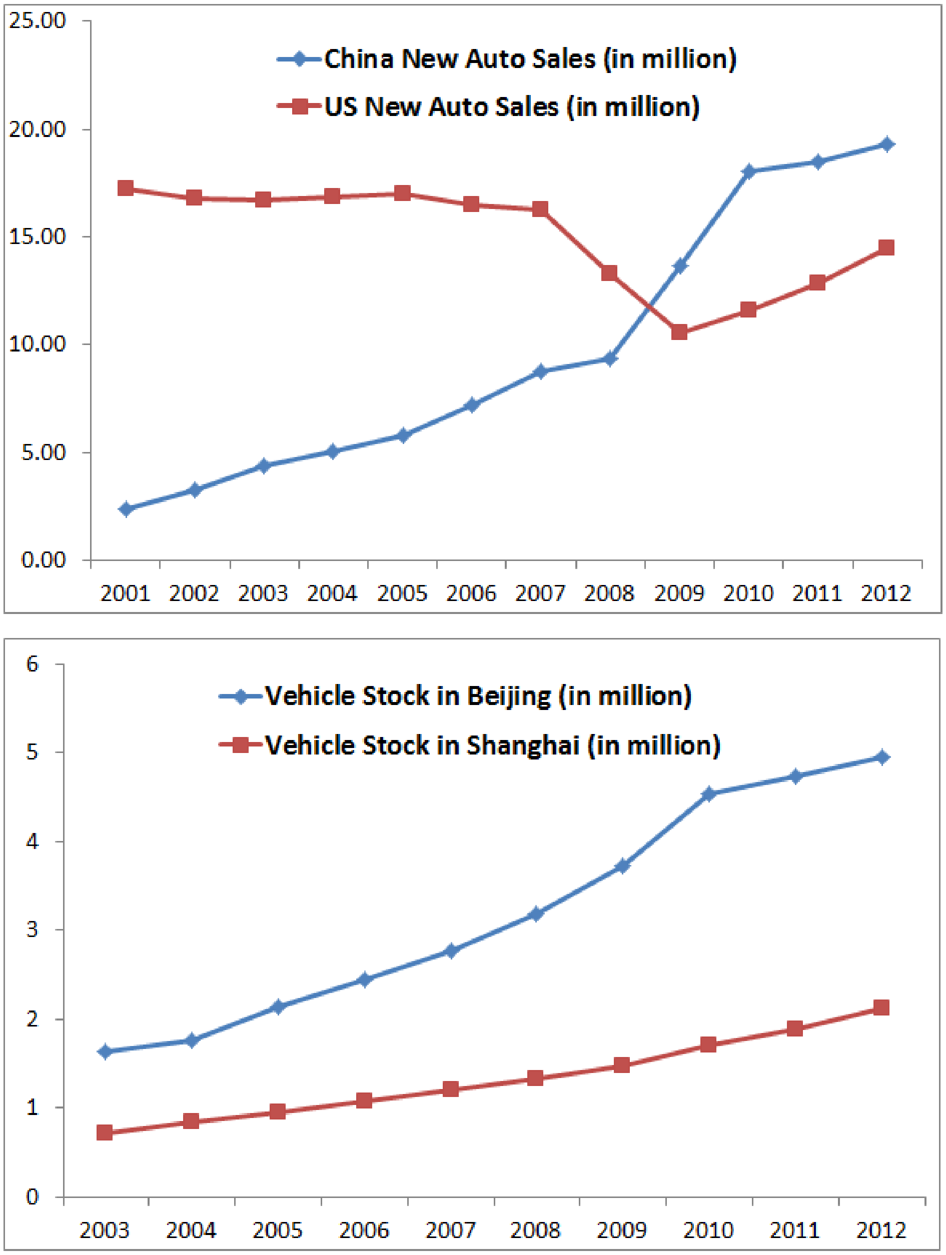
Zheng, Siqi and Matthew E. Kahn, "Understanding China's Urban Pollution Dynamics," *Journal of Economic Literature*, 2013, *51* (3), 731–772.

Figure 1: Welfare Comparison of Lottery and Auction with Externalities



Notes: In the top panel, the downward sloping curve $V_0Q_1$is the WTP schedule (demand curve) for licenses. The line below depicts the external costs from automobile usage by the consumers with the corresponding WTP on the WTP schedule. It shows that consumers with a high WTP also generate larger negative externalities when using the license. The efficiency comparison between lottery and auction depends on the relationship between WTP and the net social benefit (WTP minus external costs). If the external costs rise sufficiently faster than the increase in the WTP (e.g., the external cost curve surpasses WTP at the high WTP range) as shown in the bottom panel, allocating the licenses to the consumers with highest WTP will lead to a lower social welfare.

