

When it Hurts:

Demand Estimation in the U.S Analgesics Market

Reid Falconer^a, Eric Beckwith^a, Felix Adam^a, Maryam Rahbaralam^a

^a*Barcelona Graduate School of Economics, Barcelona, Spain*

Keywords: Demand Estimation, OTC Analgesics, Nested Logit Model

1. Introduction

Demand analysis of the market for over the counter (OTC) analgesics, better known as painkillers, offers valuable insights into what drives the demand for painkillers, how consumers tend to be influenced by prices and product differentiation. Furthermore, these insights have potential use for marketing and management strategies.

The market for painkillers can be considered vastly different from other consumer goods available in supermarkets such as soda or cereals. It is not entirely clear, whether OTC painkillers should be regarded as normal or necessary goods. It could be, that the demand for painkillers rises with falling prices, due to greater affordability. On the other hand, one could argue that, given proper use, painkillers are almost independent of prices, since consumers only use them when needed. Moreover, the market for painkillers is clearly segmented by natural groupings, given by the active substance in the drug.

In this work we present an analysis of demand patterns for analgesics in the U.S market, using scanner data. We build on the extensive Dominicks Dataset which provides weekly scanner data from 56 stores of Dominicks Finer Foods in the Chicago area from 1990 to 1997. Besides weekly sales and price data, the data set includes extensive information on customer demographics as well as many economic indicators.

Email addresses: reid.falconer@bracelonagse.eu (Reid Falconer),
eric.beckwith@bracelonagse.eu (Eric Beckwith), felix.adam@bracelonagse.eu (Felix Adam),
maryam.rahbaralam@bracelonagse.eu (Maryam Rahbaralam)

The focus of our work lies on estimating price elasticities for painkillers as well as identifying other drivers of painkiller sales. As aforementioned, painkillers can be naturally grouped by their active ingredient and consequently, we estimate own- and cross-price elasticities using a nested logit model that is close in spirit to that of Berry (1994).

Our results show that the grouping by active ingredients carries over to the observed demand patterns. Cross-price elasticities obtained from the nested logit model display low rates of substitution across different substances, while elasticities inside the groups are significantly larger. Since observed market shares remain stable over time, we assume that consumers have strong preferences towards active substances, but not for certain products within a group.

The outline of this paper is as follows: firstly, an overview of the U.S analgesics market is provided along with a description and statistical analysis of the data. Section 3 provides an in-depth discussion of the methodological approach and Section 4 presents the results. Finally, Section 5 concludes.

2. The U.S OTC Analgesics Industry and Data

Over-the-counter (OTC) analgesics or painkillers are non-prescription drugs to treat fever and pain. Painkillers come in three main active ingredients: acetaminophen (called paracetamol in the U.K.) aspirin and ibuprofen. The active ingredients may vary in the kinds of pains they relieve and in their side effects. Acetaminophen treats most pains and fevers and is known for having little side effects. Ibuprofen also treats most pains and fevers and is often used to reduce inflammations, but it may have side effects on the stomach. Aspirin also has a blood-diluting effect, which has both advantages and disadvantages. While each active substance may, therefore, relieve pain and reduce fever in different ways and with varying side effects, consumer perceptions of the brands (branded vs generic) may also be important. This is evident from a large amount of advertising in the sector. So it is ultimately an empirical question to which extent brands within different active ingredients are substitutes.

The OTC analgesics industry is concentrated such that five firms account for approxi-

mately 95% of the market share in dollar terms ¹. Tylenol is the industry leader with a market share of approximately 42% while generic firm, Dominick's, only accounts for 23.9% of the market. Table 1 gives a breakdown of the sample used in this study and shows the market shares of the three active ingredients. With a market share of 52%, acetaminophen is by far the most important ingredient. Ibuprofen and aspirin each have a market share of approximately 34% and 14% respectively.

Table 1: Market Shares by Active Ingredients

Active Ingredient	Market Share
Aspirin	14.30%
Ibuprofen	33.80%
Acetaminophen	51.90%

Source: Dominick's dataset

Interestingly, each firm predominately specialises in one active ingredient. Table 2 presents the market shares of the firms, together with their products and active substance. This illustrates that Tylenol and Dominick's are the only two firms within the acetaminophen segment: Tylenol sells X/S Caplets as its main product, whereas Dominick's sells ibuprofen as its primary product. Moreover, Advil is the main firm within the ibuprofen segment while Anacin and Bayer are the only two firms that are solely based in the aspirin market.

¹Excedrin is omitted in this analysis, given that their main product includes both aspirin and acetaminophen as active ingredients.

Table 2: Market Shares by Firms, Products and Active Ingredients

Firm	Firm Market Share	Product	Active Ingredient	Market Share
Dominicks	23.90%	Dom Coated Aspirin	Aspirin	5.20%
		Dom E/S Non-Asp Caplet	Acetaminophen	4.80%
		Dom Ibuprofen	Ibuprofen	9.30%
		Dom X/S Non-Asp	Acetaminophen	4.60%
Tylenol	42.50%	Tylenol X/S Caplet	Acetaminophen	15.00%
		Tylenol X/S Gelcaps	Acetaminophen	13.80%
		Tylenol X/S Tablets	Acetaminophen	9.40%
		Tylenol X/S Tabs	Acetaminophen	4.30%
Advil	24.40%	Advil	Ibuprofen	14.80%
		Advil Caplet	Ibuprofen	9.60%
Anacin	4.50%	Anacin Tabs	Aspirin	4.50%
Bayer	4.60%	Bayer Aspirin	Aspirin	4.60%

Source: Dominick's dataset

2.1. Data

The primary data used in this study is drawn from the Dominick's dataset. The dataset covers store-level scanner data collected at Dominick's Finer Foods over a period of seven years. Specifically, the dataset consists of weekly observations of the prices, sales and market shares of OTC analgesics in the Chicago metropolitan area for 343 weeks, from January 1992 to December 1998. The observations are defined as product-week-store combinations, and the data contains essential product attributes: active ingredient, pill type (caplet, tablet, gelcap, etc.) and the number of pills contained in the product. Additionally, the data is supplemented with demographic data at the store level.

