Range-Based Subsidies and Product Upgrading of Battery Electric Vehicles in China *

Xiaoyue Zhang†

May 2, 2023

The Latest Version

Abstract

This paper estimates the impact of notched driving-range-based (DRB) subsidies to consumers on Chinese battery-electric-vehicle (BEV) manufacturers' incentives to reduce their production costs of driving ranges. Chinese consumers received generous subsidies if the driving ranges were above certain thresholds. Using a dynamic structural model to infer unobserved investment decisions on the cost reduction of driving ranges by manufacturers, I find that the discontinuous incentives around the range thresholds of Chinese DRB subsidies increased Chinese BEV manufacturers' probability of investing in reducing production costs of driving ranges in 2019 by up to 50 percentage points. Compared with counterfactual DRB subsidies that are linear in driving ranges and provide the same total amount of subsidy in 2019, the notched scheme induces lower investment probabilities but is a lot cheaper in future periods for the government. This dynamic impact on production costs implies that the environmental benefits and welfare gains of notched DRB subsidies are very likely larger than the estimates of existing literature. It also implies that notched subsidies can be used to induce technological adoption or product upgrading.

Keywords: Industrial policy, product subsidy, technological catch-up, electric vehicle, dynamic discrete choice model

*I am indebted to Jaap Abbring, Jeffrey Campbell, and Christoph Walsh for their patience, support and guidance throughout this project. This paper has benefited from insightful discussions with and comments from Nicola Pavanini, Jo van Biesebroeck, Frank Verboven, Reyer Gerlagh, Sjak Smulders as well as participants at the Tilburg Environmental Group. I am also thankful for the never-ending support and confidence from Edi Karni. This project is financially supported by CentER at Tilburg University.

†Tilburg University, email: x.zhang_11@uvt.nl

1 Introduction

This paper estimates the impact of driving-range-based (DRB) subsidies to consumers on Chinese battery-electric-vehicle (BEV)[1] manufacturers' incentives to reduce their production costs of driving ranges. Under the pressure of negative growth of automobile sales due to the 2008 financial crisis, the Chinese government implemented a package of policies in 2009 to revitalize its automobile industry through promoting the innovation by its domestic EV manufacturers.[2] The DRB subsidies to EV consumers were a crucial part of this ambition.[3] However, little is known about whether and how the DRB subsidies contribute to the 14 times increase in the Chinese EV sales in 2016-2022 and its domestic producers' newly-acquired presence in the global market in 2022. In fact, Chinese EV sales account for 65% of the global EV sales in 2022.[4] More specifically, it is unclear whether the Chinese DRB subsidies induced Chinese EV manufacturers to actively reduce their production costs of driving ranges, such as building more energy-efficient models or improving their supply chains of batteries, or whether the manufacturers only responded by installing more batteries and building smaller cars to increase driving ranges.

Using a dynamic structural model to explain and infer firms' unobserved decisions regarding the cost reduction of driving ranges, I find that Chinese DRB subsidies, which offer a generous amount of subsidies to consumers once the driving ranges are above certain range thresholds, i.e. a notched subsidy scheme, increased Chinese BEV manufacturers' probabilities of investing in reducing production costs of ranges in 2019 by up to 50 percentage points relative to the scenarios of no subsidies. In the latter, the invest probabilities of firms most likely to invest are less than 25%. Compared with counterfactual DRB subsidies that are linear in driving ranges and provide the same total amount of subsidy in 2019, the notched scheme induces lower investment probabilities but is a lot cheaper in the future periods for the government. This result implies that the environmental benefits and welfare gains of notched DRB subsidies are very likely larger than the estimates in existing literature that ignores this channel of reducing future production costs of ranges. The machinery built to estimate the impact on the production costs of ranges offers a possible tool for future research on devising better DRB subsidies for incentivizing cost reductions over ranges.

[1]BEVs are a type of electric vehicles who relies exclusively on rechargeable battery packs, without a secondary source of propulsion such as an internal combustion engine.

[2]The announcement is aimed at new energy vehicles, which are mainly BEVs and plug-in hybrid electric vehicles. Source: http://www.gov.cn/zwgk/2009-03/20/content_1264324.htm

[3]Source: https://www.chinanews.com.cn/ny/2010/10-13/2583130.shtml

[4]Sources: https://www.marklines.com, https://www.ev-volumes.com/, and http://ex.chinadaily.com.cn/exchange/partners/82/rss/channel/language/columns/v0m20b/stories/WS6389a97da31057c47eba25e3.html

From 2016 to 2020, BEVs in China became cheaper with longer ranges and used higher-density batteries, but there was no clear trend of a decline in BEV model sizes. Furthermore, there is large variation in prices and driving ranges in each given year, and the trend of changing prices and ranges varied across models and manufacturers. Underlying changes in battery technology alone cannot explain the variation. Possible explanations for the variation are different BEV model characteristics such as whether it is a high-end model, manufacturers' capability of exercising market power, and manufacturers' decisions to reduce their production costs of ranges. Manufacturers can reduce their production costs by designing more energy efficient models, so that the same BEV model with the same battery capacity can travel a longer distance. Manufacturers can also reduce their production costs by improving their exposure to the technological progress in batteries, such as sourcing better battery suppliers, vertically integrating with battery suppliers, or carrying out in-house R&D. I use a dynamic structural model to disentangle the channel of a reduction in the production costs of ranges from the remaining channels so that I can quantify the impact of Chinese DRB subsidies on manufacturers' decisions to reduce the production costs of ranges. I focus on BEVs because they are the most targeted EVs by Chinese DRB subsidies and have the largest sales among all the EVs.

In my structural model, homogeneous consumers make their purchase decisions in each period based on observed and unobserved model characteristics, and unobserved consumer-model-period-specific taste shocks. They can also choose not to buy a car. Manufacturers maximize their expected sum of current and discounted future profits by choosing the optimal prices and ranges in each period and make dynamic decisions on whether to reduce the production costs of ranges in the next period under an adaptive expectation about future prices and ranges of other BEV models and future subsidy schemes. If they decide to reduce the production costs, they pay an investment cost in this period. Their production costs of ranges are lower in the next period by one step with certainty and will stay at this value until further investments are made in the future periods. I assume that firms have rational expectations about the current prices and ranges and observe the current subsidy scheme, but have adaptive expectations about future prices, ranges, and subsidy schemes because predicting what happens in the current year is likely easier and more reliable than predicting what will happen in all future years with an infinite horizon. In the rapidly growing BEV industry, it is likely challenging to predict what happens next year.

I follow Berry et al. (1995) and Crawford et al. (2019) by assuming that unobserved consumer-model-period-specific taste shocks are independent and identically distributed so that I can use BEV-model-level market shares in each year to estimate parameters of consumer preferences. One period in the structural model is one year in the data. I parameterize manufacturers' production costs as linear in exogenous observed and unobserved model characteristics and a polynomial of the endogenous control variable, the driving range. The first order conditions of manufacturers' profit maximization together with BEV model prices and consumers' preference parameters give the estimated production cost functions. To account for endogeneity caused by the unobserved model characteristics in both demand and supply, I construct instruments following Berry et al. (1995).

Since the driving range is an endogenous control variable, my estimation produces BEV-model-year-specific cost parameters of ranges like Crawford et al. (2019). I infer manufacturers' investment decisions from the changes in the cost parameters of ranges. A reduction in a BEV models' cost parameters this period means that its manufacturer invested in reducing the production cost of ranges in the previous period for this BEV model. A manufacturer invests when the increase in its expected discounted sum of future profits is larger than the investment cost. If a manufacturer has multiple BEV models, the manufacturer is assumed to make individual investment decisions for each of its BEV models not taking into account the impact of the investment decisions on its other BEV models. The investment cost is unobserved, and is estimated from the inferred investment decisions and from manufacturers maximizing their expected sum of discounted profits under their beliefs about the future. The investment cost is assumed to be constant across firms and time.

I estimate the investment cost by solving a non-stationary dynamic problem with an infinite time horizon for each BEV model. It is non-stationary because the subsidy scheme changes over time. It has an infinite time horizon because firms are assumed to live forever. To solve these dynamic problems, I discretize the estimated cost parameters of ranges and assume that the space of the cost parameters is a finite set with a known lower bound, that the cost parameter of ranges decreases by 1 unit when a firm invests and otherwise stays constant, and that there are only a finite amount of possible subsidy schemes. These assumptions allow me to solve for the optimal investment policies in the dynamic problems by recursion over the state space of the cost parameters from the lowest value to the highest, similar to the solution for finite-state directional dynamic games proposed by Iskhakov et al. (2016). To simplify the dynamic problem, I assume firms to have an extreme type of adaptive expectations about the future in the sense that the prices and ranges of other models and the subsidy scheme are believed to stay forever at their current values. This simplifies the optimal investment policies for each model in each period to the solution to an auxiliary single-agent stationary dynamic problem. Using these optimal investment policies and the inferred investment decisions, I can estimate the investment cost using maximum likelihood estimation, similar to Rust (1987)'s estimation for the bus engine replacement problem. However, Rust (1987) observes the dynamic decisions, i.e. whether an engine is replaced, while I infer the dynamic decisions from the estimation of the static part of the problem. The lower bound is low enough so that no firm in my data managed to reach it. Firms do not foresee themselves surpassing this lower bound due to, for example, the difficulties in predicting the changes in the technological frontier in the future, or due to current technology bottlenecks that prevent the cost parameters from dropping below certain values.

Applying my structural model to the data, which is a newly constructed rich dataset that contains national-level vehicle model sales in 2010-2021 and the technical description of all vehicle model variations available or once available on the market,[5] I find that the existing notched investment scheme in 2019 increased the BEV manufacturers' investment probabilities by up to 50 percentage points.

I assume that consumers are homogeneous. Although consumers' heterogeneous tastes and price sensitivity may contribute to the dispersion in prices and ranges in the sense that different models may target different types of consumers, assuming away consumer heterogeneity allows for analytical theoretical results, provides better intuition, and reduces the computational burden in the dynamic part. I argue that compared to the technological progress in electric vehicles, consumers' tastes are relatively more stable over time. Therefore, consumer heterogeneity cannot explain the variation in the model-specific trends in ranges and prices. Because this variation is used for identifying the changes in the production costs of ranges and the investment cost, assuming away consumer heterogeneity should have a limited impact on the estimated investment cost.

This paper provides the first empirical evidence that notched attribute-based subsidies can produce efficiency gains. Existing literature on attribute-based subsidies, such as Ito and Sallee (2018) and Jia et al. (2022), almost universally criticizes notched subsidy schemes for the distortions they create around the thresholds and therefore recommend against notched ones for the sake of efficiency. My results show that the discontinuity in firms' profits around the thresholds can incentivize firms to put more effort in reducing production costs and consequently creates efficiency gains.

This paper contributes to the literature on the effects of industrial policies on innovation and technological progress by providing empirical evidence on whether and how subsidies on the demand side can encourage innovation and technological upgrading. Studies have shown that reducing the costs of investment can raise R&D (Takalo et al. (2013) and Criscuolo et al. (2019)) because this increases the net returns to investment. In theory, such an increase in net returns can also be achieved by promoting demand. The field experiments in Bold et al. (2022) show that higher demand can encourage farmers' technological adoptions, suggesting

[5]The sources are Chezhu Home (https://www.16888.com) and Auto Home (https://www.autohome.com.cn)

that policies targeting the demand side can be effective. However, there is no empirical evidence that the same holds in the EV market and this paper fills that gap.

