Does Merger Simulation Work?
Evidence from the Swedish Analgesics Market

Jonas Björnerstedt and Frank Verboven*

October 2015

Abstract

We analyze a large merger in the Swedish market for analgesics (painkillers). The merging firms raised prices by 40 percent, and some outsiders raised prices by more than 10 percent. We confront these changes with predictions from a merger simulation model. With basic supply side assumptions, the models correctly or moderately underpredict the merging firms' price increase. However, they predict a larger price increase for the smaller firm, which was not the case in practice, and they underpredict the outsiders' responses. We consider several supply side explanations: a plausible cost increase after the merger and the possibility of partial collusion.

Keywords: merger simulation, ex post merger evaluation, constant expenditures logit, analgesics or painkillers

**Jonas Björnerstedt*: Department of Economics, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. jonas.bjornerstedt@kuleuven.be. *Frank Verboven*: Department of Economics, University of Leuven, Naamsestraat 69, B-3000 Leuven, Belgium. frank.verboven@kuleuven.be. We acknowledge financial support from KU Leuven Program Financing and the Swedish Competition Authority. We are grateful for helpful comments from two anonymous referees, and from Douglas Lundin, Oivind Anti Nilsen, Tommaso Duso, Ariel Pakes, Amil Petrin, André Romahn, and seminar participants at Mannheim University, Northwestern University, University of Amsterdam, University of Bergen, and several conferences.

There is an ongoing debate on the usefulness of structural econometric models to predict counterfactual outcomes. Angrist and Pischke (2010) document the recent successes of "design-based" or "treatment effects" approaches in various fields, such as labor and development economics. They suggest that industrial organization would also greatly benefit from these approaches, taking empirical merger analysis as a test case example. At a minimum, they write, empirical evidence should be provided that structural econometric models can deliver reasonably accurate predictions. In a response, Nevo and Whinston (2010) acknowledge that the treatment effects approach may be useful to estimate the effects from mergers. But they also point out limitations, and discuss several circumstances where a structural model and merger simulation can be more useful. The most obvious instance arises when a competition authority has to evaluate the likely price effects of a proposed merger, and does not have information from closely comparable past mergers in the same or related markets. Both Angrist–Pischke and Nevo–Whinston agree that more retrospective merger analysis is clearly needed.

In this paper we provide such an analysis based on a large recent merger between AstraZeneca Tica (AZT) and GlaxoSmithKline (GSK) in the Swedish market for over-the-counter analgesics (painkillers). The merger raised competition concerns, since AZT and GSK were the only companies in the largest market segment, which contains products based on the active substance paracetamol (called acetaminophen in the U.S.). The competition authority nevertheless decided to clear the merger in April 2009. One argument was that consumers care more about brands than the active substance of painkillers, so that consumers still had sufficient substitution possibilities. Another argument in favour of the merger was the upcoming deregulation of the pharmacy monopoly, which was hoped to encourage new entry and competition. Prohibiting the merger would involve a costly court case to dismiss these arguments.

We start with an ex post analysis of the merger effects. We find that the merger raised prices of the merging firms in the paracetamol segment by on average 40%, and both firms raised prices by similar magnitudes. The price increases materialized almost immediately, just one month after the merger, and they remained for the entire two-year window after the merger. Furthermore, prices of products in one competing segment (ASA) also increased considerably, by on average more than 10%. Prices in the third segment (ibuprofen) did not increase.

Because of the unusually large merger effects, we can test a rich set of merger predictions. We develop and estimate an oligopoly model with differentiated products to see how well merger simulation performs in predicting the various observed effects. We start from a basic supply side, which assumes multiproduct Bertrand-Nash pricing and no changes other than the firms' ownership structure after the merger (hence a pure "loss in competition effect"). The demand side considers two discrete choice models: a nested logit and a random coefficients logit model. For both models, we introduce a constant expenditures specification as an alternative to the traditional unit demand specification (where price enters logarithmically instead of linearly and market shares are in values instead of volumes). The constant expenditures specification turns out to be more appropriate than the unit demand specification in our application: it results in a more plausible range of elasticities, more reasonable markups, and yields more realistic average predicted price effects for the merging firms (with still a moderate underprediction under the random coefficients variant in this case).

However, at the more detailed level of the individual firms and segments we find several interesting discrepancies between the predicted and the observed price increases. First, both the nested and the random coefficients model predict that the smaller firm in the merger, GSK, would raise its prices by much more than the larger firm, AZT, while in reality both companies raised their prices by approximately the same percentage. Second, our model predicts that the outsiders raise prices by only a small amount after the merger, while in practice some of the outsiders responded with a fairly large price increase.

We therefore augment the basic supply model to incorporate the role of two factors. First, we take into account that the price increase coincided with a reduction in the package size, i.e. several firms who raised their prices essentially reduced their package sizes drastically without sufficiently reducing the price of a package. Using pre-merger data, we find that marginal costs are decreasing in package size, so that marginal costs increased for those firms who reduced package sizes after the merger. This first modification gives higher average price predictions for the merging firms, resulting in a small average overprediction for the nested logit and a roughly correct average prediction for the random coefficients logit. Furthermore, the augmented model also partly explains some of the outsider firms' price increases. However, the model still predicts a much larger price increase for the smaller merging firm, and it cannot explain the price increases for some of the outsiders who did not reduce their package sizes.

Our second supply side modification consists of accounting for the possibility that firms engage in partial coordination. We consider a simple setting where firms differ from Bertrand-Nash behavior by partly taking into account their competitors' profits, both before and after the merger. This modification results in a smaller gap between the predicted price increase for the large and small merging firm, especially under the random coefficients logit model. Furthermore, this modification partly helps to explain the outsiders' responses. Although both supply-side modifications contribute to explaining the observed effects, some significant differences remain. We show this by estimating the remaining unobserved cost changes that would be required to rationalize the merger's price effects. Analogously, one might also interpret these as unobserved changes in coordinated behavior not captured in the model (which are inherently difficult to predict given the multiplicity of equilibria in a repeated game setting). See Miller and Weinberg (2015) for an approach that aims to distinguish between both cost and conduct changes to rationalize the merger price effects.

Our paper contributes to three related strands in the literature: merger simulation, ex post merger evaluation and especially to ex post evaluation of merger simulation.

Merger simulation Merger simulation as a tool for competition policy was introduced by Hausman, Leonard and Zona (1994) and Werden and Froeb (1994). Subsequent research has looked at a variety of issues, such as alternative demand models, e.g. Nevo (2000), Epstein and Rubinfeld (2001) or Ivaldi and Verboven (2005). Some of this work has explicitly compared different demand models and showed how different functional forms may result in rather different price predictions, see Crooke, Froeb, Tschantz and Werden (2003), Huang, Rojas and Bass (2008) and Slade (2009). While these comparisons are informative, it is difficult to disentangle the sources of the differences since the compared models differ in many respects. In contrast, we compare different demand specifications in a unified framework: for both a nested and a random coefficients logit model, we introduce the constant expenditures specification as an alternative to the typical unit demand model. This enables us to consider the role of the functional form of the price variable (linear versus logarithmic) in an integrated framework.

Quite surprisingly, the constant expenditures logit model has not been used before in empirical work, although it is equally tractable as the unit demand model. It can easily be integrated in Berry's (1994) nested logit and in Berry, Levinsohn and Pakes' (1995) random coefficients logit model. Only three simple modifications of the typical unit demand set-up are required: (i) price enters logarithmically instead of linearly, (ii) market shares are expressed in values instead of volumes, and (iii) the potential market size refers to the potential aggregate expenditures (in values) instead of the potential number of consumers or households. Apart from the additional flexibility from a new functional form for the price variable, the constant expenditures specification has a particular feature that may also be relevant in other applications: the pattern of price elasticities across models is quasi-independent of price, instead of quasi-linearly increasing in price as in unit demand specifications.

Our simulation model also provides greater flexibility on the supply side. We do not only allow for a standard multi-product Bertrand Nash model. We also allow for the possibility that firms partially coordinate, already before the merger. We introduce a partial coordination parameter, the weight that firms give on their competitors' profits when setting prices. This enables one to better calibrate the premerger marginal costs if reliable outside information on cost is available. Miller and Weinberg (2015) also use a partial coordination parameter, though in a different way: they scale it to zero before the merger and estimate the extent to which partial coordination between the merging firms and the outsiders increased after the merger.

Ex post merger evaluation Ex post merger analysis has moved in parallel with merger simulation, and mainly aimed to evaluate the relevance or effectiveness of competition policy towards mergers. Early work focused on mergers in major industries, such as airline markets (Borenstein, 1990; Kim and Singal, 1993), banking (Facacelli and Panetta, 2003), petroleum (Hastings, 2004; Gilbert and Hastings, 2005; Hosken, Silvia and Taylor, 2011) and appliances (Ashenfelter, Hosken and Weinberg, 2013). Ashenfelter and Hosken (2010) take advantage of scanner data to assess mergers in five different branded goods industries. They find moderate but significant price effects in the range of 3–7%. Among other things, they argue that their estimates may be viewed as a lower bound on price increases that would have occurred for other mergers that were blocked.

Ex post evaluation of merger simulation There is only a small recent literature that combines both traditions to compare the predictions from merger simulations with the actual merger effects. Peters (2006) looks at the simulated and actual price increases by the merging firms' in several airline mergers. Weinberg (2011) and Weinberg and Hosken (2013) look at the price increases of both the merging firms and their competing rivals. Friberg and Romahn (2015) look at price effects after a merger with divestiture. Relative to this interesting work, we consider a large merger between firms that are the only ones in their market segment. This resulted in large price effects, which enables us to test a richer set of predictions, in particular focus on the predicted price effects of the individual merging firms and the responses of the various outsiders. More broadly speaking, testing a broader set of predictions is of interest beyond evaluating the performance of merger simulations. It sheds light on the relevance of policy counterfactuals in a variety of other oligopoly settings with differentiated products (such as environmental policies, trade policies, taxation, etc).

The paper is organized as follows. Section I discusses the industry background, including the merger decision and the dataset. Section II develops the framework for merger simulation, as developed during the investigation. Section III discusses the empirical results for the demand model and merger simulations. Section IV provides the ex post analysis. We first present additional predictions from the merger simulations, not presented during the case but based on the same methodology. Next we confront these predictions with what actually happened in terms of prices and market shares of the merging firms and their competitors.

I. The market, the merger and its effects

In April 2009, the Swedish competition authority cleared the acquisition of AstraZeneca Tika (AZT) by GlaxoSmithKline (GSK). In this section we provide the relevant industry background, the data, the merger and its effects. These facts will motivate our analysis on the performance of merger simulation in the next sections.

