UNIVERSITEIT GENT

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

# PRICE RIGIDITY IN EUROPE AND THE US: MEASUREMENT AND EXPLANATION USING SCANNER DATA

Benjamin Verhelst

2012

Supervisor: Prof. dr. Freddy Heylen

Co-Supervisor: Prof. dr. Gerdie Everaert

Submitted to the Faculty of Economics and Business Administration of Ghent University in Fulfillment of the Requirements for the Degree of Doctor in Economics

# Price Rigidity in Europe and the US: Measurement and Explanation using Scanner Data

by

**Benjamin Verhelst**

## DOCTORAL JURY

**Prof. dr. Marc De Clercq**
(Ghent University, Dean)

**Prof. dr. Patrick Van Kenhove**
(Ghent University, Academic Secretary)

**Prof. dr. Freddy Heylen**
(Ghent University, Supervisor)

**Prof. dr. Gerdie Everaert**
(Ghent University, Co-Supervisor)

**Prof. dr. Gert Peersman**
(Ghent University)

**Prof. dr. Emmanuel Dhyne**
(National Bank of Belgium and Université de Mons-Hainaut)

**Dr. Sarah Lein**
(Swiss National Bank)

## Acknowledgements

This dissertation has been in the making for approximately five years now, and even though doing research is essentially an individual endeavor, I would certainly not have made it through if it wasn't for the professional and personal support of numerous people. Now that my PhD project is completed with satisfying results, I would like to take the opportunity to express my gratitude towards all of them.

First of all, I would like to thank my thesis supervisor Freddy Heylen (Ghent University) for guiding me along from start to finish. His efforts to read my papers multiple times and to provide useful comments and suggestions for improvement, have certainly bolstered the quality of my dissertation. I would also like to thank Gerdie Everaert (Ghent University) for his feedback along the way, and Dirk Van den Poel (Ghent University) for providing me with a very rich dataset that forms the backbone of this dissertation.

I would like to express my appreciation for the other members of my doctoral jury, Gert Peersman (Ghent University), Emmanuel Dhyne (National Bank of Belgium) and Sarah Lein (Swiss National Bank), for taking a keen interest in my work. They sacrificed some of their precious time to thoroughly read and comment on my writings, which has no doubt resulted in a leap forward in the quality of my dissertation between the pre-defense and the final version. During the course of my PhD, a lot of useful remarks on the structure and the content of my papers were provided by Maarten Dossche (National Bank of Belgium), Punnoose Jacob (École Polytechnique Fédérale de Lausanne) and Julio Carrillo (Ghent University). I have also benefited greatly from the discussions with colleagues at internal seminars, and from the suggestions of several participants at the international conferences that I was able to attend. I also acknowledge the financial support from the Research Foundation Flanders (FWO), which made this dissertation possible in the first place.

A special thank you goes out to my colleagues from the departments of Social and General Economics. It is safe to say that we became a group of friends during the course of the last five years. They always provided me with ample opportunity to relieve my feelings over a (couple of) beer(s), or by showing them how to play cards properly. I will certainly miss our daily lunch break in the Brug, and I am convinced that we will keep in touch for a long time to come. I would also like to express my particular gratitude to Erna. Her door is always open to help us through the red tape, or just for a casual chat.

Although it becomes increasingly difficult to meet up on a regular basis, I would nonetheless like to thank some good friends that I still cherish from college, and my fellow students from the Bachelor and Master in Economics at Ghent University. It is very nice to see that every one of you, how different your chosen life path may be, turns out all right. It gives me great pleasure to get together every once in a while, and catch up on stories and anecdotes from past and present.

Finally, my utmost gratitude goes out to my parents and grandparents, who have been supporting me all my life and stood by my side at every stage of my personal and professional development. I would also like to thank my brother Bernd and his wife Joke, who always go out of their way to be there when we need them. A heartfelt thank you goes out to my girlfriend Valerie, who has been my love and support during the course of my PhD project. She knows firsthand about the blessings and the pitfalls of being a researcher, and got me through any difficult moments I might have had along the way. I hope I have been able to support her in an equally rewarding way, and I will be on the front row when she finishes her PhD in the near future.

Benjamin Verhelst
Ghent, September 7, 2012

**Nederlandstalige samenvatting**

Prijsrigiditeit is een belangrijk concept binnen de economische literatuur, omdat het een primordiale rol speelt in de transmissie van nominale schokken naar de reële economie. Het ligt aan de basis van de niet-neutraliteit van geld. Indien lonen en prijzen perfect flexibel zouden zijn, dan zou een wijziging in de omvang van de geldhoeveelheid zich onmiddellijk doorzetten in een equivalente prijsverandering, waardoor de reële output niet beïnvloed wordt. Omwille van de vertraagde reactie van lonen en prijzen op veranderingen in de economische context, hebben wijzigingen in de geldvoorraad wel degelijk een effect op reële variabelen. Dit geeft centrale banken de mogelijkheid om het conjunctuurverloop te sturen via hun rentebeleid. Deze hefboom is uitermate belangrijk om zowel onderbenutting als oververhitting van de economie te voorkomen en de negatieve effecten hiervan te compenseren.

Hoewel prijsrigiditeit altijd als een vitale factor in rekening wordt gebracht bij het opstellen en simuleren van dynamische conjunctuurmodellen, is tot op heden de empirische onderbouwing van de mate en de oorzaken van prijsrigiditeit relatief beperkt gebleven. Het voorliggende doctoraat probeert in dit domein een belangrijke bijdrage te verschaffen door bepaalde hypotheses in verband met prijszetting te toetsen op basis van gedetailleerde supermarkt scanner data. Op basis van een transactiedataset van een belangrijke Europese retailer trachten we meer inzicht te verwerven in de omvang en de oorsprong van prijsrigiditeit. We nemen met andere woorden onze toevlucht tot prijsgegevens op het micro niveau om bepaalde terugkerende patronen te onderzoeken, en aanbevelingen te doen voor het opstellen en evalueren van theoretische modellen op het macro niveau.

De data die door supermarkten worden verzameld via het scanning systeem aan de kassa, bieden een ongelooflijke rijkdom aan gegevens op een zeer gedetailleerd productniveau. Bovendien kunnen we via het systeem van klantenkaarten op zoek gaan naar specifieke patronen in het koopgedrag van individuele consumenten. Deze kunnen dan op hun beurt gelinkt worden aan bepaalde theoretische concepten die in de literatuur naar voor worden geschoven ter verklaring van prijsrigiditeit. De data stellen ons in staat om een aantal aspecten van micro prijszetting onder de loep te nemen die voorheen enkel als theoretisch idee bestonden, of die tot hiertoe alleen via enquêtes werden ondersteund. We trachten dus een empirische onderbouwing te voorzien voor enkele belangrijke theoretische concepten aan de basis van prijsrigiditeit.

In de eerste paper onderzoeken we de graad van prijsrigiditeit in de data van onze Europese supermarkt, en vergelijken deze met een Amerikaanse tegenhanger. We komen tot de conclusie dat aggregatie van prijsdata over de tijd een belangrijke impact heeft op de gemeten prijsrigiditeit, en op de vergelijking van afzonderlijke regio's. Onze basisanalyse gebruikt de prijsdata van zowel de Europese als de Amerikaanse supermarktketen in hun oorspronkelijke frequentie, wat betekent dat we geen prijswijzigingen missen. Als alternatieve analyse zetten we de data eerst om in maandelijkse gegevens om de structuur te kopiëren van CPI-data. Deze laatste worden gebruikt om de consumptieprijsindex op te stellen, en zijn de meest gangbare gegevensbron voor onderzoek naar prijsrigiditeit. In de resultaten van onze basisanalyse blijken de prijzen van de Amerikaanse retailer heel wat flexibeler te zijn dan die van de Europese keten. Dit verschil gaat echter volledig verloren wanneer we de maandelijkse datasets tegenover elkaar plaatsen. De oorsprong van dit fenomeen moeten we zoeken in de prijspolitiek van de retailers in kwestie. Daar waar de Amerikaanse supermarktketen zijn prijzen elke week evalueert en eventueel aanpast, doet de Europese keten dat slechts om de twee weken. Als gevolg hiervan worden de statistieken van de Amerikaanse retailer sterker vertekend dan degene van de Europese supermarkt bij de omzetting van basis- naar maandelijkse frequentie. Op het macro-economische niveau impliceert dit dat regio's waar een sterke korte termijn prijspolitiek wordt gevoerd, onterecht als rigide kunnen worden aanzien wanneer data worden bestudeerd die niet in hun basisfrequentie beschikbaar zijn.

Het tweede project onderzoekt het bestaan van impliciete contracten in retailing. Uit enquêtes is meermaals gebleken dat bedrijven deze ongeschreven akkoorden beschouwen als de belangrijkste reden om hun prijzen stabiel te houden. Op die manier kunnen zij hun vaste klanten aan zich binden. Deze waarderen dan weer de stabiele prijspolitiek en kunnen zich aldus de moeite besparen om op zoek te gaan naar alternatieven. In periodes met sterke vraag zullen de bedrijven of retailers hun prijzen niet verhogen, omdat ze hun vaste klanten niet tegen de borst willen stoten. Wanneer de vraag zwak is, zullen ze hun prijzen niet verlagen omdat dit enkel koopjesjagers aantrekt die heel wat minder interessant zijn voor hun winstmarge. Gebruik makend van de gedetailleerde data op het niveau van de individuele consument die we ter beschikking hebben, slagen we erin om twee hypotheses te staven die primordiaal zijn binnen de impliciete contract theorie. Ten eerste vinden we dat de prijselasticiteit van loyale klanten lager is dan van sporadische klanten. Als de retailer er in slaagt om via een stabiele prijspolitiek de klantentevredenheid en –loyaliteit te verhogen, dan gaat automatisch de prijselasticiteit van hun klantenbestand omlaag en de winstmarge omhoog. Ten tweede tonen onze resultaten ook aan dat de vraagcurve van de loyale klanten concaver is dan deze van de niet-loyale consumenten. Dit impliceert dat het vaste cliënteel negatiever reageert op prijsstijgingen, en minder positief op prijsdalingen, in vergelijking met sporadische klanten. De verklaring voor het eerste punt moeten we zoeken in het feit dat vaste klanten de huidige prijzen kunnen vergelijken met wat ze betaalden in de vorige periode, en aldus sterk zullen reageren op een hogere prijs. Sporadische klanten daarentegen bezoeken de supermarkt niet frequent genoeg om een dergelijke prijsvergelijking te kunnen maken en reageren dus minder sterk. De niet-loyale klanten reageren dan weer sterker op prijsdalingen omdat zij het profiel hebben van koopjesjagers die actief lage prijzen zoeken, daar waar de vaste klanten sowieso gekocht zouden hebben aan de oorspronkelijke prijs. Het overtuigende empirische bewijs dat we vinden voor deze hypotheses onderschrijft de belangrijke rol die impliciete contracten spelen als bron van prijsrigiditeit.

De derde paper focust op de impact van gewoontevorming en sociale invloed op het consumptiegedrag van mensen. Onze bedoeling is om het theoretische concept van gewoontevorming en sociale invloed op productniveau aan een empirische test te onderwerpen, via het schatten van vraagmodellen op verschillende niveaus van productaggregatie. Hierbij maken we gebruik van zeer recente spatial econometrics technieken om niet alleen de tijdsdimensie maar ook de geografische dimensie van het consumptiepatroon van gezinnen in overweging te nemen. Om gewoontevorming te meten, nemen we het eigen koopgedrag van de vorige periode op in een model ter verklaring van het huidige consumptiepatroon. De sociale invloed nemen we dan weer in rekening door het consumptiegedrag van een geografische referentiegroep in het model op te nemen. De resultaten die we bekomen bij het schatten van onze vraagmodellen tonen een gemengd beeld met betrekking tot gewoontevorming. Bij de huidige consumptiebeslissingen ondervinden gezinnen een beperkte positieve invloed van hun eigen consumptie in de vorige periode op het niveau van productgroepen en -categorieën. Dit effect is echter grotendeels afwezig op het niveau van de individuele producten. We vinden daarentegen wel overtuigende bewijzen van sociale invloeden op het koopgedrag van mensen, ongeacht het niveau van productaggregatie. Voor zowat alle producten in onze sample vinden we een positieve invloed van het consumptieniveau van de referentiegroep op dat van de consument onder beschouwing. Keeping up with the Joneses blijkt dus wel degelijk een belangrijke factor te zijn in het koopgedrag van de doorsnee consument.

# Contents

# CHAPTER 1

## Introduction

# Introduction

In this introductory chapter, we first discuss the general context of this dissertation in order to position it in the economics literature and to underscore the importance and relevance of price rigidity for business cycle analysis. The second section looks in more detail into the specific research questions that we tackle, and the broad theoretical and empirical frameworks that are used throughout the analysis. We will then also offer a glimpse of our results and their implications.

## 1 Orientation and Motivation

The important role of price rigidity in explaining business cycle movements is well documented in the economics literature. Sticky prices are an integral part of most macro models that try to mimic the available data. It is one of the basic features of economic behaviour at the basis of the non-neutrality of money. Without price and wage stickiness, there would be no impact of money or monetary policy on real output. The delayed reaction of prices and wages leads to persistence of nominal shocks, and it gives central banks a powerful leverage mechanism to stabilize the economy and control inflation. Given its vital role for business cycle analysis, a thorough understanding of the extent and the sources of price stickiness is crucial.

This dissertation deals with different aspects of micro price setting and explores the possibilities of supermarket scanner data as a valuable source of micro data to measure price stickiness and study its origins. Only for the last decade or so, micro price data have been used extensively to measure and explain sticky prices. A decisive step in this process was taken in 2003, when the Eurosystem of Central Banks decided to set up the Inflation Persistence Network (IPN). Extensive consumer and producer price data were used to study the price setting practices in different European countries. Dhyne *et al.* (2005) give an overview of the results of this program. Across the Atlantic, similar initiatives led to the analysis of comparable CPI data, collected by the Bureau of Labor Statistics (BLS). Bils and Klenow (2004), Nakamura and Steinsson (2008) and Klenow and Kryvtsov (2008) gradually refined the measurement methodology and presented high quality empirical results on price stickiness.

Although CPI data offer an invaluable source of information on micro pricing decisions, this dissertation relies extensively on another type of micro data to study the drivers and consequences of retail price setting practices. When an individual product is bought in a supermarket, its bar code is scanned at the counter and information about this transaction is saved to a large database. These scanner data usually cover a more limited array of products compared to CPI data, but they have the major advantage that pricing points are available at a much higher frequency, and that each transaction can be traced back to an individual customer through a system of loyalty cards. Consequently, scanner data give information about which individual card holder buys which product at what time and what price. This opens up new possibilities to measure and explain price rigidity at a very detailed level.

The common theme running through each chapter of this dissertation is the use of a scanner dataset made available to us by an anonymous European retailer [1]. The highly detailed level of product definition and the individual customer dimension of the data based on compulsory loyalty cards implies that we can study certain aspects of micro price setting that would be hard or impossible to tackle in any other data context. We are able to empirically test theoretical concepts about the origins of price stickiness that have remained unexplored before, or that have only been backed up by survey results. The main contribution of this dissertation is to take certain features of the price setting process out of the realm of theoretical constructs and into the world of empirically supported facts. The next section spells out the different research questions that we consider, gives an overview of the theoretical and empirical methodologies used in the analysis, and offers a broad overview of the main results that we obtain.

## 2 Research questions, results, and contributions

The **first paper** of this dissertation concentrates on estimating the extent of price rigidity using scanner data from two distinct retailers, one in Europe and one in the US. Based on results of Dhyne *et al.* (2005) and Bils and Klenow (2004), among others, a consensus had been growing that prices in Europe are much more rigid than in the US. This result is also consistent with the smaller reaction of inflation to changes in real marginal cost in Europe, as found by Gali *et al.* (2001). Nonetheless, Nakamura and Steinsson (2008) and

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

Klenow and Kryvtsov (2008) find that this difference between Europe and the US may be largely due to the differential use of temporary sales. Looking at regular, sales-filtered prices, they find quite similar price stickiness parameters in European and US CPI data, pointing to a more extensive use of temporary sales promotions in the US.

The purpose of the first paper is to shed some additional light on the regional comparison of price rigidity by estimating stickiness parameters for both posted and regular prices on the basis of supermarket scanner data from a European and US retailer. As a benchmark for our anonymous European retailer, we use the publicly available scanner data of Dominick's Finer Foods, an important supermarket chain in the Chicago area. The main asset of both datasets is that, contrary to CPI data, the price quotes are available in their base frequency, which means that prices do not change in between observation points. Consequently, we do not lose any short-term pricing information. This is important due to the fact that many price changes, especially temporary ones, are very short-lived and hence not picked up by monthly data. Hosken and Reiffen (2004), Kehoe and Midrigan (2012), and Midrigan (2010) have clearly shown this in their scanner data studies. It is also an important factor in our data. Without proclaiming to derive macro patterns from the pricing information of specific supermarket scanner data, we can prove that ignoring potentially different short-term pricing tactics across regions can severely impact the spatial comparison of price stickiness.

First, we calculate the frequency of price change in both the European and US scanner data at base frequency. Then, we transform the original data into monthly frequency by only retaining one price quote each month, thus mimicking the structure of CPI data. In the second stage of the analysis, we recalculate the price change frequency based on those monthly price series. Comparing both sets of parameters enables us to determine the influence of short-term pricing decisions on estimated price stickiness, and above all shows us if the comparison between retailers leads to the same outcome in base and monthly frequency. If not, we can conclude that performing comparative analysis between regions based on data that are not in their base frequency, can lead to severely biased results if short-term pricing behaviour is different across those regions.

The results from our analysis show that time aggregation indeed influences the cross-regional comparison of price stickiness. Whereas prices at base frequency are much more flexible in our US scanner data than in our European data, this difference disappears in the derived monthly price series. This is true for both posted and regular prices, irrespective of the type of sales filter that we apply. At the gist of this result, there is a different price review policy on behalf of the retailers. Whereas the European retailer reviews prices every two weeks, the US retailer does so every week. The more flexible price review policy of the US retailer gets lost in the time-aggregation process from base to monthly frequency, leading to a more pronounced downward bias in the price change frequency of the US retailer. Extending this argument to the macro level, it implies that regions where short-term pricing tactics are widely used will appear comparatively more rigid when time-aggregated data are studied.

The **second paper** of this dissertation assesses the influence of customer loyalty on the shape of the demand curve, by exploring the unique availability of price and quantity data at the level of the individual customer in our European scanner dataset. Securing and enhancing customer loyalty is a major objective for any firm that wants to succeed in a competitive market environment. When we look at survey results based on interviews and questionnaires completed by a large sample of firms, long-term firm-customer agreements or implicit contracts systematically outweigh other declared sources of price stickiness, including menu costs and sticky information. This is true for the surveys of Blinder (1991, 1994) and Blinder *et al.* (1998) for the United States, Hall *et al.* (2000) for the United Kingdom, Apel *et al.* (2005) for Sweden, and Fabiani *et al.* (2005) for the euro area. Although the survey evidence is overwhelming, analytical support for implicit contracts based on actual consumption data is non-existent. We try to fill this gap by exploiting the detailed customer-level data availability of our European retailer.

To undertake this task, we revert to the implicit contract theory of price stickiness put forward by Okun (1981). The essence of his argument is that price rigidity originates from an invisible handshake between a firm and its customers. A stable price setting policy is in the interest of the firm because it discourages their loyal clientele from shopping elsewhere. It also serves the interests of the customers, because they can trust the price level to be fair, so that they can cut back on search efforts and minimize their shopping costs. The development of long-term relationships in customer markets creates a mutual advantage of sticky prices. If market conditions are tight, the firm will not increase prices because they do not want to risk antagonizing their regular customers. If demand is weak, they do not decrease prices because this would only attract the less interesting bargain hunters.

