

On Dynamics in Discrete Choice Models of Product Differentiation with Market Level Data: A State-Dependence Approach to Myopic Habits

Lapo Filistrucchi
European University Institute & University of Siena
October 2006

Preliminary

ABSTRACT

This paper investigates the possibility to introduce and identify a habit component in the structural model of utility at the basis of discrete choice models of product differentiation. In doing so it discusses the wider issue of persistence in consumers' choices, which can be due not only to habits (or more generally switching costs) but also to persistence in unobserved consumers' characteristics and/or the presence of network effects.

A model which takes (myopic) habits or variety seeking into account is first developed then estimated on data on the market for daily newspapers in Italy. In order to estimate it an approximation is proposed, whose performance is compared in a simulation to a dynamic AR1 model obtained by adding a lag of the dependent variable to the aggregate demand equation derived from a logit model. If the true model is the proposed one, in order to identify the price both models perform better in case of variety-seeking than in case of habits. However, the approximation is reasonably good in order to estimate the habit/variety seeking parameter and outperforms the AR1 model in recovering the price parameter only when the habit/variety seeking parameter is not too big, that is when habits or variety seeking are not too important.

Keywords: discrete choice models, differentiated products, dynamics, myopic habits, switching costs, variety seeking, market level data, daily newspapers.

JEL Classification: C22, C23, L40, C25

1. INTRODUCTION

In many markets consumers' behaviour is characterized by habits. As habits typically give rise to switching costs, their measurement is not only an interesting topic of research per se, but also relevant to understand the diffusion of new products, the intertemporal pricing behaviour of firms and the level of competition in those markets.

In particular, assessing market power in a market with differentiated goods requires, among other things, the measurement of the residual price elasticity of demand. In order to estimate the latter, discrete choice models of product differentiation are used more and more often by antitrust authorities and their counterparts around the world.

As consumer level data are not always available, these models are often estimated on aggregate level data, in such a way as explained in Berry (1994) and Nevo (2000). Given that the usual available micro panel derives from the observation of sales of many products in many markets but at few points in time, so far the literature has been treating observations relating to the same market at different points in time as observations of different markets. Although this might not appear problematic in the case of durable goods, such as cars¹, when consumers in the same market at two points in time are probably different, it is clearly not realistic in the case of non-durable goods, such as cereals², daily newspapers³ or magazines⁴.

The issue appears relevant because recent empirical studies in marketing using consumer level data have been able to identify the effect of habits and show the bias that ignoring it can cause in the estimates of the elasticity of demand with respect to prices, although results on the direction of the bias are ambiguous and depend on the model used⁵. Whether the bias is present also in models that use market level data

¹ As in Brenkers & Verboven (2002).

² As in Nevo (2001).

³ As in Filistrucchi(2005), where I recognise the issue, but proposes a solution which is not satisfactory since it is not consistent with the random utility model, as discussed below.

⁴ As in Kaiser (2003), who does not discuss the issue but uses the same method I use in Filistrucchi(2005).

⁵ See Erdem (1996), Roy et al (1996), Seetharaman (1999) and Che et al.(2004).

and, if so, what are its direction and magnitude is crucial to the assessment of market power: if the estimated price elasticity is higher than the true one, the observation of a high price may lead to the conclusion that there is collusion in the market; if instead the estimated price elasticity is lower than the true one, observing low prices may mislead into judging that there is competition.

In this paper I therefore investigate the possibility to introduce and distinguish different sources of dynamics in the structural model of utility at the basis of the discrete choice model of product differentiation. In particular I focus my attention on myopic habits (and variety seeking). After a brief description in the next section of the simplest model, namely the logit model, in section 3 the usual ways to deal with the issue of dynamics in the estimation of demand are discussed and shown to be unsatisfactory in a structural approach. In section 4, I propose a new way to introduce (myopic) habits in the random utility model, derive the econometric model of demand and an approximation for it; in section 5 I estimate the approximation on data for daily newspapers in Italy and in section 6 I compare the performance of the approximation to an AR1 specification when the true model is the model I propose. Section 7 concludes.

2. THE LOGIT MODEL

In order to exemplify the issue at stake, I introduce the simplest discrete choice model of product differentiation, namely the logit model.

The starting assumption, which is common to fixed coefficients models of product differentiation in general, is the following functional form of consumer i indirect utility from consuming good j at time t in market m :

$$u_{ijtm} = \alpha(y_{itm} - p_{jtm}) + \bar{x}_{jtm}\bar{\beta} + \xi_{jtm} + \varepsilon_{ijtm} \quad (1)$$

where y_{itm} is the income of consumer i at time t in market m ⁶, p_{jtm} is the price of good j at time t , \bar{x}_{jtm} is a vector of observed characteristics, ξ_{jtm} is an unobserved (by the econometrician) characteristic, ε_{ijtm} is a mean-zero stochastic term, α is consumers

⁶ Many models also include a sub index, say f , for firm when firms are multi-product (allowing therefore for the possibility of a firm, or logo, fixed effect). The indexes used here are simply those necessary to deal with the dataset used in the empirical example below.

marginal utility from income (and marginal disutility from price) and $\bar{\beta}$ is a vector of taste coefficients⁷.

The marginal utility from income and the taste parameters are therefore assumed fixed across consumers and, as a result, consumers heterogeneity enters only through the separable additive random shock ε_{ijtm} .

As consumers may decide not to consume any of the goods considered, an outside good is introduced⁸, consuming which yields to consumer i at time t the indirect utility:

$$u_{iotm} = \alpha y_{itm} + \xi_{otm} + \varepsilon_{iotm} \quad (2)$$

Since the outside good is a composite one, its price and its characteristics are not defined. The price of the outside good is then assumed to be equal to zero⁹ and all its characteristics are assumed to be unobservable¹⁰. But as ξ_{otm} is not identified, the standard practice is to set it equal to 0, which, as the term y_{itm} eventually vanishes because it is common to all products, is equivalent to normalizing the mean utility from the outside good to zero¹¹.

Consumers mean utility, δ_{jtm} from consuming good j at time t in market m is instead given by:

$$\delta_{jtm} = E_i[u_{ijtm}] = \alpha(y_{itm} - p_{jtm}) + \bar{x}_{jtm}\bar{\beta} + \xi_{jtm} \quad (3)$$

Consumers are then assumed to purchase the good which gives them the highest utility and to be never indifferent between buying one or another good. For

⁷ Such an indirect utility specification assumes a quasi-linear utility function, free therefore of wealth effects, which sounds plausible for many goods, although not for all. It also assumes that both observed and unobserved product characteristics are the same across all individuals in the market, and thus rules out for instance different prices for different consumers in the same market.

⁸ In the absence of an outside good, the model would assume consumers to be forced to choose one of the inside goods. and demand would depend only on differences in prices. Therefore, we would assume that a general increase in prices would not decrease aggregate goods sales, which would be unfortunate when estimating the elasticity of demand with respect to prices, since in general, when price increases, some consumers buy something else but some other do not buy anything.

⁹ In other words, the consumer is assumed to be choosing between buying one of the above goods or not buying it, not between buying one of the goods above or buying something else. The decision of whether to buy something else is not simultaneous.

¹⁰ Or equivalently, both the price and the characteristics are assumed to be unobservable and therefore included in ξ_{otm} .

¹¹ So that neither the market-shares of the outside good nor those of the inside goods respond to changes in the price of the outside good, unless time fixed effects are used. Also, the price of the outside good does not respond to the prices of the inside goods. Nor the prices of the inside goods respond to the price of the outside good.

convenience, it also assumed that consumers do not choose more than one good, although this behaviour can sometimes be observed.¹²

Decomposing $\xi_{jtm} = \zeta_j + \zeta_t + \zeta_m + \eta_{jtm}$ ¹³, with η_{jtm} a random shock independent of ε_{ijtm} , allows to model product, market and time specific unobserved characteristics. Assuming ζ_j , ζ_t and ζ_m to be unknown parameters, leads to a fixed effects model¹⁴.

Using product fixed-effects allows to better estimate product differentiation, as in this type of models the product fixed effects are usually believed to capture also the vertical component¹⁵.

The inclusion of time fixed effects is instead justified by the necessity to control for the change through time in the utility of the outside good. Given that it is by definition a composite good, there are many reasons why the latter may change in time. It might, for instance, be due to the appearance of new goods or to the change in the characteristics of the goods which are included in the composite outside good. But also changes in the average consumer taste may change the relative utility of the choice to buy the inside goods with respect to the outside one. As its utility is by construction normalised to zero, the absence of time fixed effects or some equivalent control might often raise questions of identification for the estimated coefficients¹⁶.

It is then assumed that ε_{ijtm} is i.i.d. across consumers and products and that it is distributed according to a type I extreme value distribution. Assuming ε_{ijtm} to be i.i.d.

¹² This assumption is common to most empirical studies on differentiated products markets with market level data, the usual justification being that assuming otherwise is econometrically very cumbersome and the assumption is instead, at worst, a reasonable approximation. That's because multiple purchases, though by no means uncommon, are not a rule and in any case even if two products are bought together they are then often consumed at different times, so that the multiple purchase is just an organisational device. Furthermore, if the potential market size is defined large enough, we might also claim that observed multiple purchases by the same individual are the result not only of his choice but also of somebody else's decision. For instance, if potential market size is defined as total population instead of number of households, the observation that an individual buys two goods might, at least to a certain extent, reflect the fact that he is buying one good for himself and another for another member of his households who asked him to.

¹³ Alternatively one could decompose it as $\xi_{jtm} = \zeta_{jt} + \zeta_m + \eta_{jtm}$ or $\xi_{jtm} = \zeta_j + \zeta_{tm} + \eta_{jtm}$ or $\xi_{jtm} = \zeta_{jm} + \zeta_t + \eta_{jtm}$. The former would allow for product ranking to vary in time, the second for the utility of the outside goods to vary across markets and the third for product ranking to change across markets.

¹⁴ If instead ζ_j , ζ_t and ζ_m were assumed to be random variables, then we would have a random effects model. Whenever the assumption of no correlation between the observed product characteristics and the unobserved characteristics, which lie at the basis of the random effect estimation, do not appear plausible, a fixed effect specification should be used.

¹⁵ See Nevo(2001).

¹⁶ An alternative is to use a time trend as in Filistrucchi and Argentesi (2004).

across consumers rules out, in particular, the possibility that individual specific random shocks are correlated across products or equivalently only allows shocks to demand to be correlated across products if they are not individual specific.

All the assumptions above lead to an aggregate logit model. In fact in this model, at a given point in time t in a market m , each individual i is defined by a vector of random shocks $\bar{\varepsilon}_{itm} = (\varepsilon_{i0tm} \varepsilon_{i1tm} \dots \varepsilon_{iJtm})$. As a result, the set of individuals who choose product j

at time t is implicitly defined as $I_{jtm}(X_{tm}, \bar{p}_{tm}, \bar{\xi}_{tm}, \alpha, \bar{\beta}) = \{\bar{\varepsilon}_{itm} \mid u_{ijtm} \geq u_{iktm} \forall k \neq j\}$ where

$X_{tm} = (\bar{x}_{1tm}, \dots, \bar{x}_{Jtm})$, $\bar{p}_{tm} = (p_{1tm}, \dots, p_{Jtm})$ and $\bar{\xi}_{tm} = (\xi_{1tm}, \dots, \xi_{Jtm})$.

The market share of product j at time t in market m is therefore given by:

$$s_{jtm}(X_{tm}, \bar{p}_{tm}, \bar{\xi}_{tm}, \alpha, \bar{\beta}) = \text{Prob}\{u_{ijtm} \geq u_{iktm} \forall k \neq j\} = \text{Prob}\{\varepsilon_{iktm} - \varepsilon_{ijtm} \leq \delta_{ijtm} - \delta_{iktm} \forall k \neq j\} = \int_{I_{jtm}} dP_{\varepsilon}$$

which leads to¹⁷

$$s_{jtm} = \frac{\exp(\delta_{jtm})}{1 + \sum_{k \neq 0} \exp(\delta_{ktm})} \text{ for any good } j^{18}$$

(4)

and

$$s_{0tm} = \frac{1}{1 + \sum_{k \neq 0} \exp(\delta_{ktm})} \text{ for the outside option}^{19}.$$

(5)

As a result the own and cross marginal effects of price on market shares are:

$$\frac{\partial s_{jtm}}{\partial p_{jtm}} = -\alpha(1 - s_{jtm})s_{jtm}$$

(6)

and

$$\frac{\partial s_{jtm}}{\partial p_{ktm}} = \alpha s_{ktm}s_{jtm} \text{ with } k \neq j.$$

(7)

So that the own and cross price elasticities of the market shares are respectively:

$$\eta_{jjtm} = \frac{\partial s_{jtm}}{\partial p_{jtm}} \frac{p_{jtm}}{s_{jtm}} = -\alpha p_{jtm}(1 - s_{jtm})$$

(8)

¹⁷ See Train (1993).

¹⁸ Note that the term αy_{it} drops out as it is common to all options.

¹⁹ It should be noted that the presence of an outside good with market share s_{0tm} means that observations of goods sales are not sufficient to calculate market shares. As a result it is necessary to introduce the concept of potential market size as distinct from the observed market size which would simply be the sum of the observed goods sales. Potential market size is then assumed or estimated by parameterising it as depending on some market level data. Thus the definitions of market size and market shares are different from the ones commonly used.

and

$$\eta_{jkm} = \frac{\partial s_{jtm}}{\partial p_{ktm}} \frac{p_{ktm}}{s_{jtm}} = \alpha p_{ktm} s_{ktm} \quad \text{with } k \neq j. \quad (9)$$

The model thus predicts a different demand and therefore different marginal effects and elasticities for each product j , at each time t and in each market m .

