Estimating counterfactuals as a test of structural models: The case of the German coffee cartel

Raphael Johannes Kuhlmann*†

November 14, 2019

Abstract

This paper introduces a way of testing the predictive power of structural models. Based on retail market scanner data for ground coffee in Germany during the period 2002 to 2012, I compare marginal cost estimates for coffee store brands from an out-of-sample prediction to estimates based on a structural model of retail competition. The comparison reveals substantial differences. While the marginal cost estimates using an out-of-sample prediction closely follow the world market prices for coffee beans representing the main driver for changes in the cost of ground coffee, the estimates based on a structural model are considerably higher following a cartel-induced increase in wholesale prices for national brands. Given that the cost of coffee store brands are not affected by the cartel-induced increase in wholesale prices, the increase in marginal costs, based on the structural model, reflects the model's failure to capture retailers' pricing choices in a multiproduct environment.

JEL-Classification: L13, L66

Keywords: Demand estimation, nested logit models, structural models

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

†I thank my advisor Roman Inderst and my colleague Marco Kotschedoff for their valuable comments.

1 Introduction

The prediction of counterfactual outcomes is a key element in various fields of industrial organization and there is an ongoing debate over the accurateness of different prediction methods. This concerns, in particular, the application of so-called structural models. Angrist and Pischke (2010), for instance, favor the application of empirical "treatment effects" approaches and criticize the application of "complex simulation-based analyses," such as the application of structural models, to estimate counterfactual outcomes. Nevo and Whinston (2010) argue instead that there are circumstances in which structural models prove to be particularly useful.

Only a few papers, however, provide empirical evidence on the predictive power of structural models in the field of industrial organization. One of the first papers to compare the results of different estimation approaches using structural models on the one hand and more "direct" empirical techniques on the other is Hausman and Leonard (2002), who compare price estimates following the introduction of a new brand of toilet paper.[1] The focus of other papers evaluating the accurateness of structural models lies on ex-post analyses of mergers.[2] An example of such an ex-post analysis is Peters (2006), who compares the observed post-merger prices of five different airline mergers to estimates based on structural merger simulations.[3] In a more recent study, in the context of the Swedish market for analgesics, Bjoernerstedt and Verboven (2016) compare the actual price increase of a merger with the predicted price increase using a structural merger simulation.[4]

I contribute to the literature by providing a different way to analyze the predictive power of structural models. The basic test idea is to compare marginal cost estimates using a structural model to those resulting from a more direct out-of-sample prediction. The German market for coffee provides an ideal setting to compare different (cost) estimation methods given that the costs (and prices) for (ground) coffee largely depend on observable input factors such as the coffee bean world market prices. Furthermore, focusing on store

[1] Hausman and Leonard (2002) find that one of the three presented structural models produces reasonably similar results compared to the results of the direct approach.

[2] Ashenfelter et al. (2009) provide an overview of the few empirical retrospective studies for merger forecast evaluation.

[3] Peters (2006) finds that standard simulation methods do not provide accurate forecasts.

[4] Bjoernerstedt and Verboven (2016) find that the predicted price increase after the merger "is of a similar magnitude" compared to the actual price increase.

brands allows me to infer notably changes in retail marginal costs without the need to infer wholesale prices, thereby excluding one layer of possible misspecification. Another important feature of the market for coffee, making it particularly suitable for testing the predictive power of structural models, is the presence of a manufacturer cartel between 2000 and 2008, resulting in a significant variation in prices. Within this framework of observable input cost factors and sufficient variation in prices, I will compare the results of marginal cost estimates using a structural model to an out-of-sample prediction. The difference between the results of the cost estimates based on the structural model and the out-of-sample prediction will provide a measure of the predictive power of the selected structural model.

The paper is organized as follows. Section 2 provides an overview of the German market for coffee and the data used for the analysis. Section 3 presents the results of a demand estimation. Section 4 introduces a structural model to back out retailers' marginal costs from retail prices using the parameters of the demand estimation. Section 5 introduces a test of the predictive power of the used structural model of retailer pricing by establishing a comparison between the results of an out-of-sample prediction and those based on the structural model alone.

2 Data and the Market for Coffee

In my analysis, I make use of data of the German coffee market from 2002 to 2012 provided by the Nielsen Homescan panel. The panel encompasses information on the coffee purchases of up to 13,500 households.[5] While the panel also covers data on other types of coffee (for example, whole bean or instant coffee), in my analysis, I focus on 500g packages of ground coffee which cover around 50% of total consumption.[6]

[5]The Nielsen Homescan panel contains information on the expenditures of German households on fast-moving consumer goods. In addition to monitoring each participant's expenditures using a specific scanning device, the panel also monitors further characteristics on each participant, such as age and income. The panel used in this analysis covers expenditure data for the period 2002 to 2012. While the number of participants varies from 6,000 in 2002 to around 13,500 households in 2012, annual weighting factors are applied to ensure the representativeness of the sample.

[6]The market shares are calculated using standardized cup equivalents. Instant coffee, which accounts for roughly 40% of consumption in the panel, is considered to be part of a separate market and is hence not included in the analysis (which follows the argument of the German competition authority, BKartA (2014)).

