Advanced Microeconometrics - Project 3

Group 21
Lukas Malte Kemeter (xtj271),
Felix Heuer (pfv999),
Malte Jacob Rattenborg (rnp562)

December 17, 2020

Note: The text for Part 2 amounts to four normal pages (Font size = 12p, 1.5 line spacing, and margins of 2.5 cm) according to the restrictions. For sake of readability, we include the tables and figures into the text.

Problem 1

(1):

To derive the joint probability density function of the pair (Y_{i1}, Y_{i2}), we first calculate the probability that both $Y_{i1}$and Y_{i2} are observed separately. Given the independence of $\{(Y_{i1}, Y_{i2})\}_{i=1}^{N}$, their joint probability is then represented by their combined density function and can be stated as:

$$f(Y_{1i} \cup y_{2i}) = f(Y_{1i})f(Y_{2i})$$

Considering both $\{Y_{i1}\}_{i=1}^{N}$ and $\{Y_{i2}\}_{i=1}^{N}$ being normally distributed, with a mean of μ_i and a homogeneous variance of σ^2, their individual probability density is given by :

$$f(y_{ti}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{Y_{ti}-\mu_i}{\sigma})^2} \quad , for\ both\ t = 1,2$$

Accordingly, the multiplication of the individual densities of $Y_{i1}$and Y_{i2} is given by:

$$f(y_{1i} \cup Y_{2i}) = \frac{1}{\sigma^2 2\pi} e^{-\frac{1}{2}(\frac{Y_{1i}-\mu_i}{\sigma})^2} e^{-\frac{1}{2}(\frac{Y_{2i}-\mu_i}{\sigma})^2}$$

This constitutes the probability density function of the pair (Y_{i1}, Y_{i2}).

(2):

The combined Log-Likelihood function of $\{(Y_{i1}, Y_{i2})\}_{i=1}^{N}$ is then given by the sum of the logarithm-transformations of the joint probability density function:

$$
\begin{aligned}
L(\mu_i, \sigma^2) &= \sum_{i=1}^{N} log f(Y_{1i} \cup y_{2i}) \\
&= \sum_{i=1}^{N} [log\Big(\frac{1}{\sigma^2 2\pi}\Big) - \frac{1}{2}\Big(\frac{Y_{1i}-\mu_i}{\sigma}\Big)^2 - \frac{1}{2}\Big(\frac{Y_{2i}-\mu_i}{\sigma}\Big)^2] \\
&= \sum_{i=1}^{N} \Big[log(1) - log(\sigma^2) - log(2\pi) - \frac{1}{2\sigma^2}\Big[\big(Y_{1i}-\mu_i\big)^2 + \big(Y_{2i}-\mu_i\big)^2\Big]\Big] \\
&= -\sum_{i=1}^{N} \Big[log(\sigma^2) + log(2\pi) + \frac{1}{2\sigma^2}\Big[\big(Y_{1i}-\mu_i\big)^2 + \big(Y_{2i}-\mu_i\big)^2\Big]\Big]
\end{aligned}
$$

(3):

The MLEs $(\hat{\mu}_i, \hat{\sigma}^2)$ can be derived by taking the corresponding first order derivatives of the combined Log-Likelihood function w.r.t. μ_i & σ^2 :

$$
\begin{aligned}
\frac{\partial L}{\partial \mu_i} : \frac{1}{2\hat{\sigma}^2}[2(Y_{1i}-\hat{\mu}_i) + 2(Y_{2i}-\hat{\mu}_i)] &\stackrel{!}{=} 0 \\
\frac{1}{2\hat{\sigma}^2}[2Y_{1i} - 2\hat{\mu}_i + 2Y_{2i} - 2\hat{\mu}_i] &= 0 \\
\frac{1}{\hat{\sigma}^2}[Y_{1i} - 2\hat{\mu}_i + Y_{2i}] &= 0 \\
\frac{1}{\hat{\sigma}^2}[Y_{1i} + Y_{2i}] &= \frac{2}{\hat{\sigma}^2}\hat{\mu}_i \\
\hat{\mu}_i &= \frac{1}{2}[Y_{1i} + Y_{2i}]
\end{aligned}
$$

This shows that the expected value of (Y_{i1} and Y_{i2}) can be estimated by taking the mean of (Y_{i1} and Y_{i2}) for each individual.