As it is well-known to the empirical industrial organization literature²⁰, the use of an aggregate logit model to estimate demand places restrictive assumptions on own and cross price elasticities or equivalently on own and cross marginal effects of price.²¹ In particular cross derivatives will be symmetric and all other products will have the same cross elasticity with respect to the price of a given product, while own derivatives depend on own market shares only and own elasticities on own market shares and own prices only.²²

Dividing each good market share by the outside good market share, simplifying and taking natural logarithms leads to the following market shares estimation equation:

$$\delta_{jtm} \equiv \ln(s_{jtm}) - \ln(s_{otm}) = \bar{x}_{jtm} \bar{\beta} - \alpha p_{jtm} + \xi_j + \xi_m + \xi_t + \eta_{jtm} \quad (10)$$

Given estimates of the relevant parameter α , obtained with instruments if endogenous, the marginal effects and/or the elasticities above can be calculated. From these it is then possible to recover estimates of the mark-ups, for each product j at each time t and in each market m , under different hypothesis on the level and form of competition in these markets²³. If a rough measure of the real mark-ups, such as the

²⁰ See for instance Berry (1994), Berry et al. (1995) and Nevo (2000)

²¹ The same restrictions are placed on the marginal effects and elasticities with respect to any characteristic.

²² That's because additive separability together with the i.i.d. structure of the random shocks, when the amount of a positive (negative) characteristic of a good is raised (decreased), requires consumers to substitute towards other products in proportion to market shares, regardless of the other products' characteristics See Berry (1994) and Nevo (2000)

²³ For instance, if the choice variable is p , the benchmark case of Bertrand-Nash competition would imply a price-cost margin

$$p_{jtm} - c_{jtm} = - \frac{s_{jtm}}{\frac{\partial s_{jtm}}{\partial p_{jtm}}} = 1/\alpha(1 - s_{jtm})$$

While the opposite benchmark of joint profit maximization would imply

$$p_{jtm} - c_{jtm} = - \frac{s_{jtm}}{\frac{\partial s_{jtm}}{\partial p_{jtm}}} - \sum_{k \neq j} (p_{ktm} - c_{ktm}) \frac{\partial s_{ktm}}{\partial p_{jtm}} \frac{\partial s_{jtm}}{\partial p_{jtm}} = 1/\alpha(1 - s_{jtm}) + \sum_{k \neq j} (p_{ktm} - c_{ktm}) s_{ktm} / (1 - s_{jtm})$$

Substituting the estimated parameter α in the expressions above yields estimates of the mark-ups under Bertrand Nash competition and under joint profit maximization.

average mark-up in the industry in a given period, is available²⁴, they can be used to choose the most likely supply model, as discussed in Nevo (2001). Alternatively the demand equation above can be estimated simultaneously first with a pricing equation under Bertrand Nash competition, then with a pricing equation under perfect collusion. It is then possible to test statistically which of the two models of competition better fits the data, using for instance Vuong(1989) tests for model selection and non-nested hypothesis.

Alternatively in competition policy the estimates for the price elasticities obtained from the estimates of the parameter α are used to simulate mergers between firms who produce two or more products or to define a product relevant market using the SSNIP test.

In all cases it is crucial that the estimated price parameter $\hat{\alpha}$ is not biased or at least consistent.

3. TRADITIONAL WAYS TO DEAL WITH DYNAMICS

Yet, it is clear that estimating an equation such as

$$\ln(s_{jtm}) - \ln(s_{om}) = \bar{x}_{jtm} \bar{\beta} - \alpha p_{jtm} + \zeta_j + \zeta_m + \zeta_t + \eta_{jtm} \quad (11)$$

leaves no scope for dynamics in aggregate demand.

There are however many reasons why dynamics should matter. Persistence in aggregate demand can theoretically be explained by persistence of unobserved consumers' characteristics, by habits (or equivalently switching costs) and/or by consumption externalities (or network effects), the latter affecting either consumers' evaluation of the product (that is the indirect utility of consuming it) or consumers' knowledge about the product (that is the choice set). Interestingly, the same forces might also help to explain diffusion of a product in a market, whereas a model which does not take them into account would unsatisfactorily predict the instantaneous jump of a new product to its "right" market share.

The estimated residual price elasticity of demand, and therefore the assessment of market power, on the basis of the model above might be strongly biased in markets

²⁴ Of course, the more disaggregated is the rough measure, the better it is, as a mark-up is calculated for each product j at each point in time t and in each market m .

where such dynamic aspects are relevant. That's because whether habits, network effects or unobserved consumers' characteristics are at play, market shares at time t would not only be a function of goods' prices, characteristics and random shocks to demand at time t , but also of market shares in the past. Similarly for the ratio of any inside good market share with respect to that of the outside good, which is the dependent variable in the above equation. So that there might then be an omitted variable problem. Such an omitted variable problem, when estimating the static equation above as it is, might lead to substantial autocorrelation in the residuals, even if shocks to demand are not correlated through time.²⁵

From a purely econometric point of view, the realization that current market shares should be a function of past market-shares would naturally lead to estimating the equation with the inclusion of one or more lags of the dependent variable and/or of the explanatory variables or, alternatively, starting from the observation of substantial serial correlation in the residuals, to the assumption of an auto-correlated random shock, all of which are usual ways to take into account dynamics, and therefore also habits, when estimating a structural model of demand which starts from the specification of an aggregate demand equation²⁶. Yet these specifications raise critical theoretical issues, when trying to interpret the resulting dynamic aggregate demand in terms of the random utility model.

In particular including one lag of the dependent variable in the estimation equation leads to

$$\ln s_{jtm} - \ln s_{otm} = \alpha(y_{it} - p_{jtm}) + x_{jtm}\bar{\beta} + \xi_{jtm} + \rho(\ln(s_{jt-1m}) - \ln(s_{ot-1m})) + \varepsilon_{ijtm} \quad (12)$$

but such a specification only imposes the short and medium-run marginal effects of a characteristic x to be such that

²⁵ So that a non structural econometric approach to the problem would simply correct the standard errors, for instance using Newey-West estimator, as in Filistrucchi(2005).

²⁶ Alternatively one could include age and age² of the product (and/or firm if the firm is multi-product) as regressors in the aggregate equation so as to estimate

$$\ln S_{jtm} - \ln S_{otm} = \alpha(y_{it} - p_{jtm}) + x_{jtm}\bar{\beta} + \xi_{jtm} + \rho_1 age_j + \rho_2 age_j^2 + \varepsilon_{ijtm}$$

which is a necessary condition of the following specification of indirect consumers utility:

$$u_{ijtm} = \alpha(y_{it} - p_{jtm}) + x_{jtm}\bar{\beta} + \xi_{jtm} + \rho_1 age_j + \rho_2 age_j^2 + \varepsilon_{ijtm} \text{ for each good } j$$

and

$u_{iotm} = \alpha y_{it} + \xi_{iotm} + \varepsilon_{iotm}$ for the outside good, a specification where there is no clear reason why we should think that the age of a product per se affects consumers utility, unless we claim it to be only a proxy.

$$\frac{\partial s_{jmt+T}}{\partial x_{jmt}} = \rho^T \beta s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \quad \forall T \geq 0 \quad 27 \quad (13)$$

and therefore the long-run cumulated marginal effects of a transitory change and the long-run marginal effects of a permanent change in x to be respectively such that

$$\sum_{s=0}^{\infty} \frac{\partial s_{jmt+s}}{\partial x_{jmt}} = \beta \sum_{s=0}^{\infty} \rho^s s_{jmt+s} + \sum_{s=0}^{\infty} \frac{s_{jmt+s}}{s_{0mt+s}} \frac{\partial s_{0mt+s}}{\partial x_{jmt}} \quad (14)$$

and

$$\lim_{T \rightarrow \infty} \left(\sum_{s=0}^T \frac{\partial s_{jmt+T}}{\partial x_{jmt+T-s}} \right) = \lim_{T \rightarrow \infty} \left(\beta \frac{1 - \rho^{T+1}}{1 - \rho} s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \sum_{s=0}^T \frac{\partial s_{0mt+T}}{\partial x_{jmt-s}} \right). \quad 28 \quad (15)$$

There is therefore more than one structural random utility model which satisfies the condition above.

In fact, it is true that an aggregate demand specification with a lagged dependent variable such as the one above is a necessary condition of the following specification of indirect consumers' utility:

$$u_{ijm} = \alpha(y_{it} - p_{jmt}) + x_{jmt} \bar{\beta} + \xi_{jmt} + \rho \ln(s_{jt-lm}) + \varepsilon_{ijm} \quad \text{for each good } j \quad (16)$$

and

$$u_{iotm} = \alpha y_{it} + \xi_{otm} + \rho \ln(s_{0t-lm}) + \varepsilon_{iotm} \quad \text{for the outside good} \quad (17)$$

and such a specification of consumers utility might be a way to model a (lagged) network effect but not habits.

If instead one lag of the explanatory variables is introduced, the estimated equation would then be

$$\ln s_{jmt} - \ln s_{otm} = \alpha_0(y_{it} - p_{jmt}) + \alpha_1(y_{it-1} - p_{jt-lm}) + x_{jt-lm} \bar{\beta}_1 + x_{jmt} \bar{\beta}_0 + \xi_{jmt} + \varepsilon_{ijm} \quad (18)$$

but such a specification only imposes the short and medium-run marginal effects of a characteristic x to be such that

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 s_{jmt} + \frac{s_{jmt}}{s_{0mt}} \frac{\partial s_{0mt}}{\partial x_{jmt}} \quad ,^{29} \quad (19)$$

²⁷ See appendix A.1 for a proof.

²⁸ One could be tempted to say that $\sum_{s=0}^{\infty} \frac{\partial s_{jmt+s}}{\partial x_{jmt}} = \lim_{T \rightarrow \infty} \sum_{s=0}^T \frac{\partial s_{jmt}}{\partial x_{jmt-s}} = \frac{1}{1 - \beta} \frac{\partial s_{jmt}}{\partial x_{jmt}} \forall l, j$ as in a linear

AR model. But that is not true because the logit model is non-linear and such an expression does not satisfy the stated conditions. The above would be true if market shares were constant through time.

²⁹ See appendix A.2 for a proof.

$$\frac{\partial s_{jmt+1}}{\partial x_{jmt}} = \beta_1 s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}},^{30} \quad (20)$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jmt}} = \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \quad \forall T > 1,^{31} \quad (21)$$

and therefore the long-run cumulated marginal effects of a transitory change and the long-run marginal effects of a permanent change in x to be respectively such that

$$\sum_{s=0}^{\infty} \frac{\partial s_{jdt+s}}{\partial x_{jtd}} = \beta_0 s_{jmt} + \beta_1 s_{jmt+1} + \sum_{s=0}^{\infty} \frac{s_{jmt+s}}{s_{0mt+s}} \frac{\partial s_{0mt+s}}{\partial x_{jtm}} \quad (22)$$

and

$$\lim_{T \rightarrow \infty} \left(\sum_{s=0}^T \frac{\partial s_{jmt+T}}{\partial x_{jmt+T-s}} \right) = \lim_{T \rightarrow \infty} \left(\beta_0 s_{jmt+T} + \beta_1 s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \sum_{s=0}^T \frac{\partial s_{0mt+T}}{\partial x_{jmt+T-s}} \right).^{32} \quad (23)$$

There is again more than one structural random utility model which satisfies the condition above.

In fact, such an aggregate demand specification with lagged explanatory variables is a necessary condition of the following specification of indirect consumers' utility:

$$u_{ijm} = \alpha_0 (y_{it} - p_{jmt}) + x_{jmt} \bar{\beta}_0 + \alpha_1 (y_{it-1} - p_{jmt-1}) + x_{jmt-1} \bar{\beta}_1 + \xi_{jmt} + \varepsilon_{ijm} \quad \text{for each good } j \quad (24)$$

and

$$u_{iot} = \alpha_0 y_{it} + \alpha_1 y_{it-1} + \xi_{otm} + \varepsilon_{iotm} \quad \text{for the outside good,} \quad (25)$$

a specification which is however difficult to interpret.

If finally the error term is assumed to be auto correlated of order one, the estimated equation would be

$$\ln s_{jmt} - \ln s_{0mt} = \alpha_0 (y_{it} - p_{jmt}) + \alpha_1 (y_{it-1} - p_{jmt-1}) + x_{jmt} \bar{\beta}_0 + x_{jmt-1} \bar{\beta}_1 + \rho \ln(s_{jmt-1}) + \xi_{jmt} + \varepsilon_{ijm}, \quad (26)$$

a specification which appears to derive from

$$u_{ijm} = \alpha (y_{it} - p_{jmt}) + x_{jmt} \bar{\beta} + \xi_{jmt} + \varepsilon_{ijm} \quad \text{for each good } j \quad (27)$$

and

$$u_{iotm} = \alpha y_{it} + \xi_{otm} + \varepsilon_{iotm} \quad \text{for the outside good,} \quad (28)$$

³⁰ See appendix A.3 for a proof.

³¹ See appendix A.4 for a proof.

³² These two latter effects would coincide if the market shares were constant through time.

where still $\xi_{jtm} = \zeta_j + \zeta_t + \zeta_m + \eta_{jtm}$ but now $\eta_{jtm} = \rho\eta_{jt-1m} + \varphi_{jtm}$ with φ_{jtm} i.i.d, which simply implies that the average utility from consuming a given good in a given market is correlated through time and has no direct structural interpretation.