We empirically test two key hypotheses of Okun's theory of customer markets. Firstly, the price elasticity of demand of loyal customers should be lower than that of non-loyals. This determines the gain of stable prices for the firm or retailer. If they succeed in increasing customer loyalty through a stable price setting policy, they can decrease the price elasticity of demand of their customer base and increase their profit margins. Secondly, the curvature of demand of loyal customers should be higher than that of non-loyals, where the curvature is defined as the concavity of the demand curve. It would imply that regular customers react more negatively to price increases, and less positively to price decreases than non-loyals. Okun (1981) explains the first aspect because loyal customers are very responsive to prices that exceed the price level observed in the past. Random shoppers do not have a personal purchase history and can only act upon the current observed price, so their response will be smaller. When the relative price level decreases, these random shoppers will react more strongly because they are more likely to be bargain hunters that actively search for low prices. The regular customers on the other hand were ready to buy at the previous price anyway, so the impact of a price decrease on their consumption level will be smaller.

Before we can test these hypotheses, we need to divide the pool of customers according to their behavioural loyalty towards the retailer. With the frequency and monetary value of purchases as the input variables, we use segmentation techniques and clustering analysis to split the customer base in three segments. The top and bottom segment can subsequently be used in a comparative demand analysis in order to test our hypotheses outlined above. For the estimation of the price elasticity and curvature parameters, we use the Behavioural Almost Ideal Demand System (B-AIDS) of Dossche *et al.* (2010), which is an extension of the standard AIDS model of Deaton and Muellbauer (1980), capable of fully capturing non-linearities in the demand curve. We estimate the model at the individual product level for a wide range of different product categories, using Seemingly Unrelated Regression (SUR) as the preferred estimation methodology.

Although the estimated elasticity and curvature parameters vary considerably across product categories, the results of our empirical demand analysis confirm the validity of the two hypotheses underlying Okun's theory of implicit contracts. We find that loyal customers are indeed less price elastic than non-loyals, although the result is insignificant in the aggregate. More importantly, the demand curve is more concave for regular customers than for random shoppers. The former value a stable price, whereas the latter are merely looking for bargains. These results hold in the aggregate and most of the individual product categories in our sample. These findings provide an important incentive to the firm or retailer to commit to a sticky price. Price increases are a deal breaker for the highly valuable regular customers, whereas price decreases only attract bargain hunters that will shop elsewhere when prices go back to their regular level. These results support the role of implicit contracts as source of price stickiness.

The **third paper** of this dissertation focuses on the estimation of internal and external deep habits, and their relevance as a driving force of price stickiness. Ravn *et al.* (2006) introduce the idea of deep habits that are formed on a product-by-product basis, as opposed to the more standard superficial habits that are formed at the aggregate level over a composite good. Internal deep habits imply that an individual household's consumption level of a product is positively influenced by the choice behaviour of the own household in the past, whereas external deep habits entail that people are guided in their purchasing behaviour towards products that are being bought by people around them. Ravn *et al.* (2006) show that deep habits make the optimal pricing problem of the firm dynamic, and lead to a countercyclical mark-up, in line with the empirical evidence. Deep habits also help to mimic procyclical labour and real wage dynamics that are present in the data. Nakamura and Steinsson (2011) argue that in the presence of deep habit formation, firms benefit from committing to a sticky price. Their theoretical model leads to a firm-preferred equilibrium with prices at or below a price cap. Hence, deep habits can help to explain why retail data often exhibit a rigid regular price and frequent sales promotions.

Although the concept is gaining momentum in the theoretical literature, there is no readily available empirical support for the deep habit hypothesis. The purpose of our third paper is to empirically investigate the importance of deep habit formation in consumption, using the detailed customer-level scanner data from our European retailer. We estimate an expenditure model that includes both a time and a spatial lag capturing internal and external deep habits, respectively. The former captures inertia or persistence in consumption, whereas the latter captures preference interdependence across households, or keeping up with the Joneses. The model is estimated at different levels of product aggregation, in order to compare the strength of superficial and deep habits, and to see at which level deep habit formation is most pronounced.

We estimate our expenditure models at the zip code level. In other words, the spatial unit that we consider is the zip code, and all individual customer data are aggregated to this level of analysis. We choose to estimate a time-space simultaneous model, which tests if the expenditure share of a certain good at time $t$ in zip code area $i$ depends on the expenditure share of that good at time $t-1$ in zip code area $i$, i.e. internal deep habit, and/or on the expenditure share of that good at time $t$ in all zip code areas $j$ that form the reference group for area $i$, i.e. external deep habit. This part of the analysis most closely resembles the estimation methodology used by Korniotis (2010) for his examination of internal and external superficial habits in US state-level consumption. An integral aspect to determine preference interdependence among zip code areas is the way in which the reference group is established. We choose for a contiguity-based spatial weight matrix, which implies that expenditure in a certain zip code area $i$ can potentially be affected by expenditure in all zip code areas $j$ that share a common border with $i$. Given the two-directionality of the neighbour relationship in space, we have to take into account the simultaneity of the spatial lag in our estimation. Due to the presence of spatial fixed effects, we also have to correct our estimates for the incidental parameter bias. We therefore use 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.

Our results from estimating the model at the product group and product category level point to the existence and significance of both internal and external deep habits for all product groups and the majority of product categories in our sample. The results of our demand analysis at the individual product level provides strong evidence for external habit formation, but internal habits are largely absent at this level of product aggregation. Deep habits are present in most of our empirical set-ups, implying that households derive utility from comparisons with their own past consumption pattern, and with the consumption behaviour of a reference group. The deep habit parameters are however below the superficial habit benchmark, and they decrease in strength when the level of product aggregation becomes more detailed. The parameters that we obtain are also below the calibrated values that have been used in the literature.

## References


Apel, M., Friberg, R. and Hallsten, K. (2005). Microfoundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms. *Journal of Money, Credit, and Banking*, **37** (2), 313–338.

Bils, M. and Klenow, P. (2004). Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy*, **112** (5), 947–985.

Blinder, A. (1991). Why are Prices Sticky? Preliminary Results from an Interview Study. *American Economic Review*, **81** (2), 89–96.

— (1994). On Sticky Prices: Academic Theories Meet the Real World. In G. Mankiw (ed.), *Monetary Policy, NBER Studies in Business Cycles*, Chicago: The University of Chicago Press, pp. 117–154.

—, Canetti, E., Lebow, D. and Rudd, J. (1998). Asking About Prices: A New Approach to Understanding Price Stickiness. *Russell Sage Foundation*, 380p.

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

Dhyne, E., Alvarez, L., Le Bihan, H., Veronese, G., Dias, D., Hoffman, J., Jonker, N., Lünneman, P., Rumler, F. and Vilmunen, J. (2005). Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data. *ECB Working Paper*, **No. 524**, 50p.

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.

Fabiani, S., Druant, M., Hernando, I., Kwapil, C., Landau, B., Loupias, C., Martins, F., Mathä, T., Sabbatini, R., Stahl, H. and Stockman, A. (2005). The Pricing Behaviour of Firms in the Euro Area: New Survey Evidence. *ECB Working Paper*, **No. 535**, 47p.

Gali, J., Gertler, M. and Lopez-Salido, D. (2001). European Inflation Dynamics. *European Economic Review*, **45** (7), 1237–1270.

Hall, S., Walsh, M. and Yates, A. (2000). Are UK Companies' Prices Sticky? *Oxford Economic Papers*, **52** (3), 425–446.

Hosken, D. and Reiffen, D. (2004). Patterns of Retail Price Variation. *RAND Journal of Economics*, **35** (1), 128–146.

Kehoe, P. and Midrigan, V. (2012). Prices Are Sticky After All. *Federal Reserve Bank of Minneapolis - Research Department Staff Report*, **No. 413**, 41p.

Klenow, P. and Kryvtsov, O. (2008). State-Dependent or Time-Dependent Pricing: Does It Matter for Recent US Inflation? *Quarterly Journal of Economics*, **123** (3), 863–904.

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.

Midrigan, V. (2010). Menu Costs, Multi-Product Firms and Aggregate Fluctuations. *Econometrica*, **79** (4), 1139–1180.

Nakamura, E. and Steinsson, J. (2008). Five Facts about Prices: A Reevaluation of Menu Cost Models. *Quarterly Journal of Economics*, **123** (4), 1415–1464.

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

Okun, A. (1981). Prices and Quantities: A Macroeconomic Analysis. *The Brookings Institution*, 371p.

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

# CHAPTER 2

## Estimated Price Rigidity in European and US Scanner Data: The Influence of Time Aggregation

# Estimated Price Rigidity in European and US Scanner Data: The Influence of Time Aggregation *

Benjamin Verhelst †  Dirk Van den Poel ‡

**Abstract**

Using scanner data from a European and a US retailer, this paper qualifies earlier results on price stickiness based on monthly CPI data. The European retailer reviews prices every two weeks, whereas the US retailer does so on a weekly basis. When we study the original data in base frequency, regular (sales-filtered) prices are far more flexible in the US data. This finding collapses however when we study monthly price series derived from our high frequency scanner data. Then, regular prices show the same degree of flexibility in the US and the European data. This result proves that time aggregation of the price series from base to monthly frequency may wipe out the more flexible price review strategy of the US retailer and imply upward biased estimates of price rigidity. At the macro level, it implies that a comparison of countries or regions where different short-term pricing strategies are used, based on data that are not in their base frequency, might lead to spurious conclusions.

**JEL**: C33, D4, E3, L66

**Keywords**: Price setting, Scanner data, Frequency of price change, Sales filtering

*The authors would like to thank Freddy Heylen, Gerdie Everaert, Gert Peersman, Emmanuel Dhyne, Sarah Lein, Maarten Dossche and Punnoose Jacob for valuable comments. We have also benefited from comments received at the 16th International Conference on Computing in Economics and Finance (London, 2010) and the 22nd Annual Meeting of the Society for Economic Dynamics (Ghent, 2011). 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 Flemish Fund for Scientific Research (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

A large literature documents the persistent effects of nominal shocks on real output and inflation (Peersman, 2004; Christiano *et al.*, 2005). Given the key role of price rigidity to explain this persistence, many micro-based models of price setting have been developed for macroeconomic models. The empirical assessment of the price-setting process at the micro level, however, has until recently been limited. This paper makes an empirical contribution to the literature in this field by estimating price stickiness from two high frequency supermarket scanner datasets, one in Europe and one in the US. Our results qualify earlier findings of Bils and Klenow (2004), Dhyne *et al.* (2005), Nakamura and Steinsson (2008) and Klenow and Kryvtsov (2008), and underscore the influence of the frequency at which price quotes are available on estimated price stickiness parameters. We show that neglecting this frequency issue can lead to spurious results and potentially blur cross-sectional comparisons of price rigidity.

Until the start of the 2000s, most micro-based price studies were based on relatively small samples. Although these studies show that price changes are infrequent, they offer limited understanding of micro price setting at large because they concentrate on a small number of goods sold in one geographical area (Cecchetti, 1986; Lach and Tsiddon, 1992; Kashyap, 1995). A major step forward in the empirical study of price setting at the micro level was made in 2003, when the Eurosystem of Central Banks set up the Inflation Persistence Network (IPN). This led to the collection of large datasets of consumer and producer prices from various euro area countries, used to construct the Consumer and Producer Price Indices (CPI and PPI). Dhyne *et al.* (2005) summarize the results obtained in a number of national studies and build upon these to deduce the specifics of the price setting process in the euro area.

At roughly the same time of the IPN formation, Bils and Klenow (2004) compiled similar price records underlying the computation of the CPI for the United States, collected by the Bureau of Labor Statistics (BLS). Consequently, it became an obvious research objective to compare the price flexibility in Europe and the US based on large datasets of consumer prices, as differences in price stickiness call for different policy measures [1].

[1] Comparable studies with similar datasets have been conducted by Baharad and Eden (2004) for Israel, Masahiro and Yumi (2007) for Japan, Karadi and Reiff (2007) for Hungary, Coricelli and Horvath (2010) for Slovakia and Gagnon (2009) for Mexico.

As sticky prices raise the output cost of stabilising inflation, a central bank needs to take the stickiness of prices into account when deciding which weights to assign to output and inflation variability in the monetary policy reaction function.

Dhyne *et al.* (2005) find an average euro area monthly frequency of price change of around 15 percent, and a weighted median price duration of 10.6 months. The results of Bils and Klenow (2004) for the US, based on a comparable set of product categories, show an average monthly frequency of price change in the neighbourhood of 25 percent and a weighted median price duration of 4.6 months. Based on these estimates, which are calculated from millions of monthly price observations over several years for a broad array of products, euro area prices appear far stickier than US prices. The observation of higher micro price rigidity in Europe than in the US was fully consistent with earlier macroeconomic estimates showing a smaller reaction of inflation to changes in real marginal cost in Europe (Gali *et al.*, 2001).

Nakamura and Steinsson (2008) analyze the frequency of price change more thoroughly and emphasize the importance of product turnover and especially temporary sales promotions in US CPI data. If the associated temporary price changes are believed to be orthogonal to macroeconomic conditions, then regular sales-filtered prices are more interesting than posted prices when studying macroeconomic issues. Consequently, an adequate filtering of the product turnover and sales episodes from the posted price series is required. Whereas Bils and Klenow (2004) eliminate sales by using the proportion of sales in the overall data, Nakamura and Steinsson (2008) correct for sales directly using the detailed price level information of each separate product [2]. After filtering out price changes associated with product turnover and temporary sales promotions, they obtain a frequency of price change in their US data that is much closer to the ones found in comparable euro area data. Hence, the gap in price stickiness between Europe and the US that is found in the posted price series could simply reflect that US retailers use sales promotions more often than their European counterparts. Klenow and Kryvtsov (2008) look at this issue in a slightly different way, as they consider price changes associated

[2] The BLS Commodities and Services Substitution Rate Table that is used by Bils and Klenow (2004) to study price rigidity does not contain information on the magnitude of price changes during sales periods. It only contains the share of price quotes that involve some change in price. Because the BLS collects prices net of sales and other promotions, Bils and Klenow (2004) only have the proportion of sales in the overall data at their disposal to construct the regular price series.

to product turnover to be price changes fair and square, and therefore only filter out temporary sales promotions. On top of that, they also take into account non-adjacent price quotes, whereas Nakamura and Steinsson (2008) only consider adjacent price observations. Nonetheless, the results of Klenow and Kryvtsov (2008) point to the same conclusion as obtained by Nakamura and Steinsson (2008). Whereas the probability of a posted price change is higher in the US than in the euro area, the probability of a regular price change is highly similar in both regions.

Besides CPI data, there is another important source of micro data that offers a vast amount of pricing points over time and for a broad range of products, that being scanner data. Unlike CPI data, which are gathered by government officials, scanner data are obtained through the scanning process of product-specific bar codes that is in common use in supermarkets and drugstores. Scanner data therefore contain all price quotes of items that are actually being bought by consumers. The retailing sector is huge and consumer-retailer interactions mimic conditions present in many other types of purchases, making supermarket data highly valuable to study consumer and producer behaviour (Chevalier *et al.*, 2003; Dossche *et al.*, 2010) [3].

Although CPI data are in general much broader in terms of product coverage than scanner data from any given supermarket, the latter have a very important asset in that they are generally measured on a weekly basis, whereas CPI data are measured monthly, at best. Our purpose is to use the higher frequency of scanner data to revisit the results on price stickiness obtained in CPI-studies. We analyze price records in their base frequency, which means that the retailer does not change the price of its products in between data points. We compare the frequency of price change that we obtain from our datasets at base frequency with the ones from datasets at monthly frequency that we artificially derive from the original high frequency data. As such, we are able to gauge the effects of short-term pricing tactics on the stickiness of prices and on the validity of the comparison between Europe and the US.

Our results show that posted and regular prices of the US retailer are far more flexible than those in the European data when we use the full amount of information contained in our scanner datasets. However, switching from base frequency data to derived monthly

[3] Scanner data have been used extensively in macroeconomic research, see e.g. Levy *et al.* (1997), Dutta *et al.* (1999), Eichenbaum *et al.* (2011), Midrigan (2010) and Kehoe and Midrigan (2012).

data, the difference largely disappears, leaving the frequency of price change at a similar level. This dichotomy stems from a more aggressive price review policy of the US retailer, and highlights the importance of the frequency at which price quotes are available in the data. In a macro setting, it implies that the use of monthly data to compare price stickiness between countries or regions might lead to spurious results in case of differences in short-term pricing tactics.

The remainder of this paper is organized as follows. In section 2, we describe the scanner data and give some descriptive statistics. Section 3 describes the estimation methodology and presents the main results with respect to the frequency of price change. The robustness of these results with respect to the sales filtering method is ascertained. In section 4, we decompose price changes into increases and decreases, and study their frequency. Section 5 looks at the size of price changes. Section 6 concludes.

## 2 Data

We use two scanner datasets, both originating from large and geographically dispersed retailers. For the US, we use the publicly available scanner data from 86 stores of Dominick's Finer Foods, an important supermarket chain in the Chicago metropolitan area. The time span of these data runs from September 1989 to May 1997. The data have already been used extensively for similar research purposes (Peltzman, 2000; Chevalier *et al.*, 2003; Rotemberg, 2005; Ray *et al.*, 2006; Midrigan, 2010; Kehoe and Midrigan, 2012). For the euro area, we use unique scanner data from 6 stores of an anonymous European retailer [4]. These data run from January 2002 to April 2005. They have only been used before by Dossche *et al.* (2010) to test the existence of the kinked demand curve and to estimate its curvature.

The different timing of the data should not pose any problem arising from varying growth or inflation rates. During a recession, price setters might adapt their prices more swiftly to the economic environment, because they become more attentive and revise their prices more often during tough times (Klenow and Malin, 2010). In this respect, it is important to note that our datasets include both prosperous and recessionary episodes, so that the comparison between them is not affected by a possible increase in the frequency of price

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

change during an economic downturn [5]. Inflation can also spur additional price changes, as price setters want to keep up with the changing macro price level. However, both our data episodes are situated in a low inflation environment, the average aggregate inflation rate during the sample period being 3.43 percent in the US and 2.60 percent in Europe. This small difference can hardly explain any potential gap in price stickiness. Although the composition of price increases and decreases changes with inflation, these movements almost entirely cancel out leaving the frequency of price change unaffected (Dhyne *et al.*, 2005). Hence, low to moderate inflation only shows up at the intensive margin, i.e. the size of price changes, not at the extensive margin (Klenow and Kryvtsov, 2008). Only during high inflation episodes, the frequency of price change might be affected as well (Gagnon, 2009; Wulfsberg, 2009).

An important aspect of these datasets is that they are available in their base frequency. We know that the retailers do not change their prices in between data points. For the US data, the base frequency is weekly, the European data are available at biweekly frequency. Consequently, we do not lose any price information in between price quotes, so no measurement error occurs when estimating price change frequencies. We restrict the scanner datasets by selecting only those product categories that are represented in both, as to make them comparable. This is very important because the frequency of price change is known to be extremely heterogeneous across products, implying that a proper assessment and comparison of price stickiness must be based on similar product baskets (Bils and Klenow, 2004; Dhyne *et al.*, 2005; Nakamura and Steinsson, 2008). By doing so, we end up with 16 product categories containing 9261 products for the US dataset, and 18 categories containing 1270 products in the European data [6]. In appendix I, we list all product categories that we include in our European and US datasets, with their respective number of products. All categories contain a large number of fairly homogeneous products, defined at a very detailed level by their Universal Product Code. One UPC will for example be connected to a package of 20 scrolls of white toilet paper of a certain brand. The raw data consist of individual price trajectories, i.e. sequences of elementary price quotes for a specific product in a specific outlet.

[5] The US economy suffered from a downturn in 1990-1991, whereas the European data include the downturn of 2001-2003.

[6] The difference in the number of product categories stems from the fact that some of them are defined more broadly in US data compared to the classification scheme of the euro area retailer. The product coverage of both our constructed datasets is however highly similar.

# 3 Frequency of Price Change

In this section, we present statistics on the frequency of price change for both the US and European retailer. An important aspect of the analysis is the type of price series that we study, as previous research on price setting behaviour using micro data has shown that excluding promotional prices from the data has a major impact on inference about the stickiness of prices. Not surprisingly, there is an ongoing debate in the literature about how to define a sales promotion and whether one should treat regular and sales prices asymmetrically (Eichenbaum *et al.*, 2011).