A. The market for OTC painkillers

Substances and forms Over-the-counter analgesics or painkillers are non-prescription drugs to treat pain and fever. Painkillers come in three main active substances: paracetamol (called acetaminophen in the U.S.), ibuprofen and acetylsalicylic acid (ASA or aspirin). There are also two less important active substances: diclofenak and naproxen. The active substances may differ in the types of pains they relieve and in their side effects. Paracetamol treats most pains and fevers, and is known for having little side effects (except that it may damage the liver). Ibuprofen also treats most pains and fevers and is often used to reduce inflammations, but it may have side effects on the stomach. The ASA substance also has a blood-diluting effect, which has both advantages and disadvantages. Each active substance may therefore relieve pain and reduce fever in different ways and with different side effects.

Painkillers also come in various administrative forms. Tablets are the most important form, followed by fizzy tablets. There are also some other forms (such as liquid, suppository and powder), but these are much less important. Table 1 shows the market shares of the three main substances and the two main administrative forms, according to the total value of sales in 2008. With a market share of 42%, paracetamol is by far the most important substance. Ibuprofen and ASA each have a comparable market share of 29%. Paracetamol and Ibuprofen are mainly sold as tablets, whereas ASA is dominantly sold as fizzy tablets.

Firms and brands All companies specialize in one or at most two active substances. They typically sell one main brand per active substance, and sometimes an additional smaller brand. Table 2 shows the 2008 market shares of the companies and their brands, broken

Table 1: Market shares in 2008, by form and active substance

Form	Paracetamol	Ibuprofen	ASA	Total
Tablet	36.1	29.0	2.6	67.7
Fizzy tablet	6.0		26.3	32.3
Total	42.1	29.0	28.9	100

Note: This table shows the market shares of the main administrative forms and active substances, according to the total value of sales in 2008. Paracetamol is known as acetaminophen in the U.S.

down by active substance. This shows that the two merging companies AZT and GSK are the only companies in the paracetamol segment: AZT sells Alvedon as its main brand and Reliv as a smaller brand, whereas GSK sells the popular brand Panodil. McNeil (selling Ipren) and Nycomed (selling Ibumetin) are the main companies in the Ibuprofen segment. McNeil (selling Treo) is by far the largest company in the ASA segment. There are two other companies with much smaller market shares: Meda and Bayer.

While consumers may base their purchasing decision on the active substance and its associated medical effects, their perceptions regarding the companies' brands may also be important. This is evident from the large amount of advertising in the sector. So it is ultimately an empirical question to which extent brands with different active substances are substitutes.

Note that none of the active substances are still under patent, so in principle any firm can produce painkillers with one of the active substances for pain relief. In practice, entry in the painkiller market is infrequent, which may partly be due to the advertising costs to establish a new brand. It might also be partly due to the fact that the distribution of pharmaceuticals has historically been controlled by a state-owned monopoly. We turn to this next.

Distribution Until the deregulation of 2009, the companies distributed all their drugs through the state-owned pharmacy monopoly, Apoteket AB. In 2008 Apoteket operated 850 community pharmacies, 76 hospital pharmacies and 30 shops for over-the-counter and health care services. The pharmaceutical companies determined the wholesale prices, but indirectly also the retail prices, since Apoteket applied a fixed percentage markup on the wholesale prices. Apoteket also did not supply its own private labels.

After a market investigation, the Swedish government decided to deregulate the distribution of pharmaceutical products at the end of 2009. Several state pharmacies were sold to private companies, and non-pharmacy retail outlets became entitled to sell non-prescription drugs. The reforms also gave more freedom to the pharmacies in various respects (e.g. removal of universal supply and uniform price obligations). The government expected that the deregulation of the distribution system would increase competition, encourage entry of new products and lead to the establishment of new distribution channels, such as sales in supermarkets. In our analysis below, the post-deregulation data covers sales from the previous state pharmacies and other pharmacy chains, but no sales from supermarkets. We do not expect this to have an impact on our results, since the post-regulation period is short and started with a considerable delay after the merger.

Table 2: Market shares in 2008, by brand and active substance

Firm	Brand	Paracet.	Ibupr.	ASA	Total
AZT	Alvedon	29.3			31.5
	Reliv	2.2			
GSK	Panodil	10.6			10.6
McNeil	Ipren		19.1		44.7
	Treo			22.5	
	Magnecyl			3.1	
Nycomed	Ibumetin		9.2		9.2
Meda	Alindrin		0.7		3.6
	Albyl			0.2	
	Bamyl			2.7	
Bayer	Aspirin			0.4	0.4
	Alka-seltzer			0.0	
Total		42.1	29.0	28.9	100

Note: This table shows the market shares of the main firms and brands and active substances, according to the total value of sales in 2008. Paracetamol is known as acetaminophen in the U.S.

B. The merger

GSK notified its planned acquisition of AZT's painkiller brands (Alvedon and Reliv) on December 22, 2008. Although the merging firms were the only competitors in the paracetamol segment, the Swedish competition authority formally cleared the merger on April 3, 2009.[1] In its Decision, the competition authority described that it based its analysis on a large number of contacts in the industry and on its own analysis, which included a merger simulation study that we had conducted during the investigation; for more details, see our earlier working paper, Björnerstedt and Verboven (2012).

The competition authority justified its Decision to clear the merger on the grounds that consumers base their decisions more on the brand than on the active substance (where the main brands are Alvedon, Ipren and Treo). Furthermore, and probably more importantly, the competition authority stated that it expected increased competition because of the coming deregulation of the state-owned pharmacy monopoly. This view is well summarized in the competition authority's 2009 Annual Report:

> "GSK and AZT were the only companies providing over-the-counter (OTC) pharmaceuticals on the Swedish market that included the active substance "paracetamol", i.e. Alvedon, Reliv and Panodil. Much of the work associated with the investigation involved assessing the potential effects of the pending deregulation of the pharmacy market. Deregulation would mean that players other than Apoteket would be able to provide OTC pharmaceuticals and at the same time pharmaceutical companies would no longer be able to determine prices for customers. Deregulation would also enable new pharmaceutical stakeholders to enter the Swedish self-care market with their brands; for example including the paracetamol substance. In this way, the buying power of pharmacies and retailers would improve, which could possibly result in improved price competition between the different products available in the self-care market. After conducting a special investigation, the Swedish Competition Authority found that GSK's

[1] The justification of the Decision was very short, see p. 5-6 on http://www.kkv.se/upload/Filer/Konkurrens/2009/Beslut/beslut_08_0706_2008.pdf (in Swedish).

acquisition of AZT would not manifestly impede effective competition and no action was taken regarding this concentration."[2]

Not all of these arguments may be economically well-grounded, but this quote makes it clear that the competition authority hoped the deregulation would result in more competition. Finally, the competition authority wrote that the simulation study showed that mergers would not lead to significant price increases. However, the simulation study only predicted insignificant price increases in one scenario with large marginal cost savings (of at least 25%). This suggests that the competition authority may implicitly have had in mind rather large efficiencies.

Whatever the competition authority's economic arguments to clear the merger (brands more important than active substance, competition from deregulation or large efficiencies), there may also be political factors explaining the decision: the upcoming deregulation which had not been finalized, or a costly court case when a merger is prohibited.

C. Dataset

Our main dataset comes from the national distributor Apoteket AB for the period before the deregulation, January 1995 until December 2009, and from Apotekens Servicebolag AB for the period afterwards, January 2010 until May 2011. For the period after the deregulation, the data includes other pharmacy chains than the former state monopoly Apoteket AB, except for sales from supermarkets. The dataset and contains product-level sales information for Sweden, at a monthly frequency during the period January 1995 until May 2011. A product is defined as a brand, form, package size and dose. For example, one of AZT's products is Alvedon tablet, 30 pieces, 500 mg/piece. An observation for product j in month t contains information on the total sales value or revenue across all pharmacies in Sweden, r_{jt}, and the total sales volume, q_{jt}, from which we compute the price per unit $p_{jt} = r_{jt}/q_{jt}$.

The sales dataset was combined with two other datasets: one on marketing expenditures by brand and month (collected by Sifo RM), and one on macro-economic variables (from

[2] See http://www.kkv.se/t/Page____5925.aspx.

Statistics Sweden), such as nominal and real GDP, the number of men and women on sick leave (all monthly) and total population of men and women (yearly).

Note that there is no unambiguous measure for the unit of consumption in the market for painkillers, and hence no obvious measure for the sales volume q_{jt} and the price per unit p_{jt} of each price. In particular, it is not appropriate to measure q_{jt} as the number of sold packages and p_{jt} as the price per sold package, since the products are sold in different package sizes (number of tablets) and in different doses (mg per tablet). We consider three different measures for the unit of consumption: the "tablet" (or fizzy tablet), the "defined daily dose" (ddd) as defined by the World Health organization, and the "normal dose", i.e. the number of doses used on a normal single consumption occasion. We thus have three measures of the sales volume q_{jt} and three corresponding measures of the price p_{jt}: price per tablet, price per ddd, and price per normal dose. Note that these price measures correspond to the actual transaction price paid by every consumer, since pharmacy chains set uniform prices across all pharmacies in Sweden.

Table 3 presents summary statistics of the main variables over the pre-merger period 1995-2008. We focus on products from the three main active substances (paracetamol, ibuprofen and ASA) and the two main administrative forms (tablets and fizzy tablets). This covers about 90% of the total value of sales of analgesics. The total number of observations is 7,240, which amounts to an average of 43 products per month. Total sales value r_{jt} per product/month is on average 1.24 million SEK. The number of tablets is on average 1.11 million across products and months, so the average price per tablet is 1.1SEK. The average price per normal dose is slightly higher, 1.6 SEK, and the average price per defined daily dose (ddd) is 6.0 SEK. We will focus our discussion on the results from the first measure (price per tablet and number of sold tablets), but the other three measures gave similar results.

D. The price and market share effects of the merger

We can now consider the price and market share effects following the merger. It will be useful to first summarize the results by segment, since the merging firms are the only firms in one of the segments (paracetamol) and these firms are not active at all in the other two

Table 3: Summary statistics for the Swedish market for analgesics, 1995-2008

Variable	Mean	St. Dev.	Min.	Max.
revenue ($r_{jt} = p_{jt}q_{jt}$)	1.24	2.56	.00	22.95
number of tablets(q_{jt})	1.11	2.19	.00	16.61
number of defined daily doses (q_{jt})	.21	.43	.00	3.07
number of normal doses (q_{jt})	.77	1.57	.00	11.08
price per tablet (p_{jt})	1.06	.46	.27	2.55
price per defined daily dose (p_{jt})	6.02	2.21	1.74	15.50
price per normal dose(p_{jt})	1.61	.60	.43	3.88
marketing	564.1	1445.7	0	13536
sickwomen	822.9	197.0	391	1204
sickmen	524.5	108.0	254	763
GDPnom (in billions)	621.6	107.4	443.2	859.7
popwomen (in thousands)	4524.2	54.8	4471.4	4652.6
popmen (in thousands)	4437.4	72.5	4366.1	4603.7

Note: 7240 observations (products, years, months). Sales value or revenue (r_{jt}) is in 1 million SEK (including VAT), price per unit (p_{jt}) is in SEK, sales volume (q_{jt}) is in 1 million. 1€ = 10.8 SEK, 1$ = 8.0 SEK in December 2008.

segments (ibuprofen and ASA). But we will also consider more detailed results by firm and segment.