The quantity sold was computed by summing weekly total units sold within each store and dividing that by the size of the package (i.e. size adjusted units sold, given that some firms have different package sizes). Additionally, since the actual prices were missing, unit values were used as a proxy for them. To compute unit values, prices were calculated at the store level for each product, as total revenue for the product in the store in a week divided by the total number of products sold in that week ². Furthermore, all prices (unit

²Note that expenditures are further adjusted by subtracting the value of coupons (if any)

values) were adjusted for inflation by dividing them by the weekly regional Consumer Price Index (CPI).

Over the study period, the average total weekly amounts of analgesic sold by Dominick's, Tylenol, Advil, Anacin and Bayer, were 408, 251, 215, 331 and 369 units, respectively, indicating that Dominicks was the leading firm followed by Bayer, Anacin, Tylenol and Advil. Furthermore, this firm ordering is identical when evaluating the firm's median weekly prices per unit with Dominick's and Bayer having the two lowest median weekly prices per unit within the market respectively (See: Figure 1).

Figure 1: Median Price vs Average Sales by Firm

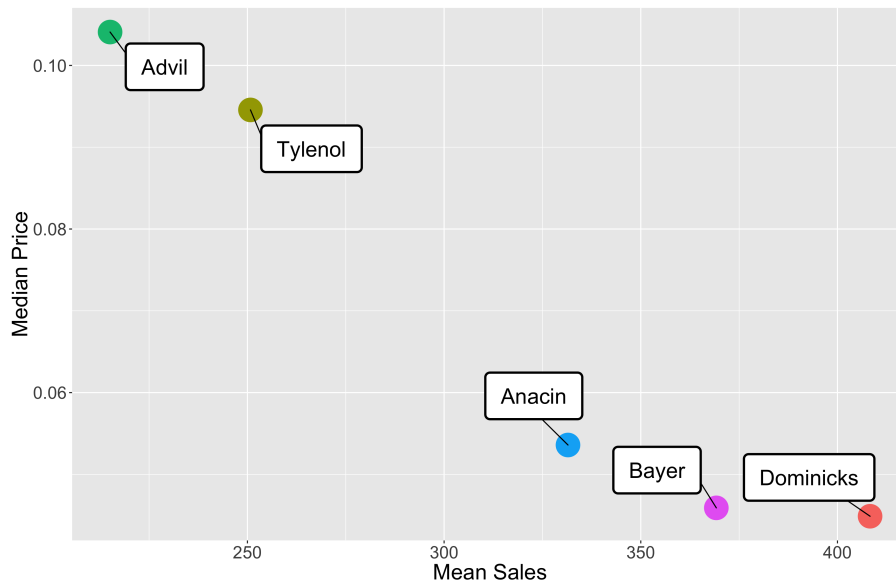
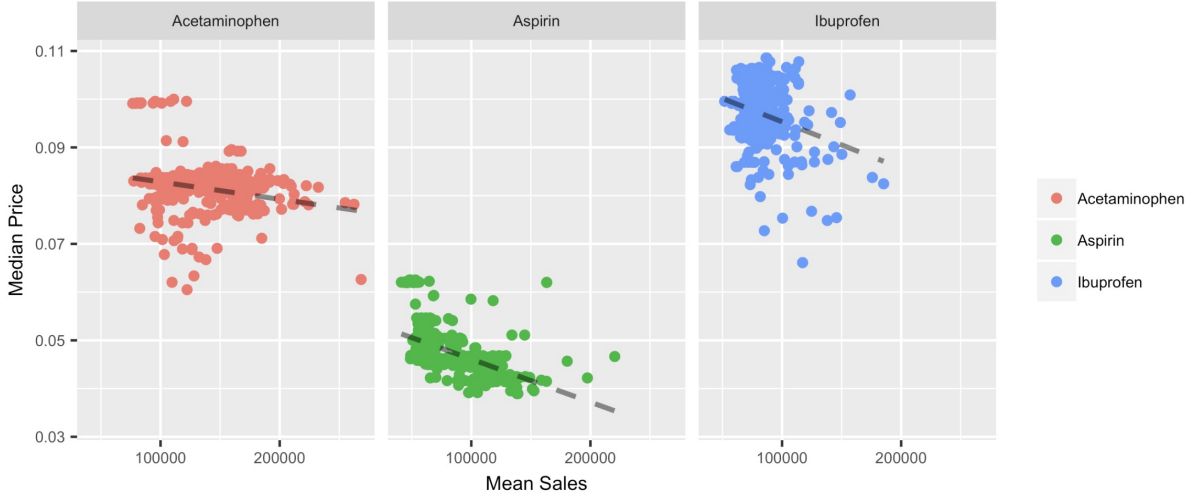


