FACULTEIT ECONOMIE
EN BEDRIJFSKUNDE

TWEEKERKENSTRAAT 2
B-9000 GENT
Tel. : 32 - (0)9 – 264.34.61
Fax. : 32 - (0)9 – 264.35.92

WORKING PAPER

Deep Habits in Consumption: A Spatial Panel Analysis Using Scanner Data

Benjamin Verhelst [1]
Dirk Van den Poel [2]

[1] SHERPPA, Ghent University
[2] Department of Marketing, Ghent University

December 2012
2012/823

D/2012/7012/56

Deep Habits in Consumption: A Spatial Panel Analysis Using Scanner Data *

Benjamin Verhelst †   Dirk Van den Poel ‡

Abstract

Using scanner data from a large European retailer, this paper empirically assesses deep habit formation in consumption. Deep habit formation constitutes a possible source of price stickiness and helps to mimic procyclical labour and real wage dynamics that are present in macro data. To gauge the existence and the extent of deep habits in consumption, we estimate a dynamic time-space simultaneous model for consumption expenditure at different levels of product aggregation. This spatial panel model enables us to test for both internal and external deep habit formation at the same time. The former captures inertia or persistence in consumption, and is included in the empirical specification as a time lag. The latter captures preference interdependence across households and is captured by a spatial lag. Our results show mixed evidence with respect to internal habit formation, whereas the external habit effect is almost always positive and significant.

JEL: C33, C38, D12, L14

Keywords: Deep habits, Preference interdependence, Spatial panel

*The authors would like to thank Freddy Heylen, Gerdie Everaert, Gert Peersman, Emmanuel Dhyne, and Sarah Lein for valuable comments. We have also benefited from comments received at the 18th International Conference on Computing in Economics and Finance (Prague, 2012) and the 6th World Conference of the Spatial Econometrics Association (Salvador de Bahia, 2012). We acknowledge support from the Belgian Program on Interuniversity Poles of Attraction, initiated by the Belgian State, Federal Office for scientific, technical and cultural affairs, contract UAP No. P 6/07. Benjamin Verhelst acknowledges financial support from the Research Foundation Flanders (FWO). Any remaining errors are ours.

†Corresponding author, SHERPPA and Department of Social Economics, Ghent University, Benjamin.Verhelst@UGent.be

‡Department of Marketing, Ghent University, Dirk.VandenPoel@UGent.be

1 Introduction

Standard models of consumer behaviour generally assume that individual consumption decisions are independent from the past spending pattern of the own household or the choice behaviour of other households. Preferences are assumed to be separable across time and across households (Alvarez-Cuadrado *et al.*, 2012). There is however a large literature documenting the importance of both internal and external habit formation in consumption choices. Internal habits are formed when the consumption choices of a household in the current period are influenced by those of the same household in previous periods, i.e. inertia/persistence in consumption choices. The reasons for internal habit formation can be very diverse, ranging from addiction or a sense of brand loyalty to switching costs or unknown quality of other products (Nakamura and Steinsson, 2011). External habits or interdependent preferences on the other hand refer to the dependence of consumption behaviour of one household on the known decisions of a certain reference group of other households, i.e. keeping up with the Joneses. People who identify with a particular group often adopt the preferences of that group, which results in interdependent choices. This can be driven by brand credibility, learning, or social concerns such as identification or compliance (Yang and Allenby, 2003).

Ravn *et al.* (2006) introduce the concept of deep habits that are formed over narrowly defined individual varieties, as opposed to the more traditional superficial habits that are formed at a more aggregate level over a composite consumption good. The introduction of internal and external deep habits in consumption implies that consumers form habits on a good-by-good basis. They derive utility not only from their current level of consumption of a certain good, but also from how this consumption level compares to their own past consumption level and that of people around them for that particular good. Through the introduction of deep habits in consumption, the optimal pricing problem of the firm becomes dynamic. They have to take into account that the price they charge today impacts future sales through the effect of current demand on future demand. If aggregate demand is high, firms will lower prices to capture excess demand, build a habit stock and hence also increase future demand. Mark-ups are hence countercyclical, whereas they are constant in the superficial habit model. Through its effects on output and labour demand, deep habits help to mimic procyclical labour and real wage dynamics that are present in macro data (Ravn *et al.*, 2006).

The theoretical demand function that arises when internal and external deep habits are introduced separately into the standard consumer demand model is very similar. Both specifications share the same mark-up, labour, and real wage dynamics. The only notable difference between the two specifications is of purely analytical nature. The firm's pricing problem ceases to be consistent under internal deep habits, rendering the consumer problem more complex (Ravn *et al.*, 2006). In the empirical analysis, we will introduce internal and external deep habit formation into the same model, captured by a time lag and a spatial lag, respectively. This will enable us to assess the strength of both types of habit, and give direction to the choice of the most appropriate theoretical deep habit formulation.

The purpose of this paper is to empirically investigate the importance of deep habit formation in consumption, using point-of-sale scanner data from six stores of an anonymous European retailer [1]. To the best of our knowledge, this paper is the first to estimate deep habit parameters using micro data for consumption at different levels of product aggregation. More precisely, we estimate consumer demand systems of the AIDS type at the product group, product category and individual product level. Appendix A gives an overview of the product aggregation structure that we use in the empirical analysis. The standard AIDS specification of Deaton and Muellbauer (1980) is extended with a time lag and a spatial lag, capturing internal and external deep habits, respectively.

We analyse the transaction data at the zip code level by aggregating across individual households. Each zip code area is therefore treated as a separate spatial unit, which amounts to a representative household framework at the zip code level. The time lag in this model specification measures the effect of the past expenditure level of a certain good or category of goods in a certain zip code area on the current expenditure on the same good or category in that area. The spatial lag then captures the effect of the expenditure share of a certain good in neighboring areas on the expenditure share of that good in the focal area. Although we are the first to estimate deep habit parameters at a multitude of product aggregation levels, the methodology and the specific definition of internal and external habits is most closely related to the study of aggregate consumption in the US by Korniotis (2010). Our model set-up also resembles the empirical study of internal and external superficial habits in Ravina (2007) and Alvarez-Cuadrado *et al.* (2012).

[1] Due to a strict confidentiality agreement, we cannot disclose the identity of the retailer.

The analysis at the zip code level consists of estimating a dynamic spatial panel data model with time and spatial fixed effects. Our preferred model is of the time-space simultaneous type, in which the expenditure share of a product or product category in one zip code area is jointly determined with its expenditure share in the neighboring areas (Anselin *et al.*, 2008). The codependence of the zip code areas is determined by the non-zero elements of a spatial weight matrix W, which serves as the spatial lag operator and has a similar role in spatial models as the shift operator $(t-1)$ in a time series model (Bradlow *et al.*, 2005). The two-directionality of the neighbour relation in space gives rise to a simultaneity issue, whereas the presence of the spatial fixed effects in our dynamic demand model leads to a dynamic panel data bias on the internal habit coefficient. We address both issues in our estimation by using the Bias-Corrected Least Squares Dummy Variables (BCLSDV) estimator of Lee and Yu (2010), which is specifically devised to estimate dynamic spatial panel data models with both time and individual fixed effects.

Our estimation results at the product group and product category level point to the existence and significance of both internal and external deep habit formation for virtually all product groups and for more than half of the product categories that we consider. At the product group level, internal and external habits are quite sizable, with an average parameter value across product groups of 0.20 and 0.31, respectively. The habit parameters are considerably smaller in size when we estimate the model at the product category level. Average parameter values across categories then amount to 0.08 and 0.12, respectively. Nonetheless, the majority of product categories exhibit positive and significant internal and external habit effects. The results of our demand analysis at the individual product level, performed for the top four products in ten selected categories, paint a slightly different picture. The external habit effect is still present for the majority of individual products that we study, with an average parameter value across products of 0.10. The internal habit on the other hand is largely absent at the individual product level, with an average of only 0.02. We can conclude from this elaborate demand analysis that deep habits are present in most of the empirical set-ups that we use. However, they are below the superficial habit level that we calculate as a benchmark, and they appear to decrease in strength when the level of product definition becomes more detailed. In most of our model set-ups, the degree of deep habit formation that we find is considerably lower than the calibrated values that have been used in the literature.