Price effects Figure 1 shows the price evolution for the three main segments: paracetamol, ibuprofen and ASA, comparing two years before the merger (April 2007–April 2009) with two years after the event (May 2009–May 2011). The results are striking. In the paracetamol segment, where the merging firms AST and GSK are the only competitors, average prices sharply increased from about 1.5 SEK to 2 SEK, already one month after the merger. The price increase is especially striking since prices only showed a small increase in the six months prior to the merger, and they remained more or less constant after the sharp increase just after the merger. Only near the end of the period, there is a slight tendency of a price drop, perhaps associated with new entry threats following the deregulation (although entry was surprisingly slow). In contrast, in the ibuprofen segment prices remained stable after the merger, whereas in the ASA segment they appear to increase by a modest amount (from 1.4 SEK to 1.55 SEK). This suggests that the sharp price increase by the merging firms was not

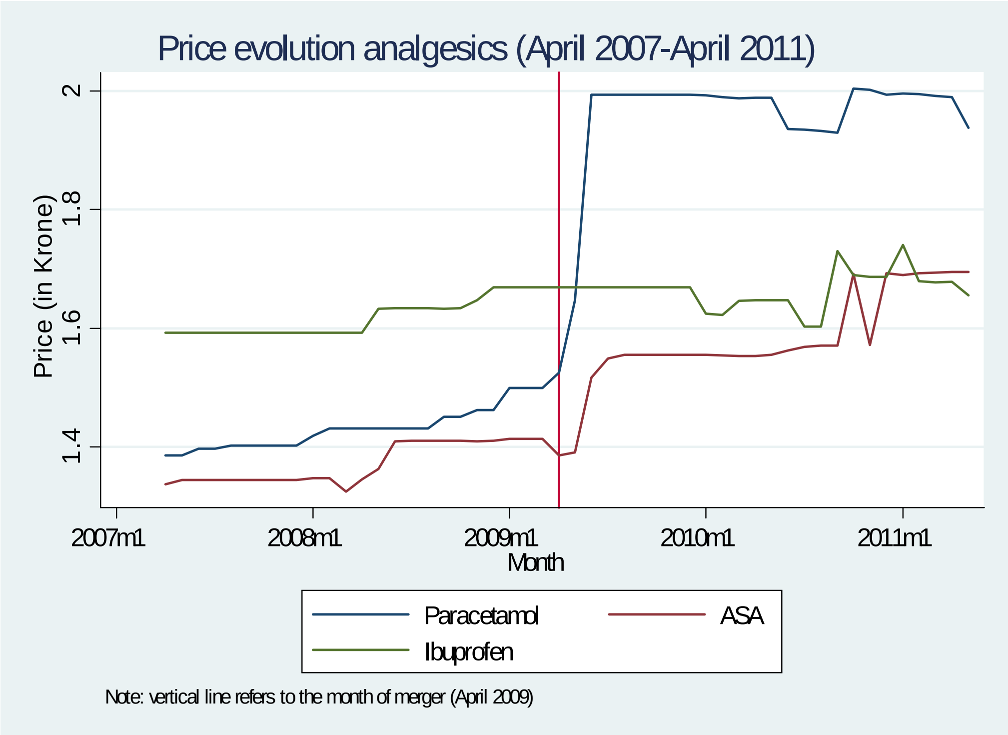

Figure 1: Price evolution analgesics (April 2007 - April 2011)

due to a general cost or demand shock unrelated to the merger.

To gain further insights on this, we estimate the following regression, in line with Ashenfelter and Hosken (2010) and other recent work:

$$\ln p_{it} = \alpha_i + \beta_i PostMerger_t + \varepsilon_{it}, \tag{1}$$

where p_{it} is the average price of "product group" i, and $PostMerger_t$ is a dummy variable equal to 1 after the merger event.[3] The literature sometimes assumes that the merger does not have an impact on the competitors' prices. If this assumption is satisfied, one can interpret this regression as a difference-in-difference estimator, where the difference between the merging firms' β_i and the competitors' β_i measures the merger price effect. In practice, it is possible that the merger raises the competitors' prices (under Bertrand competition,

[3] Our specification is slightly more general than Ashenfelter and Hosken (2008) and other work. They typically constrain the same effect for the control group after the merger, whereas we allow different product groups i to have different price changes.

but especially if there is some coordination). If this is the case, the difference between the merging firms' and the competitors' β_i's can be viewed as a lower bound for the merger price effect.

We define the product group i in the above regression at two levels: the substance and the substance×firm, and use a comparison window of one year before and after the merger (to exclude possible changes after the deregulation). Table 4 shows the results. According to the top left panel, the merger led to an average price increase in the merging firms' paracetamol segment of 39.7%. At the same time, the merger left prices in the ibuprofen segment essentially unchanged (+0.1%). But the prices in the ASA segment increased by 13.3%.

The bottom left shows the estimated price effects at the level of the substance×firm. The merging firms, who are the only ones in the paracetamol segment, raised their prices substantially and more or less proportionately: AZT by 39.2% and GSK by a slightly larger amount of 40.9%. The competitors raised their prices by much lower amounts. In the ibuprofen segment, price increases were very low: McNeil raised its prices by only 0.4% and Nycomed by only 0.9%, while Meda did not change its prices. In the ASA segment, firms raised prices by higher amounts: McNeil by 17.9%, Bayer by 6.8% and Meda by 8.9%.

Why did the large and sudden price increase by the merging firms not raise a significant amount of controversy in Sweden? In fact, the merged firm AZT-GSK implemented the price increase by reducing their main package sizes from 30 to 20 tablets, while reducing the price per package by only a small amount, for example from 41.5 crowns to 38.5 crowns for one of their most selling products. The reduction in the maximum package size had been required by the Swedish medical products agency (Läkemedelsverket), because of concerns with an increase in cases of painkiller overdoses. The firms argued that the resulting increase in the price per tablet was warranted because of the increased costs with the reduced package size. However, it is rather implausible that this explains the entire price increase of 40%, because the companies in the ASA segment had also been required to lower their package sizes and they only raised prices by on average 13%. In our merger simulation analysis below, we will consider more systematically how reduced package size may have raised costs and to which

Table 4: Actual price and market share effects, two year window

	Price (% change)		(%)	Market share (% point change)	
	Coeff	St. Err	Before	Coeff	St Err
Regressions at the level of the substance					
substance fixed effects	yes				yes
Paracetamol×merger	39.7	(1.0)	47.0	-3.6	(0.6)
Ibuprofen×merger	0.1	(1.7)	27.3	4.3	(0.2)
ASA×merger	13.3	(4.7)	25.8	-0.8	(0.5)
R^2	.969			.986	
Regressions at the level of the firm × substance					
firm×substance fixed effects	yes				yes
Paracetamol					
AZT×merger	39.2	(1.2)	34.2	-3.6	(0.4)
GSK×merger	40.9	(2.1)	12.7	0.0	(0.3)
Ibuprofen					
McNeil×merger	0.4	(1.1)	17.1	2.0	(0.2)
Meda×merger	0.0	(0.1)	0.7	0.0	(0.1)
Nycomed×merger	0.9	(1.1)	9.5	1.5	(0.2)
ASA					
McNeil×merger	17.9	(4.8)	21.4	-2.3	(0.4)
Meda×merger	8.9	(5.4)	3.7	1.1	(0.2)
Bayer×merger	6.8	(1.0)	0.6	0.4	(0.1)

Notes: This table shows actual price and market share effects, based on the regression model (1) for price and analogous model for market share. The percentage price effects are obtained from a transformation of the parameters β_i using $\exp(\beta_i) - 1$ (and a corresponding adjustment of the standard errors using the delta method). Robust standard errors are reported.

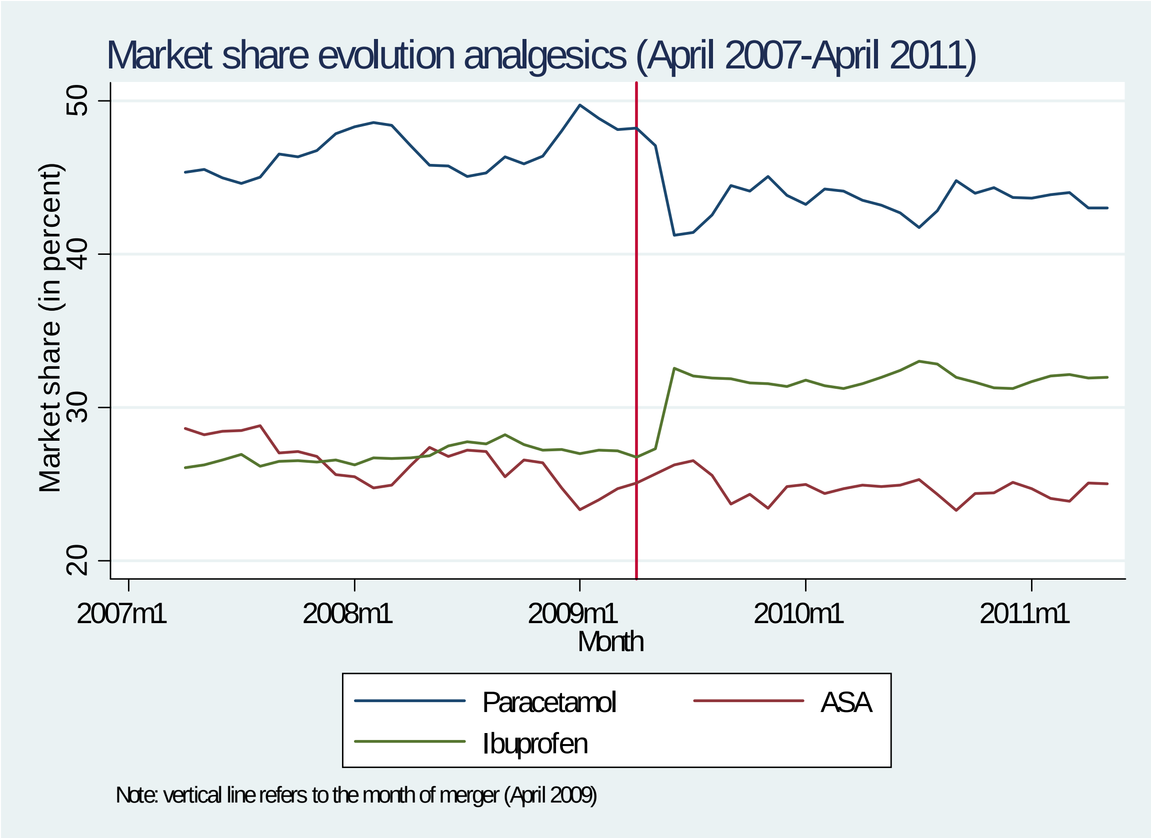

Figure 2: Market share evolution analgesics (April 2007 - April 2011)

extent this may have been responsible for the price rises. Before doing this, we first consider how the changes in market shares following the merger.

