Kristina Birch, Jørgen Kai Olsen and Tue Tjur

Regression Models for Market–Shares

Summary. On the background of a data set of weekly sales and prices for three brands of coffee, this paper discusses various regression models and their relation to the multiplicative competitive–interaction model (the MCI model, see Cooper 1988, 1993) for market–shares. Emphasis is put on the interpretation of the parameters in relation to models for the total sales based on discrete choice models.

Key words and phrases. MCI model, discrete choice model, market–shares, price elasticity, regression model.

1. Deterministic models.

1.1. The simple model.

Let $b = 1, \ldots, B$ denote the brands of a given fmcg. In our example, the fmcg is coffee. Let Q_b denote the sales of brand b in a given short period (a week, say) in some welldefined area (e.g. a city or a state), and let p_b denote the (average) price of the brand b during that period. Conceptually, our starting point is the mathematically nice — though somewhat unrealistic — model

$$Q_b = e^{\alpha_b} p_b^{\beta_b} = \exp(\alpha_b + \beta_b \log(p_b))\,. \tag{1.1}$$

describing the demand as a standard Cobb–Douglas function of the price. Here, β_b is the — usually negative — *price elasticity*, which has the interpretation that a small relative increase of the price, say from p_b to $(1+\Delta)p_b$, is reflected by a decrease of the sales from Q_b to $(1+\beta_b\Delta)Q_b$. The parameter α_b, the "brand strength", aggregates all sorts of time–independent (or at least slowly varying) properties of the brand that are not related to the price, like quality, supply efficiency, advertising and other marketing efforts.

To see why this model is unrealistic on a market with more or less substitutable brands, it suffices to consider the case where all β_b are equal. In this case, the total sales of the fmcg takes the form

$$Q_{\cdot} = Q_1 + \cdots + Q_B = \sum_{b=1}^{B} \exp(\alpha_b + \beta \log(p_b))$$

where β is the common value of the β_b. It follows that a small relative increase of *all* brand prices $p_1, \ldots, p_B$ by a factor $1+\Delta$ would result in a relative decrease of the total sales $Q_{\cdot}$ by the factor $1+\beta\Delta$. This means that the total sales of the fmcg is determined by the same price elasticity as the sales of a single brand. This is obviously not realistic. If we think of coffee, a considerable increase of all prices is not likely to imply more than a moderate decrease of coffee consumption. This actually happened in 1994 and 1997, where a poor coffee harvest increased the prices on the Danish market by about 50%, without any observable change of the coffee consumption. Whereas, if we imagine a single brand increasing its price by 50% independently of all other brands, we would certainly expect that brand to disappear very quickly from the market.

Quite generally, price elasticities that describe an entire fmcg–category will usually be smaller in absolute value than the elasticities for single brands in that category on a market with competing brands. This can be handled by extension of the simple model by terms that take *cross–elasticities* into account. We return to this later, but let us first take a look at the consequences of the simple model when interpreted as a model for market–shares only. By this we mean the following. Consider the market–shares $S_1, \ldots, S_B$ defined by

$$S_b = \frac{Q_b}{Q_1 + \cdots + Q_B} .$$

According to (1.1), these market–shares are

$$S_b = \frac{\exp(\alpha_b + \beta_b \log(p_b))}{\sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(p_k))} \tag{1.2}$$

This is what Cooper (1988, 1993) calls the MCI–model (where MCI stands for multiplicative competitive–interaction). As a model for market–shares, it appears to be quite realistic, with the obvious reservation that the interpretation of the parameters β_b is different from and more complicated than indicated above. On the micro–level, the market share S_b has a simple interpretation as the probability that a (random) purchase occasion results in the choice of brand b. Moreover, the multinomial model coming out of this — a socalled *discrete choice model* — can be interpreted as a random utility model, where the customer's choice on each purchase occasion is the result of a maximization of a utility function with a random error term, see McFadden (1986).

The closest competitor to the MCI model is the MNL (multinomial logit) model, which can be written in exactly the same way, except that the prices enter directly instead of the logarithmized prices. We prefer the MCI model to the MNL model because it has the canonical property that a proportional increase of all brand prices will leave the market–shares unchanged, provided that all β_b are equal. This is not necessarily a realistic property, but it seems more realistic than the corresponding property for the MNL model, which is that a common *absolute* increase of all prices will leave the market–shares unchanged, provided that all β_b are equal. Anyway, most of what is said in the second part of this paper about the regression models derived from the MCI model applies as well to the MNL model when logarithmized prices are replaced with prices. For this reason, we lose very little by restricting our attention in the following to the MCI model. For our illustrating data set the difference between the two models is small, since the prices vary in a rather narrow interval (from 22 to 37 DKK) where the log–curve is almost linear.

Until now we have only been talking about one single short period. The estimation of the parameters α_b and β_b requires, of course, that we have observations over many such periods with varying prices. Let Q_{bt} and p_{bt} be observed sales and prices, respectively, for brands $b = 1, \ldots, B$ over a longer period divided into short periods (typically weeks) $t = 1, \ldots, T$. An immediate extension of our basic model to this scenario is

$$Q_{bt} = \exp(\delta_t + \alpha_b + \beta_b \log(p_{bt})) . \tag{1.3}$$

Here, the interpretation of the parameters α_b and β_b is the same as before, and with the same reservation that the model is unrealistic and the interpretation of β_b will change in a moment when we write down the induced model for market–shares. The parameter δ_t is added to take care of all sorts of time dependent effects that act in parallel on all brands. In the case of coffee (and similarly for many other fmcg's), the weekly sales are affected by such things as weather, season of year, television programs, non–working days etc. The implication of our assumption of a log–additive or multiplicative effect of period and brand is that such influences act proportionally on the sales figures for all brands, which means that the market–shares are not affected by any time trend.