In addition, under the restriction $\beta_1 = -\rho\beta_0$ and $\alpha_1 = -\rho\alpha_0$, this latter specification is econometrically equivalent to³³

$$\ln s_{jtm} - \ln s_{0tm} = \alpha_0(y_{it} - p_{jtm}) + \alpha_1(y_{it-1} - p_{jt-1m}) + x_{jtm}\bar{\beta}_0 + x_{jt-1m}\bar{\beta}_1 + \rho[\ln(s_{jt-1m}) - \ln(s_{0t-1m})] + \xi_{jtm} + \varepsilon_{ijtm} \quad (29)$$

which only imposes the short and medium-run marginal effects of a characteristic x to be such that

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 s_{jmt} + \frac{s_{jmt}}{s_{0mt}} \frac{\partial s_{0mt}}{\partial x_{jmt}}, \quad (30)$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jmt}} = (\beta_1 + \rho\beta_0) s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}}, \quad (31)$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jmt}} = \rho^T \beta_0 s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \quad \forall T > 1 \quad (32)$$

and therefore the long-run cumulated marginal effects of a transitory change and the long-run marginal effects of a permanent change in x to be respectively such that

$$\sum_{s=0}^{\infty} \frac{\partial s_{jmt+s}}{\partial x_{jmt}} = \beta_1 s_{jmt+1} + \sum_{s=0}^{\infty} \rho^s s_{jmt+s} + \sum_{s=0}^{\infty} \frac{s_{jmt+s}}{s_{0mt+s}} \frac{\partial s_{0mt+s}}{\partial x_{jmt}} \quad (33)$$

and

$$\lim_{t \rightarrow \infty} \left(\sum_{s=0}^T \frac{\partial s_{jmt+T}}{\partial x_{jmt+T-s}} \right) = \lim_{t \rightarrow \infty} \left(\beta_1 s_{jmt+T} + \beta_0 s_{jmt+T} \frac{1 - \rho^{T+1}}{1 - \rho} + \frac{s_{jmt+T}}{s_{0mt+T}} \sum_{s=0}^T \frac{\partial s_{0mt+T}}{\partial x_{jmt+T-s}} \right) \quad (34)$$

So that there is again more than one structural random utility model which satisfies the equation above.

Such a specification is of course a necessary condition of the following specification of indirect consumers' utility:

$$u_{ijtm} = \alpha_0(y_{it} - p_{jtm}) + x_{jtm}\bar{\beta}_0 + \alpha_1(y_{it-1} - p_{jt-1m}) + x_{jt-1m}\bar{\beta}_1 + \rho \ln(s_{jt-1m}) + \xi_{jtm} + \varepsilon_{ijtm} \quad (35)$$

and

³³ See Greene (2003), pag.581.

³⁴ See appendix A.4 for a proof.

³⁵ See appendix A.5 for a proof.

³⁶ See appendix A.6 for a proof.

³⁷ These two latter effects would coincide only if the market shares are constant through time.

$u_{iot} = \alpha_0 y_{it} + \alpha_1 y_{it} + \xi_{otm} + \rho \ln(s_{0t-1m}) + \varepsilon_{iotm}$ for the outside good, (36)
a specification which is just a combination of the first two above and thus raises the structural issues already discussed.

Finally, one should notice that in all three cases above the long-run effects are unfortunately not easy to calculate, in particular because they depend on the future market shares of the products, so that one needs to make assumptions on their behaviour extra-sample (or at least on those of their characteristics) and even so one can make forecasts only up to some finite period in the future.³⁸

4. A STATE-DEPENDENCE APPROACH TO MYOPIC HABITS

I propose here and discuss a way to include explicitly a consumer specific habit component in the utility specification at the basis of the logit model with market level data.

Let

$$u_{ijtm} = \alpha(y_{itm} - p_{jm}) + \bar{x}_{jm} \bar{\beta} + \gamma_{ijtm} + \xi_{jtm} + \varepsilon_{ijtm} \text{ for any good } j \quad (37)$$

and

$$u_{iotm} = \alpha y_{itm} + \gamma_{iotm} + \xi_{otm} + \varepsilon_{iotm} \text{ for the outside option,} \quad (38)$$

where $\gamma_{ijtm} \neq 0$ if consumer i in market m already chose product j in time $t-1$, whereas $\gamma_{ijtm} = 0$ if consumer i did not choose product j at time $t-1$ ³⁹. The underlying idea under such a specification of utility is that consumers enjoy a shift in utility if they consume again the same product they consumed in the previous period. If $\gamma_{ijtm} > 0$ consumer i is, if $\gamma_{ijtm} < 0$ he likes variety. Here heterogeneity does not enter anymore only through the additive random shock ε_{ijt} but also through γ_{ijtm} ; however, given γ_{ijtm} , it does enter through the additive random shock ε_{ijt} only.

Let then $\gamma_{ijtm} = \gamma_{jtm} \neq 0$ if consumer i in market m already chose product j in time $t-1$, that is let heterogeneity among consumers who already bought the product in that market in the previous period enter only through the additive random shock ε_{ijt} . If

³⁸ An extreme simplifying assumption could of course be that the market shares are constant in time.

³⁹ One could of course introduce state-dependence on more than one choice in the past, but as it will become clear below it would complicate much the analysis (and the data requirements).

such a shift is positive, habits on average prevail, if it is negative instead consumers on average seek variety.

Then the market share of product j at time t in market m is given by:

$$\begin{aligned}
s_{jtm} &= \text{prob}\{u_{ijtm} \geq u_{iktm} \forall_{k \neq j}\} = \\
&= \text{prob}\{u_{ijtm} \geq u_{iktm} \forall_{k \neq j} / u_{ij(t-1)m} \geq u_{ikt-1m} \forall_{k \neq j}\} \text{prob}\{u_{ij(t-1)m} \geq u_{ikt-1m} \forall_{k \neq j}\} + \\
&+ \text{prob}\{u_{ijtm} \geq u_{iktm} \forall_{k \neq j} / \exists_{k \neq j} : u_{ij(t-1)m} \leq u_{ikt-1m}\} \text{prob}\{\exists_{k \neq j} : u_{ij(t-1)m} \leq u_{ikt-1m}\} = \\
&= \text{prob}\{u_{ijtm} \geq u_{iktm} \forall_{k \neq j} / u_{ij(t-1)m} \geq u_{ikt-1m} \forall_{k \neq j}\} \text{prob}\{u_{ij(t-1)m} \geq u_{ikt-1m} \forall_{k \neq j}\} + \\
&+ \sum_{s \neq j} \text{prob}\{u_{ijtm} \geq u_{iktm} \forall_{k \neq j} / u_{ist-1m} \geq u_{ikt-1m} \forall_{k \neq j, k \neq s}\} \text{prob}\{u_{ist-1m} \geq u_{ikt-1m} \forall_{k \neq j, k \neq s}\}
\end{aligned}$$

that is

$$\begin{aligned}
s_{jtm} &= \frac{\exp(\delta_{jtm} + \gamma_{jtm})}{1 + \exp(\delta_{jtm} + \gamma_{jtm}) + \sum_{k \neq 0, k \neq j} \exp(\delta_{ktm})} s_{jt-1m} + \sum_{s \neq j, s \neq 0} \frac{\exp(\delta_{jtm})}{1 + \exp(\delta_{stm} + \gamma_{stm}) + \sum_{k \neq 0, k \neq s} \exp(\delta_{ktm})} s_{st-1m} + \\
&+ \frac{\exp(\delta_{jtm})}{\exp(\gamma_{0tm}) + \sum_{k \neq 0} \exp(\delta_{ktm})} s_{0t-1m} \text{ for any good } j
\end{aligned} \tag{39}$$

and

$$s_{0tm} = \frac{\exp(\gamma_{0tm})}{\exp(\gamma_{0tm}) + \sum_{k \neq 0, k \neq j} \exp(\delta_{ktm})} s_{0t-1m} + \sum_{s \neq 0} \frac{1}{1 + \exp(\delta_{stm} + \gamma_{stm}) + \sum_{k \neq 0, k \neq s} \exp(\delta_{ktm})} s_{st-1m} \tag{40}$$

for the outside good⁴⁰.

Therefore, dividing each good market share by the outside good market share, simplifying and taking natural logarithms leads to the following market shares estimation equation:

⁴⁰ In fact s_{jtm} and s_{0tm} are functions of only $(J+1)-1$ lagged market-shares, as the latter sum to one.

$$\begin{aligned}
 \ln s_{jtm} - \ln s_{0tm} = & \\
 & \ln \left[\frac{\exp(\psi_{jtm})}{1 + \exp(\delta_{jtm} + \gamma_{jtm}) + \exp(\delta_{stm})} s_{jt-1m} + \sum_{s \neq 0, j} \frac{1}{1 + \exp(\delta_{jtm}) + \exp(\delta_{stm} + \gamma_{stm})} s_{st-1m} + \frac{1}{\exp(\psi_{0tm}) + \exp(\delta_{jtm}) + \exp(\delta_{stm})} s_{0t-1m} \right] + \\
 & - \ln \left[\frac{1}{1 + \exp(\delta_{jtm} + \gamma_{jtm}) + \exp(\delta_{stm})} s_{st-1m} + \sum_{s \neq 0, j} \frac{1}{1 + \exp(\delta_{jtm}) + \exp(\delta_{stm} + \gamma_{stm})} s_{st-1m} + \frac{\exp(\psi_{0tm})}{\exp(\psi_{0tm}) + \exp(\delta_{jtm}) + \exp(\delta_{stm})} s_{0t-1m} \right] + \\
 & + \delta_{jtm}
 \end{aligned} \tag{41}$$

the latter showing the mistake you incur into when you estimate

$$\ln s_{jtm} - \ln s_{0tm} = \delta_{jtm} \quad (\text{where always } \delta_{jtm} = \bar{x}_{jtm} \bar{\beta} - \alpha p_{jtm} + \zeta_j + \zeta_m + \zeta_t + \eta_{jtm}) \tag{42}$$

which does not take the habit or variety seeking component into account,

but also the difference from when you just include one lag of the dependent variable

$$\ln s_{jtm} - \ln s_{0tm} = \rho (\ln s_{jt-1m} - \ln s_{0t-1m}) + \delta_{jtm} \tag{43}$$

and suggesting it is possible to distinguish habits (variety-seeking) from a direct (lagged) network effect deriving from

$$u_{ijm} = \alpha (y_{it} - p_{jtm}) + x_{jtm} \bar{\beta} + \xi_{jtm} + \rho \ln(s_{jt-1m}) + \varepsilon_{ijm} \quad \text{for each good } j \tag{44}$$

and

$$u_{i0tm} = \alpha y_{it} + \xi_{0tm} + \rho \ln(s_{0t-1m}) + \varepsilon_{i0tm} \quad \text{for the outside good.} \tag{45}$$

The equation above is however difficult to estimate. One difficulty is due to the presence of more than one random term and to the way the random terms $\eta_{stm} \forall s$ enter the equation, as they are inside each $\delta_{stm} \forall s$ and thus they are in general not additive (only the last term $\delta_{jtm} = \bar{x}_{jtm} \bar{\beta} - \alpha p_{jtm} + \zeta_j + \zeta_m + \zeta_t + \eta_{jtm}$ is additive). In addition, the model is non linear, so that for instance the terms that multiply the lags vary, in time and both across product and across markets, depending on the characteristics, the prices and the random shocks⁴¹. Last, it is clearly not possible for a fixed γ to vary across all the three dimensions j , t and m .

In order to estimate the parameters of interest in the equation above $(\alpha, \bar{\beta}, \gamma)$, let then γ be constant across products, so that $\gamma_{jtm} = \gamma_{0tm} = \gamma_{tm} \forall j = 1 \dots J$ ⁴². If we then

⁴¹ These two features make the system of equations for the market shares different from a linear VAR model, to which it might resemble at first sight.

⁴² From an economics point of view one could be interested in estimating a different γ for each product j rather than for each time period t or market m . Yet the approximation I propose below can be derived only assuming γ to be constant across products.

approximate the equation by a first order Taylor approximation around $\gamma_m = 0$ ⁴³, we get

$$\ln s_{jtm} - \ln s_{0tm} = \gamma(s_{jt-1m} - s_{0t-1m}) + \delta_{jtm} \quad (46)$$

where still

$$\delta_{jtm} = \bar{x}_{jtm} \bar{\beta} - \alpha p_{jtm} + \zeta_j + \zeta_t + \zeta_m + \eta_{jtm}.$$

Such a linear equation is then easy to estimate, although it is now impossible to distinguish habits (variety-seeking) from a direct (lagged) network effect⁴⁵.

Once the parameters are estimated it is then possible to recover estimates for the short-run and long-run elasticities.

Given the market-shares' equations for s_{jt} and s_{0t} (39 e 40), the short-run own and cross marginal effects of price on market shares are respectively:

$$\begin{aligned} \frac{\partial s_{jt}}{\partial p_{jt}} = & -\alpha \frac{\exp(\delta_{jt} + \gamma_t)}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \left(1 - \frac{\exp(\delta_{jt} + \gamma_t)}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \right) s_{jt-1} + \\ & -\alpha \sum_{s \neq j, s \neq 0} \frac{\exp(\delta_{jt})}{1 + \exp(\delta_{st} + \gamma_t) + \exp(\delta_{jt}) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \left(1 - \frac{\exp(\delta_{jt})}{1 + \exp(\delta_{st} + \gamma_t) + \exp(\delta_{jt}) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \right) s_{st-1} + \\ & -\alpha \frac{\exp(\delta_{jt})}{\exp(\gamma_t) + \exp(\delta_{jt}) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \left(1 - \frac{\exp(\delta_{jt})}{\exp(\gamma_t) + \exp(\delta_{jt}) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \right) s_{0t-1} \end{aligned} \quad (47)$$

from which the elasticity $\frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jtm}}$ can be worked out $\forall j$,

⁴³ Notice that, if the original function depends on k variables, in the k-dimensional space $\gamma_m = 0$ defines not one point, but an hyper plane of dimensions k-1. Since for $\gamma_m = 0$ the function is flat in all directions except along the γ axis, as all the other first derivatives are zero at $\gamma_m = 0$, the approximation is valid whatever the values of the other k-1 parameters.