Figure 2: New Vehicle Sales in China and Vehicle Stocks in Beijng and Shanghai



Notes: The top panel depicts annual sales of new passenger vehicles in China and the U.S. from 2001 to 2012. The bottom panel shows the vehicle stock of passenger vehicles in Beijing and Shanghai from 2003 to 2012.

Figure 3: Monthly Vehicle Sales and Number of licenses in Beijing and Shanghai

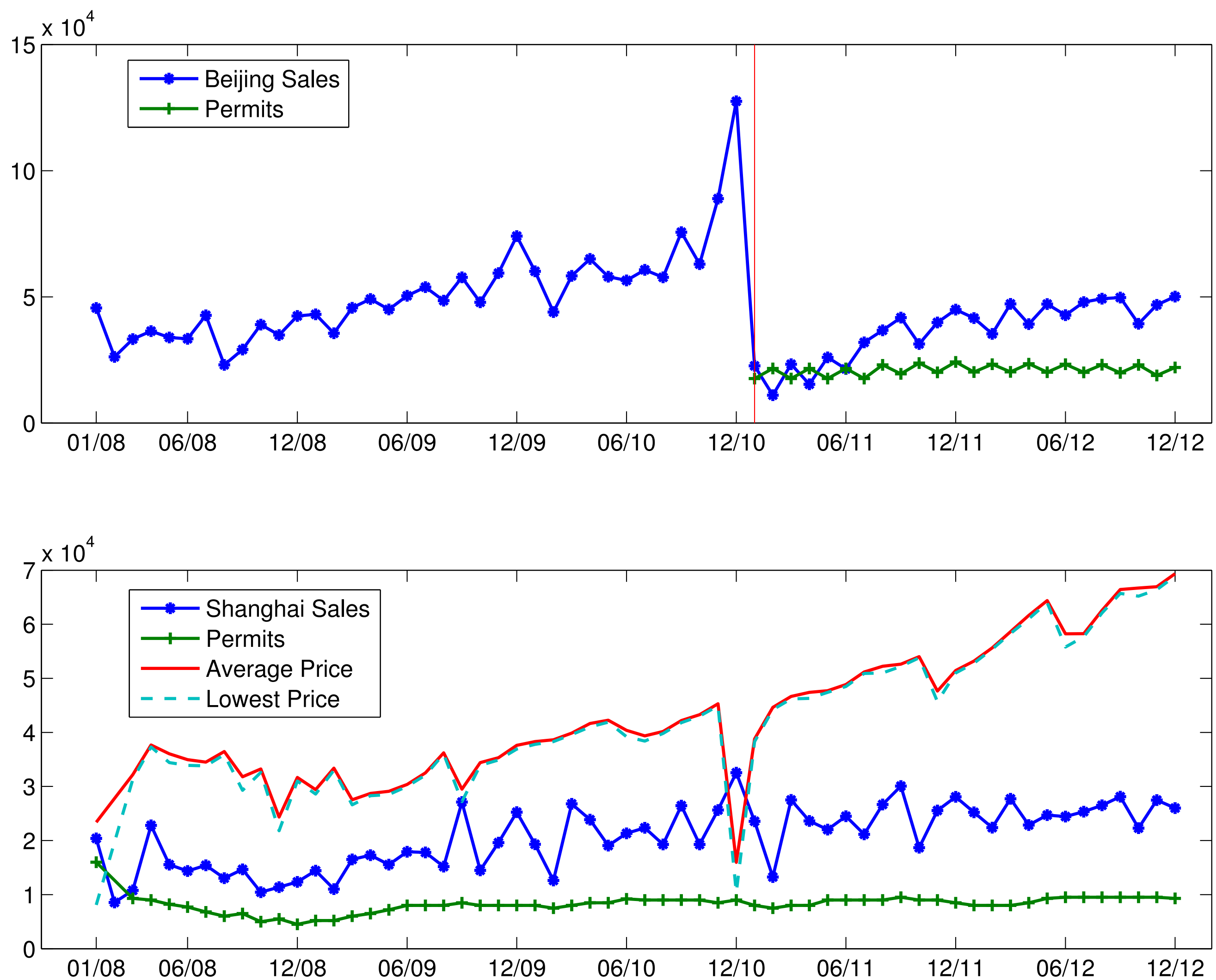