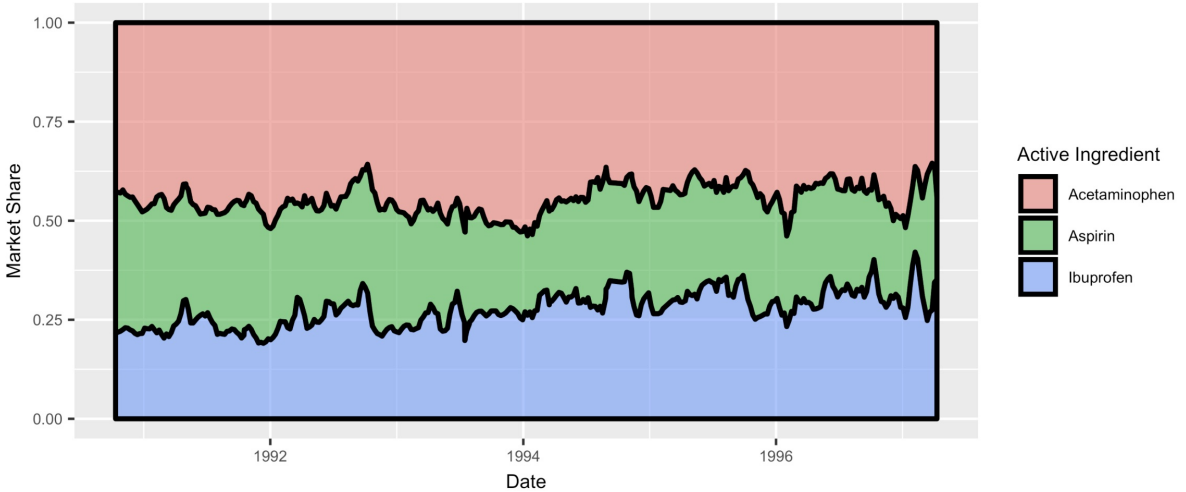
Figure 2 shows that the grouping by the active substance can also be found in the price and sales data. The three groups clearly stand out, with ibuprofen having the highest median prices, followed by acetaminophen and aspirin. Moreover, all plots show signs of downward sloping demand curves.

Figure 2: Median Price vs Average Sales by Active Ingredient



The market shares by percentage of total quantity sold can be seen in Figure 3. Here we see, that the market shares by active ingredient don't vary substantially over time. Acetaminophen has the highest market share, followed by ibuprofen and aspirin. The shares almost remain constant over the whole time frame, with only little deviations.

Figure 3: Market Shares by Active Substance (%Of Qty Sold)



3. Methodological Approach

3.1. Econometric Set Up

The findings described in Section 2 lead us to our econometric setup. We make use of a nested logit model as described by McFadden (1978), Cardell (1997) and Berry (1994).

We begin by motivating the use of the nested logit model, then describe our econometric setting and finally discuss instrumental and control variables.

McFadden (1978) analyses the behaviour of consumers with regard to choosing a residential location. While the choice of painkillers and housing seem very distant at first, we argue that the nested logit model used by McFadden offers some valuable insights into the painkiller market. McFadden (1978) argues, that consumers choose their residential location by maximising the utility they receive from a particular house. This utility depends on the characteristics, such as location, number of rooms and the community. Additionally, the consumer is faced with choosing exactly one house or none (McFadden, 1978).

The idea that consumers maximise utility with respect to product characteristics translates easily to the painkiller market. Consumers choose painkillers dependent on their tolerance for specific active ingredients (i.e. aspirin, acetaminophen and ibuprofen) and their brand preferences (i.e. branded vs generic). The second assumption, consumers only buy one good or none (the outside good), can also be motivated through properties specific to painkillers.

While it is possible to take combinations of certain painkillers, it is generally not recommended to do so, only if indicated by medical advice. The simultaneous intake of aspirin and ibuprofen is associated with an increased risk of stomach bleeding, or other side effects (Miller, 1981). Furthermore, we assume that consumers buy painkillers to tackle a specific problem, such as headaches or light fever.

Under rational behaviour, consumers will, therefore, buy the product which best suits their needs. This phenomenon is inherently different from other product categories, such as sodas, where one could imagine that consumers purchase different types of sodas at the same time. Moreover, including the choice of the outside good makes sense from the modelling and consumer perspective. It seems sensible to assume that in any given week, many consumers choose not to buy any painkillers at all, since they don't experience any health issues. Including the outside good also simplifies the estimation of the structural demand parameters, as discussed in Berry (1994).

Another critical assumption made by McFadden (1978) is that consumer tastes are cor-