## 3.1 Temporary sales

Before going into the identification and filtering of sales prices and the comparison of results on the frequency of price change for posted and regular price series of the European and US retailer, we digress on the phenomenon of temporary sales that is apparently very important for price setters, especially in retailing. There is an ongoing discussion about which prices should be taken into account when thinking about the macroeconomic implications of price rigidity and calibrating macroeconomic models.

Assuming that temporary sales have macro content, they have to remain part of the price series in order to find the appropriate measure of price stickiness for macro policy purposes. If on the other hand they are considered orthogonal to macroeconomic aggregates, they do not represent an actual form of rigidity and can be left out. Standard staggered price models focus on permanent price changes and either drop temporary price changes from the data, the temporary-changes-out approach, or leave them in and treat them as permanent, the temporary-changes-in approach. Kehoe and Midrigan (2012) propose a menu cost model that includes motives for both permanent and temporary price changes by including a parameter on the technology of price adjustment.

To decide about the macro content of sales prices, the obvious thing to do is to look at the motives of price setters to offer temporary sales, a topic that has been studied extensively in the industrial organization literature. Adapting to idiosyncratic shocks that hit the firm is a first straightforward motive to set a sales price. In other words, it is a simple way for the firm to adapt prices continuously to fluctuations in its costs or demand. A second often cited motive for temporary sales is intertemporal price discrimination, i.e. retailers lower their markup and price of a certain product in periods of high price elasticity of demand for that product, so as to maximize the opportunity to gain market share (Kehoe and Midrigan, 2012). Loss-leader behaviour on behalf of competing retailers is a closely related motive for temporary sales, where they discount items in high relative demand, even if this does not coincide with a period of high aggregate demand, as this signals to consumers that all unadvertised products are sold at their reservation price (Chevalier *et al.*, 2003).

All three of these sales motives can broadly be described as orthogonal to macroeconomic aggregates. Consequently, if these motives are believed to be important, using regular prices to calibrate macroeconomic models is advocated. Additionally, due to the transience of price adjustment associated with sales, a given number of price changes due to sales yield much less aggregate price adjustment than the same number of regular price changes, further strengthening the case for studying regular price series. However, there might also be motives for temporary sales that do contain macro content, for example when inflation has been low or when excess inventory builds up (Klenow and Kryvtsov, 2008). In the latter case, clearance sales serve as a perfect method for inventory management in case of shifts in aggregate demand or unpredictable shifts in tastes (Nakamura and Steinsson, 2008). When inventory is large, reducing inventory reduces costs, so that the expected marginal cost of selling an item is lower in periods of high inventory (Hosken and Reiffen, 2004). In this case, the magnitude and duration of temporary sales do respond to macroeconomic shocks, hence they should not be excluded from the analysis (Bils and Klenow, 2004).

As the macroeconomic relevance of temporary sales is not the main issue of the paper, we will provide estimates of the frequency of price change for both posted and regular prices. This has the advantage that the results can be used in a flexible way in a variety of economic settings, and are easy to compare with previous studies. An additional advantage is that statistics for both posted and regular prices are important if one wants to calibrate a hybrid model in which firms face a different cost for a temporary versus permanent price change, as in Kehoe and Midrigan (2012).

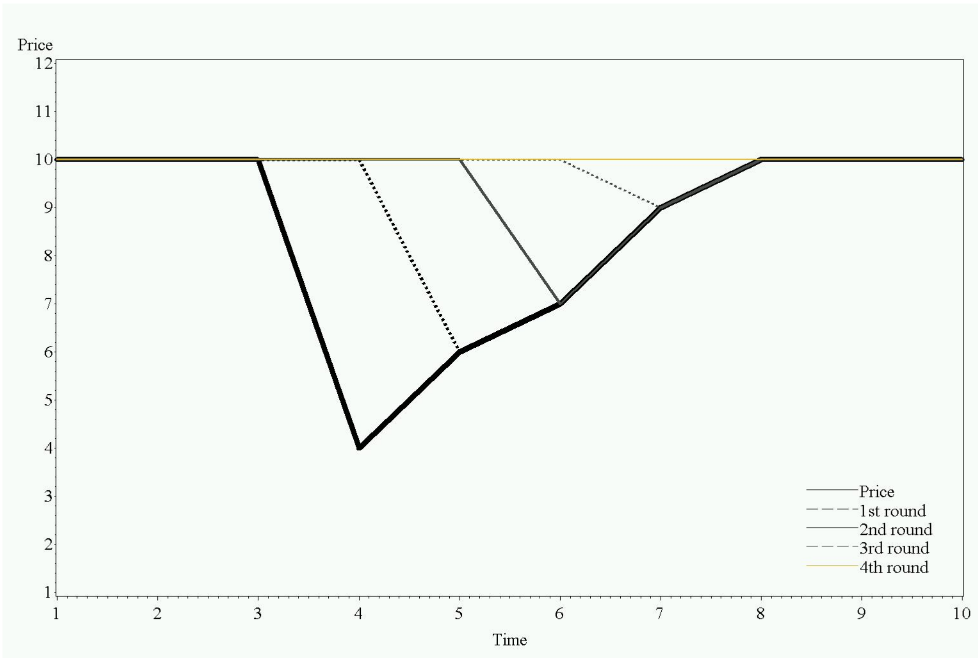
To filter temporary sales from the posted price series, we need a specific definition of what constitutes a sale before we can design a filter algorithm. We therefore have to decide on three different dimensions of a sales definition. The first aspect to be considered is the symmetry of a sale, i.e. should the price before and after the sales episode be the same or is it allowed to be higher after the sale? The second facet of the definition is the maximum number of periods that a sale can be in effect. The third aspect to consider is if a sale price should remain fixed or is allowed to vary during the sales episode. The sales filter can then be parameterized based on the choice with respect to these three dimensions of the definition.

Klenow and Kryvtsov (2008) at some point define a temporary sale quite simply as a price change in one month that is changed back to the original price the next month. To be viewed as a sale, a price decrease thus has to be followed by an identical price increase in the next period, i.e. only symmetrical V-shaped price changes lasting one period are classified as sales. Hosken and Reiffen (2004) define a sale in a similar, straightforward way as occurring if the price falls by at least some fixed percentage (e.g. 10 or 20 percent) between periods t-1 and t and then rises by at least that percentage between periods t and t+1. The difference with the definition of Klenow and Kryvtsov (2008) is that both symmetrical and asymmetrical V-shaped price changes are covered by this definition. A product is thus recorded as being on sale in month t if the prices in month t-1 and month t+1 are both significantly higher than the price charged in month t, but not necessarily the same. It counts brief discounts followed by changes in the regular price as sales. Price decreases that last more than one month are again not treated as sales.

Campbell and Eden (2010) identify a sale in a similar but slightly more restrictive fashion as a price decline of 10 percent or more in a given week that the store completely reverses within two weeks, instead of one month. All prices between the initial decline and the reversal are then branded as sale prices. The more restrictive nature of this sale classification is only possible because they have weekly instead of monthly data. Midrigan (2010) also uses weekly price quotes but defines a temporary sale in a less restrictive way as a price decrease of any size that is reversed in one of the four weeks following the original price cut. This definition not only covers V-shaped price changes but also gradual price decreases, provided these are eventually followed by a price increase after at most four weeks following the first price cut. The price is therefore allowed to vary during the sales episode. A price decrease that lasts more than four weeks is always treated as a regular price change, even if it is reversed afterwards. This filter artificially introduces a new sale in case a price cut is gradually reversed. To correctly identify and filter out all sales, the algorithm has to be repeated three times in order to eliminate sales that have been gradually implemented.

To visualize this peculiarity of the filtering process, we show in figure 1 a fictitious example of a price pattern in which a sale is gradually reversed. After the first loop of the filtering algorithm, the sale price at time 4 is replaced with the last observed regular price. We now have a sale price at time 5, which will be filtered out when the algorithm is repeated a second time. This in turn generates a new sale at time 6 and the procedure continues until each successively introduced sale is filtered out. The number of times the algorithm has to be repeated to correctly filter out all sales depends on the maximum duration of a temporary sale, in this case four periods.

Figure 1: Filtering out a gradually reversed temporary sale



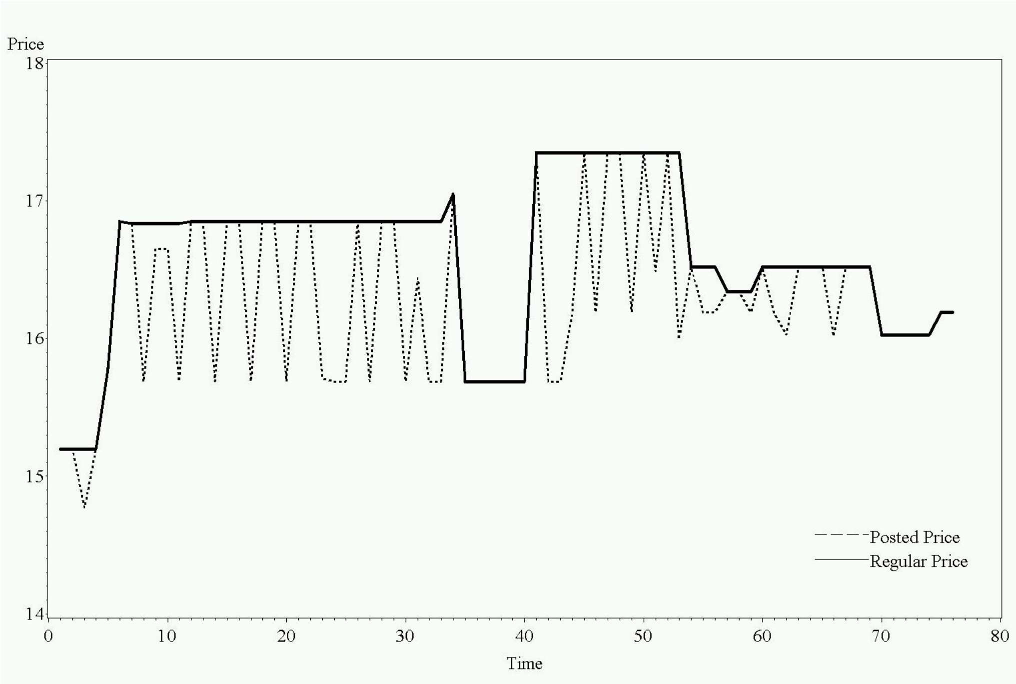

We choose to apply four different sales filters to the posted price series, all similar to the one used by Midrigan (2010). All our filters define a sale as a price decrease of any size that lasts a maximum of four weeks following the initial price cut. As a robustness check, we also tested several filters that allow sales to last six or eight weeks, but the results were not significantly different from the ones that we will present below. Apparently, the bulk of temporary sales are very short-lived. Capping off the length of a sales episode at four weeks, we then design a symmetric and an asymmetric filter for the case where we impose a fixed sales price for the duration of the sale on the one hand, and for the case where the sale price is allowed to vary during the sale on the other hand. We thus end up with four different filtered series. When presenting the results, we will refer to these four filters as *symmv4fix*, *asymmv4fix*, *symmv4flex* and *asymmv4flex*, respectively. If a price quote is defined as a temporary sale by the filter of our choosing, we replace this sale price with the last observed regular price before the sale. Intuitively, we would suggest the asymmetric filter that allows for a varying sales price, i.e. *asymmv4flex*, as the one that most effectively distills a regular price from the posted price series. The latter also most closely resembles the one of Midrigan (2010).

To visualize the effect that this sales filter has on the price trajectory of an arbitrary product, we show in figure 2 the posted and regular price series of a product in the laundry detergent category of our European scanner dataset over a time period of 76 biweeks. The dotted line gives the posted price, whereas the full line signifies the regular price. The difference between both clearly visualizes the temporary sales. In this case, the majority of sales are one-period symmetric V-shaped sales. The pattern shows that the typical retail product has a stable regular price, and deviations from that regular price are downward and short-lived (Hosken and Reiffen, 2004).

Whereas Klenow and Kryvtsov (2008) fill up gaps in their price series, we choose to work with adjacent prices only, following Nakamura and Steinsson (2008). Another difference in their analysis is the problem of voluntary and forced product substitution that is typical for CPI data [7]. Klenow and Kryvtsov (2008) include price changes that are associated with forced product turnover in their analysis, whereas Nakamura and Steinsson (2008) only compare regular prices in between item substitutions, hence excluding all substitution-related price changes. In our scanner data, we can steer clear of this discussion, as there is no voluntary product substitution due to item rotation, and forced product substitution leads to a new price trajectory attached to a different Universal Product Code.

[7]Voluntary product substitution originates from item rotation in the data set, e.g. each product could be included in the dataset for a period of five years before being replaced, or when the statistical office decides to revise its sample of products because the consumption habits have changed (Dhyne *et al.*, 2005). Forced product substitution occurs when a product is no longer sold and the government official who collects the price quotes chooses a replacement item, in many cases a newer version of the discontinued product.

Figure 2: Posted and regular price series (asymmv4flex filter)




### 3.2 Frequency vs. Duration Approach

As to make our results comparable with previous research, we will capture price rigidity by the duration of a price spell in months [8]. There are however two approaches to obtain this price stickiness parameter. The duration approach measures the duration of price spells directly from the data, whereas the frequency approach aims at calculating the frequency of price change which can then be inverted to obtain the duration of a price spell [9].

We will follow the latter approach for two reasons. First of all, gaps in price trajectories are a serious concern in the duration approach, as it is impossible to derive the spell length directly if price quotes are missing. One way to tackle this problem is to carry forward the last available price to fill in gaps, but this generates an upward bias for the duration of price spells (Baudry *et al.*, 2006; Klenow and Kryvtsov, 2008). In the frequency approach, a long and uninterrupted span of time series is not a necessary

[8] A month is defined here as a period of exactly four weeks. Our results will therefore slightly overestimate duration. This can however easily be taken into account if one wants to compare our results to the ones obtained in earlier studies that work with monthly data.

[9] See Baudry *et al.* (2006) and Fabiani *et al.* (2010), among others, for an analysis of both approaches.

condition to calculate the frequency of price change. This is a very important asset of the frequency approach in our setting, as scanner data usually contain a reasonable amount of gaps in their price series, for example due to stockouts.

Secondly, the duration approach suffers from serious censoring issues that complicate the measurement of price spell durations. The problem of censoring reflects the fact that in the process of price collection the true time of beginning/ending of the first/last price spell might not correspond to the one observed in the dataset, as it comes before/after the first/last price observation, leading to a downward bias in the average duration of a price spell when you ignore censoring (Fabiani *et al.*, 2010). Dropping the first and last price spell is not a satisfactory solution for the problem, because these spells are more likely to be long, again implying a downward bias in the average duration of a price spell. In the frequency approach on the other hand, no explicit treatment of censoring is called for, in turn allowing us to use all available information in the dataset as only observations that involve transitions from or to unobserved prices have to be dropped. Appendix II details the computation process of the different statistical measures that we apply to calculate the frequency of price change.

In order to obtain estimates for the duration of a price spell using the frequency approach, we need to invert the estimates of the frequency of price change. In this respect, it is important to note that we use data in their base frequency. As a consequence, we can calculate the duration of a price spell simply by inverting the frequency parameter, call it $\lambda$. The implied duration of a price spell therefore equals $1/\lambda$, expressed in time units corresponding to the intervals at which price quotes are available [10].

## 3.3 Parameter Choice

There are a number of statistics that can be used to capture the implied duration of a price spell. We have to take two important aspects into account when choosing a statistic. The first one is the choice between the mean or median duration. Both represent a valuable measure of price stickiness, although their value can be significantly different depending on the structure of the data.

[10] If prices can change at any time in between data points, then the instantaneous probability of a price change is $-ln(1-\lambda)$ and the implied mean time between price changes is $-1/ln(1-\lambda)$, see Bils and Klenow (2004).

As we can see in figures 3 and 4, the distribution of price spell durations in both the European and the US data is very right skewed, containing a lot of very short price spells and a long tail at the right with outliers. Consequently, the median duration will be much lower than the mean. We expect the skewness of the distribution to diminish when we gradually filter out sales, bringing the mean and median durations closer to each other. Nonetheless, in an environment with this type of uneven distribution, the median captures the true amount of rigidity better than the mean.

In figures 3 and 4, we overlay the histogram of duration with the log-normal distribution where the threshold parameter $\theta$ is assumed to be zero, and the scale and shape parameters are estimated from the data using maximum likelihood. The histogram of duration can clearly be captured by a log-normal distribution. Taking all of this into account, we therefore use the median duration as our preferred statistic. Checking the means as well, as a robustness check of our results, we found the conclusions of our analysis to remain unaffected (results available upon request).

Figure 3: Distribution of biweekly price spell duration (EU)

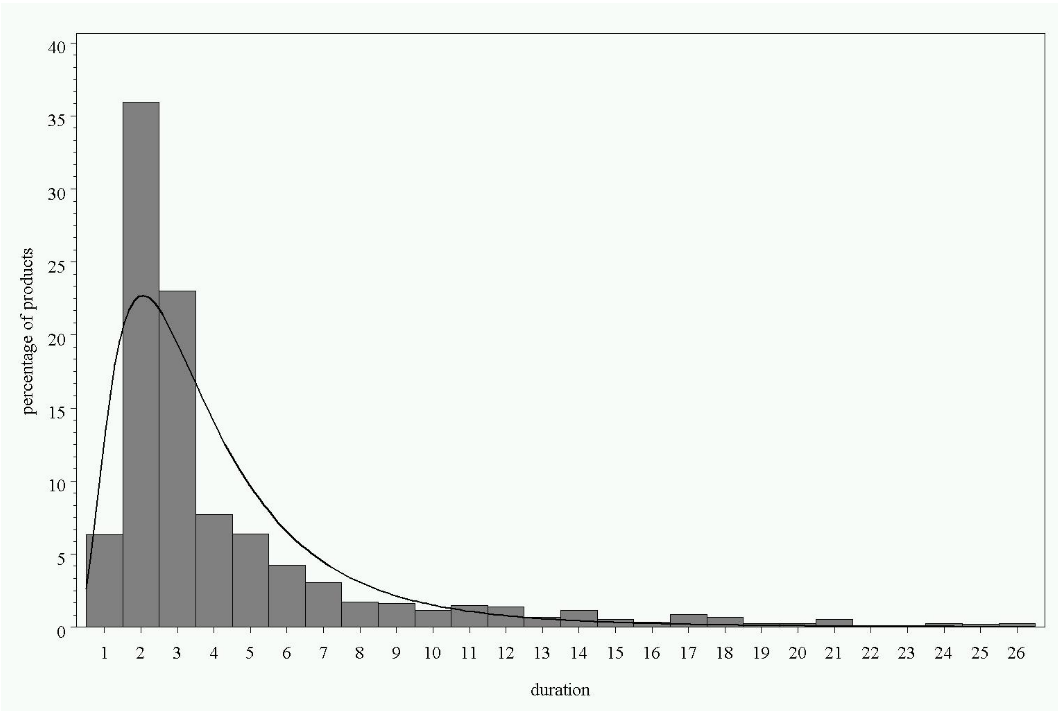
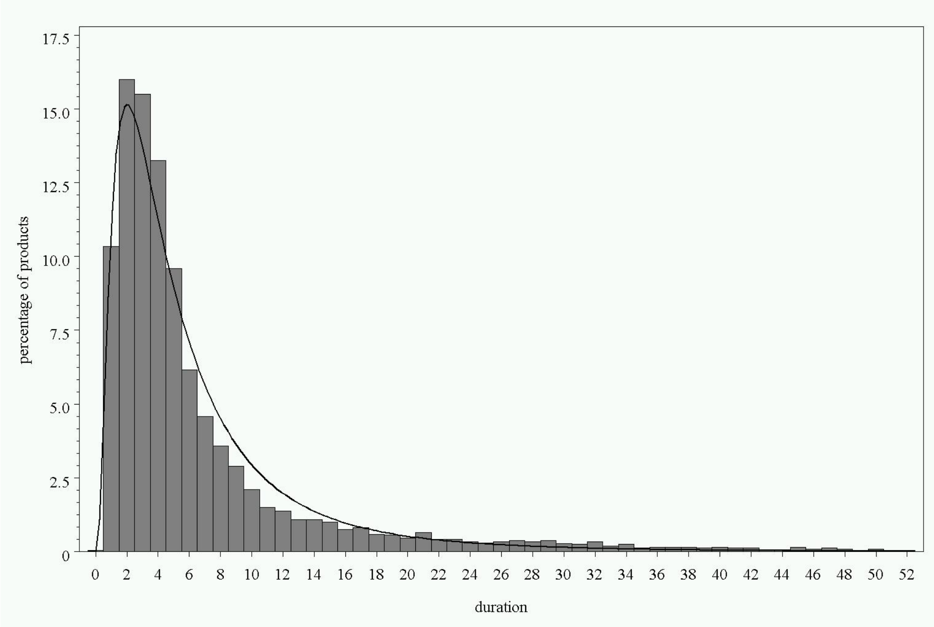


Figure 4: Distribution of weekly price spell duration (US)




A second aspect that we have to take into account is the level of aggregation at which we analyse our data. We have two workable levels of aggregation at our disposal, namely the product and product category level. At the product category level, the number of frequency parameters from which to calculate medians is rather limited, putting the soundness of the estimates at risk. We therefore propose to calculate medians at the product level. To this end, we compute the mean frequency of price change for each separate product across time and stores, take the median of all those frequency parameters and invert it to obtain the median implied duration of a price spell at the product level. Appendix II shows how we compute the mean frequency of price change for a specific product $i$ across $n$ stores and $\tau$ time periods.