The estimator for the variance σ^2 can be calculated by taking the derivative w.r.t σ, setting it to 0 and solving for $\hat{\sigma}^2$.

$$
\frac{\partial L}{\partial \sigma} : \sum_{i=1}^{N} \Big[-\frac{2\hat{\sigma}}{\hat{\sigma}^2} + \frac{4\hat{\sigma}}{(2\hat{\sigma}^2)^2}[(Y_{1i}-\hat{\mu}_i)^2 + (Y_{2i}-\hat{\mu}_i)^2] \Big] \stackrel{!}{=} 0
$$

$$
\begin{aligned}
\sum_{i=1}^{N} \frac{2}{\hat{\sigma}} &= \sum_{i=1}^{N} \frac{1}{\hat{\sigma}^3} \left[(Y_{1i} - \mu_i)^2 + (Y_{2i} - \hat{\mu}_i)^2 \right] \\
\sum_{i=1}^{N} \hat{\sigma}^2 &= \sum_{i=1}^{N} \frac{1}{2} \left[(Y_{1i} - \hat{\mu}_i)^2 + (Y_{2i} - \hat{\mu}_i)^2 \right] \\
\frac{1}{N} \sum_{i=1}^{N} \hat{\sigma}^2 &= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \left[(Y_{1i} - \hat{\mu}_i)^2 + (Y_{2i} - \hat{\mu}_i)^2 \right] \\
\hat{\hat{\sigma}}^2 &= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \left[(Y_{1i} - \hat{\mu}_i)^2 + (Y_{2i} - \hat{\mu}_i)^2 \right]
\end{aligned}
$$

This shows that the estimator for the variance is the average of the sample variances, as we take the average over the squared difference between the variable of interest and the estimator. The expression can be be further simplified by substituting our mean-estimator $\hat{\mu}$:

$$
\begin{aligned}
\hat{\hat{\sigma}}^2 &= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \left[\left(\frac{Y_{1i} - Y_{2i}}{2} \right)^2 + \left(\frac{Y_{2i} - Y_{1i}}{2} \right)^2 \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \frac{1}{2} \frac{1}{4} \left[\left(Y_{1i} - Y_{2i} \right)^2 + \left(Y_{2i} - Y_{1i} \right)^2 \right] \\
&= \frac{1}{N4} \sum_{i=1}^{N} \left(Y_{1i} - Y_{2i} \right)^2
\end{aligned}
$$

(4):

The underlying bias of estimating $\hat{\mu}_i$ amounts to the difference between the expected value of the estimator and its true value:

$$
\begin{aligned}
Bias &= E[\hat{\mu}_i] - \mu_i \\
Bias &= E\left[\frac{Y_{1i} + Y_{2i}}{2} \right] - \mu_i
\end{aligned}
$$

Now, we use the distributional assumption of $E[Y_{1i}]$ & $E[Y_{2i}]$, which are both assumed to be μ_i, such that:

$$
Bias = \frac{1}{2}(\mu_i + \mu_i) - \mu_i = 0
$$

Thus the MLE of the expected value is unbiased, irrespectively of N. This means that as $N \to \infty$, the bias does not change as it does not depend on the number of included individuals.

(5):

As shown in (4), the MLE for μ has a bias of zero and accordingly is unbiased irrespectively of the number of observations N. This stems from the fact, that under the given example every person i is assigned an individual estimator for an individual expected value of Y_{1i} &Y_{2i}, similar to a case with individual fixed effects. Therefore, the number of used observations is limited to the number of time periods, in this case $t = 2$ and given the independence of Y_{1i} &Y_{2i}, increasing the number of observations does not contribute additional information when estimating the corresponding expected values.
The implied probability limit is given by:

$$\lim_{N\to\infty} \hat{\mu}_i = \hat{\mu}_i = \frac{1}{2}[Y_{1i} + Y_{2i}]$$

The expression for $\hat{\mu}_i$ is independent of N. This means the probability limit of the estimator $\hat{\mu}_i$ is the estimator itself and converges in $N \to \infty$ to the fixed expression $\frac{1}{2}[Y_{1i} + Y_{2i}]$. This implies that in the probability limit the estimator does not necessarily converge to the true value of the mean, which implies inconsistency. The reason for this result is the incidental parameters problem. As for each individual estimator of the mean, we only use T observations for each parameter. This means additional individuals does not yield more information for each parameter and leads to an increase in estimated parameters.

(6):

Equivalently to (4), the bias in estimating $\hat{\sigma}^2$ is given by the difference between the expected value of the estimator and the true variance. In the following derivation we use the definition of the variance:

$$\begin{aligned} Var(X) &= E(X^2) - E(X)^2 \\ E(X^2) &= Var(X) + E(X)^2 \end{aligned}$$