Figure 4: New Vehicle Monthly Sales (in logarithm) in the Four Cities

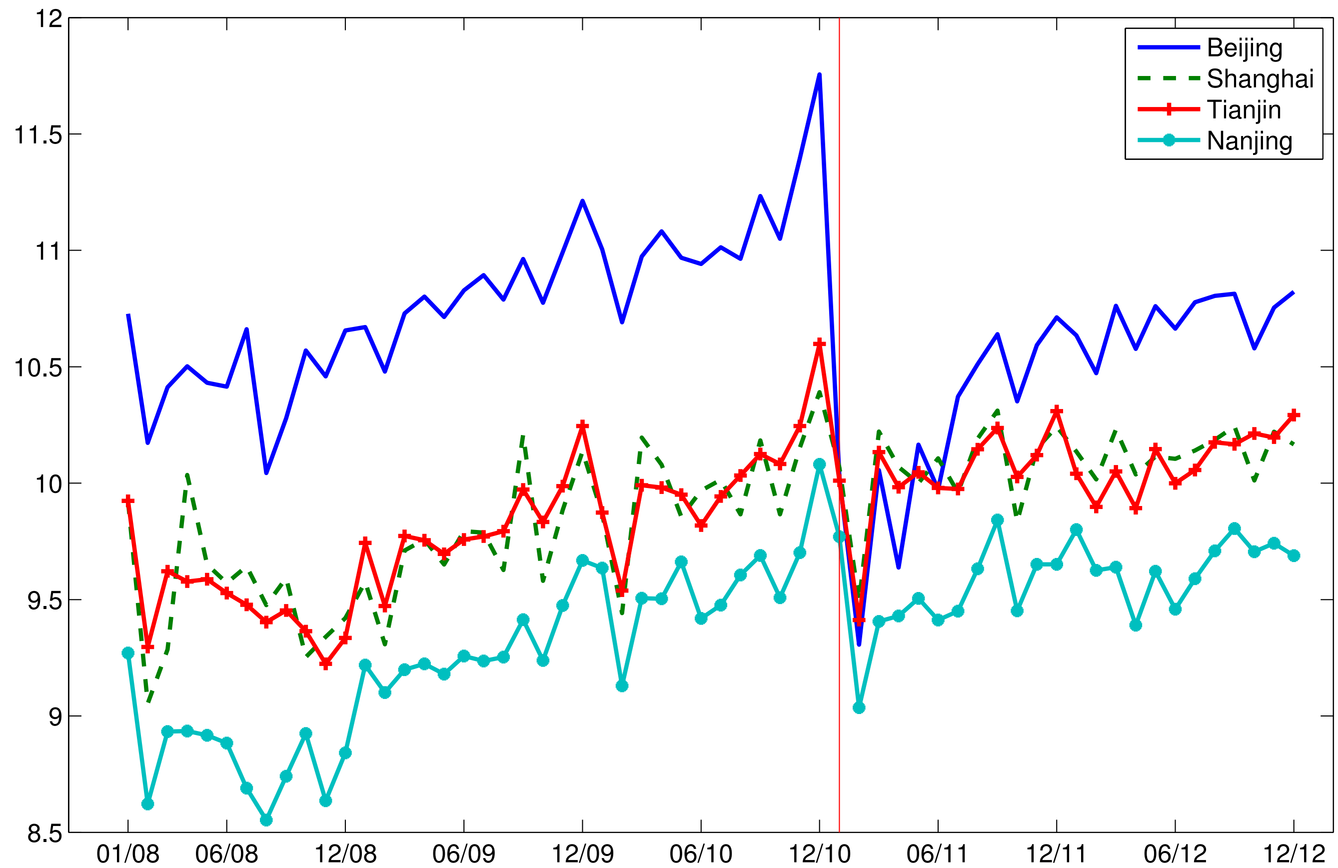
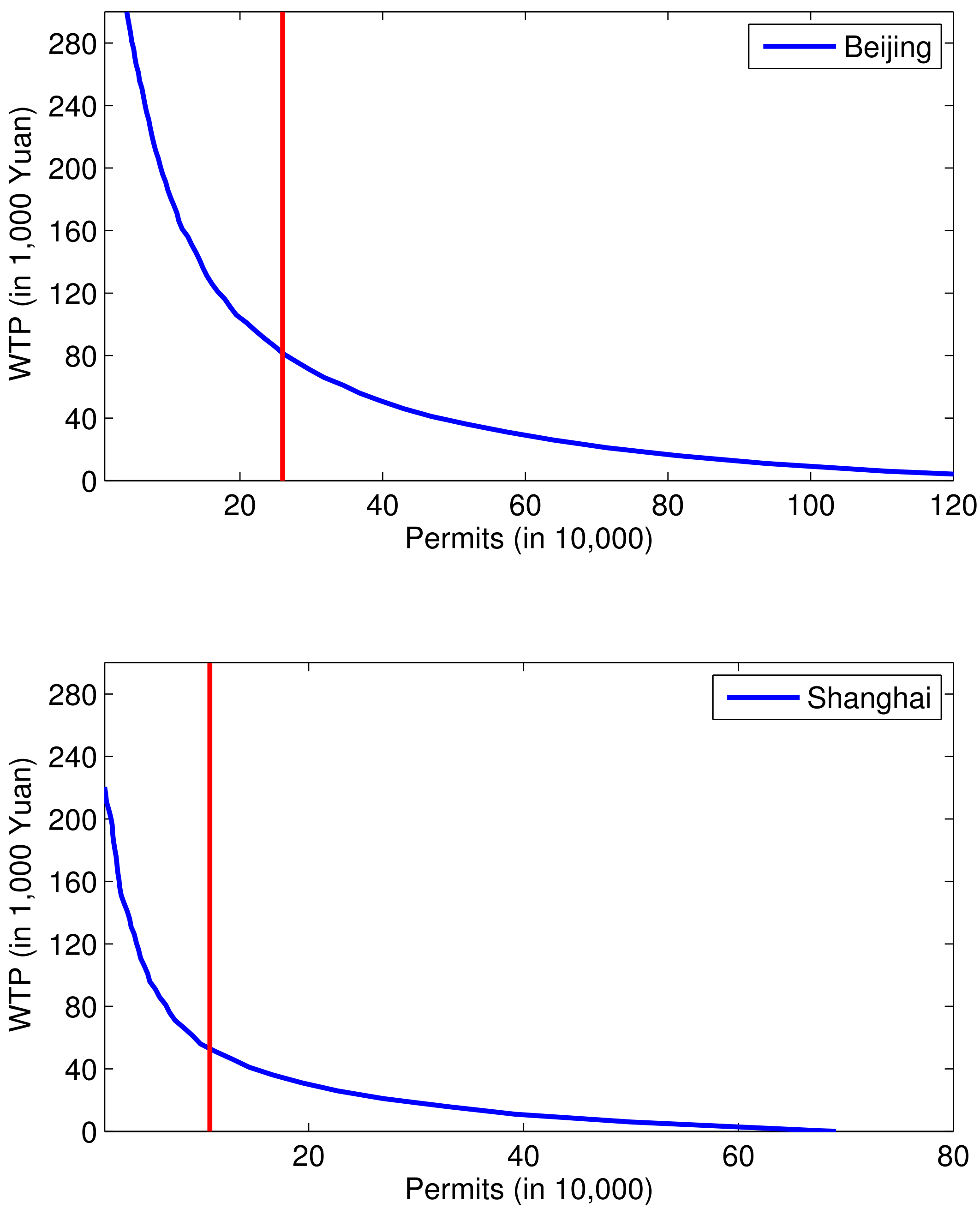
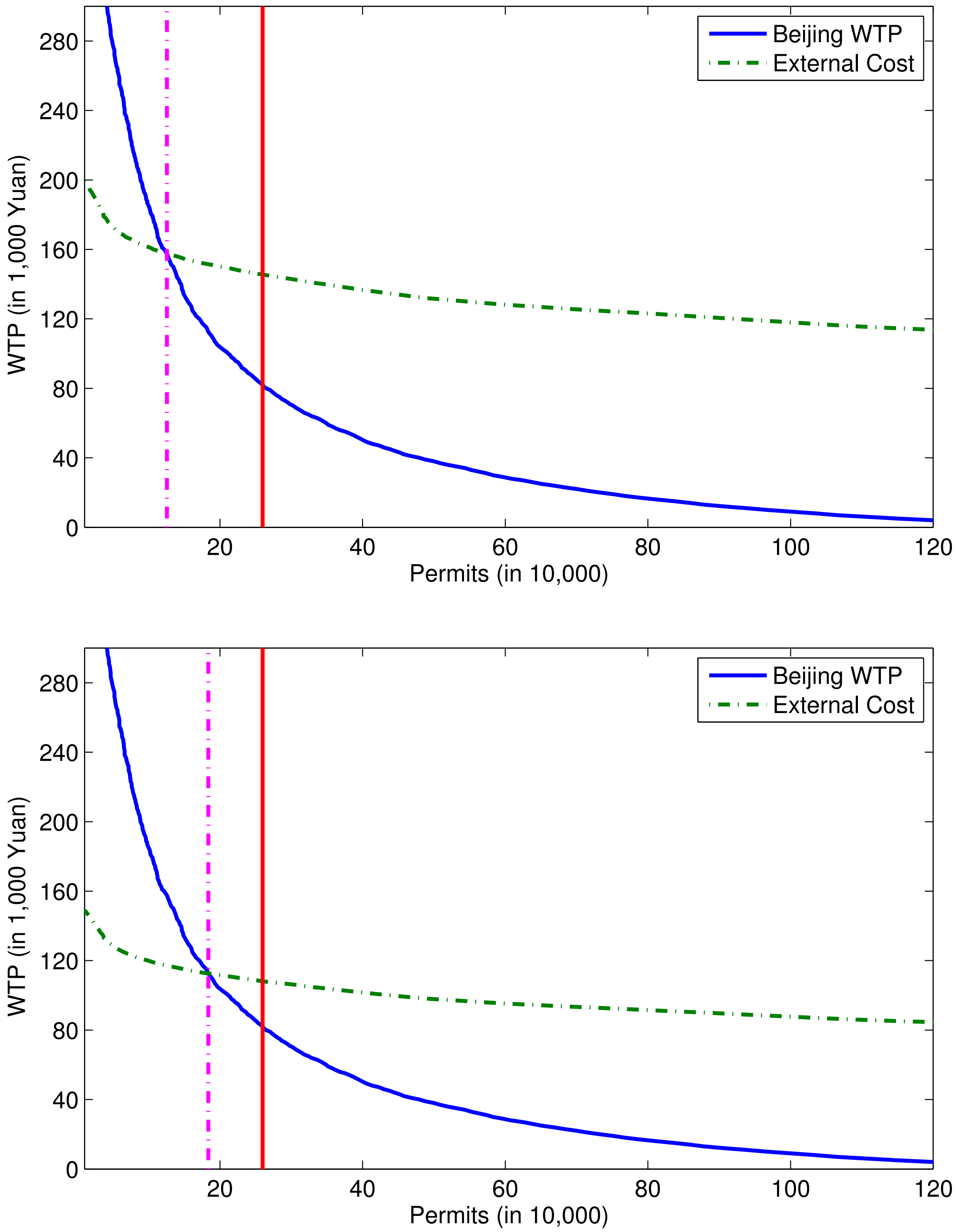