Market share effects Did the large price increase of the merging firms also affect market shares? Figure 2 shows the market share evolution (in volumes), using the same comparison window as Figure 1. This shows that the market share of the merging firms' paracetamol segment suddenly dropped by a sizeable 5%, down from about 47% to about 42%. The market share of ibuprofen (where prices did not change) increased sharply, from about 27% to 32%. The market share of ASA (where prices moderately increased) remained more or less unchanged. It is less clear from Figure 2 whether these market share changes were permanent, since they show some volatility over the sample. We therefore estimated a regression similar to (1), but with the log of price replaced by the market share as the dependent variable.

The right panel of Table 4 shows the results. The market share of the merging firms' paracetamol segment dropped by a significant 3.6% over the considered period (95% confidence interval of 2.4%–4.8%). This loss was entirely in favor of the ibuprofen market share, which increased by a substantial 4.3%. The market share of ASA decreased by 0.8%, consistent with our earlier finding that ASA prices increased rather substantially after the merger (in contrast with ibuprofen prices).

Interesting additional findings obtain for the market shares at the level of the substance×firm (bottom right panel in Table 4). Despite the fact that prices increased slightly more for GSK than for AZT products, only AZT experienced a market share drop (by –3.6%); the market share of GSK remained more or less unchanged, perhaps because GSK was able to give its brand greater visibility after the merger. In the ibuprofen segment, only McNeil experienced a market share increased (while Meda's market share remained unchanged). Finally, in the ASA segment, McNeil (-2.6%) lost market share to Meda (+2.0%), consistent with the earlier finding that McNeil raised its prices by a larger amount than Meda.

Summary The merger led to a large price increase by the merging firms in the paracetamol segment, and a corresponding market share drop (although entirely stemming from the largest company, AZT). Prices of the competitors in the ASA segment also partly increased after the merger, but only McNeil experienced a corresponding market share drop. Finally, prices in the ibuprofen segment remained more or less unchanged, and market shares increased (mainly for McNeil). In the next sections we evaluate how well merger simulation predicts these facts, and what factors should be taken into account to rationalize the facts.

II. Framework for merger simulation

We now present the framework for the merger simulation. We first motivate and discuss our adopted demand model, used to estimate the substitution patterns across products. We then present the model of oligopolistic price-setting behavior, used to uncover premerger marginal costs and to predict post-merger prices.

A. Demand model

Our demand model is a discrete choice model, which incorporates unobserved consumer heterogeneity in the valuation of various product characteristics. While discrete choice models were initially developed for estimation with micro-level choice data, Berry (1994) and Berry, Levinsohn and Pakes (1995), henceforth BLP, show how they can be estimated with aggregate sales data to generate rich substitution patterns. Popular models include the nested logit and the random coefficients logit model, and we consider both models.

Our first demand model is a two-level nested logit model, which allows preferences to be correlated along two discrete product dimensions: the products' active substance (paracetamol, ibuprofen and ASA) and their administrative form (tablet or fizzy tablet). The nested logit model allows cross-price elasticities to be greater between products that have the same substance and/or form. Segmentation according to substance is particularly relevant, since the merging companies are the only ones active in the paracetamol segment. However, the nested logit model is also potentially restrictive: it restricts substitution in a hierarchical way (with nests and subnests) and it does not allow for correlated preferences along continuously measured product characteristics. We therefore also consider a second demand model: BLP's random coefficients logit model. This gives greater flexibility in substitution patterns though at an increased computational cost and with practical numerical difficulties as recently documented in for example Knittel and Metaxoglou (2014).

While the discrete choice literature focuses on incorporating sufficient consumer heterogeneity to obtain rich substitution patterns, we were especially concerned with the typically adopted and potentially restrictive functional form for the price variable. The aggregate discrete choice literature since Berry (1994) and Berry, Levinsohn and Pakes (1995) has adopted a utility specification where price enters additively with income. This specification has the property that consumers buy one unit of their preferred product. While this may be an appealing property for some commodities such as automobiles, it may be less realistic for many frequently purchased consumer items. More importantly, the price specification implies that the price elasticities of different products are quasi-linearly increasing in prices: if product A is twice as expensive as product B, it also tends to have a price elasticity that is twice as high. This property does not only hold in the logit and nested logit model; it may also be present to some extent in the random coefficients logit model.

For example, in an interesting paper on the same industry, Chintagunta (2002) estimates a random coefficients logit model for five (U.S.) painkiller brands.[4] Although he finds significant heterogeneity for the price parameter, the estimated own-price elasticities show an increasing relationship with prices across products.[5] This pattern is not unrealistic per se, but it does follow from the linear price specification. In our application, we were particularly concerned with the linear price specification because, unlike Chintagunta (2002), we have many brands and, as shown in Table 3, prices vary by a factor of more than nine.

We therefore consider an alternative possible utility specification, where price (as well as income) enters logarithmically instead of linearly, implying that consumers do not buy one unit but a constant expenditure of their preferred product. We build on the work of Hanemann (1984), who proposed a framework to model discrete-continuous choices, and showed how to estimate such models with micro-level choice data, as applied in e.g. Hendel (1999) and Dubé (2004). We show how our utility specification leads to a natural extension of the aggregate discrete choice demand models of Berry (1994) and BLP, with three differences: price enters logarithmically instead of linearly, market shares are measured in values instead of volumes, and the potential market refers to the potential aggregate budget instead of the potential number of consumers. The implied own- and cross-price elasticities are quasi-constant instead of quasi-linearly increasing in price, as in the unit demand model. To our knowledge, no other work has departed from the unit demand model in discrete choice models with aggregate sales data.

In the discussion below, we compare the unit demand and the constant expenditure specification in an aggregate nested logit model. In the Appendix, we show how this also extends to a random coefficients logit model.

[4] To our knowledge, there are no other papers estimating discrete choice models for painkillers at the brand level. Chevalier, Kashyap and Rossi (2001) estimate a log-log demand model at the category level, and obtain an estimated price elasticity for the painkiller category equal to -1.87.

[5] Tables 2 and 5 in Chintagunta (2002) show the following relationship between own-price elasticities and average prices: Advil -2.996 vs. 7.41; Tylenol, -2.69 vs. 6.16; Motrin -2.66 vs. 5.95; Bayer -2.25 vs. 4.95; Store -1.81 vs. 3.55. This pattern is also present in other logit or random coefficients logit applications.

Utility There are L consumers, $i = 1, \ldots, L$. Each consumer chooses one out of $J + 1$ differentiated products, $j = 0, \ldots, J$; good 0 is the outside good or no-purchase alternative. Suppose consumer i has the following conditional indirect utility for good $j = 0, \ldots, J$:

$$u_{ij} = \mathbf{x}_j \boldsymbol{\beta} + \xi_j + \alpha f(y_i, p_j) + \varepsilon_{ij}, \tag{2}$$

where $\mathbf{x}_j$ is a vector of observed product characteristics of product j, p_j is price, ξ_j captures unobserved product characteristics, y_i is income of individual i, $\boldsymbol{\beta}$ and α are utility parameters, and ε_{ij} is a random utility term or an individual-specific taste parameter for good j.

Conditional on buying product j, a consumer i's demand for product j follows from Roy's identity, $d_j(y_i) = -(\partial f / \partial p_j) / (\partial f / \partial y_i)$. We consider the following two specifications for $f(y_i, p_j)$:

$$\begin{array}{llcll} \text{Unit demand} & f(y_i, p_j) & = & y_i - p_j & \Rightarrow \quad d_j(y_i) = 1 \\ \text{Constant expenditures} & f(y_i, p_j) & = & \gamma^{-1} \ln y_i - \ln p_j & \Rightarrow \quad d_j(y_i) = \gamma \frac{y_i}{p_j} \end{array} \tag{3}$$

Conditional on choosing j, an individual buys one unit in the first specification, and spends a constant fraction of her budget, γ, in the second specification. The first specification is typically adopted in aggregate discrete choice models, sometimes under a variant such as BLP's Cobb Douglas specification, $f(y_i, p_j) = \ln(y_i - p_j)$, which also implies unit demand since income and price still enter additively. The second specification is a special case of Hanemann's framework for micro-level discrete choice models, and we will show here how it can be incorporated in an aggregate discrete choice framework.

For the two specifications (3), we can write utility (2) more compactly as follows

$$u_{ij} = K_i + \delta_j + \varepsilon_{ij}, \tag{4}$$

where in the unit demand specification $K_i = \alpha_i y_i$ and $\delta_j \equiv x_j \beta - \alpha p_j + \xi_j$; and in the constant expenditures specification, $K_i = \alpha_i \gamma^{-1} \ln y_i$, and $\delta_j \equiv \mathbf{x}_j \boldsymbol{\beta} - \alpha \ln p_j + \xi_j$. Intuitively, one can interpret δ_j as the mean utility component of product j. In both specifications, we normalize the mean utility of the outside good to zero, $\delta_0 = 0$.

Choice probabilities Each consumer i chooses the product j that maximizes her random utility u_{ij}. Assume that the random utility terms follow the extreme value distributional assumptions of a two-level nested logit model. Partition the set of products into G groups, $g = 0, \ldots, G$ (where group 0 consists of the outside good 0) and further partition each group g into H_g subgroups, $h = 1, \ldots, H_g$. Each subgroup h of group g contains J_{hg} products, so that $\sum_{g=1}^{G} \sum_{h=1}^{H_g} J_{hg} = J$.

Given random utility maximization, the probability that a consumer i chooses product $j = 1, \ldots, J$ takes the following well-known form:

$$s_j = s_j(\boldsymbol{\delta}, \sigma) \equiv \frac{\exp(\delta_j/(1-\sigma_1))}{\exp(I_{hg}/(1-\sigma_1))} \frac{\exp(I_{hg}/(1-\sigma_2))}{\exp(I_g/(1-\sigma_2))} \frac{\exp(I_g)}{\exp(I)}, \tag{5}$$

where I_{hg}, I_g, and I, are the inclusive values or "log sum" formulas (see Appendix), $\boldsymbol{\delta}$ is a $J \times 1$ vector containing the mean utilities δ_j, and $\sigma = (\sigma_1, \sigma_2)$ are the nesting parameters associated with the nested logit distribution. Note that the separable terms K_i cancel out from the choice probabilities (5).

The nesting parameters capture the preference correlation across products of the same subgroup (σ_1) or group (σ_2), and should satisfy $1 \geq \sigma_1 \geq \sigma_2 \geq 0$ (McFadden, 1978). When σ_1 is high, preferences are strongly correlated across products of the same subgroup, and when σ_2 is high, preferences show additional correlation across products of the same group. If $\sigma_1 = \sigma_2 = 0$, the model reduces to a simple logit model, so that preferences are not correlated across products from the same subgroups or groups.

