# An Empirical Investigation of Store Brands and their Role to Mitigate Brand Manufacturer Price Increases

*Click here for the latest version*

Marco J. W. Kotschedoff *

This draft: November 11, 2018

**Abstract**

Using individual coffee purchase data, this paper analyzes how retailers can use their store brands to mitigate the effect of brand manufacturer wholesale price increases. The empirical analysis exploits an asymmetric rise in wholesale prices for store brands and national brands to reveal consumers' substitution patterns. Combining the estimated consumer preferences with a structural model of retail competition allows to measure changes in retailers' unobserved marginal costs and margins. Multi-brand retailers can increase their category profits by 2-10% if they re-adjust margins after the asymmetric rise in wholesale prices and divert more demand towards their store brands. Another finding is that the positioning of a retailer's store brand dampens the increase in wholesale prices. A store brand that is perceived as a close substitute (cross-price elasticity approaches one) dampens the increase in wholesale prices for national brands by approximately 16%, on average, compared to a fully differentiated store brand. This finding provides evidence that "me-too" store brands work as partial insurance against upstream market structure shocks, favoring store brand positioning close to the leading national brand.

**Keywords:** Store brands, Category management, Pass-through, Retail competition

**JEL classification:** D4, L11, L41, L81, M31

*Johann Wolfgang Goethe University, Theodor-W.-Adorno-Platz 3, 60323 Frankfurt, Germany; kotschedoff@econ.uni-frankfurt.de

I thank my supervisor Roman Inderst for continuous support and guidance. The paper has benefited from conversations with Ajay K. Kohli, Thomas Otter, Fiona M. Scott Morton, Max Pachali and Christine Zulehner. It received valuable comments from audiences at Goethe University Frankfurt, MaCCI 2018, EARIE 2018, EMAC Research Camp 2018. Any errors are my own.

# 1 Introduction

Retailers design and introduce store brands with two conflicting objectives: market segmentation versus improving their bargaining position against national brand manufacturers. Positioning a store brand at a different location in product space than prominent national brands enables retailers to increase category profits by segmenting the market for instance between price-sensitive consumers and brand-name loyal consumers. On the other hand, retailers can enhance their bargaining power in wholesale price negotiations with manufacturers when they position their store brands close to national brands. Thus, retailers with such "me-too" store brands get better wholesale terms compared to other non-imitating retailers.

This paper contributes to the store brand positioning discussion by providing a further argument for retailers to imitate the leading national brands: insuring against upstream market structure shocks. This involves answering the following two questions: First, how should a retailer react to an upstream market structure shock on a subset of brands in its product category? Second, how does the positioning of store brands help to lower the wholesale price increase resulting from an upstream market structure shock?

I am able to estimate consumers' substitution patterns between store brands and national brands for substantial price differences that are rarely observed in practice. In particular, I exploit a unique source of price variation, due to a series of coordinated wholesale price increases by major ground coffee brand manufacturers. The cross-price effects between national brands and store brands are based on revealed preferences and proxy retailers' store brand positioning. I combine the estimated consumer preferences with a structural model of retail competition to extract changes in retailers' unobserved marginal costs that are due to higher wholesale prices. This allows me to answer both questions. First, I examine why retailers should strategically alter their prices and margins to divert more demand towards their own store brands after the increase in wholesale prices on national brands. Second, I show that the positioning of a retailer's store brand dampens the increase in wholesale prices.

The analysis is based on consumer panel data between 2004 and 2005 that track German households' coffee purchases. I use the substantial price variation to estimate a discrete choice demand model allowing for consumer preference heterogeneity. Due to the suddenness of the manufacturers' wholesale price increases and the relatively short sample period, the characteristics and the perception of coffee brands remain unchanged. Quality improvements of store brands or global income and wealth shocks are thus ruled out as explanations for store brands having higher market shares in 2005.[1] I employ this setting as a quasi-controlled environment to study how retailers should react to higher wholesale prices on national brand and how the positioning of retailers' store brand dampens the increase in wholesale prices. Multi-brand retailers can increase their category profits by 2-10% if they re-adjust margins after the asymmetric rise in wholesale prices and divert more demand towards their store brands. A store brand that is perceived as a close substitute (cross-price elasticity approaches one) dampens the increase in wholesale prices for national brands by approximately 16%, on average, compared to a fully differentiated store brand. This provides evidence that "me-too" store brands work as partial insurance against upstream market structure shocks such as upstream cartel formation, merger waves or tacit collusion. It favors a store brand positioning close to the leading national brand if market factors facilitate both explicit and as well as tacit collusion. Another relevant factor is if competition authorities cannot or do not effectively limit market power increases of upstream mergers.

The rest of this paper is organized as follows. The relevant literature is reviewed in Section 2. Section 3 presents the data and details on the German coffee market. Section 4 sets up the demand model and explains the identification strategy. The demand estimation results are shown in Section 5. An examination of how the positioning of retailers' store brand dampens the increase in wholesale prices follows in Section 6. Finally, I draw conclusions in Section 7. Additional material is contained in an Appendix.

## 2 Related literature

Considerable work has already been done to study the strategic role of store brands for retailers. Among the first were Raju, Sethuraman, and Dhar (1995), who identify three conditions under which store band introduction raises category profits: low cross-price sensitivity among national brands, high cross-price sensitivity between a national brand and store brand, a large number of national brands in that category. Once a retailer decides to introduce a store brand, however, it faces the following store brand positioning problem.

One channel to increase category profits with store brands is market segmentation. Thus, a retailer should position a store brand rather away from the national brand. The typical suggestion for store brand positioning would be to introduce a perceived "low-quality" store brand. In fact, most store brands have a lower perceived quality than the corresponding national brands. The

[1] Lamey, Deleersnyder, Dekimpe, and Steenkamp (2007) and more recently Dubé, Hitsch, and Rossi (2018) show that country-wide negative income and wealth shocks, for instance during a recession, increase demand for store brands.

general notion is if category sales are expandable by market segmentation, a retailer may be better off positioning the store brand away from the national brand to target the price-sensitive market segment while the brand-manufactures focus on the advertising/quality-sensitive market segment. This type of market segmentation is also common in pharmaceutical markets where entry by generics occurs after patent expiry to further segment the market.[2] Returning to the store brand positioning problem, Sayman, Hoch, and Raju (2002), identify conditions when store brands should imitate the leading national brand and when store brands should target a different consumer segment. Similarly, Gabrielsen and Sørgard (2007) examine why low-quality private labels are introduced in some product categories and not in others. Other studies also support the differentiation argument by investigating how (differentiated) store brand increase store loyalty (e.g. Corstjens and Lal (2000); Ailawadi, Pauwels, and Steenkamp (2008); Seenivasan, Sudhir, and Talukdar (2016)).

The other channel how store brands increase retailers' category profits is the bargaining argument. Retailers can enhance their bargaining power in wholesale price negotiations with manufacturers when they position their store brands close to national brands. Following this line of reasoning, retailers should position their store brands as closely as possible to the national brands. Various theoretical studies have pointed to the benefits of store brands to enhance retailers' position vis-á-vis national brands (e.g. Mills (1995); Bontems, Monier-Dilhan, and Réquillart (1999); Scott Morton and Zettelmeyer (2004); Villas-Boas and Chambolle (2015)). Following empirical studies support this argument: Chintagunta, Bonfrer, and Song (2002) consider the pricing implications when retailers introduce store brands and find that retailer's margins increase after store brand introduction. Similarly, Ailawadi and Harlam (2004) show empirical results that relate higher retail margins on national brands in the presence of store brands to greater bargaining power. Meza and Sudhir (2010) examine how store brands strengthen a retailer's bargaining power and whether this retailer can strategically influence the negotiations by favoring store brands more than what would be implied by optimal category pricing. Cohen and Cotterill (2011) provide a counterfactual analysis under different structural models of wholesale price determination to quantify the implications for retailer profits when their store brands are hypothetically deleted. All this work suggests that store brand introduction result in better wholesale prices for retailers.

As a novelty, this paper contributes to the literature by examining how store brands "perform" across different retail chains in response to an extraordinary wholesale price increase stemming from an upstream market structure shock. My analysis provides therefore an additional argu-

[2]See for instance Ching (2010), Frank and Salkever (1992).

ment for introducing store brands that are close substitutes to national brands as they build a partial insurance against brand-specific wholesale price shocks on manufacturer brands. The power of retailers' store brands in mitigating upstream market power shocks depends on their store brand positioning. As imitating store brands can limit the manufacturer wholesale price increases, they are strategically important in product categories where upstream market structure shocks are more likely.

## 3 Coffee market details and data

Germany is the world's second largest coffee market (after the US) and coffee is the country's most popular beverage.[3] The grocery retail market is the main sales channel for coffee. Although coffee houses abound in Germany, most coffee is sold through grocery stores. For this study, I use data on German households' coffee purchases from the Nielsen Company (Germany) GmbH Homescan consumer panel data on fast moving consumer goods (FMCG).[4] Each year, there is a representative sample of German households recording their fast-moving consumer goods purchases. All participants are surveyed once a year and several demographic variables such as age, household size, and income are recorded. I match the representative sample of households with purchases made in the coffee category. The focus is on ground coffee, which constitutes around 60% of total coffee purchases in the sample, and of which more than 95% is accounted for by the sales of 500g packages.

Figure 1 shows the sales-weighted monthly average price on an aggregate of store brands and the price of an aggregate of national ground coffee brands from 2002 to 2012. In addition, it shows how the world market price for the main raw coffee bean types, Arabica and Robusta, evolved.[5] As Bettendorf and Verboven (2000) explain, 1.19 kg of raw coffee beans are required to produce 1 kg of roasted ground coffee. I therefore adjust raw coffee bean prices by the factor 1.19 in Figure 1 to obtain prices per 500g of roasted coffee. The left panel depicts average prices from 2002 to 2012 and the right panel zooms in on the years 2004 and 2005.

In December 2004 and April 2005, the leading coffee manufacturers announced higher wholesale prices to the retailers, citing higher raw coffee bean prices.[6] In fact, after receiving information

[3]According to the German Coffee Association, 162 liter coffee per capita were consumed in 2014. This is more than the per capita consumption of mineral water (143,5 liter) and beer (107 liter) in 2014.

[4]Any analysis calculated (or derived) and conclusion drawn from the Nielsen data are those of the researcher and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

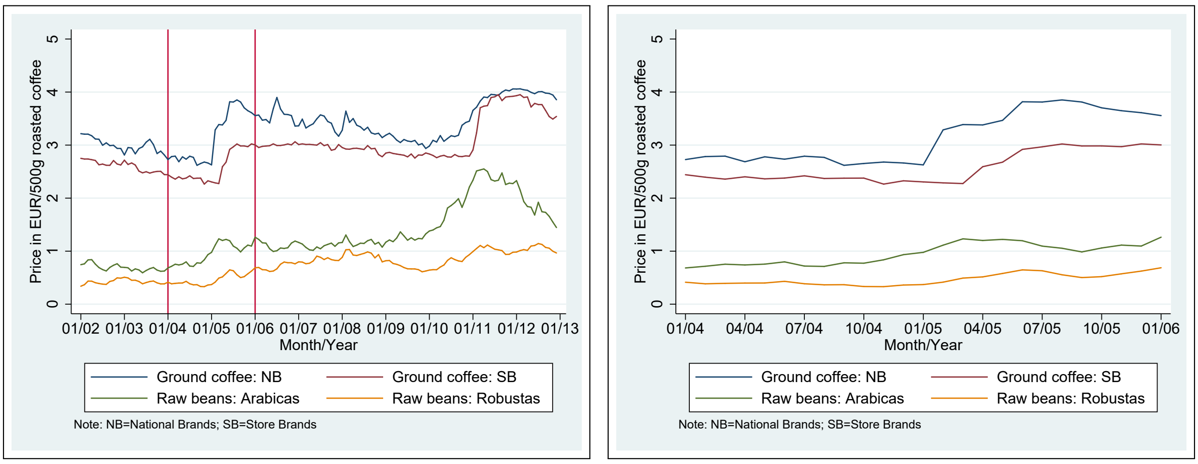
[5]The International Coffee Organization (ICO) provides a composite indicator price for raw coffee beans and the respective subgroups. See `http://www.ico.org/coffee_prices.asp?section=Statistics`; accessed 13 November 2015.

[6]See Bundeskartellamt (2010).

from a whistle-blower and undertaking a dawn raid in July 2008, the German antitrust agency found the four largest coffee manufacturers guilty of coordinating prices from 2000 to 2008. Over this period, five jointly coordinated increases in wholesale prices were verified, including two in in 2005. Coffee store brands are almost exclusively produced by fringe coffee manufacturers and procured in a competitive process. Some retail chains even own coffee roasting facilities themselves.[7] This unique setting of asymmetric changes in retailers' marginal costs across brands makes the German coffee market especially well suited to study the strategic role of store brands as a tool to mitigate wholesale price increases by national brand manufacturers.

After raising the prices on branded coffee, retail chains raised their prices for store brands with some delay in mid-2005, also citing higher input costs. The increase in store brand prices, however, is not of the same magnitude as that for national brands. Thus, the price gap between store brands and national brands widened. In this paper, I examine the economic mechanism behind this price divergence.

Figure 1: Time series of monthly (sales-weighted) average prices for store and national coffee brands




In order to generate empirical evidence on how retailers can use their store brands to mitigate manufacturers' (joint) pricing power, I examine the change in prices, margins and demand due to the national coffee brands' wholesale price increases in 2004 and 2005. I restrict the data to two years to ensure that substitution to store brands is driven only by price differentials and not by changes in customer perceptions of store brand quality. I did not observe any introductions of new quality tiers in this two year period that might have driven a change in market share or price.

[7]See Bundeskartellamt (2010) and Bundeskartellamt (2014).