$$
\begin{aligned}
Bias &= E\left[\hat{\sigma}^2\right] - \sigma^2 \\
&= E\left[\frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}(Y_{1i}-\hat{\mu}_i)^2 + \frac{1}{N}\sum_{i=1}^{N}(Y_{2i}-\hat{\mu}_i)^2\right)\right] - \sigma^2 \\
&= E\left[\frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^{N}(Y_{1i}-\frac{1}{2}(Y_{1i}-Y_{2i}))^2 + \frac{1}{N}\sum_{i=1}^{N}(Y_{2i}-\frac{1}{2}(Y_{1i}-Y_{2i}))^2\right)\right] - \sigma^2 \\
&= \frac{1}{2}E\left[\frac{1}{N}\sum_{i=1}^{N}(Y_{1i}-\frac{1}{2}(Y_{1i}-Y_{2i}))^2 + (Y_{2i}-\frac{1}{2}(Y_{1i}-Y_{2i}))^2\right] - \sigma^2 \\
&= \frac{1}{2}E\left[\frac{1}{N}\sum_{i=1}^{N}(\frac{1}{2}(Y_{1i}-Y_{2i}))^2 + (\frac{1}{2}(Y_{2i}-Y_{1i}))^2\right] - \sigma^2 \\
&= \frac{1}{4}E\left[\frac{1}{N}\sum_{i=1}^{N}(Y_{1i}-Y_{2i})^2\right] - \sigma^2 \\
&= \frac{1}{4}E\left[\frac{1}{N}\sum_{i=1}^{N}Y_{1i}^2 - 2Y_{1i}Y_{2i} + Y_{2i}^2\right] - \sigma^2 \\
&= \frac{1}{4}\left(\frac{1}{N}\sum_{i=1}^{N}E[Y_{1i}^2] - 2E[Y_{1i}Y_{2i}] + E[Y_{2i}^2]\right) - \sigma^2 \\
&= \frac{1}{4}\left(\frac{1}{N}\sum_{i=1}^{N}\sigma^2 + \mu^2 - 2\mu^2 + \sigma^2 + \mu^2\right) - \sigma^2 \\
&= \frac{1}{2}\left(\frac{N}{N}\sigma^2\right) - \sigma^2 \\
&= -\frac{1}{2}\sigma^2
\end{aligned}
$$

This shows that the estimator of the variance has a non-zero bias and underestimates the true variance. Furthermore, we see that the bias in the variance is constant for all N. The bias does not converge to 0 for $N \to \infty$. The reason for this result is that additional individuals do increase the number of observations for the estimation of the sample variance, but also increase the number of estimated parameters as we have a parameter for the mean for each individual. This means that each included individual increases the estimation variance, which leads to a biased variance estimation even in the probability limit. This leads to a bias in the variance estimation due to the incidental parameters problem.

(7):

To derive the probability limit of the variance estimator $\hat{\sigma}^2$, we use the derived expression of the estimator and derive the limit when $N \to \infty$. The law of large numbers assumes that averages converge to expected values when $N \to \infty$. This is used in the following:

$$\begin{aligned}
\lim_{N\to\infty} \hat{\sigma}^2 &= \lim_{N\to\infty} \frac{1}{N4}\sum_{i=1}^{N}\left[(Y_{1i}-Y_{2i})^2\right] \\
&= \lim_{N\to\infty} \frac{1}{N4}\sum_{i=1}^{N}\left[Y_{1i}^2-2Y_{1i}Y_{2i}+Y_{2i}^2\right] \\
&= \frac{1}{4}\left[E[Y_{1i}^2]-2E[Y_{1i}Y_{2i}]+E[Y_{2i}^2]\right] \\
&= \frac{1}{4}\left[\sigma^2+\mu^2-2\mu^2+\sigma^2+\mu^2\right] \\
&= \frac{1}{2}\sigma^2
\end{aligned}$$

This shows that as $N \to \infty$ the variance estimator $\hat{\sigma}^2$ does not converge to the true variance. Hence, the estimator is inconsistent. Instead the estimator converges towards the true value divided by two. Our variance estimator $\hat{\sigma}^2$ converges in probability as the number of observations (N) goes towards infinity. This is because both Y_{1N} & Y_{2N} are distributed with the same variance across all individuals, resulting in a single variance estimator for the whole sample. Therefore, increasing the number of observations increases the number of usable information for variance. Nevertheless, as for each additional individual an additional parameter is estimated, the estimator $\hat{\sigma}^2$ does not converge to the true value.

(8):

In the following we consider the given model as a panel model with t=1,2 and argue that y_{i1} and y_{i2} can be seen as outcomes for different time periods. The panel variance will be estimated according to the variance estimator derived in exercise (3). Accordingly, we derive:

$$\begin{aligned}
\hat{\sigma}^2_{PANEL} &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{T}\sum_{i=1}^{T}\left[(y_{ti}-\hat{\mu}_i)^2\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left[(y_{1i}-\hat{\mu}_i)^2+(y_{2i}-\hat{\mu}_i)^2\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left[(\hat{\varepsilon}_{i1})^2+(\hat{\varepsilon}_{i2})^2\right]
\end{aligned}$$

Next, we consider the standard linear panel model with individual fixed effects (FE), for Y_{1i} & Y_{2i}:

$$y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it}$$

The predicted value of y in this model is then:

$$E[y_{it}] = \alpha_i + x_{it}'\beta$$

Accordingly, the given variance of y_{it} is given by:

$$Var(y_{it}) = E\left[\left(y_{it} - E[y_{it}]\right)^2\right] = E\left[\left(\varepsilon_{it}\right)^2\right]$$