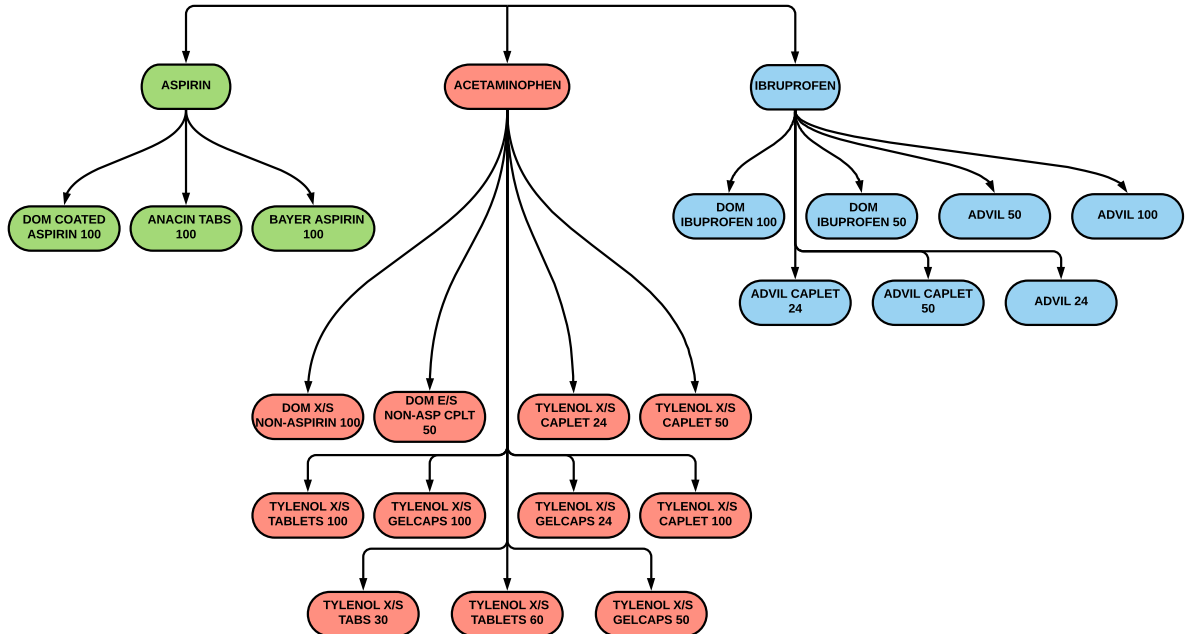
related across products. In the case of housing, the similarities are shared across communities, where houses in the same community share specific characteristics (McFadden, 1978)). Given that painkillers can be easily grouped by their active ingredient, this assumption translates well to our setting. We expect that consumer and product characteristics are correlated inside the groups. For example, some consumers might be more inclined to buy painkillers containing ibuprofen than aspirin, be it for medical or digestibility reasons.

Now that the conceptual framework has been set, we present the empirical model.

3.2. The Empirical Model

As described, the nested logit model arises from the choice a consumer faces when buying painkillers. Each consumer bases their choice of painkiller on the utility they receive from it. We group products into four groups $g \in G = \{A, P, I, O\}$ where A aspirin, P is paracetamol, and I is ibuprofen. (See Figure 4).

Figure 4: Tree structure of the active ingredient nested logit model



Group $g = O$ denotes the choice the outside good. Consumers i 's utility u_{ij} of consuming product j is defined as:

$$u_{ij} = X_j\beta - \alpha p_j + \xi_j + \xi_{ig} + (1 - \sigma)\epsilon_{ij} \quad (1)$$

Where p_j denotes the price of product j , ξ_i are product characteristics specific to product j , ξ_{ig} are common properties for all products in group g and ϵ_{ij} is the error term assumed to have an extreme value distribution. The Parameter σ , which is $0 \leq \sigma < 1$, determines the “within group” correlation. If σ approaches one, the correlation among brands within a group becomes one, meaning that the consumer is only concerned with choosing a group, not products within the group. If σ approaches zero, the “within group” correlation becomes zero, which is equivalent to the simple multinomial logit model (Berry, 1994).

We assume that consumer i purchases brand j if $u_{ij} > u_{ik}$, Where $\forall k \neq j$. If ϵ_{ij} follows a generalized extreme value distribution, the market share of product $j \in g$ within group g is derived as follows:

$$s_{j|g}(\vec{\delta}, \sigma) = \frac{\exp(\delta_j/(1 - \sigma))}{D_g} \quad (2)$$

where $D_g \equiv \sum_{j \in \mathcal{G}_g} \exp(\delta_j/(1 - \sigma))$. Similarly, the probability of choosing group g is:

$$s_g(\vec{\delta}, \sigma) = \frac{D_g^{(1-\sigma)}}{\left[\sum_k D_k^{(1-\sigma)} \right]} \quad (3)$$

Accordingly, the market share of product $j \in g$ can be derived from the following relationship:

$$s_j = s_{j/g} \cdot S_g \quad (4)$$

Meanwhile, the market share of the “outside good” is:

$$s_0(\vec{\delta}, \sigma) = 1 / \left[\sum_k D_k^{(1-\sigma)} \right] \quad (5)$$

In this nested logit framework, the independence of irrelevant alternatives (IIA) assumption holds. In the following, we show the calculations of the price elasticities. Taking the

logs of market shares we derive our empirical model:

$$\begin{aligned}\ln(s_j) - \ln(s_0) &= \delta_j + \sigma \ln(s_{j|g}) \\ &= \vec{x}_j \vec{\beta} - \alpha p_j + \sigma \ln(s_{j|g}) + \xi_j\end{aligned}\tag{6}$$

We can then derive the own- and cross price elasticities from our estimated parameters:

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

However, the presence of unobserved characteristics creates the well-known problem of price endogeneity: if consumers and firms know the values of unobserved product characteristics ξ_{jt} , while we (the econometricians) do not, then prices (p_j) are likely to be correlated with these unobservables. Additionally, in the nested logit model, the nesting term $\ln(s_{j|g})$ is endogenous by construction, since any shocks to ξ_{jt} that increase the dependent variable $\ln(s_j) - \ln(s_0)$ will also increase the nesting term $\ln(s_{j|g})$.

3.3. Endogeneity Bias and Instrumental Variables

Our primary interest lies in estimating the causal effect of price changes on market shares, as this effect will allow us to obtain meaningful elasticities estimates.

In addition to the above-mentioned sources of endogeneity within our model, prices and quantities are determined by the intersection of demand and supply curves. However, it is not clear, whether demand or supply shocks drive the changes observed in quantities and prices. It could be, that a sudden positive demand shock, due to a flu pandemic, shifts the demand curve outward and therefore drives up prices. Similarly, sudden supply shocks could also lead to price and quantity changes. In other words, prices are determined within the system. However, we are interested in the effect of exogenous price changes on demand (see Angrist and Pischke (2008) for a detailed discussion).