The methodological contribution of this paper is to combine Berry et al. (1995) and Crawford et al. (2019)'s model for structurally estimating demand and supply using aggregate data, such as automobile models' annual sales at the national level rather than individual consumer's decisions, with the literature of dynamic discrete choice models (Rust (1987)) where structural model primitives of dynamic decisions are estimated using data on observed individual decisions, so that one can do counterfactual analysis that involves dynamic decision making without observing the individual dynamic decisions.

The remainder of the paper is organized as follows. I present historical background and stylized facts in Section 2 and describe the dataset in Section 3. I then introduce the model and demonstrate theoretical results in Section 4. I explain the estimation procedure in Section 5. Results are provided in Section 6. Section 7 concludes. Appendix explains how I merge data from multiple sources to one panel.

2 Historical background and stylized facts

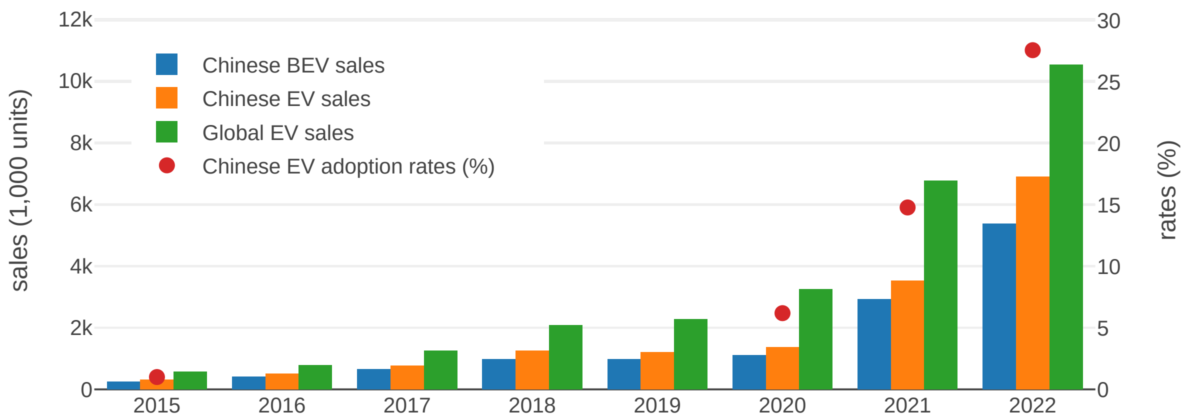
China's ambition of improving its global presence in automobile manufacturing dates back to the 1980s. Despite a large amount of resources spent in the automobile industry, China was far from gaining prominence in the global automobile manufacturing until the 2020s. Under the pressure of reduced automobile sales due to the shocks from the 2008 financial crisis, China decided in 2009 to pursue its ambition by promoting electric vehicle (EV) manufacturing, especially the innovation by domestic EV manufacturers.[6] To achieve this goal, the Chinese government unleashed a series of industrial policies. One important part of these policies was offering a series of generous range-based subsidies to consumers with the aim of stimulating the domestic EV manufacturers' product upgrading.[7] Since then, Chinese domestic EV sales started to grow at an increasing rate. The sales increased more than 20 times in 2015-2022, reaching 65% of the global EV sales in 2022. The Chinese adoption rate, i.e. the ratio of new EV sales to total new vehicle sales, increased from 1% in 2015 to 28% in 2022 (Figure 1). Li et al. (2022) shows that this rapid growth of the EV sales in China was largely due to the generous consumer subsidies. Furthermore, Chinese domestic EV manufacturers also started to gain a global presence in recent years. For example, Chinese EV export almost doubled in 2022 compared with 2021.[8].

[6] Source: http://www.gov.cn/zwgk/2009-03/20/content_1264324.htm

[7] Source: https://www.chinanews.com.cn/ny/2010/10-13/2583130.shtml

[8] Source: http://ex.chinadaily.com.cn/exchange/partners/82/rss/channel/language/columns/v0m20b/stories/WS6389a97da31057c47eba25e3.html

Figure 1: Trends of BEV and EV sales



Notes: Adoption rates are the ratios of new EV sales to total new vehicle sales. BEVs and EVs include both passenger cars and commercial cars.
Source: The Chinese annual sales of BEVs and EVs in 2015 are from CAAM and in 2016-2022 are from www.marklines.com. Chinese adoption rate in 2015 is calculated by the author using the Chinese EV sales from CAAM mentioned above and the total Chinese vehicle sales from www.marklines.com. The adoption rates in 2020 are from www.ce.cn. Those in 2021 and 2022 are from CPCA. Global EV sales are from www.ev-volumes.com.

Table 1: Range-based subsidy for BEVs by the central government (10,000 RMB)

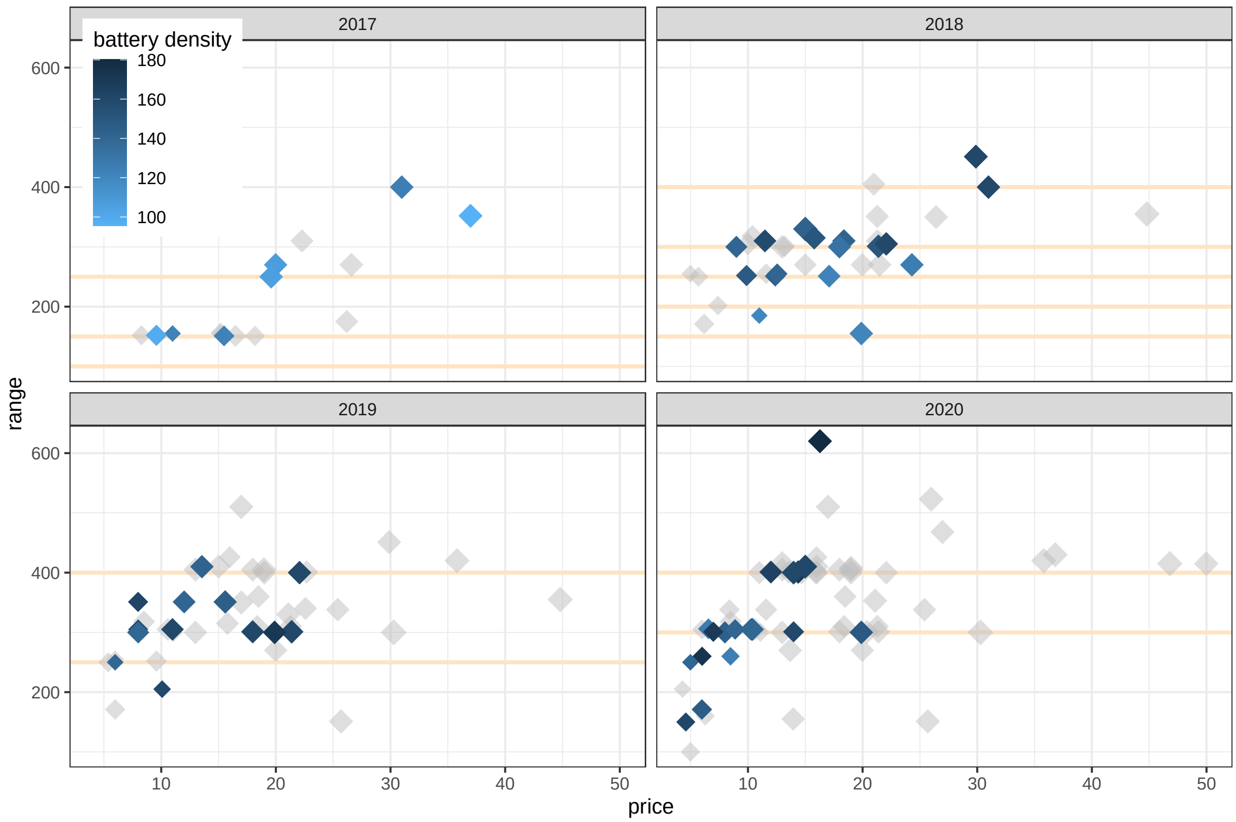
range (km)	2016	2017	2018	2019	2020	2021	2022
$[0, 100)$	0	0	0	0	0	0	0
$[100, 150)$	2.5	2	0	0	0	0	0
$[150, 200)$	4.5	3.6	1.5	0	0	0	0
$[200, 250)$	4.5	3.6	2.4	0	0	0	0
$[250, 300)$	5.5	4.4	3.4	1.8	0	0	0
$[300, 400)$	5.5	4.4	4.5	1.8	1.62	1.3	0.91
$[400, \infty)$	5.5	4.4	5	2.5	2.25	1.8	1.26

Table 1 demonstrates the range-based subsidies to consumers purchasing BEVs. Here I only list the subsidies of BEVs because they are the focus of this paper. It is a notched subsidy scheme because the amount of subsidies consumers receive is a discontinuous function of the driving ranges. Both the thresholds of the driving ranges and the amount of subsidies for a given threshold change over time. In general, the thresholds become higher and the subsidies become smaller. While the subsidies that consumers receive decline in 2016-2021, the annual sales of BEV in China increases as shown in Figure 1.

Figure 2 demonstrates prices (10,000 RMB), ranges (km), model sizes (length (m) $\times$ width (m)), and battery density (battery capacity/battery weight (wh/kg)) of new BEV models or new releases of existing BEV models in 2017-2020. I do not include 2016 because there were too few BEV models. The orange lines are the thresholds introduced in that year. Each diamond represents a BEV model. This figure shows that BEV models become cheaper with longer ranges, higher-density batteries, but almost no change in sizes, and that there is large variation in the trend of prices and ranges over the years.

There are several possible reasons behind the large variation: differences in BEV model characteristics other than ranges, different markups in the model prices, and different production costs of ranges. Examples of different production costs of ranges are how energy efficient are the designs of models in the sense that whether a model with the same size can travel over a longer distance, and different access to existing battery technology. In the next two sections, I will build a model that takes into account all these channels and disentangle the channel of differences in production costs of ranges from the remaining channels and then quantify how firms' production costs of ranges respond to the consumer subsidies.

FIGURE 2: Trends of BEV attributes in China



Notes: Battery density is the unit of kw/kg. Ranges are in the unit of km. Prices are measured in 10,000 RMB. The sizes of the diamonds are the sizes of the BEV model which are measured as length (m) × width (m). The color gray means no information on battery density is available. Orange lines are the range thresholds of the notched subsidy scheme introduced in that year.
Source: Auto Home

3 Data

I collect sales data from Chezhu Home (https://www.16888.com) and technical description from both Chezhu Home (https://www.16888.com) and Auto Home (https://www.autohome.com.cn). Additional variables about BEVs, such as battery energy density (the ratio between battery capacity and battery weight), are collected using the Lists of Recommended New Energy Vehicles (hereafter the NEV Lists) published by the Chinese government in 2017-2020, which are in total 49 lists. These variables are required when calculating each product's subsidy value.

In addition to the direct purchase subsidy, the exemption of purchasing tax is also an important factor for consumers' purchase decisions. I use the Lists of NEVs Eligible for Purchasing Tax Exemption (hereafter the NEVPTE List) and the Lists of NEVs Removed from the NEVPTE Lists (hereafter the RNEVPTE List) published by the Chinese government in 2016-2020 to decide whether a product receives tax exemption in a year. There are 32 NEVPT Lists and 10 RNEVPTE Lists in 2016-2020. The tax exemption is considered when estimating the demand parameters.