The derived model for market–shares thus becomes (since the factor $\exp(\delta_t)$ cancels out)

$$S_{bt} = \frac{\exp(\alpha_b + \beta_b \log(p_{bt}))}{\sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(p_{kt}))} \tag{1.4}$$

1.2. A model including cross–elasticities.

Cross–elasticities can be viewed as an attempt to introduce the idea that the sales of a single brand is more price–sensitive than the aggregated sales in the entire fmcg–category. By the introduction of cross–elasticities we build into the model the property that a decrease of a brand's price will not only increase the sales of that brand, it will also decrease the sales of the competing brands. To account for this, parameters γ_{bk}, $b, k = 1, \dots B$, $b \neq k$, are introduced in the model as follows.

$$Q_{bt} = \exp\left(\delta_t + \alpha_b + \beta_b \log(p_{bt}) + \sum_{k:k\neq b} \gamma_{bk} \log(p_{kt})\right). \tag{1.5}$$

The interpretation of these *cross–elasticities* γ_{bk} is very similar to that of elasticities. If brand k increases its price from p_k to $(1+\Delta)p_k$, the sales of brand b will change from Q_b to $(1+\gamma_{bk}\Delta)Q_b$. Hence, we would expect cross–elasticities to be non–negative.

As a model for the total sales Q_{bt}, this model appears a lot more realistic than the simple model. But it has a rather peculiar property in the case of large price variations within brands, which indicates that it should be used with some care.

As we can easily see, the derivative of the total sales in the fmcg–category with respect to $\log(p_{bt})$ is

$$\frac{d(Q_{1t} + \cdots + Q_{Bt})}{d\log(p_{bt})} = \beta_b Q_{bt} + \sum_{k:k\neq b} \gamma_{kb} Q_{kt}\,.$$

Consider a situation where all brands on the market have the same strength parameter α, the same negative price elasticity β, and where also all the cross–elasticities equal a common positive value γ. If all prices are also equal, the total sales will, of course, take a common value Q_t, and the expression for the derivative above will take the form $\beta Q_t + (B-1)\gamma Q_t$. If this derivative is negative — which we expect it to be, since an isolated increase of the price of a single brand b is not likely to *increase* the total sales in the fmcg — we must have

$$\gamma < \frac{-\beta}{B-1} \, .$$

More generally, in order to ensure that this derivative is negative in a realistic domain of variation for the set of prices, the cross elasticities must be small compared to the direct elasticities. But the general expression for the derivative shows that if all cross–elasticities are positive, this derivative becomes positive when Q_{bt} is sufficiently small.

This property of the model can be described by the following scenario. Suppose, in a situation with B brands which are relatively equal in all respects, that (for some obscure reason) a single brand b suddenly increases its price by (say) a factor 2. If we have a market with highly substitutable brands and the values of the parameters are realistic, the sales Q_{bt} of that brand will decrease to something very small. Most of its sales will be taken over by the competitors, but the total sales of the fmcg will probably decrease slightly. Up to this point, the model is quite realistic, since there may be a small fraction of customers that are extremely loyal to this brand and prefer to buy it less frequently rather than switching to another brand. But assume now that the (by now almost bankrupt) producer of brand b desperately decides to increase the price again, this time by a factor 10! Now, that brand will essentially disappear from the market (in the real world it *would* disappear), but according to the model the sales of the remaining brands will increase even more than they did when p_{bt} was doubled, due to the positive cross–elasticities γ_{kb}. And since the sales of brand b were approximately zero also before this desperate action, the result is that the total sales of the fmcg will *increase*! This is certainly not realistic, and it shows that the model cannot be taken as more than a local approximation. The problem is that the price of a brand with an extremely small market–share can still have a high influence on the sales of the others. The immediate solution is to let very small brands influence other brands by very small cross elasticities. But if a brand develops from large to small (or vice versa), we have the problem. Hence, it is the assumption of *constant* cross–elasticities that creates the problem.

As a model for market–shares, this model takes the complicated form

$$S_{bt} = \frac{\exp\Big(\alpha_b + \beta_b \log(p_{bt}) + \sum_{k:k\neq b} \gamma_{bk} \log(p_{kt})\Big)}{\sum_{b'=1}^{B} \exp\Big(\alpha_{b'} + \beta_{b'} \log(p_{b't}) + \sum_{k:k\neq b'} \gamma_{b'k} \log(p_{kt})\Big)} \tag{1.6}$$

Notice that this model is overparameterized in the following sense. If for a given brand b, a constant, say κ, is added to the elasticity β_b and to all the cross–elasticities γ_{kb}, $k \neq b$, then the value of S_{bt} will remain unchanged, because this change of parameters is equivalent to the multiplication of the nominator and all terms of the denominater by the factor $\exp(\kappa \log(p_{bt}))$. This merely reflects the fact that a distinction between an increase of the sales of brand b and a proportional decrease of the sales of its competitors is impossible on the basis of market shares alone. In principle, this overparametrization can be removed by the (somewhat counter–intuitive) assumption that $\beta_b = 0$ for all b, thus pretending that the market–shares are determined by cross–elasticities only. However, there is no need to make this assumption, as long as we remember that the identifiable parameters are differences $\beta_b - \gamma_{kb}$ and $\gamma_{kb} - \gamma_{hb}$ between coefficients to $\log(p_{bt})$ in the model. The actual elasticities and cross–elasticities are not identifiable in the model for market shares.