The remainder of the paper is organized as follows. Section 2 spells out the specifics of our supermarket scanner data. Section 3 covers the dynamic spatial panel analysis. Special attention is given to the specification, identification and estimation of the model, and results on internal and external habits at different levels of product aggregation are presented. We also test some alternative empirical specifications, with an emphasis on the dynamics of the habit formation process over time. Section 4 concludes.

2 Data

The data that we use are derived from the daily transactions database of six different stores of an anonymous European retailer, and run from January 2002 until November 2004. Price policy is centralized, so prices and price changes are identical across the six stores of the retailer. All prices are set at the beginning of each bi-weekly period, and remain unchanged for at least two weeks. The bar code of each individual product that leaves one of those stores is scanned at the counter and the purchase transaction is saved to the database. The retailer covers a wide range of products, accounting for approximately 40% of euro area CPI, and all of them are registered at a very detailed level through their Universal Product Code (UPC). This dataset has been used before to study price setting and customer behaviour at the micro level. Dossche *et al.* (2010) use it to test the existence of the kinked demand curve and to estimate its curvature. Verhelst and Van den Poel (2010) combine this European scanner data with the publicly available US data from Dominick's Finer Foods to stress the importance of time aggregation for measured price stickiness, and its potential impact on cross-sectional comparisons. Verhelst and Van den Poel (2012) resort to the individual customer dimension of the dataset to quantify loyalty-induced differences in estimated elasticities and curvatures, and relate this to the role of implicit contracts as a source of price rigidity.

The current paper focuses on habit formation and preference interdependence in micro price data. The European scanner dataset offers the necessary information at the individual customer level to allow for such an analysis. Through a system of compulsory loyalty cards, the purchasing pattern of each individual household can be tracked over time. On top of that, we know in which zip code area each loyalty card holder lives, and we can devise a shape file that specifies the geographical location of each area. These assets of the dataset give us the opportunity to relate the purchasing behaviour of households to their own past spending pattern and that of a geographical reference group. In other words, we can test for both internal and external habit formation. The detailed level at which products are registered in the database gives us the opportunity to test for habits at different levels of product aggregation, hence to distinguish between superficial and deep habits in consumer behaviour.

As the start of our analysis, we aggregate the daily transactions data at the UPC level across time and products. We randomly create 68 product categories by aggregating sales across all its component items. These 68 categories are in turn allocated to 7 broad product groups. Appendix A gives an overview of all the product categories that we consider, the number of items in each category, and the composition of the product groups. The ten underlined product categories in appendix A are withheld for the demand analysis at the individual product level. These categories have been selected because we consider them to be representative and relevant for deep habit formation. For each of those categories, we select the four top products based on availability on the shelf and market share inside their product category. More specifically, we first select all products that are available on the shelf at least 95% of the time, and among those products, we select the 4 items with the highest market share inside their category. The rest of the items are bundled in a fifth composite good, constructed as a weighted average of all other items in the category, to make sure that we capture all possible substitution opportunities for the selected top products. This structure is then used to test an AIDS system at the individual product level for ten selected categories. The distinction between different levels of product aggregation offers the chance to test the level at which internal and external deep habit formation is more pronounced.

With respect to time aggregation, we transform the data from daily to monthly frequency [2]. Given the available data period, the time dimension is $t = 37$. As mentioned earlier, we also aggregate the transaction data across households to the level of the zip code area, and we treat the latter as the spatial unit of analysis. The number of zip codes that are available in the shape file of our dataset is equal to 589, so the spatial dimension of the data is $N = 589$. The average size of a zip code area is approximately 500 km^2, and the average number of customers per zip code area amounts to roughly 2200.

[2]One month in fact amounts to exactly four weeks in our context. This is due to the fact that the retailer reviews prices every two weeks, so that it makes more sense to aggregate the data into stretches of four weeks rather than one month. We will ignore this timing issue in our analysis, and use the terms month and four-week period interchangeably.

3 Dynamic spatial panel analysis

We test for deep habit formation at the zip code level and define the internal and external deep habit accordingly. The former is defined as the effect of previous period consumption in the same zip code area on the current consumption level, whereas the latter captures the effect of contemporaneous consumption in a geographical reference group of other zip code areas. These definitions of internal and external habits are in line with the analysis of habit formation at the US state level of Korniotis (2010).

We extend the standard AIDS demand system of Deaton and Muellbauer (1980) with a time lag and a spatial lag to capture internal and external habits, respectively. The main analysis focuses on a single lag framework, whereas we introduce additional dynamics over time and across space in section 3.4. To ensure that each demand system at the different levels of product aggregation is a valid representation of consumer preferences, we impose the standard adding-up, homogeneity and symmetry conditions.

To keep estimation manageable, we assume weak stationarity, which implies that consumers use a multi-stage budgeting procedure in the allocation of their resources. A typical household first decides on the total expenditure level during a certain month in a certain store. Then, they allocate this amount to different product groups. Once this is done, the household decides for each single product group how to split the allotted amount across its constituent product categories, without reference to prices or consumption levels in any of the other product groups. The allotted expenditure share for each product category is in turn divided among its constituent individual products, regardless of price or consumption levels in any of the other product categories.

3.1 Empirical specification

Following the taxonomy of Anselin *et al.* (2008), we estimate a dynamic time-space simultaneous demand model for each of the product groups and categories in our sample, and for a limited number of individual products within selected product categories. The expenditure share of a random product category inside its product group in zip code area i at time t is specified in the following way [3]:

[3]The empirical specification at the product group and individual product level is a straightforward variant of equation (1), taking into account the different level of product aggregation.

$$s_{it} = \gamma s_{i,t-1} + \rho W s_{jt} + \sum_{n=1}^{m} \lambda_n \ln p_{nt} + \beta \ln \left(\frac{X_t}{P_t^*} \right) + \mu_i + \tau_t + \eta_{it} \quad (1)$$

where s_{it} and $s_{i,t-1}$ are defined as the expenditure share of the product category under consideration in zip code area i at time t and $t-1$, respectively. $W = \sum_{j=1}^{N} w_{ij}$ is an $N \times N$ spatial weight matrix with $w_{ij} = 1$ if zip code areas i and j are spatially dependent, and $w_{ij} = 0$ otherwise. s_{jt} is the expenditure share of the category in zip code area j at time t, p_{nt} is the Stone price index of category n, X_t is total expenditure inside the covering product group, and P_t^* is the Stone price index of the product group. μ_i is the spatial fixed effect that captures time-invariant characteristics of a specific zip code area, τ_t is the time fixed effect absorbing any time-specific shocks to consumption common to all spatial units, and η_{it} is an iid error term with zero mean and variance σ_η^2 [4].

The cross-sectional specific effects μ_i are treated as fixed effects in the estimation of equation (1). The fixed effects model is more appropriate than the random effects model when the cross-sectional units are fixed and not sampled (Hsiao, 2003). In our spatial setting, adjacent zip code areas serve as the cross-sectional units. Hence, inference should be conditional on the observed spatial units, and a random effects specification would be inappropriate in this context (Elhorst, 2010b).

When trying to infer the influence of social interactions in a certain geographical area on the consumption decisions of the individual households that live in that area, we should take into account the reflection problem that was raised first by Manski (1993). He distinguishes between three types of interaction effects that can explain why individual households that belong to the same group or live in the same area would behave similarly. First of all, he defines correlated effects that originate from the fact that households face the same shocks, or are subject to certain unobserved environmental characteristics. In our retailer context, these could for example come from advertising decisions of the retailer that affect all customers alike. Correlated effects should be captured in equation (1) through the common time effects τ_t. Secondly, Manski (1993) defines contextual effects

[4] We do not incorporate potential spatial dependence in the error terms, as this would create severe identification problems. However, ignoring possible spatial autocorrelation among the errors only makes the estimates of the explanatory variables less efficient, preserving their unbiasedness and consistency (Elhorst, 2010a).

when common unobserved characteristics of the group lead people to behave in a similar way. Households with a comparable social profile could for example be clustered in specific areas, and behave similarly not because they live close to each other but because they share a similar income, age, or education level. These exogenous interaction effects should be accounted for in our estimation specification by the spatial fixed effects μ_i. The third type of interaction effect, and the one that we are interested in, is the endogenous interaction effect, i.e. the true impact of social interactions among customers. The choice decisions of households directly depend on the known decisions of other people around them. We capture these endogenous interaction effects by including the spatial lag term as a regressor in the empirical specification.