In line with other studies using aggregate demand models – see for example Bonnet and Dubois (2010) or Bonnet and Villas-Boas (2016) – I calculate volume-weighted average prices and aggregated market shares by product, for which I define a product as an individual combination of brand, retailer, and retail format (e.g., discounter). Since retailers offer a portfolio of different products with different taste varieties (e.g., regular, mild or decaffeinated) within the same brand, I follow the literature and aggregate over these varieties but control for the share of each variety in the following analysis.[7]

After aggregating over the different varieties, the dataset encompasses a maximum of 44 products sold at the five major retail chains covering 68% or more of all total ground coffee sales in the retail market for each geographical area. The regional segmentation is based on Nielsen's geographical classification separating the regions in Germany into seven (promotional) markets, for which I subsequently calculate volume-weighted average prices and market shares.[8]

With the five leading retail groups accounting for more than 70% of the market, the retail market for coffee is highly concentrated. Similarly, in the upstream market, the four largest brand manufacturers account for roughly two thirds of the overall ground coffee sales of national brands (excluding store brands). These four manufacturers were found guilty by the German competition authority of operating a cartel between 2000 and 2008. The cartel was detected in July 2008. The resulting price variation will subsequently be used in my analysis.

Store brands account for almost 30% of all ground coffee sales. While there is typically not much information available on the production of store brands, the German competition authority conducted a sector inquiry, noting that private label coffee is mainly produced by smaller manufacturers that specialize in private label production, rather than by large national brand manufacturers (see BKartA (2014) for more details).

To provide an overview of the data, Table 1 comprises a separate account of the summary statistics of the key variables for three periods: the full sample period 2002–2012, the post-cartel period 2009–2012, and the cartel period from 2002 to 2008. From this, it becomes evident that the average retail prices as well as the world market prices for

[7]See for example Cohen and Cotterill (2011) or Holler and Rickert (2019).
[8]See for example The Nielsen Company GmbH (2017), p. 67.

raw beans (for 500g packages) were lower in the cartel period compared to the post-cartel period. In Figure 1 I provide a graphical representation of the volume-weighted average price for both the store brands and national brands (in cents per cup) as well as the world market prices for Arabicas and Robustas, which are the two mainly consumed bean types.

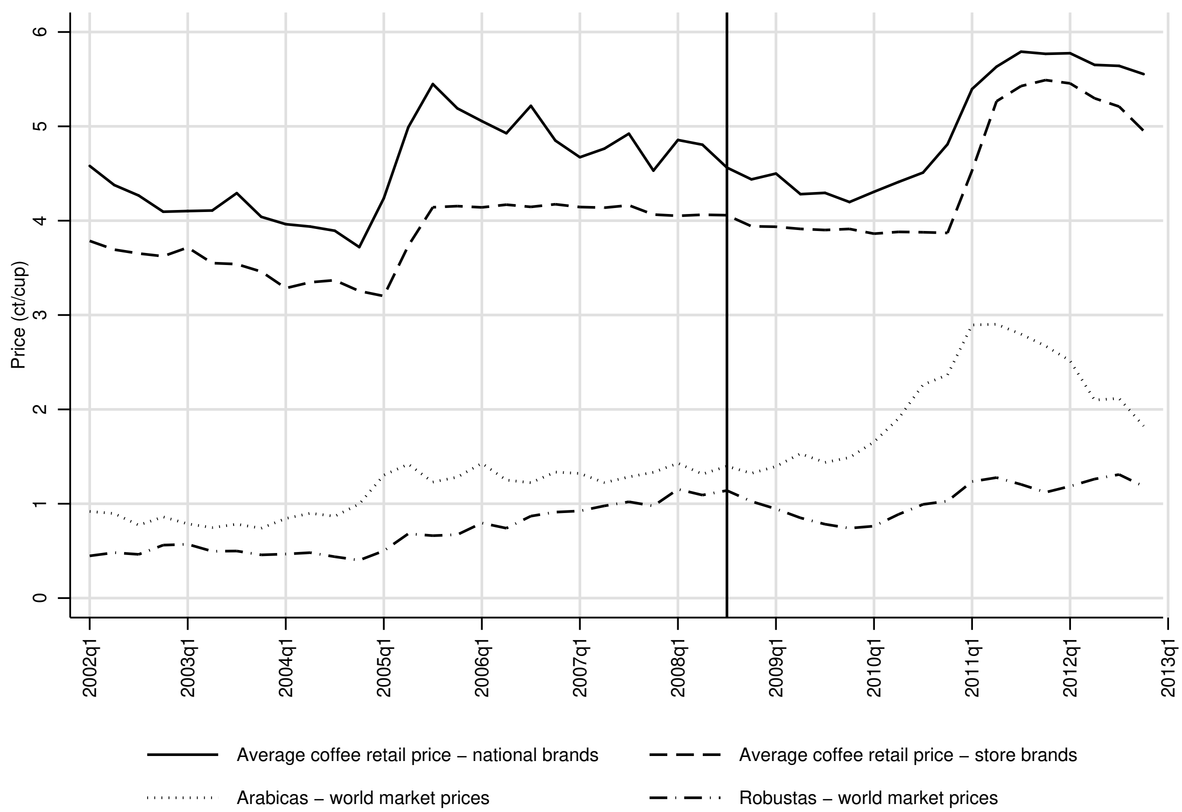
With respect to the development of prices in Figure 1, I would like to emphasize the following two main observations. First, in the post-cartel period, I observe how both the retail prices of store brands as well as the retail prices of national brands react in a similar magnitude to increases in the world market prices for coffee beans. Second, during 2004 and 2005 (i.e., in the cartel period), I observe a substantial increase in retail prices (national brands and store brands), while the world market prices for beans do not increase to the same extent. The price increase in 2005 followed after announcements made by the manufacturers of the coffee cartel that there would be a substantial increase in the wholesale prices for the leading six national brands.[9] Accordingly, it becomes evident that retailers not only passed on the wholesale price increases of national brand products to their customers, but also the increased prices for store brands, although there was no substantial increase in costs, given the description of private label coffee production. Within the structural model analyzed below, this must be interpreted as an optimal reaction of retailers to increase their margin on store brands in the presence of higher retail prices for national brands. Eventually, the test of the model is thus whether and to what extent this multiproduct price adjustment can be accounted for.