This has the additional consequence, that the problem we had with the model as a model for total sales more or less disappears. Since we cannot measure the size of the cross–elasticities, we can — at any time — assume that they are small compared to the direct elasticities. The only exception from this occurs when the cross–elasticities describing a certain brand's influence on the other brands appear to be very different. In this case, they cannot all be close to zero. Moreover, this means that the brand in question can have a very high influence on the *proportion* between other brands' market shares, and this property persists even if the brand becomes vanishing due to a suicidal price policy. For this reason, the hypothesis that all cross elasticities γ_{kb} ($k \neq b$, b fixed) are equal, is an important one to be tested in the regression models in the second half of this paper. Under this hypothesis, we may (due to the overparametrization) formally assume that all cross–elasticities are zero, which means that we are back in the model (1.4).

In this context, it is important to notice that (1.4) can be derived from a model for the total sales, which appears a lot more realistic than (1.3). This is the topic of the next section.

1.3. An alternative interpretation of the simple model.

Consider the model for the total sales

$$Q_{bt} = \exp(\delta_t) \frac{\exp(\alpha_b + \beta_b \log(p_{bt}))}{\exp(\alpha_0) + \sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(p_{kt}))} \tag{1.7}$$

In this equation, the right hand side looks very much like the expression (1.4) for the market shares. The important difference is the term $\exp(\alpha_0)$ in the denominator and the multiplicative term $\exp(\delta_t)$ in front of the fraction. An easy way of understanding this model is in terms of a "pseudo brand" named brand 0. "Buying brand 0" means "not buying". In this interpretation,

$$Q_{0t} = \exp(\delta_t) \frac{\exp(\alpha_0)}{\exp(\alpha_0) + \sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(p_{kt}))}$$

is the number of units (e.g. 500g bags of coffee) that were *not* bought (because they were too expensive), and the multiplicative term $\exp(\delta_t) = Q_{0t}+Q_{1t}+\cdots+Q_{Bt}$ is the upper limit of the total sales $Q_{1t}+\cdots+Q_{Bt}$ in period t. Or — if you wish — the total sales of period t in a hypothetical situation, where all prices are so low that they have no influence on the consumers' decisions. It is easy to derive this model from a discrete choice (multinomial) model where the choice "no brand selected" (i.e. no purchase performed) is included.

An extended interpretation of this model could be that Q_{0t} includes the sales of brands that are not observed, if any. Provided that these unobserved brands keep their prices fixed (which we are more or less forced to assume if we don't know anything about them), this is just a way of collecting the terms $\exp(\alpha_k + \beta_k \log(p_{kt}))$ for such unobserved brands in the single constant term $\exp(\alpha_0)$. And, in fact, the interpretations of "buying brand 0" as "not buying" or "buying an unknown brand" are only the two extremes on a scale that could include psuedo–substitutes such as buying instant coffee (or tea or cocoa) instead of coffee.

The selection probabilites in this model are the "generalized market shares"

$$S'_{bt} = \frac{Q_{bt}}{Q_{0t} + Q_{1t} + \cdots + Q_{Bt}}\,, \; b = 0, 1, \ldots B\,.$$

The market shares S_{bt} (in the usual sense, i.e. without the additional term for brand 0 in the denominator) take the role as *conditional* probabilities of selecting a brand, given that a brand (on our list) is actually selected. Accordingly, the model for market shares derived from this model is also (1.4). But as we shall see now, (1.7) has much more realistic properties than the simple Cobb–Douglas model (1.3) which was our original justification of (1.4).

The (direct) price elasticities in this model are

$$e_{bb} = \frac{d\log(Q_{bt})}{d\log(p_{bt})} = \beta_b\,(1 - S'_{bt})$$

and the cross–elasticities are

$$e_{bk} = \frac{d\log(Q_{bt})}{d\log(p_{kt})} = -\beta_k S'_{kt}$$

Provided that all β_b are negative (which is assumed in the following), these elasticities and cross–elasticities (which are no longer constant) have the signs they should have. In addition, the cross–elasticities have the property requested in connection with the discussion of model (1.5) (where the cross–elasticities were constant) that small brands influence other brands via small cross–elasticities. This suggests that the model may — as opposed to (1.6) — have the desirable property that an increase of a single brand's price can never increase the total sales on the market. And, indeed, the derivative of the total sales with respect to the log–price of brand b turns out to be

$$\frac{d(Q_{1t} + \cdots + Q_{Bt})}{d\log(p_{bt})} = \exp(\delta_t)\beta_b S'_{bt} S'_{0t}$$

which is negative when β_b is negative.