The spatial lag term $\rho W s_{jt}$ captures preference interdependence and deserves some further attention. In demand models that are estimated at the macro level, the interdependence of preferences is present implicitly, but it is an omitted variable when you estimate demand at the micro level (Alessie and Kapteyn, 1991). It should therefore be taken into account explicitly in micro models, and we do so by including the spatial lag term as an explanatory variable in the empirical specification. It formalizes the role of the spatial location of a customer in his/her choice behaviour. Typically, nearby locations generate similar outcomes (Bradlow *et al.*, 2005). The role of the spatial map is similar to the role of time in time series models, as proximity on the map implies high correlation in the response variables. However, in contrast to time, space is not defined on a single dimension, and does not run in a single direction (Bronnenberg, 2005). This peculiarity of spatial processes will have severe consequences for identification and estimation of the dynamic spatial panel model.

The spatial lag operator constructs a new variable that consists of the weighted average of the neighbouring observations (Anselin *et al.*, 2008). The neighbour relation among zip code areas is operationalized through the spatial weight matrix W, which needs to be specified in advance [5]. In our analysis, we opt for geographical reference groups, implying that preference interdependence is derived from physical proximity (Bell and Song, 2004). More specifically, W takes the form of a binary contiguity matrix, i.e. $w_{ij} = 1$

[5]Since the spatial weight matrix is endogenous, estimating it from the data would lead to severe identification problems (Korniotis, 2010).

if zip codes i and j share a common border, and $w_{ij} = 0$ otherwise [6]. By construction, all diagonal elements w_{ii} are equal to 0. W is row-normalized so that the rows of the matrix sum to one, and obviously remains constant over time.

3.2 Identification and estimation

The estimation of the dynamic spatial panel model given in equation (1) is subject to two important issues. The main concern in estimating any type of spatial model is the endogeneity of the spatial lag term, due to the two-directionality of the neighbour relation in space, i.e. 'I am my neighbour's neighbour'. This contrasts with the one-directionality of time dependence (Anselin *et al.*, 2008). The value of the dependent variable for one agent is jointly determined with that of the neighbouring agents, and this simultaneity issue must be accounted for in the estimation of the spatial model.

The second concern with respect to the estimation of equation (1) is the inclusion of time and spatial fixed effects in our regression specification. We need these fixed effects to correct for potential unobserved heterogeneity across time periods and zip code areas. But it is a well-known fact that the inclusion of cross-sectional fixed effects in a dynamic model leads to an incidental parameter bias, and the standard within-group estimates are inconsistent for large N and fixed T (Nickell, 1981).

Taking these estimation issues into account, we estimate equation (1) using the Bias-Corrected Least Squares Dummy Variables (BCLSDV) approach of Lee and Yu (2010), which is specifically developed to estimate dynamic spatial panel data models with both time and spatial fixed effects. The standard LSDV estimator is a Maximum Likelihood (ML) estimator for a dynamic fixed effects panel model that includes endogenous interaction effects. It is based on the conditional log-likelihood function of the model, i.e. conditional upon the first observation in each zip code area, due to the presence of the lagged dependent variable as a regressor (Elhorst, 2012). In what follows, we will derive the LSDV estimator that accounts for the endogeneity of the spatial lag term Ws_{jt} [7].

[6] Yang and Allenby (2003) show that geographic reference groups are more important than demographic reference groups in determining individual preferences.

[7] This analysis builds on Anselin *et al.* (2008) and Elhorst (2010c), who give a detailed description of the estimation methodology for the model with only spatial fixed effects, and on Lee and Yu (2010) who extend the methodology to a model with both time and spatial fixed effects.

We start from the log-likelihood function of equation (1)

$$
\begin{aligned}
LogL &= -\frac{NT}{2} log(2\pi\sigma^2) + Tlog\,|I_N - \rho W| \\
&\quad -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(s_{it} - \gamma s_{i,t-1} - \rho W s_{jt} - x_{it}\beta - \mu_i - \tau_t)^2
\end{aligned}
\tag{2}
$$

where $x_{it}\beta$ incorporates the price and total expenditure effects of equation (1). The second term on the right-hand side represents the Jacobian term of the transformation from η to the dependent variable s taking into account the endogeneity of the spatial lag. The Jacobian can be decomposed in terms of the eigenvalues of the spatial weights matrix to reduce its dimension and facilitate estimation, thus $log\,|I_N - \rho W| = \sum_i log(1 - \omega_i)$ with ω_i as the eigenvalues of W (Anselin *et al.*, 2008). The time fixed effect τ_t can be concentrated out by taking the partial derivative of (2) with respect to τ_t:

$$
\frac{\partial LogL}{\partial \tau_t} = \frac{1}{\sigma^2}\sum_{i=1}^{N}(s_{it} - \gamma s_{i,t-1} - \rho W s_{jt} - x_{it}\beta - \mu_i - \tau_t) = 0 \tag{3}
$$

Equation (3) can be solved for τ_t:

$$
\tau_t = \frac{1}{N}\sum_{i=1}^{N}(s_{it} - \gamma s_{i,t-1} - \rho W s_{jt} - x_{it}\beta - \mu_i) \tag{4}
$$

This solution for τ_t is then substituted into the log-likelihood function (2) to obtain the concentrated log-likelihood function, where the time fixed effect is concentrated out:

$$
\begin{aligned}
LogL &= -\frac{NT}{2} log(2\pi\sigma^2) + Tlog\,|I_N - \rho W| \\
&\quad -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(\bar{s}_{it} - \gamma \bar{s}_{i,t-1} - \rho W \bar{s}_{jt} - \bar{x}_{it}\beta - \mu_i)^2
\end{aligned}
\tag{5}
$$

where variables with a bar are demeaned along the spatial dimension by

$$
\bar{s}_{it} = s_{it} - \frac{1}{N}\sum_{i=1}^{N} s_{it} \tag{6}
$$

and the same for the variables $s_{i,t-1}$, Ws_{jt} and x_{it}. Similarly to the time fixed effect, the spatial fixed effect μ_i can in turn be concentrated out by taking the partial derivative of (5) with respect to μ_i:

$$\frac{\partial LogL}{\partial \mu_i} = \frac{1}{\sigma^2}\sum_{t=1}^{T}(\bar{s}_{it} - \gamma\bar{s}_{i,t-1} - \rho W\bar{s}_{jt} - \bar{x}_{it}\beta - \mu_i) = 0 \tag{7}$$

and equation (7) can be solved for μ_i:

$$\mu_i = \frac{1}{T}\sum_{t=1}^{T}(\bar{s}_{it} - \gamma\bar{s}_{i,t-1} - \rho W\bar{s}_{jt} - \bar{x}_{it}\beta) \tag{8}$$

Substituting the solution for μ_i into the log-likelihood function (5), we obtain the concentrated log-likelihood function, where both the time and spatial fixed effect are concentrated out:

$$\begin{aligned} LogL &= -\frac{NT}{2}log(2\pi\sigma^2) + Tlog\,|I_N - \rho W| \\ &\quad -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(s^*_{it} - \gamma s^*_{i,t-1} - \rho W s^*_{jt} - x^*_{it}\beta)^2 \end{aligned} \tag{9}$$

where variables with an asterisk denote variables that are demeaned along both the time and space dimension by

$$s^*_{it} = s_{it} - \frac{1}{T}\sum_{t=1}^{T}s_{it} - \frac{1}{N}\sum_{i=1}^{N}s_{it} \tag{10}$$

and the same for the variables $s_{i,t-1}$, Ws_{jt} and x_{it}.