[9]See BKartA (2010).

Table 1: Summary statistics of key variables: During and post-cartel

	Mean	Stand. Dev.	Min	Max	N
2002–2012					
Price EUR/500g	3.25	0.59	1.91	6.01	10164
Mild share	0.43	0.28	0.00	1.00	10164
Organic label share	0.01	0.03	0.00	1.00	10164
Decaf share	0.07	0.11	0.00	1.00	10164
Arabica bean price EUR/500g	1.06	0.43	0.53	2.07	10164
Robusta bean price EUR/500g	0.60	0.20	0.29	0.94	10164
Discounter	0.20	0.40	0.00	1.00	10164
Store brand	0.20	0.40	0.00	1.00	10164
# of products in store	5.97	2.07	1.00	8.00	10164
Cartel (2002–2008)					
Price EUR/500g	3.11	0.51	1.91	5.17	6468
Mild share	0.43	0.28	0.00	1.00	6468
Organic label share	0.00	0.03	0.00	1.00	6468
Decaf share	0.08	0.12	0.00	1.00	6468
Arabica bean price EUR/500g	0.80	0.18	0.53	1.02	6468
Robusta bean price EUR/500g	0.51	0.18	0.29	0.82	6468
Discounter	0.20	0.40	0.00	1.00	6468
Store brand	0.20	0.40	0.00	1.00	6468
# of products in store	5.97	2.07	1.00	8.00	6468
Post-Cartel (2009–2012)					
Price EUR/500g	3.50	0.65	2.45	6.01	3696
Mild share	0.43	0.29	0.00	1.00	3696
Organic label share	0.01	0.03	0.00	0.53	3696
Decaf share	0.06	0.10	0.00	1.00	3696
Arabica bean price EUR/500g	1.51	0.37	1.00	2.07	3696
Robusta bean price EUR/500g	0.75	0.14	0.53	0.94	3696
Discounter	0.20	0.40	0.00	1.00	3696
Store brand	0.20	0.40	0.00	1.00	3696
# of products in store	5.97	2.07	1.00	8.00	3696

Figure 1: Average retail and world market prices for ground coffee and coffee beans



3 Demand

In the following section I present a discrete choice demand model providing the framework to estimate demand parameters with aggregate market-level data. The estimated demand parameters will later be used as input parameters for the estimation of marginal costs using a structural model of retail competition.

The discrete-choice literature encompasses various methods to estimate demand parameters using aggregate market-level data.[10] In my analysis, I will harness a simple logit as well as a nested logit demand model. In the following section, I consequently derive the nested logit demand model. The simple logit demand model follows from the nested logit specification.

3.1 Demand Specification

Following the discrete choice literature (notably Berry (1994)) I model consumer i's conditional indirect utility for product $j = 1, ..., J$ in a nested logit framework. Consumer i's conditional indirect utility from any product j is specified by:

$$u_{ij} = x_j\beta - \alpha p_j + \xi_j + (1 - \sigma)\epsilon_{ij}, \tag{1}$$

where x_j is the vector of observed product characteristics, p_j is its price, ξ_j is an unobserved (by the econometrician) characteristic for product j and ϵ_{ij} is an individual-specific component of utility. The vectors β, and α are the corresponding parameters to be estimated. I can write the model as

$$u_{ij} = \delta_j + (1 - \sigma)\epsilon_{ij}, \tag{2}$$

where $\delta_j = x_j\beta - \alpha p_j + \xi_j$ is the mean or common part of the consumers' utility. Since the econometrician only observes the quantities demanded, rather than the utilities, the mean utility of the outside good is normalized to 0 (i.e., $\delta_j = 0$). For the nested logit model specification, it is assumed that the individual-specific error term ϵ_{ij} follows a generalized

[10] Most notably by Berry (1994), Berry et al. (2004), or Petrin (2002).

extreme value distribution with a cumulative distribution of the following form:

$$F(\epsilon_{ij}) = -exp(-\sum_{g=1}^{G}(\sum_{j\in g}\exp(-\epsilon_{ij}/(1-\sigma))^{(1-\sigma)}). \tag{3}$$