Only purchases at the top eleven retail chains are explicitly considered for demand estimation, because infrequent shopping trips to smaller chains make it difficult to impute prices for non-chosen alternatives and to cross-validate data. The imputation of prices for non-chosen alternatives and the construction of choice sets are described in the Appendix A.1. These top eleven retail chain purchases represent approximately 65% of all ground coffee purchases made in the sample. A common practice in the demand estimation literature with household data, e.g. Dubé, Hitsch, and Rossi (2010), Geyskens, Gielens, and Gijsbrechts (2010) or Erdem, Keane, and Sun (2008), is to only include households that purchase within the product category at least several times. I follow this practice and only consider households that purchase ground coffee brands at least four times. Table 1 compares the sample of ground coffee purchasing households with the full household sample. I conclude that in terms of observable demographics both samples seem comparable in both years.[8] In total, we have 3,736 ground coffee brand purchasing households we can include in our analysis.[9]

The demand estimation technique in this paper is computationally very burdensome and computing time increases with the number of households (and products) included in the estimation. In order to make the estimation feasible in an adequate amount of time, I draw a sub-sample of coffee purchasing households for estimation. In particular, I draw (without replacement) 1,000 households (approximately a 27% sample) from the 3,736 coffee purchasing households with weights constructed from the projection factors in order to obtain a representative sample. Similar papers, such as Osborne (2011), also take samples of households to reduce the computational burden in the estimation.

To estimate the demand model, I define a product as a coffee-brand-mildness-retail chain combination as most brands further differentiate their coffee products with a standard version and a mild version. Typically, the mild version has a higher price. In the demand estimation literature there are two common approaches to define the outside good. One approach is to use each households' shopping trip to a grocery store or supermarket as a choice observation (e.g. Besanko, Gupta, and Jain (1998); Pavlidis and Ellickson (2017)). With this approach, however, the outside good has often a very high market share, around 90%, depending on the product category. As a consequence, price-sensitivity estimates are very low because in the majority of choice observations consumers do not react to price reductions as they are actually not intending

[8]Note that the income variable was originally a categorical variable. Thirteen different income ranges are recorded. In order to simplify the comparison of household samples I constructed a continuous income variable for each household by drawing from a uniform distribution within the respective income range.

[9]There are 297 households that we observe only in 2004 and 236 households only in 2005.

Table 1: Description of household samples and characteristics

| | 2004 | | | | | 2005 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variables | Mean | Stand. Dev. | Median | Min | Max | Mean | Stand. Dev. | Median | Min | Max |
| Full sample | $N$ =6387; $N_{balanced}$ =5543; $N_{only2004}$ =844 | | | | | $N$ =6342; $N_{balanced}$ =5543; $N_{only2005}$ =799 | | | | |
| Household size | 2.40 | 1.17 | 2.00 | 1.00 | 9.00 | 2.36 | 1.17 | 2.00 | 1.00 | 9.00 |
| Monthly net income in EUR | 2264.11 | 1233.84 | 2046.00 | 502.00 | 7996.00 | 2265.71 | 1289.10 | 2020.00 | 500.00 | 7986.00 |
| Urban dummy | 0.63 | 0.48 | 1.00 | 0.00 | 1.00 | 0.63 | 0.48 | 1.00 | 0.00 | 1.00 |
| Age oldest household member | 54.09 | 14.29 | 54.00 | 20.00 | 98.00 | 54.12 | 14.57 | 54.00 | 20.00 | 99.00 |
| Childen in household | 0.27 | 0.44 | 0.00 | 0.00 | 1.00 | 0.26 | 0.44 | 0.00 | 0.00 | 1.00 |
| Female share in household | 0.52 | 0.27 | 0.50 | 0.00 | 1.00 | 0.51 | 0.28 | 0.50 | 0.00 | 1.00 |
| Ground coffee sample | $N$ =3500; $N_{balanced}$ =3203; $N_{only2004}$ =297 | | | | | $N$ =3439; $N_{balanced}$ =3203; $N_{only2005}$ =236 | | | | |
| Household size | 2.49 | 1.14 | 2.00 | 1.00 | 7.00 | 2.46 | 1.13 | 2.00 | 1.00 | 7.00 |
| Monthly net income in EUR | 2279.31 | 1190.58 | 2074.00 | 506.00 | 7996.00 | 2292.18 | 1258.57 | 2058.00 | 500.00 | 7966.00 |
| Urban dummy | 0.60 | 0.49 | 1.00 | 0.00 | 1.00 | 0.60 | 0.49 | 1.00 | 0.00 | 1.00 |
| Age oldest household member | 55.70 | 13.62 | 55.00 | 23.00 | 98.00 | 56.29 | 13.75 | 56.00 | 23.00 | 99.00 |
| Childen in household | 0.27 | 0.45 | 0.00 | 0.00 | 1.00 | 0.26 | 0.44 | 0.00 | 0.00 | 1.00 |
| Female share in household | 0.52 | 0.24 | 0.50 | 0.00 | 1.00 | 0.52 | 0.24 | 0.50 | 0.00 | 1.00 |
| Estimation sample | $N$ =960; $N_{balanced}$ =909; $N_{only2004}$ =51 | | | | | $N$ =949; $N_{balanced}$ =909; $N_{only2005}$ =40 | | | | |
| Household size | 2.30 | 1.13 | 2.00 | 1.00 | 7.00 | 2.27 | 1.13 | 2.00 | 1.00 | 7.00 |
| Monthly net income in EUR | 2071.40 | 1136.48 | 1838.00 | 512.00 | 7706.00 | 2094.07 | 1205.74 | 1836.00 | 500.00 | 7940.00 |
| Urban dummy | 0.60 | 0.49 | 1.00 | 0.00 | 1.00 | 0.61 | 0.49 | 1.00 | 0.00 | 1.00 |
| Age oldest household member | 56.04 | 14.68 | 57.00 | 24.00 | 91.00 | 56.52 | 14.82 | 58.00 | 25.00 | 94.00 |
| Childen in household | 0.24 | 0.43 | 0.00 | 0.00 | 1.00 | 0.23 | 0.42 | 0.00 | 0.00 | 1.00 |
| Female share in household | 0.53 | 0.28 | 0.50 | 0.00 | 1.00 | 0.53 | 0.29 | 0.50 | 0.00 | 1.00 |

to buy within that product category regardless of the price movements.[10] Another approach is to condition on purchasing within the product category and use for instance an aggregate of other (fringe) products as the outside good (e.g. Villas-Boas (2007)). I follow the latter approach and use an aggregate of ground coffee purchases at the remaining fringe retail chains to form the outside good.

Table 2: Data description: by brand and by retailer

| | Average price (EUR) | | | Market share (%) | | |
|---|---|---|---|---|---|---|
| | 2004-2005 | 2004 | 2005 | 2004-2005 | 2004 | 2005 |
| By Brand | | | | | | |
| Outside good | 3.06 | 2.77 | 3.46 | 33.92 | 36.44 | 31.00 |
| Brand 1 | 3.20 | 2.87 | 3.63 | 5.73 | 6.03 | 5.37 |
| Brand 2 | 2.98 | 2.62 | 3.37 | 5.45 | 5.32 | 5.61 |
| Brand 3 | 2.99 | 2.66 | 3.41 | 11.06 | 11.51 | 10.54 |
| Brand 4 | 2.76 | 2.36 | 3.26 | 7.41 | 7.67 | 7.12 |
| Brand 5 | 2.40 | 2.13 | 2.87 | 5.04 | 6.02 | 3.90 |
| Brand 6: SB | 2.48 | 2.25 | 2.66 | 27.23 | 22.34 | 32.90 |
| Brand 7 | 3.46 | 3.14 | 3.97 | 4.15 | 4.67 | 3.55 |
| By Retailer | | | | | | |
| Outside good | 3.06 | 2.77 | 3.46 | 33.92 | 36.44 | 31.00 |
| Retailer 1 | 2.45 | 2.24 | 2.64 | 9.73 | 8.58 | 11.06 |
| Retailer 2 | 2.48 | 2.27 | 2.67 | 10.64 | 9.04 | 12.49 |
| Retailer 3 | 2.90 | 2.53 | 3.40 | 5.19 | 5.59 | 4.73 |
| Retailer 4 | 2.90 | 2.62 | 3.16 | 8.36 | 7.53 | 9.33 |
| Retailer 5 | 2.85 | 2.50 | 3.21 | 10.66 | 10.20 | 11.21 |
| Retailer 6 | 3.30 | 2.96 | 3.62 | 1.10 | 1.00 | 1.21 |
| Retailer 7 | 2.93 | 2.73 | 3.32 | 1.55 | 1.88 | 1.15 |
| Retailer 8 | 2.88 | 2.57 | 3.31 | 4.90 | 5.33 | 4.40 |
| Retailer 9 | 2.83 | 2.58 | 3.13 | 2.50 | 2.54 | 2.45 |
| Retailer 10 | 2.79 | 2.49 | 3.20 | 9.44 | 10.12 | 8.66 |
| Retailer 11 | 3.19 | 2.84 | 3.50 | 2.01 | 1.76 | 2.30 |
| No. of households | | | | 1000.00 | 960.00 | 949.00 |
| No. of purchase incidents | | | | 24701.00 | 12659.00 | 12042.00 |

Table 2 displays market-share weighted average prices for the products considered for estimation summarized by brands and by retail chains. For reasons of confidentiality, I kept all brand and retailer names anonymous in this article. I also compare average prices and market shares between the years 2004 and 2005. Average prices for all coffee brands and across all retailers increase from 2004 to 2005, but the increase is smaller for store brands. As a consequence, store brands (Brand 6) gain market shares in 2005. Notably, Retailers 1 and 2, which exclusively stock store brands in the coffee category, expand their market shares. Going back to Table

[10]With data tracking consumers within a store, one could condition only on those choice observations in which the consumer stands actually in front of the product category shelves for some seconds. This approach requires, however, more information

Table 3: Store brand share description

| | Share of store brand sales (%) | | |
|---|---|---|---|
| Retailer | 2004-2005 | 2004 | 2005 |
| Retailer 1 | 100.00 | 100.00 | 100.00 |
| Retailer 2 | 100.00 | 100.00 | 100.00 |
| Retailer 3 | 4.67 | 1.56 | 8.93 |
| Retailer 4 | 37.29 | 27.93 | 46.05 |
| Retailer 5 | 9.61 | 7.84 | 11.47 |
| Retailer 6 | 18.55 | 14.36 | 22.55 |
| Retailer 7 | 6.77 | 3.00 | 13.92 |
| Retailer 8 | 7.82 | 4.72 | 12.18 |
| Retailer 9 | 47.08 | 41.21 | 54.13 |
| Retailer 10 | 3.97 | 0.81 | 8.25 |
| Retailer 11 | 11.52 | 8.16 | 14.51 |

1, we observe that summary statistics of observable consumer characteristics in all samples do not change substantially from 2004 to 2005. Therefore, changes in observable consumer characteristics cannot explain the increase in store brand market shares in 2005. This is an important observation because Lamey et al. (2007) and more recently Dubé et al. (2018) have found that country-wide negative income and wealth shocks, for instance during a recession, increase store brand market shares.

In Table 3, the focus is on the within-retail chain substitution from national brands to store brands by comparing the share of ground coffee store brand sales on retailers' total ground coffee sales over both years. As Retailers 1 and 2 exclusively stock coffee store brands, their share of store brand sales is always 100%. For the remaining retailers, we observe that the share of store brand sales increases from 2004 to 2005, but with different magnitude across retailers. While for Retailer 5 the share of store brand sales increases only by around 4%, the share of store brand sales in Retailer 4 increases by approximately 18%.

# 4 Demand model and identification

I specify a demand model that assumes a household makes a discrete choice among $J_t$ ground coffee products at each purchase occasion.[11] Household $i$'s indirect utility from purchasing coffee

[11] I follow the modeling approach of other studies estimating the demand for ground coffee, such as Villas-Boas (2009), Nakamura and Zerom (2010), Draganska, Klapper, and Villas-Boas (2010) and Bonnet, Dubois, Villas Boas, and Klapper (2013). The assumption of single unit demand holds for the majority of purchases (75%) and seems to be an adequate approximation. For the remaining observations where households purchase several packages of the same brand, I assume that the household has already decided on the number of packages in a first stage, for instance because of the household size or special occasion, and on the second stage chooses a brand based on price per package, reducing the problem to a discrete choice setting again. If households purchase more than one package due to consumer stockpiling behavior at price discounts and not due to generally higher

product $j$ at period $t$ is

$$U_{jit} = \beta_{i,b} + \alpha_i p_{jt} + \delta_i \mathbf{1}\{j = mild\} + \psi_{i,l} + \varepsilon_{jit}, \tag{1}$$

where $b \in \{Brand\ 1, ..., Brand\ 7\}$ and $l \in \{1, \ldots, 10\}$ in this application. The price is given by $p_{jt}$ and I normalize the mean utility of the outside option to zero, $u_{0it} = 0$. The indicator variable $\mathbf{1}\{j = mild\}$ denotes whether the product $j$ is the mild variant of the coffee brand $b$. The coefficient $\psi_{i,l}$ is an indicator variable denoting that a product $j$ is offered at retail chain $l$.

The specification of the demand model implicitly assumes that all households have complete information about the coffee products offered by the eleven retail chains and the final purchase decision is not only determined by the brand itself but also by the preference for the specific retail chains where the product is offered. These chain-specific preference parameters can capture unobserved retail chain characteristics such as proximity to a household or reputation. This full information choice set assumption still dominates the applied demand estimation literature, especially if consumers' consideration sets are not directly observed. (Notable exceptions are Goeree (2008), Draganska and Klapper (2011) and Honka, Hortaçsu, and Vitorino (2017).) The structure of the data combined with enough preference heterogeneity allow us to infer individual retail chain parameters $\psi_{i,l}$ that effectively locate the alternatives most likely to be considered and offer valuable information on the level of differentiation across retail chains required for the supply model following this section. This approach aims to soften the full information assumption by altering individual retail chain parameters.

For the individual demand specification in Equation 1, I expect coefficients of retail chains from which a household never purchased any coffee to approach large negative values. A large negative retail chain parameter effectively drops all products within that retail chain from the household's choice set since their choice probabilities approach zero. On the contrary, coffee brands in retail chains a household often purchases will have positive choice probabilities since they are in the household's individual choice set. This is a crucial property as it diminishes a likely bias in the estimate of $(\alpha_i, \delta_i, \beta_{i,Brand\ 1}, \ldots, \beta_{i,Brand\ 7})'$ caused by including alternatives a household did not consider at any time of purchase while not controlling for that heterogeneity in choice sets.