Aggregate and inverted aggregated demand Aggregate demand for a product j is the probability that a consumer buys that product, multiplied by the quantity purchased,

$d_j(y_i)$, aggregated over all L consumers according to income distribution P_y:

$$\begin{aligned} q_j &= \int s_j(\boldsymbol{\delta},\sigma)\, d_j(y)\, \mathrm{d}P_y(y)\, L \\ &= s_j(\boldsymbol{\delta},\sigma) \int d_j(y)\, \mathrm{d}P_y(y)\, L. \end{aligned}$$

The second equality follows from the fact that the choice probability $s_j(\boldsymbol{\delta},\sigma)$, given by (5), does not depend on income. Using (3), we can solve the remaining integral. For the unit demand specification, we simply have $\int d_j(y) \mathrm{d}P_y(y)\, L = L$, whereas for the constant expenditures specification we have $\int d_j(y) \mathrm{d}P_y(y)\, L = \gamma Y/p_j$, where $Y \equiv \int y \mathrm{d}P_y(y)\, L$ is total income of all consumers. Substituting and rearranging then gives expressions for the choice probabilities in terms of observables:

$$\begin{array}{ll} \text{Unit demand} & \frac{q_j}{L} = s_j(\boldsymbol{\delta},\sigma) \\ \text{Constant expenditures} & \frac{p_j q_j}{B} = s_j(\boldsymbol{\delta},\sigma) \end{array} \tag{6}$$

where we define $B = \gamma Y$ as the total potential budget allocated to the differentiated products in the economy, a constant fraction γ of total income of all consumers Y. Hence, the choice probabilities are equal to the market shares in volume terms for the familiar unit demand specification, whereas they are equal to market shares in value terms for the constant expenditures specification.

The goal is to estimate the parameters (α, β, σ) entering the demand system (6). The econometric error term ξ_j enters non-linearly through the mean utility terms δ_j. To obtain a tractable model, we can follow the same approach as proposed by Berry (1994) for both specifications, i.e. invert the system of choice probabilities $s_j = s_j(\boldsymbol{\delta},\sigma)$, $j = 1,\ldots,J$, to solve for the mean utilities $\delta_j = \delta_j(\mathbf{s},\sigma)$. Following Berry (1994) for the one-level nested logit and Verboven (1996) for the two-level nested logit, we obtain an analytical solution for the inverted choice probability system:

$$\ln(s_j/s_0) = \sigma_1 \ln(s_{j|hg}) + \sigma_2 \ln(s_{h|g}) + \delta_j, \tag{7}$$

where $s_{j|hg}$ is the market share of j within subgroup hg, and $s_{h|g}$ is the market share of subgroup hg in group g.

In the familiar unit demand specification, one can substitute $\delta_j \equiv \mathbf{x}_j\boldsymbol{\beta} - \alpha p_j + \xi_j$, and the market shares are in volume terms and relative to the total number of consumers L. In the constant expenditures specification, there are three differences. First, one should substitute $\delta_j \equiv \mathbf{x}_j\boldsymbol{\beta} - \alpha \ln p_j + \xi_j$, so price enters logarithmically instead of linearly. Second, one should substitute the market shares in value terms, as evident from (6). Third, the potential market is now the total potential budget as a fixed fraction γ of GDP, $B = \gamma Y$, instead of the total number of buyers, L.[6] We will not estimate γ, but impose a specific value (or range), similar to the practice of imposing values for L in unit demand specifications.

Both variants of (7) are linear in the error term ξ_j. They can be estimated using an instrumental variable regression of volume or value market shares (relative to outside good market shares) on product characteristics, price (or log price) and subgroup and group market shares, where the endogenous variables are price and the (sub)group market shares. Note that the same type of data that identify the unit demand specification, exogenous product characteristics, also identify the constant expenditures demand specification (except for requiring a different specification of the potential market).

In the Appendix, we show how to extend the constant expenditure specification to BLP's random coefficients model: (6) still applies, where $s_j(\boldsymbol{\delta}, \sigma)$ is now replaced by BLP's usual market share integral. The Appendix also provides further details for the nested logit, where we show that the price elasticities of demand are quasi-constant in price for the constant expenditure specification, instead of quasi-linear in price for the unit demand specification.

[6] Some other papers have used a logarithmic price term, for example Peters (2006) or Gowrisankaran and Rysman (2012). Verboven (1996) uses a Box-Cox transformation of the price term, $\left(p_j^{\mu} - 1\right)/\mu$ to nest both the linear and logarithmic specifications. While these approaches are useful to obtain a more flexible functional form for price, they are not consistent with utility maximization. As we show here, the logarithmic specification can be made consistent after some simple adjustments regarding the computation of market shares and the potential market (and it is straightforward to generalize this to the Box-Cox transformation, but the model is then no longer linear in the parameters).

B. Oligopoly model

The oligopoly model serves two purposes. First, in combination with the demand parameters it enables one to uncover the premerger marginal costs. Second, based on the demand parameters and uncovered marginal costs, it can be used to predict the price effects of the merger.

Each firm f owns a portfolio of products F_f. Its total variable profits are given by the sum of the profits for each product $k \in F_f$:

$$\Pi_f(\mathbf{p}) = \sum_{k \in F_f} (p_k - c_k)\, q_k(\mathbf{p}) \tag{8}$$

where c_k is the constant marginal cost for product k and $q_k(\mathbf{p})$ is demand, as given by (6), now written as a function of the $J \times 1$ price vector $\mathbf{p}$. The profit-maximizing price of each product $j = 1, \ldots, J$ should satisfy the following first-order condition:

$$q_j(\mathbf{p}) + \sum_{k \in F_f} (p_k - c_k)\, \frac{\partial q_k(\mathbf{p})}{\partial p_j} = 0. \tag{9}$$

A price increase affects profits through three channels. First, it directly raises profits, proportional to current demand $q_j(\mathbf{p})$. Second, it lowers the product's own demand, which lowers profits proportional to the current markup. Third, it raises the demand of the other products in the firm's portfolio, which partially compensates for the reduced demand of the own product. If the first-order conditions (9) hold for all products $j = 1 \cdots J$, a multiproduct Bertrand-Nash equilibrium obtains.

To write this system of J first-order conditions in vector notation, define the $J \times J$ matrix $\boldsymbol{\theta}^F$ as the firms' product ownership matrix, a block-diagonal matrix with a typical element $\theta^F(j,k)$ equal to 1 if products j and k are produced by the same firm and 0 otherwise. Let $\mathbf{q}(\mathbf{p})$ be the $J \times 1$ demand vector, and $\boldsymbol{\Delta}(\mathbf{p}) \equiv \partial \mathbf{q}(\mathbf{p})/\partial \mathbf{p}'$ be the corresponding $J \times J$ Jacobian matrix of first derivatives. Let $\mathbf{c}$ be the $J \times 1$ marginal cost vector. Using the operator $\odot$ to denote element-by-element multiplication of two matrices of the same dimension, we have

$$\mathbf{q}(\mathbf{p}) + \left(\boldsymbol{\theta}^F \odot \boldsymbol{\Delta}(\mathbf{p})\right)(\mathbf{p} - \mathbf{c}) = 0.$$

This can be inverted to give the following expression:

$$\mathbf{p} = \mathbf{c} - \left(\boldsymbol{\theta}^F \odot \boldsymbol{\Delta}(\mathbf{p})\right)^{-1} \mathbf{q}(\mathbf{p}), \tag{10}$$

which decomposes the price into two terms: marginal cost and a markup, which depends on the own- and cross-price elasticities of demand.

It is straightforward to generalize this expression to allow for (partial) coordinated behavior. Suppose that firms put a weight $\phi \in (0, 1)$ on the profits of their competitors and modify the objective function (8) accordingly. The same expression (10) then obtains, where the zeros in the matrix $\boldsymbol{\theta}^F$ are replaced by the parameter ϕ. We will focus on the non-cooperative case where $\phi = 0$. However, in an extension we also consider a case where $\phi > 0$, to see whether this brings the merger predictions closer to reality. One could also allow ϕ to vary across products, but since there is little information about the possibility and the extent of coordination we keep a simple specification.

Equation (10) serves two purposes. First, it can be rewritten to uncover the pre-merger marginal cost vector $\mathbf{c}$ based on the pre-merger prices and estimated price elasticities of demand, i.e.

$$\mathbf{c}^{pre} = \mathbf{p}^{pre} + \left(\boldsymbol{\theta}^{F,pre} \odot \boldsymbol{\Delta}(\mathbf{p}^{pre})\right)^{-1} \mathbf{q}(\mathbf{p}^{pre}). \tag{11}$$

Second, (10) can be used to predict the post-merger equilibrium. The merger involves two possible changes: a change in the product ownership matrix from $\boldsymbol{\theta}^{F,pre}$ to $\boldsymbol{\theta}^{F,post}$ and, if there are cost changes, a change in the marginal cost vector from $\mathbf{c}^{pre}$ to $\mathbf{c}^{post}$. To simulate the new price equilibrium, we used fixed point iteration on (10), where we apply a dampening factor less than 1 to the last term in case of no convergence. We also considered the Newton method and this gave the same results.

III. Empirical analysis

In this section we present the empirical results from various demand models, and we compare their predicted price effects under the most standard merger simulation where there are no other changes except firm ownership. In the next section we then focus on the demand models with price predictions closest to the actual price effects, and we discuss how various supply side assumptions may explain the differences between predicted and actual effects.

A. Specification and estimation

The nested logit model has form as the upper nest and substance as the lower nest, implying consumers are most likely to substitute to another product of the same form and substance, and would substitute more to another substance than to another form. The random coefficients logit model includes four random coefficients: for fizzy tablet, paracetamol, branded product (Alvedon, Ipren and Treo) and the constant. This model has the advantage that it does not rely on an a priori hierarchical nesting structure for the form and substance, and it enables us to incorporate the role of substitution between branded products as suggested to be relevant by the competition authority.[7] For both the nested logit and the random coefficients logit we estimate the unit demand and the constant expenditures specification.

For these various demand models, we define a product j as a brand, form, package size and dose. We include the following variables as determinants of mean utility (relative to the outside good): price (unit demand) or log of price (constant expenditures), marketing expenditures, the fraction of sick women and sick men in the total population, a time trend and monthly dummy variables capturing seasonal effects. In addition, since we observe a panel of multiple periods (all months during 1995-2008), we also include a set of fixed effects per product j. These fixed effects account for time-invariant unobserved product characteristics affecting mean utility, such as package size and dose. We can estimate the effects of these characteristics using a second stage regression of the fixed effects on these

[7] We use the univariate nested quadrature rules of Heiss and Winschel (2008) with a high accuracy level of 11 to approximate the market share integral. It is interesting to note that this integration method turned out to be very robust to the use of different starting values for the parameters.

product characteristics (as in Nevo, 2000).

Aggregate discrete choice models require one to determine the size of the potential market, i.e. the total number of potential consumers L in the unit demand and the total potential budget B in the constant expenditures specification. For both variants, we assume that the potential market is twice the average amount spent over the entire period (in units for the first and in values for the second specification). We performed a sensitivity analysis with alternative factors: 1.5, 2 (base), 4 and 6 and obtained similar results.