To investigate this model's properties further it is useful to study its behaviour when the prices vary from period to period in such a way that the proportions between them is kept fixed. Thus, assume that

$$p_{bt} = \lambda_t p_b$$

where λ_t is the common proportionality factor. Then the expression for the total sales of $b \neq 0$ becomes

$$Q_{bt} = \exp(\delta_t)\frac{\exp(\alpha_b + \beta_b \log(\lambda_t p_b))}{\exp(\alpha_0) + \sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(\lambda_t p_k))}$$

and the market shares are

$$S_{bt} = \frac{Q_{bt}}{Q_{1t} + \cdots + Q_{Bt}} = \frac{\exp(\alpha_b + \beta_b \log(\lambda_t p_b))}{\sum_{k=1}^{B} \exp(\alpha_k + \beta_k \log(\lambda_t p_k))}$$

From the last expression we see, first of all, that if the β_b are equal, the factor $\exp(\beta \log(\lambda_t))$ cancels out, which means that the market shares are constant. If all prices become extremely small, the total sales will approach the maximal value $\exp(\delta_t)$, and if the prices grow beyond all limits, the total sales will fall to zero. But the market shares of the brands are constant all the way in both cases, as long as the prices are kept proportional.

For our interpretation of the β–parameters, a more interesting question is what happens under extreme prices when the β's are not equal. The expression for S_{bt} shows that if the β_b are negative and pairwise distinct and λ_t tends to 0, then the market share of the brand with the highest value of $|\beta_b|$ will tend to 1, all the others will tend to zero. Conversely, if all prices tend to ∞, the brand with the smallest value of $|\beta_b|$ will be alone on the (very small) market in the limit. More generally, we can say that a brand with a small value of $|\beta_b|$ is rather robust (as far as market shares are concerned) in a situation where the prices increase proportionally to a high level, whereas a brand with a high value of $|\beta_b|$ is robust in a situation where the prices decrease proportionally to zero. It is tempting to characterize the former as discount brands (even though we cannot conclude that they have low prices) which are taking over when all prices are so high that nobody can afford to by a decent brand, and the latter as high quality brands which most people prefer when they can afford it. In short, $|\beta_b|$ may in some respects be interpretable as a measure of quality.

2. Statistical models.

In the following, we discuss a number of multiple regression models and illustrate them by computations related to a data set. First of all, let us introduce the data.

In a period of 122 weeks from June 1994 to November 1996, the marketing firm Millward Brown Denmark collected a data set of weekly sales and prices of the three major coffee brands on the Danish market, which are Gevalia (G), Merrild (M) and Karat (K). In addition, the weekly sales and average price of all other brands ("Other (O)") were somehow estimated. This data set can be downloaded from

`http://www.mes.cbs.dk/~sttt/Coffee.txt`

Let $Q_{\mathrm{G}t}$, $Q_{\mathrm{M}t}$, $Q_{\mathrm{K}t}$ and $Q_{\mathrm{O}t}$ denote the sales of the four brands (in the following we use the term "brand" also for the remainder group O). Similarly, we let $p_{\mathrm{G}t}$, $p_{\mathrm{M}t}$, $p_{\mathrm{K}t}$ and $p_{\mathrm{O}t}$ denote the prices.

2.1. Regression models for logarithmized sales–ratios.

A consequence of the deterministic models (1.3) and (1.7) is that the proportion between market shares of two brands, say G and M, will satisfy the relation

$$\frac{S_{\mathrm{G}t}}{S_{\mathrm{M}t}} = \exp\left((\alpha_{\mathrm{G}} + \beta_{\mathrm{G}} \log(p_{\mathrm{G}t})) - (\alpha_{\mathrm{M}} + \beta_{\mathrm{M}} \log(p_{\mathrm{M}t}))\right)$$

or

$$\log \frac{S_{\mathrm{G}t}}{S_{\mathrm{M}t}} = \alpha_{\mathrm{G}} - \alpha_{\mathrm{M}} + \beta_{\mathrm{G}} \log(p_{\mathrm{G}t}) - \beta_{\mathrm{M}} \log(p_{\mathrm{M}t})\,.$$

This equation suggests a multiple regression model with the logarithmized proportion between the two brand sales as the dependent variable and the two logarithmized prices as explanatory variables. Thus, we consider the model

$$\log \frac{S_{\mathrm{G}t}}{S_{\mathrm{M}t}} = \alpha_{\mathrm{G}} - \alpha_{\mathrm{M}} + \beta_{\mathrm{G}} \log(p_{\mathrm{G}t}) - \beta_{\mathrm{M}} \log(p_{\mathrm{M}t}) + \varepsilon_t \tag{2.1}$$

where the error terms ε_t are assumed to be i.i.d. $N(0, \sigma^2)$. The estimates in this model become

Parameter	Estimate	Std.dev.	T	P
$\alpha_{\mathrm{G}} - \alpha_{\mathrm{M}}$	-0.09	1.342	-0.065	0.948122
β_{G}	-5.026	0.5723	-8.782	0.000000
β_{M}	-4.971	0.5991	-8.298	0.000000

As we can see, the estimates of β_{G} and β_{M} are very close to each other. A simple T–test for the hypothesis $\beta_{\mathrm{G}} = \beta_{\mathrm{M}}$ results in a T–value of -0.1384, which is certainly insignificant. Thus, we can reduce to the simple regression model with only one explanatory variable $\log(p_{\mathrm{G}t}/p_{\mathrm{M}t})$, in which the estimate of the common β becomes -5.004. Which means, for example, that if Gevalia doubled its price independently, the proportion $S_{\mathrm{G}}/S_{\mathrm{M}}$ would shrink by the factor $2^{-5.004} \approx 1/32$.