The results at the product category level are calculated in the same way, computing the mean frequency of price change for each product category across time, stores and products, then taking the median over all categories and inverting it to obtain the median implied duration of a price spell at the product category level. We will present the latter results as a robustness check, because the statistics at the product level could potentially be biased as they overweigh categories which have a large number of very similar products (Eichenbaum *et al.*, 2011).

### 3.4 Results

The results in table 1 show that the posted price of the median product lasts slightly longer in our European data compared to the US data, 1.5 versus 1.3 months respectively. In other words, half of prices last less than 1.5 and 1.3 months in the European and US scanner data, respectively. In the European data, after every period of two weeks, 34 percent of products witness a price change. The US retailer, on the other hand, changes the price of 18.5 percent of their products at the weekly revision time. When we look at the median product category, the difference in price spell duration between the retailers is more pronounced, 1.6 months versus 1.0 months respectively.

Table 1: Median implied duration of a price spell in months, base frequency

| | Product level | | Category level | |
|---|---|---|---|---|
| | EU | US | EU | US |
| No filter | 1.5 | 1.3 | 1.6 | 1.0 |
| Symmv4fix | 5.0 | 2.0 | 3.1 | 1.6 |
| Asymmv4fix | 5.6 | 2.3 | 4.0 | 2.0 |
| Symmv4flex | 5.2 | 2.7 | 3.3 | 2.4 |
| Asymmv4flex | 5.7 | 3.3 | 4.5 | 3.3 |

Note: Symmv4fix filters all symmetric V-shaped sales of up to 4 weeks with a fixed sales price. Asymmv4fix filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed sales price. Symmv4flex filters all symmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price. Asymmv4flex filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price.

If we take out sales using our four preferred filters, we see that also regular prices are clearly more flexible in the US scanner data, irrespective of the strictness of the filtering method. With the most extensive method, filtering out all symmetric and asymmetric sales with a fixed or flexible sales price of up to 4 weeks, the US regular price for the median product lasts 3.3 months, whereas the median duration in Europe amounts to 5.7 months. At every biweekly price revision, only 8.8 percent of European products witness a regular price change, whereas 7.6 percent of US regular prices change at the weekly price revision. The difference in the extent of price stickiness is also present at the product category level, albeit less pronounced, 3.3 months versus 4.5 months. Using the full informational content of our pricing data, we can thus conclude that both posted and regular prices are more flexible in our US data compared to the data of the European retailer.

Compared to previous research on the extent of price stickiness, our estimates of duration are very low. The most important reason to explain the high frequency of price change is that we study retail prices. This implies that some very sticky classes of consumer goods are not included in our sample, mainly services but also apartment rents or restaurant meals for example [11]. Another factor is the type of retailer, as large supermarkets like the ones we study change their prices more swiftly compared to corner shops (Dhyne *et al.*, 2005). As the duration estimates of regular prices are closer to what has been found in the literature compared to the results for posted prices, our retailers also seem to use temporary sales more extensively compared to the average in the economy.

In table 2, we present the results of a similar analysis on monthly price series that we directly derived from our high frequency scanner data by withholding only the first price observation of each month. As such, we mimic the structure of typical monthly data sets like the one used to form the CPI. When we now look at the median duration of a price spell for posted prices, there is no significant difference between the retailers. This result holds both at the product and the category level. The conclusion from table 1 that the prices of the US retailer are more flexible than those of its European counterpart therefore collapses in this setting.

On top of that, prices in general appear to be more sticky when monthly data are studied. This is a direct consequence of the fact that a lot of posted prices have short duration, which is not picked up in monthly data. It is therefore an inherent problem of price data that are not available in their base frequency, as they do not provide direct evidence about the critical issue of how many price changes happen in between data points (Kehoe and Midrigan, 2012).

[11] Services display much more price rigidity due to lower volatility of consumer demand and input costs. The latter is due to a higher share of labour as an input factor, the cost of which is much more stable than that of intermediate goods (Bils and Klenow, 2004; Dhyne *et al.*, 2005).

Table 2: Median implied duration of a price spell in months, monthly frequency (time aggregation before filtering)

| | Product level | | Category level | |
|---|---|---|---|---|
| | EU | US | EU | US |
| No filter | 2.7 | 2.9 | 2.6 | 2.6 |
| Symmv4 | 5.3 | 4.7 | 3.6 | 4.0 |
| Asymmv4 | 5.7 | 5.2 | 4.5 | 4.7 |

Note: Symmv4fix filters all symmetric V-shaped sales of up to 4 weeks with a fixed sales price. Asymmv4fix filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed sales price. Symmv4flex filters all symmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price. Asymmv4flex filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price.

Switching our attention to regular prices, it is important to note that in this stage of the analysis, we filter sales at a monthly frequency, i.e. we first transform the high frequency data into monthly price series and filter out sales afterwards. Consequently, there is no longer a difference between fixed and flexible sales episodes, as temporary sales of up to 4 weeks last just one period when filtering sales at a monthly frequency. For the median product, we see that also regular prices display more or less the same flexibility in our European and US data, duration being only slightly higher in Europe than in the US. For the median product category, prices of the US retailer even appear to be more rigid than in the European data, albeit only marginally so, contrary to our conclusions for the data in base frequency. This clearly signifies a non-negligible amount of measurement error when using monthly data. Comparing countries or regions with respect to price stickiness based on this kind of data should therefore be performed with due caution.

Our data also offer the possibility to filter out sales at weekly/biweekly frequency and then time-aggregate the filtered weekly data into monthly observations. Midrigan (2010) uses this method to find an estimate for price stickiness that can be used to calibrate a macroeconomic model. The reason to do this is that time-aggregating first and then eliminating sales, as we have done to obtain the results of table 2, can produce spurious price changes if stores periodically put their prices on sale, at regular intervals. Consequently, the latter method is to be preferred to the one above, but take into account that this is not possible when one starts out with monthly data, say the CPI data that have been used widely to study price stickiness.

Looking at the results in table 3, we see that the duration of a regular price for the median product is lower in the US data compared to the European data, as we have found in our high frequency data, but the difference here is smaller in size. The estimate of duration for the European data using our most strict sales filter is 6.6 months, slightly higher than the one we found in table 1. The estimated duration of the median product in the US data on the other hand, is now 5.2 months, considerably higher than the 3.3 months from table 1. At the product category level, prices appear slightly more rigid in the US than in Europe, contrary to what we obtained using data in base frequency.

Table 3: Median implied duration of a price spell in months, monthly frequency (filtering before time aggregation)

| | Product level | | Category level | |
|---|---|---|---|---|
| | EU | US | EU | US |
| No filter | 2.7 | 2.9 | 2.6 | 2.6 |
| Symmv4fix | 5.9 | 3.7 | 3.8 | 3.1 |
| Asymmv4fix | 6.4 | 3.9 | 4.3 | 3.3 |
| Symmv4flex | 6.3 | 4.7 | 4.0 | 4.0 |
| Asymmv4flex | 6.6 | 5.2 | 4.5 | 4.7 |

Note: Symmv4fix filters all symmetric V-shaped sales of up to 4 weeks with a fixed sales price. Asymmv4fix filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed sales price. Symmv4flex filters all symmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price. Asymmv4flex filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price.

We can therefore conclude that the method of filtering out sales at base frequency and then time-aggregating, does a better job in calculating the frequency of price change and the duration of a price spell than the method that does the job the other way around, but that it is still a shaky basis to obtain reliable estimates upon which regional comparisons can be drawn. Apparently, the problem remains that there is a lot of information contained in the short-term price series that gets lost when the high frequency data are aggregated into monthly price series. If this lost information differs across regions, the cross-sectional comparison of price duration will be biased after aggregation over time. Our results highlight an important caveat in the recent empirical literature that studies price stickiness using micro data. If the latter are collected at a monthly frequency, a lot of price changes get lost in between data points leading to a potentially severe bias in the estimation of the frequency of price change and the duration of a price spell. A comparison between two countries or regions based on monthly data gives misleading results in those cases where there is a difference in short-term pricing tactics, mimicked in our data by the difference in price review frequency between the US and European retailer. Time-aggregation takes away a lot of relevant information to take this frequency issue into account. This caveat not only holds true for posted prices, but also for regular prices because a typical sales filter takes away a lot of variation on the downside of the regular price, but not on the upside. Consequently, retailers that change their prices more actively both upwards and downwards will appear more rigid relative to retailers with a stable price setting policy, when the data being used are not in their base frequency.

Looking at our data, the flexibility of US retail prices lies in their short-term, weekly movements. Because these are absent in the European data as a policy choice by the retailer, studying the data in base frequency leads to significantly different results compared to the same data in monthly frequency. Taking this result to the macroeconomic level, it implies that flexible short-term pricing strategies do not show up in the statistics, and countries where these are extensively used to react vividly to changing market conditions, could therefore mistakenly be classified as rigid.

A more insistent short-term pricing strategy in one region compared to another might be due to a number of reasons. First, it can be an optimizing choice on behalf of the retailer in response to more flexible factor markets. Additionally, higher perceived price elasticity of demand, for example due to more competitive product markets with lower wholesale markups, can force the price setter to adapt prices more swiftly in response to price changes at competing retailers. A third potential rationale for adapting prices more promptly could lie in a lower menu cost, for example due to a more technologically efficient, electronic price tagging system.

# 4 Frequency of Price Increases vs. Decreases

Looking at the median duration of a price spell does not give any insight into the composition of the price changes. Therefore, we break up all price changes into increases and decreases, and table 4 presents the median frequency of both. We only present these results for an analysis of the scanner data in base frequency, as we have shown this to be the only accurate basis to perform frequency estimation. Therefore, the presented US and European frequency statistics are weekly and biweekly, respectively.

Table 4: Median frequency of price increases and decreases

| | EU | | US | |
|---|---|---|---|---|
| | Up | Down | Up | Down |
| No filter | 0.174 | 0.165 | 0.098 | 0.087 |
| Symmv4fix | 0.062 | 0.042 | 0.067 | 0.055 |
| Asymmv4fix | 0.056 | 0.035 | 0.057 | 0.049 |
| Symmv4flex | 0.059 | 0.039 | 0.049 | 0.038 |
| Asymmv4flex | 0.053 | 0.032 | 0.041 | 0.031 |

Note: Symmv4fix filters all symmetric V-shaped sales of up to 4 weeks with a fixed sales price. Asymmv4fix filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed sales price. Symmv4flex filters all symmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price. Asymmv4flex filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price.

Price increases are slightly more frequent than decreases when we look at the posted, unfiltered data. This is true for both our European and US data. At the European retailer, 17.4 percent of products increase in price at the biweekly revision, whereas 16.5 percent witness a price decrease. For the US retailer, the equivalent frequencies at the weekly revision time are 9.8 percent and 8.7 percent. Interestingly, as the frequency of price change decreases by filtering out temporary sales along our four preferred dimensions, the percentage of price decreases vis-á-vis increases in the retail data falls in both the European and US data, but they nonetheless remain important for the flexibility of prices.

These results establish the prevalence of price decreases in retailing, even when temporary sales are filtered out. Hence, we find no support for downward rigidity in the price level. Price decreases are thus not just a reflection of temporary sales episodes, but a robust stylized fact in retail pricing. This confirms earlier results from the literature (Baudry *et al.*, 2006; Dhyne *et al.*, 2005).

# 5 Size of Price Changes

Besides the frequency of price change, the size of price changes is another dimension of the pricing behaviour of firms. It is not possible to give some sensible insight into the micro foundation of the inflation rate without looking at the frequency and size of price changes in conjunction. Different combinations along those two dimensions can lead to the same inflation rate, as it can be decomposed in the following way (Klenow and Kryvtsov, 2008):

$$\pi_t = \lambda_t^+ . dp_t^+ - \lambda_t^- . dp_t^-$$

with $\pi_t$ the inflation rate in period t, $\lambda_t^+$ and $dp_t^+$ the frequency and size of price increases, and $\lambda_t^-$ and $dp_t^-$ the frequency and size of price decreases. In this section, we will look into the size of price changes. Combining this analysis with the frequency of price increases and decreases from section 4 will give some insight into the micro level roots of the inflation rate.

Table 5 displays the mean and median size of price increases and decreases in our European and US scanner data. In appendix II, we show how these size parameters are computed. It is important to note that we compute the size of price increases and decreases as the difference of logarithm, so that the two successive price changes recorded during a temporary sale are equal in absolute terms (Dhyne *et al.*, 2005). A first point to note in table 5 is that posted price changes are sizable if you compare them with the low inflation rates that are prevalent in western economies during the last decades. In line with the more aggressive short-term pricing strategy with respect to the frequency of price change, we see that the US retailer, which reconsiders its prices more aptly, not only changes its prices more often, but also by more than the European retailer. Focusing on the size of a price increase for the median product, we find it to be 10.6 percent in the US data versus 7.2 percent in the European data. The same applies for the median size of a price decrease, with an even higher dispersion of 12.8 percent versus 7.2 percent. These results are in the same ballpark as the ones found in previous research on the size of price changes (Dhyne *et al.*, 2005; Nakamura and Steinsson, 2008; Klenow and Kryvtsov, 2008).

Table 5: Size of price changes

| | EU | | | | US | | | |
|---|---|---|---|---|---|---|---|---|
| | Increase | | Decrease | | Increase | | Decrease | |
| | Mean | Median | Mean | Median | Mean | Median | Mean | Median |
| No filter | 9.630 | 7.208 | 9.654 | 7.155 | 17.963 | 10.648 | 19.888 | 12.783 |
| Symmv4fix | 8.480 | 5.841 | 8.527 | 5.585 | 14.669 | 7.372 | 17.251 | 9.554 |
| Asymmv4fix | 7.344 | 4.927 | 7.790 | 4.564 | 12.460 | 6.100 | 13.639 | 6.720 |
| Symmv4flex | 8.536 | 5.859 | 8.581 | 5.580 | 14.326 | 6.761 | 17.336 | 8.952 |
| Asymmv4flex | 7.287 | 4.853 | 8.060 | 4.689 | 13.055 | 5.749 | 16.641 | 8.058 |

Note: Symmv4fix filters all symmetric V-shaped sales of up to 4 weeks with a fixed sales price. Asymmv4fix filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed sales price. Symmv4flex filters all symmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price. Asymmv4flex filters all symmetric and asymmetric V-shaped sales of up to 4 weeks with a fixed or flexible sales price.

Secondly, we notice that the median size of a price change is lower than the mean, implying that the distribution of price changes is skewed towards small price changes with a limited number of big price changes in the tail. This is especially the case for the US retailer. A third point to note when looking at both the mean and median size of posted price changes is that price increases and decreases are of the same order of magnitude in the European data, whereas price decreases are slightly larger in size than price increases in the US data.

Turning our attention to regular prices, we clearly see that the size of regular price changes is smaller than the size of posted price changes. In other words, temporary sales are larger in size than regular price changes, a feature that is common to both retailers. The difference between the mean and median size is even more pronounced for regular than for posted price changes, pointing to the fact that the bulk of regular price changes are small. This applies to both price increases and decreases. For the US retailer, the median size of regular price decreases is larger than that of regular price increases. This is most apparent when we also filter out all asymmetric sales, because large price increases that follow immediately after a temporary sales price and overshoot the previous regular price are in these cases partially filtered out, transforming a lot of large price increases into small ones. Along with the high frequency of price decreases that we showed in section 4, the relatively large size of price changes signifies an important role for large but relatively transient idiosyncratic shocks in the pricing strategy of a typical retailer (Bils and Klenow, 2004). That poses a serious problem for a lot of price rigidity models which assume that price changes only occur in response to aggregate shocks [12]. In an inflationary environment, almost all price changes should then have to be increases (Nakamura and Steinsson, 2008).

Although the mean and median price change are quite sizable, a lot of price changes are small. To visualize this, figures 5 and 6 show a histogram for the size of posted and regular price changes, respectively, in the US dataset [13]. We see in figure 5 that more than 30 percent of posted price changes is smaller than 5 percent in absolute value, and 19 percent of them is smaller than 3 percent. The histogram for posted price changes is very symmetrical, and is captured quite well by a normal distribution with an estimated mean $\mu = 0,202$ and standard deviation $\sigma = 17,708$. This pattern is mainly due to the fact that symmetrical V-shaped sales are quite common, and the price decreases and increases due to this type of temporary sales are equal in size. Figure 6 provides a histogram of the size of regular price changes in the US dataset. Small price changes are even more prevalent in the regular price series, thereby confirming that price changes attached to temporary sales are larger in size than regular price changes. Slightly less than half of regular price changes are smaller than 5 percent in absolute value, and 28 percent of them is smaller than 3 percent. The distribution of the size of regular price changes is again quite symmetrical, with an estimated mean of the overlaying normal distribution equal to 0.260 and a standard deviation of 13.834.

[12] See for example Taylor (1980), Calvo (1983), Caplin and Spulber (1987), Dotsey *et al.* (1999), Chari *et al.* (2000), Mankiw and Reis (2002), Burstein (2006) among others. Golosov and Lucas (2007), Gertler and Leahy (2008) on the other hand propose models that do incorporate a role for idiosyncratic shocks.

[13] The pattern for the size of price changes is highly similar in the European dataset (results available upon request).