The resulting nested logit individual choice probability of consumer i choosing product j out of product group g, which, at the aggregate level, equals the market share function of product j from group g, is given by:

$$\bar{s}_{j|g}(\delta,\sigma) = (exp(\delta_j/(1-\sigma)))/D_g, where \tag{4}$$

$$D_g \equiv \sum_{j\in g}\exp(\delta/(1-\sigma)). \tag{5}$$

Similiarly, the probability of choosing one of the products in group g is given by

$$\bar{s}_g(\delta,\sigma) = \frac{D_g^{(1-\sigma)}}{\sum_g D_g^{(1-\sigma)}}. \tag{6}$$

Consequently, the market share of product j follows from equations 4 and 6:

$$s_j = \bar{s}_{j|g}(\delta,\sigma)\bar{s}_g(\delta,\sigma) = \frac{exp(\delta/(1-\sigma))}{D_g^{\sigma}(\sum_g D_g^{(1-\sigma)})}. \tag{7}$$

The market share for the outside good (with $\delta_j = 0$) is given by

$$s_0 = 1/exp(D_g^{\sigma}). \tag{8}$$

Taking logs of the market shares and subtracting the market share of the outside good from product j yields

$$ln(s_j) - ln(s_0) = \delta_j/(1-\sigma) - \sigma ln(D_g). \tag{9}$$

Taking the log of the group share (equation 6), plugging into equation 9 and solving for δ_j yields

$$\delta_j(s,\sigma) = ln(s_j) - ln(\bar{s}_{j|g}) - ln(s_0). \tag{10}$$

Plugging 10 into $\delta_j = x_j\beta - \alpha p_j + \xi_j$ gives the final demand estimation specification

$$ln(s_j) - ln(s_0) = x_j\beta - \alpha p_j + \sigma ln(\bar{s}_{j|g}) + \xi_j, \tag{11}$$

where estimates of β, α, and σ can be obtained using a linear regression model. As will be discussed later, prices and within-market shares are endogenous and the corresponding coefficients will therefore be estimated using an Instrumental Variable (IV) approach.

The corresponding price elasticities of demand, which measure the variation in demand in response to a change in price, follow from the derivatives of equation 11 and are given by:

$$\eta_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j} = \begin{cases} -\alpha p_j \frac{1-\sigma s_{j|g}-(1-\sigma)s_j}{1-\sigma} & \text{, if j=k,} \\ \alpha p_k \frac{\sigma s_{k|g}+(1-\sigma)s_k}{1-\sigma} & \text{, if } j \neq k \text{ and } j,k \in g \\ \alpha p_k s_k & \text{, otherwise} \end{cases} \tag{12}$$

3.2 Identification and Results

3.2.1 Identification

As briefly mentioned in the previous section, retail prices and the logarithm of the within-market shares are likely endogenous and the corresponding coefficients should therefore be estimated using an IV approach. The endogeneity of prices arises due to the fact that the econometrician does not observe all (relevant) product characteristics that are known to consumers and producers. These product characteristics are, however, taken into account when producers set prices. Similarly, prices and market shares are (usually) determined simultaneously, resulting in the endogeneity of within-market shares.

As a consequence, I will employ the world market prices of Arabica and Robusta beans as instruments in the following analysis to deal with the endogeneity of prices. Exogenous cost-shifts, such as the world market prices, are naturally good instruments, since these are correlated with prices (satisfying the rank condition) and are independent of the error term (i.e., the unobserved product characteristics). Following, amongst others, Villas-Boas

(2009) and Bonnet et al. (2013), I interact the world market prices with product indicator variables to obtain product-specific instruments.

In addition, I will employ the sum of rival firms' brand characteristics within a nest (i.e., the sum of the shares of the taste varieties of all other firms in a given nest), and the number of products as instruments.[11] The intuition behind using rival firms' brand characteristics is that the more similar the characteristics of two products are, the more these two products are substitutable and hence good instruments (for the within-market shares). The characteristics of other firms are considered to be appropriate instruments, since these characteristics are excluded from the utility function (see equation 1, u_{ij} does not depend on x_k if $j \neq k$).[12]

3.2.2 Results

The parameter estimates of both the logit and the nested logit demand model are presented in Table 2. The first column depicts the results of the logit estimation. The second column shows the result (of the logit estimation) using IVs. The third column displays the results of the nested logit specification. The fourth column shows the results (of the nested logit estimation) using an IV approach. The price coefficient in all four specifications has the expected sign and is significantly different from zero. For both models (i.e., simple logit and nested logit), I observe that instrumenting the price variable leads to a higher price coefficient in absolute terms. Thus, our instruments seem to be able to reduce the endogeneity bias. Similarly, the coefficient on the natural logarithm of market shares within a nest decreases after employing instrumental variables. The significance of the relevant tests for endogeneity, under- and overidentification also suggest that the selected instruments reduce the endogeineity bias.

Table 3 shows a comparison of the resulting average price elasticities using the parameter estimates from the empirical results of the different demand models. I observe that products within the same nest, i.e., within the same retailer, exhibit higher cross-price elasticities than products offered in different stores. Note that I obtain elasticities of similar magnitude

[11] These instruments follow Berry et al. (1995) and Berry (1994).