Anselin and Hudak (1992) spell out a two-step estimation procedure to find the maximum for the parameters of the cross-sectional model, while Elhorst (2010c) extends the procedure for spatial panels. If we stack the observations as successive cross-sections for $t = 1, ..., T$ to obtain $NT \times 1$ vectors for Y^*, $(I_T \otimes W)Y^*$ and Y^*_{-1}, where $\otimes$ denotes the

Kronecker product, and an $NT \times K$ matrix for X^*, then the estimates for γ, β and σ^2 can be expressed as a function of the sole unknown parameter ρ. A unique, numerical solution for ρ is obtained by maximizing the concentrated log-likelihood function

$$LogL = C - \frac{NT}{2} log \left[(e_0^* - \rho e_1^*)'(e_0^* - \rho e_1^*) \right] + Tlog \left| I_N - \rho W \right| \tag{11}$$

where C is a constant not depending on ρ, and e_0^* and e_1^* are the residuals of successively regressing Y^* and $(I_T \otimes W)Y^*$ on $\tilde{X}^*$, where $\tilde{X}^* = [Y_{-1}^* X^*]$. Using the numerical estimate for ρ, the other parameters γ, β and σ^2 are derived as follows:

$$\left[\hat{\gamma}\hat{\beta} \right]' = \left(\tilde{X}^{*\prime} \tilde{X}^* \right)^{-1} \tilde{X}^{*\prime} \left[Y^* - \rho (I_T \otimes W) Y^* \right] \tag{12}$$

$$\hat{\sigma}^2 = \frac{1}{NT} \left(Y^* - \rho(I_T \otimes W)Y^* - \tilde{X}^* \left[\hat{\gamma}\hat{\beta}' \right]' \right)' \left(Y^* - \rho(I_T \otimes W)Y^* - \tilde{X}^* \left[\hat{\gamma}\hat{\beta}' \right]' \right) \tag{13}$$

Due to the incidental parameter bias, the standard LSDV estimator is inconsistent in a fixed effect dynamic spatial panel model. More specifically, the coefficient of the lagged dependent variable will suffer from a negative dynamic panel data bias. Lee and Yu (2010) devise an analytical bias correction procedure for the LSDV estimator that corrects for the incidental parameter bias. The bias corrected estimates are constructed by first estimating the asymptotic bias and then subtracting it from the parameter estimates of the uncorrected approach. Lee and Yu (2010) provide analytical proof that when $N/T^3 \to 0$ and $T/N^3 \to 0$, the bias corrected estimates are $\sqrt{NT}$ consistent and asymptotically centered normal. They support this with a Monte Carlo experiment showing that their bias correction approach succeeds in strongly reducing the bias for various values of N and T.

Elhorst (2010b) investigates the small sample properties of this BCLSDV estimator in terms of bias and root mean squared error (RMSE) using Monte Carlo experiments for small values of T and finds that the small sample bias in the response parameter of the endogenous interaction effect is rather small, in contrast to the Arellano and Bond (1991) system GMM estimator extended to include endogenous interaction effects. The BCLSDV estimator also greatly reduces the negative bias in the coefficient of the lagged dependent variable, i.e. the Nickell bias, compared to the uncorrected LSDV estimator.

3.3 Results

We estimate model (1) at three different levels of product aggregation to check at which level internal and external habit formation are most pronounced. As a benchmark, we also estimate superficial habit formation in the aggregate scanner data by estimating a simple consumption specification:

$$s_{it} = \gamma s_{i,t-1} + \rho W s_{jt} + \mu_i + \tau_t + \eta_{it} \tag{14}$$

where s_{it} and $s_{i,t-1}$ are defined as the log sales at the retailer in zip code area i at period t and $t-1$, respectively, s_{jt} is period t consumption in zip code area j, and all other variables are as defined before [8].

Table 1 presents the estimation results of the dynamic, time-space simultaneous demand model specified in equation (1), estimated using the BCLSDV approach of Lee and Yu (2010). The table shows the internal habit parameter γ and the external habit parameter ρ at the different levels of product aggregation. At the product group, product category and individual product level, we present average parameters across the different groups, categories and products, respectively. The habit parameters at the aggregate level in the first column are based on the estimation of equation (14) and are given as a benchmark to check the strength of deep versus superficial habits. Standard errors for the average parameters at the group, category and product level are given between parentheses below the respective parameter values. They are calculated by dividing the standard deviation across the individual parameters by the square root of the number of sampled groups, categories and products, respectively.

The results show that the size of both the internal and external habit parameter is increasing with the level of product aggregation. Habit formation and preference interdependence are more pronounced at the more aggregate levels of product classification. All the habit parameters are significant at the 5% level, although the internal habit parameter at the individual product level is economically insignificant. To get more insight

[8] In comparing the external deep habit parameters with their superficial benchmark, we should take into account the different definition of s across specifications. Whereas ρ in the aggregate model can be interpreted as an elasticity, this is only the case for the disaggregate models if the time-averaged value of s is identical across contiguous zip code areas.

into the habit formation process, the left panel of the table in Appendix B presents the internal and external habit coefficients for all the product groups and categories that are behind the average parameters in table 1. All the coefficients at the group level are positive and highly significant, except for the external habit in the leisure and eduction group. At the product category level, the internal deep habit parameter is positive and significantly different from zero at the 5% significance level for 48 out of 68 categories, whereas external deep habits are positive and significant for 34 categories.

Table 1: Average habit parameters (time-space simultaneous model, 1-month time lag)

	Aggregate	Group	Category	Product
γ	0.269***	0.200***	0.116***	0.018**
	(0.023)	(0.041)	(0.016)	(0.007)
ρ	0.437***	0.306***	0.082***	0.102***
	(0.028)	(0.074)	(0.015)	(0.014)

Note: The parameters at the group, category and product level are obtained by taking the average over all individual habit parameters of the product groups, product categories and individual products, respectively. Standard errors are given between parentheses; * p<0.10, ** p<0.05, *** p<0.01.

The left panel of the table in Appendix C gives the coefficients of the individual product regressions for the four top products in ten selected categories. The evidence on internal habits is mixed, with only 11 out of 40 products displaying a positive and significant parameter, whereas all other coefficients are statistically insignificant at the 5% level. The evidence on preference interdependence at the individual product level is more convincing, with 25 out of 40 products displaying positive and statistically significant external habit formation. Nonetheless, the value of the external habit parameters is relatively small in economic terms. The main message from this analysis is that external deep habits are statistically significant for the majority of groups, categories and individual products that we consider. Internal deep habits are present at the product group and category level, whereas at the individual product level they are very weak at best.

To assess the economic significance of our estimated deep habit parameters, we compare our empirical results with the parameters as calibrated in the literature. In most of our model set-ups, the degree of deep habit formation that we find in our data is considerably lower than the calibrated values that have been used in recent macroeconomic models. In their seminal study on deep habit formation, Ravn *et al.* (2006) calibrate a deep habit parameter of 0.86 and apply it to a standard consumption model. We find no evidence of such a strong habit formation mechanism in our scanner data. Our results at the product group level do however lend some support to the calibration value of 0.52 that Ravn *et al.* (2012) use in their two-country model to explain the effects of government spending shocks, and the value of 0.50 that Jacob (2012) imposes in his analysis of the consumption response to government spending. However, if the product category or the individual product level are deemed to be the most relevant aggregation levels for deep habit formation, then even these calibrated values are out of sync with the empirical evidence that we present.

Although our empirical evidence points to a more limited extent of deep habit formation, we have to be cautious in comparing our estimates with calibrated values from macro models. First of all, these models study consumption as a whole, whereas our estimates are based on a selection of retail products. Our empirical results clearly show a reasonable amount of cross-product and cross-category heterogeneity with respect to the strength of habit formation. Consequently, there might be some disconnect between the deep habit parameters that we find in our dataset and the ones we would obtain if we were able to study a more complete consumption basket. Secondly, we should be careful in comparing habit formation parameters across different demand specifications. More specifically, our external deep habit parameter is defined as a contemporaneous effect, whereas in the calibration of Ravn *et al.* (2006), Ravn *et al.* (2012) and Jacob (2012), this parameter is lagged one period in time. In the robustness section, we show that the contemporaneous effect dominates the lagged effect in all our model set-ups, lending support for the time-space simultaneous specification of our preferred model. This result implies that the speed of adjustment of the external habit stock is very high. This may in turn be due to the nature of the products that we study, as the habit formation process for a typical retail product is probably more prone to short-term habits than for example services, cars or real estate.