Figure 5: WTP Schedules for Licenses in Beijing and Shanghai



Notes: The top panel depicts the demand curve or the WTP schedule for licenses in Beijing in 2012 while the graph in the bottom panel is for Shanghai in 2012. The vertical lines are the quotas, 259,800 and 108, 100 in Beijing and Shanghai, respectively.

Figure 6: WTP Schedules and External Costs



Notes: The external costs depicted in the top panel (the downward sloping dashed line) are calculated over a 15-year time span of using the license while those in the bottom panel are based on a 10-year time span. The solid vertical lines denote the observed quota level of 259,800. The dashed vertical lines are the optimal levels of quota of 125,612, and 183,796 for the top and bottom, respectively.

Table 1: City Characteristics in Beijing, Nanjing, Shanghai, Tianjin

Year	City	No. of Households (mil.)	Average Household Income (in Yuan)	New vehicle Sales	National Vehicle Sales(mil.)	% in nation
2008	Beijing	605.99	69,230	419,703	6.76	6.21
2009	Beijing	636.29	74,866	610,076	10.33	5.91
2010	Beijing	668.10	81,404	815,211	13.76	5.93
2011	Beijing	687.86	88,838	346,207	14.47	2.39
2012	Beijing	704.89	98,466	536,216	15.50	3.46
2008	Nanjing	224.68	62,431	83,459	6.76	1.24
2009	Nanjing	230.84	68,861	131,316	10.33	1.27
2010	Nanjing	237.00	76,442	177,105	13.76	1.29
2011	Nanjing	240.08	86,940	166,770	14.47	1.15
2012	Nanjing	241.62	98,069	187,194	15.50	1.21
2008	Shanghai	774.12	79,225	169,678	6.76	2.51
2009	Shanghai	799.21	84,495	212,139	10.33	2.05
2010	Shanghai	825.10	92,330	268,507	13.76	1.95
2011	Shanghai	841.23	105,067	284,693	14.47	1.97
2012	Shanghai	853.06	116,545	303,095	15.50	1.96
2008	Tianjin	347.88	56,131	160,221	6.76	2.37
2009	Tianjin	356.92	61,737	223,774	10.33	2.17
2010	Tianjin	366.20	69,477	276,716	13.76	2.01
2011	Tianjin	383.35	76,455	278,336	14.47	1.92
2012	Tianjin	399.92	84,138	292,442	15.50	1.89

Notes: The sales data are from R.L. Polk & CO. and other variables are from various issues of Annual Social and Economic Development Report by each of the cities. The average income is nominal. New vehicle sales include passenger cars and light trucks.