[12] The idea to use demand-side instruments as an alternative to cost variables was developed by Berry et al. (1995).

compared to those of other studies with retail coffee data, see for instance Villas-Boas (2009).

Table 2: Empirical results from demand estimation

	(1)	(2)	(3)	(4)
Price CENT/cup	-0.657***	-1.178***	-0.239***	-1.020***
	(-12.29)	(-21.40)	(-6.84)	(-19.76)
Mild share	0.115	0.153**	0.184*	0.175***
	(0.96)	(3.00)	(2.10)	(3.95)
Organic label share	1.439*	2.939***	1.616**	2.942***
	(2.17)	(9.12)	(2.86)	(10.55)
Decaf share	0.0141	0.320***	-0.165	0.250**
	(0.07)	(3.41)	(-1.04)	(3.05)
$ln(s\_j/g)$			0.869***	0.288***
			(26.12)	(10.41)
Constant	-1.606***	0.674**	-1.804***	0.523*
	(-5.03)	(2.67)	(-9.39)	(2.37)
Quarter and Year fixed effects	Yes	Yes	Yes	Yes
Product fixed effects	Yes	Yes	Yes	Yes
Market fixed effects	Yes	Yes	Yes	Yes
Sample size	10164	10164	10164	10164
R^2	0.475	0.448	0.722	0.586

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Average implied price elasticities by demand model

	(1)	(2)	(3)	(4)
Own Price Elasticity	-2.992	-5.362	-6.821	-6.129
	(0.554)	(0.992)	(2.145)	(1.278)
Cross Price Elasticity within nest			1.566	0.438
			(1.482)	(0.417)
Cross Price Elasticity	0.022	0.040	0.008	0.034
	(0.031)	(0.056)	(0.011)	(0.048)

Standard deviation in parentheses

4 Structural model for retail competition

In the following section I derive a structural model which, in combination with the parameters from the demand estimation, provides the framework to infer marginal costs using information on prices, elasticities, and market shares.[13] I consider a Nash-Bertrand retail pricing model in a differentiated product market with several (oligopolistic) retailers selling multiple products. In the model, the retailers' marginal costs, c^r, are assumed to be independent of quantity. Dropping the time-subscript to simplify notation, retailer r's profits are denoted by:

$$\pi_r = \sum_{j \in S_r} \left[p_j - p_j^w - c_j^r \right] s_j(p) D \text{ for r=1,...,R,} \tag{13}$$

where $s_j(p)$ equals the market share of product j, D denotes the potential market for coffee, and S_r is the set of products offered at the retailer. Retailers internalize the effect of a change in prices of one product on the demand of their other products. In order to facilitate the analysis I follow the literature and introduce a JxJ ownership matrix T_r, which equals 1 if the retailer offers both product m and product j (and zero otherwise):

$$T_r = \begin{cases} 1 & \text{if same retailer offers products j and m,} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

[13]The structural model follows the linear pricing model in Villas-Boas (2007).

The resulting first-order condition for each product is

$$s_j + \sum_{m \in S_r} T_r(m,j) \left[p_m - p_m^w - c_m^r\right] \frac{\partial s_m}{\partial p_j} = 0 \text{ for } j = 1, ..., N.. \tag{15}$$

Changing to matrix notation, price-cost margins at the retailer level are thus determined by[14]

$$\underbrace{p - p^w - c^r}_{m_r} = - \left[T_r * \Delta_r\right]^{-1} s(p), \tag{16}$$

where $\Delta_r(i,j) = \frac{\partial s_j}{\partial p_i}$ representing the retailer's response matrix. The response matrix contains first derivatives of the market shares with respect to all retail prices. Note that in the case of store brands, p^w is set to zero, so that I can solve for the retailer's marginal costs for store brands $c^{r,sb}$:

$$c^{r,sb} = p + \left[T_r * \Delta_r\right]^{-1} s(p). \tag{17}$$

Equation 17 shows that I obtain the marginal costs from observed prices and the demand parameters (in the form of the response matrix). The accuracy of the marginal costs calculation therefore largely depends on the results of the demand estimation.

The average of the results of the marginal cost calculation using the structural model can be retrieved from Table 5, in which the average marginal costs across products are displayed separately for the period 2002 to 2008 (which will be denoted as Period I in the following) and from 2009 to 2012 (Period II) for each of the different demand specifications (see columns 3 and 5 depicting the costs inferred from the structural model).

[14]Note that I denote an element-by-element matrix multiplication between two matrices A and B of the same dimension as $A * B$.

5 Testing the structural model

5.1 Procedure

In the following section, in order to test the structural model, I will juxtapose the marginal cost estimates inferred from the presented structural model with those estimates based on an out-of-sample prediction for Period I. The out-of-sample prediction will subsequently be carried out using the marginal cost inferred from the structural model for Period II (2009–2012). In a first step, I will harness the results from the demand estimation to infer the marginal costs for the entire period (see equation 17). Based on the inferred marginal costs for Period II, I will, in a second step, "backcast" the marginal costs, $\widehat{c}^{r,sb}$, for Period I using the results presented in section 5.2. In order to test the results of the structural model, I will, in a final step, compare the backcasted costs to those inferred from the structural model alone.