⁴⁴ See Appendix A.7 for a proof.

⁴⁵ Unless we assume that the network effect is absent or that the network effect is absent for the outside good while the habit (variety seeking) component is present.

$$\begin{aligned}
 \frac{\partial s_{jt}}{\partial p_{lt}} = & \alpha \frac{\exp(\delta_{jt} + \gamma_t)}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \right) s_{jt-1} + \\
 & + \alpha \sum_{s \neq j, s \neq 0, s \neq 1} \frac{\exp(\delta_{jt})}{1 + \exp(\delta_{st} + \gamma_t) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{1 + \exp(\delta_{st} + \gamma_t) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \right) s_{st-1} + \\
 & + a \frac{\exp(\delta_{jt})}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt} + \gamma_t)}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \right) s_{lt-1} + \\
 & + \alpha \frac{\exp(\delta_{jt})}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \right) s_{0t-1}
 \end{aligned} \tag{48}$$

from which the elasticity $\frac{\partial s_{jt}}{\partial p_{lt}} \frac{p_{lt}}{s_{jt}}$ can be calculated $\forall j, l \neq j$

and

$$\begin{aligned}
\frac{\partial s_{0t}}{\partial p_{lt}} = & -\alpha \frac{1}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \right) s_{jt-1} + \\
& -\alpha \sum_{s \neq j, s \neq 0, s \neq 1} \frac{1}{1 + \exp(\delta_{st} + \gamma_t) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{1 + \exp(\delta_{st} + \gamma_t) + \sum_{k \neq 0, k \neq s} \exp(\delta_{kt})} \right) s_{st-1} + \\
& -a \frac{1}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt} + \gamma_t)}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \right) s_{lt-1} + \\
& -\alpha \frac{\exp(\gamma_t)}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \left(\frac{\exp(\delta_{lt})}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \right) s_{0t-1}
\end{aligned} \tag{49}$$

from which $\frac{\partial s_{0t}}{\partial p_{lt}} \frac{p_{lt}}{s_{0t}}$ can be calculated $\forall l$.

From the market-shares' equations the following derivatives are also easily calculated

$$\frac{\partial s_{jt}}{\partial s_{jt-1}} = \frac{\exp(\delta_{jt} + \gamma_t)}{1 + \exp(\delta_{jt} + \gamma_t) + \sum_{k \neq 0, k \neq j} \exp(\delta_{kt})} \quad \forall t, j \neq 0 \tag{50}$$

$$\frac{\partial s_{jt}}{\partial s_{lt-1}} = \frac{\exp(\delta_{jt})}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \quad \forall j \neq 0, l \neq 0 \tag{51}$$

$$\frac{\partial s_{jt}}{\partial s_{0t-1}} = \frac{\exp(\delta_{jt})}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \quad \forall j \neq 0 \tag{52}$$

$$\frac{\partial s_{0t}}{\partial s_{0t-1}} = \frac{\exp(\gamma_t)}{\exp(\gamma_t) + \sum_{k \neq 0} \exp(\delta_{kt})} \quad \forall t \tag{53}$$

$$\frac{\partial s_{0t}}{\partial s_{lt-1}} = \frac{1}{1 + \exp(\delta_{lt} + \gamma_t) + \sum_{k \neq 0, k \neq l} \exp(\delta_{kt})} \quad \forall l \tag{54}$$

which show that the marginal effects of market-shares at time $t-1$ on those at time t are non linear.

In addition, as each product market-share at time t is a function of all market-shares at time $t-1$, the medium-run own and cross marginal effects of price on market shares are:

$$\Delta_{(J+1^*J)T} = \left(\prod_{h=1}^T \Delta_{(J+1^*J+1)T}^{SS} \right) \Delta_{(J+1^*J)T}^{SP} \quad \forall T > 1 \quad \text{(marginal effects of a temporary change in price)} \quad (55)$$

$$\Delta_{(J+1^*J)t+T}^* = \sum_{v=0}^T \left(\prod_{h=1}^v \Delta_{(J+1^*J+1)t+h}^{SS} \right) \Delta_{(J+1^*J)t}^{SP} \quad \forall T > 1 \quad \text{(cumulated marginal effects of a temporary change in price)} \quad (56)$$

and

$$\Delta_{(J+1^*J)t+T}^{**} = \Delta_{(J+1^*J)t+T}^{SP} + \sum_{v=0}^{T-1} \left(\prod_{h=v+1}^T \Delta_{(J+1^*J+1)t+h}^{SS} \right) \Delta_{(J+1^*J)t+v}^{SP} \quad \forall T > 1 \quad \text{(marginal effects of a permanent change in price)} \quad (57)$$

where

$$\delta_{Tj,k} = \frac{\partial s_{jt+T}}{\partial p_{kt}}, \quad \delta_{Tj,k}^{SS} = \frac{\partial s_{jt+h}}{\partial s_{kt+h-1}} \quad \text{and} \quad \delta_{Tj,k}^{SP} = \frac{\partial s_{jt}}{\partial p_{kt}}$$

So that the long-run own and cross marginal effects of price on market shares are respectively:

$$\lim_{T \rightarrow \infty} \Delta_{(J+1^*J)T} = \left(\prod_{h=1}^{\infty} \Delta_{(J+1^*J+1)T}^{SS} \right) \Delta_{(J+1^*J)T}^{SP} \quad (58)$$

$$\lim_{T \rightarrow \infty} \Delta_{(J+1^*J)t+T}^* = \sum_{v=0}^{\infty} \left(\prod_{h=1}^v \Delta_{(J+1^*J+1)t+h}^{SS} \right) \Delta_{(J+1^*J)t}^{SP} \quad \forall T > 1 \quad (59)$$

and

$$\lim_{T \rightarrow \infty} \Delta_{(J+1^*J)t+T}^{**} = \lim_{T \rightarrow \infty} \left(\Delta_{(J+1^*J)t+T}^{SP} + \sum_{v=0}^{T-1} \left(\prod_{h=v+1}^{\infty} \Delta_{(J+1^*J+1)t+h}^{SS} \right) \Delta_{(J+1^*J)t+v}^{SP} \right) \quad (60)$$

all of which are unfortunately not easy to calculate, not only because the marginal effects spread in time through all products, but also because, given that all the marginal effects are non linear, they depend on the future characteristics of the product, so that one needs to make assumptions on the behaviour extra-sample of these characteristics and even so one can make forecasts only up to some finite period in the future.

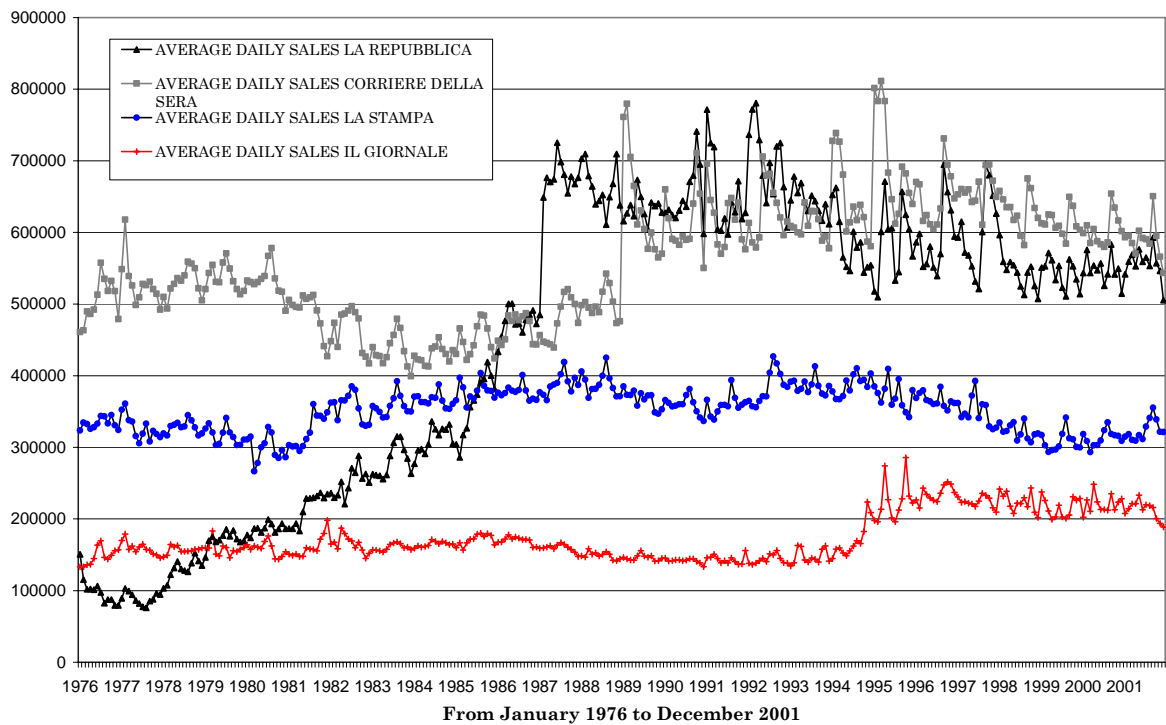
5. AN EXAMPLE

Let us consider the case of daily newspapers. It is well known in the business that readers tend to exhibit strong habits when it comes to choosing a daily newspaper. Figure 1 below shows the average daily sales in each month from 1976 to 2001 of the four main national dailies in Italy, namely *Corriere della Sera*, *La Repubblica*, *La Stampa* and *Il Giornale*⁴⁶. There is a strong persistence in the data, although many important features of daily newspapers characteristics, such as news themselves, vary through time and across products. Such a persistence might therefore be due to switching costs due to habits⁴⁷. In addition, one should note that the daily newspaper *La Repubblica* was first issued on 14 January 1976. As a result its graph had the S-shaped diffusion graph, which is not easily explained with a static structural model of demand. The persistence is clearly not removed as the data are transformed to obtain the dependent variable in the static aggregate equation normally obtained from a logit with market level data.

Figure 1 – Corriere della Sera, La Repubblica, La Stampa and Il Giornale on paper

⁴⁶ The data on sales I use in the analysis mainly draws from the data collected every year, from 1976 onwards, by the association Accertamento Diffusione Stampa (ADS), which collects them from publishers and certifies them for advertising purposes. Newspapers are free to choose whether to have their data certified or not, but if they choose so they are obliged to provide all the information required and the truthfulness of the reported information can be verified by the ADS. The latter reports average daily sales in each month and average daily prints in each month and for each different day of the week in each month (i.e. for the Fridays in June 1997). We then added to the database other useful information, mainly obtained by newspaper publishers themselves, such as the nominal prices of the newspapers, the dates of editors changes and their names, the dates the different supplements first appeared, the list of all promotions with the corresponding periods and the dates the different local chronicles were added to some of the national newspapers or the national newspaper was bundled to a local one.

⁴⁷ The data are sales at the newsstand not subscriptions, so that they are the result of a choice readers make everyday.



The equation I propose to estimate is

$$\ln(s_{jtm}) - \ln(s_{otm}) = \rho(s_{jtm} - s_{otm}) + \bar{x}_{jtm}\bar{\beta} - \alpha p_{jtm} + \zeta_j + \zeta_m + \zeta_t + \eta_{jtm} \quad (61)$$

where j is the daily newspaper, t is the month in a year, m is the day of the week⁴⁸.

Average daily market shares s_{jtd} were calculated as the number of average daily sales⁴⁹ of newspaper j at time t in day m of the week over total Italian population above 14 years of age at time t ⁵⁰. The outside good market share was calculated as

$$s_{otd} = 1 - \sum_j s_{jtd} \quad ^{51}.$$

⁴⁸ Observations on Mondays are dropped in order to have a more balanced panel. *La Repubblica* started to have a Monday issue in January 1994, *La Stampa* in January 1992 (though there was a Monday issue of *La Stampa Sera* up to December 1991) and *Il Giornale* in January 1980.

⁴⁹ The data I use for the econometric estimation are average daily sales for each different day of the week in each month obtained from the ADS data under the assumption that the ratio of sales to prints does not change across weekdays in a given month. Such an assumption allows me to enjoy a higher disaggregation of the data in order to better identify the price effect and the effect of weekly supplements, as the latter are bundled to a given newspaper only on given day of the week and, due to suspicious price coordination among editors, on that day only the price varies (is higher) compared to the price of the same newspaper in other days and of other newspapers on that day. See Argentesi(2005) & Filistrucchi(2005). I could alternatively have used data on prints, which are however include also paid subscriptions, gift subscriptions, gift copies and sales abroad. Clearly these, and in particular subscriptions (which are a very small percentage of prints) could offer additional sources of persistence other than the ones of interest here.

⁵⁰ Total market size was assumed constant across different weekdays in a month.

⁵¹ Clearly the outside good enjoys the highest market share.

Prices p_{jtm} are average real prices obtained dividing average nominal prices by the Italian monthly CPI⁵².