By demeaning x_{it} & y_{it} in order to eliminate the FE, we obtain:

$$E[\ddot{y}_{it}] = \ddot{x}_{it}'\beta$$
$$Var(\ddot{y}_{it}) = E\left[\left(\ddot{y}_{it} - E[\ddot{y}_{it}]\right)^2\right] = E\left[\left(\ddot{\varepsilon}_{it}\right)^2\right]$$

We apply the estimator $\hat{\sigma}_{PANEL}$ equivalently to the FE case derived here. This gives us:

$$\begin{aligned}\hat{\sigma}^2_{FE} &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left[(\ddot{y}_{i1} - \ddot{x}_{i1}'\hat{\beta})^2 + (\ddot{y}_{i2} - \ddot{x}_{i2}'\hat{\beta})^2\right] \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left[(\hat{\ddot{\varepsilon}}_{i1})^2 + (\hat{\ddot{\varepsilon}}_{i1})^2\right] \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{1}{2}\left[\sum_{t=1}^{2}(\hat{\ddot{\varepsilon}}_{i1})^2\right] \\ &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{\ddot{\varepsilon}}_{i1})^2 \quad , for\ T = 2\end{aligned}$$

Hence, the given equation for $\hat{\sigma}^2_{FE}$ is equivalent to $\hat{\sigma}^2_{PANEL}$ when using demeaned variables. It follows that a consistent estimation of σ^2_{PANEL} with demeaned data is sufficient to consistently estimate σ^2_{FE} . Nevertheless, as shown in (7) the variance estimator is inconsistent, which implies that we cannot consistently estimate σ^2_{FE}. This implies why we do cannot estimate valid standard errors in this case, due to the incidental parameters problem. We have argued that the two outcomes can be seen as observations in different periods. As we estimate a μ for each individual, we would have to estimate N parameters. For each parameter we would have T observations to estimate a given parameter. The number of estimated parameters increases with the sample size. The estimator of the variance can, according to Cameron and Trivedi (2005, p.781), be defined as:

$$\frac{T-1}{T}\sigma^2$$

This is in line with our example of T=2 and the estimator $\frac{1}{2}\sigma^2$. This implies that consistent estimation of σ^2 can only be achieved by including more years such that $T \to \infty$.

To overcome this inconsistency of the estimated variance, we consider the solution suggested by Group 20 (Project 3). They suggest that the estimation of the expected values μ reduces the remaining degrees of freedom, which should be accounted for. As we estimate N $\hat{\mu}$'s, we should correct for N degrees of freedom in the denominator. This would result in the following estimator:

$$\begin{aligned}\sigma^2 &= \frac{1}{2N-N}\sum_{i=1}^{N}\left[(y_{1i}-\hat{\mu}_i)^2+(y_{2i}-\hat{\mu}_i)^2\right] \\ &= \frac{1}{N}\sum_{i=1}^{N}\left[(y_{1i}-\hat{\mu}_i)^2+(y_{2i}-\hat{\mu}_i)^2\right]\end{aligned}$$

We calculate the bias of the variance estimator with corrected degrees of freedom in the following:

$$\begin{aligned}Bias &= E\left[\hat{\sigma}^2\right]-\sigma^2 \\ &= E\left[\frac{1}{N}\sum_{i=1}^{N}(y_{1i}-\hat{\mu}_i)^2+\frac{1}{N}\sum_{i=1}^{N}(y_{2i}-\hat{\mu}_i)^2\right]-\sigma^2 \\ &= \frac{1}{2}E\left[\left(\frac{1}{N}\sum_{i=1}^{N}[(y_{1i}-y_{2i})^2]\right)\right]-\sigma^2 \\ &= \frac{1}{2}E\left[\left(\frac{1}{N}\sum_{i=1}^{N}[(y_{1i}^2-2y_{1i}y_{2i}+y_{2i})^2]\right)\right]-\sigma^2 \\ &= \frac{1}{2}\left[\left(\frac{1}{N}\sum_{i=1}^{N}[(E[y_{1i}^2]-2E[y_{1i}y_{2i}]+E[y_{2i})^2]]\right)\right]-\sigma^2 \\ &= \frac{1}{2}\left[\left(\frac{1}{N}\sum_{i=1}^{N}[(\sigma^2+\mu^2-2\mu^2+\sigma^2+\mu^2]\right)\right]-\sigma^2 \\ &= \frac{N}{N}\sigma^2-\sigma^2 \\ &= 0\end{aligned}$$

This shows that the correction of the degrees of freedom eliminates the bias and ensures consistent estimation of the standard errors, which makes it possible to calculate valid standard errors in the fixed effects model. For consistent estimation of σ^2 we assume homoskedastic error terms and no within-i correlation. This implies that the variance of the error terms can be estimated by:

$$\hat{\sigma}_{\varepsilon}^2 = \frac{1}{NT-N-K}\sum_i\sum_t \ddot{\varepsilon}_{it}^2$$