In order for the comparison to be meaningful and not merely the result of incorrectly backcasted marginal costs, it is essential to obtain good estimates of $\widehat{c}^{r,sb}$. As shown in equation 16, retailers' marginal costs depend on wholesale prices for coffee and retailer-specific marginal costs. As noted above, for store brands it is reasonable to not consider wholesale margins. Hence, focusing on store brands enables me to obtain accurate estimates of the marginal costs without the need for information on wholesale prices. Changes in marginal costs for store brands thus ultimately depend on raw bean prices. The market for ground coffee, and coffee store brands in particular, is therefore particularly suitable for obtaining accurate predictions of the marginal costs for coffee store brands using an out-of-sample prediction.

5.2 Backcasting of store brand cost

In order to subsequently estimate or "backcast" costs of store brands for Period I, I regress retailer costs for store brands from Period II, as retrieved from equation (17), on observable cost-shifters and product characteristics:

$$\overline{c}^{r,sb} = \gamma_j + \sum_{b=1}^{B} \lambda_b \Omega_t + \sum_{r=1}^{R} \phi_r \chi_t + X_j \delta + \epsilon_{jt}, \tag{18}$$

where γ_j is a product-specific constant, Ω_t is a matrix of cost shifters (world market coffee bean prices), χ_t is a matrix of cost-shifters (average retail labor gross salaries in Germany) interacted with retailer indicator variables, and X_j denotes observable product characteristics. When applied to "backcast" costs for Period I, recall that I denote the respective prediction by $\widehat{c}^{r,sb}$. The regression estimates are depicted in Table 4, where each column refers to the cost estimates resulting from the respective specification of the demand estimation (e.g., the first column is based on the costs inferred from the parameters of the simple logit demand specification – see column 1 of Table 2). The resulting backcasted cost estimates for store brands are presented in Table 5 (see column showing backcasted cost for Period I).

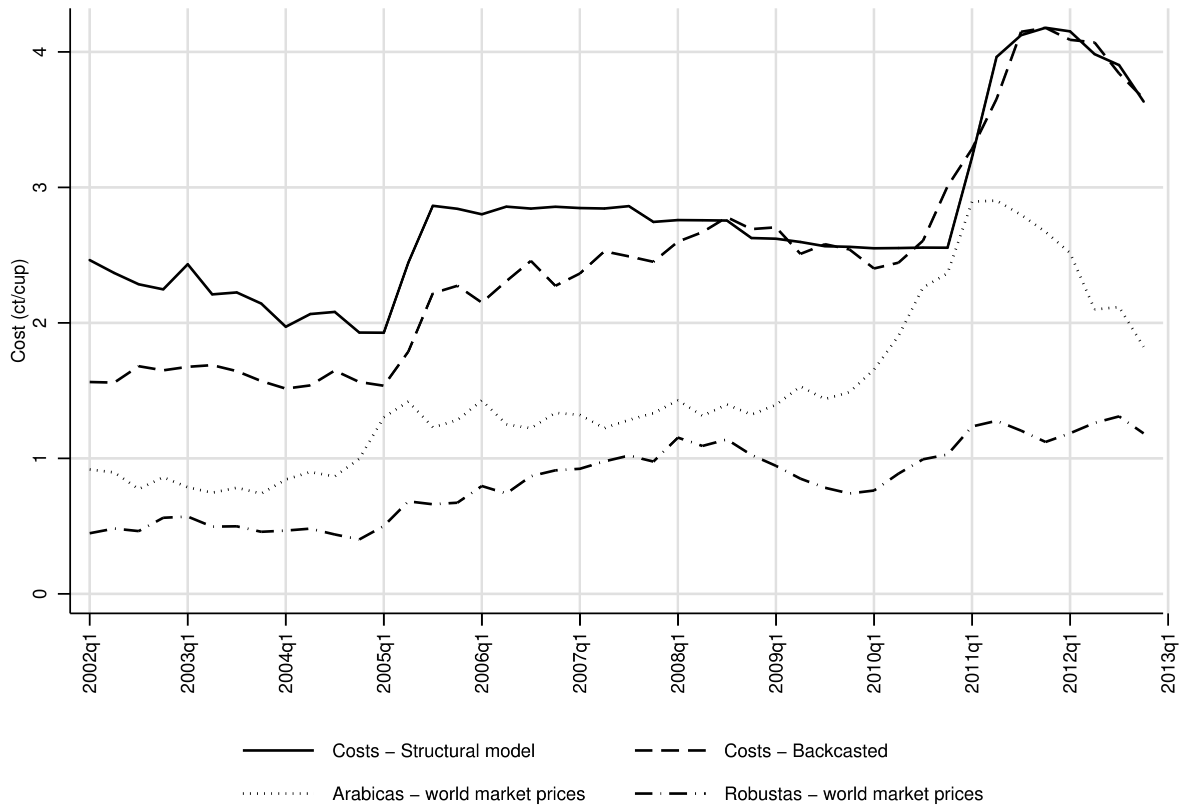
Table 4: Empirical results of cost estimation