I do not observe the exact amount of subsidy per transaction, but I collect the formulas for calculating subsidies announced by the Chinese government in 2010-2020 and then calculate the subsidies using product characteristics accordingly. I take the number of households and consumer price index in each year in 2010-2020 from the Chinese Year Books.

According to the State Grid Corporation of China, the monopolistic electricity supplier in China, the electricity price remains at 0.542 RMB/kwh during my sample. I use the prices of 92 and 95 gasoline, and diesel in Beijing 2010-2021 from https://data.eastmoney.com/cjsj/oil_default.html to approximate the national average prices.

I limit my sample to passenger vehicles with no more than 5 doors and include only BEVs and gasoline vehicles. Table 2 shows the summary statistics of all the products in my sample in 2010-2020. The observation unit is a model in a year. Table 3 shows the summary statistics of BEVs. Since there were no passenger BEVs before 2012 in my data, Table 3 only covers 2012-2020. The data described in Table 2 and Table 3 is used for the estimation in the static part. Table 4 gives the summary statistics for BEVs in 2019. There were 44 BEV products in 2019. To simplify the estimation in the dynamic part, I use only firms in 2019 for that part of the estimation.

Table 2: Summary statistics of the entire sample (2010-2020)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
price [1]	2,138	11.57	6.75	3.04	7.37	13.85	53.80
power/weight (kw/kg)	2,138	0.10	0.01	0.05	0.09	0.11	0.17
cost per km (RMB)	2,138	0.40	0.11	0.05	0.33	0.46	0.91
size (m^2) [2]	2,138	8.24	0.71	5.30	7.79	8.72	10.27
torque (N·m)	2,138	200.61	61.88	88	150	240	553
luxury level[3]	2,138	13.57	20.24	1	4.0	14.4	151

Each observation is a model-year.
[1] Prices are deflated by the annual consumer price index and are in units of 10,000 RMB.
[2] Size is measured using length (m) × width (m).
[3] Luxury level is an index constructed as the sum of several dummy covariates such as whether the vehicle model has a rain sensor or a key-less start.

Table 3: Summary statistics of BEVs (2012-2020)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
range (km)	171	304.38	94.16	80.00	255.00	380.00	620.00
price[1]	171	14.15	7.29	3.35	8.79	17.25	39.03
price$-\tau$[1]	171	13.24	7.56	2.27	7.37	16.79	39.03

[1] Prices and subsidies (τ) are deflated by the annual consumer price index and are in units of 10,000 RMB. Purchasing tax is also deducted in price-τ if the product is eligible for tax exemption.

Table 4: Summary statistics of BEVs (2019)

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
range (km)	44	330.64	72.04	151	300	400	510
price [1]	44	13.47	6.61	4.31	8.71	16.92	35.86
price$-\tau$[1]	44	12.44	7.20	2.27	6.46	15.40	35.86

[1] Prices and subsidies (τ) are deflated by the annual consumer price index and are in units of 10,000 RMB. Purchasing tax is also deducted in price-τ if the product is eligible for tax exemption.

4 Model

We start with the static part of the structural model which is similar to the setup in Berry et al. (1995) and Crawford et al. (2019). Starting from this section, a manufacturer is called a firm. To separate the usage of model as a vehicle model from the usage as a structural model, a vehicle model will henceforth be called a product.

4.1 Demand

Consumers maximize their indirect utility by deciding whether to purchase a product, i.e. a car, and which product to purchase if they decide to purchase one. If consumer i chooses to buy a product j from firm f in the year t, the indirect utility from such a purchase is U_{ijft}:

$$U_{ijft} = \delta_{jft} + \epsilon_{ijft} \ , \ \text{ for product } j \text{ from manufacturer } f \text{ at time } t$$

which consists of the mean utility of purchasing this product δ_{jft} and consumer i's taste shock for this product in the year t, ϵ_{ijft}. Since consumers are homogeneous, the mean utility is product-year specific and is the same for all the consumers. I assume δ_{jft} to be:

$$\delta_{jft} = \alpha p_{jft} + x'_{jft}\beta + \beta^R R_{jft} \cdot \mathbb{1}[\text{j is a BEV}] + \xi_{jft} + \zeta_t + \iota_{jf}$$

where α and β are parameters of consumers' price sensitivity and tastes for observed model characteristics, x_{jft}. Model-and-year two-way fixed effects are ζ_t and ι_{jf}. $\mathbb{1}[\cdot]$ takes the value 1 if the statement inside is true and 0 otherwise. For BEVs, β^R measures consumers' marginal utility of driving ranges. ξ_{jft} is consumers' utility derived from product j's unobserved characteristics in the year t.

If the consumer chooses not to purchase a product or, in other words, chooses an outside option, the indirect utility is U_{i0t} with mean normalized to 0:

$$U_{i0t} \ = \epsilon_{i0t} \ , \ \text{ for the outside option}$$

where ϵ_{i0t} is a mean-zero taste shock. All the taste shocks, ϵ_{ijft} and ϵ_{i0t}, are independent and identically distributed type-I extreme value with mean zero.

Demand for product j from firm f in the year t is

$$N_t \cdot \frac{\exp(\delta_{jft})}{1 + \sum_{j,f} \exp(\delta_{jft})}$$

where N_t is the total number of households in the year t, which is my measure of the number of consumers considering whether to purchase a car in the year t and, if so, which model to purchase. This measure of the total number of consumers is the same as Berry et al. (1995). This product's market share is

$$s_{jft} = \frac{\exp(\delta_{jft})}{1 + \sum_{j,f} \exp(\delta_{jft})} \tag{1}$$

4.2 Supply

Firm f's profits at time t, π_{ft}, are the sum of all its products' profits:

$$\pi_{ft} = \sum_{j \in \mathcal{J}_{ft}} N_t s_{jft}(p_{jft} - mc_{jft} + \tau(R_{jft}))$$

where $\mathcal{J}_{ft}$ is the set of firm f's products at time t, N_t is the total amount of household, s_{jft} is product j's market share, p_{jft} is the price, and $\tau(R_{jft})$ is the subsidy product j receives. $\tau(\cdot)$ represents the subsidy scheme that firms face and it specifies the amount of subsidy a product can receive based on its range R_{jft}. The marginal cost functions are parametrized as linear in exogenous model characteristics and the endogenous range squared:

$$mc_{jft} = c_{jft} R^2_{jft} \cdot \mathbb{1}[\text{j is a BEV}] + w'_{jft}\mu + \eta_{jft}$$

If a model is a gasoline vehicle, then $\mathbb{1}[\text{j is a BEV}] = 0$ and range R_{jft} does not enter the marginal cost functions. The vector w_{jft} represents the observed covariates, which include power-weight ratio, size, luxury level, miles per gallon (equivalent),[9] and torque. η_{jft} captures the unobserved model characteristics that affect production costs.

Firms maximize the sum of their current and discounted expected future profits by choosing the optimal prices and ranges in each period, and by deciding whether to pay a fixed investment cost λ for each of its products this period t, to reduce the cost parameter of ranges, $c_{jf,t+1}$, of its products next period by 1 step for sure. The cost parameter c_{jft} of range takes a value from a known finite set $\mathcal{C} = \{c_1, c_2, \cdots, c_L\}$ that satisfies $c_{l-1} < c_l$ and $\log(c_l) - \log(c_{l-1})$ constant for $l \in \{2, 3, ..., L\}$. This means the reduction in c_{jft} due to investment is lower for lower values of c_{jft}. If a model's current c_{jft} takes the value c_l and this firm invests, then in the next period, $c_{jf,t+1}$ decreases by 1 step for sure to c_{l-1}. Because c_{jft} can either stay constant or become smaller, this assumption of $\mathcal{C}$ being a finite set allows me to solve the dynamic problem using backward induction over the state space $\mathcal{C}$. Since

[9]Miles per gallon (mpg) is the distance, measured in miles, that a gasoline car can travel per gallon of fuel. Miles per gallon equivalent (mpge) is the electric vehicle version of mpg, which is measured as the distance an EV can travel on 33.7kWh of electricity

prices and ranges are set in every period, they are static optimization problems. The first order conditions (FOCs) of prices and ranges are:

$$\frac{\partial \pi_{ft}}{\partial p_{jft}} = s_{jft} + \sum_{k \in \mathcal{J}_{ft}} (p_{kft} - mc_{kft} + \tau(R_{kft})) \frac{\partial s_{kft}}{\partial p_{jft}} = 0 \tag{2}$$

$$\frac{\partial \pi_{ft}}{\partial R_{jft}} = s_{jft} \left(-\frac{\partial mc_{jft}}{\partial R_{jft}} + \frac{\partial \tau(R_{jft})}{\partial R_{jft}} \right) + \sum_{k \in \mathcal{J}_{ft}} (p_{kft} - mc_{kft} + \tau(R_{kft})) \frac{\partial s_{kft}}{\partial R_{jft}} = 0 \tag{3}$$

I do not include a linear term of R_{jft} because it is difficult to identify the cost parameters of ranges when there are two parameters related to ranges in the marginal cost function.

If $\tau(\cdot)$ is a notched scheme, $\tau(\cdot)$ is discontinuous. The optimal range may not satisfy the FOC in Equation (3) because $\frac{\partial \tau(R_{jft})}{\partial R_{jft}}$ does not exist at the thresholds. For notched schemes, $\frac{\partial \tau(R_{jft})}{\partial R_{jft}} = 0$ when R_{jft} is not at the thresholds. Therefore, for notched scheme, I first find the ranges that satisfy the Equation (3) where $\frac{\partial \tau(R_{jft})}{\partial R_{jft}} = 0$. These solutions are called interior solutions. If these interior solutions' ranges are smaller than the thresholds, I calculate the profits at these interior solutions' ranges and the profits if firms set their ranges at the range thresholds, i.e. the corner solution. The one that gives higher profits is the optimal range.

Using the market shares from the demand model, I can derive the responses of market shares to changes in prices and ranges:

$$\frac{\partial s_{kjt}}{\partial d_{jft}} = \begin{cases} -s_{kft} s_{jft} \frac{\partial \delta_{jft}}{\partial d_{jft}} \text{ , if } k \neq j \\ (1 - s_{jft}) s_{jft} \frac{\partial \delta_{jft}}{\partial d_{jft}} \text{ , if } k = j \end{cases}$$

and $d_{ift} \in \{p_{ift}, R_{ift}\}$, $\frac{\partial \delta_{ift}}{\partial p_{ift}} = \alpha$, $\frac{\partial \delta_{ift}}{\partial R_{ift}} = \beta^R$.