Figure 5: Size distribution of posted price changes (US)

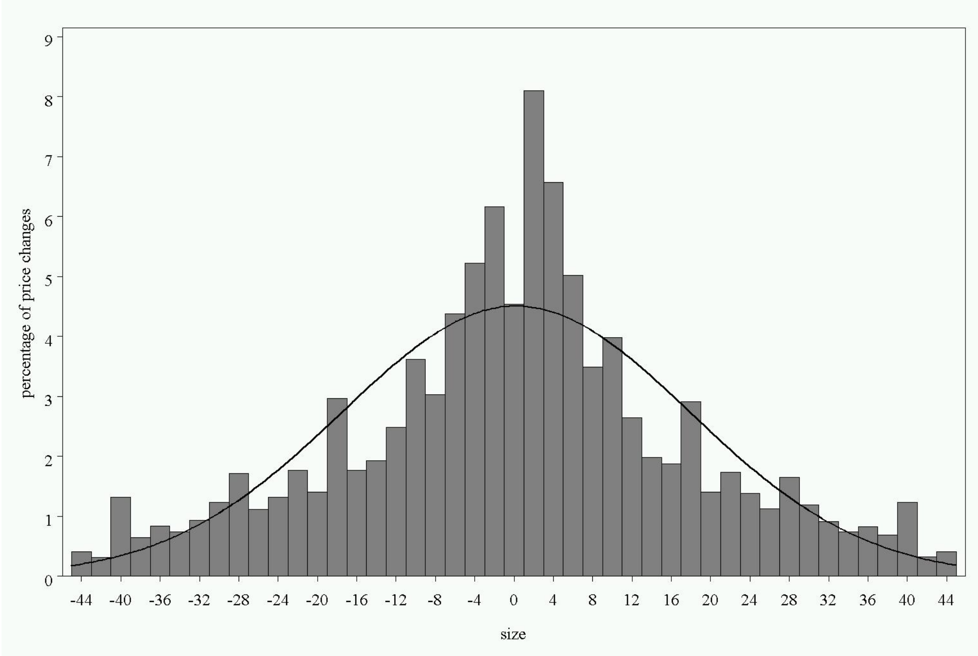
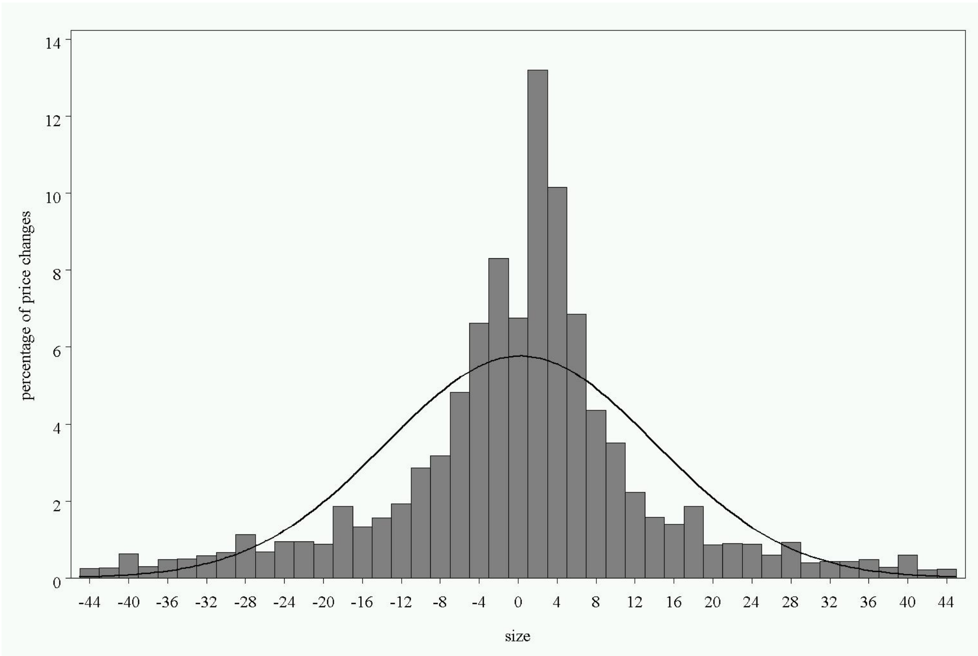



Figure 6: Size distribution of regular price changes (US)



The higher incidence of small price changes in the regular price series compared to the posted price series is also apparent when we look at the kurtosis of both distributions. The size distribution of posted price changes has a kurtosis coefficient of 3.033, which is very close to the coefficient of 3 for the normal distribution. The size distribution of regular price changes, on the other hand, is clearly leptokurtic, with a kurtosis coefficient of 4.572. The latter distribution therefore displays more weight in the vicinity of zero and the tails are less fat compared to the normal distribution and the size distribution of posted price changes. The high prevalence of small price changes in the data is hard to reconcile with a standard menu cost model (Midrigan, 2010).

# 6 Conclusion

Price rigidity is a catalyst that transforms nominal shocks into real effects, and therefore plays a vital role in macroeconomic theory and policy. This paper adds to the recent progress in empirical research by estimating price rigidity from high-frequency scanner data of two large retailers, one in Europe and one in the US. The results on the degree of price stickiness found in micro data can be highly useful to calibrate business cycle models and support the policy making process.

Recent empirical results based on CPI data show that the frequency of price change is higher in the US than in Europe, but that the difference largely disappears when regular prices are analyzed. The scanner data at our disposal have an important advantage in that they are available in their base frequency, i.e. prices do not change in between data points. Consequently, there is no source of measurement error in the data gathering process. This enables us to incorporate all short-term movements into the measured price stickiness parameters.

A comparative analysis between our scanner data in base frequency on the one hand, and data at monthly frequency that we generated from our high-frequency price series on the other hand, shows us that the use of monthly data in a regional comparison of price rigidity potentially leads to spurious conclusions. When we study the original data in base frequency, regular prices are far more flexible in the US data. This finding collapses however when we study monthly price series derived from our high frequency scanner data. Then, regular prices show the same degree of flexibility in the US and the European data. This discrepancy between the results from the base frequency data and the ones we obtain from the derived monthly data stems from a different price setting policy on behalf of the retailers. The European retailer reviews prices every two weeks, whereas the US retailer does so on a weekly basis. This salient characteristic of the data is used as a natural experiment to analyse the effects of differential short-term pricing strategies on the value of the price stickiness parameter. Our results show that the more flexible price review strategy of the US retailer gets lost in the transition of the price series from base to monthly frequency. The monthly data fail to incorporate short-term pricing information into the reported level of price stickiness, and these short-term dynamics may very well be vastly different across countries or regions.

This result is robust to the type of sales filter that we apply to obtain our regular price series, filtering symmetric and/or asymmetric V-shaped sales with a fixed or flexible sales price and a duration of the sales episode of up to four weeks. The result also holds irrespective of the statistic used to capture price rigidity, be it the mean or the median across products or product groups. The analysis of our scanner data also confirms some stylized facts often cited in the literature: sales are an important driving force of micro price flexibility, at least in retailing; the frequency of price change is extremely heterogeneous across products; price changes are on average much larger than needed to keep up with aggregate inflation, although many small price changes do occur; price decreases are not merely a reflection of sales episodes, but a robust fact in micro price setting, providing no proof of downward rigidity in retail prices.

## References


Baharad, E. and Eden, B. (2004). Price Rigidity and Price Dispersion: Evidence from Micro Data. *Review of Economic Dynamics*, **7** (3), 613–641.

Baudry, L., Le Bihan, H., Sevestre, P. and Tarrieu, S. (2006). Price Rigidity in France: Information from Consumer Price Collections. *Economie et Statistique*, **386** (1), 37–57.

Bils, M. and Klenow, P. (2004). Some Evidence on the Importance of Sticky Prices. *Journal of Political Economy*, **112** (5), 947–985.

Burstein, A. (2006). Inflation and Output Dynamics with State-Dependent Pricing Decisions. *Journal of Monetary Economics*, **53** (7), 1235–1257.

Calvo, G. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics*, **12** (3), 383–398.

Campbell, J. and Eden, B. (2010). Rigid Prices: Evidence from US Scanner Data. *Federal Reserve Bank of Chicago - Working Paper Series*, **No. 2005-08**, 45p.

Caplin, A. and Spulber, D. (1987). Menu Costs and the Neutrality of Money. *Quarterly Journal of Economics*, **102** (4), 703–726.

Cecchetti, S. (1986). The Frequency of Price Adjustment: A Study of the Newsstand Prices of Magazines. *Journal of Econometrics*, **33** (3), 255–274.

Chari, V., Kehoe, P. and McGrattan, E. (2000). Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem? *Econometrica*, **68** (5), 1151–1179.

Chevalier, J., Kashyap, A. and Rossi, P. (2003). Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data. *American Economic Review*, **93** (1), 15–37.

Christiano, L., Eichenbaum, M. and Evans, C. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of Political Economy*, **113** (1), 1–45.

Coricelli, F. and Horvath, R. (2010). Price Setting and Market Structure: An Empirical Analysis of Micro Data. *Managerial and Decision Economics*, **31** (2-3), 209–233.

Dhyne, E., Alvarez, L., Le Bihan, H., Veronese, G., Dias, D., Hoffman, J., Jonker, N., Lünneman, P., Rumler, F. and Vilmunen, J. (2005). Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data. *ECB Working Paper*, **No. 524**, 50p.

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.

Dotsey, M., King, R. and Wolman, A. (1999). State-Dependent Pricing and The General Equilibrium Dynamics of Money and Output. *Quarterly Journal of Economics*, **114** (2), 655–690.

Dutta, S., Bergen, M., Levy, D. and Venable, R. (1999). Menu Costs, Posted Prices, and Multiproduct Retailers. *Journal of Money, Credit and Banking*, **31** (4), 683–703.

Eichenbaum, M., Jaimovich, N. and Rebelo, S. (2011). Reference Prices, Costs, and Nominal Rigidities. *American Economic Review*, **101** (1), 234–262.

Fabiani, S., Gattulli, A., Veronese, G. and Sabbatini, R. (2010). Price Adjustment in Italy: Evidence from Micro Producer and Consumer Prices. *Managerial and Decision Economics*, **31** (2-3), 93–104.

Gagnon, E. (2009). Price Setting During High and Low Inflation: Evidence from Mexico. *Quarterly Journal of Economics*, **124** (3), 1221–1263.

Gali, J., Gertler, M. and Lopez-Salido, D. (2001). European Inflation Dynamics. *European Economic Review*, **45** (7), 1237–1270.

Gertler, M. and Leahy, J. (2008). A Phillips Curve with an Ss Foundation. *Journal of Political Economy*, **116** (3), 533–572.

Golosov, M. and Lucas, R. (2007). Menu Costs and Phillips Curves. *Journal of Political Economy*, **115** (2), 171–199.

Hosken, D. and Reiffen, D. (2004). Patterns of Retail Price Variation. *RAND Journal of Economics*, **35** (1), 128–146.

Karadi, P. and Reiff, A. (2007). Menu Costs and Inflation Asymmetries: Some Micro Data Evidence. *MPRA Paper*, **No. 7102**, 28p.

Kashyap, A. (1995). Sticky Prices: New Evidence from Retail Catalogs. *Quarterly Journal of Economics*, **110** (1), 245–274.

Kehoe, P. and Midrigan, V. (2012). Prices Are Sticky After All. *Federal Reserve Bank of Minneapolis - Research Department Staff Report*, **No. 413**, 41p.

Klenow, P. and Kryvtsov, O. (2008). State-Dependent or Time-Dependent Pricing: Does It Matter for Recent US Inflation? *Quarterly Journal of Economics*, **123** (3), 863–904.

— and Malin, B. (2010). Microeconomic Evidence on Price-Setting. *Prepared for the Handbook of Monetary Economics (Benjamin Friedman and Michael Woodford, editors).*

Lach, S. and Tsiddon, D. (1992). The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data. *Journal of Political Economy*, **100** (2), 349–389.

Levy, D., Bergen, M., Dutta, S. and Venable, R. (1997). The Magnitude of Menu Costs: Direct Evidence from a Large US Supermarket Chain. *Quarterly Journal of Economics*, **112** (3), 791–825.

Mankiw, G. and Reis, R. (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, **117** (4), 1295–1328.

Masahiro, H. and Yumi, S. (2007). Price Setting in Japan: Evidence from CPI Micro Data. *Bank of Japan Working Paper Series*, **No. 07-E-20**, 61p.

Midrigan, V. (2010). Menu Costs, Multi-Product Firms and Aggregate Fluctuations. *Econometrica*, **79** (4), 1139–1180.

Nakamura, E. and Steinsson, J. (2008). Five Facts about Prices: A Reevaluation of Menu Cost Models. *Quarterly Journal of Economics*, **123** (4), 1415–1464.

Peersman, G. (2004). The Transmission of Monetary Policy in the Euro Area: Are the Effects Different Across Countries? *Oxford Bulletin of Economics and Statistics*, **66** (3), 285–308.

Peltzman, S. (2000). Prices Rise Faster than They Fall. *Journal of Political Economy*, **108** (3), 466–502.

Ray, S., Chen, H., Bergen, M. and Levy, D. (2006). Asymmetric Wholesale Pricing: Theory and Evidence. *Marketing Science*, **25** (2), 131–154.

Rotemberg, J. (2005). Customer Anger at Price Increases, Changes in the Frequency of Price Adjustment and Monetary Policy. *Journal of Monetary Economics*, **52** (4), 829–852.

Taylor, J. (1980). Aggregate Dynamics and Staggered Contracts. *Journal of Political Economy*, **88** (1), 1–23.

Wulfsberg, F. (2009). Price Adjustments and Inflation - Evidence from Norwegian Consumer Price Data 1975-2004. *Norges Bank Working Paper*, **No. 2009/11**, 34p.

# Appendices

## A. Description of the datasets

Table 6: Description of the datasets

| EU | | US | |
|---|---|---|---|
| Product Group | # Products | Product Group | # Products |
| Bath soap | 182 | Bath soap | 600 |
| Bathroom tissues | 30 | Bathroom tissues | 128 |
| Canned soup | 5 | Bottled juices | 511 |
| Canned tuna | 46 | Canned soup | 445 |
| Cereals | 49 | Canned tuna | 278 |
| Coke | 39 | Cereals | 490 |
| Dish detergent | 43 | Cheeses | 667 |
| Dishwasher tablets | 43 | Crackers | 330 |
| Emmental | 56 | Dish detergent | 287 |
| Fruit juice | 54 | Grooming products | 1599 |
| Grooming products | 243 | Laundry detergent | 582 |
| Gruyere | 19 | Refrigerated juices | 228 |
| Laundry detergent | 94 | Snack crackers | 425 |
| Lemonade | 33 | Soaps | 337 |
| Smoked salmon | 18 | Soft drinks | 1746 |
| Snack crackers | 136 | Toothpastes | 608 |
| Soaps | 34 | | |
| Toothpastes | 146 | | |
| Total | 1270 | Total | 9261 |

## B. Definition of the different statistical measures

In order to calculate our statistics with respect to the stickiness of prices in Europe and the US, we create four binary variables based on the price trajectories in our scanner datasets, in a similar way as in Dhyne *et al.* (2005).

Assume that $p_{ijt}$ is the price of a product $i$, sold at store $j$, at time $t$. Because we only include adjacent price observations in our analysis, we first create an indicator variable $I_{ijt}^{obs}$, attached to $p_{ijt}$, that takes a value of 1 if a price quote for product $i$ at store $j$ from the previous period is observed and 0 otherwise:

$$I_{ijt}^{obs} = \begin{cases} 1 & \text{if } p_{ij,t-1} \text{ is observed} \\ 0 & \text{if } p_{ij,t-1} \text{ is not observed} \end{cases}$$

Studying only adjacent price observations, we will only withhold price quotes for which $I_{ijt}^{obs} = 1$ to calculate the price stickiness parameters. The second binary variable that we create is a price change indicator for product $i$ at store $j$ at time $t$:

$$I_{ijt} = \begin{cases} 1 & \text{if } p_{ijt} \neq p_{ij,t-1} \\ 0 & \text{if } p_{ijt} = p_{ij,t-1} \end{cases}$$

In the same way, we also create a price increase and price decrease indicator for product $i$ at store $j$ at time $t$:

$$I_{ijt}^{+} = \begin{cases} 1 & \text{if } p_{ijt} \succ p_{ij,t-1} \\ 0 & \text{if } p_{ijt} \leq p_{ij,t-1} \end{cases}$$

$$I_{ijt}^{-} = \begin{cases} 1 & \text{if } p_{ijt} \prec p_{ij,t-1} \\ 0 & \text{if } p_{ijt} \geq p_{ij,t-1} \end{cases}$$

Adding these four indicator variables to our datasets, calculation of the frequency of price change for product $i$ across all $n$ stores and $\tau$ time periods is straightforward:

$$\lambda_i = \frac{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}}{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}^{obs}}$$

The frequency of price increase and decrease for product $i$ across all $n$ stores and $\tau$ time periods are again calculated in a similar way:

$$\lambda_i^{+} = \frac{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}^{+}}{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}^{obs}}$$

$$\lambda_i^{-} = \frac{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}^{-}}{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}^{obs}}$$

The size of price changes, price increases and price decreases are then easily calculated as follows:

$$dp_i = \frac{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt} \; (ln \; p_{ijt} - ln \; p_{ij,t-1})}{\sum_{j=1}^{n} \sum_{t=2}^{\tau} I_{ijt}}$$

$$dp_i^+ = \frac{\sum_{j=1}^{n}\sum_{t=2}^{\tau} I_{ijt}^+ \; (ln\ p_{ijt} - ln\ p_{ij,t-1})}{\sum_{j=1}^{n}\sum_{t=2}^{\tau} I_{ijt}^+}$$

$$dp_i^- = \frac{\sum_{j=1}^{n}\sum_{t=2}^{\tau} I_{ijt}^- \; (ln\ p_{ij,t-1} - ln\ p_{ijt})}{\sum_{j=1}^{n}\sum_{t=2}^{\tau} I_{ijt}^-}$$

# CHAPTER 3

## Customer Loyalty and the Curvature of Demand:

## Evidence from Scanner Data

# Customer Loyalty and the Curvature of Demand: Evidence from Scanner Data *

Benjamin Verhelst † Dirk Van den Poel ‡

**Abstract**

This paper uses scanner data at the individual customer level, compiled from the loyalty card database of a European retailer, to estimate the price elasticity and curvature of demand for loyal and non-loyal customer segments. Drawing from Customer Relationship Management (CRM), we use segmentation techniques and cluster analysis to split up the customer base in three groups according to their behavioural loyalty to the retailer. We then perform a demand analysis on the loyal and non-loyal segments in parallel, discarding the large middle cluster. Our results from estimating a Behavioural Almost Ideal Demand System (B-AIDS) for numerous product categories reveal that loyal customers have a considerably more concave demand curve than non-loyals. This result holds true in the aggregate, and for all but some individual product categories. The more pronounced asymmetry in the price elasticity of demand for loyal customers provides a major incentive to the retailer to commit to a sticky price.

**JEL**: C33, C38, D12, L14

**Keywords**: Customer loyalty, Clustering, Curvature of demand

---

*The authors would like to thank Freddy Heylen, Gerdie Everaert, Gert Peersman, Emmanuel Dhyne, Sarah Lein and Julio Carrillo for valuable comments. We have also benefited from comments received at the 8th International Conference on Computational Management Science (Neuchatel, 2011) and the 17th International Conference on Computing in Economics and Finance (San Francisco, 2011). 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

Securing and enhancing the loyalty of their customers is a major objective for any firm that wishes to thrive in a competitive market environment. The importance of long-term firm-customer agreements or implicit contracts has been supported by survey evidence from Blinder (1991, 1994) and Blinder *et al.* (1998) for the United States, Hall *et al.* (2000) for the United Kingdom, Apel *et al.* (2005) for Sweden, and Fabiani *et al.* (2005) for the euro area. All these studies have also highlighted the primary role of implicit contracts in the price setting policy of firms. Implicit contracts systematically outweigh all other declared sources of price stickiness, including menu costs and sticky information that are traditionally the go-to sources of sticky prices in macroeconomics. However, these survey results are based on interviews and questionnaires, completed by a large sample of firms. There is no readily available and thorough analytical support for the hypothesis that implicit contracts or firms' objective to secure customer loyalty would affect the price setting policy of the firm.

In this paper, we empirically establish a link between customer loyalty and the shape of the demand curve. We are the first to make a rigorous comparative analysis of the elasticity and curvature of demand of loyal and non-loyal customer segments. The main result of this analysis establishes a considerably more concave demand curve for loyal customers compared to non-loyals. After price increases loyals reduce their demand more than non-loyals, while after price cuts loyals raise their demand less than non-loyals. Loyal customers seem to value a stable price, whereas this is less important for infrequent shoppers. This difference is a new and important result which provides a clear incentive for firms to keep prices stable. Our empirical analysis does not prove the implicit contract theory of price rigidity, but it does empirically confirm one of its building blocks as made explicit in Okun (1981).

We use point-of-sale scanner data of an anonymous European retailer with a unique availability of price and quantity data at the level of the individual customer [1]. Through the compulsory nature of the retailer's loyalty card program, we can perfectly detect for

[1] Due to a strict confidentiality agreement, we cannot disclose the identity of the retailer. This dataset has been used before by Dossche *et al.* (2010) to test the existence of the kinked demand curve, and by Verhelst and Van den Poel (2010) to stress the importance of time aggregation for measured price stickiness and regional comparisons.

every individual at any point in time which products he or she buys in what quantity and at what price. To allow for a comparison of the price elasticity and curvature of demand between loyal and non-loyal customers, the individual card holders have to be divided according to some loyalty metric. In this paper, we will focus exclusively on behavioural loyalty because supermarket scanner data are less appropriate for studying the attitudinal dimension of loyalty (Allaway *et al.*, 2006).