Problem 2

(1): Table 1 presents estimates about the effect of log-prices and home-market sales on overall log sales.[1] Car prices are measured in a common currency (SDR) to allow for cross-country comparison. We decided against price per income per capita to ensure a precise identification of the price effect, distinguished from an income effect. Table 1 shows a significant negative log-price effect on log sales across all specifications. Car sales in the firms' home country are on average significantly higher than in foreign markets except for regression (6). Yet, by step-wise introducing dummies capturing market, year, car and segment specific fixed effect (FE), estimates differ notably at times. This might suggest an omitted variable bias (OVB), also indicated by an increased explanatory power of models (2)-(6). Given that exogeneity implies the unconditional mean assumption, which is the identifying assumption of OLS, an OVB violates identification through endogeneity. Hence, we introduce a set of controls which explain sales but are also correlated with price. There are controls for car and country specific characteristics. Model (6) accounts for a large share of the overall variation, however, as each observation is given an estimate, OLS identification is violated through high dimensionality. For a given sample the number of parameters increases the mean squared prediction error through a reduction in the effective number of observations[2]. By estimating a coefficient for each car for each year and market, each observation is assigned a single estimate (similar to Q1). Hence, the coefficients are likely to be inconsistent such that we cannot identify the true parameters. For home and price to be identified we have to balance the trade-off between OVB and high dimensionality. As a solution, we propose to use controls to avoid OVB and chose to only employ segment rather than car type FE, to avoid reducing the variation too much. Thus, the price effect captures variation within segments, either sold at home or abroad, across countries, car types and years. When including additional FE, Lasso estimation could be used to restore prediction accuracy under high dimensionality through a penalty that decreases the overall variance by excluding irrelevant regressors. The found inconsistency of the variance estimator under FE in Q1.8 is resolved by degrees of freedom correction. We use dummies to estimate the FE regression and correct the degrees of freedom by the number of included regressors, which is equivalent to the correction in Q1.8.

[1]Home market sales is represented by a dummy indicating whether a car is sold in its country of origin.
[2]Defined by the number of observation over the number of parameters

Table 1: POLS Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	10.36***	10.46***	15.5***	9.87***	7.61***	16.2***	-17.27***	-1.00
	(0.173)	(0.164)	(0.247)	(0.284)	(0.2)	(0.567)	(0.776)	(0.78)
Log Price	-0.23***	-0.28***	-0.96***	-0.21***	-0.12***	-1.1***	-0.04***	-0.10***
	(0.02)	(0.019)	(0.03)	(0.028)	(0.023)	(0.073)	(0.038)	(0.04)
Home	2.19***	1.88***	2.26***	2.18***	2.13***	-1.87***	18.75***	1.87***
	(0.033)	(0.032)	(0.032)	(0.024)	(0.032)	(0.023)	(0.028)	(0.028)
Population[4]							0.02***	-0.02***
							(0.001)	(0.001)
VAT							-0.9***	-0.87***
							(0.21)	(0.209)
IncSharep50p90							23.07***	2.02*
							(10.384)	(1.033)
IncSharep90p100							31.62***	3.35***
							(0.628)	(0.628)
Dest.Exch.Rate[2]							-0.01***	-0.03***
							(0.002)	(0.002)
Exp.Exch.Rate[3]							0.03***	-0.001***
							(0.002)	(0.002)
GDP[1]							0.07***	0.14**
							(0.065)	(0.065)
Fuel Efficiency							-0.09***	-0.10***
							(0.01)	(0.01)
Horsepower							-0.03***	-0.03***
							(0.001)	(0.001)
Width							0.05***	0.05***
							(0.003)	(0.003)
Length							-0.28***	-0.32
							(0.062)	(0.073)
Height							0.18	0.25
							(0.236)	(0.249)
Seats							0.47***	0.46***
							(0.032)	(0.032)
Market FE	-	(YES)	-	-	-	(YES)	-	
Year FE	-	-	(YES)	-	-	(YES)	-	-
Car FE	-	-	-	(YES)	-	(YES)	-	-
Segment FE	-	-	-	-	(YES)	-	-	(YES)
obs	11504	11504	11504	11504	11504	11504	11504	11504
R^2	0.289	0.365	0.343	0.672	0.343	0.734	0.537	0.543

[1] average exchange rate of destination country relative to SDR
[2] average exchange rate of export country relative to SDR
[3] nominal GDP in common currency (SDR), in billion
[4] In million
$p < 0.1*, p < 0.05**, p < 0.01***$ & standard errors in parenthesis