Characteristics x_{jtm} include dummies for supplements⁵³ (both of generalist and women's kind⁵⁴), for having a Monday issue⁵⁵, for newspaper editors as a proxy for editorial line, for having a website⁵⁶, and for games of the "lotto" kind played simply and only by buying the newspaper. Finally, newspaper-day of the week fixed effects were used as well as month and year fixed effects.

Results are reported in Table 1, together with results from the estimation of an AR1 model and a static model which uses Newey-West standard errors. All three models are estimated by OLS⁵⁷. As a result the mean prediction error is, by construction, zero in all three cases. However, if we use the parameters estimated by the approximation to calculate the mean prediction error with respect to the original non approximated equation, we get a value of approximately +0.28 (which is +5.8% of the average of the dependent variable), suggesting that the approximation on average tends to slightly over predict the market shares of the J observed products and under predict that of the outside good. The root of the mean square error is instead respectively 0.1398 (2.8%), 0.0716 (1.4%) and 0.15864 (3.2%) for the three models above and when we evaluate the fit of the approximation on the original model is +0.2980065 (6.1%).

Table 1– Estimated taste parameters

Explanatory variable	Approximation to habits model	AR1 model	Static model with Newey-West s.e.
Real price	-0.1531*10 ^{-3***} (0.0231**10 ⁻³)	-0.0848*10 ^{-3***} (0.0118*10 ⁻³)	-0.1694*10 ^{-3***} (0.0312*10 ⁻³)

⁵² Average nominal prices for each day of the week in each month were obtained by averaging over the official nominal prices of the newspaper with weights given by the number of each day of the week in the month.

⁵³ In particular, I control for the effect of the supplement on the day it is issued but also for the promotional effect on the other days of the week. See Argentesi(2004)

⁵⁴ Women's supplements are *Io Donna* for *Corriere della Sera* and *D-Donna della Repubblica* for *La Repubblica*. Generalist supplements are instead *Il Venerdì* of *La Repubblica*, *Sette* of *Corriere della Sera* and *Specchio* of *La Stampa*. The day in which they are issued has for some of them changed through time.

⁵⁵ Although observations for the Mondays were dropped, this dummy was included in order to account for a possible effect on demand in other days of the week.

⁵⁶ See Filistrucchi(2005).

⁵⁷ Nickell(1981) showed that OLS estimates a model with fixed effects and lags of the dependent variable leads to inconsistent estimates for $N \rightarrow \infty$, but given that here the number of products J is 4, of markets M is 6 and the number of periods T is 312 the relevant asymptotics should be that for $T \rightarrow \infty$.

Website	-0.1910*** (0.0163)	-0.0239*** (0.0085)	-0.2391*** (0.0343)
Games	+0.1791*** (0.00998)	+0.0352*** (0.0052)	+0.2365*** (0.0212)
Generalist supplement (day of issue)	+0.1644*** (0.0154)	+0.06743*** (0.0079)	+0.2579*** (0.0218)
Women's Supplement (day of issue)	+0.0792*** (0.0192)	+0.04980*** (0.0098)	+0.0925*** (0.0224)
Generalist supplement (all days)	+0.5259*** (0.0108)	+0.0745*** (0.0064)	+0.6764*** (0.0262)
Women's Supplement (all days)	+0.0118 (0.0189)	+0.0167* (0.0097)	+0.0151 (0.0326)
γ or ρ	+28.3371*** (0.6217)	+0.8822*** (0.0052)	n.a.
number of observations	7372	7372	7440
number of regressors	70	70	70
R ²	0.9319	0.9821	0.9126

dependent variable $\ln(s_{jd} / s_{oid})$; *** 1% **5% * 10%;

estimated coefficients for editors and Monday issues and estimated time and newspaper-day of the week fixed effects are not reported

The estimated values for γ is +28.33715, while the value for ρ in the AR1 model is +0.8821897 and both are significant at a 99% confidence level. The price parameter α is -0.0001531 for the approximated specification, -0.0000848 for the AR1 specification and -0.0001694 for the static model. In all three cases the coefficients are significantly different from 0 at a 99% confidence level. As for all the other parameters the estimates obtained by the proposed approximation are much higher than those obtained by the AR1 model and a bit lower than those obtained by the static one. Yet, since all the three models are non linear, these differences have no particular meaning unless they translate in differences in marginal effects or elasticities.

Table 2 therefore reports the average of the short-run⁵⁸ own elasticities of price, calculated from the parameters estimated in the three models above, together with their standard deviations. The estimates from the proposed model are approximately -0.25/-0.26, only slightly higher than those of -0.21/-0.22 from the static model. The

⁵⁸ The model I propose and the AR1 model both have also long-run elasticities while the static model clearly implies only one elasticity.

lowest short-run elasticities are those, around -0.10/-0.11, from the dynamic AR1 model.

Table 2– Estimated short run average own price elasticities of demand

	Approximated model	AR1 model	Static model
Corriere della Sera	-0.2528 (0.05819)	-0.1081 (0.0259)	-0.2160 (0.0517)
La Repubblica	-0.2564 (0.0531)	-0.1083 (0.0258)	-0.2163 (0.0515)
La Stampa	-0.2510 (0.0512)	-0.1049 (0.0229)	-0.2097 (0.0457)
Il Giornale	-0.2533 (0.04184)	-0.1033 (0.0169)	-0.2065 (0.0337)

standard deviation in parenthesis

6. SIMULATIONS

In order to evaluate the goodness of the proposed approximation when one wants to estimate the parameters of the original equation and to compare it an AR1 specification for aggregate demand I conducted a few simulations.

First, I drew 1000 times a value for γ from a uniform distribution on the interval (-50,+50), kept the other parameters fixed at the value estimated in the previous section, took the first observations in the sample as initial values for the market shares and used the observed value for the other regressors. For each draw of γ I generated the time series of the data starting from the observed initial values, I estimated the approximated model and the traditional ones and compared the estimated parameters with those that generated the data.

Figure 2, 3, 4 and 5 below plot the draws for γ vs. their estimate $\hat{\gamma}$ for, respectively, $\gamma > 0$, $\gamma < 0$, $0 < \gamma < -10$ and $-10 < \gamma < 0$. It appears that a positive γ is on average well estimated up to the value of +6.5, while a negative one is well estimated only down to a value⁵⁹ of -1.5.

⁵⁹ Clearly these numerical values might depend on the values of the other parameters and on the values of the regressors.

Table 3– Draws and estimates for the habit parameter γ
(draws: 1000 γ * 1 α)

	Average	Standard deviation	Min	Max
γ	-0.0512	+28.6652	-49.9826	+49.90336
$\hat{\gamma} \forall \gamma$	+19.6526	+43.2549	-1.9547	+125.9647
γ if $\gamma > 0$	+24.4596	+13.958	+0.3183	+49.90336
$\hat{\gamma}$ if $\gamma > 0$	+40.2902	+53.011	+0.3059	+125.9647
γ if $\gamma < 0$	-25.4605	+14.221	-49.9826	-0.1406
$\hat{\gamma}$ if $\gamma < 0$	-1.7417	+0.2026	-1.9548	-0.0604
γ if $0 < \gamma < 6.5$	+3.1675	+1.7594	+0.3183	+6.2954
$\hat{\gamma}$ if $0 < \gamma < 6.5$	+3.1813	+1.6805	+0.3059	+6.4055
γ if $-1.5 < \gamma < 0$	-0.8762	+0.4535	-1.4933	-0.1406
$\hat{\gamma}$ if $-1.5 < \gamma < 0$	-0.7473	+0.3486	-1.2195	-0.0604

Table 4– Percentages of estimated intervals which contain γ
(draws: 1000 γ * 1 α)

	Pg 90	Pg 95	Pg 99
$\forall \gamma$	0.0340	0.0410	0.0460
if $\gamma < 0$	0.0122	0.0122	0.0163
if $\gamma > 0$	0.0550	0.0688	0.0748
if $-1.5 < \gamma < 0$	0.4286	0.4286	0.5714
if $0 < \gamma < 6.5$	0.4909	0.6182	0.6727

Figure 2- Estimated values of γ for $\gamma > 0$

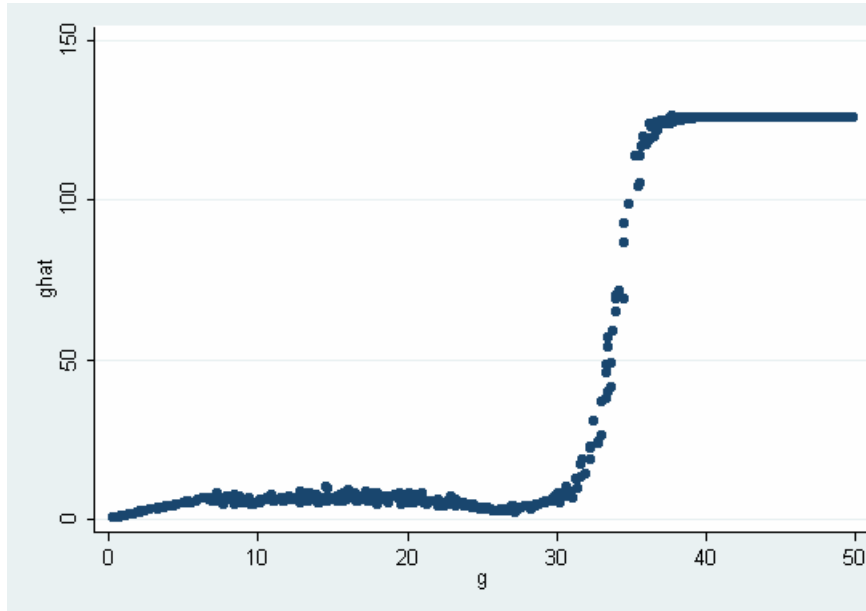


Figure 3- Estimated values of γ for $\gamma > 0$

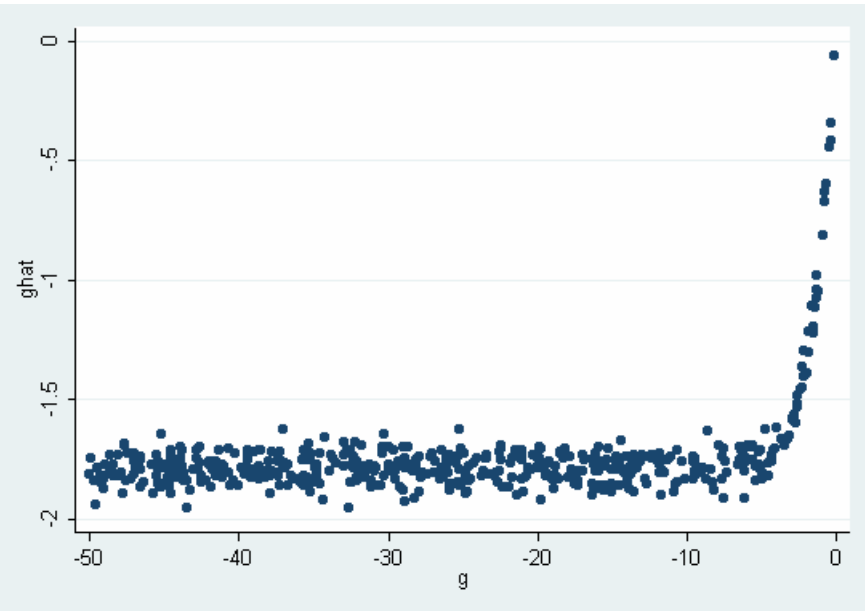


Figure 4- Estimated values of γ for $0 < \gamma < 10$

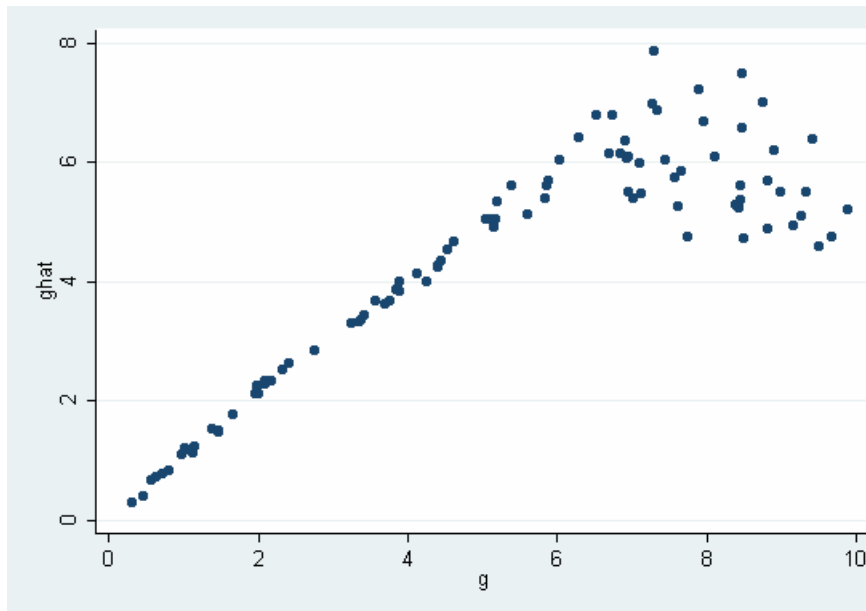


Figure 5 - Estimated values of γ for $-10 < \gamma < 0$

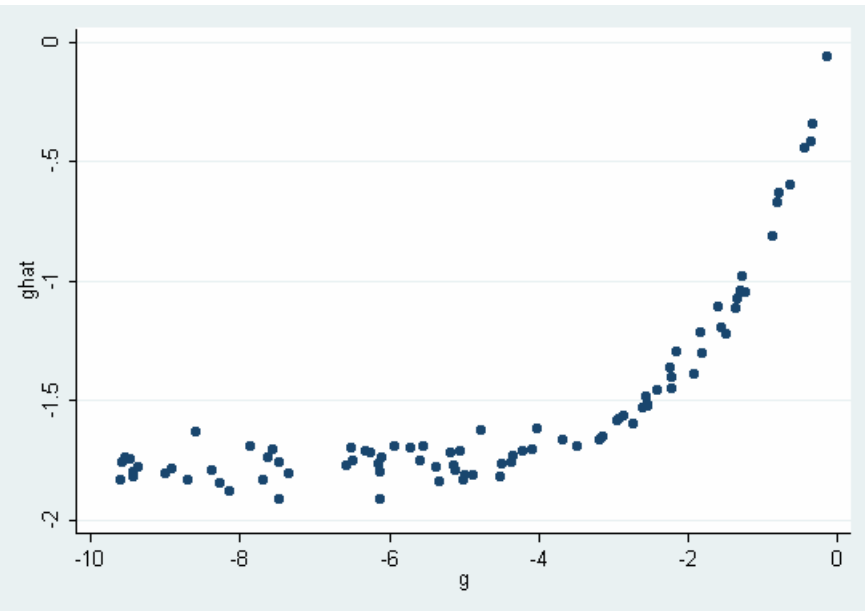


Table 3 reports instead summary statistics for the habit parameter γ and its estimate $\hat{\gamma}$ while Table 4 reports the frequency pg90 pg95 pg99 with which the estimated confidence interval contained the true value at 90%, 95% and 99% confidence levels for respectively every γ , $\gamma > 0$, $\gamma < 0$, $0 < \gamma < 6.5$ and $-1.5 < \gamma < 0$. Again the approximation seems to work quite well only around $\gamma = 0$, whereas for values of γ far away from zero, the estimates get very far from the true value (slightly farther for $\gamma < 0$ than for $\gamma > 0$).