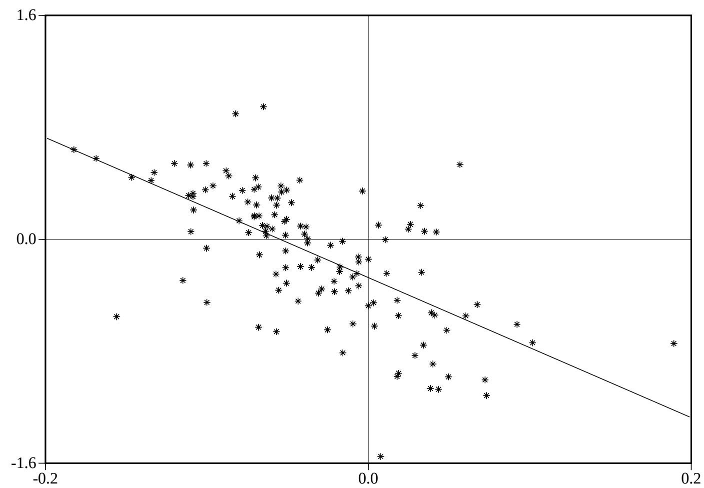

Figure 1. $\log(Q_{\mathrm{G}t}/Q_{\mathrm{M}t})$ against $\log(p_{\mathrm{G}t}/p_{\mathrm{M}t})$

Figure 1 shows the scatter plot corresponding to this simple regression. Notice that there is no indication of variance heterogeneity. This is a topic that we have not mentioned before, but actually it is not a triviality that a model with constant variance should be used here. If we had reasons to believe in the multinomial model, derived from the interpretation of the counts Q_{bt} as sums of independent binary variables (the discrete choice model), we would expect a variance derived in some complicated manner from the variance–covariance matrix of the multinomial distribution. Thus, if variance heterogeneity had been a problem, the standard solution would be to use a weighted regression with weights defined as suitable functions of the prices. Or, perhaps, a more advanced model with a variance specified as a function of the mean. However, the multinomial variation is not the dominating source of variation in this case, since the counts Q_{bt} are very large and the statistical variation stems from many other sources than multinomial variation. In other words, if we estimated the multinomial model directly we would almost certainly run into problems with heavy over–dispersion (we actually tried it and found an over–dispersion corresponding to a scale parameter of 120!). For this reason, there is no point in introducing more complicated variance structures as long as the variance appears to be constant. The scatter plot clearly confirms this, and so do various diagnostic tests that we shall not report here. We shall not mention this subject again for the models that are analyzed in the following, but we can assure the reader that all sorts of model checks have been made, and they all confirm that variance homogeneity can be assumed in these models.

We are not going to repeat this analysis for all the remaining pairs of brands, but for illustrative purposes we take a look at one of them, namely the model for the logarithm of the proportion between the sales of Gevalia and Karat. In this model, we obtain the estimates

Parameter	Estimate	Std.dev.	T	P
$\alpha_{\mathrm{G}} - \alpha_{\mathrm{K}}$	4.48	1.333	3.359	0.001051
β_{G}	−4.719	0.6383	−7.393	0.000000
β_{K}	−3.597	0.5870	−6.127	0.000000

The difference between the estimates of β_{G} and β_{K} is greater this time, and actually the T–test for $\beta_{\mathrm{G}} = \beta_{\mathrm{K}}$ results in a rather convincing rejection ($T = -2.828$, $P = 0.0055$). According to our interpretation of the model as a consequence of model (1.7), a possible explanation of this could be that most customers consider the quality of Karat lower than the quality of Gevalia. This is consistent with the fact that the average price of Karat over the period was 4% lower than the corresponding average for Gevalia, whereas the market share of Karat was, on average, only 57% of Gevalia's share. However, another — probably more realistic — explanation is that Karat is a local product, produced in Aalborg, and accordingly protected by a large group of highly loyal consumers in Northern Jutland.

With the cross–elasticity model (1.5) as our starting point, we obtain —

in exactly the same way — the regression model

$$\begin{aligned}
\log \frac{S_{\mathrm{G}t}}{S_{\mathrm{M}t}} &= \alpha_{\mathrm{G}} - \alpha_{\mathrm{M}} + (\beta_{\mathrm{G}} - \gamma_{\mathrm{MG}}) l_{\mathrm{G}t} - (\beta_{\mathrm{M}} - \gamma_{\mathrm{GM}}) l_{\mathrm{M}t} \\
&\quad + (\gamma_{\mathrm{GK}} - \gamma_{\mathrm{MK}}) l_{\mathrm{K}t} + (\gamma_{\mathrm{GO}} - \gamma_{\mathrm{MO}}) l_{\mathrm{O}t} + \varepsilon_t
\end{aligned} \tag{2.2}$$