Using segmentation techniques and clustering analysis, we divide the customer base into three loyalty segments. Discarding the large middle cluster, a comparative demand analysis of the top and bottom segment of the customer base highlights potential loyalty-induced differences in consumer behaviour. To obtain elasticity and curvature estimates for both segments in parallel, we estimate the Behavioural Almost Ideal Demand System (B-AIDS) of Dossche *et al.* (2010) with Stone price index approximation on a large number of different product categories. This model is an extension of the standard AIDS model developed by Deaton and Muellbauer (1980), capable of fully capturing potential non-linearities in the demand curves for individual product categories. It is of the utmost importance to be able to freely estimate the curvature of demand if we want to test for loyalty-induced effects of price change.

Our result of a more concave demand curve for loyal customers provides useful evidence that supports implicit contracts as a source of price stickiness, as price setters will want to commit to a stable price in order to preserve the relations with their loyal customers. The asymmetric price elasticity of demand for repeat customers provides the retailer with ample incentive to keep prices stable in the wake of demand fluctuations, preventing the market from clearing. Sticky prices are the natural outcome of this process.

The remainder of the paper is organized as follows. Section 2 describes the scanner data of our European retailer in more detail. Section 3 introduces the behavioural loyalty concept used in this paper and presents the segmentation analysis that divides the customer base according to the level of loyalty. In section 4, we estimate a demand model and derive elasticity and curvature estimates for a large number of product categories, and perform a comparative analysis of the loyal and non-loyal customer segments. Section 5 concludes.

## 2 Data

We use point-of-sale scanner data, gathered from six stores of an anonymous euro area retailer. Whenever a product is scanned at the counter of one of these stores, the scanning device registers the purchase transaction and saves it in a transactions database. The retailer offers a very wide range of approximately 15000 stock-keeping units, covering 40% of euro area CPI. They are registered at a very detailed level through their Universal Product Code (UPC). In this project, we use daily transactions data running from January 2002 until November 2004. The data period can be divided into 76 bi-weekly periods, during which the price is kept constant as a policy choice of the retailer. In other words, every two weeks there is a price review of all the items, at which point the retailer can choose for each individual item to change its price or not. Between price reviews, all prices remain unchanged, so the retailer does not continuously adapt its prices to changes in market conditions. It is important to note that price policy is centralized, so prices and price changes are identical across the different stores of the retailer.

This dataset has been used before by Dossche *et al.* (2010) and Verhelst and Van den Poel (2010), albeit at a more aggregate level. Their results confirm a number of stylized facts of price setting: regular prices are sticky, although the frequency of price change is extremely heterogeneous across products; price decreases are quite common, even after filtering out temporary sales promotions, providing no proof of downward rigidity in retail prices; the size of price changes is large compared to aggregate inflation, although many small price changes do occur. These findings are in line with previous results on price setting in the euro area based on broader CPI-data (Dhyne *et al.*, 2005). For more detailed statistics, we refer to Dossche *et al.* (2010) and Verhelst and Van den Poel (2010).

This paper explores for the first time the individual customer dimension of the data. Through a system of compulsory loyalty cards, each purchase transaction is linked to an individual customer. Consequently, the invoice lines of the transactions database show us for each day of the considered data period who buys what product in what quantity and at what price. The full dataset contains slightly more than 1.3 million unique customers who visited one of the stores at least once during the period 2002-2004. Together, they paid approximately 72 million visits to one of the stores in the three-year period under consideration, during which they bought a total of nearly 658 million items in varying quantities. A typical item is defined at a very detailed level by its UPC. The extensive nature of the database and its level of detail offer a unique possibility to study differences in consumer behaviour according to individual characteristics, more specifically the behavioural loyalty of the customer.

To keep estimation of the demand models in section 4 manageable, we limit the scope of this dataset along the product dimension. As a starting point, we randomly select 75 broad product categories. We follow the methodology of Dossche *et al.* (2010) and include four items per product category plus a fifth composite item, which is constructed as a weighted average of all other items in the category. This fifth 'item' is introduced to fully capture substitution opportunities for the four selected items. We choose to limit the category-level demand analysis to four main items in order to keep estimation manageable, both with respect to the number of parameters to be estimated and the availability of observations. In the specific demand model that we will estimate, the expenditure share of a certain item depends on the price of the other items inside the same category. If too many items are included individually, we risk losing observations due to shorter or non-overlapping data availability.

The selection of the four top products for each category is inspired by data availability and market share requirements. We first select the items which are available for purchase in each of the six stores at least 95% of the time, i.e. the item has to be on the shelf in each store in at least 73 out of 76 bi-weekly periods. This ensures a minimal loss of observations. Among the remaining items, four are selected based on highest market share in the category. All other items are bundled in the fifth, composite item called 'other'. If there are different top products across stores, we select those with the best ranking in most stores.

Once this is done for all 75 product categories, we impose two requirements on the categories for them to be withheld in the comparative elasticity and curvature analysis. First, we require the four selected top products to jointly represent at least a 20% market share of the product category to which they belong. This should ensure that these items are pivotal and can be viewed as representative for their category. Second, we need price variation for each of the selected items in order to ensure that a demand curve can be estimated accurately. In this respect, we only select those product categories for which each of the four selected products witnesses at least one regular price change. Out of the 75 product categories that we start with, only 20 satisfy both requirements. They are listed in Appendix A, followed by the number of items in each category. The surviving 20 categories contain a total of 961 items. The number of unique customers that consume at least one item from those categories during the period 2002-2004 drops from 1.3 million to slightly less than 1 million.

Verhelst and Van den Poel (2010) show that temporary price markdowns are quite common in the dataset. It is important to take this into account when estimating elasticity and curvature parameters, because the promotional price elasticity is generally much higher than the regular price elasticity (Bijmolt *et al.*, 2005). We emphasize that every temporary price markdown of an item is accompanied by the item being mentioned in the retailer's circular. Our dataset contains an indicator variable that is equal to 1 if it is included in the circular and 0 otherwise. When estimating our demand models, this indicator variable will be very useful as a dummy variable capturing the effect of price promotion and increased visibility of the item on the elasticity and curvature of demand.

# 3 Customer segmentation

The segmentation of the customer base in homogeneous groups according to the level of loyalty is a recurring subject in Customer Relationship Management (CRM). Customer value analysis is important, as retaining a loyal customer is much cheaper than attracting a new one (Cheng and Chen, 2009). Hence, it pays off for the retailer to concentrate marketing efforts on detecting and rewarding their most loyal customers to solidify their allegiance (Allaway *et al.*, 2006). The RFM model, in which the behavioural loyalty of customers is measured by the recency, frequency and monetary value of past purchases, has been used extensively by marketeers over the past fifty years (Hughes, 2005) [2]. Once each customer is scored on the three dimensions of the model, clustering techniques can be applied to split the customer base into segments based on some distance measure over the quantitative input attributes R, F and M. This procedure is generally used to find the optimal number of clusters with as much inter-cluster heterogeneity and intra-cluster homogeneity as possible (Jonker *et al.*, 2004). It usually serves direct marketing purposes, by distinguishing a limited number of highly loyal customers that are likely to react to targeted marketing programs, hence maximizing profit for the retailer (Pauler and Dick, 2006).

[2] A good description of the RFM model can be found in Bult and Wansbeek (1995), Jonker *et al.* (2004), Pauler and Dick (2006) and Cheng and Chen (2009), among others.

The aim in our analysis is slightly different, as we simply want to separate the loyal from the non-loyal customers based on their spending behaviour in the 20 selected product categories. We therefore discard the recency coordinate, as it does not convey the loyalty of a customer over a prolonged period of time. The frequency (F) and monetary value (M) of purchases are therefore the quantitative variables that we use as yardsticks to measure store loyalty [3]. To this end, we calculate for each individual customer the number of times he or she visited the store and bought at least one item from our selected categories over the considered data period (F), and the total amount of money spent on those items during that period, in euro (M). In other words, all customers are scored on the F and M attributes, and these numeric scores serve as the input for a clustering analysis that splits the customer base into a predetermined number of homogeneous and disjoint clusters, taking into account the different scale of the F and M scores and their correlation structure.

The purpose of this procedure is to obtain a grouping in which customers are similar within the same cluster, but dissimilar to the customers in any of the other clusters (Cheng and Chen, 2009). Both attributes are given equal weight in the analysis that follows. Although there is a clear positive correlation between the frequency and monetary value of purchases in our data, with a correlation coefficient of 0.47, they do measure different dimensions of loyalty and deserve to be treated as two separate aspects of customer loyalty.

In order to measure the similarity or dissimilarity between customers with respect to F and M, we need to define a certain distance metric (Ryu and Eick, 2005). The most commonly used metric is the Euclidean distance, which is defined between two customers $i$ and $j$ as follows

$$d_E(i,j) = \sqrt{(i-j)^T(i-j)} \tag{1}$$

where in our two-variable case $i = (i_F, i_M)^T$ and $j = (j_F, j_M)^T$ and $i_F$, $i_M$, $j_F$ and $j_M$ are the scores of customers $i$ and $j$ on frequency and monetary value, respectively.

[3] We will stick to the notion of store loyalty, even though we have to keep in mind that our segmentation is based on a selection of 20 product categories from the gamma offered by the retailer. We assume that store loyalty and supermarket loyalty are equivalent, because the six stores in our dataset are located far from each other and individual customers do not visit more than one store.

However, this distance measure requires that the F and M scores are measured at the same scale, and are uncorrelated [4]. As both requirements are violated in our setting, we work with Mahalanobis distance instead, which is scale-invariant and allows for positive correlation between F and M. It is defined between two customers $i$ and $j$ as follows

$$d_M(i,j) = \sqrt{(i-j)^T \Sigma^{-1} (i-j)} \tag{2}$$

where $\Sigma$ is the within-cluster covariance matrix and all other variables are defined as before. As the composition of the clusters is not known beforehand, we estimate $\Sigma$ from the initial F and M values of each customer using the methodology of Art *et al.* (1982).

The clustering procedure itself is in fact an optimization problem. The eventual partitioning of the customers is based on the least-squares criterion, i.e. the minimization of the sum of square-error for all customers in the database, with the errors defined as the Mahalanobis distances between each customer and the cluster centers

$$E = \sum_{c=1}^{k} \sum_{i \in C_c} |i - m_c|^2 \tag{3}$$

where $k$ is the number of clusters, $i$ is the point in the F-M plane representing individual customer $i$, and $m_c$ is the mean of cluster $C_c$. To minimize (3), the algorithm will assign each customer to the cluster with the closest mean.

We choose to apply the K-means clustering procedure of MacQueen (1967), in which each cluster is represented by the center of the cluster. The number of clusters $k$ needs to be chosen in advance, here we fix it at three. By forming three clusters and discarding the middle one, we can highlight the potential differences in elasticity and curvature between the loyal and non-loyal segment, without worrying about customers that are

[4] The different scale of F and M implies that the variance of the monetary variable is much larger than the variance of the frequency variable, therefore M has more effect on the resulting clusters than F if the distances are not normalized. The positive correlation between the F and M attributes implies that customers $i$ and $j$ are distributed around their cluster center in a non-spherical manner, so that allocation of the customers towards the different clusters should take into account not only the distance to the cluster center, but also the direction.

somewhere in between and could potentially add a lot of noise to the comparison. The k-means clustering procedure with $k = 3$ will produce exactly three clusters of the greatest possible distinction, as compact and as detached as possible (Pauler and Dick, 2006).

There are two major steps when performing this method, the assignment step and the reestimation step (Wu *et al.*, 2009). In the first step, the algorithm chooses three customers as a first guess of the initial cluster centers. Then, all other customers are assigned to the cluster that minimizes the Mahalanobis distance between the customer under consideration and the cluster mean over the coordinates F and M. The clusters that we obtain at this stage are only temporary. Once all customers are assigned, the cluster means are recalculated during the second step and the algorithm repeats by examining each customer again and placing it in the cluster with the closest mean. This iterative process continues until there is no longer a reassignment of customers among the different clusters.

We have to take into account that the results of the k-means clustering procedure are sensitive to the presence of outliers in the data (van der Laan *et al.*, 2003). When we look at our customer base, we notice that there are a limited number of heavy buyers. Both the choice of initial cluster centers in the assignment step and the iterative formation of the clusters in the reestimation step will be impacted by these outliers. We will now explain how we adapt the standard k-means clustering procedure to deal with both sources of outlier distortion in the final clusters.

The initialization method of the algorithm is designed to find reasonably good clusters during the assignment step, i.e. even before any reestimation occurs. As such, outliers have a higher probability to be chosen as an initial cluster center. To avoid this inherent pitfall of the procedure, we perform a preliminary cluster analysis with 100 clusters. In the next step, each cluster containing less than 20 customers is deleted, and the centers of the remaining clusters are used as input seeds in an assignment step with $k = 3$. In this way, outliers are excluded as potential cluster centers. The rationale is that outliers will be far away from most other customers with respect to F and M, so that they end up in a low-frequency cluster during the preliminary cluster analysis. Once the cluster centers are set, the excluded outliers are reintroduced into the dataset before the reestimation step starts.

Although the influence of severe outliers on the initial cluster centers is dealt with by the deletion of low frequency clusters in the preliminary analysis, they still threaten to distort the natural formation of the clusters during the multiple iterations of the k-means algorithm by pulling the cluster centers towards them. To avoid that, we resort to the $L_p$ clustering criterion of Späth (1985) with maximum reduction of outlier effects. This method does not minimize the mean square difference between customers and their respective cluster means, as in the standard case of k-means clustering without outlier correction, but the mean absolute difference between customers and their respective cluster medians. Instead of equation (3), the partitioning criterion therefore becomes

$$E = \sum_{c=1}^{k} \sum_{i \in C_c} |i - m_c| \tag{4}$$

with $m_c$ now defined as the median of cluster $C_c$ and all other variables as defined before. This criterion introduces a weighting scheme in the algorithm that favors customers close to the cluster centers when recomputing the latter. In other words, it minimizes the influence of outliers on the reestimation of the cluster centers.

Table 1 contains some descriptive statistics of the three clusters that we obtain from the adapted k-means clustering procedure described above and compares them with those of the joint dataset [5]. As can be seen in the table, we obtain a non-loyal cluster containing approximately 80% and a loyal cluster with less than 4% of the customers [6]. The remaining 16% end up in the middle cluster and will be discarded for the demand analysis of section 4. Although the loyal customers make up only a small part of the customer base, they are in fact responsible for one third of total expenditure at the six stores under consideration, only slightly less than the much larger non-loyal segment. The cluster median of the frequency variable for the non-loyal and loyal customer groups are 5 and 51 times, respectively. For the monetary value of the purchases, the respective cluster medians are 48 and 1564 euro.

[5]This analysis is based on the limited dataset containing the 20 product categories that we select based on market share and price variation requirements, see section 4.2 for more details.

[6]This outcome is in line with the finding of Allaway *et al.* (2006) that only a small percentage of loyalty card users demonstrate behaviour that can be considered truly loyal.

Table 1: Descriptive statistics of clusters

| | Joint | Loyal | Neutral | Non-loyal |
|---|---|---|---|---|
| # of customers | 998086 | 35212 | 161512 | 801362 |
| Median frequency | 5.00 | 50.83 | 23.61 | 4.80 |
| Median monetary value | 73.38 | 1564.47 | 495.85 | 47.66 |

# 4 Demand analysis

## 4.1 The model

Now that the customers are segmented, we can test for loyalty-induced differences in their spending behaviour through a comparative demand analysis of the loyal and non-loyal clusters. Two important aspects of the analysis are the elasticity and the curvature of demand. The elasticity measures the responsiveness of the quantity demanded of a good to a change in its price, whereas the curvature measures the price elasticity of the price elasticity, i.e. the sensitivity of the price elasticity of demand for a good to a change in its relative price (Dossche *et al.*, 2010). The curvature is indicative for the change in the slope of the demand curve at different levels of the relative price.

Differences in elasticity and curvature of demand between loyal and non-loyal customers are a key building block of the implicit contract theory for price rigidity of Okun (1981). This theory conjectures that consumption of loyal shoppers is more elastic at high relative prices, because these repeat customers are very responsive to price levels that exceed what they experienced previously. Random buyers react less strongly to high relative prices because they only observe the current price and therefore have no terms of reference. On the other hand, consumption of the random shoppers is more elastic at low relative price levels because they are typically bargain hunters that search actively for low prices. Repeat customers were ready to buy at the previous price, so their reaction to a price reduction is bound to be low. Through a stable pricing policy, the firm can discourage its regular customers from shopping elsewhere, whereas these customers can minimize their shopping costs. In other words, the firm commits to an implicit agreement with its regular customers not to increase prices when market conditions are tight (Nakamura and Steinsson, 2011). In exchange, the firm does not decrease its prices when demand is weak, because this would only attract less interesting bargain hunters, who will shop elsewhere when demand picks up and the price returns to its original level.

In order to derive the elasticity and curvature parameters for both segments, we resort to the Behavioural Almost Ideal Demand System (B-AIDS) of Dossche *et al.* (2010). This is an extension of the workhorse AIDS model of Deaton and Muellbauer (1980), capturing the potential non-linearity in the demand curve, and allowing for flexible estimation of the required parameters. The standard AIDS model is based on the PIGLOG class of consumer preferences, it permits exact aggregation over consumers, and provides a local first-order approximation to any true demand system (Deaton and Muellbauer, 1980). Although the AIDS model is perfectly suited for an analysis of supermarket scanner data, and is known for its flexibility, transparency and ease of estimation, the standard version of the model does not allow for free estimation of the curvature parameter (Dossche *et al.*, 2010). The behavioural extension of the AIDS model offers a solution to this shortcoming by introducing a quadratic effect of the relative price of a good on top of the usual direct price effects. The basic specification of the B-AIDS model in budget share form is the following,

$$s_i = \alpha_i + \sum_{j=1}^{N} \gamma_{ij} \ln p_j + \beta_i \ln\left(\frac{X}{P}\right) + \sum_{j=1}^{N} \delta_{ij} \left(\ln\left(\frac{p_j}{P}\right)\right)^2 \tag{5}$$

where $i$=1,...,N is the number of items included in the demand system, $s_i$=$(p_i q_i)/X$ is the expenditure share of item $i$ in the product category, $p_j$ is the price of item $j$, $X$ is total category expenditure, and $P$ is the translog price index for this particular product category. It is defined by Deaton and Muellbauer (1980) as

$$\ln P = \alpha_0 + \sum_{j=1}^{N} \alpha_j \ln p_j + \frac{1}{2} \sum_{j=1}^{N} \sum_{i=1}^{N} \gamma_{ij} \ln p_i \ln p_j \tag{6}$$

The behavioural extension of the model concerns the quadratic term in the relative price at the right-hand side of equation (5). The original AIDS model has $\delta_{ij}$=0, and as we demonstrate below, the curvature parameter then becomes a restrictive function of the price elasticity. Hence, the original AIDS does not allow us to freely estimate the curvature parameter, whereas the added quadratic term in equation (5) allows to capture both concave and convex asymmetry in consumer reaction to price changes (Dossche *et al.*, 2010). The extended equation (5) is a valid representation of consumer preferences as long as the standard adding-up ($\sum_{i=1}^{N} \alpha_i = 1, \sum_{i=1}^{N} \gamma_{ij} = 0, \sum_{i=1}^{N} \beta_i = 0, \sum_{i=1}^{N} \delta_{ij} = 0$), homogeneity ($\sum_{j=1}^{N} \gamma_{ij} = 0$), and symmetry ($\gamma_{ij} = \gamma_{ji}$) restrictions hold (Dossche *et al.*, 2010). For ease of computation, Deaton and Muellbauer (1980) propose an approximation of the nonlinear translog price index given in (6) by the Stone price index

$$\ln P^* = \sum_{i=1}^{N} s_i \ln p_i \tag{7}$$

where $s_i$ is again the expenditure share of good $i$ in the product category, and $p_i$ is the price of item $i$. The use of the Stone index as an approximation to the true nonlinear price index makes the demand model linear in the parameters and facilitates its estimation. The linear approximation is highly accurate when the different price series of the demand model are collinear (Alston *et al.*, 1994). We estimate our demand models at the product category level, based on item price series inside relatively narrow product categories. These can be expected to show considerable collinearity, making the linear approximation very accurate in our context.