3.4 Robustness

Although most empirical analyses of habit formation in consumer behaviour include only one time lag in the model specification, there may be more persistence in the consumption decisions of households than can be captured by a single time lag. There can be a fundamental divergence between the time at which an individual buys a good and the time of the actual consumption of that good. Especially for non-perishable products, people generally buy in larger quantities and stockpile. The actual consumption is then spread out over a longer period of time. Since we use sales data with a relatively high frequency, this type of behaviour can potentially have a pronounced influence on our estimated internal habit parameters.

To address this issue, we incorporate additional time dynamics into the empirical specification of our model. More specifically, we replace the one-period time lag of the expenditure share $s_{i,t-1}$ by the average expenditure share over the previous six months:

$$s_{it} = \gamma \bar{s}_{i,t-r} + \rho W s_{jt} + \sum_{n=1}^{m} \lambda_n \ln p_{nt} + \beta \ln \left(\frac{X_t}{P_t^*} \right) + \mu_i + \tau_t + \eta_{it} \tag{15}$$

where $\bar{s}_{i,t-r} = \frac{1}{6} \sum_{r=1}^{6} s_{i,t-r}$ and all other variables are as defined before. Table 2 presents the internal and external habit parameters γ and ρ from estimating equation (15) at the product group, product category and individual product level. Standard errors are again provided between parentheses below the average parameter values. The superficial habit parameters are given as a benchmark and are based on the estimation of equation (14), in which the one-period time lag is replaced by the six-month average time lag:

$$s_{it} = \gamma \bar{s}_{i,t-r} + \rho W s_{jt} + \mu_i + \tau_t + \eta_{it} \tag{16}$$

The results in table 2 show that the external habit parameter ρ remains largely unaffected by the transformation of the time lag specification, which is of course not surprising. The internal habit parameter is not affected in the aggregate case, but there are some notable changes in the average parameter values at the group, category and product level. At the product group level, γ increases from 0.200 to 0.325. At the product category and individual product level on the other hand, γ decreases quite sizably. Hence, the internal habit parameter is lower when current expenditure is compared to the average in the previous six months than if we compare it with the expenditure level in the previous month. At the category level, the average parameter becomes insignificant, whereas at the individual product level, it is now significantly negative.

Table 2: Average habit parameters (time-space simultaneous model, 6-month time lag)

	Aggregate	Group	Category	Product
γ	0.265***	0.325***	0.010	-0.086***
	(0.041)	(0.062)	(0.022)	(0.017)
ρ	0.481***	0.296***	0.091***	0.102***
	(0.029)	(0.080)	(0.016)	(0.015)

Note: The parameters at the group, category and product level are obtained by taking the average over all individual habit parameters of the product groups, product categories and individual products, respectively. Standard errors are given between parentheses; * p<0.10, ** p<0.05, *** p<0.01.

The right panel of the table in Appendix B gives a detailed breakdown of the parameter values across the different product groups and categories. Both the internal and external habit coefficients are positive and significant for all product groups, except for leisure and education. At the product category level, the external habit parameter ρ is positive and significant for 33 out of 68 categories, in line with the results of section 3.3. However, the internal habit parameter γ is positive and significant at the 5% level for only 16 out of 68 categories, whereas it is significantly negative for 11 other categories. This translates into an insignificant average parameter at the product category level. The right panel of the table in Appendix C provides similar output for the individual products of our ten selected categories. The evidence on external habit formation is in line with the results from section 3.3, but the internal habit effect now becomes significantly negative for 15 out of 40 products, whereas not a single individual product displays significantly positive internal habit formation. The picture that emerges from this analysis is that products that are very popular at some point in time, do not succeed in holding on to their market share for more than a couple of months. This is especially true for products in the equipment and clothing groups, where new product introductions take place at a relatively high frequency.

To get more insight into the time dynamics of the internal habit, we estimate a model with up to six individual time lags:

$$\begin{aligned} s_{it} &= \gamma_1 s_{i,t-1} + \gamma_2 s_{i,t-2} + \gamma_3 s_{i,t-3} + \gamma_4 s_{i,t-4} + \gamma_5 s_{i,t-5} + \gamma_6 s_{i,t-6} \\ &\quad + \rho W s_{jt} + \sum_{n=1}^{m} \lambda_n \ln p_{nt} + \beta \ln\left(\frac{X_t}{P_t^*}\right) + \mu_i + \tau_t + \eta_{it} \end{aligned} \tag{17}$$

The results in table 3 strengthen the conclusion of the previous analysis based on the six-month average time lag that positive time dependence is relatively short-lived at the product category and the individual product level. Whereas the coefficients of the lagged expenditure share remain positive and significant at the product group level for up to five months, they turn negative and significant at the product category and individual product level after two to three months already.

Table 3: Average habit parameters (time-space simultaneous model, 6 time lags)

	Aggregate	Group	Category	Product
ρ	0.456***	0.279***	0.085***	0.097***
	(0.029)	(0.072)	(0.015)	(0.016)
γ_1	0.229***	0.128***	0.111***	0.017
	(0.025)	(0.031)	(0.016)	(0.011)
γ_2	0.089***	0.070***	-0.007	-0.018**
	(0.026)	(0.024)	(0.009)	(0.008)
γ_3	0.039	0.040**	-0.017**	-0.027***
	(0.026)	(0.018)	(0.007)	(0.007)
γ_4	0.035	0.021**	-0.015**	-0.028***
	(0.026)	(0.011)	(0.007)	(0.007)
γ_5	-0.058**	0.076***	-0.011	-0.019***
	(0.026)	(0.021)	(0.007)	(0.007)
γ_6	-0.046*	0.011	-0.027***	-0.019***
	(0.025)	(0.012)	(0.007)	(0.007)

Note: The parameters at the group, category and product level are obtained by taking the average over all individual habit parameters of the product groups, product categories and individual products, respectively. Standard errors are given between parentheses; * p<0.10, ** p<0.05, *** p<0.01.

A final robustness check with respect to the dynamic specification of the model concerns the time dimension of the external habit effect. Again following the taxonomy of Anselin *et al.* (2008), we estimate a time-space dynamic model that takes the following form for a random product category:

$$s_{it} = \gamma s_{i,t-1} + \rho W s_{jt} + \delta W s_{j,t-1} + \sum_{n=1}^{m} \lambda_n \ln p_{nt} + \beta \ln\left(\frac{X_t}{P_t^*}\right) + \mu_i + \tau_t + \eta_{it} \qquad (18)$$

where $s_{j,t-1}$ is the expenditure share of the category inside its product group in zip code area j at time $t-1$ and all other variables are as defined before. Compared to the time-space contemporaneous model, equation (18) includes a time-space lag as an additional regressor. The consumption behaviour of a typical household can potentially be influenced by previous period behaviour of other households, e.g. because it takes some time to witness their choices and act upon them.

The results in table 4 show that the internal and external habit parameters γ and ρ remain largely unaffected by the inclusion of the time-space lag. The latter itself is positive and significant at the 5% level in the aggregate, but it is statistically and economically insignificant at the group, category and product level. This confirms that internal and external habits in retailing are short-lived, and most of the effects die out relatively quickly.

Table 4: Average habit parameters (time-space dynamic model, 1-month time lag)

	Aggregate	Group	Category	Product
γ	0.266***	0.208***	0.123***	0.021**
	(0.026)	(0.042)	(0.018)	(0.010)
ρ	0.453***	0.332***	0.101***	0.128***
	(0.028)	(0.082)	(0.021)	(0.019)
δ	0.105**	0.046	0.059*	0.006
	(0.041)	(0.077)	(0.033)	(0.026)

Note: The parameters at the group, category and product level are obtained by taking the average over all individual habit parameters of the product groups, product categories and individual products, respectively. Standard errors are given between parentheses; * p<0.10, ** p<0.05, *** p<0.01.