Finally Table 5 & 6 report respectively summary statistics for the price parameter α and for the different specifications of the estimating equation its estimates $\hat{\alpha}$ and the frequency pa90 pa95 pa99 with which the estimated confidence interval contains the true value at 90%, 95% and 99% confidence levels for respectively every γ , $\gamma > 0$, $\gamma < 0$, $0 < \gamma < 6.5$ and $-1.5 < \gamma < 0$. Overall for values of γ between -50 and +50, when aiming at estimating the price parameter α the approximation seems to perform worst then the traditional AR1 model, although both of them are highly unsatisfactory, since for positive values of γ the probabilities with which the estimated confidence intervals contain the true values are far from the confidence levels (for negative value of γ they are instead much closer). As γ gets near to 0 however this latter difference closes down significantly and, what's more, the approximation performs better than the traditional model.

**Table 5- Draws and estimates for the price parameter α
(draws:1000 $\gamma * 1 \alpha$)**

	Model	Average	Standard deviation	Min	Max
α		-0.1531	0	-0.1531	-0.1531
$\hat{\alpha} \forall \gamma$	approx	-0.3971	+1.0593	-5.0326	+2.3646
	AR1	-0.0728	+0.1984	-0.8957	+0.6933
$\hat{\alpha}$ if $\gamma > 0$	approx	-0.7149	+1.4057	-5.0326	+2.3646
	AR1	+0.0103	+0.1926	-0.8957	+0.6933
$\hat{\alpha}$ if $\gamma < 0$	approx	-0.0678	+0.1600	-0.475	+0.3280
	AR1	-0.1590	+0.1648	-0.6377	+0.2682
$\hat{\alpha}$ if $0 < \gamma < 6.5$	approx	-0.1653	+0.1393	-0.5134	+0.2470
	AR1	-0.0459	+0.1459	-0.2685	+0.3584
$\hat{\alpha}$ if $-1.5 < \gamma < 0$	approx	-0.1425	+0.1548	-0.4206	+0.1497
	AR1	-0.1782	+0.1672	-0.4696	+0.1323

all values are multiplied by 1000

**Table 6 - Fraction of estimated intervals which contain α
(draws: 1000 γ * 1 α)**

	Pa 90 (approx)	Pa 95 (approx)	Pa 99 (approx)	Pa 90 (AR1)	Pa 95 (AR1)	Pa 99 (AR1)
$\forall \gamma$	0.4940	0.5400	0.5860	0.5550	0.6020	0.6640
if $\gamma < 0$	0.8289	0.9022	0.9674	0.8615	0.9226	0.9857
if $\gamma > 0$	0.1709	0.1906	0.2181	0.2593	0.2927	0.3536
if $-1.5 < \gamma < 0$	0.9286	1 ⁶⁰	1	0.8571	1	1
if $0 < \gamma < 6.5$	0.8364	0.9091	0.9636	0.7454	0.8182	0.9091

Second, I repeated the exercise above for 100 random draws of the other parameters. Each time for each parameter $v \in (\alpha, \bar{\beta}, \bar{\xi})$ a value was drawn from a uniform distribution on the interval $(0, 2\hat{v})$ where \hat{v} is the value estimated in the previous section. For each draw of the other parameter the same sequence of values was drawn for γ . As a result the obtained dataset has 10,000 observations with the same 100 values of γ for each v and the same 100 values of v for each γ .

The results of this second simulation qualitatively confirm those obtained in the first simulation. For each value of the other parameters the plots of γ against its estimates look very similar to the corresponding to Figure 2, 3, 4 and 5 with just a slight change in the length of the interval around 0 where the approximation performs satisfactorily⁶¹. Table 7 & 8 report again summary statistics for the habit parameter γ and its estimate $\hat{\gamma}$ and the frequency pg90 pg95 pg99 with which the estimated confidence interval contained the true value at 90%, 95% and 99% confidence levels for respectively every γ , $\gamma > 0$, $\gamma < 0$, $0 < \gamma < 6.5$ and $-1.5 < \gamma < 0$. Again the estimates get slightly farther away from the true value for $\gamma < 0$ than for $\gamma > 0$.

⁶⁰ There are only 14 observations in this interval.

⁶¹ In the tables below however for presentational reasons we keep the intervals $(-1.5, 0)$ and $(0, 6.5)$ as the interval were the proposed approximation performs best.

Finally Table 9 & 10, as Table 5 & 6 above, report respectively summary statistics for the price parameter α and, for each specification of the estimating equation, its estimates $\hat{\alpha}$ and the frequencies pa_{90} pa_{95} pa_{99} with which the estimated confidence intervals contain the true value at 90%, 95% and 99% confidence levels for respectively every γ , $\gamma > 0$, $\gamma < 0$, $0 < \gamma < 6.5$ and $-1.5 < \gamma < 0$. Again, when aiming at estimating the price parameter α the approximation seems to perform worst then the traditional AR1 model, although both of them are highly unsatisfactory for positive values of γ , since the probabilities with which the estimated confidence intervals contain the true values are far from the confidence levels. As we get closer to 0 however the performance of both models improves considerably and in addition the approximation performs better than the traditional model.

**Table 7– Draws and estimates for the habit parameter γ
(draws: 100 γ * 100 α)**

	Average	Standard deviation	Min	Max
γ	+1.5946	+28.0623	-49.964	+49.7412
$\hat{\gamma} \forall \gamma$	+16.4766	+39.5273	-2.1477	+125.8777
γ if $\gamma > 0$	+23.1955	+14.4514	+2.0968	+49.7412
$\hat{\gamma}$ if $\gamma > 0$	+30.8472	+48.1746	-0.1378	+125.8777
γ if $\gamma < 0$	-25.8975	+13.1980	-49.9645	-4.3398
$\hat{\gamma}$ if $\gamma < 0$	-1.81312	+0.1062	-2.1477	-1.2482
γ if $0 < \gamma < 6.5$	+3.7454	+1.2149	+2.0968	+5.6026
$\hat{\gamma}$ if $0 < \gamma < 6.5$	+2.9760	+0.7414	+1.7954	+5.4895
γ if $-1.5 < \gamma < 0$	n.a. ⁶²	n.a.	n.a.	n.a.
$\hat{\gamma}$ if $-1.5 < \gamma < 0$	n.a.	n.a.	n.a.	n.a.

**Table 8– Fraction of estimated intervals which contain γ
(draws: 100 γ * 100 α)**

	PG 90	PG 95	PG 99

⁶² In the second simulation, only 100 values of g are drawn and there are unfortunately no draws in the interval $(-1.5, 0)$.

$\forall \gamma$	0.0153	0.0179	0.0213
if $\gamma < 0$	0	0	0
if $\gamma > 0$	0.0273	0.0320	0.0380
if $-1.5 < \gamma < 0$	n.a. ⁶³	n.a.	n.a.
if $0 < \gamma < 6.5$	0.1686	0.2029	0.2386

**Table 9– Draws and estimates for the habit parameter α
(draws: 100 γ * 100 α)**

	Model	Average	Standard deviation	Min	Max
α		-0.1584	0.0881	-0.3038	-0.0007
$\hat{\alpha} \forall \gamma$ ⁶⁴	approx	-0.8245	1.6721	-8.9442	+2.5243
	AR1	-0.1754	0.2740	-1.8259	+0.7977
$\hat{\alpha}$ if $\gamma > 0$	approx	-1.3772	2.0674	-8.9442	+2.5243
	AR1	-0.1447	0.3258	-1.8259	+0.7977
$\hat{\alpha}$ if $\gamma < 0$	approx	-0.1212	0.1781	-0.6061	+0.4201
	AR1	-0.2145	0.1811	-0.662	+0.3631
$\hat{\alpha}$ if $0 < \gamma < 6.5$	approx	-0.2474	0.2078	-0.8497	+0.2549
	AR1	-0.0615	0.2032	-0.4388	+0.6279
$\hat{\alpha}$ if $-1.5 < \gamma < 0$	approx	n.a. ⁶⁵	n.a	n.a	n.a
	AR1	n.a.	n.a	n.a	n.a

all values are multiplied by 1000

**Table 10– Fraction of estimated intervals which contain α
(draws:100 γ * 100 α)**

	PA 90 (approx)	PA 95 (approx)	PA 99 (approx)	PA 90 (AR1)	PA 95 (AR1)	PA 99 (AR1)
$\forall \gamma$	0.4872	0.5217	0.5694	0.5582	0.6287	0.7026

⁶³ See note 62.

⁶⁴ $\forall \gamma$ the same series of values for α was drawn. So $\forall \gamma$ summary statistics of $\hat{\alpha}$ should be compared with those of $\alpha \forall \gamma$.

⁶⁵ See note 62.

if $\gamma < 0$	0.8898	0.9252	0.9811	0.8543	0.9482	0.9982
if $\gamma > 0$	0.1709	0.2046	0.2459	0.3255	0.3777	0.4703
if $-1.5 < \gamma < 0$	n.a ⁶⁶	n.a	n.a	n.a	n.a	n.a
if $0 < \gamma < 6.5$	0.7371429	0.8671429	0.9442857	0.53	0.6042857	0.76

7. CONCLUSION

In many markets consumers' behaviour is characterized by habits. Whereas recent empirical studies using consumer level data are able to identify the effect of habits and show the bias that ignoring it can cause in the estimates of the elasticity of demand with respect to prices, and therefore on competition policy discussions, discrete choice models of product differentiation which use aggregate level data usually treat observations relating to the same market at different points in time as observations of different markets. Although this might not be problematic in the case of durable goods when consumers in the same market at two points in time are probably different, it is clearly not realistic in the non durable goods case. The econometric issue is relevant for economic policy too as these models are being used more and more often for market power assessment in antitrust cases and usually with market level data, which are easier to recover than consumer level ones.

I showed that the traditional solutions to add lags of the dependent variable and/or lags of the explanatory variables to the aggregate demand equation are unsatisfactory when the aggregate demand equation is the outcome of a discrete choice model of product differentiation with market level data. In doing so I discussed the wider issue of persistence in consumers' choices (and therefore in products' sales), which can be due not only to habits (or more generally switching costs) but also to persistence in unobserved consumers' characteristics and/or the presence of network effects.

I then propose a model which takes (myopic) habits or variety seeking behaviour into account and estimate it on data on the market for daily newspapers in Italy. I calculate the implied short run price elasticities of demand and I compare them with those obtained by both a static and an AR1 specification for the aggregate demand equation. In order to estimate the new model an approximation is proposed, whose

⁶⁶ See note 62.

performance I compare in simulations to the dynamic AR1 specification. The simulations show that if the true model is the proposed one, the approximation is reasonably good in order to estimate the habit/variety seeking parameter and outperforms the AR1 model in recovering the price parameter, only when the habit/variety seeking parameter is not too big, that is when habits or variety seeking are not too important. In addition, in order to recover the price parameter, both models perform much better in case of variety-seeking than in case of habits and quite unsatisfactorily in the presence of strong habits.

Unfortunately, the estimated value for the habit component on the dataset on daily newspapers in Italy is quite high and falls in the range where the approximation does not perform well.

Further research on the issue could not only attempt larger sample (and larger sets of parameters' values) simulations, attempt the estimation of a second order approximation or try another way to estimate the model above, but also take an heterogeneity approach to the issue in specifying a random utility model or vice versa try to identify which consumers utility specifications are compatible with adding lags to the aggregate equation.