(2): Model (8) in Table 1 estimates a linear coefficient vector γ for the effect of covariates on aggregate demand, which does not differ between consumers, years and cars. Thus, price and other factors can only be modeled as having a single homogeneous effect on aggregate car demand across car types, markets, years and individuals. The estimated coefficient for log prices can be interpreted as own-price elasticity, i.e. the percentage change in log-quantity sold if the log-price increases by one percent. The model assumes a homogeneous own-price sensitivity, i.e. demand for all cars depend equally on price changes. When including prices of other cars as an independent variable, the model assumes a homogeneous cross-price sensitivity for all car types. As the outcome is a continuous variable, the estimated coefficients and price elasticities have a cardinal interpretation. There is no restriction on the kind of variables included in this model. Including country, time or car invariant variables, does only affect the interpretation of the intercept. The discrete choice model (2) considers individual car choices. Consumers are assumed to choose the car that maximizes their utility function: $u_{citj} = x'_{itj}\beta_c + \varepsilon_{citj}$. As indicated by β_c, consumers are generally allowed to have individual-specific returns to certain attributes. The total quantity can depend differently on different factors, representing individual preferences and circumstances. As the outcome is a discrete variable, the coefficients do not provide a cardinal interpretation. The estimates can be interpreted as the change in the probability to buy a given car relative to an alternative due to an attribute. As the estimates are allowed to differ between consumers, different people can have different own- and cross price elasticities, which are not directly estimated by the coefficient β. The model does not allow for an intercept, so only variables that vary across individuals, time and cars are valid. Due to the independence of irrelevant alternative assumption (IIA) in the Logit model, additional cars reduce the expected demand for all cars equally, not just the closest substitutes. A cross-price elasticity unequal to 0 for very far substitutes is possible but counterintuitive.

(3): To estimate β consistently, we use the conditional logit (CL) model in market shares (Table 4 Col. 1 and 2). The choice probability is given by the probability that a given car is bought conditional on controls x_{itj}. As in the discrete choice model, we assume that the probability that a given car is chosen can be modeled as:

$$Pr(y_{citj} = j|X'_{it};\beta) = \left[\frac{exp(x'_{itj}\beta)}{\sum_{k\in J_{it}} exp(x'_{itk}\beta)}\right] \quad if \quad j \in J_{it}$$

By taking the ratio of choice probabilities of two cars l & j,we calculate the relative market shares:

$$\frac{s_{jit}}{s_{lit}} = \frac{\frac{exp(x'_{itj}\beta)}{\sum_{k\in J_{it}} exp(x'_{itk}\beta)}}{\frac{exp(x'_{itl}\beta)}{\sum_{k\in J_{it}} exp(x'_{itk}\beta)}} = \frac{exp(x'_{itj}\beta)}{exp(x'_{itl}\beta)}$$

Taking the logarithm on both sides derives the market share of car j relative to car l:

$$log[s_{jit}] - log[s_{lit}] = (x_{itj} - x_{itl})'\beta$$

Table 2: Results for question 3 and 4

Model Y=	Cond. logit quantity	Cond. logit quantity	NLS market share	NLS market share
log price (eurpr)	-2.31***	-1.04***	-0.80	-0.40
	(-70.36)	(-8.52)	(-0.86)	(-0.11)
home	3.25***	3.02***	1.80*	1.70*
	(60.56)	(60.67)	(1.78)	(1.73)
obs.	11,504	11,504	11,504	11,504
panel	unbalanced	unbalanced	balanced	balanced
controls		✓		✓
country FE				
car class dummies		✓		✓
exit flag			1	1

t values in parenthesis and p<0.1*, p<0.05**, p<0.01***.
Note that unbalanced panel was used.
Car class dummies (class 1 excluded) are included.
Note that NaN's are skipped when summing over individual criterion functions (q) to get Q.

Hence, β is identified through isolation in the regression and is estimated consistently under the exogeneity assumption of OLS. Exogeneity hinges on the uncorrelatedness of the regressors with the error term, ruling out reversed causality and OVB. I.e. car attributes cannot be adapted depending on market share and omitted variables cannot be correlated with attributes or market shares. This is highly questionable. β can be interpreted as the marginal effect (ME) of a one unit change in the attribute x_{itj} relative to x_{itl} on the relative market share of car j. This model can be estimated by OLS, which uses the difference between market share in logs as the dependent variable and the difference in prices as the independent variable. It is not crucial which cars are chosen for this analysis to estimate β. The model assumes a homogeneous coefficient in the population. The marginal effect of changing prices is a linear effect, which can be calculated at any two given points of the function, by extrapolating the linear effect to all cars. The reference car in our analysis is the car with the lowest quantity of sales in a given year and market .