Therefore, our identification assumption is the existence of instruments z_t . An instrument is a variable that is correlated with one of our endogenous variables p_j or $\ln(s_{j|g})$ (rele-

vance) but is not correlated with the error term ξ_{jt} (exogeneity). Formally, they have to satisfy the conditional expectation $E[\xi_{jt}|z_t] = 0$.

We use two sets of instrumental variables. First, to obtain unbiased estimates of the price coefficient (α) we follow Chintagunta et al. (2003) who use wholesale prices as instruments for retail prices. Wholesale prices are likely to be driven by factors which are exogenous to store-level demand, such as manufacturing prices. As seen in Appendix D, wholesale prices are also strongly correlated with retail prices. Thus, it makes sense to use wholesale prices as an instrument in our two-stage least squares (2SLS) model.

However, to avoid risks of wholesale prices being endogenous with retail prices, we first predict wholesale prices and then use that prediction as an instrument. For this prediction model, we use base product characteristics such as package size, the primary ingredient, and firm. In addition, since the base active ingredients are no longer subject to any patents, they can in many cases be manufactured in other countries at lower cost. For example, the production of paracetamol in China and India occupied about 84% of the global production in 2016 (Wong, 2018). Although these percentages were lower during the time period of our study, this should still be significant. We, therefore, include historical cross currency rates for China, India, and Europe to allow for different bulk purchasing choices by the firms.

Second, to obtain unbiased estimates of the nesting term $\ln(s_{j/g})$, we instrument using the number of competitor products on sale within each group at a given time. Nevertheless, a possible problem is weak identification, which occurs when instruments are only weakly correlated with the endogenous variable. Since there are two endogenous variables, the conventional first-stage F-statistic is no longer suitable to test for weak instruments. Therefore, we utilise the Sanderson and Windmeijer (2016) F-statistic to check whether each particular endogenous variable is weakly identified.

3.4. Control Variables

Similar to the issue of endogeneity bias, in the absence of true experiments we need to control for various other confounding factors, such as consumer heterogeneity Chintagunta et al. (2003)).

The use of store-level data allows us to exploit the variability of store-level sales and prices. Aggregating over different regions would potentially lead to the averaging out of a large share of quantity and price movements. Fortunately, the data set contains a rich set of demographic and economic variables.

To control for household wealth, we include the median income, the mean household value, the percentage of the population that is retired and the percentage of the population that is unemployed. We expect that wealthier households are more inelastic to price changes, especially for goods that could be considered necessary such as painkillers. On the other hand, an impoverished household will presumably react strongly to price changes, due to its tighter budget.

Further, we include the percentage of the population below nine years, the average household size, the percentage of ethnic minorities and the percentage of people with no vehicle. We need to control for the percentage of children in the population since certain analgesics are not recommended below the age of 16. This restriction could potentially influence substitution patterns, especially across substances. Furthermore, household size will likely affect market shares through various channels. For example, larger households could react more abruptly to prices due to having constrained budgets or more apparent needs for multiple painkillers. The percentage of the population without a vehicle will generally affect the ability to shop and travel far distances which should, therefore, impact the ability of consumers to substitute by switching to different stores.

4. Results

As a baseline, we first ran a multinomial logit model using our selected demographics and sales data - with our predicted wholesale price as the instrumental variable (Table 3, column 1). We obtain a price coefficient of -11.48, which allows us to calculate the own price elasticities shown in Appendix B. These results are consistent with what we would expect logically; the premium brands such as Advil and Tylenol are more elastic than generics, i.e. Dominicks. Additionally, the smaller count packages are more elastic than the bulk 100 counts. We also note that the cross-price elasticities in this model are negligible, with the largest across all products of the order 0.003.

Moving to the nested logit model framework, we compare our choice of nesting by ingredient to not only the baseline multinomial logit but additionally, an alternative grouping on generic vs branded (See Figure A.1). In all nestings, we continue to use our predicted wholesale prices as an instrument for prices, with the addition of competitor products on sale with a group as instruments for the group shares - with standard errors are clustered at the store-level³. We found that nesting by ingredient (Table 3, column 3) provided a significant improvement over the basic multinomial logit model as well as over the generic nesting (Table 3, column 2) as can be seen in the estimation results. Also, since the dataset only has one brand of generic, Dominicks, and it is the store chains brand as well, there are concerns around using this as a nesting split in this case. Since the nesting formula can be extended to the case of multiple levels of nests with different correlation parameters for different groups (Berry, 1994) we looked at the possible extension to a dual level nest. We ran a model with two nesting levels generic vs branded and ingredient (See Figure A.2 and Table 3, column 4), and while the coefficients looked sensible, the model fit did not improve over our ingredient nesting. It seems to us that a more logical future extension would be to try moving to a cross-nested logit model or random coefficient model than to try extending to multi-level nested logit.

Using our ingredient nesting model (Table 3, column 3), we obtain a price coefficient of -6.37 with a group correlation of utility levels coefficient of 0.39. As this group coefficient parameter approaches one, the within-group correlation of utility levels goes to one, and as it approaches zero, the within-group correlation goes to zero. Using these parameters, we compute own and cross-price elasticities within each of our ingredient nests (See Appendix C). We see a similar pattern in own price elasticities to the multinomial in regards to brand and package size though the numbers are now larger. However, we also now have significant positive cross-price elasticities within each group to represent the more relative consumer indifference between products within an ingredient nest. Again we also note that the cross-price elasticities outside of a grouping are effectively zero.