Table 2: Summary Statistics of Vehicle Sales and Characteristics

Variable	Mean	Median	S.D.	Min	Max
Price (in 2012 RMB)	315.22	197.79	276.58	36.49	1148.38
Monthly sales by model in Beijing	125.93	31.00	240.66	1.00	3464.00
Monthly sales by model in Nanjing	34.71	9.00	65.36	1.00	964.00
Monthly sales by model in Shanghai	57.01	11.00	131.23	1.00	3036.00
Monthly sales by model in Tianjin	57.33	12.00	125.75	1.00	2295.00
Vehicle size (m^2)	8.01	8.09	1.01	4.20	10.97
Displacement (liter)	2.09	1.87	0.77	0.90	5.70
Liters per 100 kilometers	9.10	8.80	2.09	2.90	17.20
Yuan per 100 kilometers	66.17	63.62	16.10	20.61	129.95
Passenger car dummy	0.63	1.00	0.48	0.00	1.00
SUV dummy	0.22	0.00	0.41	0.00	1.00
Minivan dummy	0.15	0.00	0.36	0.00	1.00
Large dummy	0.14	0.00	0.35	0.00	1.00
Luxury dummy	0.02	0.00	0.16	0.00	1.00
Medium dummy	0.35	0.00	0.48	0.00	1.00
Mini dummy	0.03	0.00	0.17	0.00	1.00
Small dummy	0.21	0.00	0.41	0.00	1.00
Upper medium dummy	0.24	0.00	0.43	0.00	1.00

Notes: The observation is at the vehicle model-year-month level. Vehicle prices include vehicle sales tax which is 10 percent in 2008, 2011 and 2012. The tax varied across vehicles with different engine size in 2009 and 2010. There are 21,228 observations from 2008 to 2012 with 1,769 models (vintage-nameplate) in the data set.

Table 3: The Share of Households by Income among Vehicle Buyers and All Households

City	Annual Household	New Vehicle buyers					All Households				
	Income (in Yuan)	2008	2009	2010	2011	2012	2008	2009	2010	2011	2012
Beijing	Under 48k	14.24	18.93	14.92	11.8	7.89	26.69	18.24	12.35	9.34	6.7
Beijing	48k to 96k	34.71	30.73	31.11	26.15	23.67	59.82	63.3	62.68	61.67	60.03
Beijing	96k to 144k	30.73	32.71	32.82	34.47	36.51	9.63	13.64	19.41	20.15	18.93
Beijing	Over 144k	20.3	17.63	21.16	27.59	31.92	3.86	4.82	5.56	8.84	14.34
Nanjing	Under 48k	15.72	10.16	5.87	6.92	1.4	40.6	37.07	30.56	12.96	7.83
Nanjing	48k to 96k	36.75	36.57	26.08	34.15	14.16	46.39	44.55	45.26	55.32	48.46
Nanjing	96k to 144k	28.21	36.59	40.05	33.25	49.76	9.27	13.18	17.11	22.87	31.98
Nanjing	Over 144k	19.33	16.68	27.99	25.68	34.66	3.74	5.21	7.07	8.84	11.73
Shanghai	Under 48k	2.03	4.56	1.54	1.11	0.91	23.96	13.42	9.53	6.69	5
Shanghai	48k to 96k	7.86	11.82	8.79	8.94	5.9	51.67	60.21	59.02	47.93	42.42
Shanghai	96k to 144k	43.78	40.44	33.84	34.11	18.76	16.58	17.16	19.88	29	29.33
Shanghai	Over 144k	46.34	43.19	55.83	55.83	74.44	7.79	9.22	11.57	16.38	23.26
Tianjin	Under 48k	15.88	21.69	20.37	10.35	4.81	49.25	36.09	28.08	17.42	10.04
Tianjin	48k to 96k	41.78	42.35	39.45	37.55	22.87	42.49	54.31	57.9	60.95	60.51
Tianjin	96k to 144k	27.62	25.45	30.43	36.45	52.13	6.11	7.01	10.04	16.6	23.02
Tianjin	Over 144k	14.74	10.51	9.74	15.64	20.17	2.14	2.6	3.97	5.03	6.42

Notes: The income data for vehicle buyers come from national representative surveys on new vehicle buyers by Ford Motor Company. The data on all households are from Annual Statistical Yearbook for each city by Bureau of Statistics.