	(1)	(2)	(3)	(4)
Mild share	0.388***	0.401***	0.332***	0.395***
	(5.57)	(5.83)	(4.32)	(5.71)
Organic label share	2.293***	2.327***	2.138***	2.311***
	(11.85)	(11.89)	(11.19)	(11.88)
Decaf share	0.353**	0.346**	0.384**	0.349**
	(3.09)	(3.08)	(3.04)	(3.08)
Constant	-16.94**	-16.57**	-18.61**	-16.75**
	(-2.78)	(-2.77)	(-2.62)	(-2.77)
World market bean prices + 2 lags	Yes	Yes	Yes	Yes
Labor costs (interacted with retailer)	Yes	Yes	Yes	Yes
Product fixed effects	Yes	Yes	Yes	Yes
Market fixed effects	Yes	Yes	Yes	Yes
Sample size	752	752	752	752
R^2	0.920	0.921	0.907	0.921

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure 2: Average cost backed out from structural model and backcasted costs (based on results of specification 4)



5.3 Comparison

In the following section, I compare the results for Period I using the cost estimation based on the structural model alone to the backcasted costs. Figure 2 shows the average cost across all store brands based on the structural model alone compared to the average backcasted costs (dashed line) over time (the results are based on the demand estimation parameters of the instrumented nested logit specification number (4)).[15]

While the results of both estimation approaches are evidently very similar during Period II, they substantially differ during Period I. On average, during Period I, the marginal cost estimates based on the structural model are about 26% above the backcasted marginal costs.

[15]See Figure 3 in the appendix for a product-specific illustration of the cost estimates.

I find similar results when looking at the product level. Table 5 shows the average cost (and resulting average gross margins expressed as $1 - \frac{cost}{price}$) of the two estimation approaches for the instrumented logit and the instrumented nested logit specifications (specifications (2) and (4)) for each product for Period I and Period II. Again, I observe that the estimation results are almost identical during Period II, but differ considerably during Period I.

In order to put the estimation results, which are expressed in cost per cup, into perspective, Table 5 also illustrates the implied gross margins (see figures in parentheses). Looking at the results based on the fourth specification for Period II, the implied margins for both estimation approaches range from 27% to 32% depending on the product. During Period I, the implied margins resulting from the structural models alone range from 32% to 36%, while the results of the backcasting approach range from 41% to 53%.

The large difference between the backcasted marginal costs and the estimates based on the structural model alone must be interpreted in light of the retailers' pricing reaction of their store brands in response to the presence of higher wholesale prices of national brands. Recall that the cost of store brands were not affected by the cartel and that, between 2004 and 2005, I observe a large increase in retail prices while at the same time world market prices for beans did not increase quite as much (see Figure 1). Looking at the estimation results from the structural model alone during Period I, higher observed prices for store brands are "accomodated" by an increase in marginal costs. Instead, the increase in the marginal cost estimates using the backcasting approach is much more similar to the increase of world market prices for Arabica beans and hence to the main cost input factor.[16]

Admitting that the difference in cost estimates between the two estimation approaches might still be the result of a misspecified out-of-sample prediction (for example, due to unobserved cost-shifters). However, given the nature of the product, a more likely reason for the difference in the marginal cost estimates is that the set-up of the structural model does not properly take into account the retailers' multiproduct pricing following a general increase in the price level (due to the cartel-induced increase in wholesale prices of national

[16] On average, the prices for store brands increase by 0.87 cent per standardized cup while the world market prices for Arabica beans only increase by 0.65 cent per cup. The cost estimated using the structural model increase by 0.86 ct per standardized cup, while the backcasted costs increase by only 0.75ct per standardized cup.

Table 5: Costs (ct/cup) and (gross) Margins (in %)

Specification	Product no.	Structural model - Period I	Backcasted - Period I	Structural model - Period II	Backcasted - Period II
	1	2.82 (26.3)	2.39 (37.6)	3.60 (21.5)	3.60 (21.5)
	2	2.84 (26.1)	2.57 (33.1)	3.69 (21.2)	3.69 (21.2)
	3	2.62 (27.9)	2.02 (44.4)	3.28 (24.3)	3.28 (24.3)
(2)	4	2.84 (26.3)	2.46 (36.2)	3.57 (22.2)	3.57 (22.2)
	5	2.67 (27.6)	2.21 (40.1)	3.36 (23.2)	3.36 (23.2)
	6	2.88 (25.4)	2.25 (41.7)	3.53 (21.3)	3.53 (21.3)
	7	2.94 (24.9)	2.58 (34.1)	3.80 (20.7)	3.80 (20.7)
	8	2.79 (25.8)	2.17 (42.3)	3.39 (22.6)	3.39 (22.6)
	1	2.52 (34.2)	2.10 (45.1)	3.31 (27.8)	3.31 (27.8)
	2	2.55 (33.6)	2.27 (40.9)	3.40 (27.4)	3.40 (27.4)
	3	2.32 (36.1)	1.70 (53.2)	2.97 (31.5)	2.97 (31.5)
(4)	4	2.53 (34.4)	2.15 (44.2)	3.27 (28.7)	3.27 (28.7)
	5	2.36 (36.0)	1.90 (48.5)	3.06 (30.1)	3.06 (30.1)
	6	2.59 (32.9)	1.97 (49.0)	3.24 (27.7)	3.24 (27.7)
	7	2.65 (32.3)	2.29 (41.5)	3.51 (26.7)	3.51 (26.7)
	8	2.50 (33.5)	1.88 (50.0)	3.10 (29.2)	3.10 (29.2)