The deterministic part of utility is defined as $V_{jit} = \beta_{i,b} + \alpha_i p_{jt} + \delta_i \mathbf{1}\{j = mild\} + \psi_{i,l}$ for household $i$ and every choice alternative $j$ at time $t$. Assuming that $\varepsilon_{ijt}$ follows a type I extreme

consumption rates (e.g. because of household size or special occasions), household price parameter estimates are biased as shown in Hendel and Nevo (2006) and Erdem, Imai, and Keane (2003).

value distribution, individual choice probabilities are given by a multinomial logit model

$$Pr_{it}\left\{j|p\right\} = s_{jit} = \frac{e^{V_{jit}}}{1+\sum_{k=1}^{J_t} e^{V_{kit}}}. \tag{2}$$

I have nineteen parameters to estimate on the household level, denoted as $\theta_i = (\alpha_i, \delta_{i,mild}, \beta_{i,Brand\ 1}, \ldots, \beta_{i,Brand\ 7}, \psi_{i,1}, \ldots, \psi_{i,10})'$.[12] To estimate individual demand parameters $\theta_i$ I rely on a hierarchical Bayesian multinomial logit model with a mixture of normals first-stage prior. This approach not only allows approximate deviations from standard normal heterogeneity distributions as described in Rossi, Allenby, and McCulloch (2005), but is also well suited for the purpose of estimating individual level coefficients when the amount of data provided by each panel unit is rather small. Table 4 shows the varying amount of information provided by each household in the sample. The hierarchical Bayesian approach effectively

Table 4: Distribution of the number of coffee purchase incidents across $N$ =1000 households in the estimation sample

| | Min. | 1st Qu. | Median | Mean | 3rd. Qu. | Max. |
|---|---|---|---|---|---|---|
| Purchases | 4 | 13 | 21 | 25 | 32 | 109 |

pools information across households through the prior, shrinking extreme coefficient estimates (implied by a short history of observations on the individual level) towards the sample mean.

The model should provide estimates of $\{\theta_i\}$ in line with basic economic theory. While brand and store coefficients can take any sign and value, price coefficients should be constrained to be negative for every household, i.e. $\alpha_i \leq 0$.[13]

The Bayesian implementation of the demand model follows the approach in Pachali et al. (2017). They propose a Markov Chain Monte Carlo (MCMC) algorithm similar to Rossi et al. (2005) that effectively samples from the posterior distribution of a model while allowing sign and/or order constraints on some coefficients. The basic idea is that unconstrained coefficients have a standard normal prior while sign and order constraints are imposed through a log-normal distribution. MCMC inference is performed on a transformed space exploiting the property that

[12] I estimate individual retail chain preferences relative to a baseline retail chain (retailer 11) in order identify the likelihood. For $l \neq 11$, $\hat{\psi}_{i,l} = \psi_{i,l} - \psi_{i,11}$ measures household $i$'s preference for the $l$th retail chain relative to the eleventh retail chain as the baseline level.

[13] In an unconstrained model, the marginal posterior distribution of the price coefficient has potentially non-negligible support for positive values in the right tail due to the small number of observations on the individual level. This is problematic for computing counterfactual prices because it would be optimal to charge infinitely high prices and only keep consumers with weakly positive price coefficients in the market. Before researchers were able to include sign constraints on the price parameter, they trimmed ex-post few remaining households with positive price coefficients for counterfactual simulations. See for instance Dubé, Hitsch, Rossi, and Vitorino (2008), p. 423. More recent studies implement sign constraints on the price coefficient. See for instance Allenby, Brazell, Howell, and Rossi (2014) and Pachali, Kurz, and Otter (2017).

coefficients are jointly normally distributed after the transformation.

I specify the constraints on $\theta_i$ by defining the functional form $g : \mathbb{R}^k \rightarrow \mathbb{R}^k_c$ mapping conditionally normally distributed variates $\theta_i^*$ to sign constrained coefficients $\theta_i$ that enter the likelihood, where $k$ denoting the number of coefficients in $\theta_i$. The hierarchical prior is specified as follows in this application

$$
\theta_i^* = \begin{pmatrix} \alpha_i^* \\ \delta_i^* \\ \beta^*_{i,Brand\ 1} \\ \vdots \\ \beta^*_{i,Brand\ 7} \\ \psi^*_{i,1} \\ \vdots \\ \psi^*_{i,10} \end{pmatrix} = g^{-1}\left(\theta_i\right) = \begin{pmatrix} ln(-\alpha_i) \\ \delta_i \\ \beta_{i,Brand\ 1} \\ \vdots \\ \beta_{i,Brand\ 7} \\ \psi_{i,1} \\ \vdots \\ \psi_{i,10} \end{pmatrix} \sim N\left(\bar{\theta}^*, V_{\theta^*}\right)^{(\text{ind}_i)}, \tag{3}
$$

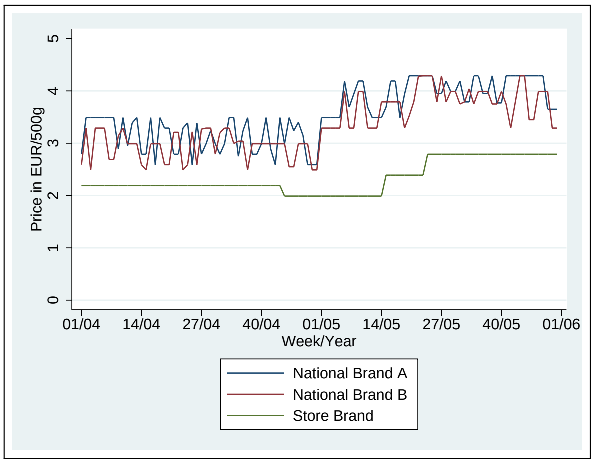
for the mixture of $S$ multivariate normals as a first-stage prior model on the transformed coefficients and $\text{ind}_i$ is the latent indicator variable denoting component membership of household $i$, with $\text{ind}_i \in \{1, \ldots, S\}$. Appendix A.2 provides more details on the MCMC approach and information about prior specifications.

The analysis depends on obtaining realistic substitution patterns between national brands and store brands. The variation that identifies these substitution patterns has two different sources: short-term as well as long-term variation in relative prices. Figure 2 shows the typical price variation that we can observe within a retail chain.[14] First, national coffee brands regularly offer price discounts with different frequencies over time. Most of the time not all national brands are simultaneously on sale. Store brands typically follow a every day low price strategy. This means price differences across brands fluctuate within a retail chain over time. Second, after the coffee manufacturers increased their wholesale prices in 2005, the gap in average retail prices between national brands and store brands widened. Afterward, it narrowed again as retailers adjusted their store brand pricing. The price differential in terms of average retail prices for national brands and store brands, however, was still larger after the retailers' price adjustment for store brands in the second half of 2005 than in 2004, as can be observed in Figure 1 as well as in Figure 2. Given that preferences did not dramatically change within these two years, this price divergence suggests a long-term adjustment in average retail margins due to

[14] The exact retail chain is not stated here, also the brands get different labels to support anonymity.

category profit optimization. The change in average retail prices from 2004 to 2005 was driven by both higher wholesale margins for manufacturer brands and also higher input costs for all coffee brands. From a retailer's perspective this was an asymmetric cost shock because marginal costs of stocking national brands increase more than those of stocking store brands. As a result, the data exhibit both short-term as well as long-term price variation between coffee brands. Both sources of price variation are helpful for identifying consumers' brand preferences, especially as we are interested in how consumers choose between national brands and store brands.

Figure 2: Typical price variation within a retail chain




Price endogeneity is a major concern in almost all demand estimation applications with observational data if firms adjust prices to (by the econometrician) unobserved demand shocks resulting in biased price parameter estimates. Typically, we expect that firms raise (or lower) prices if they observe positive (or negative) demand shocks. This leads to a bias of the price parameter towards zero, the so-called attenuation bias. Alternatively, firms can also lower prices if they observe a positive demand shock for that product and compete more intensively for a larger market for instance if more shoppers are on the market. Thus, the price parameter would have a negative bias. In general, changes in (by the econometrician) unobserved product characteristics that correlate with prices and influence demand, such as advertisement or shelf position, are problematic as they generate a bias with often unknown direction. The aim is to control for all possible demand factors that are correlated with prices.

Much of the demand estimation literature using individual homescan data of grocery purchases assumes that price variation in the data is exogenous after controlling for brand and/or product fixed-effects.[15] According to Erdem et al. (2008), the reasoning for that assumption stems from a typical pricing pattern of grocery products that we observe in retail chains. Typically, we observe a regular price as price ceiling with rotating sharp price reductions for one or few other

[15]Examples are Dubé et al. (2010) or Erdem et al. (2008).

products within the same category. These short-term price reductions are then followed by a return to former regular prices afterward. Only if these price cuts are retailers' responses to consumers' taste shocks do we face a problem of price endogeneity. This is, however, implausible for the following three reasons: First, short-term price promotions are usually negotiated between retailers and manufactures over a larger time period. The negotiations do not only include the level of price promotions but also the timing. Thus, they cannot be regarded as sudden reactions to consumers' taste shocks because they are pre-determined. Second, I argue that brand-specific taste shocks are a quite rare event unless there is some kind of brand-specific food scandal or product safety issue. I would rather expect taste shocks for the entire category but then we would observe almost exclusively simultaneous price cuts for all products within the category which does not match the observed retail pricing pattern. Third, as pointed out by Pesendorfer (2002), this retail pricing pattern can be best described as retailers playing mixed strategies in order to inter-temporally price discriminate between consumers. Thus, short-term price fluctuations caused by mixed strategies are unrelated to taste shocks and can be regarded as exogenous.

Nevertheless, we might be still concerned about the long-term variation in prices as shown in Figure 1. If the long-term price variation is related to changes in unobserved product characteristics, prices are endogenous. One possible approach to deal with endogenous variables is to apply instrumental variable techniques. As Petrin and Train (2010) explain, individual discrete choice demand models are non-linear estimations and per individual there are often only few or even no purchases of each product at a given time, the usual product-market control approach of aggregate demand estimations in combination with a two-stage least squares instrumental variable technique, such as in the seminal papers by Berry (1994) and Berry, Levinsohn, and Pakes (1995), cannot be applied. Petrin and Train (2010) suggest using a control function approach in such settings instead. I follow Petrin and Train (2010) and regress the potential endogenous price variable on control variables and instruments. As suggested, I retrieve the error terms and add them into the demand estimation equation.

Raw coffee bean world market prices are the first set of instruments that I consider for the control function approach. With minor modifications, I follow, amongst others, Villas-Boas (2009) and Bonnet et al. (2013), by interacting brand dummy variables with the monthly International Coffee Organization (ICO) composite indicator price, a price index that reflects the world market price for raw coffee beans.[16] It is a weighted average of the commodity prices for Colombian Mild Arabicas, Other Mild Arabicas, Brazilian and Other Natural Arabicas and Robusta raw coffee beans. The idea behind this approach is that the exact cost of inputs may differ for each

[16] See `http://www.ico.org/coffee_prices.asp?section=Statistics`; accessed 13 November 2015.

brand due to different suppliers of raw coffee beans and different cultivation regions. The interaction allows for such differences.[17] World market prices for raw coffee beans are considered to be exogenous since the world market trading volume is too large to be influenced by local demand shocks in Germany.

The manufacturer price agreement itself can be used as a second set of instruments. Similar to changes in input prices, an increase in wholesale prices due to price coordination (i.e. more upstream market power) can be regarded as a supply side instrument. In other words, it cannot directly influence the demand side except indirectly through its impact on prices.[18] The price agreement affected brand manufacturer wholesale pricing differently than store brand producers. And even among the national brands the effect might vary. Thus, I have another brand-specific exogenous price variation to exploit. Due to the extensive investigation by the German antitrust agency, I know ex-post that the main coffee manufacturers announced price increases in December 2004 and April 2005. For the control function approach, I therefore use indicator variables denoting the first and second price increase time span in the sample period interacted with brand dummy variables as additional instruments.[19]

Following the reasoning on price endogeneity with respect to supermarket pricing patterns in Erdem et al. (2008), I consider short-term price variation due to temporary price cuts to be exogenous. Therefore, I include a price promotion dummy variable indicating a price below the regular price into the first stage regression. The first stage regressions of the control function approach are summarized in Table A.1 in the Appendix. I use the error terms of the forth specification to include into the demand estimation equation. Note that around 91 % of the variation in prices can be explained by the instruments and exogenous demand variables.

## 5 Estimation results

I estimate the model described in Section 4 and Appendix A.2 with a successively larger number of mixture components to compare models with more flexible prior heterogeneity specifications and compute their log marginal likelihoods using the Newton-Raftery method on the values trimmed by 1 % on the bottom and the top of likelihood draws, as suggested in Dubé et al. (2010) and Gamerman and Lopes (2006).[20] According to Table 5, the one normal component model

[17] In contrast to Villas-Boas (2009), I use brand dummy variables instead of brand-retailer dummy variables for interaction because I think this is the right level of observation.

[18] The price agreement was not detected until July 2008 and only by evidence from a whistle-blower.

[19] The first price increase dummy ranges from January 2005-April 2005 and the second price increase dummy ranges from May 2005 -December 2005.

[20] I run the sampler for $R = 50,000$ iterations and keep every fifth draw. I decided to burn the first $5,000$ kept draws after inspecting time series plots of individual level posterior distributions. Posterior inference is based on $5,000$ draws from the converged posterior distribution.

dominates in terms of model fit based on the log marginal likelihood. Figure 3 shows marginal

Table 5: Log marginal likelihood values for demand models

| | Value |
|---|---|
| One normal component | -33255.9 |
| Five normal component | -33605.8 |
| Ten normal component | -33744.7 |

posterior densities of the ground coffee preference coefficients implied by the one component model. The left panel of the figure illustrates the posterior densities of the price coefficient. As restricted in the hierarchical prior, the density only supports values in the negative domain. The shape of the marginal posterior density of the price coefficient is therefore uni-modal. The density represents several households in its left tail with a low parameter value being very sensitive to changes in the prices.