(4): The given estimator is a non-linear least squares estimator as it minimizes the squared difference between the observed outcomes and the prediction of a non-linear estimation. To argue consistency, we show that the true parameter is the unique minimum of the criterion function, which implies that the model is identified. As we assume that the error has an expected value of 0, we can use that the law of iterated expectation which implies $E[\varepsilon|x] = 0$. Thus, the expected value of the error term given any function of X is 0. Identification is shown by defining the least squares as the criterion function: $q = [s - Pr(j|X;\beta)]^2$.

$$
\begin{aligned}
q &= [(s - Pr(j|X;\beta_0)) + (Pr(j|X;\beta_0) - Pr(j|X,\beta))]^2 \\
&= [(s - Pr(j|X;\beta_0)]^2 + 2(Pr(j|X;\beta_0) - Pr(j|X,\beta))\varepsilon + [Pr(j|X;\beta_0) - Pr(j|X,\beta)]^2
\end{aligned}
$$

We now take the expected value of the criterion function to show that it is identified.

$$E[q] = E[(s - Pr(j|X;\beta_0)]^2 + 2(Pr(j|X;\beta_0) - Pr(j|X,\beta))\varepsilon + [Pr(j|X;\beta_0) - Pr(j|X,\beta)]^2$$

$$E[[s - Pr(j|X;\beta)]^2] = E[(s - Pr(j|X;\beta_0)]^2 + [Pr(j|X;\beta_0) - Pr(j|X,\beta)]^2$$

Thus, the true parameter at $\beta = \beta_0$ is the unique minimizer of the criterion function, so the parameter is identified. The considered estimator is an m-estimator, as it minimizes a criterion function that is a sample average. Theorem 5.1 of m-estimators requires identification and that the criterion function converges to the true criterion function for consistency. Convergence is implied by the law of large numbers. Hence, model estimates are consistent. Column 3 and 4 in Table 2 show the corresponding results, where the effect of price on market share is negative as expected, while the effect of the home dummy is positive.

(5): Both elasticites are calculated by changing x marginally and using the resulting change in the predicted choice probability by the CL model in the following formula:

$$e_{itj} = \frac{\partial Pr(j|X_{it})}{\partial x_{itjk}} \frac{x_{itjk}}{Pr(j|X_{it})}$$

Table 3 shows the corresponding average own- and cross-price elasticities. The own-price elasticity describes the percentage change in demand for car j given a one percent change in price for car j. Home-market demand is more inelastic than overall demand to . The opposite holds for foreign markets. This seems to confirm that firms enjoy a home market

Table 3: Average elasticities

	eurpr
Own-price elasticity	-3.49 %
Own-price home	-3.38 %
Own-price foreign	-3.52 %
Cross-price elasticity	0.05 %
Cross-price home	0.13 %
Cross-price foreign	0.03 %

qvars = 'logpr', 'li', 'hp', 'wi', 'le', 'he', 'pl', 'home';
Class dummies included (1st class excluded).
Note that beta coefficients come from thetahat from Q4.
Note that NaN's are skipped when summing over cars.

advantage, as demand is less elastic w.r.t prices. This might imply that consumers have a home bias and care less about the price of domestic cars. Cross-price elasticity describes the responsiveness of demand for car j to a change in the price of another car $i \neq j$. The effect of a price increase of another car on demand is positive but neglectable. As we lack standard errors for the elasticities we cannot do inference and test for significant differences between home and foreign markets.[3] When interacting price with home, Table 4 shows that the negative price effect on market shares is a lot smaller in home markets, thus confirming the

[3] Such standard errors could potentially be computed by using the delta method.

Table 4: Results for heterogeneuos price effects

	Y = log market share
constant	-2.89***
	(-15.45)
log price (eurpr)	-0.29***
	(-13.00)
log price*home	0.20***
	(57.42)
obs.	11,504
R^2	0.23
Country FE	x
Car class dummies	✓

t values in parenthesis and p<0.1*, p<0.05**, p<0.01***.
Note that unbalanced panel was used.

results from Table 3. Increasing the price at home leads to a smaller reduction of demand compared to price increases abroad. The demand of car j decreases in its own price as predicted by standard economic theory.

(6): When looking at the overall mean difference between home and foreign market prices the POLS results in Table 5 indicate a significant but small difference. Home prices are on average 8% lower. In model (1), car makers' pricing would not be optimal as they do not exploit the less elastic demand in the home market. However, when examining within

