Project 3: Conditional Logit

AØKA08084U Advanced Microeconometrics

University of Copenhagen

Character Count (incl. spaces): 12274

Sebastian Hørlück, Matias Piqueras

March 18, 2022

Contents

1 Introduction

When companies decide on a price for their products they face a trade-off when maximizing profits, as is most companies primary goal. Low prices secure a larger market share, but also lowers revenue per unit sold, and vice versa. Thus, a key goal for companies is to estimate the demand curve of customers, as to have the best knowledge base for deciding prices.

The aim of this project is to estimate the home bias in demand for cars and investigate how the "home bias" affects the price elasticity of cars. We do this through a conditional logit model. We use a dataset consisting market shares for the 40 most sold cars in five different countries through 30 years.

Our results show that cars sold in their home countries increase market shares by 1.97 percentage points on average and statistically significant. Further we see that the price elasticity for cars sold in their home countries is on average 0.06 percentage points lower than other cars and statistically significant.

2 Methods

In this section we outline the methods used to estimate two quantities of interest (1) the effect of "home bias" and (2) its effect on own-price elasticity of demand. To do so we focus on two models, first, the conditional logit and then an OLS method which is used as a baseline comparision.

For each market $i = c, t$, given by a country c and a year t, in our data the outcome variable $\boldsymbol{y}_i$ is the vector of observed market shares for the 40 most sold cars $j = 1, \ldots, J$. Thus, y_{ji} is the scalar outcome for one car model. For each market-year i we also have the matrix of individual car characteristics $\boldsymbol{X}_i \in \mathbb{R}^{J \times K}$ which is further detailed in section 3.

2.1 Conditional Logit

Since the data is aggregated on market level, we use that the expected market share of some car model s_{ji} is asymptotically equivalent to the average choice probabilities of household h choosing j as $H \to \infty$. This means that the utility u_{ih} for a car can be estimated from our data. If we further parameterize the problem as $u_{ijh} = \boldsymbol{x}_{ij}\boldsymbol{\beta} + \epsilon_{ij}$, where $\epsilon_{ij} \sim$ IID Extreme Value Type 1 and assume that an individual choice is given by $y_{ijh} = \text{argmax}_{j \in 1,\ldots J} u_{ijh}$ then $\boldsymbol{\beta}$ can recovered by applying the logit link function

$$s_{ji}(\boldsymbol{X}_i, \boldsymbol{\beta}) \equiv \mathrm{P}[y_i = j|\boldsymbol{X}_i] = \frac{\exp\left(\boldsymbol{x}_{ij}\boldsymbol{\beta}\right)}{\sum_{k=1}^{J}\exp\left(\boldsymbol{x}_{ik}\boldsymbol{\beta}\right)} \tag{1}$$

The non-linear functional form of (1) does not offer a closed form solution to estimate $\boldsymbol{\beta}$. Instead, we use the M-estimation framework and more specifically, maximum likelihood estimation (MLE) to recover the parameters of interest. Let $\boldsymbol{\Theta}$ denote the parameter space and $\hat{\boldsymbol{\beta}}$ the estimator of the true $\boldsymbol{\beta}_0$ then an M-estimator is the solution to the sample problem

$$\hat{\boldsymbol{\beta}} \in \underset{\boldsymbol{\beta}\in\Theta}{\operatorname{argmin}} \frac{1}{N}\sum_{i=1}^{N} q(\boldsymbol{y}_i, \boldsymbol{X}_i, \boldsymbol{\beta}). \tag{2}$$

Where $q(\cdot)$ is a criterion function to be minimized. In (Wooldridge, 2010, p.400) it is shown that while (2) implies identifiability it does not necessarily imply identification of a unique solution. In the case of the conditional logit model we use the results from economist Daniel McFadden who showed that while the constant needs to be normalized, $\boldsymbol{\beta}_0$ is identified in (1) under our distributional assumption about ϵ_{ij} as long as the variable corresponding to the parameter being estimated varies across alternatives. Under the further condition of uniform law of large numbers which holds if the parameter space $\boldsymbol{\Theta}$ is compact and the criterion $q(\cdot)$ is continuous in $\boldsymbol{\beta}$ (Theorem 12.1 in Wooldrige), it turns out that $\hat{\boldsymbol{\beta}}$ is consistent as it convergence in probability to $\boldsymbol{\beta}_0$ as $N \to \infty$. In practice, we will use that any MLE is also an M-estimator and that $q(\cdot)$ in our setting is given by the negative loglikelihood contribution that takes the particular form $-\sum_{j=1}^{J} y_{ij} \log s_j\left(\mathbf{X}_i, \boldsymbol{\beta}\right)$. In summary, this means that we can estimate $\hat{\boldsymbol{\beta}}$ by minimizing the following expression

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \; -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{J} y_{ij} \log s_j\left(\mathbf{X}_i, \boldsymbol{\beta}\right). \tag{3}$$