Combining the formulas for $\frac{\partial s_{kjt}}{\partial p_{jft}}$ and $\frac{\partial s_{kjt}}{\partial R_{jft}}$ with Equations (2) and (3) gives:

$$\begin{aligned} 0 = & 1 + \sum_{k \in \mathcal{J}_{ft}, k \neq j} (p_{kft} - mc_{kft} + \tau(R_{jft})) \cdot (-s_{kft} \alpha) \\ & + (p_{jft} - mc_{jft} + \tau(R_{jft})) \cdot (1 - s_{jft}) \alpha \end{aligned} \tag{4}$$

$$\begin{aligned} 0 = & \left(-\frac{\partial mc_{jft}}{\partial R_{jft}} + \frac{\partial \tau(R_{jft})}{\partial R_{jft}} \right) + \sum_{k \in \mathcal{J}_{ft}, k \neq j} (p_{kft} - mc_{kft} + \tau(R_{jft})) \cdot (-s_{kft} \beta^R) \\ & + (p_{jft} - mc_{jft} + \tau(R_{jft})) \cdot (1 - s_{jft}) \beta^R \end{aligned} \tag{5}$$

Combining these two equations and using the parametric form of marginal costs gives:

$$-2c_{jft} R_{jft} + \tau'(R_{jft}) - \frac{\beta^R}{\alpha} = 0 \tag{6}$$

where $\tau'(R^*_{jft}) = \frac{\partial \tau(R_{jft})}{\partial R_{jft}}$.

If the market share of each individual product is very close to zero, i.e. $s_{ijt} \approx 0$, which turns out to be the case in my data, the FOCs in Equations (4) and (5) can be approximated as:

$$
\begin{aligned}
0 \approx & 1 + (p_{jft} - mc_{jft} + \tau(R_{jft})) \cdot \alpha \\
0 \approx & -\left(\frac{\partial mc_{jft}}{\partial R_{jft}} + \frac{\partial \tau(R_{jft})}{R_{jft}}\right) + (p_{jft} - mc_{jft} + \tau(R_{jft}))\beta^R
\end{aligned}
$$

So firm j's optimal profits and optimal prices are approximately

$$
\pi^*_{ft} \approx -\frac{\sum_{j \in \mathcal{J}_{ft}} s^*_{jft}}{\alpha} \tag{7}
$$

$$
p^*_{jft} \approx -\frac{1}{\alpha} + c_{jft}(R^*_{jft})^2 + w'_{jft}\mu + \eta_{jft} - \tau(R^*_{jft}) \tag{8}
$$

where s^*_{jft} is the market share when all the firms choose their optimal prices p^*_{jft} and ranges R^*_{jft}. From these equations, it can be seen that firms' optimal profits, prices, and ranges in each period are functions of $\vec{c}_t$ and all the remaining model characteristics. Since the remaining model characteristics are exogenous and taken as given, we express the optimal profits, prices, and ranges as $\pi^*_t(\vec{c}_t)$, $p^*_t(\vec{c}_t)$, and $R^*_t(\vec{c}_t)$.

In the next section, I will explain firms' investment decisions on reducing the cost parameters of ranges. Since all the profits, prices, ranges in the next section are the optimal values, I drop the $*$ in the notation.

4.3 Firms' dynamic investment problems

Firm f will invest in reducing c_{jft} for its product j if the expected returns to investment according to firm f's belief about the future market conditions are higher than the costs of investment. The returns to investment on product j are the increase in the sum of this product's expected discounted profits with an infinite time horizon. To simplify the analysis of the dynamic part, I assume that the investment decision on product j does not take into account the impact on the profits of f's other products and does not affect other products' cost parameters of ranges. The assumptions about firms' beliefs about the future are as follows: when considering investment for product j, firm f believes the prices and ranges of all the other products both inside and outside firm j in future periods will remain at their current values and that the exogenous product characteristics of all products including j will stay constant; firm f also believes that the future subsidy scheme and the future total number of consumers will remain the same as the current one. In each period, firms update their product-specific beliefs after observing that period's prices, ranges, exogenous product characteristics, number of consumers, and subsidy scheme.

I deviate from rational expectations in dynamic games to avoid equilibrium and identification problems of dynamic games among heterogeneous agents with rational expectations. The deviation is also due to the difficulties in justifying that firms can perfectly predict the infinite future in the rapidly involving EV industry because firms lack the necessary information or capabilities, similar to the discussions in Pesaran (1989) in a more general context. In fact, it is sometimes difficult to even predict what happens next year in the EV industry. For example, it has been reported that the changes in Chinese BEV subsidies came out as surprises to firms in some years. Therefore, firms in the EV industries likely make decisions based on intuitive predictions, and I assume their expectations to be adaptive. In standard adaptive expectations, future values are believed to be a linear function of historical values.[10]. To simplify the analysis for the dynamic part, I use a special case of adaptive expectations and assume that future values are believed by firms to equal the current values, in other words, the weights on historical values in standard adaptive expectations are set to 0. This special case of adaptive expectations implies that the investment decision for each model is a single-agent dynamic-discrete-choice problem.

Putting the dynamic part and the static part of the structural model together, firms' beliefs about other products' prices and ranges in the current period are rational expectations, whereas beliefs about future prices and ranges are adaptive expectations. The intuition behind these assumptions about firms' beliefs is that predicting what happens in the current year is likely easier and more reliable than predicting what will happen in all future years with an infinite horizon. In addition, assuming rational expectations is a common practice in static empirical structural models, and adaptive expectations are a widely-used alternative to rational expectations to guarantee existence of a unique equilibrium.

Since all model primitives in firms' beliefs are constant, in each period and for each product, firms solve a stationary problem with an infinite horizon. In other words, there is a stationary dynamic problem for each product in each period. The observed investment decisions are the decisions in the first periods of all these auxiliary stationary dynamic problems. Denote the value of product j from firm f at time t under f's belief formed at time t in the relevant auxiliary stationary dynamic problem by $v_{jf}^t(c_l)$:

$$v_{jf}^t(c_l) = \mathbb{E}[\max\{U_{jf}^t(c_l, 1) + \epsilon_{jf}(1), U_{jf}^t(c_l, 0) + \epsilon_{jf}(0)\}] \tag{9}$$

[10] Adaptive expectations are widely used in studies on inflation and monetary policies as well some applications for oligopoly (Okuguchi (1970) studies stability of oligopoly equilibrium under adaptive expectation) and firms' responses to technology shocks (Huang et al. (2009)).

where $c_{jft} = c_l$. The superscript t indicates the time when a belief is formed. The subscript jf indicates that the belief is formed by firm f for product j and that the value is about that product from that firm. There is no time subscript because of stationarity. $U^t_{jf}(c_l, 0)$ and $U^t_{jf}(c_l, 1)$ are the choice-specific values that satisfy

$$U^t_{jf}(c_l, 1) = -\lambda + \pi^t_{jf}(c_l) + \rho v^t_{jf}(c_{l-1}) \tag{10}$$

$$U^t_{jf}(c_l, 0) = \pi^t_{jf}(c_l) + \rho v^t_{jf}(c_l) \tag{11}$$

where λ is the investment cost. $\epsilon_{jf}(0)$ and $\epsilon_{jf}(1)$ are type-I extreme-value shocks with mean zero and are independent and identically distributed across time and products. The expectation is taken over these choice-specific shocks. ρ is the discount factor.

Equation (9) is the Bellman equation of this auxiliary stationary problem. The first part in the maximization is the value of investment, and the second part is that of no investment. Firms will invest when the first part is larger than the second part:

$$\begin{aligned}\mathbb{P}(a^t_{jf}(c_l) = 1) =& \mathbb{P}(U^t_{jf}(c_l, 1) + \epsilon_{jf}(1) > U^t_{jf}(c_l, 0)) + \epsilon_{jf}(0) \\ =& \frac{\exp(U^t_{jf}(c_l, 1))}{\exp(U^t_{jf}(c_l, 1)) + \exp(U^t_{jf}(c_l, 0))}\end{aligned} \tag{12}$$

When $c_{jft} = c_l$, the investment probability of product j at time t in the data is $\mathbb{P}(a^{jft}_{jf}(c_l) = 1)$. Each period t's auxiliary stationary problem can be solved by backwards induction from $c_{jft} = c_1$.[11] This also produces the optimal investment decisions, which are used to calculate the investment probabilities in the auxiliary stationary problem and the investment probabilities in the data.

At c_1, since there is no further reduction possible and prices and ranges of all the other products are constant according to firms' beliefs, value of c_1 under belief jf formed at time t is:

$$v^t_{jf}(c_1) = \frac{\pi^t_{jf}(c_1)}{1 - \rho}$$

$\pi^t_{jf}(c_1)$ is the profits of j under the belief formed by f for j at t if $c_{jf\cdot} = c_l$. For $1 < l \leq L$, $v^t_{jf}(c_l)$, $U^t_{jf}(c_l, 0)$, $U^t_{jf}(c_l, 1)$, and $\mathbb{P}(a^t_{jf}(c_l) = 1)$ can be solved by backward recursion according to Equations (9), (10), (11), and (12).

[11]There is no investment decision to be made at c_1 but the continuation value of c_1 is needed to solve the optimal investment decisions at c_2. The backward induction is over the space of the state variables.

4.4 Investment probability under different subsidy scheme scenarios

Proposition 1 *Consider a case where the dynamic problem defined by Equations (9), (10), and (11) with the state space $\{c_1, c_2, \cdots, c_L\}$ is implemented under two subsidy schemes called n and o. These subsidy schemes can be linear, notched, or no subsidy. Denote the value functions v^t_{jf}, the profits π^t_{jf}, and investment probability $\mathbb{P}(a^t_{jf} = 1)$ under the scheme n as $v^{t,n}_{jf}$, $\pi^{t,n}_{jf}$, and $\mathbb{P}(a^{t,n}_{jf} = 1)$, and under the scheme o as $v^{t,o}_{jf}$, $\pi^{t,o}_{jf}$, and $\mathbb{P}(a^{t,o}_{jf} = 1)$. Then for $c_l \in \{c_2, \cdots, c_L\}$, the following results hold:*

- *if $\pi^{t,n}_{jf}(c_l) - \pi^{t,0}_{jf}(c_l) = (1-\rho)(v^{t,n}_{jf}(c_{l-1}) - v^{t,0}_{jf}(c_{l-1}))$, then $\mathbb{P}(a^{t,n}_{jf}(c_l) = 1) = \mathbb{P}(a^{t,o}_{jf}(c_l) = 1)$;*
- *if $\pi^{t,n}_{jf}(c_l) - \pi^{t,0}_{jf}(c_l) < (1-\rho)(v^{t,n}_{jf}(c_{l-1}) - v^{t,0}_{jf}(c_{l-1}))$, then $\mathbb{P}(a^{t,n}_{jf}(c_l) = 1) > \mathbb{P}(a^{t,o}_{jf}(c_l) = 1)$;*
- *if $\pi^{t,n}_{jf}(c_l) - \pi^{t,0}_{jf}(c_l) > (1-\rho)(v^{t,n}_{jf}(c_{l-1}) - v^{t,0}_{jf}(c_{l-1}))$, then $\mathbb{P}(a^{t,n}_{jf}(c_l) = 1) < \mathbb{P}(a^{t,o}_{jf}(c_l) = 1)$.*

In other words, the investment probability under the scheme n is larger than the one under the scheme o if the difference in the current profits is small enough compared to the difference in the value of investment.