The right panel of the figure plots the marginal posterior densities of the ground coffee brand intercepts. Households with different brand preferences are located at different positions within these distributions. Table 6 summarizes quantiles and the first two moments of the marginal posterior distributions of all estimated coefficients implied by the one component model given the control function approach. The relatively high standard deviations of the brand intercepts indicate heterogeneous preferences for the different coffee brands. Preferences over retail chains exhibit substantial variation as well.

In the Appendix A.3, I show in Table A.2 as a robustness check the estimation results for the one component model without using the control function approach. For the price coefficients, the posterior mean is slightly smaller (by around 0.25) and the standard deviation is larger. The other demand parameters do not vary much from the estimates in Table 6. As most of the price variation can be explained in the first stage regression by adding exogenous demand variables and instruments, it is not surprising that the estimates do not dramatically change. In a further robustness check, I include a store brand trend variable which is an interaction of a store brand dummy variable with the year 2005. This variable aims to capture a possible change in preferences with respect to store brands over both years which might have been an alternative explanation for why store brands gain market shares in 2005. The posterior mean coefficient, however, is very close to zero (see Table A.3 in the Appendix) and the other parameter estimates do not really change after introducing that trend variable. This suggests that there is no dramatic change in store brand preferences and the increase in store brand market shares in 2005 is mainly driven by the higher price differential between store brands and national brands.

Figure 3: Marginal posterior densities of price sensitivity and coffee brand preference coefficients for the one component model

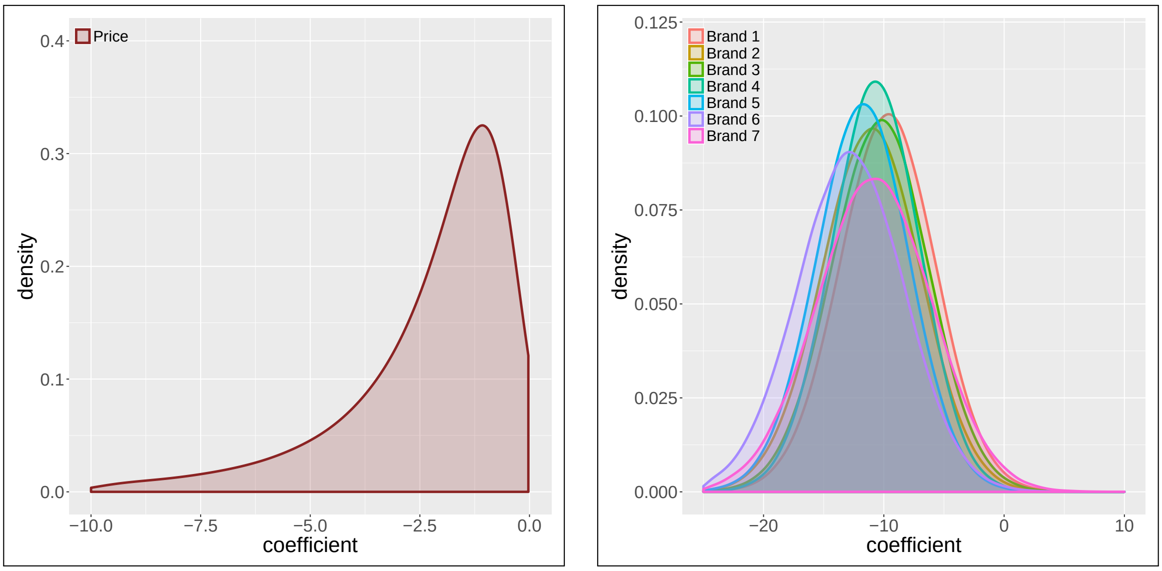



Table 6: Quantiles and first two moments of the the marginal posterior densities for the one component model

| Coefficients | Quantiles | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5% | 25% | 50% | 75% | 95% | Mean | Stand. Dev. |
| Price (EUR/500g) | -7.385 | -3.105 | -1.703 | -0.936 | -0.392 | -2.535 | 2.801 |
| Mild | -5.181 | -2.448 | -0.551 | 1.323 | 4.047 | -0.561 | 2.805 |
| Brand 1 | -16.263 | -12.369 | -9.691 | -7.040 | -3.226 | -9.714 | 3.969 |
| Brand 2 | -17.888 | -13.837 | -11.057 | -8.301 | -4.347 | -11.080 | 4.115 |
| Brand 3 | -16.986 | -13.036 | -10.310 | -7.635 | -3.761 | -10.338 | 4.021 |
| Brand 4 | -16.740 | -13.139 | -10.684 | -8.243 | -4.761 | -10.705 | 3.643 |
| Brand 5 | -18.153 | -14.353 | -11.757 | -9.185 | -5.475 | -11.781 | 3.855 |
| Brand 6: SB | -20.145 | -15.770 | -12.777 | -9.785 | -5.516 | -12.792 | 4.449 |
| Brand 7 | -18.725 | -14.025 | -10.811 | -7.598 | -3.011 | -10.826 | 4.773 |
| Retailer 1 | -3.385 | 1.380 | 4.665 | 7.966 | 12.812 | 4.685 | 4.924 |
| Retailer 2 | -4.846 | 1.332 | 5.582 | 9.866 | 16.107 | 5.607 | 6.375 |
| Retailer 3 | -5.877 | -1.489 | 1.544 | 4.574 | 8.991 | 1.551 | 4.526 |
| Retailer 4 | -2.501 | 1.704 | 4.625 | 7.568 | 11.892 | 4.652 | 4.378 |
| Retailer 5 | -6.474 | -1.385 | 2.135 | 5.660 | 10.794 | 2.142 | 5.248 |
| Retailer 6 | -7.643 | -2.813 | 0.492 | 3.793 | 8.607 | 0.489 | 4.937 |
| Retailer 7 | -6.834 | -2.411 | 0.586 | 3.606 | 8.063 | 0.603 | 4.534 |
| Retailer 8 | -6.206 | -1.821 | 1.221 | 4.263 | 8.699 | 1.228 | 4.533 |
| Retailer 9 | -6.653 | -1.488 | 2.022 | 5.526 | 10.714 | 2.021 | 5.280 |
| Retailer 10 | -5.216 | -0.658 | 2.520 | 5.707 | 10.352 | 2.536 | 4.739 |
| CF | -2.470 | -1.263 | -0.423 | 0.417 | 1.625 | -0.423 | 1.246 |

The Hierarchical Bayesian model has the advantage that inference at the household level comes naturally with the estimation output. In order to illustrate that individuals are differently located within the unconditional posterior parameter distribution based on their revealed brand choices, I show in Table A.4 and A.5 in the Appendix the individual posterior parameter distribution for a rather less price sensitive and a rather more price sensitive household. The household in Table A.4 purchases almost exclusively the rather expensive coffee product "Brand 7". This chart shows that this household has a small price parameter (in absolute amount) and values "Brand 7" comparatively more than other brands. In contrast, the household in Table A.5 prefers one coffee product, "Brand 2", but also regularly switches between almost all other coffee brands. As a consequence, the estimated price parameter is relatively high (in absolute amount) since we have a rather price sensitive household in this case. Figure A.1 in the Appendix illustrates this pattern graphically by showing where these two households lie in the unconditional population price parameter distribution. This exercise intended to demonstrate that actual individual choices determine the demand parameter estimates and reveal households' preferences. The composition of these heterogeneous households influences market demand and drive the upcoming results.

Demand parameters in isolation are hard to interpret in discrete choice models. Therefore, I compute own- and cross-price elasticities in order to understand the substitution patterns across products. For the purpose of this analysis, it is particularly relevant to analyze the substitution patterns between national brands and store brands within a retail chain. Given the demand estimates, I follow Pachali et al. (2017) and rely on a procedure defined as lower level model non smoothed (n.s.) in order to approximate the preference distribution $\Theta$ representing the relevant population of German households. Intuitively, this approach integrates over individual level posterior distributions and takes the uncertainty from the demand estimation into account. In particular, I adapt the procedure to the context of fast moving consumer goods, where regular consuming households contribute more to the aggregated market demand. Thus, I construct household weights and divide the number of ground coffee purchases a household made in a year by the total number of yearly ground coffee purchases in the estimation sample. Each draw in the preference distribution therefore represents one purchase incident for the given time span. Obtaining $H$ coffee demand parameter draws $\Theta$ that represent the aggregated German coffee demand involves the following steps:

1. Draw a household $i_h$ with replacement from the estimation sample based on the constructed weights that take into account how much a household contributes to the aggregated demand

2. From household $i_h$ obtain a random demand parameter draw $\theta_h = (\alpha_h, \delta_{h,mild}, \beta_{h,Brand\ 1}, \ldots, \beta_{h,Brand\ 7}, \psi_{h,1}, \ldots, \psi_{h,10})'$ from household $i_h$'s set of MCMC demand parameter draws $\theta_{i_h} = [\theta_{i_h,1}, ..., \theta_{i_h,R}]$

3. Repeat step one and two $H = 10,000$ times and save all draws in matrix $\Theta = [\theta_1, ..., \theta_H]$

Elasticities, markets shares, and other equations will then be evaluated for each parameter draw in $\Theta$. Integrating over all draws yields to the aggregate elasticities and market shares. I evaluate market shares, own-price and cross-price elasticities for each draw at average regular coffee brand prices in 2004. Using the $H$ coffee demand parameter draws $\Theta$ that represent the aggregated German coffee demand, the implied market share estimate is given by

$$s_j(p) = \int Pr\left\{j|p, \theta_h\right\} f(\theta) d\theta, \tag{4}$$

where $f(\theta)$ is the density of demand parameters in the German aggregate demand for ground coffee. Market share $s_j$ of product $j$ is therefore a function depending on consumers' preferences, prices and other product characteristics.

Relating to Nevo (2000) and using the distribution of demand parameters, the aggregate price elasticities are given by

$$\eta_{jk} = \frac{p_k}{s_j}\frac{\partial s_j}{\partial p_k} = \begin{cases} \frac{p_j}{s_j} \int \alpha_h s_{jh}(1 - s_{jh}) f(\theta) d\theta & \text{if j=k} \\ -\frac{p_k}{s_j} \int \alpha_h s_{jh} s_{kh} f(\theta) d\theta & \text{otherwise} \end{cases} \tag{5}$$

Table 7 illustrates the substitution patterns between national brands and store brands within and across retail chains for a subset of coffee products. In contrast to simple logit and nested logit demand estimations, the high degree in heterogeneity allows for very flexible substitution patterns. We observe that cross-price elasticities are generally higher within a retail chain as indicated by the off-diagonal elements within the same retailer. This pattern indicates a share of consumers who regularly switch between brands within their preferred retail chain, for instance during a brand-specific price discount. We also observe, that cross-price elasticities across retail chains are higher for the same coffee brand. This suggests, that there are also households with strong preferences for a specific brand for which they shop across retail chains. In contrast, cross-price elasticities for different brands sold at different retail chains are closer to zero, indicating a weaker substitution relationship. While the own-price elasticities are high, they are in a similar range as in other studies that examined ground coffee demand (e.g. Villas-Boas (2009), Nakamura and Zerom (2010), Bonnet et al. (2013) and Bonnet and Villas-Boas (2016)). This is an quite interesting result because these prior studies mainly used aggregate sales and price data instead of individual household purchase decisions. Such a robustness regardless of using individual household data or aggregate market data for demand estimation is remarkable. In addition, compared to demand estimations on other consumer packaged goods, the own-price elasticities are plausible (e.g. ready-to-eat-cereals in Nevo (2001) or yoghurt in Villas-Boas (2007)). With respect to the cross-price elasticities, however, it seems that my estimates exhibit more heterogeneity. This can be explained by the fact that individual household purchase data are usually better for revealing substitution patterns across products than aggregate market data.

## 6 How store brand positioning dampens higher wholesale price

In the previous section, I estimated households' preferences for differentiated ground coffee brands. The demand estimation shows that cross-price elasticities between coffee brands within a retail chain are non-negligible. The aim of this section is first to illustrate how the upstream price agreement of brand manufacturers result in an asymmetric increase in retail chains' marginal costs. Second, I examine how the positioning of retailers' store brand dampens the increase in wholesale prices. Researchers typically do not directly observe marginal costs. To deduce those marginal costs, I introduce a structural supply model of retail competition and derive them from the downstream retail prices and implied markups.

The focus lies on measuring persistent (asymmetric) changes in marginal costs and the long-run adjustment of retail margins due to category profit maximization. In order to approximate retail chains' long-term category management and pricing decisions I abstract from the more complicated pricing policy where some retail chains play mixed strategies involving occasional price cuts on some products to inter-temporarily price discriminate between consumers. In particular, I compute retail chains' average shelf prices per product for 2004 and 2005 as an ingredient for the model of retail competition.[21]

The supply model closely follows the established notation in the empirical industrial organization and quantitative marketing literature as in Nevo (2001). The $L = 11$ multi-product retail chains belong to $M = 7$ retail companies which supply in total $J = 106$ products. This represents the retail market structure in 2004/2005 where some retail companies had different retail formats in their portfolio such as discounters and full-line supermarkets. Each retail chain $l$ belongs to a company $m$ and offers a set of coffee brands. Each retail company maximizes its profits

[21] In fact, I am not aware of any structural analysis on the grocery retail sector that explicitly models retail chains playing mixed strategies.