Finally, to estimate the model it is necessary to specify a reasonable set of instruments. We start from the commonly used identification assumption that the product characteristics, other than price, are uncorrelated with the error term. The products' own characteristics are then natural instruments, but additional instruments are required to identify the price coefficient and the distributional parameters (the nesting parameters in the nested logit and the standard deviations of the random coefficients in the random coefficients logit). BLP suggest to use functions of the other product characteristics as additional instruments.[8] For the nested logit model, our instrument set includes the products' own characteristics and counts of the number of other products: overall, by group, by subgroup, by firm, by firm and group and by firm and subgroup. The Appendix shows summary statistics on the variation of these instruments (due to product entry and exit), and also presents the first stage regressions of the endogenous variables (price and the shares $\ln s_{j|hg}$ and $\ln s_{h|g}$) on the instrument set. For the random coefficients logit model, we use the same instruments as in BLP in a first stage (sums of other product characteristics of the same firm and of other firms for each variable with a random coefficient), and optimal instruments in a second stage following Chamberlain (1987) and Berry, Levinsohn and Pakes (1999); Reynaert and Verboven (2014) provide detailed Monte Carlo evidence to demonstrate that optimal instruments may considerably improve the efficiency of the estimator. Note that, as in Chintagunta (2002), we treat marketing expenditures as an exogenous variable in both models. This assumption may

[8] More specifically, they suggest to use counts and sums of the characteristics of the other products of the same firm and of the other products of the other firms. For the nested logit model, Verboven (1996) suggested to take counts and sums by subgroups and groups as additional instruments. Bresnahan, Stern and Trajtenberg (1997) followed a similar approach for their "principles of differentiation" GEV model.

be justified to the extent that the set of product fixed effects takes away the main source of correlation with the error term. We also considered a specification where marketing expenditures are treated as endogenous (using the same instrument set), and this gave closely comparable results.

B. Parameter estimates, elasticities and predicted price effects

Parameter estimates Table 5 presents the estimated demand parameters for the four demand models: two-level nested logit and random coefficients logit, both under the unit demand and constant expenditures specification.

Consider first the results from the nested logit model (first two columns of Table 5). As in Chintagunta (2002), marketing expenditures have a positive effect on the products' demands. Relative to ASA, consumers have a higher mean valuation for paracetamol and ibuprofen. Unsurprisingly, consumers also have a higher valuation for the branded products. Consumers have a significantly higher valuation for products with a higher dosage. Finally, consumers do not value package size per se: they do not have a significantly different mean valuation for products that come in a higher or lower package size.

In both specifications the price coefficient α has the expected sign. The subgroup and group nesting parameters are fairly comparable ($\sigma_1 = 0.93$ and $\sigma_2 = 0.79$ in the unit demand specification, and $\sigma_1 = 0.86$ and $\sigma_2 = 0.69$ in the constant expenditures specification). These estimates satisfy the requirements for the model to be consistent with random utility theory, $1 \geq \sigma_1 \geq \sigma_2 \geq 0$. In both specifications, the inequalities are strict (in a statistically significant way). This implies that consumers perceive products of the same form and substance as the closest substitutes, products of a different substance but the same form as weaker substitutes, and products of different form as the weakest substitutes.[9]

Now consider the results from the random coefficients logit model (last two columns of Table 5). While the mean valuation parameters usually have the same sign and significance,

[9] We also estimated a model with the reverse nesting order (where consumers would substitute more to another form than to another substance), but this led to estimates of the nesting parameters $\sigma_1 < \sigma_2$, inconsistent with random utility theory. Following common practice (e.g. Goldberg, 1995), we therefore rule out this reverse nesting structure.

Table 5: Demand parameter estimates

	Nested logit		Random coefficient logit	
	Const Exp	Unit	Const Exp	Unit
constant	-6.941	-5.054	-31.682	-33.909
	(1.247)	(0.786)	(6.481)	(7.086)
price ($-\alpha$)	-0.289	-2.041	-1.616	-11.838
	(0.089)	(0.149)	(0.228)	(1.778)
marketing expenditures	13.384	8.905	92.304	113.806
	(2.456)	(1.782)	(11.189)	(22.373)
log(dosage)	0.757	0.813	3.540	5.067
	(0.195)	(0.123)	(1.015)	(1.110)
log(package size)	-0.025	-0.184	0.132	-0.397
	(0.082)	(0.051)	(0.425)	(0.464)
fizzy	-0.024	-0.277	-1.408	-4.931
	(0.097)	(0.061)	(0.503)	(0.550)
paracetamol	0.323	0.118	0.911	0.766
	(0.128)	(0.081)	(0.668)	(0.730)
ibuprofen	0.585	0.671	2.496	3.819
	(0.174)	(0.109)	(0.902)	(0.986)
branded	0.533	0.381	3.286	2.795
	(0.122)	(0.077)	(0.632)	(0.691)
subgroup (σ_1)	0.861	0.927		
	(0.017)	(0.013)		
group (σ_2)	0.690	0.791		
	(0.017)	(0.010)		
fizzy (σ)			4.396	7.470
			(0.504)	(2.060)
paracetamol (σ)			2.907	2.428
			(0.969)	(1.445)
branded(σ)			3.464	5.921
			(0.560)	(3.632)
constant (σ)			3.504	2.524
			(0.632)	(4.980)

Notes: 7240 observations for 1995-2008. Specifications also include the variables sickwomen and sickmen, monthly fixed effects and 56 product fixed effects. Robust standard errors are reported. Mean-utility coefficients of time-invariant variables are based on second stage regression of the product fixed effects.

their magnitudes tend to be much larger. This is because of a different scaling: parameters in the nested logit become of a more comparable order of magnitude if they are divided by $(1 - \sigma_1)$. The standard deviations of the random coefficients are statistically significant and have a large magnitude compared to the mean valuations. For example, in the constant expenditures specification, the standard deviation of the valuation is 4.4 for fizzy tablet and 2.9 for paracetamol, which are both about three times larger than the mean valuations (in absolute value). There is also significant consumer heterogeneity in the valuation of branded products, and there is considerable heterogeneity in the value of painkillers relative to the outside good.

Price elasticities and markups Table 6 summarizes what these parameter estimates imply for the price elasticities, the markups and the basic predicted price effects of the merger. The top part of Table 6 provides summary information on the own-price and cross-price elasticities at the level of the brand, i.e. the effect on brand demand after a price increase of all products of the same brand (all forms, dosages and package sizes). The numbers refer to the average (and range) across products during December 2008, the last month of the dataset used to estimate the demand model.

For the two nested logit models, we find the following. In the constant expenditures specification, the own-price elasticity is on average -2.4, and it ranges between -3.05 and -2.00. Furthermore, the cross-price elasticities are much larger for products of the same substance (on average 0.30) than for products of a different substance (0.05). There is a similar pattern in the unit demand specification, but the level of all elasticities is considerably higher. More interestingly, the range of price elasticities is much higher (varying between -15.5 and -5.2), and is here essentially proportional to the wide range in prices across brands. Note that the range of price elasticities is even higher at the product level (where the lowest elasticity is about 10 times lower than the highest), since at that level the price range is also much wider.[10]

[10] It is of interest to compare these estimates with the ones from a unit demand (random coefficients) logit, obtained by Chintagunta (2002). As discussed above, his estimated price elasticities for the five analgesics brands range between -1.8 and -3.0. These elasticities are also proportional to prices (but the range is smaller than in our case, since the price range is smaller).

Table 6: Price elasticties, markups, predicted price effects

	Nested logit		Random coefficients logit	
	Const Exp	Unit	Const Exp	Unit
	Brand-level price elasticities			
Own	-2.43	-8.84	-2.05	-3.61
Own (range)	(-3.05 ; -2.00)	(-15.45 ; -5.16)	(-2.61 ; -1.51)	(-6.52 ; -1.99)
Cross - same substance	0.30	2.12	0.12	0.36
Cross - different substance	0.05	0.28	0.05	0.19
	Markups (percent)			
Paracetamol	48.5	17.3	59.0	44.7
Ibuprofen	45.0	11.8	55.7	32.3
ASA	59.1	22.9	66.5	44.5
	Basic predicted price effects (percent)			
Paracetamol	41.1	15.9	21.29	9.34
Ibuprofen	1.2	0.5	0.75	0.76
ASA	1.5	0.6	0.81	0.45

Notes: All numbers are averages across products for December 2008, except for the numbers in parentheses, which refer to the range.

Now consider the random coefficients logit models. In the constant expenditures specification, the own-price elasticities are comparable to the nested logit model. The cross-price elasticities are also larger for products of the same substance (on average 0.12) than for products of a different substance (0.05). The difference seems less pronounced, but this is not the case if we consider only the cross-price elasticities within the paracetamol segment (which are on average 0.5).

Table 6 also shows the implied equilibrium markups under Bertrand-Nash behavior. The average markups vary somewhat across segments: they tend to be highest in the ASA segment, and are of comparable size in the paracetamol and ibuprofen segment. The competition authority's investigation suggested that the actual markups for paracetamol are in fact higher than the estimates in Table 6.[11] In particular, the estimated markups appear to be implausibly low in the unit demand specification of the nested logit (between 11.8% and 22.9%). In the other specifications, they seem of a more reasonable order of magnitude, especially for

[11] We cannot give precise details on the exact marginal cost information used to compute the markups (and this is likely not perfectly accurate in any case). Broadly speaking, the marginal cost information was broken down between the cost of the active substance paracetamol and the costs of packaging.

the two constant expenditures specifications, though they are still underestimated to some extent.

Predicted price effects Finally, the bottom part of Table 6 shows the predicted price effects of a basic merger simulation. This is based on the non-cooperative multi-product pricing oligopoly model of section II.B., where there is only a change in the merging firms' product ownership. Hence, this assumes a pure "loss of competition effect" from the merger between AZT and GSK and no cost changes or other supply side effects. Since in such a simple setting the predicted merger effects only depend on the own-price and cross-price elasticities, this is also a simple way to summarize the combined role of these elasticities. We present the average predicted price increases for each of the three active substances. Recall that the merging firms are only active in the paracetamol segment, and no other firms are active in that segment. Hence, the merging firms' average price increase coincides with the price increase in the paracetamol segment, while the outsiders' price increases correspond with the price increases in the other segments.

The nested logit models predict the following. For the constant expenditures specification, there is a quite substantial price increase of 41.1% in the merging firms' paracetamol segment. This follows from the strong market segmentation by substance ($\sigma_1 > \sigma_2$), which implies low cross-price elasticities between products of the merging firms and their rivals who sell different substances. For the unit demand specification, the predicted price increase is only 15.9%. Compared with the actually observed price increase of 39.7% (obtained earlier in Table 4), the constant expenditure specification gives a quite accurate prediction of the average price increase, while the unit demand specification considerably underestimates the price effects.

The random coefficients logit model generally results in lower predicted price increases, by 21.3% for the constant expenditures specification and by 9.3% for the unit demand specification. This is due to two factors. First, the random coefficient for the paracetamol dummy, while significant, is apparently quantitatively less important than the nesting parameter in the nested logit model. Second, there are other sources of consumer heterogeneity which may raise the extent of substitution to other products with different active substances (e.g. the random coefficient for branded products).