Proof. Define $\mu_l \equiv v^{t,n}_{jf}(c_l) - v^{t,o}_{jf}(c_l)$ and $\theta_l \equiv \pi^{t,n}_{jf}(c_l) - \pi^{t,o}_{jf}(c_l)$. Rewriting the Bellman equations in (9) using the Mcfadden surplus gives the following Bellman operators:

$$
\begin{aligned}
\Psi^{t,n,l}_{jf}(v) =& \pi^{t,n}_{jf}(c_l) + \log(\exp(\rho v) + \exp(-\lambda + \rho v^{t,n}_{jf}(c_{l-1}))\\
\Psi^{t,o,l}_{jf}(v) =& \pi^{t,o}_{jf}(c_l) + \log(\exp(\rho v) + \exp(-\lambda + \rho v^{t,o}_{jf}(c_{l-1}))
\end{aligned}
$$

for $l = 2, \cdots, L$. Since these Bellman operators are contraction mappings defined over a space of bounded functions with a finite state space, they all have a unique fixed point by the contraction mapping theorem. I denote the fixed points as $v^{t,n}_{jf}(c_l)$ and $v^{t,o}_{jf}(c_l)$. Rewriting Equation (12) gives:

$$
\mathbb{P}(a^t_{jf}(c_l) = 1) = \frac{\exp[-\lambda + \rho(v^t_{jf}(c_{l-1}) - v^t_{jf}(c_l))]}{1 + \exp[-\lambda + \rho(v^t_{jf}(c_{l-1}) - v^t_{jf}(c_l))]}
$$

This shows the sign of $\mathbb{P}(a^{t,n}_{jf}(c_l) = 1) - \mathbb{P}(a^{t,o}_{jf}(c_l) = 1)$ is the same as the sign of $[v^{t,n}_{jf}(c_{l-1}) - v^{t,n}_{jf}(c_l)] - [v^{t,o}_{jf}(c_{l-1}) - v^{t,o}_{jf}(c_l)]$. Because $[v^{t,n}_{jf}(c_{l-1}) - v^{t,n}_{jf}(c_l)] - [v^{t,o}_{jf}(c_{l-1}) - v^{t,o}_{jf}(c_l)] = [v^{t,n}_{jf}(c_{l-1}) - v^{t,o}_{jf}(c_{l-1})] - [v^{t,n}_{jf}(c_l) - v^{t,o}_{jf}(c_l)] = \mu_{l-1} - \mu_l$, the sign of $\mathbb{P}(a^{t,n}_{jf}(c_l) = 1) - \mathbb{P}(a^{t,o}_{jf}(c_l) = 1)$ is the same as the sign of $\mu_{l-1} - \mu_l$.

Evaluating the Bellman operator $\Psi_{jf}^{t,n,l}(v)$ at $v = v_{jf}^{t,o}(c_l) + \mu_{l-1}$ gives:

$$\begin{aligned}
\Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) =& \pi_{jf}^{t,n}(c_l) + \log[\exp(\rho(v_{jf}^{t,o}(c_l) + \mu_{l-1})) + \exp(-\lambda + \rho v_{jf}^{t,n}(c_{l-1}))] \\
=& \pi_{jf}^{t,n}(c_l) + \log[\exp(\rho(v_{jf}^{t,o}(c_l) + \mu_{l-1})) + \exp(-\lambda + \rho(v_{jf}^{t,o}(c_{l-1}) + \mu_{l-1}))] \\
=& \pi_{jf}^{t,n}(c_l) + \log[\exp(\rho v_{jf}^{t,o}(c_l)) + \exp(-\lambda + \rho v_{jf}^{t,o}(c_{l-1}))] + \rho\mu_{l-1} \\
=& \pi_{jf}^{t,n}(c_l) - \pi_{jf}^{t,o}(c_l) + v_{jf}^{t,o}(c_l) + \rho\mu_{l-1} \\
=& v_{jf}^{t,o}(c_l) + \theta_l + \rho\mu_{l-1}
\end{aligned}$$

The second equation uses the definition of μ_{l-1}, and the third equation brings the common $\rho\mu_{l-1}$ out of the log operator. The fourth equation uses the fact that $v_{jf}^{t,o}(c_l)$ is the fixed point of $\Psi_{jf}^{t,o,l}(v)$.

If $\pi_{jf}^{t,n}(c_l) - \pi_{jf}^{t,0}(c_l) = (1-\rho)(v_{jf}^{t,n}(c_{l-1}) - v_{jf}^{t,0}(c_{l-1}))$, i.e. $\theta_l = (1-\rho)\mu_{l-1}$, then

$$\Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) = v_{jf}^{t,o}(c_l) + (1-\rho)\mu_{l-1} + \rho\mu_{l-1} = v_{jf}^{t,o}(c_l) + \mu_{l-1}$$

So $v_{jf}^{t,o}(c_l) + \mu_{l-1}$ is a fixed point of $\Psi_{jf}^{t,n,l}(v)$. Since $\Psi_{jf}^{t,n,l}(v)$ has a unique fixed point, it must be $v_{jf}^{t,o}(c_l) + \mu_{l-1}$. So $\mu_{l-1} = \mu_l$ and $\mathbb{P}(a_{jf}^{t,n}(c_l) = 1) = \mathbb{P}(a_{jf}^{t,o}(c_l) = 1)$.

If $\pi_{jf}^{t,n}(c_l) - \pi_{jf}^{t,0}(c_l) < (1-\rho)(v_{jf}^{t,n}(c_{l-1}) - v_{jf}^{t,0}(c_{l-1}))$, i.e. $\theta_l < (1-\rho)\mu_{l-1}$, then

$$\Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) < v_{jf}^{t,o}(c_l) + (1-\rho)\mu_{l-1} + \rho\mu_{l-1} = v_{jf}^{t,o}(c_l) + \mu_{l-1}$$

Therefore, $v_{jf}^{t,n}(c_l) < \cdots < (\Psi_{jf}^{t,n,l})^3(v_{jf}^{t,o}(c_l) + \mu_{l-1}) < (\Psi_{jf}^{t,n,l})^2(v_{jf}^{t,o}(c_l) + \mu_{l-1}) < \Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) < v_{jf}^{t,o}(c_l) + \mu_{l-1}$, where $v_{jf}^{t,o}(c_l)$ is the fixed point of $\Psi_{jf}^{t,n,l}(v)$. This means $v_{jf}^{t,o}(c_l) + \mu_{l-1} > v_{jf}^{t,n}(c_l)$, so $\mu_{l-1} > \mu_l$ and $\mathbb{P}(a_{jf}^{t,n}(c_l) = 1) > \mathbb{P}(a_{jf}^{t,o}(c_l) = 1)$.

If $\pi_{jf}^{t,n}(c_l) - \pi_{jf}^{t,0}(c_l) > (1-\rho)(v_{jf}^{t,n}(c_{l-1}) - v_{jf}^{t,0}(c_{l-1}))$, i.e. $\theta_l > (1-\rho)\mu_{l-1}$, then

$$\Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) > v_{jf}^{t,o}(c_l) + (1-\rho)\mu_{l-1} + \rho\mu_{l-1} = v_{jf}^{t,o}(c_l) + \mu_{l-1}$$

Therefore, $v_{jf}^{t,n}(c_l) > \cdots > (\Psi_{jf}^{t,n,l})^3(v_{jf}^{t,o}(c_l) + \mu_{l-1}) > (\Psi_{jf}^{t,n,l})^2(v_{jf}^{t,o}(c_l) + \mu_{l-1}) > \Psi_{jf}^{t,n,l}(v_{jf}^{t,o}(c_l) + \mu_{l-1}) > v_{jf}^{t,o}(c_l) + \mu_{l-1}$, which means $\mu_{l-1} < \mu_l$ and $\mathbb{P}(a_{jf}^{t,n}(c_l) = 1) < \mathbb{P}(a_{jf}^{t,o}(c_l) = 1)$. ■

When firms' profits are increased, this affects investment probabilities through both the value of investing and the value of not investing. Higher profits at a c lower than the current c means a higher value for investment. However, higher profits at the current c imply a higher value for not investing. The investment probability is higher when switching from one scheme to another if the increase in the value for investing is higher than the increase in the value for not investing. As Proposition 1 shows, this is equivalent to comparing the increase in the values for investing with the increase in the current period profits. This proposition also shows that increasing the profits is not the key to boost investment, the key is the disproportionally larger increase in the profits gained from investment, or in other words, the profits at a lower c, which can be reached at some point in the future through investment.

This proposition also implies that not all the subsidy scheme can increase investment probabilities. Therefore, in Section 6, I will compare the investment probabilities under the notched scheme implemented in China in 2019 to a counterfactual scenario of no subsidies to evaluate the impact of the notched scheme on investment probabilities. I then construct a counterfactual linear scheme where the amount of total subsidies spent by the government in 2019 is the same as the one spent under the 2019 notched scheme. I then compare the investment probabilities under the 2019 notched scheme with those under this counterfactual linear scheme assuming that these subsidies schemes will stay forever.

5 Estimation

The setup of the model allows me to first estimate the static part of the model, i.e. the demand side and the static part of the supply side. This estimates all the demand parameters and the marginal cost functions, and it closely follows Berry et al. (1995) and Crawford et al. (2019). I then discretize the cost parameter of ranges, c_{jft} which is already estimated when estimating the marginal cost functions, and infer firms' investment decisions using changes in the discretized c_{jft}, in the sense that firms invest when the discretized c_{jft} decreases. Solving the dynamic problem in Equation (9), the conditional choice probability of investing is a function of the unknown investment cost λ because all the other parameters have been estimated in the static part. I then use maximum likelihood estimation to find the value of the investment cost that maximizes the inferred investment decisions. The first part of this section describes the estimation procedure in the static part and the second part explains the dynamic part.

5.1 The static part

According to the structural model, the mean utility of product j at time t is:

$$\delta_{jft} = \alpha p_{jft} + x'_{jft}\beta + \xi_{jft}$$

Then the market share of model j from firm f at time t is:

$$s_{jft} = \frac{\exp(\delta_{jft})}{1 + \sum_{j,f} \exp(\delta_{jft})}$$

The market share in logarithms relative to the outside option is:

$$\ln(s_{jft}) - \ln(s_{0t}) = \alpha p_{jft} + x'_{jft}\beta + \beta^R R_{jft} \cdot \mathbb{1}[\text{j is a BEV}] + \xi_{jft} + \zeta_t + \iota_{jf}$$

The left-hand side and the p_{jft} and x_{jft} from the right-hand side are known from data. I use the instruments constructed by Berry et al. (1995) to account for the endogeneity of prices and estimate α and β using the generalized method of moments (GMM). x_{jft} includes size, power-weight ratio, cost per km, torque, luxury level, whether released this year, and years the model exists in the data. Cost per km measures the expenditure on fuel per kilometer travelled, which reflects the fuel efficiency. The luxury level is a numerical indicator that sums over several dummy variables, such as whether a product contains a rain sensor or has a key-less start. All parameters except β^R, the marginal utility of ranges, are estimated using the full data that includes both gasoline vehicles and BEVs. β^R is estimated using only BEVs.

I estimate mc_{jft} and $\frac{\partial mc_{jft}}{\partial R_{jft}}$ using the first-order conditions of firms' profit maximization in Equations (2) and (3). Using the estimated $\hat{\alpha}$ and $\hat{\beta}^R$, only mc_{jft} and $\frac{\partial mc_{jft}}{\partial R_{jft}}$ are unknown in Equations (2) and (3). I follow Crawford et al. (2019) to calculate $\widehat{mc}_{jft}$ and $\widehat{\frac{\partial mc_{jft}}{\partial R_{jft}}}$ using matrix inversion. Since the marginal cost functions is $mc_{jft} = c_{jft}R^2_{jft} + w'_{jft}\mu + \eta_{jft}$, $\frac{\partial mc_{jft}}{\partial R_{jft}}$ satisfies:

$$\frac{\partial mc_{jft}}{\partial R_{jft}} = 2c_{jft}R_{jft}$$

Thus, the estimated $\hat{c}_{jft}$ is:

$$\hat{c}_{jft} = \left(\widehat{\frac{\partial mc_{jft}}{\partial R_{jft}}}\right) \cdot \frac{1}{2R_{jft}}$$

The remaining parameters in the marginal cost function can also be estimated using GMM with Berry et al. (1995) style of instruments to account for endogenous prices and ranges. However, I skip this step because I only need $\hat{c}_{jft}$.