Table 4: Pre-policy Trend Analysis

Variables	Specification 1		Specification 2		Specification 3	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Ln(price+bid)	-4.846	0.352	-5.089	0.367	-5.095	0.368
Beijing*2009	-0.033	0.085	-0.033	0.085	-0.032	0.085
Beijing*2010	0.021	0.080	0.021	0.080	-0.026	0.082
Nanjing*2009	0.056	0.071	0.056	0.071	0.056	0.071
Nanjing*2010	0.081	0.069	0.081	0.069	0.081	0.069
Shanghai*2009	-0.054	0.094	No		No	
Shanghai*2010	-0.071	0.088	No		No	
Vintage-model fixed effects	Yes		Yes		Yes	
City-segment fixed effects	Yes		Yes		Yes	
Year-month fixed effects	Yes		Yes		Yes	
Shanghai year-month fixed effects	No		Yes		Yes	

Notes: The dependent variable is ln(market shares). Specifications 1 and 2 use all the 47,232 observations from 2008 to 2010. Specification 3 drops observations in Nov. and Dec. of 2010 in Beiijing to remove anticipation effect and has 46,462 observations. Tianjin is the base group and 2008 is the base year. The standard errors are clustered at the model level.

Table 5: Sales Impacts of the Lottery Policy in Beijing

Variables	Specification 1			Specification 2			Specification 3		
	Coef.	S.E.	Impacts	Coef.	S.E.	Impacts	Coef.	S.E.	Impacts
Ln(price+bid)	-3.213	0.174		-3.229	0.1744		-3.229	0.175	
Lottery in 2011	-0.938	0.116	-60.6%	-0.787	0.122	-54.1 %	-0.803	0.046	-55.1%
Lottery in 2012	-0.730	0.213	-50.7 %	-0.543	0.2216	-40.5%	-0.547	0.046	-42.1%
Vintage-model fixed effects	Yes			Yes			Yes		
Year-month fixed effects	Yes			Yes			Yes		
Shanghai * year-month effects	Yes			Yes			Yes		
City-segment fixed effects	Yes			Yes			Yes		
City trend (quadratic)	Yes			Yes			No		

Notes: The dependent variable is ln(market shares). Specifications 1 use all the 84,912 observations from 2008 to 2012. Specifications 2 and 3 drop the last two months in 2010 and first two months of 2011 for Beijing to remove anticipation effect. The standard errors are clustered at the model level.

Table 6: Parameter Estimates from GMM

Variables	Benchmark		Alternative 1		Alternative 2	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Parameters in the household-specific utility (θ_2)						
Ln(price+bid)	-17.519	1.927	-23.687	2.041	-8.132	2.425
Ln(income)*Ln(price+bid)	1.558	0.193	2.141	0.256	-1.215	0.403
Ln(income)	-6.442	0.897	-7.108	1.170	5.850	1.926
σ for ln(price+bid)	0.095	0.044	0.195	0.059	0.248	0.106
σ for constant	0.907	0.261	-0.011	0.133	1.164	0.543
σ for Fuel costs per 100km	0.000	0.094	0.000	0.060	0.299	0.208
σ for vehicle size	0.022	0.326	-0.020	0.155	-0.385	0.254
σ for engine displacement	1.835	0.493	2.954	0.337	2.990	0.981
Auxiliary parameters for allocation mechanism (θ_3)						
ρ	0.202	0.037	0.395	0.030	0.205	0.041
γ_0	-1.001	0.018	-0.925	0.042	-1.002	0.038
γ_1	0.252	0.005	0.136	0.006	0.248	0.007

Notes: The benchmark model is the preferred model. Alternative specification 1 does not include the first set of moment conditions (common-trend moments). Alternative 2 does not include the second set of moment conditions (micro-moments) .