Based on the coefficients of the linear approximate B-AIDS demand model, Dossche *et al.* (2010) show how to derive the elasticity and curvature parameters of the demand model. The computation of the price elasticity of demand is based on the approximation approach of Alston *et al.* (1994) and Buse (1994), which they show to be superior to many other methods using Monte Carlo simulations. This approach leads to the following expression for the positive uncompensated price elasticity of demand

$$\varepsilon_i = -\frac{\partial \ln q_i}{\partial \ln p_i} = 1 - \frac{\gamma_{ii}}{s_i} + \beta_i - \frac{2\delta_{ii} \ln(p_i/P^*)}{s_i} + 2\sum_{j=1}^{N} \delta_{ij} \ln\left(\frac{p_j}{P^*}\right) \tag{8}$$

where $q_i{=}s_i X/p_i$. This expression for the elasticity in the B-AIDS model clearly incorporates nonlinear effects of the relative price, which were introduced by the quadratic term in equation (5). Since $s_i$ is typically far below 1, observing $\delta_{ii} < 0$ will most likely imply a concave demand curve, with $\varepsilon_i$ rising in the relative price $p_i/P^*$. When $\delta_{ii} > 0$, the demand curve will more likely be convex. At steady state, all relative prices are equal to 1, and the expression for the elasticity of demand is identical to the one from the standard AIDS model

$$\varepsilon_i = 1 - \frac{\gamma_{ii}}{s_i} + \beta_i \tag{9}$$

Logarithmic derivation of equation (8) with respect to price $p_i$ gives the following implied curvature parameter at steady state [7]

$$\epsilon_i = \frac{\partial \ln \varepsilon_i}{\partial \ln p_i} = \frac{1}{\varepsilon_i}\left[(\varepsilon_i - 1)(\varepsilon_i - 1 - \beta_i) - \frac{2\delta_{ii}(1-s_i)}{s_i} + 2\left(\delta_{ii} - s_i \sum_{j=1}^{N} \delta_{ij}\right)\right] \tag{10}$$

The impact of $\delta_{ii}$ on the curvature of demand is negative for reasonable values of $s_i$, i.e. the lower $\delta_{ii}$, the higher the curvature of demand, ceteris paribus. The implied curvature parameter is still positively correlated with the price elasticity of demand, albeit much less restrictively than in the standard AIDS model, in which $\delta_{ii} = \delta_{ij} = 0$.

## 4.2 Identification and Estimation

The invoice lines in the loyal and non-loyal daily transactions databases that we created in the segmentation analysis are transformed into datasets with a bi-weekly frequency containing the price and total quantity sold for all items. We do this in parallel for the loyal and non-loyal datasets. Selecting the appropriate product codes, these datasets in bi-weekly frequency are split up in 20 category-specific datasets each, which are then transformed in order to contain price and quantity data for the four top products and the composite good 'other'. This is the category-specific data structure that we use as input for the demand analysis. We therefore choose to estimate separate demand models at the product category level, limiting the risk of aggregation bias (Fisher *et al.*, 2001). Consumer preferences are assumed to be weakly separable, i.e. consumption decisions in one category are independent from price changes in other categories. Intra-category allocation of expenditure is made without reference to outside prices (Baltas, 2002).

The empirical demand specification that we use is directly derived from the B-AIDS demand equation (5) and resembles the expression used by Dossche *et al.* (2010):

$$\begin{aligned} s_{imt} =& \alpha_{im} + \sum_{j=1}^{5} \gamma_{ijm} \ln p_{jt} + \beta_{im} \ln\left(\frac{X_{mt}}{P^*_{mt}}\right) + \sum_{j=1}^{5} \delta_{ijm} \left(\ln\left(\frac{p_{jt}}{P^*_{mt}}\right)\right)^2 \\ &+ \sum_{j=1}^{5} \varphi_{ijm} C_{jt} + \tau_{it} + \lambda_{it} + \nu_{imt} \end{aligned} \tag{11}$$

[7]See appendix B of Dossche *et al.* (2010) for the mathematical derivation.

where $i$=1, ..., 5 is the item identifier, $m$=1, ..., 6 is the store identifier and $t$=1, ..., 76 is the time subscript defined in bi-weekly periods. $s_{imt}$ is the expenditure share of item $i$ at store $m$ and time $t$. $p_{jt}$ is the price of item $j$ at time $t$ common to all stores, $X_{mt}$ is total category expenditure at store $m$ and time $t$, and $P^*_{mt}$ is the store-specific Stone price index at time $t$. The quadratic term in the relative price captures potential non-linearities in the demand curve. The circular dummy $C_{jt}$ is equal to one when item $j$ is advertised and on display in the retailer's circular at time $t$. There is no subscript $m$ as the circular is common to all stores. The time trend variable $\tau_{it}$ captures long-term shifts in expenditure share of one item relative to the other items in the category. Three separate holiday dummies $\lambda_{it}$ are included for Easter, Christmas and New Year, which should capture shifts in spending behaviour linked to holiday festivities. The time trend and the dummies will capture broad demand shocks that are common across all stores. The intercept $\alpha_{im}$ captures item- and store-specific fixed effects, and controls for fixed product item characteristics and structural heterogeneity in consumer preferences across stores, for example due to a different regional demographic structure [8].

We estimate a separate demand model for each category-store combination. Given that we select four items per category and bundle all other items in a fifth, composite item 'other', this would imply that we estimate six store-specific systems of five equations for each product category. However, when estimating each system of equations, we have to drop one of the five equations in order to avoid singularity of the contemporaneous variance-covariance matrix of the disturbances (Buse, 1994). We drop the equation for 'other'. Its parameters can be recovered using the adding-up conditions ($\sum_{i=1}^{5} \alpha_{im} = 1, \sum_{i=1}^{5} \gamma_{ijm} = 0, \sum_{i=1}^{5} \beta_{im} = 0, \sum_{i=1}^{5} \delta_{ijm} = 0, \sum_{i=1}^{5} \varphi_{ijm} = 0$). We therefore estimate six store-specific systems of four equations, where we impose homogeneity ($\sum_{j=1}^{5} \gamma_{ijm} = 0$) and symmetry ($\gamma_{ijm} = \gamma_{jim}$) restrictions upon each system. Symmetry is also imposed on the effects of the circular dummies ($\varphi_{ijm} = \varphi_{jim}$). Each SUR estimation will thus result in four elasticity and four curvature parameters for each individual store, and for the loyal and non-loyal segment in parallel. Consequently, we obtain 24 estimated elasticity and curvature parameters in total for each product category and customer segment [9].

[8] To control for item-specific fixed effects, we have also demeaned $ln(p_{jt}/P^*_{mt})$ before introducing it into the quadratic term in the regression (Dossche *et al.*, 2010).

[9] The elasticity and curvature parameters of the composite item 'other' are not withheld in the comparative analysis due to its continuously changing composition.

Studying store level data requires aggregation over individual consumers. Although this could wash out interesting differences between individual shoppers, it tends to average out individual stochastic behaviour, reducing noise in the dependent variable (Hoch *et al.*, 1995). We are only interested in potential differences in spending behaviour between the loyal and non-loyal segment of the customer base, hence the loss of information at the individual level is not problematic in our context.

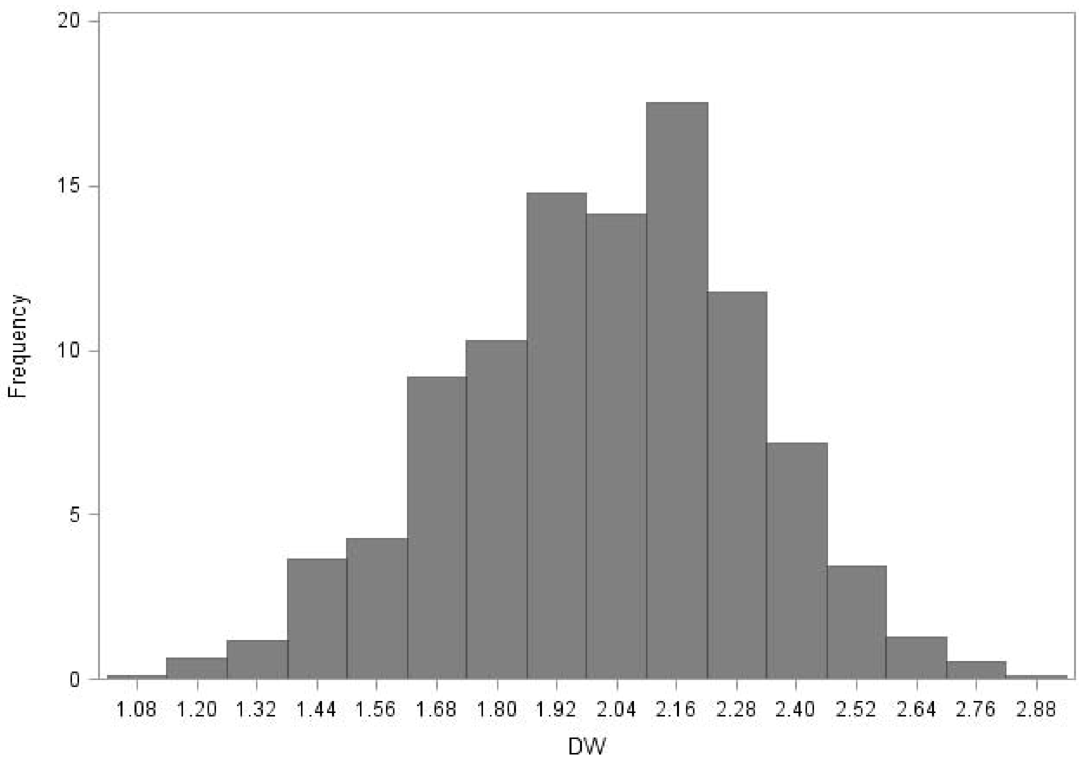
The estimation methodology that we apply is Seemingly Unrelated Regression (SUR). At first sight, estimating a price-quantity relationship offers the classic example of an endogeneity problem, and resorting to IV estimation techniques would seem to be the most straightforward way of dealing with this problem. Nonetheless, we believe that there are no identification problems for the demand curve in our setting due to two specific characteristics of the data that prevent prices $p_{it}$ to be correlated with the error term $\nu_{imt}$ (Dossche *et al.*, 2010). First of all, prices are set at the beginning of each bi-weekly period and remain unchanged for at least two weeks. Hence, the retailer introduces a source of nominal price rigidity in its price series, and does not continuously change its price to equilibrate supply and demand. Prices $p_{it}$ are then predetermined with respect to equation (11), and should therefore not be correlated with the contemporaneous error term $\nu_{imt}$.

The second data characteristic that avoids correlation between the price and the error term and enhances identification of the demand curve, is the chain-wide price setting policy of the retailer. Every two weeks, all prices are reviewed across all store locations. When they decide to change the price of a certain item $i$, they do so for every store simultaneously. There is no reason to suspect that chain-wide prices $p_{it}$ would be correlated with the store-specific error term $\nu_{imt}$. Specific demand shocks at the store-level that end up in the error term will not have enough weight to incite a price change at the chain-wide level.

One could of course argue that forward-looking retailers will take into account anticipated demand fluctuations when they decide which price they will charge for their products in the current period. This could potentially lead to correlation between the price and the contemporaneous error term. However, due to the price setting policy of our retailer, only anticipated demand shocks at the chain-wide level will be incorporated in their pricing decisions. Those should be captured by the circular and holiday dummies in the regressions, and hence not show up in the error term. The same argument applies to time-invariant item-specific characteristics, which will be captured in the regressions by the fixed effect $\alpha_{im}$.

Another argument against SUR arises if the retailer reacts to past demand shocks when setting the current price, for example by charging a higher price in the current period in reaction to a stockout in the previous period. However, this only leads to correlation between the price and the error term if demand shocks are correlated over time. Autocorrelation in the error term in our regressions is very weak however. So this condition does not seem to be fulfilled. In figure 1, we present the distribution of the Durbin-Watson statistic across all our 960 regressions (20 product categories, 4 top items per category, 6 stores, and all of this in parallel for the loyal and non-loyal customer segments).

Figure 1: Distribution of Durbin-Watson statistic



The distribution clearly supports the null hypothesis of no autocorrelation in the error term for the majority of the regressions. The median Durbin-Watson statistic across all regressions is 2.03, with a standard deviation of 0.29. The lowest value of the Durbin-Watson statistic is 1.06, whereas the highest is 2.88. The first quartile is 1.81 and the third quartile 2.22. The inter-quartile range therefore equals 0.41, implying that half of the regressions have a Durbin-Watson statistic between 1.81 and 2.22. This analysis proves that autocorrelation is largely absent from the data, so that current demand shocks are hard to predict based on past observations. Therefore, the retailer will not be able to exploit this in its price setting strategy.

Taking all of these arguments into account, we can safely assume that endogeneity is not a threat to our results, and SUR is an appropriate estimation method. And even when the specific characteristics of our data were to be ignored, Buse (1994) shows that the modest advantage in bias of an IV estimator is more than offset by its larger variance. Dossche *et al.* (2010) re-estimate their B-AIDS model using an IV method as a robustness check and the 3SLS estimates for the elasticities and curvatures that they obtain are very similar to their SUR estimates.

## 4.3 Results

Each category-specific demand model is estimated at the store level, and for loyal and non-loyal datasets in parallel, giving us 240 estimated demand models and 960 derived elasticity and curvature parameters, evenly split between the loyal and non-loyal customer segments. These parameters can be compared in the aggregate, and on a category-by-category basis to detect potential loyalty-induced differences in consumption behaviour.

Table 2 shows the price elasticity of demand by product category for the loyal customers in the second column and the non-loyal customers in the third column. The category- and segment-specific elasticity parameters are obtained by taking the median across the four top products of the category and across the six stores, i.e. the median across the 24 individual product-store elasticities for each category and customer segment. Standard errors are calculated by dividing the standard deviation across the 24 individual parameters by the square root of 24. They are given in parentheses below the respective parameter values. The price elasticity of demand in both the loyal and the non-loyal segment is positive for all categories, and it is significantly different from zero at the 5%

level for 16 out of 20 categories. The median price elasticity across the 20 categories, given at the bottom of the table, is 1.52 for the loyal customer segment and 1.68 for the non-loyals. Standard errors for these aggregate statistics are calculated by dividing the standard deviation across the 20 category-specific elasticity parameters by the square root of 20. Clearly, the elasticity is highly significant in the aggregate for both customer segments.

To examine the difference between loyal and non-loyal customers with respect to the elasticity of demand, we provide a difference measure in the fourth column of table 2. Each category-specific value is computed by taking the median across the 24 pairwise differences in the product- and store-specific elasticities between loyals and non-loyals. Standard errors are computed by dividing the standard deviation across the 24 pairwise differences by the square root of 24. An aggregate difference is again provided at the bottom of the table by taking the median across the 20 category-specific differences. The difference in price elasticity of demand between loyals and non-loyals is insignificant in the aggregate and for 14 out of 20 categories. Out of the six remaining difference measures, loyals have a significantly lower elasticity than non-loyals for five categories, whereas it is the other way around in only one case. The latter therefore react on average slightly more to price changes, although the difference is seldom significant.

Table 2 provides weak evidence for a loyalty-induced difference in the price elasticity of demand, and shows extensive heterogeneity in the elasticity parameters across product categories. These results are not new, and they originate in the combination of choice and quantity decisions. Loyal customers appear to be less price sensitive in the choice decision but more price sensitive in the quantity decision (Krishnamurthi and Raj, 1991). They buy certain items in any case, but adapt the quantity according to the current price level. Non-loyal customers on the other hand will only be persuaded to buy at the store when the price is low enough, so they either buy or they don't. The use of aggregate data masks the choice and quantity dimensions, making it hard to predict differences in overall price elasticities.

We switch our attention now to the curvature of demand, which is the main part of our analysis. Nonetheless, equation (10) shows that the curvature parameter $\epsilon_i$ is positively correlated with the elasticity parameter $\varepsilon_i$, so we have to take the preceding elasticity analysis into account when we compare the shape of the demand curve across segments.

Table 2: Price elasticity of demand ($\varepsilon_i$)

| Product Category | Loyals | Non-loyals | Difference |
|---|---|---|---|
| Baking flour | 0.94* | 0.97*** | 0.16 |
| | (0.56) | (0.33) | (0.33) |
| Chips | 1.53*** | 1.38*** | -0.01 |
| | (0.44) | (0.30) | (0.46) |
| Coke | 2.41*** | 1.38*** | 0.69* |
| | (0.30) | (0.26) | (0.36) |
| Detergent | 0.18 | 0.84 | -0.68 |
| | (0.62) | (0.62) | (0.67) |
| Emmental | 2.21*** | 4.64*** | -2.18*** |
| | (0.43) | (0.87) | (0.66) |
| Floorcloth | 4.58*** | 4.42*** | 0.96 |
| | (0.76) | (0.55) | (0.83) |
| Fruit Juice | 0.50** | 0.81*** | -0.12 |
| | (0.22) | (0.27) | (0.34) |
| Lemonade | 1.35*** | 1.21*** | 0.07 |
| | (0.25) | (0.35) | (0.29) |
| Margarine | 1.50*** | 2.28*** | -1.06*** |
| | (0.31) | (0.30) | (0.27) |
| Mayonnaise | 1.04*** | 0.14 | 0.67*** |
| | (0.24) | (0.17) | (0.22) |
| Mineral Water | 2.08*** | 1.91*** | 0.10 |
| | (0.24) | (0.23) | (0.29) |
| Nappies | 2.54*** | 3.69*** | -0.58 |
| | (0.88) | (0.62) | (1.01) |
| Plasters | 0.75 | 0.04 | 0.78 |
| | (0.48) | (0.42) | (0.64) |
| Potatoes | 0.79*** | 1.17*** | -0.36** |
| | (0.11) | (0.18) | (0.17) |
| Smoked Salmon | 3.32*** | 4.46*** | -1.04*** |
| | (0.59) | (0.60) | (0.35) |
| Sugar | 0.92 | 0.67 | 0.29 |
| | (0.80) | (0.94) | (0.85) |
| Toilet paper | 2.89*** | 2.96*** | 0.25 |
| | (0.59) | (0.74) | (0.54) |
| Tuna | 1.50** | 1.68*** | -0.58 |
| | (0.65) | (0.48) | (0.61) |
| Whiskey | 2.41*** | 2.56*** | -0.39 |
| | (0.53) | (0.64) | (0.61) |
| Wine | 2.11*** | 5.81*** | -3.21*** |
| | (0.71) | (0.60) | (0.65) |
| MEDIAN | 1.52*** | 1.68*** | -0.07 |
| | (0.24) | (0.37) | (0.25) |

*Note*: Loyal and non-loyal elasticity parameters in columns 2 and 3 are computed as the median across the 24 individual product-store elasticities in each category. The difference in column 4 is computed as the median across the 24 pairwise differences in the elasticities between loyals and non-loyals in each category. * p<0.10, ** p<0.05, *** p<0.01; standard errors between parentheses.

The comparison of the curvature parameter between loyal and non-loyal customers is indicative for a potential discrepancy in the way consumers react to price changes. Comparing both loyalty segments shows us if frequent shoppers indeed value a stable price more than non-loyal customers, which in turn supports the idea of implicit contracts as a source of price stickiness in retailing. Table 3 shows the curvature of demand by product category for the loyal segment in the second column and for the non-loyals in column 3. The way in which the category-specific curvature parameters and their respective standard errors are calculated is very similar to the way we obtained the elasticity parameters in table 2. The curvature of demand in the loyal customer segment is positive and significant at the 5% level for 17 out of 20 product categories, whereas the non-loyal segment displays significantly positive curvature for 12 out of 20 categories. The category-specific parameters translate into a median curvature across the 20 categories of 5.08 for the loyal customer segment and 2.13 for the non-loyals.