Table 7: Posterior mean price elasticities for a subset of products (only standard)

| Retailer | Brand | Elasticity | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | 1 | -5.43 | 0.13 | 0.25 | 0.34 | 0.19 | 0.05 | 0.07 | 0.23 | 0.02 | 0.04 | 0.00 | 0.21 | 0.02 | 0.06 | 0.02 | 0.01 | 0.00 | 0.01 |
| 3 | 2 | 0.37 | -6.42 | 0.31 | 0.40 | 0.60 | 0.05 | 0.18 | 0.06 | 0.15 | 0.06 | 0.01 | 0.06 | 0.19 | 0.16 | 0.06 | 0.04 | 0.01 | 0.01 |
| 3 | 3 | 0.26 | 0.11 | -6.51 | 0.34 | 0.22 | 0.14 | 0.03 | 0.10 | 0.05 | 0.22 | 0.02 | 0.06 | 0.04 | 0.50 | 0.04 | 0.03 | 0.02 | 0.01 |
| 3 | 4 | 0.40 | 0.17 | 0.39 | -6.51 | 0.48 | 0.05 | 0.05 | 0.08 | 0.07 | 0.05 | 0.05 | 0.05 | 0.03 | 0.09 | 0.40 | 0.05 | 0.01 | 0.01 |
| 3 | 5 | 0.32 | 0.36 | 0.37 | 0.69 | -7.86 | 0.07 | 0.06 | 0.04 | 0.05 | 0.04 | 0.04 | 0.03 | 0.02 | 0.09 | 0.06 | 0.30 | 0.01 | 0.00 |
| 3 | 6 | 0.30 | 0.10 | 0.84 | 0.29 | 0.27 | -6.76 | 0.02 | 0.02 | 0.02 | 0.03 | 0.16 | 0.02 | 0.02 | 0.15 | 0.03 | 0.02 | 0.10 | 0.00 |
| 3 | 7 | 0.19 | 0.17 | 0.09 | 0.11 | 0.09 | 0.01 | -4.77 | 0.04 | 0.04 | 0.02 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.04 |
| 4 | 1 | 0.07 | 0.01 | 0.03 | 0.02 | 0.01 | 0.00 | 0.00 | -4.47 | 0.22 | 0.20 | 0.06 | 0.16 | 0.03 | 0.06 | 0.02 | 0.01 | 0.00 | 0.01 |
| 4 | 2 | 0.01 | 0.04 | 0.03 | 0.04 | 0.02 | 0.00 | 0.01 | 0.46 | -5.35 | 0.31 | 0.13 | 0.07 | 0.12 | 0.14 | 0.05 | 0.04 | 0.01 | 0.02 |
| 4 | 3 | 0.02 | 0.01 | 0.10 | 0.02 | 0.01 | 0.00 | 0.00 | 0.28 | 0.22 | -4.31 | 0.13 | 0.03 | 0.03 | 0.30 | 0.03 | 0.02 | 0.02 | 0.01 |
| 4 | 6 | 0.00 | 0.00 | 0.01 | 0.03 | 0.02 | 0.02 | 0.00 | 0.14 | 0.15 | 0.20 | -3.67 | 0.01 | 0.01 | 0.10 | 0.09 | 0.04 | 0.09 | 0.00 |
| 5 | 1 | 0.08 | 0.01 | 0.02 | 0.02 | 0.01 | 0.00 | 0.00 | 0.19 | 0.04 | 0.03 | 0.00 | -4.66 | 0.20 | 0.58 | 0.18 | 0.12 | 0.02 | 0.10 |
| 5 | 2 | 0.01 | 0.05 | 0.03 | 0.02 | 0.01 | 0.00 | 0.00 | 0.07 | 0.15 | 0.05 | 0.01 | 0.43 | -5.14 | 0.76 | 0.23 | 0.20 | 0.06 | 0.13 |
| 5 | 3 | 0.01 | 0.01 | 0.10 | 0.02 | 0.01 | 0.01 | 0.00 | 0.04 | 0.04 | 0.14 | 0.03 | 0.32 | 0.19 | -4.32 | 0.21 | 0.30 | 0.07 | 0.06 |
| 5 | 4 | 0.02 | 0.02 | 0.03 | 0.27 | 0.03 | 0.00 | 0.00 | 0.04 | 0.05 | 0.05 | 0.09 | 0.36 | 0.22 | 0.79 | -5.48 | 0.64 | 0.09 | 0.05 |
| 5 | 5 | 0.01 | 0.01 | 0.03 | 0.05 | 0.18 | 0.00 | 0.00 | 0.03 | 0.07 | 0.05 | 0.06 | 0.32 | 0.24 | 1.43 | 0.83 | -6.72 | 0.16 | 0.03 |
| 5 | 6 | 0.00 | 0.00 | 0.02 | 0.01 | 0.01 | 0.02 | 0.00 | 0.01 | 0.02 | 0.04 | 0.15 | 0.07 | 0.08 | 0.38 | 0.15 | 0.20 | -2.66 | 0.02 |
| 5 | 7 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.02 | 0.04 | 0.04 | 0.03 | 0.01 | 0.36 | 0.22 | 0.37 | 0.10 | 0.04 | 0.02 | -4.36 |

**Note:** Cell entry $j, k$, where $j$ indexes rows and $k$ indexes columns, gives the percentage change in market share for product $j$ for a one percent change in the price of product $k$.

$$\Pi_m = \sum_{j \in S_m} [p_j - c_j] \, s_j(p) D, \tag{6}$$

for $m = 1, \ldots, M$ where $s_j(p)$ equals the market share of product $j$, $D$ denotes the market size and $S_m$ is the set of products offered by retail chains that belong to company $m$. Market share $s_j(p)$ is a function of households' preferences and prices, as specified in Equation 4.

As the model incorporates heterogeneous preferences, market shares are obtained by integrating over the distribution of households' preferences. I use the set of coffee demand parameter estimate draws $\Theta$ that I obtain in Section 5. Recall that I follow Pachali et al. (2017) and use their approach which is defined as lower level model non-smoothed (n.s.) in order to generalize on the German population of households. In particular, I adapt the procedure to the context of fast-moving consumer goods and construct weights that take into account how much a household contributes to the aggregated demand. Each draw in the preference distribution $\Theta$ therefore represents one purchase incident of the yearly demand for ground coffee.

I assume that retail chains compete in a pure-strategy Nash-Bertrand game with differentiated products and take profit cannibalization of other products in their portfolio into account. As in the rest of Europe, also the market concentration of the German grocery retail sector has increased over the past few decades. The high market share of low pricing discounters and the general notion that German consumers are relatively price-sensitive, however, indicates that the grocery retail sector remains rather competitive.[22] The Nash-Bertrand game with differentiated products should therefore be an adequate approximation of the mode of competition among German retail chains. This leads to the following first-order condition for product $j$

$$s_j(p) + \sum_{k=1}^{J} \Omega(k,j) \, [p_k - c_k] \, \frac{\partial s_k}{\partial p_j} = 0, \tag{7}$$

for $j = 1, ..., J$ where $\Omega$ is a $(J \times J)$-matrix defining the product ownership structure from the perspective of the retail company with $\Omega(k,j) = 1$ if both product $k$ and product $j$ are offered by retail chains of the same company (and zero otherwise). Matrix notation allows for a more elegant representation of the first order conditions

[22] See European Commission (2014).

$$s(p) + [\Omega * \Delta](p - c) = 0, \tag{8}$$

where $\Delta$ denotes a matrix of partial demand derivatives with respect to price with $\Delta(k, j) = \frac{\partial s_j}{\partial p_k}$. The $*$ represents an element-by-element matrix multiplication. The vectors of market shares, prices and marginal costs are represented by $s(p)$, $p$ and $c$ respectively.

For the research question, the effect of the price agreement on retailers' marginal costs is relevant. It should work as an asymmetric cost shock for retail chains. Changes in retailers' marginal costs due to higher wholesale margins for national brands or/and higher input costs are implicitly captured in the downstream marginal cost estimates. Given that we observe retail prices and estimate demand preferences, we can rearrange the first-order conditions to back out retailers' marginal costs:

$$c = p + [\Omega * \Delta]^{-1} s(p), \tag{9}$$

which yields a vector of marginal costs $c$. Naturally, a vector of retail margins $m^r = p - c = -[\Omega * \Delta]^{-1} s(p)$ can be obtained the same way. Retailers' marginal costs of supplying product $j$ can be further decomposed into the following parts:

$$c_j = c_j^r + p_j^w = c_j^r + m_j^w + c_j^w \tag{10}$$

where $c_j^r$ is the marginal cost of retailing, $p_j^w$ is the wholesale price for product $j$ which consists of the wholesale margin $m_j^w$ and the manufacturing marginal costs $c_j^w$. For ground coffee products, raw coffee bean prices constitute most of the manufacturing costs.

Manufacturing costs increased from 2004 to 2005 due to higher raw coffee bean prices. This increase should be relatively similar across products. The manufacturers' price agreement, however, affected only retailers' marginal costs for national brands through increasing wholesale margins. All else equal, retailers' marginal costs of stocking national brands should increase more than those of store brands. I measure, therefore, changes in prices, marginal costs and retail margins to analyze how retailers react on this asymmetric cost shock:

$$\Delta p_j = p_{j,t^*} - p_{j,t}, \tag{11}$$

$$\Delta c_j = c_{j,t^*} - c_{j,t}, \tag{12}$$

$$\Delta m_j^r = m_{j,t^*}^r - m_{j,t}^r, \tag{13}$$

where $t$ and $t^*$ are the periods before and after the asymmetric shock on retailers' marginal costs. Disentangling the change in marginal costs $\Delta c_j = c_{j,t^*} - c_{j,t} = \Delta c_j^r + \Delta p_j^w$, indirectly informs us on the changes in wholesale prices. Given the assumption that pure retailing marginal costs $c_j^r$ should be equal (or at least very similar) across brands within one retail chain, we can relate changes in $\Delta c_j$ mainly to changes in wholesale prices $p_j^w$ if we control for possible changes in pure retailing marginal costs for instance in a regression analysis with retailer fixed effect. This allows to examine how the positioning of store brands, proxied by measured cross-price elasticities, dampens wholesale price increases after upstream market structure shocks. Moreover, this approach has the advantage that we do not have to model the upstream market in order to compute unobserved wholesale prices (or margins). This would require more explicit assumptions on the upstream market conduct which is rather complicated given the documented price agreement between manufacturers.[23] For the purpose of this paper, however, this is not necessary, as there are no counterfactual experiments to compute. The focus lies in examining strategic margin adjustments by retailers and measuring changes in retailers' marginal cost mainly due to higher wholesale prices.

Figure 4 sheds light on changes in retailers' margins, costs and prices in response to the asymmetric cost shock (i.e. before and after the asymmetric cost shock). The left panel plots $\Delta p_j$ versus $\Delta c_j$ and the right panel plots $\Delta m_j^r$ versus $\Delta c_j$ for all retail chain-coffee-brand-mildness-combinations in the data. The data points are grouped into national brands and store brands. This change in store brand margins across retail chains is of particular interest. While we see a clear positive link between changes in marginal costs and prices, it is weaker between changes in marginal costs and margins, but still clear. Store brands exhibit a tendency toward smaller increases in marginal costs and therefore margins decrease in almost all cases. In contrast, for the majority of national brands retail margins increase. In other words, the cost-pass-through rate for the majority of store brands (with rather smaller marginal cost increase) is below 100% and for most national brands (with rather higher marginal cost increase) is above 100%.

The rationale for this observation is that retail chains typically respond to a relative higher in-

[23] Note that letting manufacturers act as one single monopolist is hardly the realistic description of that scenario.

crease in marginal costs of stocking national brands compared to store brands by increasing their margins for most national brands and decreasing their margins for store brands. This is due to retailers' category profit optimization, which implies optimal category pricing. As a consequence it is optimal for a retail chain to divert a portion of demand to the relatively more profitable store brands. In other words, retailers get price-sensitive customers to switch to store brands by lowering their margins. For branded products, with a relatively larger change in marginal cost, retailers increase their margins instead to charge more from consumers that have very strong preferences for national brands.

Figure 4: Scatter plots on changes in prices, margins and marginal costs

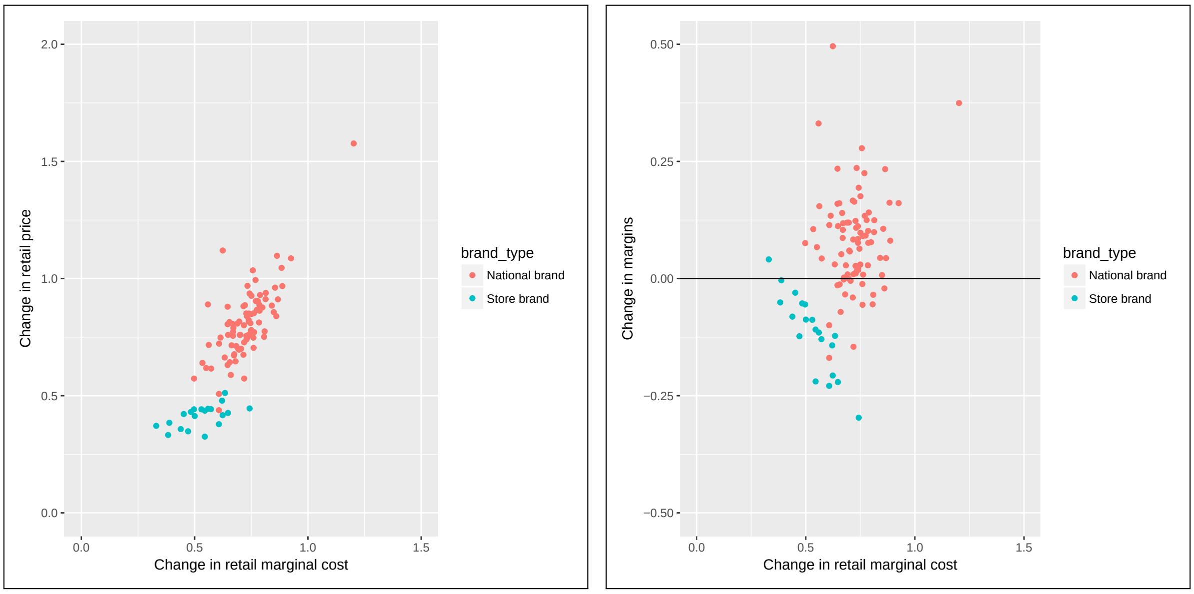



Table 8 shows how the adjustment of margins affects store brand markets shares within our retail chains and retail chains' category profits compared to a scenario of keeping margins constant. While this is not an equilibrium because there are profitable deviations for each retail chain, it is an illustrative benchmark where retail chains simply pass-on asymmetric cost increases by 100% without re-optimizing category profits. We observe that by adjusting retail margins, store brands gain market shares within multi-brand retail chains and category profits increase slightly. I expect the effect of margin adjustment on category profits to become stronger with a larger cost increase difference between national and store brands.