2.2 OLS method

Given a panel dataset of realized market share over time, it turns out that OLS can also be used as a baseline approximation of $\boldsymbol{\beta}$. We will use this to gain intuition and as a way to evaluate the results from the more complex conditional logit model. To estimate OLS we begin by noting that (1) it is easy to show that the market share of j relative to some other model l can be expressed as linear function taking the form

$$\log s_{ji} - \log s_{li} = (\boldsymbol{x}_{ji} - \boldsymbol{x}_{li})\boldsymbol{\beta}. \tag{4}$$

This suggests setting an arbitrary choice as a reference e.g. $l = 1$ and transforming the data such that $\tilde{y}_{ji} = \log s_{ji} - \log s_{1i}$ and $\tilde{\boldsymbol{x}}_{ji} = \log \boldsymbol{x}_{ji} - \log \boldsymbol{x}_{1i}$, allowing us to specify the following linear regression model $\tilde{y}_{ji} = \tilde{\boldsymbol{x}}_{ji}\boldsymbol{\beta}$ that can be estimated by OLS.

We obtain $\hat{\beta}^{OLS}$ through this method and use these estimates to find marginal effects of home bias as well as the effect of home bias, as outlined in section 2.3 and 2.4.

2.3 Marginal effects

The β-estimates obtained using either the method outlined in section 2.2 or 2.1 does not provide information on the quantity of the home bias on their own, but can be used to obtain a measurement of the home bias by providing an estimate of the marginal effects of the home bias.
We quantify the marginal effects through the following procedure: using a set of $\hat{\beta}^m$ with $m \in \{\text{OLS, CLogit}\}$ and $\boldsymbol{X}$ we predict market shares for each car j in each market i. Then we compute a set of counterfactual market shares by creating an artificilal set of covariates, $\boldsymbol{X}^a$, which is identitcal to $\boldsymbol{X}$ except for the column vector corresponding to the home-dummy, which is set to zero for all rows. We then derive the marginal effect of the homebias as:

$$\xi = \frac{1}{H} \sum_{h=1}^{H} s_h(\boldsymbol{X}_i, \hat{\beta}) - s_h(\boldsymbol{X}_i^a, \hat{\beta})$$

Where $S = \{1, 2, ..., H\}$ describes the set of j, i pairs of cars and markets of cars sold in home markets. If car A's home market is a, and car B's home market is b, then A, a and B, b is included in S, while A, b and B, a are not. We present the results of ξ in section 4.

2.4 Elasticities

We quantify the effect of home bias on price elasticites by computing own price- and cross price elasticities for all cars. The own price elasticites are computed by proposing a small increase in prices, and finding $\Delta\boldsymbol{X}$, describing the relative change in $\boldsymbol{X}$ following the proposed price change. We then compute a new set of market shares, following the price change, and find Δs describing the relative change in market shares. For each car j in market i we find the own price elasticity as:

$$\eta_{ji}^{own} = \frac{\Delta s_{ji}}{\Delta \boldsymbol{X}_{ji}}$$

Using the $N \times J$ matrix of own price elasticities we take the average of cars sold in home markets and cars not sold in home markets, to find η^{home} and η^{other}, respectively, and quantify the effect of home bias on own price elasticities as difference in average elasticities between home cars and non-home cars. We present the results of this method in section 4.

2.5 Inference

To draw inference on the estimated marginal effects and elasticites derived using the methods desribed until now, we make use of the bootstrap method.
The bootstrap method is a re-sampling method, where the initial sample, in our case the 150 markets, are treated as the population. Following Wooldridge (2010), we resample the initial sample by drawing 150 random markets from a uniform distribution with replacement. That is, each subsample in the bootstrapping process has as many observations as the initial sample, but the same observation can be repeated in bootstrap subssamples.
For each b subsample of the B subsamples we estimate the parameters $\hat{\beta}^{(b)}$ of (1) and derive the subsample marginal effects, $\hat{\xi}^{(b)}$, and difference in elasticities, $\hat{\eta}^{(b)}_{home} - \hat{\eta}^{(b)}_{other}$, using the transformations of $\hat{\beta}^{(b)}$ described in section 2.3 and 2.4.
From the bootstrap results we draw bootstrap bias' and standard errors of both marginal effects and differences in elasticities. The bootstrap bias estimate is computed as: $\bar{\hat{V}} - \hat{V}$, for $V \in \{\psi, \eta_{home} - \eta_{other}\}$, where $\bar{\hat{V}} = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{V}^{(b)}\right)$ the leftmost element of the function represents the estimates from the initial sample. The bootstrap standard error estimate is computed as

$$\mathrm{se}(\hat{V}) = \left[(B-1)^{-1}\sum_{b=1}^{B}\left[\bar{\hat{V}} - \hat{V}^{(b)}\right] \cdot \left[\bar{\hat{V}} - \hat{V}^{(b)}\right]'\right]^{\frac{1}{2}} \tag{5}$$

Using these standard errors we construct confidence intervals for $\hat{V}$. For both marginal effects and differences in elasticites we test the null hypothesis of:

$$\mathcal{H}_0 : \hat{V} = 0$$

against the alternative hypothesis of

$$\mathcal{H}_A : \hat{V} \neq 0$$

using a 95 pct. confidence interval.