To prepare for the estimation in the dynamic part, I discretize the estimated $\hat{c}_{jft}$ in logarithms over a grid whose lower bound $\log(c_1)$ is the lowest value of $\log(\hat{c}_{jft})$ in my sample in the year 2012-2020 minus 1.8, the upper bound $\log(c_L)$ is the largest value of $\log(\hat{c}_{jft})$ plus 0.1, and the distance between two adjacent grid points are 0.1. The estimated $\log(\hat{c}_{jft})$ is discretized to the closest grid point. If a model's $\hat{c}_{jft}$ is larger $\hat{c}_{jft-1}$, I set the value of $\hat{c}_{jft}$ equal to the value of $\hat{c}_{jft-1}$. I infer that there is an investment as long as c_{jft} decreases and this investment cause c_{jft+1} to be one step lower.

5.2 The dynamic part

The investment cost, the only parameter estimated in this part, is estimated using the inferred investment decisions only in 2019 to simplify the estimation in the dynamic part. I do not model entry and exit of firms and products, though entry and exit is common in the BEV industry in 2012-2020. I acknowledge that entry and exit can possibly affect the estimated investment cost. The estimation in this paper takes the set of firms and products in 2019 as exogenous. Therefore, the estimated investment cost should be interpreted as the average investment cost among the products that are still available next year and under the assumption that entry and exit is exogenous.

As explained in Section 4.3, the investment decision for model j in period t under firm f's belief formed at t for model j is the optimal choice in the first period of its auxiliary stationary dynamic problem as defined in Equation (9). Solving this problem following the steps explained in Section 4.3, gives the investment probability for model j at time t in the data. Denote the choice of investing for product j by firm f at time t in the data as $a_{jft}(c_{jft})$. When $c_{jft} = c_l$, the investment probability $\mathbb{P}(a_{jft}(c_l) = 1) = \mathbb{P}(a_{jf}^{jft}(c_l) = 1)$ is given in Equation (12), which is a function of the unknown investment cost.

I infer that there is an investment for product j at time t, i.e. $\hat{a}_{jft} = 1$ whenever $\hat{c}_{jft} > \hat{c}_{jft+1}$ because c_{jft} are assumed to decrease by one step with certainty if a firm invests. I do not use the inferred investment decisions of products whose ranges are less than 5 km away from the notched thresholds because, as explained in Section 4, the FOC in Equation (3) likely does no hold for these products and these products' inferred $\log(c_{jft})$s are biased downward. The estimated investment cost $\hat{\lambda}$ maximizes the log-likelihood function of investment decisions for all the products in 2019:

$$\hat{\lambda} = \arg\max_{\lambda} \sum_{j,f,t} \{\ln(\mathbb{P}[\hat{a}_{jft}(\hat{c}_{jft}) = 1|\lambda]) + \ln(\mathbb{P}[\hat{a}_{jft}(\hat{c}_{jft}) = 0|\lambda])\}$$

6 Results

I compare the value of c_{jft} before and after discretization to show how the discretization affects the estimated c_{jft}. Figure 3 shows the estimated c_{jft} and the difference between c_{jft} and c_{jft+1} for all the BEVs that have more than two consecutive years of observations in 2012-2020. The black dots are the $\hat{c}_{jft}$ and $\hat{c}_{jft+1} - \hat{c}_{jft}$ before the discretization, which are called the raw $\hat{c}_{jft}$ and raw $\hat{c}_{jft+1} - \hat{c}_{jft}$. The red dots are the discretized ones. The dots have darker colors when multiple dots overlap with each other. As explained in Section 5, when constructing the discretized $\hat{c}_{jft}$, I set the discretized $\hat{c}_{jft}$ equal to the discretized $\hat{c}_{jft-1}$ if the discretized $\hat{c}_{jft} > \hat{c}_{jft-1}$. It can be seen from Figure 3 that most products' raw $\hat{c}_{jft}$ satisfies $\hat{c}_{jft} \leq \hat{c}_{jft-1}$. Table 5 and Figure 4 show that most products' raw and discretized $\hat{c}_{jft+1} - \hat{c}_{jft}$ are between 0 and 0.1. Therefore, I infer that there is an investment at t when the discretized $\hat{c}_{jft+1} - \hat{c}_{jft} < 0$ and assume that $\hat{c}_{jft+1}$ is then reduced for sure by one step, i.e. 0.1.

There are in total 91 product-year observations in 2012-2020 where I can infer their investment decisions because they have observations in two consecutive years. As shown in Table 5, there are 25 observations, i.e. 27% of the 91 observations, inferred as having an investment. Among these 25 observations, the discretized $\hat{c}_{jft} - \hat{c}_{jft-1} = -0.1$ has the highest frequency, i.e. 10 observations. The discretized $\hat{c}_{jft} - \hat{c}_{jft-1} = -0.3$ has the second highest frequency, i.e. 5 observations, and followed by $\hat{c}_{jft} - \hat{c}_{jft-1} = -0.2$ with 4 observations. All the other values of the discretized $\hat{c}_{jft} - \hat{c}_{jft-1}$ have 6 observations. In terms of the distribution of the raw $\hat{c}_{jft} - \hat{c}_{jft-1}$, Figure 4 shows that $-0.05 < \hat{c}_{jft} - \hat{c}_{jft-1} < 0$ has the highest frequency. In 2019, there was investment for about 14% of the products. One possible concern over the discretization is the impact of sampling error in the estimated $\log(c_{jft})$. If the change in $\log(c_{jft})$ is due to sampling error, then an investment is inferred as existing while there is, in fact, no investment. This can be addressed by setting up the grid $\{c_1, c_2, \cdots, c_L\}$ based on the standard error of the estimated $\log(c_{jft})$. Currently, the grid is set by the author in such a way that $\log(c_l) - \log(c_{l-1}) = 0.1$ for $l \in \{2, 3, \cdots, L\}$.

TABLE 5: Frequency table of discretized $\log(c_{jft+1}) - \log(c_{jft})$

$\log(c_{jft+1}) - \log(c_{jft})$	count
0	66
-0.1	10
-0.2	4
-0.3	5
-0.4	2
-0.5	1
-0.6	3

Figure 5 shows ratios of the number of products receiving investment compared to the total number of products with two consecutive years of observations conditioned on the value of $\log(c_{jft})$. This distribution pools together investment decisions in 2012-2020 and include investment decisions under all the different subsidies in this period. Since products with

FIGURE 3: The estimated c_{jft} and its dynamics before and after the discretization

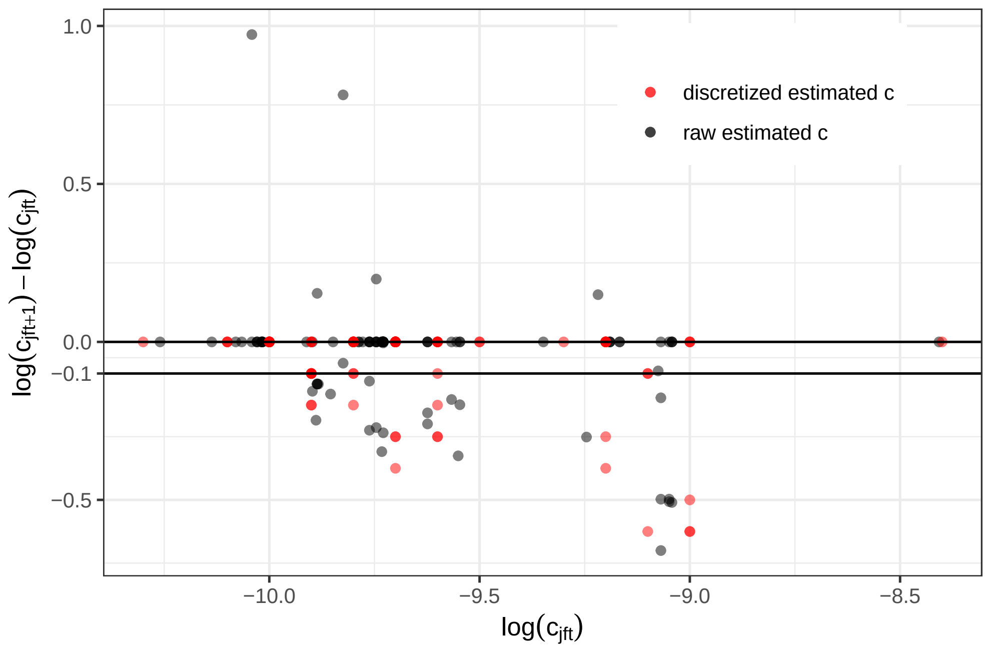
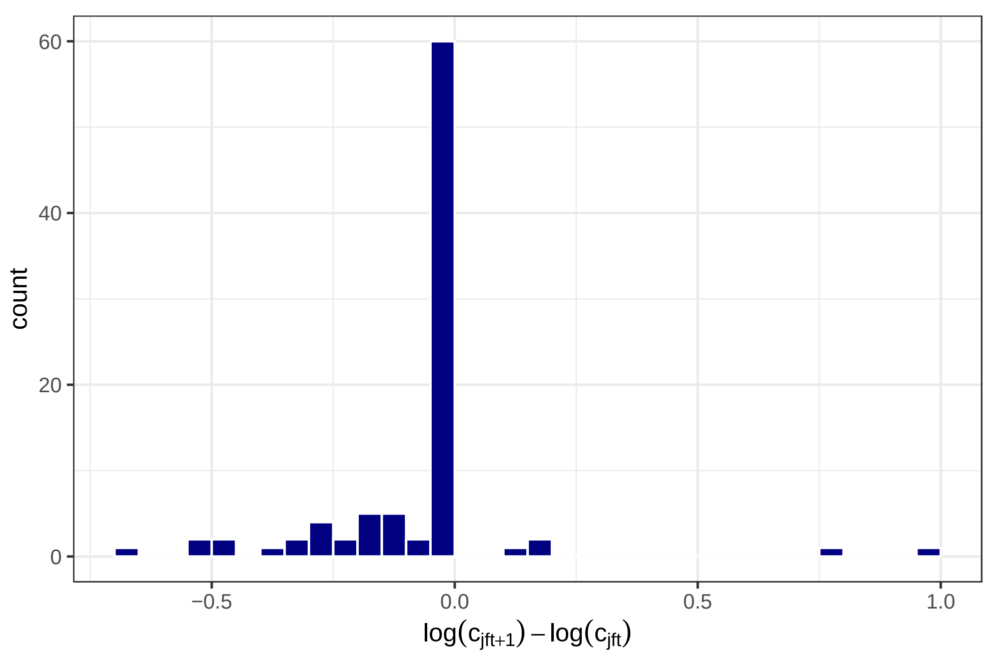


FIGURE 4: The distribution of the raw estimated $\log(c_{jft+1}) - \log(c_{jft})$



lower $\log(c_{jft})$ exists in later period and subsidies are a lot less generous in later years, the change in investment probability across different values of $\log(c_{jft})$ reflects also the change in subsidy schemes. In other words, the low investment rate at lower values of $\log(c_{jft})$ is likely driven by the less generous subsidy scheme in 2020.