for the logarithmized ratio between the market shares of Gevalia and Merrild. Here, we have introduced the short notation l_{bt} for $\log(p_{bt})$. From a statistical point of view, this is just an extension of the multiple regression model (2.1) by two additional explanatory variables which are the logarithmized prices of the other competitors on the market. But the interpretation of the parameters is different. For illustrational purposes, we prefer to consider the corresponding model for Gevalia and Karat,

$$\begin{aligned}
\log \frac{S_{\mathrm{G}t}}{S_{\mathrm{K}t}} &= \alpha_{\mathrm{G}} - \alpha_{\mathrm{K}} + (\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}}) l_{\mathrm{G}t} - (\beta_{\mathrm{K}} - \gamma_{\mathrm{GK}}) l_{\mathrm{K}t} \\
&\quad + (\gamma_{\mathrm{GM}} - \gamma_{\mathrm{KM}}) l_{\mathrm{M}t} + (\gamma_{\mathrm{GO}} - \gamma_{\mathrm{KO}}) l_{\mathrm{O}t} + \varepsilon_t
\end{aligned} .$$

The OLS–estimates in this model are

Parameter	Estimate	Std.dev.	T	P
$\alpha_{\mathrm{G}} - \alpha_{\mathrm{K}}$	4.65	1.470	3.166	0.001971
$\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}}$	−4.603	0.6977	−6.597	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{GK}}$	−3.705	0.6787	−5.459	0.000000
$\gamma_{\mathrm{GM}} - \gamma_{\mathrm{KM}}$	−0.384	0.8004	−0.480	0.632204
$\gamma_{\mathrm{GO}} - \gamma_{\mathrm{KO}}$	0.120	0.5617	0.213	0.831404

As we can see, the two explanatory variables $\log(p_{\mathrm{M}t})$ and $\log(p_{\mathrm{O}t})$ are insignificant, and the relevant F–test shows that they are actually jointly insignificant (P=0.12), which means that they can be removed from the model. Thus, we are back in the model for Gevalia and Karat corresponding to (2.1). But the interpretation of the parameters is different this time, and this suggests an alternative explanation of the significant difference between the coefficients to l_{Gt} and l_{Kt}. In the present model, this difference means that we have $\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}} < \beta_{\mathrm{K}} - \gamma_{\mathrm{GK}}$. Hence, there is no need to assume $\beta_{\mathrm{G}} < \beta_{\mathrm{K}}$, the explanation may as well be that we have $\gamma_{\mathrm{KG}} > \gamma_{\mathrm{GK}}$, which would be in accordance with our earlier remark that small brands should influence other brands by small cross–elasticities.

As to the hypothesis $\gamma_{\mathrm{GM}} - \gamma_{\mathrm{KM}} = \gamma_{\mathrm{GO}} - \gamma_{\mathrm{KO}} = 0$, which we did accept, this can be regarded as a partial version of the hypothesis that any brand influences the sales of all other brands by the same cross–elasticity. As mentioned earlier, this is a necessary property if the model should be valid in extreme situations, where a brand becomes very small due to high prices. If the price of such a brand affected two other brands with different cross–elasticities, we would be in the situation where the proportion between market shares for these two other brands would be very sensitive to the price policy of an extremely small competitor, which seems rather unrealistic. What we can see from our analysis is that neither the price of Merrild or the (average) price of the remainder group seem to be able to influence the *proportion* between the market shares of Gevalia and Karat.

By performing similar analyses for all the 6 pairs of brands, we could, in principle, perform all the pairwise comparisons required to confirm this hypothesis. But it would obviously be better to do this by a single test. For this and many other good reasons, it is better to analyze a data set like this by a model where the log–sales of *all* brands constitute the vector of responses, and where the log–prices of all brands occur as explanatory variables. This is the topic of the next section.

2.2. Regression models for the logarithmized sales.

A model for the sales of all brands can be derived from (1.5) in the following way. Instead of removing the week effect by formation of log–ratios between the sales of two brands, we simply include the week–to–week effect as a factor on 122 levels in a model for the logarithmized total sales,

$$\log Q_{bt} = \delta_t + \alpha_b + \beta_b \log(p_{bt}) + \sum_{k:k\neq b} \gamma_{bk} \log(p_{kt}) + \varepsilon_{bt}\,. \tag{2.3}$$

It is not difficult to see, that this is a generalization of (2.2) to the case $B > 2$. Actually, we get the same estimates as above of the parameters related to Gevalia and Karat. The difference is that we now have *all* brands in the same model, with the same variance. Notice that it is really a matter of taste whether we take $\log Q_{bt}$ or $\log S_{bt}$ as the reponse, since these two variates differ only by $\log Q_{.t}$ which is absorbed by the model term δ_t anyway (see Nakanishi and Cooper 1982, who propose essentially the same model under the label "dummy variable regression").