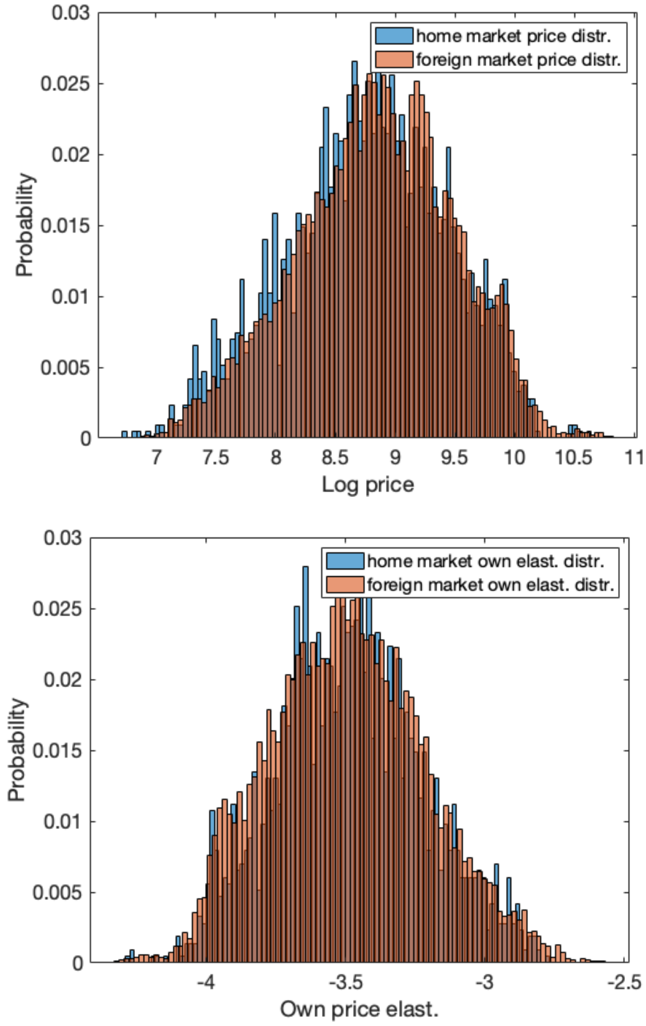
Table 5: OLS results for price on home

Y = log price	(1)	(2)	(3)	(4)
constant	8.84***	0.303	0.28	-1.28
	(1301.45)	(0.00)	(0.00)	(-0.00)
home	-0.08***	0.003	0.03***	0.01***
	(-5.29)	(0.30)	(3.18)	(3.23)
obs.	11,504	11,504	11,504	11,594
R^2	0.002	0.772	0.622	0.965
car type FE	x	✓	x	✓
year FE	x	x	✓	✓

t values in parenthesis and p<0.1*, p<0.05**, p<0.01***.
Note,that unbalanced panel was used.

variation by adding car or year FE the effect becomes positive, but insignificant in case of car FE. This could imply that differences in prices mainly come from the fact that the types of exported cars differ from home-sold cars, rather than from a home market advantage.
5 as the density for lower prices is higher among cars sold at home, which might come from different car types being sold at home and abroad. Car makers might think strategically about what cars to export, which explains why some types are not sold in the home market (or not exported). A change in the sign in model (3) indicates an unobserved time trend.

Figure 1: Density plot - prices and own elasticity



As shown in Q2.5, the difference between the home and foreign price elasticities is small, possibly explained by arbitrage under EU regulation.

(7): To identify the coefficient of interest in the CL model, we assume a homogeneous coefficient β across years and consumers. However, assuming that car demand is independent over time and consumers is critical and questions the results' validity. Given panel data, we argue that the main restriction of our analysis is that the CL model treats observations equally across time. The time dimension is not utilized and we cannot allow an unobservable time trend, which requires time dependend effects of β. Hence, the estimated price effect on the market share could also represent an underlying time trend. An example could be a general shift in preferences towards electric cars, increasing their relative market share, while their price decreases due to technological advancements. The CL model would estimate a negative relationship between price and market share, while the underlying reason is a general time trend. As a solution, we would like to allow for heterogeneous effects β, which could consist of a time-constant and time-varying part. To allow for correlation of unobservable factors over time that might determine the choice probability, Train (2009) proposes to use the mixed logit (ML) model. The ML model allows a utility model of the form (Train, 2009, p.147):

$$U_{njt} = \beta_{nt} x_{njt} + \varepsilon_{njt} \qquad with\ \beta_{nt} = b + \tilde{\beta}_{nt}$$

This allows the coefficient β_{nt} to have a time-varying term, e.g. 'climate-neutrality' could have a different effect in different years. The ML model is also a possibility to solve two other prevalent problems of the CL model. The CL model is limited by the IIA assumption, which restricts the substitution patterns, i.e. an increase in the car price for a low-price car (A) to equally affect the demand of a high (B) and low price car (C) when comparing B and C, which appears counterintuitive. Also, homogeneous effects of car attributes on the choice probability are assumed. It is reasonable to assume that the magnitude and sign of these effects differ between individuals. These restrictions are not required by the ML model according to Revelt and Train (1998) and Train (2009), which generalizes the assumed utility function of individuals.

Bibliography

Cameron, A Colin and Pravin K Trivedi (2005). *Microeconometrics: methods and applications*. Cambridge university press.

Revelt, David and Kenneth Train (1998). "Mixed logit with repeated choices: households' choices of appliance efficiency level." In: *Review of economics and statistics* 80.4, pp. 647–657.

Train, Kenneth E (2009). *Discrete choice methods with simulation*. Cambridge university press.