FIGURE 5: The conditional investment ratios in the sample of each values of the estimated c_{jft}

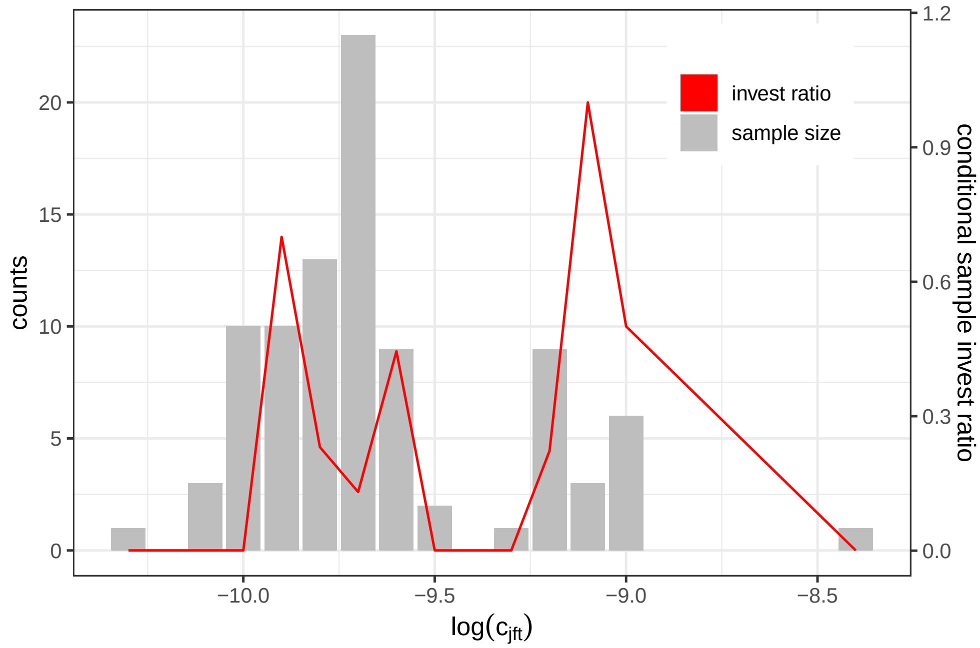


Table 6 displays the consumer taste parameters that include consumers' price sensitivity and the marginal utility from the observed product characteristics. Table 7 demonstrates the distribution of the estimated own price elasticities both in the full sample, i.e. including gasoline vehicles and BEVs, and in the sample of BEVs only. For 90% of the products, the BEV products have higher own price elasticities than those in the full sample.

The estimated investment cost is 53 billion RMB. The first two rows in Table 8 display the distributions of the BEVs' estimated range cost parameters in logarithms in 2019 and the ratios between the estimated profits and the estimated investment cost. On average, the investment cost is about 180 ($\frac{100}{0.057}$) times larger than the annual profits. The third row reports the markups of the BEVs, showing that BEVs' markups are between 1.16 and 1.30 in 2019.

Table 6: Estimated consumer taste parameters on prices and ranges

		estimated values
α		0.57 [0.14]
β	R (100 km)	2.11 [0.58]
	power-weight ratios (kw/kg)	23.28 [10.10]
	cost per km (RMB/km)	0.66 [1.11]
	size (m^2)	0.99 [0.53]
	torque (N·m)	7.61 [3.76]
	luxury level	0.04 [0.01]

Standard errors are in the brackets. Model-year two-way fixed effects are included.
All the parameters except β^R are estimated using the entire dataset, i.e. including both gasoline vehicle models and BEVs. β^R only uses BEVs.

Table 7: Summary statistics of estimated own price elasticities

	Mean	St. Dev.	min	Pctl(10)	Pctl(25)	Median	Pctl(75)	Pctl(90)	max
$\frac{\log(s_{jft})}{\log(p_{jft})}$ (full sample)	6.24	3.96	1.17	2.61	3.70	5.11	7.45	11.39	30.13
$\frac{\log(s_{jft})}{\log(p_{jft})}$ (BEV)	7.43	4.22	1.27	2.62	4.14	6.99	9.41	12.33	21.85

Table 8: Summary statistics of a selection of the supply-side parameters (2019)

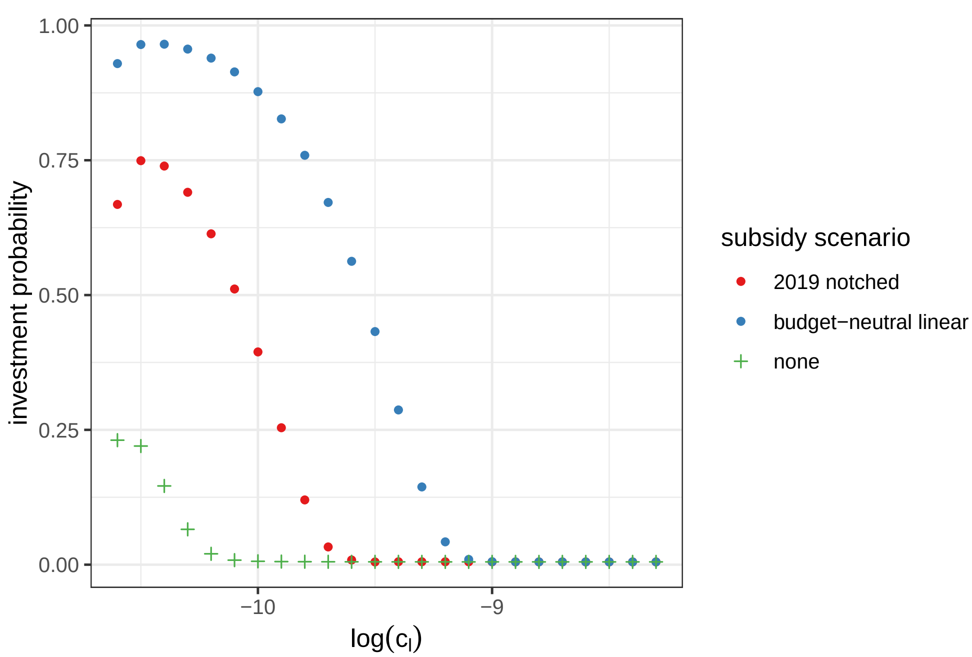
	Mean	St. Dev.	min	Pctl(10)	Pctl(25)	Median	Pctl(75)	Pctl(90)	max
$\log(c_{jft})$	−9.82	0.21	−10.30	−10	−10	−9.80	−9.70	−9.66	−9.20
$\pi_{jft} \times 100/\lambda_{jft}$	0.57	0.64	0.11	0.27	0.27	0.36	0.72	0.72	3.86
markups	1.22	0.05	1.16	1.17	1.17	1.26	1.26	1.27	1.30

6.1 Counterfactuals

Proposition 1 shows that a subsidy scheme can increase the investment probabilities if the increase in the current-period profits is small enough compared to the the increase in the value of investing. This means not all the subsidy schemes can incentivize investment. In this section, I use the estimated structural parameters to compare the investment probabilities under the existing notched subsidy scheme to a counterfactual scenario of no scheme. I also compare the notched scheme to a linear scheme that has the same budget in the current period as the notched scheme.

Figure 6 demonstrates the predicted investment probabilities at all the values of $\log(c_{jft})$ in the space $\{c_1, c_2, \cdots, c_L\}$. In 2019, $\log(c_{jf,2019})$ are between -10.3 and -9.2. The notched scenario in the figure, represented by the red dots, uses the subsidy scheme implemented in 2019, as shown in Table 1. The other two subsidy scenarios in Figure 6 are the case of no consumer subsidies, represented by the green crosses, and the case of a budget-neutral linear subsidy scheme, represented by the blue dots. The budget-neutral linear subsidy scheme has the same total amount of subsidy in 2019 as the one under the notched scheme, or in other words, the government's expenditure on subsidizing the BEV purchase in 2019 is held constant, but this does not guarantee that the expenditure will be the same for the two schemes in future periods.

FIGURE 6: Investment probabilities of BEVs in 2019 under three subsidy scenarios



Moving from $\log(c_L)$ to $\log(c_1)$, i.e. moving from the right side to the left side of Figure 6, the cost parameter decreases and the investment probability under the notched scheme first increases and then decreases slightly, but it is well above the one without a subsidy scheme. The investment probabilities under no subsidy scheme has a similar shape, but they increase at a much lower $\log(c_l)$. The investment probability in the scenario of no subsidy increases as $\log(c_l)$ decreases because firms profits are approximately proportional to $\frac{1}{c_{jft}}$. This can be seen by combining Equations (1), (6), (7), and (8). Therefore, profits increase more for smaller c_{jft} when c_{jft} decreases by a given amount. Although one step of decrease in c_{jft} becomes smaller in magnitude when c_{jft} is smaller, the profits becomes highly sensitive to small changes c_{jft} when c_{jft} is small enough so that gains to investment becomes larger at lower c_{jft}. Consequently, the investment probability is larger at lower c_{jft}.

In the linear-subsidy scenario, the investment probabilities are almost the same as those in the other two scenarios for $\log(c_l) > -9.1$ but are larger than those under the notched scheme. Since products in 2019 are in the range of -10.3 and -9.2, this implies that using a budget-neutral linear scheme can increase investment probabilities in 2019. However, these higher investment probabilities under the budget-neutral scheme are significantly more costly than the notched scheme in the future years if the schemes remain the same, or in the other words, if firms expectation about future schemes are correct. Figure 7 shows the subsidy budget ratios between the notched scheme and the linear scheme from year 0, i.e. year 2019, till year 10, i.e. year 2029. In year 10, the budget of the linear scheme is 600 times as the budget of the notched scheme. With these extra subsidies injected into the market, there are a lot more products with $c_{jft} = c_1$, the lowest value possible for c_{jft}. In that year, there would be about 60% of the products with $c_{jft} = c_1$ under the linear scheme and about 18% under the notched scheme.