REFERENCES

- Argentesi E.**, Measuring Market Power in Two-sided Markets: the Case of Newspapers, EUI Working Paper n° 2004/28, 2004.
- Argentesi E. & Filistrucchi L.**, Measuring Market Power in Two-sided Markets: the Case of Newspapers, EUI Working Paper n°2005/7, 2005.
- Anderson Simon P., de Palma André & Thisse Jacques-François**, Discrete Choice Theory of Product Differentiation, Cambridge (Massachusetts)-London (UK): The MIT Press, 1992.
- Baltagi B. H.**, Econometric Analysis of Panel Data, New York: Wiley, 2001
- Berry S.**, Estimating Discrete-choice Models of Product Differentiation, *The RAND Journal of Economics*, 1994, 25(2), 242-262.
- Berry S., Levinsohn J. & Pakes A.**, Automobile Prices in Market Equilibrium, *Econometrica*, 1995, 63, 841-890.
- Brenkers Randy & Verboven Frank**, Liberalizing a Distribution System: the European Car Market , CEPR Discussion Paper n° 3622, September 2002.
- Bresnahan Timothy P.**, Empirical Studies of Industries with Market Power in Schmalensee R. & Willig R.D. Handbook of Industrial Organisation, vol.II, Elsevier, 1989.
- Che Hai, Seetharaman P.B. & Sudhir K.**, Pricing Behaviour in Markets with State Dependence in Demand, mimeo, 2004.
- Erdem Tülin**, A Dynamic Analysis of Market Structure Based on Panel Data, *Marketing Science*, 1996, 15(4), 359-378.
- Filistrucchi Lapo**, The Impact of Internet on the Market for Daily Newspapers in Italy, mimeo, EUI, 2005.
- Greene William H.**, Econometric Analysis, Prentice Hall, 2003.
- Ivaldi Marc & Verboven Frank**, Quantifying the Effects from Horizontal Mergers on European Competition Policy, *International Journal of Industrial Organization*, 2005 forthcoming.
- Kaiser U.**, The Effects of website provision on the demand for German women's magazines, Centre for Economic and Business Research Discussion Paper n° 2003-01, 2003.

- McFadden Daniel & Train Kenneth**, Mixed MNL Models for Discrete Response, mimeo, University of California at Berkeley, May 2000.
- Nevo Aviv**, A Practitioner's Guide to the Estimation of Random-Coefficients Logit Models of Demand, *Journal of Economics & Management Strategy*, 2000,9(4), 513-548.
- Nevo Aviv**, Measuring Market Power in the Ready-to-eat Cereal Industry, *Econometrica*, 69(2), March 2001, 307-342.
- Nickell S.**, Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49, 1981, 1417-1426.
- Petrin Amil**, Quantifying the Benefits of New Products: the Case of the Minivan, *Journal of Political Economy*,2002,110(4),705-729.
- Petrin Amil & Train Kenneth**, Omitted Product Attributes in Discrete Choice Models, mimeo, University of Chicago & University of California at Berkeley, December 2002.
- Roy Rishin, Chintagunta Pradeep K. and Haldar Sudeep**, A Framework for Investigating Habits, "The Hand of the Past," and Heterogeneity in Dynamic Brand Choice, *Marketing Science*, 1996, 15(3), 280-299.
- Train Kenneth**, Qualitative Choice Analysis, Cambridge (Massachusetts)-London(UK): The MIT Press,1993.
- Train Kenneth**, Discrete choice methods with simulation, Cambridge (Massachusetts); New York(NY) : Cambridge University Press, 2003.
- Vuong Quang**, Likelihood ratio tests for model selection and non-nested hypotheses, *Econometrica*, 1989, 57,307–333.

APPENDIX – MATHEMATICAL PROOFS

A.1 FROM PAGE 10, NOTE 27

Proof that

$$\frac{\partial s_{jmt+T}}{\partial x_{jmt}} = \rho^T \beta s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \quad \forall T \geq 0 \quad (13)$$

That's because

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jmt}} = \frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \rho^T \beta$$

but

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jmt}} = \frac{\partial \ln s_{jmt+T}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jmt}}$$

so that

$$\begin{aligned} \frac{\partial \ln s_{jmt+T}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jmt}} &= \rho^T \beta \\ \frac{1}{s_{jmt+T}} \frac{\partial s_{jmt+T}}{\partial x_{jmt}} - \frac{1}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} &= \rho^T \beta \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \end{aligned}$$

or equivalently

$$\begin{aligned} \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \left(0 - \frac{\partial s_{jmt+T}}{\partial x_{jmt}} - \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \right) \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} \left(1 + \frac{s_{jmt+T}}{s_{0mt+T}} \right) &= \rho^T \beta s_{jmt+T} - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} \left(\frac{s_{0mt+T} + s_{jmt+T}}{s_{0mt+T}} \right) &= \rho^T \beta s_{jmt+T} - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} + s_{jmt+T}} - \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} (s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} + s_{jmt+T}} - \frac{s_{jmt+T}}{(s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \end{aligned}$$

A.2 FROM PAGE 10, NOTE 29

Proof that

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 s_{jmt} + \frac{s_{jmt}}{s_{0mt}} \frac{\partial s_{0mt}}{\partial x_{jmt}} \quad (19)$$

That's because

$$\frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \beta_0$$

but

$$\frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \frac{\partial \ln s_{jmt}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt}}{\partial x_{jmt}}$$

so that

$$\frac{\partial \ln s_{jmt}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt}}{\partial x_{jmt}} = \beta_0$$

$$\frac{1}{s_{jmt}} \frac{\partial s_{jmt}}{\partial x_{jmt}} - \frac{1}{s_{0mt}} \frac{\partial s_{0mt}}{\partial x_{jmt}} = \beta_0$$

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 s_{jmt} + \frac{s_{jmt}}{s_{0mt}} \frac{\partial s_{0mt}}{\partial x_{jmt}}$$

or equivalently

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 s_{jmt} + \frac{s_{jmt}}{s_{0mt}} \left(0 - \frac{\partial s_{jmt}}{\partial x_{jmt}} - \frac{\sum_{s \neq j, 0} \partial s_{stm}}{\partial x_{jmt}} \right)$$

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} \left(1 + \frac{s_{jmt}}{s_{0mt}} \right) = \beta_0 s_{jmt} - \frac{s_{jmt}}{s_{0mt}} \frac{\sum_{s \neq j, 0} \partial s_{stm}}{\partial x_{jmt}}$$

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} \left(\frac{s_{0mt} + s_{jmt}}{s_{0mt}} \right) = \beta_0 s_{jmt} - \frac{s_{jmt}}{s_{0mt}} \frac{\sum_{s \neq j, 0} \partial s_{stm}}{\partial x_{jmt}}$$

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 \frac{s_{jmt} s_{0mt}}{s_{0mt} + s_{jmt}} - \frac{s_{jmt} s_{0mt}}{s_{0mt+T} (s_{0mt} + s_{jmt})} \frac{\sum_{s \neq j, 0} \partial s_{stm}}{\partial x_{jmt}}$$

$$\frac{\partial s_{jmt}}{\partial x_{jmt}} = \beta_0 \frac{s_{jmt+T} s_{0mt}}{s_{0mt} + s_{jmt}} - \frac{s_{jmt}}{(s_{0mt} + s_{jmt})} \frac{\sum_{s \neq j, 0} \partial s_{stm}}{\partial x_{jmt}}$$

A.3 FROM PAGE 11, NOTE 30

Proof that

$$\frac{\partial s_{jmt+1}}{\partial x_{jmt}} = \beta_1 s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}} \quad (20)$$

That's because

$$\frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial x_{jtm}} = \frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jtm}} = \beta_1$$

but

$$\frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial x_{jtm}} = \frac{\partial \ln s_{jmt+1}}{\partial x_{jtm}} - \frac{\partial \ln s_{0mt+1}}{\partial x_{jtm}}$$

so that

$$\frac{\partial \ln s_{jmt+1}}{\partial x_{jtm}} - \frac{\partial \ln s_{0mt+1}}{\partial x_{jtm}} = \beta_1$$

$$\frac{1}{s_{jmt+1}} \frac{\partial s_{jmt+1}}{\partial x_{jtm}} - \frac{1}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jtm}} = \beta_1$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} = \beta_1 s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jtm}}$$

or equivalently

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} = \rho \beta s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \left(0 - \frac{\partial s_{jmt+1}}{\partial x_{jtm}} - \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jtm}} \right)$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} \left(1 + \frac{s_{jmt+1}}{s_{0mt+1}} \right) = \beta_1 s_{jmt+1} - \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} \left(\frac{s_{0mt+1} + s_{jmt+1}}{s_{0mt+1}} \right) = \beta_1 s_{jmt+1} - \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} = \beta_1 \frac{s_{jmt+1} s_{0mt+1}}{s_{0mt+1} + s_{jmt+1}} - \frac{s_{jmt+1} s_{0mt+1}}{s_{0mt+1} (s_{0mt+1} + s_{jmt+1})} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+1}}{\partial x_{jtm}} = \beta_1 \frac{s_{jmt+1} s_{0mt+1}}{s_{0mt+1} + s_{jmt+1}} - \frac{s_{jmt+1}}{(s_{0mt+1} + s_{jmt+1})} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jtm}}$$

A.4 FROM PAGE 11, NOTE 31

Proof that

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} = \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jtm}} \quad \forall T > 1 \quad (21)$$

That's because

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jtm}} = \frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jtm}} = 0$$

but

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jtm}} = \frac{\partial \ln s_{jmt+T}}{\partial x_{jtm}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jtm}}$$

so that

$$\frac{\partial \ln s_{jmt+T}}{\partial x_{jtm}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jtm}} = 0$$

$$\frac{1}{s_{jmt+T}} \frac{\partial s_{jmt+T}}{\partial x_{jtm}} - \frac{1}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jtm}} = 0$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} = \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jtm}}$$

or equivalently

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} = \frac{s_{jmt+T}}{s_{0mt+T}} \left(0 - \frac{\partial s_{jmt+T}}{\partial x_{jtm}} - \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jtm}} \right)$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} \left(1 + \frac{s_{jmt+T}}{s_{0mt+T}} \right) = - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} \left(\frac{s_{0mt+T} + s_{jmt+T}}{s_{0mt+T}} \right) = - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} = - \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} (s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jtm}}$$

$$\frac{\partial s_{jmt+T}}{\partial x_{jtm}} = - \frac{s_{jmt+T}}{(s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jtm}}$$

A.5 FROM PAGE 12, NOTE 35

Proof that

$$\frac{\partial s_{jmt+1}}{\partial x_{jmt}} = (\beta_1 + \rho\beta_0)s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}} \quad (31)$$

That's because

$$\frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial x_{jmt}} = \frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \beta_1 + \rho\beta_0$$

but

$$\frac{\partial(\ln s_{jmt+1} - \ln s_{0mt+1})}{\partial x_{jmt}} = \frac{\partial \ln s_{jmt+1}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+1}}{\partial x_{jmt}}$$

so that

$$\begin{aligned} \frac{\partial \ln s_{jmt+1}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+1}}{\partial x_{jmt}} &= \beta_1 + \rho\beta_0 \\ \frac{1}{s_{jmt+1}} \frac{\partial s_{jmt+1}}{\partial x_{jmt}} - \frac{1}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}} &= \beta_1 + \rho\beta_0 \\ \frac{\partial s_{jmt+1}}{\partial x_{jmt}} &= (\beta_1 + \rho\beta_0)s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\partial s_{0mt+1}}{\partial x_{jmt}} \end{aligned}$$

or equivalently

$$\begin{aligned} \frac{\partial s_{jmt+1}}{\partial x_{jmt}} &= (\beta_1 + \rho\beta_0)s_{jmt+1} + \frac{s_{jmt+1}}{s_{0mt+1}} \left(0 - \frac{\partial s_{jmt+1}}{\partial x_{jmt}} - \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jmt}} \right) \\ \frac{\partial s_{jmt+1}}{\partial x_{jmt}} \left(1 + \frac{s_{jmt+1}}{s_{0mt+1}} \right) &= (\beta_1 + \rho\beta_0)s_{jmt+1} - \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+1}}{\partial x_{jmt}} \left(\frac{s_{0mt+1} + s_{jmt+1}}{s_{0mt+1}} \right) &= (\beta_1 + \rho\beta_0)s_{jmt+1} - \frac{s_{jmt+1}}{s_{0mt+1}} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+1}}{\partial x_{jmt}} &= (\beta_1 + \rho\beta_0) \frac{s_{jmt+1}s_{0mt+1}}{s_{0mt+1} + s_{jmt+1}} - \frac{s_{jmt+1}s_{0mt+1}}{s_{0mt+1}(s_{0mt+1} + s_{jmt+1})} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+1}}{\partial x_{jmt}} &= (\beta_1 + \rho\beta_0) \frac{s_{jmt+1}s_{0mt+1}}{s_{0mt+1} + s_{jmt+1}} - \frac{s_{jmt+1}}{(s_{0mt+1} + s_{jmt+1})} \frac{\sum_{s \neq j, 0} \partial s_{smt+1}}{\partial x_{jmt}} \end{aligned}$$

A.6 FROM PAGE 12, NOTE 36

Proof that

$$\frac{\partial s_{jmt+T}}{\partial x_{jmt}} = \rho^T \beta_0 s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \quad \forall T > 1 \quad (32)$$

That's because

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jmt}} = \frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial(\ln s_{jmt} - \ln s_{0mt})} \frac{\partial(\ln s_{jmt} - \ln s_{0mt})}{\partial x_{jmt}} = \rho^T \beta_0$$

but

$$\frac{\partial(\ln s_{jmt+T} - \ln s_{0mt+T})}{\partial x_{jmt}} = \frac{\partial \ln s_{jmt+T}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jmt}}$$

so that

$$\begin{aligned} \frac{\partial \ln s_{jmt+T}}{\partial x_{jmt}} - \frac{\partial \ln s_{0mt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 \\ \frac{1}{s_{jmt+T}} \frac{\partial s_{jmt+T}}{\partial x_{jmt}} - \frac{1}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 \frac{s_{jmt+T}}{s_{0mt+T}} + \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\partial s_{0mt+T}}{\partial x_{jmt}} \end{aligned}$$

or equivalently

$$\begin{aligned} \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 s_{jmt+T} + \frac{s_{jmt+T}}{s_{0mt+T}} \left(0 - \frac{\partial s_{jmt+T}}{\partial x_{jmt}} - \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \right) \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} \left(1 + \frac{s_{jmt+T}}{s_{0mt+T}} \right) &= \rho^T \beta_0 s_{jmt+T} - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} \left(\frac{s_{0mt+T} + s_{jmt+T}}{s_{0mt+T}} \right) &= \rho^T \beta_0 s_{jmt+T} - \frac{s_{jmt+T}}{s_{0mt+T}} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} + s_{jmt+T}} - \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} (s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \\ \frac{\partial s_{jmt+T}}{\partial x_{jmt}} &= \rho^T \beta_0 \frac{s_{jmt+T} s_{0mt+T}}{s_{0mt+T} + s_{jmt+T}} - \frac{s_{jmt+T}}{(s_{0mt+T} + s_{jmt+T})} \frac{\sum_{s \neq j, 0} \partial s_{smt+T}}{\partial x_{jmt}} \end{aligned}$$

A.7 FROM PAGE 16, NOTE 44.