We investigate the difference between the loyal and non-loyal customer segments by providing a difference measure in column 4 of table 3. Similar to the elasticity analysis, we calculated this measure by taking the median across the 24 pairwise differences in the product- and store-specific curvatures between loyals and non-loyals. We find that the curvature of demand is significantly higher for loyals compared to non-loyals in 11 out of 20 product categories, whereas none of the other categories display a significantly higher curvature parameter for non-loyals. The aggregate difference, obtained by taking the median across the 20 category-specific differences, amounts to 2.38 and is highly significant. Hence, we find strong evidence that the loyal customer segment has a more concave demand curve.

This result supports the role of implicit contracts in the price setting process. Taking into account that it is five times cheaper to keep a loyal customer compared to adding a new one, retailers will not want to risk antagonizing their loyal shoppers (Cheng and Chen, 2009). A situation ensues in which both the buyer and the seller have an interest in stable prices. This is the perfect breeding ground for a bilateral commitment to an implicit contract [10].

[10]Nakamura and Steinsson (2011) elaborate on this issue in a context where consumers are subject to internal deep habit formation. Firms then commit to an implicit contract/sticky price to manage customer's expectations about future prices.

Table 3: Curvature of demand ($\epsilon_i$)

| Product Category | Loyals | Non-loyals | Difference |
|---|---|---|---|
| Baking flour | 4.78*** | 1.64 | 3.12*** |
| | (1.17) | (1.07) | (0.99) |
| Chips | 5.48*** | 1.73** | 3.34*** |
| | (0.71) | (0.75) | (0.84) |
| Coke | 5.53*** | 2.04*** | 3.29*** |
| | (0.90) | (0.75) | (0.94) |
| Detergent | -4.62*** | -3.04*** | -1.45 |
| | (1.12) | (0.94) | (1.07) |
| Emmental | 6.29*** | 4.49*** | 2.27*** |
| | (0.62) | (0.64) | (0.84) |
| Floorcloth | 7.64*** | 6.48*** | 1.24 |
| | (0.70) | (1.02) | (0.92) |
| Fruit Juice | 0.79 | 0.33 | 0.17 |
| | (0.86) | (0.82) | (0.69) |
| Lemonade | 3.09*** | -0.40 | 3.16*** |
| | (0.86) | (0.96) | (0.77) |
| Margarine | 6.77*** | 4.65*** | 2.51*** |
| | (0.72) | (0.96) | (0.72) |
| Mayonnaise | 6.57*** | 4.11*** | 2.49** |
| | (0.97) | (1.42) | (1.24) |
| Mineral Water | 3.22*** | -0.32 | 2.71*** |
| | (0.62) | (0.83) | (0.89) |
| Nappies | 8.44*** | 6.83*** | 0.62 |
| | (2.08) | (1.23) | (1.22) |
| Plasters | 2.91** | -2.08 | 4.78*** |
| | (1.27) | (1.40) | (1.71) |
| Potatoes | 0.33 | 0.12 | 0.29 |
| | (0.81) | (0.68) | (0.54) |
| Smoked Salmon | 1.41** | 2.22** | -0.86 |
| | (0.63) | (0.90) | (0.72) |
| Sugar | 7.98*** | 1.23 | 6.11*** |
| | (1.54) | (1.52) | (1.34) |
| Toilet paper | 3.42*** | 3.77*** | -0.44 |
| | (0.61) | (1.22) | (0.90) |
| Tuna | 4.29*** | 3.60*** | 1.03 |
| | (1.20) | (1.05) | (0.91) |
| Whiskey | 8.47*** | 5.80*** | 3.01*** |
| | (1.44) | (1.32) | (1.08) |
| Wine | 5.38*** | 6.52*** | -1.34 |
| | (1.43) | (0.88) | (1.21) |
| MEDIAN | 5.08*** | 2.13*** | 2.38*** |
| | (0.73) | (0.67) | (0.56) |

*Note*: Loyal and non-loyal curvature parameters in columns 2 and 3 are computed as the median across the 24 individual product-store curvatures in each category. The difference in column 4 is computed as the median across the 24 pairwise differences in the curvatures between loyals and non-loyals in each category. * p<0.10, ** p<0.05, *** p<0.01; standard errors between parentheses.

### 4.4 Robustness

In this section, we check the robustness of our results with respect to some alternative model specifications, and provide some additional evidence in support of a loyalty-induced effect on the curvature of demand. We test two distinct alternative specifications of the expenditure share equation by replacing the Stone price index by a different reference price in the quadratic term of equation (5). First, instead of looking at the effect of the relative price of the items inside its category on the expenditure share $s_i$ of item $i$, we check the influence of price changes by imposing the price of the item in the previous period as the reference price instead of $P^*$ in the quadratic term of equation (5)

$$s_{it} = \alpha_i + \sum_{j=1}^{N} \gamma_{ij} \ln p_{jt} + \beta_i \ln \left( \frac{X_t}{P_t} \right) + \sum_{j=1}^{N} \delta_{ij} \left( \ln \left( \frac{p_{jt}}{p_{j,t-1}} \right) \right)^2 \tag{12}$$

Besides the relative price or the change in price with respect to the previous period, customers can also react to deviations of the price from a certain regular price level. To test this specification, we replace $P^*$ in the quadratic term of equation (5) by the price that is most common in the 12-month period leading up to time $t$

$$s_i = \alpha_i + \sum_{j=1}^{N} \gamma_{ij} \ln p_j + \beta_i \ln \left( \frac{X}{P} \right) + \sum_{j=1}^{N} \delta_{ij} \left( \ln \left( \frac{p_j}{p_{j,reg}} \right) \right)^2 \tag{13}$$

Based on the coefficients of these alternative expenditure share specifications, the elasticity and curvature parameters can be calculated using expressions (9) and (10). Their median values across our 20 product categories are presented in table 4, together with the baseline results from the main analysis based on equation (5). Although there are some minor changes in the values of the elasticity and curvature parameters at the category-level, the cross-category results in table 4 confirm the key conclusions from section 4.3. The price elasticity of demand is slightly lower for the loyal customers, albeit not significantly so. More importantly, loyals have a significantly more concave demand curve than the non-loyal customers. It appears that all specifications pick up the same basic characteristic of consumer behaviour that loyal customers value a stable and fair price. They react more negatively to high or increasing relative price levels and less positively to low or decreasing relative price levels than non-loyals.

Table 4: Elasticity and curvature parameters of alternative demand specifications

| | $\varepsilon_L$ | $\varepsilon_{NL}$ | $\epsilon_L$ | $\epsilon_{NL}$ |
|---|---|---|---|---|
| Equation (5) | 1.52 | 1.68 | 5.08 | 2.13 |
| Equation (12) | 1.34 | 1.51 | 4.81 | 2.20 |
| Equation (13) | 1.65 | 1.77 | 4.90 | 2.28 |

*Note*: $\varepsilon_L$ and $\varepsilon_{NL}$ are the elasticity parameters, and $\epsilon_L$ and $\epsilon_{NL}$ are the curvature parameters for the loyal and non-loyal segments, respectively.

The comparative analysis of the curvature parameter supports a loyalty-induced effect of the relative price on the elasticity of demand. The B-AIDS model is able to incorporate this behavioural aspect of consumption through the presence of the quadratic term in the expenditure share equation (5). The value of the parameter $\delta_{ii}$ from the regression features prominently in the expression for the curvature of demand (10) and directly translates into the value of the curvature parameter. Not surprisingly, the mean value of the 480 (20×4×6) estimates of $\delta_{ii}$ in our sample is almost three times larger, in absolute value, for the loyal compared to the non-loyal segment, $-0.76$ versus $-0.27$. Table 5 gives the percentage of significant estimates of $\delta_{ii}$ at different significance levels for the loyal and non-loyal segments. Irrespective of the significance level, $\delta_{ii}$ is significant for a lot more products when we consider loyal compared to non-loyal customers. For loyal shoppers, $\delta_{ii}$ is significant at the 5% level more often than not, whereas this is the case for less than one out of three products when we consider the non-loyal customer segment. This confirms our main result of a loyalty-induced effect of the relative price level on the elasticity of demand, and the high percentages reported in table 5 lend firm support to the behavioural extension of the AIDS model as defined in equation (5).

Table 5: Significance of $\delta_{ii}$

| | Loyal | Non-loyal |
|---|---|---|
| 10% | 61% | 33% |
| 5% | 55% | 29% |
| 1% | 39% | 23% |

*Note*: The results for each customer segment are computed as the number of significant estimates of $\delta_{ii}$ divided by the total number of estimates, i.e. 480 (20 categories × 4 products × 6 stores).

## 5 Conclusion

Using supermarket scanner data from numerous product categories, we have estimated the elasticity and curvature of demand for loyal and non-loyal customer segments. We have shown analytically that loyal customers have a more concave demand curve than non-loyal customers, where we define loyalty in a behavioural way based on the frequency and monetary value of purchases over an extended period of time. The more pronounced concavity of demand shows up in the aggregate, and for all but some individual product categories. It implies that loyal customers react more negatively to price increases, and less positively to price decreases than non-loyals. The preference for a stable price is increasing with the behavioural loyalty of the customer.

This new and important stylized fact is in line with the theoretical underpinning of Okun's (1981) implicit contract theory of price rigidity. He postulates that repeat buyers are very responsive to prices that exceed the price level that they experienced in the previous period. Random shoppers only observe the current price and are less able to relate this price level to previous experiences. Their reaction is bound to be less strong. When the relative price level decreases, random shoppers will increase their consumption by more than the loyal customers. The former are more likely to be bargain hunters and search actively for a low price, whereas the latter were ready to buy at the previous price, so their reaction to the price reduction is bound to be low.

These findings provide ample incentive to the retailer to commit to an implicit contract and a stable price. Volatile prices on behalf of the retailer risk to turn off high-value regular customers and hurt long-run profits. Sticky prices are a natural outcome of this process, as retailers will try to keep their prices as stable as possible in order to preserve the trust of their clientele. This empirical result supports the conclusion of survey evidence that implicit contracts are an important source of price stickiness.

## References


ALLAWAY, A., GOONER, R., BERKOWITZ, D. and DAVIS, L. (2006). Deriving and Exploring Behavior Segments within a Retail Loyalty Card Program. *European Journal of Marketing*, **40** (11/12), 1317–1339.

ALSTON, J., FOSTER, K. and GREEN, R. (1994). Estimating Elasticities with the Linear Approximate Almost Ideal Demand System: Some Monte Carlo Results. *Review of Economics and Statistics*, **76** (2), 351–356.

APEL, M., FRIBERG, R. and HALLSTEN, K. (2005). Microfoundations of Macroeconomic Price Adjustment: Survey Evidence from Swedish Firms. *Journal of Money, Credit, and Banking*, **37** (2), 313–338.

ART, D., GNANADESIKAN, R. and KETTENRING, R. (1982). Data-based Metrics for Cluster Analysis. *Utilitas Mathematica*, **21** (A), 75–99.

BALTAS, G. (2002). An Applied Analysis of Brand Demand Structure. *Applied Economics*, **34** (9), 1171–1175.

BIJMOLT, T., VAN HEERDE, H. and PIETERS, R. (2005). New Empirical Generalizations on the Determinants of Price Elasticity. *Journal of Marketing Research*, **42** (2), 141–156.

BLINDER, A. (1991). Why are Prices Sticky? Preliminary Results from an Interview Study. *American Economic Review*, **81** (2), 89–96.

— (1994). On Sticky Prices: Academic Theories Meet the Real World. In G. Mankiw (ed.), *Monetary Policy, NBER Studies in Business Cycles*, Chicago: The University of Chicago Press, pp. 117–154.

—, CANETTI, E., LEBOW, D. and RUDD, J. (1998). Asking About Prices: A New Approach to Understanding Price Stickiness. *Russell Sage Foundation*, 380p.

BULT, J. and WANSBEEK, T. (1995). Optimal Selection for Direct Mail. *Marketing Science*, **14** (4), 378–394.

BUSE, A. (1994). Evaluating the Linearized Almost Ideal Demand System. *American Journal of Agricultural Economics*, **76** (4), 781–793.

Cheng, C. and Chen, Y. (2009). Classifying the Segmentation of Customer Value via RFM Model and RS Theory. *Expert Systems with Applications*, **36** (3), 4176–4184.

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

Dhyne, E., Alvarez, L., Le Bihan, H., Veronese, G., Dias, D., Hoffman, J., Jonker, N., Lünneman, P., Rumler, F. and Vilmunen, J. (2005). Price Setting in the Euro Area: Some Stylized Facts from Individual Consumer Price Data. *ECB Working Paper*, **No. 524**, 50p.

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.

Fabiani, S., Druant, M., Hernando, I., Kwapil, C., Landau, B., Loupias, C., Martins, F., Mathä, T., Sabbatini, R., Stahl, H. and Stockman, A. (2005). The Pricing Behaviour of Firms in the Euro Area: New Survey Evidence. *ECB Working Paper*, **No. 535**, 47p.

Fisher, D., Fleissig, A. and Serletis, A. (2001). An Empirical Comparison of Flexible Demand System Functional Forms. *Journal of Applied Econometrics*, **16** (1), 59–80.

Hall, S., Walsh, M. and Yates, A. (2000). Are UK Companies' Prices Sticky? *Oxford Economic Papers*, **52** (3), 425–446.

Hoch, S., Kim, B., Montgomery, A. and Rossi, P. (1995). Determinants of Store-Level Price Elasticity. *Journal of Marketing Research*, **32** (1), 17–29.

Hughes, A. (2005). Strategic Database Marketing. *New York: McGraw-Hill Publishing Company*, 445p.

Jonker, J., Piersma, N. and den Poel, D. V. (2004). Joint Optimization of Customer Segmentation and Marketing Policy to Maximize Long-term Profitability. *Expert Systems with Applications*, **27** (2), 159–168.

Krishnamurthi, L. and Raj, S. (1991). An Empirical Analysis of the Relationship between Brand Loyalty and Consumer Price Elasticity. *Marketing Science*, **10** (2), 172–183.

MacQueen, J. (1967). Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*, **1** (14), 281–297.

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

Okun, A. (1981). Prices and Quantities: A Macroeconomic Analysis. *The Brookings Institution*, 371p.

Pauler, G. and Dick, A. (2006). Maximizing Profit of a Food Retailing Chain by Targeting and Promoting Valuable Customers using Loyalty Card and Scanner Data. *European Journal of Operational Research*, **174** (2), 1260–1280.

Ryu, T. and Eick, C. (2005). A Database Clustering Methodology and Tool. *Information Sciences*, **171** (1-3), 29–59.

Späth, H. (1985). Cluster Dissection and Analysis. *Chichester, England: Ellis Horwood*, 226p.

van der Laan, M., Pollard, K. and Bryan, J. (2003). A New Partitioning Around Medoids Algorithm. *Journal of Statistical Computation and Simulation*, **73** (8), 575–584.

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.

Wu, H., Chang, E. and Lo, C. (2009). Applying RFM Model and K-Means Method in Customer Value Analysis of an Outfitter. *The 16th ISPE International Conference on Concurrent Engineering*, 665–672.

## Appendix A: Selected product categories

| Product Category | # Items | Product Category | # Items |
|---|---|---|---|
| Baking flour | 21 | Mineral water | 70 |
| Chips | 138 | Nappies | 65 |
| Cola | 42 | Plaster | 33 |
| Detergent | 43 | Potatoes | 26 |
| Emmental | 58 | Smoked salmon | 19 |
| Floorcloth | 11 | Sugar | 19 |
| Fruit juice | 57 | Toilet paper | 39 |
| Lemonade | 37 | Tuna | 72 |
| Margarine | 67 | Whiskey | 82 |
| Mayonnaise | 45 | Wine | 17 |

*Note*: A typical item is defined at a very detailed level by its Universal Product Code (UPC). One UPC will for example be connected to a package of 20 scrolls of white toilet paper of a certain brand.

# CHAPTER 4

## Deep Habits in Consumption:

## A Spatial Panel Analysis using Scanner Data

# 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.

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

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. $\mu_i$ is the spatial fixed effect that captures time-invariant characteristics of a specific zip code area, $\tau_t$ is the time fixed effect absorbing any time-specific shocks to consumption common to all spatial units, and $\eta_{it}$ is an iid error term with zero mean and variance $\sigma_\eta^2$ [4].

The cross-sectional specific effects $\mu_i$ 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 $\tau_t$. 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 $\mu_i$. 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 $\eta$ 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 $\omega_i$ as the eigenvalues of W (Anselin *et al.*, 2008). The time fixed effect $\tau_t$ can be concentrated out by taking the partial derivative of (2) with respect to $\tau_t$:

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

Equation (3) can be solved for $\tau_t$:

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

This solution for $\tau_t$ 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 $\mu_i$ can in turn be concentrated out by taking the partial derivative of (5) with respect to $\mu_i$:

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

and equation (7) can be solved for $\mu_i$:

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

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

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

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 $\gamma$, $\beta$ and $\sigma^2$ can be expressed as a function of the sole unknown parameter $\rho$. A unique, numerical solution for $\rho$ is obtained by maximizing the concentrated log-likelihood function

$$LogL = C - \frac{NT}{2} log \left[ (e_0^* - \rho e_1^*)'(e_0^* - \rho e_1^*) \right] + T log \left| I_N - \rho W \right| \tag{11}$$

where $C$ is a constant not depending on $\rho$, 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 $\rho$, the other parameters $\gamma$, $\beta$ and $\sigma^2$ 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 $N/T^3 \rightarrow 0$ and $T/N^3 \rightarrow 0$, 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 $\gamma$ and the external habit parameter $\rho$ 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 $\rho$ 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 |
|---|---|---|---|---|
| $\gamma$ | 0.269*** | 0.200*** | 0.116*** | 0.018** |
| | (0.023) | (0.041) | (0.016) | (0.007) |
| $\rho$ | 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 $\gamma$ and $\rho$ 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 $\rho$ 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, $\gamma$ increases from 0.200 to 0.325. At the product category and individual product level on the other hand, $\gamma$ 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 |
|---|---|---|---|---|
| $\gamma$ | 0.265*** | 0.325*** | 0.010 | -0.086*** |
| | (0.041) | (0.062) | (0.022) | (0.017) |
| $\rho$ | 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 $\rho$ is positive and significant for 33 out of 68 categories, in line with the results of section 3.3. However, the internal habit parameter $\gamma$ 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:

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

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 |
|---|---|---|---|---|
| $\rho$ | 0.456*** | 0.279*** | 0.085*** | 0.097*** |
| | (0.029) | (0.072) | (0.015) | (0.016) |
| $\gamma_1$ | 0.229*** | 0.128*** | 0.111*** | 0.017 |
| | (0.025) | (0.031) | (0.016) | (0.011) |
| $\gamma_2$ | 0.089*** | 0.070*** | -0.007 | -0.018** |
| | (0.026) | (0.024) | (0.009) | (0.008) |
| $\gamma_3$ | 0.039 | 0.040** | -0.017** | -0.027*** |
| | (0.026) | (0.018) | (0.007) | (0.007) |
| $\gamma_4$ | 0.035 | 0.021** | -0.015** | -0.028*** |
| | (0.026) | (0.011) | (0.007) | (0.007) |
| $\gamma_5$ | -0.058** | 0.076*** | -0.011 | -0.019*** |
| | (0.026) | (0.021) | (0.007) | (0.007) |
| $\gamma_6$ | -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 $\gamma$ and $\rho$ 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 |
|---|---|---|---|---|
| $\gamma$ | 0.266*** | 0.208*** | 0.123*** | 0.021** |
| | (0.026) | (0.042) | (0.018) | (0.010) |
| $\rho$ | 0.453*** | 0.332*** | 0.101*** | 0.128*** |
| | (0.028) | (0.082) | (0.021) | (0.019) |
| $\delta$ | 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 | | | |
|---|---|---|---|---|---|---|---|---|
| | $\gamma$ | | $\rho$ | | $\gamma$ | | $\rho$ | |
| 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 | | | |
|---|---|---|---|---|---|---|---|---|
| | $\gamma$ | | $\rho$ | | $\gamma$ | | $\rho$ | |
| 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