Note that for both the nested logit and the random coefficients logit, the predicted price effects are lower in the unit demand specification than in the constant expenditures specification. The reason is the functional form for the price variable: it enters linearly in the unit demand specification, implying that the price elasticity is quasi-linearly increasing in price, which constrains firms to further raise prices as consumers become more elastic. In the constant expenditures specification, price enters logarithmically, implying a quasi-constant price elasticity so that firms are less constrained to raise prices.

Note that, in all models, the predicted price increases by the competing firms in the other segments are very small, compared with the price increase in the paracetamol segment. Competitors thus respond only weakly to the price increase initiated by the merging firms. This is because of our finding of limited substitution between segments, combined with the fact that there are many competing Bertrand-Nash firms.

To summarize, the various demand models give qualitatively comparable findings, but the magnitudes differ. The unit demand specifications appear to be less appropriate in our application. First, they imply an implausibly large range of price elasticities (especially at the product level, where the differences are up to a factor of 10). Second, they also result in unreasonably low markups, especially for the nested logit. Third, the unit demand specifications considerably underestimate the mergers' price effects.

IV. Evaluating merger simulation

The previous section concluded that the two constant expenditures demand specifications are more appropriate in our application. We therefore now focus on these specifications to examine the merger predictions in more detail.

In subsection IV.A we provide a more detailed analysis on how well merger simulation performs when we only consider a change in the merging firms' product ownership (pure loss of competition effect). In section IV.B we account for other possible supply side factors after the merger (cost and conduct), to assess whether this improves the predictions from merger simulation.

A. Accounting for competition loss only

Table 7 summarizes the detailed predicted merger effects when we only account for a pure loss of competition effect. This assumes there is multiproduct Bertrand-Nash competition and the merger only changes the merging firms' product ownership. The left panel shows the predicted price effects. This is similar to the bottom part of Table 6, except that the information is now broken down by both substance and firm and that we now also present the 10–90% confidence intervals (calculated using a parametric bootstrap). The right panel shows the same information for the predicted market share effects. The predicted price and market share effects in Table 7 can be compared with the actual effects, which we presented earlier in Table 4.

Price effects As we saw before, the nested logit specification predicts the merging firms' average price increase in the paracetamol segment fairly well (41.1% compared with actual effect of 39.7%). This is confirmed when we consider the 10–90% confidence interval for the merger prediction, which is between 34.6% and 49.1%.

However, the individual predicted price increases by firm deviate quite substantially from the actual effects in several respects. First, the predicted price increase of the larger firm

Table 7: Predicted price and market share effects – standard merger simulation

	Price (% change)			Market share (% point change)		
	Mean	10% CI	90% CI	Mean	10% CI	90% CI
			Nested logit			
Paracetamol	41.1	34.6	49.1	-13.3	-14.5	-12.1
AstraZeneca	30.5	25.9	35.7	-5.7	-6.4	-5.0
GSK	73.7	59.8	90.9	-7.6	-8.2	-7.0
Ibuprofen	1.2	1.0	1.3	8.2	7.3	9.0
McNeil	1.5	1.3	1.6	4.9	4.4	5.5
Meda	0.1	0.0	0.1	0.2	0.2	0.3
Nycomed	0.5	0.5	0.6	3.0	2.6	3.3
ASA	1.5	1.3	1.5	5.2	4.8	5.6
McNeil	1.6	1.5	1.7	4.1	3.8	4.4
Meda	0.1	0.1	0.1	0.9	0.8	0.9
Bayer	0.1	0.1	0.1	0.2	0.2	0.2
			Random coefficients logit			
Paracetamol	21.3	9.6	34.5	-7.2	-10.2	-4.2
AstraZeneca	16.3	6.6	27.5	-3.2	-5.1	-1.5
GSK	36.6	18.2	57.1	-4.0	-5.2	-2.8
Ibuprofen	0.8	0.6	1.0	3.9	2.4	5.3
McNeil	0.9	0.7	1.3	2.3	1.3	3.2
Meda	0.1	0.1	0.1	0.1	0.1	0.2
Nycomed	0.5	0.3	0.7	1.5	0.9	2.0
ASA	0.8	0.5	1.1	3.3	1.9	4.9
McNeil	0.9	0.5	1.2	2.7	1.5	4.0
Meda	0.3	0.2	0.4	0.5	0.3	0.7
Bayer	0.0	0.0	0.0	0.1	0.1	0.1

Notes: The predicted price and market share effects are for December 2008, assuming Bertrand-Nash price-setting behavior. The demand models refer to the constant expenditure specifications of Table 5.

AZT (30.5%) is much lower than that of the smaller firm GSK (+68.4%),[12] whereas in reality both firms raised prices by comparable magnitudes (+39.2% versus 40.9%). Intuitively, the significantly lower predicted price increase for AZT than GSK follows from the property in nested logit models that the markups of small firms are lower than those of large firms, and these markups become equalized within a firm after a merger.[13]

Second, the outsider rivals are predicted to raise prices by relatively low amounts, with the largest price increase by the largest firm, McNeil (+1.5% in the ibuprofen segment and +1.6% in the ASA segment). In practice, the price increases were much higher for all firms in the ASA segment: McNeil (+17.9%), Meda (+8.9%) and Bayer (+6.8%).

Now consider the random coefficients logit specification. As seen before, this model underpredicts the merging firms' average price increase in the paracetamol segment. The results in Table 7 show that this underprediction is significant: the 10–90% confidence interval is 9.6%–34.5%, which is below the actual effect of 39.7%. Abstracting from this underprediction, the random coefficients logit confirms both observations for the nested logit. First, the predicted price increase of the larger firm AZT is much lower than that of the smaller firm GSK. While markup equalization within a firm is not a built-in feature of the random coefficients model, the same intuition thus continues to apply: markups tend to be larger for larger firms, and become more equal within a firm after a merger. Second, the outsiders' price responses are again small, and much below the actual price increases of the outsiders.

Market share effects The nested logit specification predicts a market share drop for the paracetamol segment of –13.3% points (confidence interval between –14.5% and –12.1%). The actual market share drop is lower in absolute value (–3.6%). The same conclusion applies for the random coefficients logit, where the predicted market share drop is –7.2%, which is again lower than the actual market share drop (despite the smaller price increase in this specification). Similar findings obtain for the other segments, and for a break-down by substance and firm.

[12] The gap is significant in the sense that the confidence interval for AZT's price increase (between 25.9% and 35.7%) does not overlap with that of GSK's price increase (between 59.8% and 90.9%).

[13] Anderson and de Palma (1992) already demonstrated that markups of the same firm are equalized in a nested logit model with symmetric firms. It can be shown that this property extends to asymmetric firms.

One explanation for the lower than predicted market share drop is a lower price elasticity because of the presence of brand loyalty; see for example Dubé, Hitsch and Rossi (2010). Another explanation is that other factors may have changed after the merger that made paracetamol more attractive. This can for example be a perceived change in quality, perhaps coinciding with the deregulation, although there is no clear indication of this. Similarly, it is possible that the acquiring but smaller firm GSK (who did not see a market share drop at all) benefited from increased visibility or from increased advertising spending after the merger. We will not formally explore these possibilities to explain the market share effects further, and instead focus in more detail on the role of additional supply side factors that may help to explain the price effects.

B. Accounting for other supply side changes

We consider two supply side factors other than the competition loss effect that may contribute to explaining the merger's price effects: the impact of a marginal cost increase due to the package size reduction; and the role of partially coordinated behavior.

Marginal cost increases due to package size reduction An important change that coincided with the merger event in 2009 was the reduction in package size by several brands. As discussed in Section I, in the paracetamol segment the merged firm AZT-GSK reduced the package size of its brands Panodil and Alvedon from 30 to 20 tablets. Moreover, in the ASA segment, McNeil and Meda removed all their large package size (containing 100 tablets). These package size reductions may have led to an increase in marginal costs and hence larger price effects, also for the outsider firms.

To assess this possibility, we used the premerger data to perform a logarithmic regression of the products' marginal costs, as backed out from the oligopoly model using (11), on the product fixed effects and a time trend; in a second stage we then regressed the product fixed effects on the same time-invariant product characteristics as those included in the demand model. The results, presented in the Appendix, show that the elasticity of marginal cost with respect to package size is negative and highly significant, equal to –0.328 in the nested logit and -0.337 in the random coefficients logit specification (with standard errors of 0.058 and 0.047 respectively). This implies that the reduction in package size implied a considerable increase in marginal costs for the concerned firms. For example, Alvedon's package size reduction from 30 to 20 tablets would raise marginal costs by 14.2% in the nested logit model ($= (20/30)^{-0.328} - 1$) and by a similar 14.6% in the random coefficients model.

To assess the role of these marginal cost increases we redid the earlier merger simulations, after accounting for the marginal cost increase for the relevant products. The left part of Table 8 presents the results. This shows that the price predictions become closer to actual prices but differences remain.

First consider the merging firms in the paracetamol segment. The random coefficients logit gives a lower underprediction of their average price increase. The nested logit model also gives a better prediction for one of the merging firms (AZT), but there is a stronger overprediction of the price increase of the other merging firm (GSK). As such, the gap between the very large predicted price increase for GSK and the smaller increase for AZT remains.

Second, for the non-merging rivals in the ASA segment, the marginal cost increases partly explain the observed price increases. For example, in the nested logit model the predicted price increase is 8.1% for Meda's ASA brands, which is close to the actual 8.9% increase. However, the predicted price increase is 6.6% for McNeil's ASA brands, which becomes closer to, but is still an underprediction of the observed 17.9% price increase. For Bayer, the model still does not explain the observed price increase (since Bayer did not reduce package size).

Third, for the outsiders in the ibuprofen segment there are no package size reductions and hence no marginal cost changes. The predicted price changes are correspondingly small, and they are in line with the insignificant actual price effects for the ibuprofen brands found above.

In sum, accounting for the marginal cost increases due to the package size reduction helps to explain the merging firms' average price increase in the random coefficients logit model, and the price increases of some outsider firms in the ASA segment. But the extended model also leaves several price increases unexplained.