4 Conclusion

Habit formation and preference interdependence are potentially important drivers in the consumption decisions that households make every day. Consumers can derive utility from aligning their current consumption choices with their own expenditure pattern in the previous period, and with the observed spending behaviour of a reference group. Consumption decisions are then interdependent across both time and space. This paper makes an empirical contribution to the consumption literature by estimating internal and external habit formation at different levels of product aggregation, using detailed scanner data of a large European retailer. We therefore test for different forms of deep habit formation, and compare its strength with the benchmark superficial habits that are more prevalent in existing literature. Deep habits render the firm's pricing problem dynamic, leading to a countercyclical mark-up, and help to mimic procyclical labour and real wage dynamics that are present in macro data.

We test for internal and external deep habits at the zip code level by estimating a dynamic time-space simultaneous model of expenditure that includes both a time and a spatial lag as regressors. The former captures inertia in consumption decisions, whereas the latter measures the impact of social interactions or neighbourhood effects on observed consumption choices. The reference group is formed on the basis of physical proximity, with a spatial weight matrix based on contiguity of zip code areas. We estimate our model at the different levels of product aggregation using the Bias-Corrected Least Squares Dummy Variables (BCLSDV) approach of Lee and Yu (2010), taking into account the simultaneity of the spatial lag and the dynamic panel data bias.

Our results provide evidence for the existence and significance of both internal and external deep habit formation for all product groups and more than half of the product categories that we consider. At the individual product level, external habits remain positive and significant, whereas internal habits are largely absent. Comparing the different aggregation levels, habit formation and preference interdependence become weaker at more disaggregate levels of product definition. The average habit parameters that we find at the product category and individual product level are well below the calibrated deep habit parameters that have been used in the literature.

References

Alessie, R. and Kapteyn, A. (1991). Habit Formation, Interdependent Preferences and Demographic Effects in the Almost Ideal Demand System. *The Economic Journal*, **101** (406), 404–419.

Alvarez-Cuadrado, F., Casado, J. M., Labeaga, J. M. and Sutthiphisal, D. (2012). Envy and Habits: Panel Data Estimates of Interdependent Preferences. *Banco de Espana Working Paper*, **No. 1213**, 48p.

Anselin, L. and Hudak, S. (1992). Spatial Econometrics in Practice: A Review of Software Options. *Regional Science and Urban Economics*, **22** (3), 509–536.

—, Le Gallo, J. and Jayet, H. (2008). Spatial Panel Econometrics. In L. Mátyás and P. Sevestre (eds.), *The Econometrics of Panel Data*, Springer-Verlag Berlin Heidelberg, pp. 625–660.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, **58** (2), 277–297.

Bell, D. and Song, S. (2004). Social Contagion and Trial on the Internet: Evidence from Online Grocery Retailing. *Wharton School Marketing Department Paper*, p. 45p.

Bradlow, E., Bronnenberg, B., Russell, G., Arora, N., Bell, D., Duvvuri, S., Ter Hofstede, F., Sismeiro, C., Thomadsen, R. and Yang, S. (2005). Spatial Models in Marketing. *Marketing Letters*, **16** (3/4), 267–278.

Bronnenberg, B. (2005). Spatial models in marketing research and practice. *Applied Stochastic Models in Business and Industry*, **21** (4/5), 335–343.

Deaton, A. and Muellbauer, J. (1980). An Almost Ideal Demand System. *American Economic Review*, **70** (3), 312–326.

Dossche, M., Heylen, F. and Van den Poel, D. (2010). The Kinked Demand Curve and Price Rigidity: Evidence from Scanner Data. *Scandinavian Journal of Economics*, **112** (4), 723–752.

Elhorst, P. (2010a). Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, **5** (1), 9–28.

— (2010b). Dynamic Panels with Endogenous Interaction Effects when T is Small. *Regional Science and Urban Economics*, **40** (5), 272–282.

— (2010c). Spatial Panel Data Models. In M. Fisher and A. Getis (eds.), *Handbook of Applied Spatial Analysis*, Springer-Verlag Berlin Heidelberg, pp. 377–407.

— (2012). Dynamic Spatial Panels: Models, Methods, and Inferences. *Journal of Geographical Systems*, **14** (1), 5–28.

Hsiao, C. (2003). Analysis of Panel Data, Second Edition. *Cambridge University Press, Cambridge*, 369p.

Jacob, P. (2012). Deep Habits, Price Rigidities and the Consumption Response to Government Spending. *Journal of Money, Credit and Banking, forthcoming.*

Korniotis, G. (2010). Estimating Panel Models With Internal and External Habit Formation. *Journal of Business & Economic Statistics*, **28** (1), 145–158.

Lee, L. and Yu, J. (2010). A Spatial Dynamic Panel Data Model with both Time and Individual Fixed Effects. *Economic Theory*, **26** (2), 564–597.

Manski, C. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, **60** (3), 531–542.

Nakamura, E. and Steinsson, J. (2011). Price Setting in Forward-Looking Customer Markets. *Journal of Monetary Economics*, **58** (3), 220–233.

Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica*, **49** (6), 1417–1426.

Ravina, E. (2007). Habit Persistence and Keeping up with the Joneses: Evidence from Micro Data. *Mimeo*, 58p.

Ravn, M., Schmitt-Grohé, S. and Uribe, M. (2006). Deep Habits. *Review of Economic Studies*, **73** (1), 195–218.

—, — and Uribe, M. (2012). Explaining the Effects of Government Spending Shocks. *Journal of Monetary Economics*, **59** (3), 215–234.

Verhelst, B. and Van den Poel, D. (2010). Price Rigidity in Europe and the US: A Comparative Analysis Using Scanner Data. *UGent Working Paper*, **No. 10/684**, 32p.

— and Van den Poel, D. (2012). Implicit Contracts and Price Stickiness: Evidence from Customer-Level Scanner Data. *UGent Working Paper*, **No. 12/776**, 26p.

Yang, S. and Allenby, G. (2003). Modeling Interdependent Consumer Preferences. *Journal of Marketing Research*, **40** (3), 282–294.

Appendix A: Product classification

The list below gives an overview of the 7 product groups (in bold), the 68 product categories classified by product group, and the number of items in each product category (between brackets). The 10 product categories that are selected for the analysis at the individual product level are underlined:

Food: potatoes (26), baking flour (18), chips (138), cornflakes (49), emmental cheese (56), smoked salmon (18), gruyere cheese (19), liver pie (98), biscuits (9), margarine (62), mayonnaise (45), minarine (2), spaghetti (30), ice cream (130), spinach (29), sugar (19), tomato soup (5), tuna (46)
Beverages: champagne (110), coke (39), jenever (43), lemonade (33), mineral water (66), port wine (54), tea (67), vermouth (11), fruit juice (54), whiskey (82), wine (17), beer (6), chocolate milk (9)
Equipment: casserole (74), digital camera (178), airing cupboard (61), dvd-recorder (20), dvd-player (121), mixing tap (25), microwave oven (108), measuring tape (15), hedge shears (32), knife (19), vacuum cleaner (115), toaster (40), washing machine (36)
Personal care: deodorant (238), shower gel (175), hairspray (7), nail polish (15), plasters (33), toothpaste (175), toilet paper (13), nappies (64), handkerchief (63)
Leisure and education: pen (123), home trainer (52), school book (34), cartoon (86), football (32), dictionary (32)
Clothing: bathing suit (522), jeans (79), jacket (88), socks (271), T-shirt (438)
Cleaning products: floorcloth (11), toilet soap (34), dishwashing detergent (43), soap powder (98)