This model has a lot of parameters (122 week parameters δ_t, 4 direct elasticity parameters β_b and 12 cross elasticity parameters γ_{bk}), but the week parameters should be regarded as nuisance parameters here, and the most interesting hypothesis is that the γ_{bk} are independent of b, which formally implies that they can be set to zero. The estimates of the parameters of interest in this model are

Parameter	Estimate	Std.dev.	T	P
$\beta_{\mathrm{G}} - \gamma_{\mathrm{MG}}$	-4.862	0.6116	-7.951	0.000000
$\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}}$	-4.603	0.6116	-7.527	0.000000
$\beta_{\mathrm{G}} - \gamma_{\mathrm{OG}}$	-3.316	0.6116	-5.423	0.000000
$\beta_{\mathrm{M}} - \gamma_{\mathrm{GM}}$	-5.764	0.7016	-8.216	0.000000
$\beta_{\mathrm{M}} - \gamma_{\mathrm{KM}}$	-6.148	0.7016	-8.764	0.000000
$\beta_{\mathrm{M}} - \gamma_{\mathrm{OM}}$	-4.500	0.7016	-6.414	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{GK}}$	-3.705	0.5949	-6.228	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{MK}}$	-3.808	0.5949	-6.402	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{OK}}$	-3.418	0.5949	-5.747	0.000000
$\beta_{\mathrm{O}} - \gamma_{\mathrm{GO}}$	-0.737	0.4923	-1.497	0.135302
$\beta_{\mathrm{O}} - \gamma_{\mathrm{MO}}$	-1.803	0.4923	-3.663	0.000288
$\beta_{\mathrm{O}} - \gamma_{\mathrm{KO}}$	-0.617	0.4923	-1.254	0.210851

Notice, for example, that the estimate -4.603 of $\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}}$ is the same as we found in the analysis of $\log(S_{\mathrm{G}t}/S_{\mathrm{K}t})$. However, the standard deviation has changed slightly, since we are now working under the assumption that the variance is the same for all 4 brands.

The tests presented in this table are not particularly relevant. The tests of interest are those corresponding to the hypothesis that the price of each brand influence the sales of the others by the same cross–elasticity, which corresponds to the phenomena in the table that the first three estimates are approximately identical, the next three also, etc. Formally, this hypothesis is equivalent to the hypothesis that the γ's can be set to zero. But the relevant F–test for this hypothesis shows that it can not be accepted (F(8,351)=5.305, P=0.000003).

From the table of estimates, we can see that the most pronounced deviation from this hypothesis is that parameters of the form $\beta_b - \gamma_{kb}$ (b fixed) tend to be smaller when $k = \mathrm{O}$ than when k is a proper brand. For example, the price of Gevalia has more influence on the market share of Merrild and Karat than it has on the share of the remainder group. This is, perhaps, not surprising, since the aggregation of all other brands to a single group has some unpredictable consequences. The remainder group seems to contain brands with very loyal consumers, like brands associated with specific supermarket chains etc.

For this reason, we repeated the analysis for the three "proper" brands only. Notice that this is consistent with the model, since the derived model for the three brands and their market shares has exactly the same form as the one where "Others" is included as an additional brand with constant price. What we are doing here is just to pretend that we do not know about the sales and prices of "Others", thus including the sales of other brands in the category "not buying" (brand 0). The only problem with this is that we may lose some information when we ignore the average price for other brands. The estimated log–price coefficients in a model with only the three proper brands are

Parameter	Estimate	Std.dev.	T	P
$\beta_{\mathrm{G}} - \gamma_{\mathrm{MG}}$	-4.864	0.6503	-7.480	0.000000
$\beta_{\mathrm{G}} - \gamma_{\mathrm{KG}}$	-4.603	0.6503	-7.078	0.000000
$\beta_{\mathrm{M}} - \gamma_{\mathrm{GM}}$	-5.139	0.6800	-7.558	0.000000
$\beta_{\mathrm{M}} - \gamma_{\mathrm{KM}}$	-5.453	0.6800	-8.020	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{GK}}$	-3.730	0.6226	-5.991	0.000000
$\beta_{\mathrm{K}} - \gamma_{\mathrm{MK}}$	-4.061	0.6226	-6.523	0.000000

As we can see, these estimates are approximately pairwise identical, and the F–test for the hypothesis

$$\gamma_{\mathrm{MG}} = \gamma_{\mathrm{KG}}, \quad \gamma_{\mathrm{GM}} = \gamma_{\mathrm{KM}} \text{ and } \gamma_{\mathrm{GK}} = \gamma_{\mathrm{MK}}$$

actually confirms this ($F(3,236) = 0.285$, $P=0.84$). Thus we are back in the simple model (1.7). If, in this multiple regression model we add a term of the form $\gamma_{b\mathrm{O}} \log(p_{\mathrm{O}t})$, it becomes slightly significant ($P=0.035$). Hence, it is more or less a matter of taste whether this term should be included or not. We prefer to draw the conclusions in the model without this term. Hence, our final model is (1.4). The estimates of the parameters of interest in this model are

Parameter	Estimate	Std.dev.	T	P
$\alpha_{\mathrm{G}} - \alpha_{\mathrm{K}}$	4.52	1.217	3.717	0.000251
$\alpha_{\mathrm{M}} - \alpha_{\mathrm{K}}$	4.82	1.263	3.815	0.000173
β_{G}	-5.013	0.4485	-11.176	0.000000
β_{M}	-5.019	0.4695	-10.691	0.000000
β_{K}	-3.880	0.4144	-9.364	0.000000

In this model, the hypothesis $\beta_{\mathrm{G}} = \beta_{\mathrm{M}} = \beta_{\mathrm{K}}$ can not be accepted ($P=0.0017$), whereas it can obviously be accepted that $\beta_{\mathrm{G}} = \beta_{\mathrm{M}}$. Again, Karat seems to behave differently from the two other brands.