On the other hand, more similar products tend to have higher cross-price elasticities. For example, Tylenol X/S Caplets 24 and Tylenol Tab 30 have an elasticity of 1.6, while

³Given that the products sold in the same store will share identical customer demographics, one must cluster the standard errors at the store-level in order to control for the potential presence of an unobserved effect on the error term (see Angrist and Pischke (2008)).

Tylenol X/S Caplets 24 has an elasticity of 0.77 with Tylenol X/S Tablets 100. So while the products all have greater own elasticities, they are also far more vulnerable to competitor price movements within their ingredient group.

This vulnerability is more complicated in our dataset and nestings since the same firm produces many of the grouped products and thus the market share they would steal from the group are often from their competing products. Another factor pointing towards this is that cross-price elasticities within groups are higher when looking at products of the same firm. On average, inside of the acetaminophen nest, Tylenol products have a cross-price elasticity of 1.01 with other Tylenol products, but only a cross-price elasticity of 0.88 with Dominicks products. The same holds in the ibuprofen nest. The average cross-price elasticity between Advil products is 0.97, while the average cross-price elasticity of Advil and Dominicks is 0.76. On the other hand, this could also reflect a preference for a brand-over a generic product.

Of the exogenous variables included in the final model, the most significant appears to be products having sales promotions in a given week, percentage of the demographic population under the age of 9, and the ethnicity breakdown of the store customers.

Table 3: Parameter Estimates from the Main Specifications

	Multinomial (1)	Nested Logit (2)	Nested Logit (3)	Nested Logit (4)
Dependent Variable	Log Difference Market Shares			
Estimator	OLS	2SLS	2SLS	2SLS
Price (weekly measure)	-11.483	-4.8581	-6.3717	-3.6426
Instrument: predicted wholesale price	(-38.957)	(-6.4456)	(-20.684)	(-4.7684)
%Population under age 9	3.8289	4.0242	4.2789	4.0891
	(2.1443)	(2.1662)	(2.2768)	(2.2003)
Log median income	-0.2045	-0.1335	-0.1791	-0.1539
	(-1.0951)	(-0.6514)	(-0.8437)	(-0.7530)
%Population that are Blacks & Hispanics	-0.46	-0.515	-0.5894	-0.5287
	(-2.9861)	(-3.1366)	(-3.3419)	(-3.1840)
Average Household Size	-0.1448	-0.1719	-0.1902	-0.1731
	(-0.8500)	(-0.9364)	(-1.0367)	(-0.9434)
% with no vehicles	-0.1981	-0.161	-0.2146	-0.1848
	(-0.6336)	(-0.4796)	(-0.6093)	(-0.5467)
Mean household value (approx)	0.0006	0.0008	0.0007	0.0008
	(0.6031)	(0.7470)	(0.6801)	(0.7611)
%Population that are Retired	-0.1228	0.0457	-0.0561	0.0197
	(-0.3126)	(0.1100)	(-0.1346)	(0.0471)
%Population that are Unemployed	-0.1411	0.2979	0.6348	0.3005
	(-0.1009)	(0.1965)	(0.4021)	(0.1972)
%Population that are College graduates	0.1832	0.0512	0.1332	0.0728
	(0.3746)	(0.0953)	(0.2325)	(0.1347)
All weekly promotions	0.4819	0.3744	0.3196	0.4007
	(28.708)	(15.869)	(17.506)	(16.308)
Nesting Parameters				
Share in Group (Generic vs Branded)		0.344		0.12
Instrument: competitor products on sale within brand type		(7.8623)		(2.3570)
Share in Group (Active Ingredient)			0.3908	
Instrument: competitor products on sale within active ingredient			(19.252)	
Share in Group (Active Ingredient within Generic vs Branded)				0.2404
Instrument: competitor products on sale within active ingredient and brand type				(18.387)
Constant	-3.1217	-3.5606	-3.0288	-3.6577
	(-1.6719)	(-1.7468)	(-1.4208)	(-1.7967)
No. Observations	344796	344796	344796	344796
Clusters	Store Level	Store Level	Store Level	Store Level
R-squared	0.3328	0.595	0.6209	0.5364
Adj. R-squared	0.3328	0.595	0.6209	0.5364
F-statistic	3276.8	4771.8	9833.7	5459.5

Notes: T-stats are reported in parentheses and robust standard errors clustered at store-level used during estimation.

5. Conclusion

Our model results support the hypothesis that the analgesics market is primarily grouped by active ingredient. The small cross-price elasticities across groups indicate a low level of cross-substance substitution. As discussed, reasons for this pattern could arise from consumer preferences for specific active ingredients. Another potential influence could be medical advice. Consumers might be advised by their doctors only to consume certain analgesics, dependent on their health issues. Support for the hypothesis that only active ingredients matter comes from the fact that in-group elasticities are significantly larger. This finding is because products within a group only differ by the size of the package or the method of intake (i.e. pills vs caplets). We assume that consumers are aware of the similarities and thus essentially care about prices. This result is also indicated by the fact that more similar products tend to have higher cross-price elasticities. Additionally, in-group elasticities are related to packaging size, with smaller packages having larger elasticities.

However, there are some hints towards brand loyalty, indicated by higher elasticities within a group and brand compared to the elasticities within a group and outside brand. This result supports the idea that consumers know which active ingredient and brand they want when they enter the store.

What does this mean for firms? In our data set, Dominicks is the only firm that is active in all three groups. While the aspirin group itself only has three products, the other segments are considerably larger and dominated by Advil or Tylenol. Given that consumers are unlikely to be induced to switch group, these brands can focus on solidifying their position inside their respective segment.