Appendix B: Detailed results group and category level

	One-month time lag				Six-month average time lag			
	γ		ρ		γ		ρ	
Aggregate	0.269	***	0.437	***	0.265	***	0.481	***
	(0.028)		(0.023)		(0.041)		(0.029)	
Group								
Food	0.084	***	0.313	***	0.189	***	0.269	***
	(0.019)		(0.026)		(0.035)		(0.030)	
Beverages	0.230	***	0.430	***	0.539	***	0.437	***
	(0.022)		(0.028)		(0.043)		(0.029)	
Equipment	0.161	***	0.216	***	0.337	***	0.209	***
	(0.020)		(0.030)		(0.050)		(0.033)	
Personal Care	0.412	***	0.509	***	0.451	***	0.543	***
	(0.019)		(0.022)		(0.028)		(0.024)	
Leisure/education	0.128	***	-0.081	**	0.080	*	-0.114	***
	(0.015)		(0.032)		(0.047)		(0.035)	
Clothing	0.139	***	0.339	***	0.236	***	0.364	***
	(0.021)		(0.028)		(0.047)		(0.031)	
Cleaning products	0.245	***	0.416	***	0.446	***	0.361	***
	(0.022)		(0.027)		(0.037)		(0.031)	
Category								
Potatoes	0.186	***	0.298	***	0.019		0.387	***
	(0.021)		(0.029)		(0.047)		(0.030)	
Baking flour	-0.176	***	0.117	***	0.018		0.144	***
	(0.024)		(0.035)		(0.070)		(0.037)	
Chips	0.046	*	0.083	**	0.036		0.063	*
	(0.024)		(0.034)		(0.060)		(0.037)	
Cornflakes	-0.008		0.188	***	-0.006		0.224	***
	(0.021)		(0.032)		(0.058)		(0.035)	
Emmental	0.066	***	0.071	**	-0.042		0.073	**
	(0.024)		(0.034)		(0.062)		(0.036)	
Smoked salmon	0.048	**	0.025		0.017		0.023	
	(0.022)		(0.033)		(0.057)		(0.036)	
Gruyere	0.003		0.066	**	0.004		0.066	*
	(0.024)		(0.033)		(0.056)		(0.036)	
Liver pie	0.109	***	0.527	***	-0.110	**	0.533	***
	(0.018)		(0.022)		(0.047)		(0.024)	
Biscuits	0.132	***	0.005		-0.063		0.015	
	(0.024)		(0.034)		(0.053)		(0.037)	

Margarine	0.086	***	0.153	***	0.179	***	0.147	***
	(0.024)		(0.034)		(0.054)		(0.037)	
Mayonnaise	0.146	***	0.117	***	0.012		0.114	***
	(0.023)		(0.034)		(0.057)		(0.037)	
Minarine	0.055	**	0.097	***	0.040		0.097	***
	(0.025)		(0.034)		(0.069)		(0.036)	
Spaghetti	0.005		-0.051		-0.167	**	-0.084	**
	(0.025)		(0.036)		(0.072)		(0.039)	
Ice cream	0.150	***	0.173	***	0.019		0.182	***
	(0.024)		(0.034)		(0.057)		(0.036)	
Spinach	0.149	***	0.293	***	-0.113	**	0.312	***
	(0.022)		(0.031)		(0.055)		(0.033)	
Sugar	0.136	***	-0.053		0.194	***	-0.042	
	(0.026)		(0.037)		(0.057)		(0.040)	
Tomato soup	0.118	***	0.074	**	0.065		0.100	***
	(0.025)		(0.034)		(0.064)		(0.037)	
Tuna	0.187	***	0.223	***	-0.285	***	0.248	***
	(0.024)		(0.033)		(0.057)		(0.035)	
Champagne	-0.079	***	0.091	***	-0.474	***	0.110	***
	(0.022)		(0.034)		(0.072)		(0.036)	
Coke	0.046	*	0.020		0.237	***	0.025	
	(0.024)		(0.037)		(0.062)		(0.039)	
Jenever	0.143	***	0.201	***	0.002		0.237	***
	(0.024)		(0.034)		(0.055)		(0.036)	
Lemonade	-0.057	**	0.015		-0.153	**	0.040	
	(0.025)		(0.036)		(0.069)		(0.039)	
Mineral water	0.065	***	0.131	***	-0.062		0.133	***
	(0.023)		(0.033)		(0.071)		(0.035)	
Port wine	0.050	**	0.266	***	0.125	**	0.279	***
	(0.024)		(0.032)		(0.061)		(0.035)	
Tea	0.134	***	0.007		0.088		0.047	
	(0.024)		(0.035)		(0.056)		(0.037)	
Vermouth	-0.105	***	-0.067	*	-0.813	***	-0.076	*
	(0.027)		(0.038)		(0.089)		(0.040)	
Fruit Juice	0.024		-0.034		-0.080		-0.031	
	(0.022)		(0.033)		(0.060)		(0.036)	
Whiskey	0.182	***	-0.061	*	0.003		0.017	
	(0.025)		(0.037)		(0.058)		(0.039)	
Wine	0.178	***	-0.011		0.054		-0.056	
	(0.025)		(0.036)		(0.063)		(0.040)	

Beer	0.146	***	-0.073	**	0.107		-0.079	**
	(0.026)		(0.036)		(0.071)		(0.039)	
Chocolate milk	0.067	***	-0.087	**	-0.101	*	-0.127	***
	(0.025)		(0.036)		(0.061)		(0.039)	
Casserole	0.072	***	0.176	***	-0.333	***	0.156	***
	(0.025)		(0.034)		(0.068)		(0.037)	
Digital camera	0.159	***	-0.021		-0.045		-0.009	
	(0.024)		(0.034)		(0.067)		(0.036)	
Airing cupboard	0.175	***	-0.078	**	0.018		-0.113	***
	(0.024)		(0.034)		(0.061)		(0.036)	
Dvd-recorder	0.200	***	-0.046		0.164	***	-0.060	
	(0.024)		(0.035)		(0.060)		(0.038)	
Dvd-player	-0.012		-0.021		0.058		0.017	
	(0.024)		(0.035)		(0.054)		(0.037)	
Mixing tap	0.143	***	-0.098	***	-0.137	**	-0.111	***
	(0.020)		(0.038)		(0.069)		(0.041)	
Microwave oven	0.129	***	0.051		-0.005		0.092	***
	(0.025)		(0.034)		(0.064)		(0.036)	
Measuring tape	0.250	***	0.069	**	0.145	**	0.043	
	(0.023)		(0.031)		(0.058)		(0.035)	
Hedge shears	0.182	***	0.210	***	0.336	***	0.131	***
	(0.023)		(0.032)		(0.053)		(0.036)	
Knife	0.456	***	0.038		0.145	**	0.051	
	(0.022)		(0.034)		(0.072)		(0.038)	
Vacuum cleaner	0.023		-0.049		0.090		-0.026	
	(0.025)		(0.034)		(0.065)		(0.037)	
Toaster	0.360	***	-0.095	***	0.010		-0.025	
	(0.019)		(0.034)		(0.049)		(0.038)	
Washing machine	0.240	***	-0.042		0.205	***	-0.026	
	(0.022)		(0.032)		(0.053)		(0.035)	
Deodorant	0.011		0.112	***	-0.104	*	0.107	***
	(0.024)		(0.035)		(0.057)		(0.037)	
Shower gel	-0.050	**	0.040		-0.092		-0.012	
	(0.023)		(0.035)		(0.062)		(0.039)	
Hairspray	0.115	***	0.010		0.033		0.009	
	(0.025)		(0.035)		(0.057)		(0.038)	
Nail polish	0.292	***	-0.015		0.238	***	-0.040	
	(0.023)		(0.031)		(0.049)		(0.033)	
Plasters	0.075	***	0.029		0.002		0.025	
	(0.023)		(0.033)		(0.056)		(0.035)	