3 Variable Selection

For the analysis we have included the following explanatory variables:

- Price over GDP in demand country. This the variable we use for price elasticities.
- Home-dummy. This is the variable we use to qunatify home bias'.
- Size = (length × height × width), following Noton (2015)'s approach to include car characteristics and avoiding collinarity.
- Motor power = (horsepower × cylinders), also following Noton (2015).
- Weight
- Fuel efficiency, measured as an average of fuel efficiency at three different speeds.
- Exchange rate between importer and exporter country.
- Class dummies at five different levels. One is excluded to avoid dummy trap.
- Brand dummies at 33 different levels. One is excluded to avoid dummy trap.

We use price over GDP as this captures differences in inflation and income changes across countries, following Noton (2015). We further include exchange rates among importer and exporter countries to capture price changes following changes only related to changes in exchange rates.

This results in a total of 43 explanatory variables included in the analysis. We use shares of market held by the top 40 selling cars among the share held by the cars, i.e. the share variable sums to 1 for each country and year, even though other cars than these 40 were sold in the country during the year.

Table 1: Main results

Measures	Results
Home bias	1.971247
	[0.0187 - 0.0207]
Price elasticity, home cars	-0.303522
	[-0.4159 - -0.1778]
Price elasticity, other cars	-0.359096
	[-0.5215 - -0.2157]
Elasticity difference	0.055573
	[-0.1087 - -0.0348]

[.] indicates 95 pct. confidence intervals based on bootstrapped standard errors

4 Results

Our main β-estimates stem from the method outlined in section 2.1. Using the estimated $\widehat{\beta}^{CLogit}_{home}$ we find the marginal effect of whether a car is sold in its home country, following section 2.3. We compute standard errors following the bootstrap method described in 2.5 with 500 bootstrap iterations. We present the results in

Determining the marginal effect of the home, we find that the home bias increases the market share by 1.97 percentage points on average, with a standard error of 0.05. Thus the results are significant on a 5 pct. significance level.

Further we investigate whether the price elastiticy of home cars differentiate from the price elasticity on other cars. We find that the price elasticity on home cars is -0.3035 and for other cars -0.3591, thus presenting a difference in elasticities of -0.0556. Thus, at the given prices, the demand for home produced cars would decrease by almost 0.06 percentage points less than other cars at a price increase of one percent. The difference in elasticities between home cars and other cars show a standard error of 0.02, thus the difference is significant.

The elasticity of on home cars show a standard error of 0.048, and the elasticity of no home cars show a standard error of 0.066, thus the two elasticities can not be shown to be differenct on a 5 pct. significance level.

4.1 Check of parameters using OLS approach

Since the approach for estimating the parameters of (1) using M-estimation requires can proof to be difficult in the respect of programming it, we use the OLS-method explained in section 2.2 to check whether the parameters stemming from our CLogit approach seem correct.

Following these estimates we find an average marginal effect of 2.01 percentage points on market shares stemming from the home bias, thus presenting an estimate close to the one derived through the Conditional Logit method.

Deriving the price elasticities using $\hat{\beta}^{OLS}$ we find an average own price elasticity for home cars of 0.2467, while for other cars the average own price elasticity is 0.2983. Thus the OLS method presents a difference in elasticities of -0.0517 percentage points. As with the marginal effects of the home bias, the elasticities based on the OLS approach are relatively close the elasticites from Conditional Logit.

Thus we find our results based on CLogit to be computed correctly.

5 Discussion

Our results show significant effects, both in terms of the home bias and the difference in price elasticities. We conclude this based on bootstrapped standard errors, but we could also have computed standard errors based on the Delta method.

We also want to draw some attention to the IIA assumption. Its a property of the CLogit model which implies that the ratio between any two alternatives will be preserved if a third alternative is added even if it is unimportant for the choice on an intuitive level (Wooldridge, 2010, p.648). Thus, when we compute elasticites by artifically changing prices for one car, all other cars in the market are affected by the price change. This assumption can seem unrealistic as the demand for luxury cars should not be affected by price changes on cheap cars. To circumvent this, a Nested Logit model could be introduced, where cars would be divided into "nests", and demand for cars would only be affected for price changes on cars within the same nest.

Anoter possible drawback of this method is the assumption of even preferences across markets. Given that the markets in our analysis vary both in location and time, it could be a reasonable assumption that customers differ between markets, thus not showing identical preferences.

References

Noton, C. (2015). On the size of home bias. *Applied Economics*, *47*(2), 123–128.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Massachusetts Institute of Technology.