This pattern can be linked to the theoretical findings in Moorthy (2005), which provide a comparative-statics analysis of cross-brand cost pass-through if retailers manage product categories with various brands and compete against each other. As implied by Moorthy (2005)'s analysis of analysis of cross-brand cost pass-through, I find empirical evidence that demand is diverted away from national brands whose costs have increased more than those of to store brands similarly. The insights from Figure 4 are also related to the work in Dubé and Gupta (2008) and Besanko, Dubé, and Gupta (2005) documenting cross-brand cost pass-through.

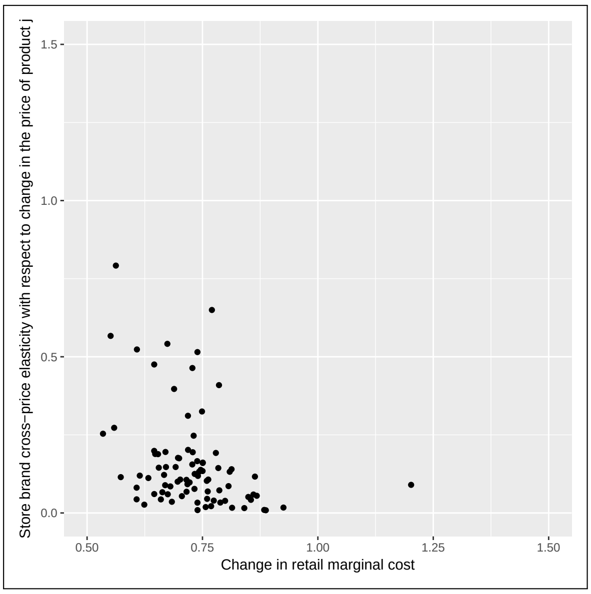
Table 8: Demand diversion through margin adjustment

| | Share of store brand sales (%) | | Change in profits |
|---|---|---|---|
| Retail chain | Constant margins | Adjusted margins | (%) |
| Retailer 1 | 100.00 | 100.00 | 0.57 |
| Retailer 2 | 100.00 | 100.00 | 2.33 |
| Retailer 3 | 5.07 | 8.86 | 4.73 |
| Retailer 4 | 30.24 | 36.21 | 2.17 |
| Retailer 5 | 5.99 | 9.16 | 2.85 |
| Retailer 6 | 19.81 | 24.73 | 2.51 |
| Retailer 7 | 5.90 | 7.92 | 10.15 |
| Retailer 8 | 5.26 | 7.48 | 3.34 |
| Retailer 9 | 34.39 | 42.41 | 5.07 |
| Retailer 10 | 3.40 | 4.75 | 4.78 |
| Retailer 11 | 13.80 | 15.91 | 1.11 |

Another question of this paper is how the positioning of retailers' store brand dampens the increase in wholesale prices. Retail chains can strategically use their store brands to divert demand to them if wholesale prices for manufacturer brands increase more than the marginal cost of supplying their store brands. The basic idea is to proxy store brand positioning in the product space by measuring the store brand cross-price elasticity within a retail chain. This approach follows Sayman et al. (2002) who conceptualize store brand positioning as cross-price effect (price substitutability) between national brand and store brand. The conjecture is that a store brand which is perceived as a close substitute (cross-price elasticity approaches one) dampens the increase in retailers' marginal costs of supplying that national brand compared to a retailer that stocks a rather differentiated store brand. As we can assume that the marginal costs of retailing is the same across all brands within a retail chain (or at least very similar), the main source of differences in retailers' marginal costs increases are the wholesale prices.

Figure 5 tries to establish this link by plotting the cross-price elasticity in terms of percentage change in store brand market share within a retail chain for a one percentage increase of price for product $j$ against the change in retailers' marginal costs for a national brand. We observe a clear tendency that a higher store brand cross-price elasticity is associated with a lower increase in retailers' marginal costs for national brands. This relationship is more formally examined in the regression model in Table 9. I regress the measured changes in retailers' marginal costs for national brands, both in absolute levels and log levels, on the store brand cross-price elasticity controlling for brand fixed effects and retailer fixed effects. The brand fixed effects intend to control for the wholesale price increase of that manufacturer in general. The retailer fixed

Figure 5: Changes in retailers' national brand marginal costs and cross-price elasticity with respect to their store brand




effects capture whether retailing marginal costs increased across all coffee brands in general for that retail chain. For this reason, the coefficient on store brand cross-price elasticity measures how a store brand with a high cross-price elasticity dampens the marginal cost increase of the respective national brand everything else equal. For this setting, we observe that a store brand that is perceived as a close substitute (cross-price elasticity approaches one) dampens the increase in wholesale prices for national brands by around 0.10 EUR (absolute level estimation) or approximately 16% (log level estimation), on average. While this effect might look negligible at first glance, it becomes economic relevant if we consider the high sales volume in that category.

Table 9: Cost increase dampening effect of national brand imitating store brands

| | Specifications | | | | | |
|---|---|---|---|---|---|---|
| | $\Delta c_j$ | | | $ln(\Delta c_j)$ | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Store brand cross-price elasticity | $-0.130^{***}$ | $-0.146^{***}$ | $-0.100^{**}$ | $-0.198^{***}$ | $-0.221^{***}$ | $-0.160^{***}$ |
| | (0.034) | (0.038) | (0.040) | (0.044) | (0.049) | (0.052) |
| Constant | $0.750^{***}$ | $0.713^{***}$ | $0.655^{***}$ | $-0.292^{***}$ | $-0.342^{***}$ | $-0.420^{***}$ |
| | (0.012) | (0.031) | (0.042) | (0.015) | (0.040) | (0.054) |
| Brand fixed effects | No | Yes | Yes | No | Yes | Yes |
| Retailer fixed effects | No | No | Yes | No | No | Yes |
| Observations | 86 | 86 | 86 | 86 | 86 | 86 |
| $R^2$ | 0.152 | 0.221 | 0.361 | 0.195 | 0.260 | 0.400 |
| Adjusted $R^2$ | 0.142 | 0.162 | 0.235 | 0.186 | 0.204 | 0.281 |

*Note:* $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01
Fixed effects are not displayed. Standard errors in parentheses.

The findings have important strategic implications for retail marketers, particularly with respect to a retailer's category management, assortment strategy and store brand design. Store brands work as partial insurance against upstream wholesale price increases, for instance due to price agreements, tacit collusion or mergers waves. Also exchange rate shocks or politically motivated import tariff adjustments can lead to higher wholesale prices for only a subset of brands. These events are rather difficult to predict for a retailer. Retail chains have only for their store brands full control of pricing, product configuration and production conditions. Therefore, it is strategically important to establish a strong store brand positioned close to manufacturer brands in product categories where manufacturer price increases are likelier. Such a "me-too" store brand can be especially valuable if the imitated national brand belongs to the retailer's top-selling products.

Focusing on the threat of upstream collusion, Ivaldi, Jullien, Rey, Seabright, and Tirole (2007) describe several factors that facilitate tacit collusion between manufacturers such as high upstream market concentration, high entry costs, high frequency of sales, stable demand, limited product differentiation, and multi-market contacts. Depending on the product category these factors are more or less prevalent. Naturally, also explicit collusion is easier to maintain under these conditions. There are product categories in the retailing industry where sales are very frequent, demand is well predictable with few multi-category manufacturers that also compete in other product markets: Examples are products such as tooth paste, laundry detergents and chocolate. On the other hand, there are product categories with rather specialized manufactures and infrequent sales. Typically products where the majority of sales occurs in seasons belong to that category such as ice cream, sparkling wine or toys. Retail marketers should have this checklist in mind when they look for an additional argument to introduce a store brand that is positioned close to the national brands.

## 7 Conclusions

Economists and marketing researchers have already been studying the success of store brands for several years. As a novelty, this paper exploits a significant rise in brand manufacturers' wholesale prices to examine the strategic role of retailers' store brands in mitigating market power-driven price increases on upstream markets. It contributes to the growing literature quantifying the strategic role that store brands play for retailers. In addition, the approach builds on recent literature on cross-brand pass-through because the higher rise in wholesale prices for national brands is an asymmetric shock to retailers' marginal costs.

My analysis provides an additional argument for introducing a store brand that is positioned close to national brands. By altering their prices and margins, retail chains can divert more demand towards their own store brands. Multi-brand retailers can increase their category profits by 2-10% if they re-adjust margins after the asymmetric rise in wholesale prices and divert more demand towards their store brands. The power of retailers' store brands in mitigating market power-driven price increases on upstream markets depends on the store brand positioning. A store brand that is perceived as a close substitute (cross-price elasticity approaches one) dampens the increase in wholesale prices for national brands by approximately 16%, on average. This has direct policy implications for retailers' store brand management and marketing strategy across product categories. Because store brands can mitigate manufacturer wholesale price increases, establishing a strong store brand in product categories where manufacturer price increases are more likely is strategically important.

The strategic role of store brands is also important with respect to the recurring disputes between retail chains and large manufacturing companies over price conditions, which lead to retailers' removing products from their assortment on a temporary basis. A store brand that is perceived as a close substitute to the removed products can limit consumers' switching to other retail chains. An aspect that should be addressed by future research. In addition, further research should examine in detail how to design store brands, for instance through similar appearance or proximity in the product characteristic space, to optimally utilize their power to mitigate the effect of brand manufacturer price increases or branded product removals.

## References


Ailawadi, K. L. and B. Harlam (2004). An empirical analysis of the determinants of retail margins: the role of store-brand share. *Journal of Marketing 68*(1), 147–165.

Ailawadi, K. L., K. Pauwels, and J.-B. E. Steenkamp (2008). Private-label use and store loyalty. *Journal of Marketing 72*(6), 19–30.

Allenby, G. M., J. Brazell, J. Howell, and P. E. Rossi (2014). Economic valuation of product features. *Quantitative Marketing and Economics 12*(4), 421–456.

Berry, S., J. Levinsohn, and A. Pakes (1995). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–890.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*, 242–262.

Besanko, D., J.-P. Dubé, and S. Gupta (2005). Own-brand and cross-brand retail pass-through. *Marketing Science 24*(1), 123–137.

Besanko, D., S. Gupta, and D. Jain (1998). Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science 44*(11-part-1), 1533–1547.

Bettendorf, L. and F. Verboven (2000). Incomplete transmission of coffee bean prices: evidence from the netherlands. *European Review of Agricultural Economics 1*, 1–16.

Bonnet, C., P. Dubois, S. B. Villas Boas, and D. Klapper (2013). Empirical evidence on the role of nonlinear wholesale pricing and vertical restraints on cost pass-through. *Review of Economics and Statistics 95*(2), 500–515.

Bonnet, C. and V. Réquillart (2013). Tax incidence with strategic firms in the soft drink market. *Journal of Public Economics 106*, 77–88.

Bonnet, C. and S. B. Villas-Boas (2016). An analysis of asymmetric consumer price responses and asymmetric cost pass-through in the french coffee market. *European Review of Agricultural Economics*.

Bontems, P., S. Monier-Dilhan, and V. Réquillart (1999). Strategic effects of private labels. *European Review of Agricultural Economics 26*(2), 147–165.

Bundeskartellamt (2010, January). Bußgeldverfahren gegen Kaffeeröster wegen Preisabsprachen. *Fallbericht B11-18/08.*

Bundeskartellamt (2014, September). Sektoruntersuchung Lebensmitteleinzelhandel. *Az. B2-15/11.*

Ching, A. T. (2010). A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review 51*(4), 1175–1207.

Chintagunta, P. K., A. Bonfrer, and I. Song (2002). Investigating the effects of store-brand introduction on retailer demand and pricing behavior. *Management Science 48*(10), 1242–1267.

Cohen, M. A. and R. W. Cotterill (2011). Assessing the impact of retailer store brand presence on manufacturer brands in an equilibrium framework. *The Journal of Industrial Economics 59*(3), 372–395.

Corstjens, M. and R. Lal (2000). Building store loyalty through store brands. *Journal of Marketing Research 37*(3), 281–291.

Draganska, M. and D. Klapper (2011). Choice set heterogeneity and the role of advertising: An analysis with micro and macro data. *Journal of Marketing Research 48*(4), 653–669.

Draganska, M., D. Klapper, and S. B. Villas-Boas (2010). A larger slice or a larger pie? an empirical investigation of bargaining power in the distribution channel. *Marketing Science 29*(1), 57–74.

Dubé, J.-P. and S. Gupta (2008). Cross-brand pass-through in supermarket pricing. *Marketing Science 27*(3), 324–333.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics 41*(3), 417–445.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2018). Income and wealth effects on private-label demand: evidence from the great recession. *Marketing Science 37*(1), 22–53.

Dubé, J.-P., G. J. Hitsch, P. E. Rossi, and M. A. Vitorino (2008). Category pricing with state-dependent utility. *Marketing Science 27*(3), 417–429.

Dubois, P., R. Griffith, and M. O'Connell (2017). The effects of banning advertising in junk food markets. *The Review of Economic Studies 85*(1), 396–436.

Erdem, T., S. Imai, and M. P. Keane (2003). Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics 1*(1), 5–64.

Erdem, T. and M. P. Keane (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science 15*(1), 1–20.

Erdem, T., M. P. Keane, and B. Sun (2008). The impact of advertising on consumer price sensitivity in experience goods markets. *Quantitative Marketing and Economics 6*(2), 139–176.

European Commission (2014). The economic impact of modern retail on choice and innovation in the eu food sector.

Frank, R. G. and D. S. Salkever (1992). Pricing, patent loss and the market for pharmaceuticals. *Southern Economic Journal 52*, 165 – 179.

Gabrielsen, T. S. and L. Sørgard (2007). Private labels, price rivalry, and public policy. *European Economic Review 51*(2), 403–424.

Gamerman, D. and H. F. Lopes (2006). *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. CRC Press.

Geyskens, I., K. Gielens, and E. Gijsbrechts (2010). Proliferating private-label portfolios: How introducing economy and premium private labels influences brand choice. *Journal of Marketing Research 47*(5), 791–807.

Goeree, M. S. (2008). Limited information and advertising in the us personal computer industry. *Econometrica 76*(5), 1017–1074.

Hendel, I. and A. Nevo (2006). Measuring the implications of sales and consumer inventory behavior. *Econometrica 74*(6), 1637–1673.

Honka, E., A. Hortaçsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *The RAND Journal of Economics 48*(3), 611–646.