FIGURE 7: The predicted annual subsidy expenditures over 10 years if the subsidy schemes do not change

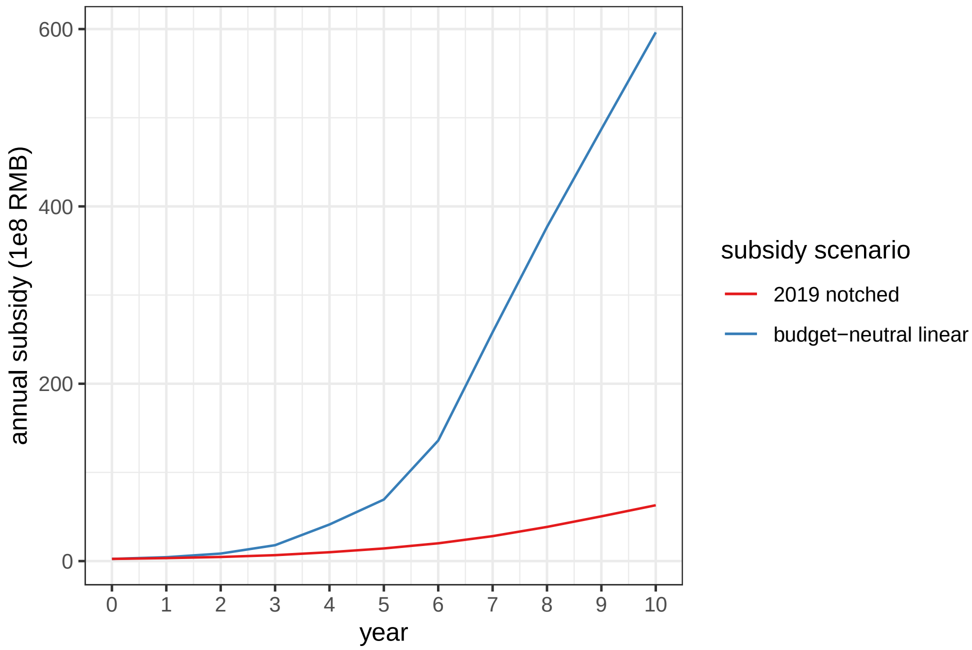
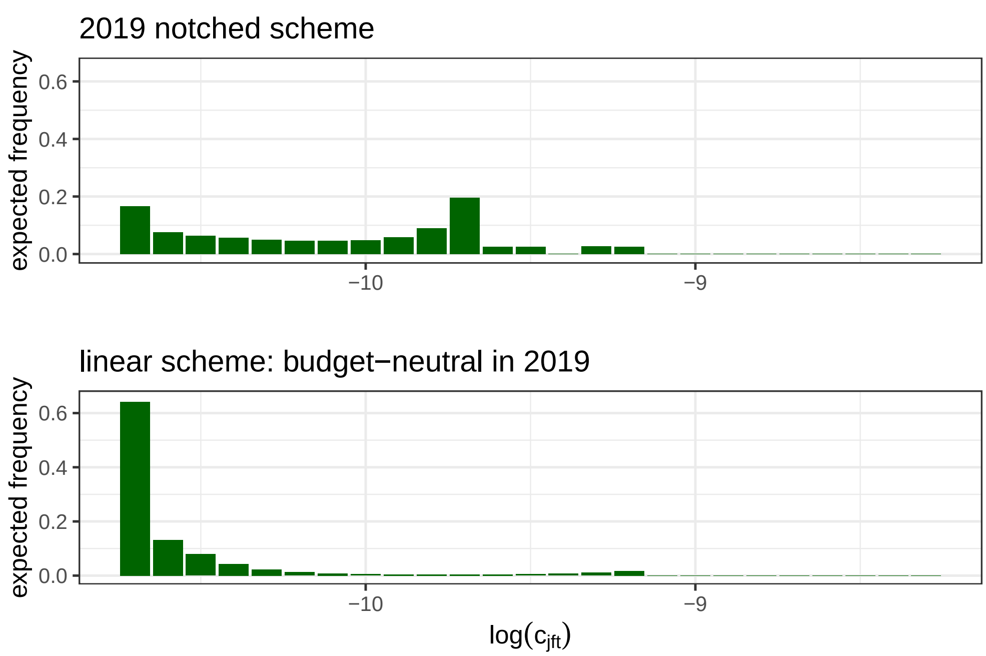


FIGURE 8: The predicted $\log(c)$ distribution of the 2019 cohort of firms after 10 years if the subsidy schemes do not change



7 Conclusion

This paper models and estimates manufacturers' incentives to reduce the production cost of ranges in response to range-based subsidies to consumers purchasing battery electric vehicles (BEVs) and finds that the notched scheme increases the investment probabilities by up to 50 percentage points in 2019. Compared to a linear scheme where the total amount of subsidy in 2019 is constant, the investment probabilities under the notched scheme are lower, but the notched scheme is much cheaper for the government in future years.

The numerical results in this paper are derived under the assumption that demand is static, in the sense that consumers' decisions about purchasing a BEV last year does not affect their decisions this year. This paper also assumes that subsidies to consumers do not generate dynamic changes in demand, such as causing consumers to postpone or expedite the decisions of buying a BEV. Other assumptions imposed are homogeneous consumers, adaptive expectations about other BEV models' future prices and ranges, adaptive expectations about the future subsidy scheme, and that investment leads to one unit of decrease in the log of the cost parameter of increasing ranges, i.e. $\log(c_l)$, with certain.

The assumption of homogeneous consumers can be relaxed by allowing the consumer taste parameters to differ across consumers and then estimate the demand side of the structural model using random-coefficient estimation as it is implemented in Berry et al. (1995). The adaptive expectation about other BEVs' prices and ranges can be relaxed by using adaptive expectations about other BEV models' range cost parameters, i.e. c_{jft}. This implies that investment decisions on one BEV take into account the price and range responses of the other BEVs. This adaptive expectation about prices and ranges can also be replaced by rational expectation so that the investment decision is a dynamic game. Under the belief that the subsidy scheme stays constant, the auxiliary dynamic problem in each period is a stationary directional dynamic game and can be solved using the method of Iskhakov et al. (2016). The assumption that $\log(c_l)$ decreases by 1 unit with certainty if a firm invests can be relaxed by allowing the change in $\log(c_{jft})$ to follow a distribution where $\log(c_{jft})$ may decreases by more than 1 unit. Although most BEVs' $\log(c_{jft})$ either decrease by one unit or do not change in 2019, there are some BEVs whose $\log(c_{jft})$ decreases by more than 1 unit. Allowing stochastic improvements over c_{jft} can improve the fit of the structural model to the data in this regard. Instead of assuming that manufacturers believe the subsidy scheme never changes, firms can be assumed to have perfect foresight about future subsidy schemes. This is probably the most difficult extension because the problem of investment decisions is then non-stationary with infinite time horizon. If the adaptive expectation about other BEVs' prices and ranges is kept, the investment decision problem is still a single-agent problem.

Since the subsidy scheme in China will eventually disappear, the periods after the subsidy scheme disappears can be modeled as stationary and the periods with subsidy scheme can be solved by backward recursion.

In all these extensions, as long as the increase in the value of investment is large enough compared to the increase in the current-period profits when introducing a notched scheme, the notched scheme increases the investment probabilities. Whether it is better to use a linear scheme or a notched scheme depends on the target and financial constraints of the government. If the government's goal is to stimulate investment and does not care the expenditure in the future periods, it may be better to use a linear scheme. If the government can not afford an expensive linear in the future periods, a notched scheme may offer a balance between stimulating investment and avoid the budget from becoming too high in the future.

References

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841.

Bold, T., S. Ghisolfi, F. Nsonzi, and J. Svensson (2022): "Market Access and Quality Upgrading: Evidence from Four Field Experiments," *American Economic Review*, 112, 2518–2552.

Crawford, G. S., O. Shcherbakov, and M. Shum (2019): "Quality Overprovision in Cable Television Markets," *American Economic Review*, 109, 956–995.

Criscuolo, C., R. Martin, H. G. Overman, and J. Van Reenen (2019): "Some Causal Effects of an Industrial Policy," *American Economic Review*, 109, 48–85.

Huang, K. X., Z. Liu, and T. Zha (2009): "Learning, Adaptive Expectations and Technology Shocks," *The Economic Journal*, 119, 377–405.

Iskhakov, F., J. Rust, and B. Schjerning (2016): "Recursive Lexicographical Search: Finding All Markov Perfect Equilibria of Finite State Directional Dynamic Games," *The Review of Economic Studies*, 83, 658–703.

Ito, K. and J. M. Sallee (2018): "The Economics of Attribute-Based Regulation: Theory and Evidence from Fuel Economy Standards," *The Review of Economics and Statistics*, 100, 319–336.

Jia, B. P., H.-s. Kwon, and L. Shanjun (2022): "Attribute-Based Subsidies and Market Power: An Application to Electric Vehicles," *working paper*.

Li, S., X. Zhu, Y. Ma, F. Zhang, and H. Zhou (2022): "The Role of Government in the Market for Electric Vehicles: Evidence from China," *Journal of Policy Analysis and Management*, 41, 450–485, publisher: John Wiley & Sons, Ltd.

Okuguchi, K. (1970): "Adaptive Expectations in an Oligopoly Model," *The Review of Economic Studies*, 37, 233–237, publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Pesaran, M. (1989): *The Limits to Rational Expectations*, Basil Blackwell.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033.

Takalo, T., T. Tanayama, and O. Toivanen (2013): "Estimating the Benefits of Targeted R&D Subsidies," *Review of Economics and Statistics*, 95, 255–272.

Appendix

A Data preparation

I need to merge all the data collected from various sources into one panel. Each observation is a model-year-month. The variables include monthly sales, technical descriptions such as horse power, eligibility for subsidies, the values of the subsidies if eligible, eligibility for purchase tax exemption, the price of electricity and gasoline, the number of charging poles or charging stations. The biggest challenge is to merge sales data, subsidy info, purchase tax info, and technical description. In this section, I will talk about how this merge is carried out.

A.1 Merging NEVPTE List, RNEVPTE List, and the NEV List

I first convert the NEVPTE List, RNEVPTE List, and the NEV List published by the government from text format into data format. I treat observations in these lists as redundant if the same model ID number in the same year-month appears more than once in a list. Among the redundant observations, I keep the observation with the lowest number of missing variables, the smallest size, and the smallest weight. Because these lists use the same model ID numbers, I merge them based on model ID numbers and use variables shared by the lists to check whether the merging is correct. More specifically, the lists all report manufacturer's name, model type, and weight. I first check whether the manufacturer's names and model types are the same. Next I check whether weights are close enough, i.e. less than 10% apart. I call the merged data "subsidy-purchase-tax" (SPT) data.

A.2 Merging sales, SPT, and the technical description data

I first take the sales data collected from Chezhu Home. I use model names to match the technical description data collected from Chezhu Home and Auto Home. I first use exact matching and then apply fuzzy matching for those unmatched by the exact matching. I then repeat the same process to match with SPT. Sales data is at the model level, but the technical description data is at the model-variation level. Therefore, for each model, there are multiple entries in our technical description data. I keep the entry that has the lowest amount of missing variables, the lowest price, and the smallest sizes, weight, and horse power, following the practice of Berry et al. (1995). For the unmatched observations, I check whether the failures of matching is due to incorrectly recorded names and correct them manually if needed.

A.3 Summary stats

Table 9 shows the differences in annual aggregate sales between my raw data, my merged data, and sales announced by the government.

TABLE 9: Differences in aggregate annual sales caused by data collection and data cleaning

year	sales merged	sales scrapped	sales web	gap merged and scrapped	gap scrapped and web
2010	7,763	11,040	13,758	0.300	0.200
2011	12,019	14,316	14,473	0.160	0.010
2012	12,844	16,455	15,494	0.220	-0.060
2013	20,481	21,135	17,929	0.030	-0.180
2014	21,856	22,424	19,701	0.030	-0.140
2015	21,755	22,322	21,146	0.030	-0.060
2016	23,304	23,788	24,377	0.020	0.020
2017	20,237	20,689	24,718	0.020	0.160
2018	19,523	19,858	23,710	0.020	0.160
2019	17,403	17,834	21,444	0.020	0.170
2020	15,761	16,311	20,178	0.030	0.190
2021	13,612	13,843	21,482	0.020	0.360

Notes:
Gap merged and scrapped = 1-sales merged/sales scrapped
Gap scrapped and web = 1-sales scrapped/sales web