Toothpaste	-0.049	**	0.056		0.002		0.036	
	(0.024)		(0.035)		(0.037)		(0.055)	
Toilet paper	0.085	***	0.089	*	0.003		0.067	*
	(0.023)		(0.034)		(0.047)		(0.036)	
Nappies	0.032		0.046		0.092		0.068	*
	(0.024)		(0.036)		(0.058)		(0.038)	
Handkerchief	0.383	***	0.250	***	0.465	***	0.281	***
	(0.022)		(0.029)		(0.040)		(0.032)	
Pen	0.245	***	0.067	*	0.340	***	0.096	**
	(0.025)		(0.036)		(0.071)		(0.039)	
Home trainer	0.231	***	0.069	**	0.088	**	0.037	
	(0.021)		(0.027)		(0.042)		(0.031)	
School book	0.375	***	0.088	**	0.246	***	0.127	***
	(0.024)		(0.034)		(0.064)		(0.038)	
Cartoon	0.294	***	0.295	***	0.196	***	0.330	***
	(0.021)		(0.030)		(0.048)		(0.033)	
Football	0.388	***	-0.021		0.040		-0.056	
	(0.022)		(0.035)		(0.062)		(0.039)	
Dictionary	0.366	***	0.094	***	0.168	***	0.190	***
	(0.023)		(0.033)		(0.052)		(0.036)	
Bathing suit	0.155	***	0.200	***	-0.029		0.189	***
	(0.022)		(0.031)		(0.055)		(0.035)	
Jeans	0.114	***	0.261	***	-0.064		0.266	***
	(0.023)		(0.031)		(0.060)		(0.034)	
Jacket	0.091	***	0.178	***	-0.117		0.165	***
	(0.027)		(0.033)		(0.111)		(0.040)	
Socks	0.004		0.204	***	0.005		0.234	***
	(0.023)		(0.033)		(0.066)		(0.035)	
T-shirt	0.109	***	0.321	***	-0.040		0.325	***
	(0.023)		(0.030)		(0.053)		(0.033)	
Floorcloth	0.016		0.165	***	0.038		0.220	***
	(0.024)		(0.033)		(0.054)		(0.034)	
Toilet soap	0.088	***	0.031		-0.194	***	0.037	
	(0.026)		(0.037)		(0.071)		(0.040)	
Dishwashing detergent	-0.011		0.071	**	0.092		0.150	***
	(0.024)		(0.035)		(0.066)		(0.037)	
Soap powder	-0.111	***	0.066	*	-0.146	**	0.100	***
	(0.025)		(0.036)		(0.068)		(0.039)	

Note: * p<0.10, ** p<0.05, *** p<0.01; standard errors between parentheses

Appendix C: Detailed results individual product level

	One-month time lag				Six-month average time lag			
	γ		ρ		γ		ρ	
Champagne_1	-0.032		-0.017		-0.108		-0.062	
	(0.022)		(0.035)		(0.073)		(0.041)	
Champagne_2	0.005		0.198	***	-0.178	**	0.266	***
	(0.025)		(0.034)		(0.070)		(0.035)	
Champagne_3	0.043	*	0.062	*	0.065		0.011	
	(0.024)		(0.036)		(0.059)		(0.040)	
Champagne_4	-0.010		0.038		-0.002		0.048	
	(0.025)		(0.037)		(0.067)		(0.041)	
Coke_1	-0.032		0.113	***	-0.134	**	0.130	***
	(0.024)		(0.035)		(0.064)		(0.037)	
Coke_2	-0.022		0.304	***	0.061		0.285	***
	(0.023)		(0.032)		(0.060)		(0.035)	
Coke_3	0.016		0.181	***	0.060		0.213	***
	(0.024)		(0.035)		(0.062)		(0.037)	
Coke_4	0.022		0.107	***	-0.007		0.147	***
	(0.023)		(0.034)		(0.057)		(0.036)	
Wine_1	0.092	***	0.076	**	0.021		0.117	***
	(0.021)		(0.032)		(0.051)		(0.034)	
Wine_2	-0.009		0.087	***	-0.201	***	0.120	***
	(0.021)		(0.033)		(0.061)		(0.035)	
Wine_3	0.071	***	0.101	***	-0.023		0.122	***
	(0.022)		(0.034)		(0.056)		(0.036)	
Wine_4	0.029		0.046		-0.049		0.037	
	(0.022)		(0.034)		(0.061)		(0.037)	
Digital_Camera_1	-0.015		0.280	***	-0.014		0.341	***
	(0.023)		(0.032)		(0.059)		(0.033)	
Digital_Camera_2	-0.004		0.058	*	0.007		0.098	***
	(0.023)		(0.035)		(0.055)		(0.038)	
Digital_Camera_3	0.070	***	0.072	**	0.049		0.062	*
	(0.023)		(0.035)		(0.077)		(0.037)	
Digital_Camera_4	0.003		0.229	***	-0.279	***	0.245	***
	(0.023)		(0.033)		(0.076)		(0.035)	
Dvd-player_1	-0.043	*	0.138	***	-0.192	***	0.082	**
	(0.022)		(0.035)		(0.048)		(0.038)	
Dvd-player_2	0.047	**	0.255	***	-0.015		0.132	***
	(0.023)		(0.033)		(0.022)		(0.041)	
Dvd-player_3	0.019		0.284	***	-0.074	**	0.206	***
	(0.022)		(0.033)		(0.035)		(0.042)	

Dvd-player_4	-0.006		0.053		-0.134	*	-0.061	
	(0.025)		(0.038)		(0.070)		(0.041)	
Washing_Machine_1	-0.017		0.091	**	-0.099		0.103	***
	(0.026)		(0.037)		(0.070)		(0.039)	
Washing_Machine_2	-0.015		-0.005		-0.204	***	-0.002	
	(0.023)		(0.036)		(0.070)		(0.039)	
Washing_Machine_3	-0.029		0.135	***	-0.111	*	0.135	***
	(0.019)		(0.032)		(0.059)		(0.035)	
Washing_Machine_4	0.039	**	0.047		-0.167	***	0.047	
	(0.019)		(0.032)		(0.055)		(0.034)	
Home_Trainer_1	0.040	*	-0.031		-0.022		-0.015	
	(0.023)		(0.037)		(0.068)		(0.041)	
Home_Trainer_2	-0.013		0.129	***	0.019		0.145	***
	(0.022)		(0.034)		(0.057)		(0.037)	
Home_Trainer_3	-0.023		-0.032		-0.117	**	0.026	
	(0.023)		(0.036)		(0.052)		(0.037)	
Home_Trainer_4	0.041	*	0.052		-0.101	**	0.116	***
	(0.022)		(0.034)		(0.051)		(0.041)	
Cartoon_1	0.093	***	0.099	***	0.034		0.126	***
	(0.025)		(0.037)		(0.024)		(0.035)	
Cartoon_2	0.164	***	0.076	**	0.102	*	0.076	**
	(0.024)		(0.035)		(0.061)		(0.038)	
Cartoon_3	0.058	**	0.087	**	-0.235	***	0.009	
	(0.026)		(0.036)		(0.069)		(0.040)	
Cartoon_4	0.066	***	0.106	***	-0.125		0.004	
	(0.022)		(0.034)		(0.055)		(0.038)	
Jeans_1	0.041	*	0.222	***	-0.004		0.242	***
	(0.023)		(0.034)		(0.066)		(0.036)	
Jeans_2	-0.032		-0.021		-0.180	***	0.063	
	(0.024)		(0.038)		(0.067)		(0.039)	
Jeans_3	-0.035		0.010		-0.086		0.042	
	(0.024)		(0.037)		(0.065)		(0.040)	
Jeans_4	-0.031		0.003		-0.323	***	0.001	
	(0.025)		(0.038)		(0.073)		(0.040)	
T-shirt_1	0.076	***	0.090	**	-0.257	***	0.083	**
	(0.024)		(0.035)		(0.067)		(0.038)	
T-shirt_2	0.071	***	0.094	***	-0.041		0.114	***
	(0.024)		(0.035)		(0.065)		(0.037)	
T-shirt_3	-0.034		0.209	***	-0.260	***	0.193	***
	(0.024)		(0.034)		(0.073)		(0.037)	
T-shirt_4	0.007		0.034		-0.109		0.039	
	(0.025)		(0.037)		(0.070)		(0.040)	

Note: * p<0.10, ** p<0.05, *** p<0.01; standard errors between parentheses