Ivaldi, M., B. Jullien, P. Rey, P. Seabright, and J. Tirole (2007). The economics of tacit collusion: implications for merger control. In *The Political Economy of Antitrust*, pp. 217–239. Emerald Group Publishing Limited.

Keane, M. P. (1997). Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business & Economic Statistics 15*(3), 310–327.

Lamey, L., B. Deleersnyder, M. G. Dekimpe, and J.-B. E. Steenkamp (2007). How business cycles contribute to private-label success: Evidence from the united states and europe. *Journal of Marketing 71*(1), 1–15.

Meza, S. and K. Sudhir (2010). Do private labels increase retailer bargaining power? *Quantitative Marketing and Economics 8*(3), 333–363.

Mills, D. E. (1995). Why retailers sell private labels. *Journal of Economics & Management Strategy 4*(3), 509–528.

Moorthy, S. (2005). A general theory of pass-through in channels with category management and retail competition. *Marketing Science 24*(1), 110–122.

Nakamura, E. and D. Zerom (2010). Accounting for incomplete pass-through. *The Review of Economic Studies 77*(3), 1192–1230.

Nevo, A. (2000). A practitioner's guide to estimation of random-coefficients logit models of demand. *Journal of Economics & Management Strategy 9*(4), 513–548.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica 69*(2), 307–342.

Osborne, M. (2011). Consumer learning, switching costs, and heterogeneity: A structural examination. *Quantitative Marketing and Economics 9*(1), 25–70.

Pachali, M., P. Kurz, and T. Otter (2017). How to generalize from a hierarchical model? *Working Paper*.

Pavlidis, P. and P. B. Ellickson (2017). Implications of parent brand inertia for multiproduct pricing. *Quantitative Marketing and Economics 15*(4), 369–407.

Pesendorfer, M. (2002). Retail sales: A study of pricing behavior in supermarkets. *The Journal of Business 75*(1), 33–66.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research 47*(1), 3–13.

Raju, J. S., R. Sethuraman, and S. K. Dhar (1995). The introduction and performance of store brands. *Management Science 41*(6), 957–978.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005). *Bayesian statistics and marketing.* John Wiley & Sons.

Sayman, S., S. J. Hoch, and J. S. Raju (2002). Positioning of store brands. *Marketing Science 21*(4), 378–397.

Scott Morton, F. M. and F. Zettelmeyer (2004). The strategic positioning of store brands in retailer–manufacturer negotiations. *Review of Industrial Organization 24*(2), 161–194.

Seenivasan, S., K. Sudhir, and D. Talukdar (2016). Do store brands aid store loyalty? *Management Science 62*(3), 802–816.

Seiler, S. (2013). The impact of search costs on consumer behavior: A dynamic approach. *Quantitative Marketing and Economics 11*(2), 155–203.

Villas-Boas, S. (2009). An empirical investigation of the welfare effects of banning wholesale price discrimination. *The RAND Journal of Economics 40*(1), 20–46.

Villas-Boas, S. B. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies 74*(2), 625–652.

Villas-Boas, S. B. and C. Chambolle (2015). Buyer power through producer differentiation. *International Journal of Industrial Organization 43*.

# A Appendix

## A.1 Imputing prices of non-chosen alternatives and constructing choice sets

A common feature of homescan consumer panel data is that we only observe prices of purchased products. In the US, Nielsen homescan data can be linked to Nielsen retail scanner data (at least for several stores).[24] Thus, prices for non-alternatives within a product category at certain stores are observed and can be matched to households' shopping trips. See for instance Erdem et al. (2008) or Hendel and Nevo (2006) as examples of studies that match homescan consumer data with retail scanner data.

For Germany, however, typically only consumer panel data sets are available. Also other researchers often have only the price information within the consumer panel at hand. Studies estimating demand based only on homescan consumer panel data, therefore, have to impute prices of non-chosen alternatives such has been done in the past by Keane (1997) and Erdem and Keane (1996) or recently by Bonnet and Réquillart (2013), Seiler (2013) and Dubois, Griffith, and O'Connell (2017) for instance. Bonnet and Réquillart (2013) compute average prices for each brand-retailer combination over a four-week period in the French soft drink market. Seiler (2013) relies on the national pricing policy of UK supermarket chains in order to construct weekly price series for each brand-package size-supermarket chain in the UK detergent market. Dubois et al. (2017) aggregate 1,800 unique product codes (UPCs) of potato chips in the UK to 37 brand-package size combinations using mean transaction prices.

In my data set of the German ground coffee consumers, approximately 95% of purchases are 500g package sizes. Therefore, I ignore package sizes other than 500g and join them to the outside good of fringe brands and purchases at other retail chains. I note, however, a different type of differentiation within a coffee-brand-retailer combination. Some brands further differentiate their coffee products between a standard version and a mild version. Typically, the mild version has a higher price. For this reason I define a product as a coffee-brand-mildness-retailer combination offered at one of the eleven retail chains. The procedure I use to impute prices of non-chosen alternatives and to construct the corresponding choice sets is as follows:

1. Compute the weekly median price for each coffee brand-mildness combination at each retail chain based on all households coffee purchases.

[24] Researchers from US departments can obtain US Nielsen data from the Kilts Center. See `https://research.chicagobooth.edu`

2. Infer from these price series the regular price for each coffee brand-mildness combination at each retail chain and define it as price ceiling.

3. Fill in empty weeks for each coffee brand-mildness combination at each retail chain with the last observed regular price as it is very unlikely that any price discount occurs during that week if the brand is not purchased at all.

4. Match these "complete" choice sets with each household's coffee purchase occasion. Construct a choice variable that denotes which coffee brand is purchased and replace the corresponding price (if it is not equal) with the household's recorded price.

### A.2 MCMC sampler and prior settings

Pachali et al. (2017) propose a Markov Chain Monte Carlo (MCMC) algorithm for hierarchical multinomial logit models with normal mixtures, based on the work in Rossi et al. (2005), which improves the sampling of posterior distributions for models with sign and/or order constraints on some coefficients. The following description heavily borrows from their work and only adapts to the current application on ground coffee brands. (Please consult Pachali et al. (2017) for more detailed information on how their MCMC sampler works.)
Parameters in the hierarchical prior from Equation 3 can be distinguished between $k_c$ constrained and $k_{uc}$ unconstrained coefficients for each household $i$ (conditional on their component membership $\text{ind}_i = s$)

$$\theta_i^* = \begin{pmatrix} \theta_i^{*c} \\ \theta_i^{*uc} \end{pmatrix} \sim N\left( \begin{pmatrix} \mu_{c_s}^* \\ \Gamma'_s \mu_{c_s}^* + z_s \end{pmatrix}, \begin{pmatrix} V_s^* & V_s^* \Gamma_s \\ \Gamma'_s (V_s^*)' & \Gamma'_s V_s^* \Gamma_s + \Sigma_s \end{pmatrix} \right), \tag{14}$$

where $\theta_i^{*c} = (\alpha_i^*)'$ and $\theta_i^{*uc} = \left(\beta_{i,Brand\ 1}^*, \ldots, \beta_{i,Brand\ 7}^*, \psi_{i,2}^*, \ldots, \psi_{i,11}^*\right)'$ holds in my application. The set of parameters $\left\{(z_s, \Gamma_s, \Sigma_s), (\mu_{c_s}^*, V_s^*)\right\}$ is characterized through two multivariate regression equations conditional on $N_s$ household parameters, $\{\Theta_s^{*uc}, \Theta_s^{*c}\}$, clustered into each of the $S$ components

$$\begin{aligned} \Theta_s^{*uc} &= \Theta_{z_s}^{*c} \Gamma_{z_s} + U \\ \Theta_s^{*c} &= \iota(\mu_{c_s}^*)' + U_{V_s^*}, \end{aligned} \tag{15}$$

with $vec(U') := u \sim N(0, I_{N_s} \otimes \Sigma_s)$, $U_{V_s^*} := u_{V_s^*} \sim N(0, I_{N_s} \otimes V_s^*)$, $(\Gamma_{z_s}, \Sigma_s)$ being a $(k_c{+}1 \times k_{uc})$ coefficient matrix with the intercept vector $z_s$ included in the first row as well as the $(k_{uc} \times k_{uc})$ variance-covariance matrix of unconstrained coefficients respectively and $(\mu_{c_s}^*, V_s^*)$ are the $k_c$-size mean vector as well as $(k_c \times k_c)$ variance-covariance matrix of constrained coefficients respectively. $\iota$ denotes a $(N_s \times 1)$-vector of 1's.

The MCMC sampler that Pachali et al. (2017) propose, is a standard "Gibbs"-style sampler with an RW-Metropolis step to draw individual level parameters $\{\theta_i^*\}$ similar to the one described in Rossi et al. (2005). Their modification is a two-stage update of the parameters entering the hierarchical prior. More specifically, the sampler draws from the following conditionals in each iteration (omitting subjective prior parameters for simplicity)

1. $\theta_i^* | (\mu^*_{c_{\text{ind}_i}}, V^*_{\text{ind}_i}), (\Gamma_{z_{\text{ind}_i}}, \Sigma_{\text{ind}_i}), y_i,\ i = 1, \ldots, N$
2. $\{\Gamma_{z_s}, \Sigma_s\} \,|\, \{\Theta_s^{*uc}, \Theta_s^{*c}\}, \{\text{ind}_i\}$
3. $\{\mu^*_{c_s}, V_s^*\} \,|\{\Theta_s^{*c}\}, \{\text{ind}_i\}$

This approach allows Pachali et al. (2017) to specify subjective priors of unconstrained and constrained coefficients separately from each other. This is necessary as the two represent distinct distributions on the re-transformed $\theta$-space. They use the natural conjugate prior to perform step 2 and the conditionally conjugate prior to perform step 3 of the MCMC sampler. More specifically,

$$\begin{aligned} p(\Gamma_{z_s}, \Sigma_s) &= p(\Gamma_{z_s}|\Sigma_s)p(\Sigma_s), \\ vec(\Gamma_{z_s})|\Sigma_s &\sim N(\bar{\gamma}_z, \Sigma \otimes A^{-1}_{\Gamma_z}) \\ \Sigma_s &\sim IW(\nu_\Sigma, \bar{\Sigma}) \text{ and} \\ p(\mu^*_{c_s}, V_s^*) &= p(\mu^*_{c_s})p(V_s^*), \\ \mu^*_{c_s} &\sim N(\bar{\mu}^*_c, A^{-1}_{\mu^*_c}) \\ V_s^* &\sim IW(\nu_{V^*}, \bar{V}^*) \end{aligned} \tag{16}$$

Explicit posteriors associated with these priors can be found in Pachali et al. (2017). The conditionally conjugate prior implies that mean and variance-covariance matrix are a priori independent which allows it to affect $\mu^*_{c_s}$ more explicitly through $A^{-1}_{\mu^*_c}$. I use a standard weakly informative subjective prior for the parameters entering the hierarchical prior of unconstrained coefficients, $\bar{\gamma}_z,\ A_{\Gamma_z},\ \nu_\Sigma,\ \bar{\Sigma}$. Note that these priors mainly affect posterior inference of $\theta_i^{*uc} = \left(\beta^*_{i,Brand\ 1}, \ldots, \beta^*_{i,Brand\ 7}, \psi^*_{i,2}, \ldots, \psi^*_{i,11}\right)'$. An "informative" specification for price parameter is used that enters the hierarchical prior of constrained coefficients, mainly affecting posterior inference of $\theta_i^{*c} = (\alpha_i^*)'$. More specifically, $\bar{\mu}^*_c = \left(0\right)'$ and $A_{\mu^*_c} = \text{diag}\left(1/4\right)$ for all mixture models.

The subjective priors entering the Inverted Wishart prior for $V_s^*$ imply $\nu_{V^*} = 50(50, 50)$ as well as $\bar{V}^* = c^* \nu_{V^*} I_{k_c}$ with $c^* = 0.5(0.4, 0.3)$ for mixture models with $S = 1(5, 10)$ components respectively where $I_{k_c}$ is the identity matrix of dimension $k_c \times k_c$.

The prior specification on the price coefficient may seem very informative or restrictive with $A_{\mu_c^*} = 1/4$. This prior, however, is set on the log-transformed space and a standard specification would imply an unreliably high prior variance of $\theta_i^{*c}$. Also the high values of $\nu_{V^*} = 50(50, 50)$ might seem restrictive. According to Allenby et al. (2014), however, low values of $\nu$ lead to very thick tails and may allow extremely small price coefficients. In fact, with lower values of $\nu_{V^*}$ one would obtain some individuals with almost infinite price sensitivity (because they always buy the cheap store brands), which is implausible. While dropping these households is not justifiable, I argue that setting a more restrictive prior and increasing the values of $\nu_{V^*}$, which shrinks such households towards more reliable estimates in a hierarchical model is an appropriate alternative. Furthermore, Allenby et al. (2014) describe that $\nu$ can be interpreted as the effective sample size which provides the foundation of the prior. Compared to the estimation sample size of 1000 households, the values of $\nu_{V^*} = 50(50, 50)$ are still relatively uninformative as the sample size is almost 20 times larger than the prior degrees of freedom. In comparison, Allenby et al. (2014) use a prior specification where the sample size is around 10 times larger than the prior degrees of freedom and also regard their specification as relatively uninformative.