Note: Implied (gross) margins in parentheses

brands). In a multiproduct setting, a (cartel-induced) change of cost for national brands – as observed in Period I – has a direct impact on the price setting of store brands, even though cost of store brands are not affected by such an increase in the wholesale prices of national brands. While in the backcasted model the costs remain within the magnitude of the observed cost changes of input factors, the marginal costs estimated from the structural model alone increase by considerably more than the observed input cost factors. One could therefore hypothesize that, in order for retailers to maximize profits across all brands (of coffee), the rise in the price level for national brands generates a new focal point and thus an opportunity for retailers to also increase the prices of their store brands. The selected structural model fails to accommodate such coordinated behavior.

6 Conclusions

There is an ongoing debate among researchers in the field of industrial organization over the accurateness of structural models as opposed to more widely used empirical approaches. But there is little evidence or testing of the validity of such structural models.

In this paper, I introduce a way of testing the predictive power of structural models. The test encompasses a comparison of marginal cost estimates using an out-of-sample prediction to those estimates backed out from a structural model of retail competition. The analysis focuses on the prices and costs for coffee store brands in the presence of a wholesale cartel between manufacturers of national brands for coffee. I exploit the cartel-induced price increase, and the fact that costs for store brands were not affected by the cartel, to test the predictive power of the selected structural model for retail competition.

The comparison between the different cost estimation procedures reveals substantial differences. In particular, I observe that during a substantial cartel-induced retail price increase between 2004 and 2005, the increase in marginal cost estimates based on the structural model alone is not only considerably higher than the increase in marginal cost estimates of the out-of-sample prediction but also considerably above the increase of world market prices for coffee beans, which represents the main cost input factor for ground coffee. Given that the costs for coffee store brands were not affected by the increase in wholesale prices, the difference in backed out "costs" should rather be interpreted as the retailer's optimal pricing reaction to an increase in the overall price level, which the selected structural model does not accommodate. While it remains open whether other types of structural models, e.g., allowing for coordinated behavior,[17] may perform differently, the introduced test to compare out-of-sample predictions and predictions based on structural models alone may also prove useful in other circumstances.

[17] See for example Pesendorfer and Schmidt-Dengler (2008) or Davis and Huse (2010).

References

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2):3–30.

Ashenfelter, O. C., Hosken, D., and Weinberg, M. (2009). Generating evidence to guide merger enforcement. Technical report, National Bureau of Economic Research.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *Journal of political Economy*, 112(1):68–105.

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

Bjoernerstedt, J. and Verboven, F. (2016). Does merger simulation work? evidence from the swedish analgesics market. *American Economic Journal: Applied Economics*, 8(3):125–64.

BKartA (2010). *Bussgeldverfahren gegen Kaffeeröster wegen Preisabsprachen; Case report B11-18/08.*

BKartA (2014). *Sektoruntersuchung Lebensmitteleinzelhandel. Inquiry Report B2-15/11.*

Bonnet, C. and Dubois, P. (2010). Inference on vertical contracts between manufacturers and retailers allowing for nonlinear pricing and resale price maintenance. *The RAND Journal of Economics*, 41(1):139–164.

Bonnet, C., Dubois, P., Villas Boas, S. B., and Klapper, D. (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 Villas-Boas, S. B. (2016). An analysis of asymmetric consumer price responses and asymmetric cost pass-through in the french coffee market. *European Review of Agricultural Economics*, 43(5):781–804.

Cohen, M. A. and Cotterill, R. W. (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.

Davis, P. and Huse, C. (2010). Estimating the 'coordinated effects' of mergers. *Competition Commission, January.*

Hausman, J. A. and Leonard, G. K. (2002). The competitive effects of a new product introduction: A case study. *The Journal of Industrial Economics*, 50(3):237–263.

Holler, E. and Rickert, D. (2019). Upstream collusion and resale price maintenance: An anatomy of the german coffee cartel. *Preliminary.*

Nevo, A. and Whinston, M. D. (2010). Taking the dogma out of econometrics: Structural modeling and credible inference. *Journal of Economic Perspectives*, 24(2):69–82.

Pesendorfer, M. and Schmidt-Dengler, P. (2008). Asymptotic least squares estimators for dynamic games. *The Review of Economic Studies*, 75(3):901–928.

Peters, C. (2006). Evaluating the performance of merger simulation: Evidence from the us airline industry. *The Journal of Law and Economics*, 49(2):627–649.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy*, 110(4):705–729.

The Nielsen Company GmbH (2017). *Nielsen Consumers Deutschland: Verbraucher - Handel - Werbung.*

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

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.

Appendix

Figure 3: Costs from a structural supply model and counterfactual costs by product (Specification 4)