6. References

- Angrist, J. D. and Pischke, J.-S. (2008), *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.
- Berry, S. T. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* pp. 242–262.
- Cardell, N. S. (1997), 'Variance components structures for the extreme-value and logistic distributions with application to models of heterogeneity', *Econometric Theory* **13**(2), 185–213.
- Chintagunta, P. K., Dubé, J.-P. and Singh, V. (2003), 'Balancing profitability and customer welfare in a supermarket chain', *Quantitative Marketing and Economics* **1**(1), 111–147.
- McFadden, D. (1978), 'Modeling the choice of residential location', *Transportation Research Record* (673).
- Miller, D. R. (1981), 'Combination use of nonsteroidal anti inflammatory drugs', *Drug intelligence & clinical pharmacy* **15**(1), 3–7.
- Sanderson, E. and Windmeijer, F. (2016), 'A weak instrument f-test in linear iv models with multiple endogenous variables', *Journal of Econometrics* **190**(2), 212–221.
- Wong, A. (2018), Global production of paracetamol reaches 147 kilo mt in 2016, Technical report, LP Information.
- URL:** <https://www.lpinformationdata.com/paracetamol-market-report/chemical-material.html>

Appendix A. Nested Logit Tree Structures

Figure A.1: Tree structure of the brand type nested logit model

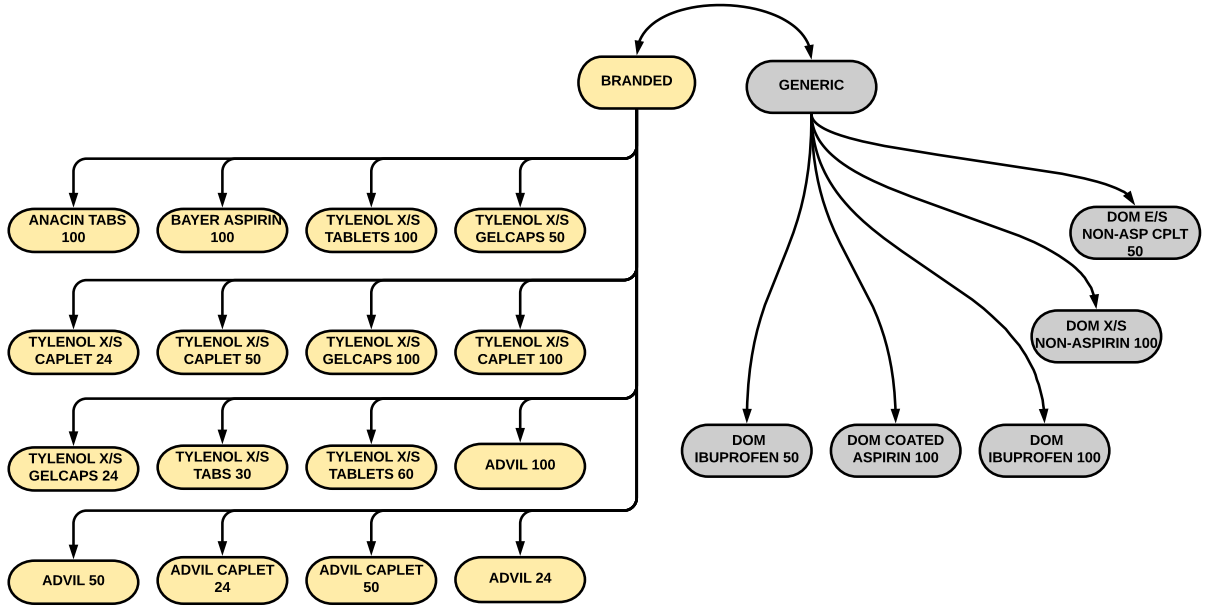
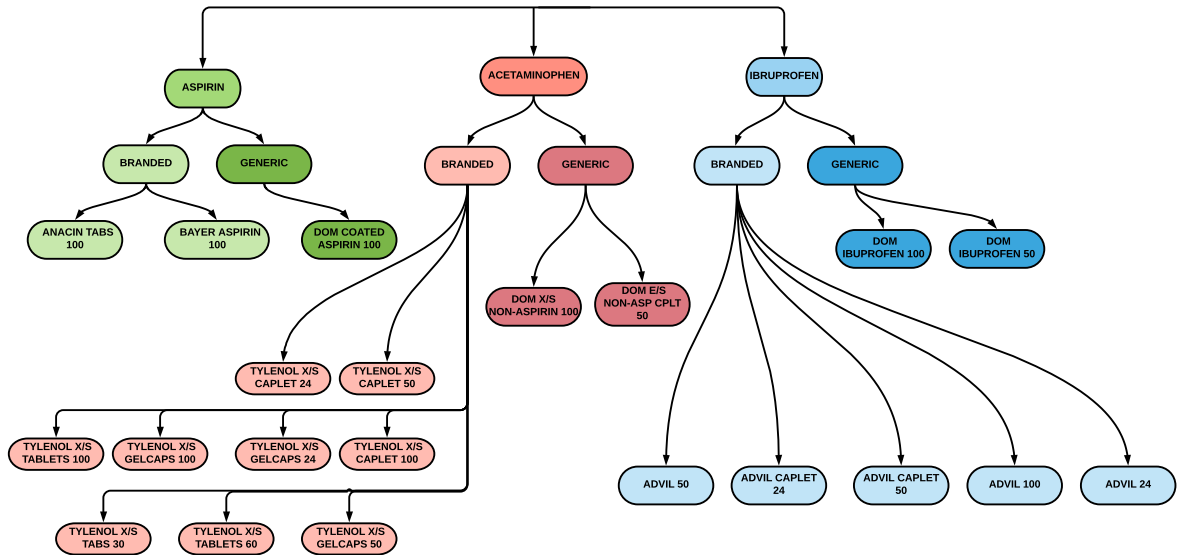