Proof that the first-order Taylor series approximation of

$$\begin{aligned} \ln s_{jtm} - \ln s_{0tm} = & \\ & \ln \left[\frac{\exp(\gamma_{tm})}{1 + \exp(\delta_{jtm} + \gamma_{tm}) + \exp(\delta_{stm})} s_{jt-1m} + \sum_{s \neq 0, j} \frac{1}{1 + \exp(\delta_{jtm}) + \exp(\delta_{stm} + \gamma_{tm})} s_{st-1m} + \frac{1}{\exp(\gamma_{tm}) + \exp(\delta_{jtm}) + \exp(\delta_{stm})} s_{0t-1m} \right] + \\ & - \ln \left[\frac{1}{1 + \exp(\delta_{jtm} + \gamma_{tm}) + \exp(\delta_{stm})} s_{st-1m} + \sum_{s \neq 0, j} \frac{1}{1 + \exp(\delta_{jtm}) + \exp(\delta_{stm} + \gamma_{tm})} s_{st-1m} + \frac{\exp(\gamma_{tm})}{\exp(\gamma_{tm}) + \exp(\delta_{jtm}) + \exp(\delta_{stm})} s_{0t-1m} \right] + \\ & + \delta_{jtm} \end{aligned}$$

around $\gamma_{tm} = 0$ is

$$\ln s_{jtm} - \ln s_{0tm} = \gamma(s_{jt-1m} - s_{0t-1m}) + \delta_{jtm} \quad (46)$$

.Let us consider without loss of generality the case of two goods j and k (plus the outside good 0).

We can write

$$\ln s_{jtm} - \ln s_{0tm} = F(\delta_{jtm}, \delta_{stm}, \gamma_{tm}, s_{jt-1m}, s_{st-1m}, s_{0t-1m}) - G(\delta_{jtm}, \delta_{stm}, \gamma_{tm}, s_{jt-1m}, s_{st-1m}, s_{0t-1m}) + \delta_{jtm}$$

or equivalently

$$\ln s_{jtm} - \ln s_{0tm} = F(a) - G(a) + \delta_{jtm}$$

where

$$a = (\delta_{jtm}, \delta_{stm}, \gamma_{tm}, s_{jt-1m}, s_{st-1m}, s_{0t-1m}).$$

Its first-order Taylor series approximation around point

$$\bar{a} = (\delta_{jtm} = \bar{\delta}_{jtm}, \delta_{stm} = \bar{\delta}_{stm}, \gamma_{tm} = 0, s_{jt-1m} = \bar{s}_{jt-1m}, s_{st-1m} = \bar{s}_{st-1m}, s_{0t-1m} = \bar{s}_{0t-1m})$$

is therefore

$$\begin{aligned} \ln \bar{s}_{jtm} - \ln s_{0tm} \cong & F(\bar{a}) - G(\bar{a}) + \delta_{jtm} + \left(\frac{\partial F(a)}{\partial \gamma_{tm}} - \frac{\partial G(a)}{\partial \gamma_{tm}} \right)_{a=\bar{a}} (\gamma_{tm} - 0) + \sum_{k=j, s} \left(\frac{\partial F(a)}{\partial \delta_{ktm}} - \frac{\partial G(a)}{\partial \delta_{ktm}} \right)_{a=\bar{a}} (\delta_{ktm} - \bar{\delta}_{ktm}) + \\ & + \sum_{k=j, s, 0} \left(\frac{\partial F(a)}{\partial s_{ktm}} - \frac{\partial G(a)}{\partial s_{ktm}} \right)_{a=\bar{a}} (s_{ktm} - \bar{s}_{ktm}) \end{aligned}$$

but both $F(a)$ - $G(a)$ and all its first-order derivatives other than the one with respect to γ_{tm} are always zero at $\gamma_{tm} = 0$ (and therefore at $a = \bar{a}$), so that

$$\ln s_{jtm} - \ln s_{0tm} \cong \delta_{jtm} + \left(\frac{\partial F(a)}{\partial \gamma_{tm}} - \frac{\partial G(a)}{\partial \gamma_{tm}} \right)_{a=\bar{a}} (\gamma_{tm} - 0)$$

In addition $\forall \bar{\delta}_{jtm}, \bar{\delta}_{stm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{ot-1m}$

$$\begin{aligned}
 \frac{\partial F(a)}{\partial \gamma_{tm}} \Big|_{\gamma_{tm}=0} &= \\
 &= \left[\frac{1}{\frac{\exp(\gamma_{jtm})}{1 + \exp(\bar{\delta}_{jtm} + \gamma_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{jt-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm} + \gamma_{stm})} \bar{s}_{st-1m} + \frac{1}{\exp(\gamma_{otm}) + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{ot-1m}} \right]_{\gamma_{tm}=0}^* \\
 &= \left\{ \left(\frac{\exp(\gamma_{tm})}{1 + \exp(\bar{\delta}_{jtm} + \gamma_{tm}) + \exp(\bar{\delta}_{stm})} \Big|_{\gamma_{tm}=0} - \frac{\exp(\gamma_{tm}) \exp(\bar{\delta}_{jtm} + \gamma_{tm})}{[1 + \exp(\bar{\delta}_{jtm} + \gamma_{tm}) + \exp(\bar{\delta}_{stm})]^2} \Big|_{\gamma_{tm}=0} \right) \bar{s}_{jt-1m} + \right. \\
 &\quad \left. - \frac{\exp(\bar{\delta}_{stm} + \gamma_{tm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm} + \gamma_{tm})]^2} \Big|_{\gamma_{tm}=0} \bar{s}_{st-1m} - \frac{\exp(\gamma_{tm})}{[\exp(\gamma_{tm}) + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm} + \gamma_{tm})]^2} \Big|_{\gamma_{tm}=0} \bar{s}_{ot-1m} \right\} = \\
 &= \left[\frac{1}{\frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{jt-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{st-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{ot-1m}} \right]^* \\
 &= \left\{ \left(\frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} - \frac{\exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \right) \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{st-1m} - \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{ot-1m} \right\} = \\
 &= \left[\frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \right]^* \\
 &= \left\{ \left(\frac{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})] - \exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \right) \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{st-1m} - \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{ot-1m} \right\} = \\
 &= [1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^* \\
 &= \left\{ \left(\frac{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})] - \exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \right) \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{st-1m} - \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{ot-1m} \right\} = \\
 &= [1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^* \\
 &= \left\{ \frac{[1 + \exp(\bar{\delta}_{stm})]}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{st-1m} - \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} \bar{s}_{ot-1m} \right\} = \\
 &= [1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^* \frac{[1 + \exp(\bar{\delta}_{stm})] \bar{s}_{jt-1m} - \exp(\bar{\delta}_{stm}) \bar{s}_{st-1m} - \bar{s}_{ot-1m}}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]^2} = \\
 &= \frac{\exp(\bar{\delta}_{stm}) (\bar{s}_{jt-1m} - \bar{s}_{st-1m}) + (\bar{s}_{jt-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})}
 \end{aligned}$$

and still $\forall \bar{\delta}_{jtm}, \bar{\delta}_{stm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{0t-1m}$

$$\begin{aligned}
& \frac{\partial G(a)}{\partial \gamma_{tm}} \Big|_{\gamma_{tm}=0} = \\
& = \frac{1}{\left[\frac{1}{1 + \exp(\bar{\delta}_{jtm} + \gamma_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{jt-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm} + \gamma_{stm})} \bar{s}_{st-1m} + \frac{\exp(\gamma_{0tm})}{\exp(\gamma_{0tm}) + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{0t-1m} \right]_{\gamma_{tm}=0}} * \\
& * \left\{ \begin{aligned} & \frac{\exp(\bar{\delta}_{jtm} + \gamma_{tm})}{[1 + \exp(\bar{\delta}_{jtm} + \gamma_{tm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm} + \gamma_{tm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm} + \gamma_{tm})]_{\gamma_{tm}=0}^2} \bar{s}_{st-1m} + \\ & + \left(\frac{\exp(\gamma_{tm})}{\exp(\gamma_{tm}) + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} - \frac{\exp(\gamma_{tm}) \exp(\gamma_{tm})}{[\exp(\gamma_{tm}) + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \right) \bar{s}_{0t-1m} \end{aligned} \right\} = \\
& = \frac{1}{\left[\frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{jt-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{st-1m} + \frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} \bar{s}_{0t-1m} \right]} * \\
& * \left\{ \begin{aligned} & \frac{\exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{st-1m} + \left(\frac{1}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} - \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \right) \bar{s}_{0t-1m} \end{aligned} \right\} = \\
& = \frac{1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]} * \\
& * \left\{ \begin{aligned} & \frac{\exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{st-1m} + \left(\frac{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm}) - 1}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \right) \bar{s}_{0t-1m} \end{aligned} \right\} = \\
& = [1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})] * \\
& * \left\{ \begin{aligned} & \frac{\exp(\bar{\delta}_{jtm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{jt-1m} - \frac{\exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{st-1m} + \frac{\exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} \bar{s}_{0t-1m} \end{aligned} \right\} = \\
& = [1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})] * \frac{[\exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})] \bar{s}_{0t-1m} - \exp(\bar{\delta}_{jtm}) \bar{s}_{jt-1m} - \exp(\bar{\delta}_{stm}) \bar{s}_{st-1m}}{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})]_{\gamma_{tm}=0}^2} = \\
& = \frac{\exp(\bar{\delta}_{jtm})(\bar{s}_{jt-1m} - \bar{s}_{0t-1m}) + \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{0t-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})}
\end{aligned}$$

so that $\forall \bar{\delta}_{jtm}, \bar{\delta}_{stm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{ot-1m}$

$$\begin{aligned}
 & \frac{\partial F(\bar{\delta}_{jtm}, \bar{\delta}_{stm}, \gamma_{tm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{ot-1m})}{\partial \gamma_{tm}} \Big|_{\gamma_{tm}=0} - \frac{\partial G(\bar{\delta}_{jtm}, \bar{\delta}_{stm}, \gamma_{tm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{ot-1m})}{\partial \gamma_{tm}} \Big|_{\gamma_{tm}=0} = \\
 &= \frac{\exp(\bar{\delta}_{stm})(\bar{s}_{jt-1m} - \bar{s}_{st-1m}) + (\bar{s}_{jt-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} + \frac{\exp(\bar{\delta}_{jtm})(\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) + \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} = \\
 &= \frac{\exp(\bar{\delta}_{stm})(\bar{s}_{jt-1m} - \bar{s}_{st-1m}) + (\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) + \exp(\bar{\delta}_{jtm})(\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) + \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} = \\
 &= \frac{\exp(\bar{\delta}_{stm})(\bar{s}_{jt-1m} - \bar{s}_{ot-1m} - \bar{s}_{st-1m} + \bar{s}_{ot-1m}) + [1 + \exp(\bar{\delta}_{jtm})](\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) + \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} = \\
 &= \frac{\exp(\bar{\delta}_{stm})(\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) - \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{ot-1m}) + [1 + \exp(\bar{\delta}_{jtm})](\bar{s}_{jt-1m} - \bar{s}_{ot-1m}) + \exp(\bar{\delta}_{stm})(\bar{s}_{st-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} = \\
 &= \frac{[1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})](\bar{s}_{jt-1m} - \bar{s}_{ot-1m})}{1 + \exp(\bar{\delta}_{jtm}) + \exp(\bar{\delta}_{stm})} = \\
 &= \bar{s}_{jt-1m} - \bar{s}_{ot-1m}
 \end{aligned}$$

As this is true $\forall \bar{\delta}_{jtm}, \bar{\delta}_{stm}, \bar{s}_{jt-1m}, \bar{s}_{st-1m}, \bar{s}_{ot-1m}$, provided $\gamma_{tm} = 0$, then

$$\ln s_{jtm} - \ln s_{otm} \cong \bar{\delta}_{jtm} + \gamma_{tm}(s_{jtm} - s_{otm})$$

and if the original function depends on k variables, the approximation is valid not around one point in the k dimensional space but around a hyper-plane of dimensions $k-1$.

Finally, it is clear from the proof above that the result can be easily extended to the case of more than 2 goods.