2.3. The inclusion of other covariates.

The general model (2.3) can easily be extended by the inclusion of other explanatory variables. If, for example, we had been in the possession of measures a_{bt} of the advertising efforts for brand b in period t, a straight-forward extension to

$$(2.3)\quad \log Q_{bt} = \delta_t + \alpha_b + \beta_b \log(p_{bt}) + \sum_{k:k\neq b} \gamma_{bk} \log(p_{kt}) + \eta_b \log(a_{bt}) + \varepsilon_{bt}$$

would enable us to analyse the effect of advertising. A hypothesis to be tested in this context could be the assumption of equal advertisment effects $\eta_{\mathrm{G}} = \eta_{\mathrm{M}} = \eta_{\mathrm{K}}$, which would mean that a proportional increase or decrease of the advertisment efforts for all brands would leave the market shares unchanged. This model is not quite realistic, since a brand with essentially no advertisment expenses might actually be able to survive on the market. Perhaps a more realistic model should have a_{bt} unlogarithmized in the linear expression. However, since we do not have such data, we shall resist from a discussion of this. This is just an example, and our point is merely that any covariate of potential relevance for the market shares can be included in the model. Covariates associated with brand and week can be included directly as above. For covariates representing market conditions that are common to all brands and depend only on the week, the inclusion makes sense only if these covariates occur in interaction with brand, i.e. with a separate coefficient for each brand, just like the term $\gamma_{b\mathrm{O}} \log(p_{\mathrm{O}t})$ that was added to the final model in the previous section. Otherwise such effects will be confounded with the general time trend parameter δ_t. A modification in a quite different direction would be to replace δ_t with week–associated effects (seasonal trends, indicators for special events etc.), but this is beyond the scope of the present paper.

As an illustrative example, we can try to extend the final model of the previous section with an explanatory variable that we do have, namely the vector of lagged prices. Thus, the idea is that the sales of a given brand in a given week is not only affected by the prices in that week, but also by the prices in the week before. This model can be written

$$\log Q_{bt} = \delta_t + \alpha_b + \beta_b \log(p_{bt}) + \eta_b \log(p_{b,t-1}) + \varepsilon_{bt}\,.$$

In fact, we have skipped a few steps here in assuming that the corresponding "lagged cross–elasticities" vanish. Anyway, the estimates in this model are

Parameter	Estimate	Std.dev.	T	P
$\alpha_{\mathrm{G}} - \alpha_{\mathrm{K}}$	4.24	1.201	3.532	0.000497
$\alpha_{\mathrm{M}} - \alpha_{\mathrm{K}}$	3.77	1.250	3.015	0.002855
β_{G}	−5.188	0.5508	−9.419	0.000000
β_{M}	−5.802	0.5793	−10.016	0.000000
β_{K}	−4.020	0.4980	−8.072	0.000000
η_{G}	1.931	0.5565	3.470	0.000619
η_{M}	2.747	0.5704	4.816	0.000003
η_{K}	1.832	0.4946	3.704	0.000265

At first sight, it may appear a bit surprising that the lagged prices are actually significant. One would hardly expect that many customers behave according to the prices they saw last week. But as we can see, the "lagged price elasticities" η_b are positive, and a straightforward explanation of this phenomenon is that customers fill up their private stocks when prices are low and consume the stocks when prices are high. Since private coffee stocks are usually rather limited in size, this means that a change of the price from one week to the next will result in a kind of overreaction. If we rewrite the model as

$$\log Q_{bt} = \delta_t + \alpha_b + (\beta_b + \eta_b)\log(p_{bt}) + \eta_b \log\left(\frac{p_{b,t-1}}{p_{bt}}\right) + \varepsilon_{bt}\,.$$

it becomes more transparent what is going on. The parameter $\beta_b + \eta_b$ is the true elasticity, in the sense that it determines the dependence of the equilibrium price in a period where the price is kept constant, so that $\log \frac{p_{b,t-1}}{p_{bt}} = 0$. The parameter η_b determines the short–term reaction on a price change. This actually means that the "long–term elasticities" are smaller in absolute value than suggested by the estimates in the model without lagged prices. A phenomenon which should certainly be taken into account when the estimated model is used as a strategic tool for optimization of price policy. We have tried to include the double–lagged prices also (i.e. the prices from two weeks before), but their estimates turned out not to be significantly different from zero.

References.

Cooper, Lee G. and Nakanishi, Masao (1988).
Market–Share Analysis.
Kluwer Academic Publishers.

Cooper, Lee G. (1993).
Market–Share Models.
in *Handbook in Operations Research and Management Science* **vol. 5** (Marketing, ed. J. Eliashberg and G. L. Lilien), 259–314.

D. McFadden (1986)
The choice theory approach to market research
Marketing Science **Vol. 5** pp 274-297.

Nakanishi, Masao and Cooper, Lee G. (1982).
Simplified Estimation Procedures for MCI Models.
Marketing Science **Vol. 1** pp 314–322.

Kristina Birch (kristina@cbs.dk), Center for Statistics
Jørgen Kai Olsen (jko.marktg@cbs.dk), Department of Marketing
Tue Tjur (tuetjur@cbs.dk), Center for Statistics

Copenhagen Business School
Solbjerg Plads 3
2000 Frederiksberg
DENMARK