Figure A.2: Tree structure of the double nested logit model



Appendix B. Multinomial Logit Elasticities

Figure B.3: Multinomial Logit Elasticities

<u>Product</u>	<u>Own Px Elasticity</u>	<u>Product</u>	<u>Own Px Elasticity</u>
ADVIL CAPLET_24	-1.63	ADVIL_100	-0.87
ADVIL_24	-1.63	TYLENOL X/S GELCAPS_100	-0.87
TYLENOL X/S GELCAPS_24	-1.62	TYLENOL X/S CAPLET_100	-0.72
TYLENOL X/S CAPLET_24	-1.6	DOM IBUPROFEN_50	-0.72
TYLENOL X/S TABS_30	-1.41	TYLENOL X/S TABLETS_100	-0.66
TYLENOL X/S GELCAPS_50	-1.17	ANACIN TABS_100	-0.62
TYLENOL X/S CAPLET_50	-1.15	DOM IBUPROFEN_100	-0.58
ADVIL CAPLET_50	-1.14	DOM E/S NON-ASP CPLT_50	-0.53
ADVIL_50	-1.14	BAYER ASPIRIN_100	-0.53
TYLENOL X/S TABLETS_60	-0.94	DOM X/S NON-ASPIRIN_100	-0.47
		DOM COATED ASPIRIN_100	-0.26

Appendix C. Nested Logit Elasticities

Figure C.4: Acetaminophen Elasticities

	DOM X/S NON-ASPIRIN 100	DOM E/S NON-ASP CPLT 50	TYLENOL X/S CAPLET 24	TYLENOL X/S CAPLET 50	TYLENOL X/S TABLETS 100	TYLENOL X/S GELCAPS 100	TYLENOL X/S GELCAPS 24	TYLENOL X/S CAPLET 100	TYLENOL X/S TABS 30	TYLENOL X/S TABLETS 60	TYLENOL X/S GELCAPS 50
DOM X/S NON-ASPIRIN 100	-0.78	0.40	1.17	0.81	0.50	0.62	1.19	0.53	1.03	0.69	0.83
DOM E/S NON-ASP CPLT 50		-0.97	1.43	1.01	0.61	0.77	1.46	0.65	1.26	0.82	1.05
TYLENOL X/S CAPLET 24			-3.31	1.28	0.77	0.98	1.85	0.82	1.60	1.05	1.31
TYLENOL X/S CAPLET 50				-2.08	0.61	0.78	1.47	0.66	1.28	0.84	1.06
TYLENOL X/S TABLETS 100					-1.09	0.61	1.16	0.52	1.00	0.65	0.82
TYLENOL X/S GELCAPS 100						-1.50	1.31	0.58	1.14	0.75	0.93
TYLENOL X/S GELCAPS 24							-3.38	0.84	1.62	1.07	1.34
TYLENOL X/S CAPLET 100								-1.09	0.85	0.56	0.68
TYLENOL X/S TABS 30									-3.01	1.12	1.39
TYLENOL X/S TABLETS 60										-1.72	1.07
TYLENOL X/S GELCAPS 50											-2.19

Figure C.5: Aspirin Elasticities

	DOM COATED ASPIRIN 100	ANACIN TABS 100	BAYER ASPIRIN 100
DOM COATED ASPIRIN 100	-0.28	0.12	0.10
ANACIN TABS 100		-0.90	0.28
BAYER ASPIRIN 100			-0.74

Figure C.6: Ibuprofen Elasticities

	DOM IBUPROFEN 100	DOM IBUPROFEN 50	ADVIL 50	ADVIL 100	ADVIL CAPLET 24	ADVIL CAPLET 50	ADVIL 24
DOM IBUPROFEN 100	-0.86	0.39	0.63	0.49	0.89	0.63	0.88
DOM IBUPROFEN 50		-1.15	0.80	0.61	1.14	0.80	1.14
ADVIL 50			-1.78	0.55	1.02	0.73	1.03
ADVIL 100				-1.27	0.88	0.63	0.88
ADVIL CAPLET 24					-3.09	1.12	1.59
ADVIL CAPLET 50						-1.98	1.32
ADVIL 24							-2.74

Appendix D. Predicting Wholesale Prices

Dependent Variable	Weekly Wholesale Prices
Size 1	-3918852.18
Size 2	-3918852.20
Size 3	-3918852.21
Firm 1	-72029832.25
Firm 2	-72029832.21
Firm 3	-72029832.23
Firm 4	-72029832.22
Firm 5	-72029832.23
Active Ingredient 1	24648424.80
Active Ingredient 2	24648424.80
Active Ingredient 3	24648424.80
USD vs CNY Firm 1	-0.00***
USD vs CNY Firm 2	-0.00***
USD vs CNY Firm 3	0.00***
USD vs CNY Firm 4	0.00***
USD vs CNY Firm 5	-0.00***
USD vs INR Firm 1	0.00
USD vs INR Firm 2	-0.00***
USD vs INR Firm 3	0.00***
USD vs INR Firm 4	0.00***
USD vs INR Firm 5	0.00***
EUR vs USD Firm 1	0.00***
EUR vs USD Firm 2	0.02***
EUR vs USD Firm 3	0.02***
EUR vs USD Firm 4	0.01***
EUR vs USD Firm 5	0.01***
N	344796.00
R^2	0.97

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure D.7: Predicted Weekly Wholesale Prices vs Actual Weekly Retail Prices