Table 7: Parameter Estimates: Further Robustness Checks

Variables	Alternative 3		Alternative 4		Alternative 5	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Parameters in the household-specific utility (θ_2:)						
Ln(price+bid)	-21.315	4.133	-17.127	2.184	-16.222	1.852
Ln(income)*Log(price+bid)	1.561	0.316	1.519	0.234	1.384	0.201
Ln(income)	-6.470	1.383	-6.387	1.140	-5.648	0.985
σ for constant	1.512	0.311	0.922	0.290	0.897	0.283
σ for ln(price)	0.215	0.084	0.080	0.063	0.091	0.056
σ for fuel costs per 100km	-0.002	0.119	0.003	0.120	0.001	0.114
σ for vehicle size	-0.065	0.193	-0.137	0.498	-0.065	0.517
σ for engine displacement	3.726	1.023	1.360	0.555	1.520	0.561
Auxiliary parameters for allocation mechanism (θ_3)						
ρ	0.283	0.036	0.182	0.033	0.186	0.041
γ_0	-0.986	0.343	-1.002	0.408	-1.004	0.065
γ_1	0.225	0.078	0.253	0.103	0.255	0.017

Notes: Alternative 3 assumes all the households as the potential buyers in a year so the market size in each month is the total number of households divided by 12. Alternative 4 assumes a certain percentage of licenses are used to buy used vehicles: 15 percent in Beijing and 5 percent in Shanghai. Alternative 5 takes the random draws for unobserved household attributes from a standard normal distribution removing draws below 2.5 and above 97.5 percentiles.

Table 8: Policy Impacts on Sales in Beijing and Shanghai

Year	Observed Sales	Counterfactural Sales		
		Benchmark	Alternative 1	Alternative 2
Beijing				
2008	411,936	411,936	411,936	411,936
2009	602,219	602,219	602,219	602,219
2010	804,355	804,355	804,355	804,355
2011	334,308	779,273	591,789	774,600
2012	520,442	1,142,064	884,541	1,114,074
Shanghai				
2008	165,298	344,955	322,379	340,356
2009	208,570	414,994	417,469	399,754
2010	264,232	577,542	538,468	567,164
2011	277,119	599,425	603,846	597,196
2012	295,047	712,215	681,951	688,113

Notes: These three specifications correspond to those in Table 6.

Table 9: Policy Impacts on Sales: Further Robustness Checks

Year	Observed Sales	Counterfactural Sales		
		Alternative 3	Alternative 4	Alternative 5
Beijing				
2008	411,936	411,936	411,936	411,936
2009	602,219	602,219	602,219	602,219
2010	804,355	804,355	804,355	804,355
2011	334,308	691,476	800,655	797,041
2012	520,442	1,004,969	1,177,813	1,170,655
Shanghai				
2008	165,298	377,179	346,277	340,970
2009	208,570	453,940	418,868	412,830
2010	264,232	651,788	575,782	568,714
2011	277,119	603,463	609,972	599,961
2012	295,047	779,833	723,197	712,205

Notes: These three specifications correspond to those in Table 7.

Table 10: Welfare Comparison Between Lottery and Auction in Beijing

	15-year Horizon		10-year Horizon	
	Lottery	Auction	Lottery	Auction
	(1)	(2)	(3)	(4)
Observed Quota Level				
Quota in 2012	259,800	259,800	259,800	259,800
Clearing price (in 1000 RMB)	0.00	82.21	0.00	82.21
Realized consumer surplus (in billion)	5.39	48.35	5.39	48.35
Unrealized consumer surplus (in billon)	42.96	0.00	42.96	0.00
Average fuel economy (liters per 100km)	8.92	9.19	8.92	9.19
Annual VMT (1000 kms)	17.76	19.80	17.76	19.80
Total external costs (in billion)	33.02	40.21	24.57	29.92
Government revenue (in billion)	0.00	21.36	0.00	21.36
Net social welfare (in billion)	-27.63	8.13	-19.18	18.43
Optimal Quota Level				
Optimal quota		125,612		183,796
Clearing price (in 1000)		157.50		112.50
Consumer surplus (in billion)		34.87		42.60
Total external costs (in billion)		20.75		21.82
Government revenue (in billion)		19.78		20.68
Net social welfare (in billion)		14.12		20.79

Notes: All monetary variables are in RMB ($1 for about 6.1 RMB). The top panel presents welfare comparison between a lottery system and a uniform price auction for Beijing in 2012. The bottom panel shows welfare outcomes with the optimal levels of quota. The optimal quotas are defined by the interactions between the vertical dashed line and the downward sloping dashed line in the two panels of Figure 6. Net social welfare is equal to realized consumer surplus minus total external costs. Columns (1) and (2) use a 15-year time span for calculating total external costs, i.e., discounted external costs during 15 years with a discount rate of 5%. Columns (3) and (4) are results based on a 10-year time span for external costs.