## A.3 Additional Tables

Table A.1: Control Function: First stage regression

| Explanatory Variables | CF Regression | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Price promotion | -0.188*** | -0.387*** | -0.424*** | -0.404*** |
| | (0.05) | (0.02) | (0.01) | (0.01) |
| Mild | | 0.086*** | 0.085*** | 0.085*** |
| | | (0.02) | (0.02) | (0.02) |
| Raw coffee bean price X Brand 1 | | | 3.404*** | 1.420*** |
| | | | (0.08) | (0.18) |
| Raw coffee bean price X Brand 2 | | | 3.080*** | 0.921*** |
| | | | (0.17) | (0.12) |
| Raw coffee bean price X Brand 3 | | | 3.451*** | 0.822*** |
| | | | (0.07) | (0.18) |
| Raw coffee bean price X Brand 4 | | | 3.607*** | 1.309*** |
| | | | (0.07) | (0.12) |
| Raw coffee bean price X Brand 5 | | | 3.488*** | 1.449*** |
| | | | (0.06) | (0.15) |
| Raw coffee bean price X Brand 6 | | | 2.037*** | 1.223*** |
| | | | (0.06) | (0.10) |
| Raw coffee bean price X Brand 7 | | | 3.935*** | 1.753*** |
| | | | (0.03) | (0.08) |
| 1. price increase: Brand 1 | | | | 0.312*** |
| | | | | (0.05) |
| 1. price increase: Brand 2 | | | | 0.308*** |
| | | | | (0.03) |
| 1. price increase: Brand 3 | | | | 0.484*** |
| | | | | (0.05) |
| 1. price increase: Brand 4 | | | | 0.355*** |
| | | | | (0.03) |
| 1. price increase: Brand 5 | | | | 0.261*** |
| | | | | (0.04) |
| 1. price increase: Brand 6 | | | | -0.120** |
| | | | | (0.06) |
| 1. price increase: Brand 7 | | | | 0.388*** |
| | | | | (0.03) |
| 2. price increase: Brand 1 | | | | 0.603*** |
| | | | | (0.05) |
| 2. price increase: Brand 2 | | | | 0.675*** |
| | | | | (0.06) |
| 2. price increase: Brand 3 | | | | 0.786*** |
| | | | | (0.05) |
| 2. price increase: Brand 4 | | | | 0.709*** |
| | | | | (0.04) |
| 2. price increase: Brand 5 | | | | 0.644*** |
| | | | | (0.03) |
| 2. price increase: Brand 6 | | | | 0.310*** |
| | | | | (0.03) |
| 2. price increase: Brand 7 | | | | 0.648*** |
| | | | | (0.02) |
| Constant | 3.217*** | 2.831*** | 3.096*** | 3.051*** |
| | (0.06) | (0.01) | (0.01) | (0.01) |
| Sample size | 12419568 | 12419568 | 12419568 | 12419568 |
| R squared | 0.021 | 0.508 | 0.882 | 0.911 |
| Month fixed effects | No | Yes | Yes | Yes |
| Brand fixed effects | No | Yes | Yes | Yes |
| Retailer fixed effects | No | Yes | Yes | Yes |

Note:* p<0.1, ** p<0.05, *** p<0.01. Fixed effects are not displayed
Standard errors in parentheses are clustered by product.

### A.3.1 Demand estimation without control function approach

Table A.2: Quantiles and first two moments of the the marginal posterior densities for the one component model

| Coefficients | Quantiles | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5% | 25% | 50% | 75% | 95% | Mean | Stand. Dev. |
| Price (EUR/500g) | -8.095 | -3.427 | -1.886 | -1.039 | -0.438 | -2.792 | 3.071 |
| Mild | -5.127 | -2.411 | -0.538 | 1.335 | 4.055 | -0.538 | 2.786 |
| Brand 1 | -15.726 | -11.937 | -9.338 | -6.755 | -3.028 | -9.353 | 3.856 |
| Brand 2 | -17.493 | -13.557 | -10.886 | -8.224 | -4.391 | -10.906 | 3.984 |
| Brand 3 | -16.391 | -12.608 | -10.014 | -7.446 | -3.751 | -10.039 | 3.847 |
| Brand 4 | -16.320 | -12.850 | -10.493 | -8.151 | -4.791 | -10.515 | 3.508 |
| Brand 5 | -17.791 | -14.119 | -11.634 | -9.176 | -5.640 | -11.665 | 3.697 |
| Brand 6: SB | -19.875 | -15.601 | -12.682 | -9.780 | -5.638 | -12.708 | 4.335 |
| Brand 7 | -17.774 | -13.326 | -10.291 | -7.269 | -2.936 | -10.314 | 4.512 |
| Retailer 1 | -3.806 | 1.054 | 4.418 | 7.788 | 12.667 | 4.422 | 5.010 |
| Retailer 2 | -5.046 | 1.005 | 5.217 | 9.458 | 15.612 | 5.246 | 6.280 |
| Retailer 3 | -5.712 | -1.515 | 1.377 | 4.278 | 8.511 | 1.387 | 4.329 |
| Retailer 4 | -2.393 | 1.615 | 4.394 | 7.190 | 11.293 | 4.415 | 4.160 |
| Retailer 5 | -6.441 | -1.483 | 1.961 | 5.394 | 10.433 | 1.966 | 5.129 |
| Retailer 6 | -8.090 | -3.197 | 0.187 | 3.583 | 8.587 | 0.208 | 5.072 |
| Retailer 7 | -6.565 | -2.223 | 0.766 | 3.759 | 8.127 | 0.773 | 4.468 |
| Retailer 8 | -6.578 | -2.136 | 0.969 | 4.076 | 8.610 | 0.979 | 4.620 |
| Retailer 9 | -6.730 | -1.775 | 1.631 | 5.070 | 10.106 | 1.653 | 5.117 |
| Retailer 10 | -5.154 | -0.749 | 2.299 | 5.352 | 9.809 | 2.307 | 4.549 |

### A.3.2 Demand estimation including a store brand trend

Table A.3: Quantiles and first two moments of the the marginal posterior densities for the one component model

| Coefficients | Quantiles | | | | | | |
|---|---|---|---|---|---|---|---|
| | 5% | 25% | 50% | 75% | 95% | Mean | Stand. Dev. |
| Price (EUR/500g) | -7.563 | -3.087 | -1.654 | -0.889 | -0.363 | -2.537 | 2.940 |
| Mild | -5.134 | -2.443 | -0.570 | 1.291 | 3.987 | -0.573 | 2.772 |
| Brand 1 | -16.210 | -12.359 | -9.729 | -7.104 | -3.335 | -9.744 | 3.916 |
| Brand 2 | -18.025 | -13.955 | -11.157 | -8.375 | -4.343 | -11.172 | 4.163 |
| Brand 3 | -17.115 | -13.119 | -10.392 | -7.678 | -3.782 | -10.413 | 4.054 |
| Brand 4 | -16.769 | -13.181 | -10.740 | -8.316 | -4.814 | -10.759 | 3.636 |
| Brand 5 | -18.411 | -14.544 | -11.926 | -9.305 | -5.527 | -11.938 | 3.915 |
| Brand 6: SB | -19.857 | -15.706 | -12.868 | -10.058 | -5.997 | -12.893 | 4.217 |
| Brand 7 | -18.406 | -13.797 | -10.653 | -7.510 | -2.970 | -10.662 | 4.688 |
| Retailer 1 | -3.371 | 1.440 | 4.763 | 8.105 | 12.969 | 4.778 | 4.967 |
| Retailer 2 | -5.377 | 1.020 | 5.437 | 9.860 | 16.304 | 5.443 | 6.597 |
| Retailer 3 | -5.702 | -1.345 | 1.673 | 4.707 | 9.122 | 1.688 | 4.509 |
| Retailer 4 | -2.344 | 1.818 | 4.709 | 7.590 | 11.846 | 4.721 | 4.313 |
| Retailer 5 | -6.217 | -1.183 | 2.298 | 5.809 | 10.872 | 2.313 | 5.200 |
| Retailer 6 | -7.827 | -2.974 | 0.414 | 3.805 | 8.788 | 0.434 | 5.055 |
| Retailer 7 | -6.335 | -2.025 | 0.941 | 3.943 | 8.341 | 0.968 | 4.463 |
| Retailer 8 | -6.497 | -1.920 | 1.251 | 4.406 | 8.997 | 1.246 | 4.712 |
| Retailer 9 | -7.474 | -1.915 | 1.880 | 5.707 | 11.251 | 1.891 | 5.687 |
| Retailer 10 | -5.073 | -0.512 | 2.650 | 5.802 | 10.419 | 2.652 | 4.713 |
| SB trend | -2.795 | -1.182 | -0.065 | 1.051 | 2.665 | -0.066 | 1.661 |
| CF | -2.431 | -1.253 | -0.436 | 0.377 | 1.560 | -0.437 | 1.213 |

### A.3.3 Illustration of inference on individual level showing heterogeneity across households

Table A.4: Individual inference for a low price sensitive household: Quantiles and first two moments of the the marginal posterior densities for the one component model

| Coefficients | | | Quantiles | | | | |
|---|---|---|---|---|---|---|---|
| | 5% | 25% | 50% | 75% | 95% | Mean | Stand. Dev. |
| Price (EUR/500g) | -1.308 | -0.873 | -0.628 | -0.439 | -0.240 | -0.686 | 0.338 |
| Mild | -1.785 | -1.327 | -1.027 | -0.734 | -0.360 | -1.041 | 0.434 |
| Brand 1 | -15.807 | -12.637 | -10.788 | -8.992 | -6.843 | -10.938 | 2.714 |
| Brand 2 | -15.842 | -12.883 | -11.119 | -9.491 | -7.261 | -11.271 | 2.556 |
| Brand 3 | -16.505 | -13.353 | -11.385 | -9.685 | -7.290 | -11.590 | 2.774 |
| Brand 4 | -15.002 | -12.573 | -10.904 | -9.276 | -7.268 | -10.977 | 2.379 |
| Brand 5 | -17.767 | -14.956 | -13.034 | -11.259 | -8.761 | -13.131 | 2.744 |
| Brand 6: SB | -19.717 | -16.691 | -14.649 | -12.719 | -10.112 | -14.742 | 2.876 |
| Brand 7 | -8.890 | -6.396 | -4.682 | -3.304 | -1.754 | -4.948 | 2.208 |
| Retailer 1 | 0.217 | 4.268 | 6.682 | 9.003 | 12.098 | 6.502 | 3.637 |
| Retailer 2 | 2.820 | 6.478 | 8.992 | 11.323 | 14.435 | 8.864 | 3.546 |
| Retailer 3 | 3.377 | 4.882 | 6.246 | 7.934 | 10.404 | 6.507 | 2.183 |
| Retailer 4 | -1.047 | 1.629 | 3.714 | 5.769 | 8.690 | 3.744 | 2.994 |
| Retailer 5 | -5.413 | -2.230 | -0.230 | 1.705 | 4.685 | -0.307 | 3.071 |
| Retailer 6 | 1.334 | 3.012 | 4.355 | 6.142 | 8.746 | 4.645 | 2.273 |
| Retailer 7 | -7.032 | -3.597 | -1.497 | 0.787 | 3.532 | -1.515 | 3.222 |
| Retailer 8 | -4.799 | -1.873 | 0.103 | 2.031 | 4.855 | 0.062 | 2.877 |
| Retailer 9 | -4.024 | -0.564 | 1.670 | 3.924 | 7.291 | 1.666 | 3.458 |
| Retailer 10 | -5.179 | -2.180 | -0.138 | 1.736 | 4.442 | -0.237 | 2.912 |
| CF | -1.276 | -0.497 | 0.033 | 0.591 | 1.422 | 0.046 | 0.809 |
| | | | Individual brand choices made | | | | |
| Brand 1 | Brand 2 | Brand 3 | Brand 4 | Brand 5 | Brand 6 | Brand 7 | Outside good |
| 0 | 0 | 0 | 0 | 0 | 0 | 26 | 6 |

Table A.5: Individual inference for a high price sensitive household: Quantiles and first two moments of the the marginal posterior densities for the one component model

| Coefficients | | | Quantiles | | | | |
|---|---|---|---|---|---|---|---|
| | 5% | 25% | 50% | 75% | 95% | Mean | Stand. Dev. |
| Price (EUR/500g) | -6.812 | -5.528 | -4.648 | -3.779 | -2.705 | -4.672 | 1.257 |
| mild | 0.426 | 1.032 | 1.431 | 1.811 | 2.379 | 1.418 | 0.587 |
| Brand 1 | -2.062 | -0.679 | 0.346 | 1.334 | 2.788 | 0.345 | 1.482 |
| Brand 2 | -1.734 | -1.020 | -0.507 | 0.017 | 0.864 | -0.488 | 0.784 |
| Brand 3 | -2.452 | -1.352 | -0.681 | 0.035 | 0.953 | -0.683 | 1.037 |
| Brand 4 | -4.174 | -3.026 | -2.297 | -1.606 | -0.643 | -2.339 | 1.079 |
| Brand 5 | -4.720 | -3.689 | -2.917 | -2.135 | -1.076 | -2.920 | 1.135 |
| Brand 6: SB | -7.161 | -5.784 | -4.863 | -3.891 | -2.559 | -4.845 | 1.403 |
| Brand 7 | -5.087 | -3.287 | -2.227 | -1.258 | 0.147 | -2.325 | 1.594 |
| Retailer 1 | -9.584 | -6.120 | -4.225 | -2.688 | -0.931 | -4.596 | 2.655 |
| Retailer 2 | -12.688 | -9.428 | -7.426 | -5.544 | -3.137 | -7.585 | 2.866 |
| Retailer 3 | -12.380 | -9.240 | -7.260 | -5.630 | -3.735 | -7.564 | 2.628 |
| Retailer 4 | -6.525 | -4.974 | -3.989 | -3.226 | -2.239 | -4.160 | 1.340 |
| Retailer 5 | -12.201 | -9.438 | -7.549 | -5.961 | -4.195 | -7.813 | 2.473 |
| Retailer 6 | -16.769 | -13.499 | -11.398 | -9.311 | -6.369 | -11.427 | 3.082 |
| Retailer 7 | -12.316 | -9.490 | -7.480 | -5.773 | -3.601 | -7.695 | 2.625 |
| Retailer 8 | -13.429 | -10.052 | -8.161 | -6.399 | -4.385 | -8.394 | 2.735 |
| Retailer 9 | -14.511 | -12.017 | -10.026 | -8.046 | -5.410 | -10.028 | 2.799 |
| Retailer 10 | -10.879 | -8.204 | -6.537 | -5.214 | -3.704 | -6.825 | 2.214 |
| CF | -2.165 | -1.273 | -0.724 | -0.145 | 0.664 | -0.726 | 0.850 |
| | | | Individual brand choices made | | | | |
| Brand 1 | Brand 2 | Brand 3 | Brand 4 | Brand 5 | Brand 6 | Brand 7 | Outside good |
| 1 | 12 | 3 | 2 | 2 | 2 | 0 | 2 |

Figure A.1: Price parameter marginal posterior distribution for two households with different price sensitivities (compared to the unconditional population distribution)