Table 8: Predicted price effects - role of supply side

	Cost Increase			Cost Increase + Part. Coord.		
	Mean	10% CI	90% CI	Mean	10% CI	90% CI
				Nested logit		
Paracetamol	57.8	49.8	67.4	48.5	43.1	54.8
AstraZeneca	46.0	40.5	52.5	41.4	37.3	46.0
GSK	93.6	77.4	113.6	70.0	60.0	82.5
Ibuprofen	1.5	1.4	1.7	4.4	4.0	4.9
McNeil	2.0	1.8	2.2	4.8	4.4	5.3
Meda	0.1	0.1	0.1	2.9	2.5	3.3
Nycomed	0.7	0.6	0.8	3.8	3.2	4.2
ASA	6.6	6.3	6.9	7.5	6.7	8.5
McNeil	6.6	6.2	6.9	7.3	6.4	8.3
Meda	8.1	8.0	8.2	9.9	9.4	10.7
Bayer	0.1	0.1	0.1	3.0	2.5	3.4
				Random coefficients logit		
Paracetamol	38.0	23.1	54.6	31.7	19.7	45.6
AstraZeneca	32.9	20.0	47.6	31.0	18.5	45.7
GSK	53.9	32.4	77.8	33.6	22.8	45.7
Ibuprofen	1.7	1.4	2.0	1.8	1.0	3.0
McNeil	2.2	1.7	2.6	2.5	1.3	4.3
Meda	0.0	0.0	0.1	0.0	-0.2	0.3
Nycomed	0.8	0.5	1.2	0.3	0.0	0.7
ASA	5.6	4.8	6.2	7.0	5.7	8.6
McNeil	5.5	4.6	6.2	6.9	5.5	8.7
Meda	7.8	7.6	8.0	8.4	8.2	8.7
Bayer	0.0	0.0	0.0	0.1	-0.2	0.3

Notes: The predicted price effects are for December 2008, under the following scenarios: (i) marginal cost increase due to package size reduction; (ii) marginal cost increase + partial coordination. The demand models refer to the constant expenditure specifications of Table 5.

Partial coordination We now further extend the model to allow firms to partially coordinate on prices, instead of behaving as multi-product Bertrand-Nash competitors. This may also explain some of the differences between the predicted and the observed price effects. First, it may help to explain why large and small merging firms raise their price by more similar amounts than predicted by a non-cooperative pricing model. Second, it may explain why some of the outsiders raise their prices so much, even if they did not experience a marginal cost increase.

To assess the role of partially coordinated behavior, we set the weight that firms put on the profits of the competitors at $\phi = 0.75$, and assume this to be the same before and after the merger. This number is somewhat arbitrary, except that it raises the pre-merger markups to a level that is more in line with the firms' estimates provided during the investigation. Setting $\phi = 0.75$ results in an average markup for paracetamol of 73.9% in the nested logit model and a similar 75.5% in the random coefficients logit model (compared with respectively 48.5% and 59.0% under Bertrand-Nash, as reported earlier in Table 6).

The right part of Table 8 shows the predicted price effects when firms partially coordinate (accounting in addition for the cost increases following the package size reduction). The merging firms' predicted price effects are on average smaller than under Bertrand-Nash, but they remain large and are of a roughly comparable order of magnitude as the actual effects.[14] Now consider the more detailed predicted price effects by firm and substance.

The gap between the predicted price increases of the merging firms becomes smaller. For the nested logit, the predicted price increase of AZT only slightly drops, from 46.0% to 41.4%, whereas the predicted price increase of GSK drops considerably, from 93.6% to 70.0%. Hence, the gap between the price increase of AZT and GSK narrows under partial collusion. Interestingly, for the random coefficients logit the gap is almost eliminated, as the predicted price increases of AZT and GSK are very similar under partial coordination: 31.0%

[14] The intuition for the lower price effects than under Bertrand-Nash is that under partial coordination firms already charge higher prices before the merger (though the effect may in principle be counteracted by the strong rival responses after the merger under partial coordination). The magnitude of the price effects remains large because apparently the price elasticities still form an important constraint when $\phi = 0.75$. To verify this, we also considered simulations with ϕ close to 1 (with implausibly large implied pre-merger markups): this resulted in small predicted price effects, as expected.

and 33.6%, respectively. This is consistent with the fact that the actual price increases of both firms were very similar (though larger at 39.2% and 40.9% respectively).

The predicted price increases of the outsiders' products give a more mixed picture. Generally speaking, because of partial coordination the outsiders respond with higher price increases than in the previous cases. This helps to better predict price increases that were previously underpredicted, mainly McNeil's and Bayer's price changes in the ASA segment (e.g. +7.3% and 3.0% in the nested logit model), though some underprediction remains. But it also implies a slightly stronger overprediction for Meda in the ASA segment (+9.9% in the nested logit) and for some brands in the ibuprofen segment.

In sum, enriching the model to allow for partial coordination gives more reasonable premerger markups and better explains the price increases of the merging firms (since they now raise their prices by more similar amounts). But it does not unambiguously improve the predictions for outsider firms.

Remaining unobserved cost changes To evaluate the performance of the merger simulation approach more systematically, we now ask whether there were remaining unobserved marginal cost changes, after accounting for the competition loss effect, the observed cost increase from the package size reduction and the role of partial collusion. More specifically, we first compute the marginal cost c_j^0 for each product as implied by the model with partial coordination and after filtering out the possible cost increase from the package size reduction.[15] We compute c_{jt}^0 for each product j and each period t during the same two-year comparison window as we used for our ex post price regression (1) in Section I. Parallel to that regression we then perform a logarithmic regression of the unobserved marginal cost of product group i on dummy variables for product groups (α_i) and interactions of these product groups with a postmerger dummy variable (β_i):

$$\ln c_{it}^0 = \alpha_i + \beta_i PostMerger_t + \varepsilon_{it}. \tag{12}$$

[15] Formally, $c_j^0 = c_j \cdot k_j^{-\gamma_k}$, where c_j is the j-th row in (11), k_j is the package size of product j and γ_k is the estimated package size coefficient in the logarithmic cost regression.

The coefficients of these postmerger effects measure the extent to which unobserved marginal costs have changed after the merger. To the extent that these estimated unobserved marginal cost changes are small compared with the initially estimated price effects, we can conclude that the merger simulation model performs well. If the unobserved marginal cost changes remain significant, then there was either a real marginal cost change after the merger, or there were some other unobservable factors not adequately captured in the model.[16]

Table 9 shows the results. For convenience, the left panel again shows the estimates from the logarithmic price regression (1), while the right panel shows the estimates from the analogous cost regressions (12), based on the nested logit and random coefficients logit models.

The top part of Table 9 shows the results from the regressions where product groups are defined at the level of the substance. According to the nested logit model, a significant unobserved marginal cost decrease of –9.1% is required to rationalize the 39.7% price increase of the merging firms in the paracetamol segment, while no significant unobserved cost change is required to rationalize the price increase in the other segments. In contrast, according to the random coefficients logit model, a significant 17.6% marginal cost increase would be required to rationalize the price increase of the merging firms. This is a translation of our earlier finding that the random coefficients logit model tends to underpredict the average price effects of the merging firms.

The bottom part of Table 9 shows the results from the regressions where the product groups are now defined at the level of the substance and firm. To explain the merging firms' price increases in the nested logit model, a larger unobserved cost reduction is required for GSK than for AZT. One interpretation is that GSK, as the acquiring firm, was able to restructure its operations to favour GSK's brand, translating in a lower economic cost of selling this brand. Similarly, to explain the merging firms' price increases in the random coefficients logit model a smaller marginal cost increase is required for GSK than for AZT.

[16] The two-year window to estimate (12) keeps out confounding effects after one year when total pharmacy sales started to decrease due to the deregulation (shift to department stores which are not in our data set). This decrease affects the calculated potential market shares and hence markups and costs. As shown in Appendix, using a longer four-year window implies a negligible change in the estimated price effects, but it does contaminate the cost effects.

Table 9: price and unobserved cost effects - two year window

	Price (% change)		Unobserved cost (% change)			
			Nested logit		RC logit	
	Coeff	St Err	Coeff	St Err	Coeff	St Err
Regressions at the level of the substance						
subst. f.e.	yes		yes		yes	
Paracetamol×merger	39.7	1.0	-9.1	2.0	17.6	5.4
Ibuprofen×merger	0.1	1.7	-2.9	1.2	5.0	3.7
ASA×merger	13.3	4.7	3.7	3.6	10.0	5.6
Regressions at the level of the firm × substance						
firm×subst f.e.	yes		yes		yes	
Paracetamol						
AZT×merger	39.2	1.2	-6.2	2.3	20.0	5.1
GSK×merger	40.9	2.1	-16.7	2.8	9.2	2.3
Ibuprofen						
McNeil×merger	0.4	1.1	-2.8	1.8	7.8	3.1
Meda×merger	0.0	0.1	-3.8	1.0	-1.7	1.2
Nycomed×merger	0.9	1.1	-3.0	1.5	-0.5	1.7
ASA						
McNeil×merger	17.9	4.8	5.6	4.4	10.4	7.1
Meda×merger	8.9	5.4	-3.6	5.2	0.4	4.5
Bayer×merger	6.8	1.0	-1.4	1.3	6.0	1.3

Notes: This table compares the estimated price effects (from first column of Table 4) with the estimated unobserved marginal cost effects, based on the regression model (12). The percentage effects are obtained from a transformation of the parameters using $\exp(\beta_i) - 1$ (and a corresponding adjustment of the standard errors using the delta method). Robust standard errors are reported.

This again suggests that the acquiring firm GSK became comparatively more efficient than AZT.

According to both demand models, either insignificant or low cost changes are required to explain most outsiders' price changes (including Meda's large price increase of 8.9% in the ASA segment). Only some significant unobserved marginal cost increases seem to be required to explain the price increases of McNeil and Bayer (mainly in the random coefficients logit). A possible interpretation is that these firms responded more cooperatively after the merger. Although there is no clear indication of such behaviour from outside sources, it could be modelled by estimating a post-merger partial coordination parameter different from the pre-merger parameter, similar to Miller and Weinberg (2015). An alternative interpretation of the price increases by McNeil and Bayer in the ASA segment is that these firms produce closer substitutes to the merging firms than implied by our demand estimates. This possibility could be explored by further extending the range of demand models we considered.

In sum, these findings suggest that the augmented supply side performs reasonably well: the nested logit model somewhat overpredicts the merging firms' price increases (as reflected in an unobserved marginal cost decrease), while the random coefficients model somewhat underpredicts their price increases (as reflected in an unobserved cost increase). Both models suggest that the smaller of the two merging firms (GSK) obtained a relative cost advantage to the target firm (AZT), and that some outsiders firms in the ASA segment may have responded more cooperatively.

V. Conclusions

We have made use of a unique merger case to evaluate the usefulness of merger simulation as a structural approach to predict the effects from mergers. The merger case is unique because it involved large players who have no other competition in their own segment. The merger led to large price increases: 40% for the merging firms and more than 10% for several outsiders. This enables us to assess a large range of predictions from merger simulation at the level of the individual firms.

The merger simulation model started from two demand models (nested and random coefficients logit), where we proposed a constant expenditures specification as a possible alternative to the typical unit demand specification. Our demand estimates show the following two key points. First, market segmentation according to active substance is a very important differentiation dimension. This implies that the two merging firms may form a strong competitive constraint on prices before the merger. Second, the constant expenditures specification entails a more plausible pattern of price elasticities across products.

Based on these two demand side findings, the model predicts relatively large price increases by the merging firms. However, the basic supply side model results in several discrepancies between the predicted and observed price effects. First, both firms raised their prices by a similar percentage, while the model predicts a much larger price increase for the smaller firm. Second, some of the outsider firms also raised price by a fairly large amount after the merger, while the model predicts only small responses by the outsiders. We consider several supply side explanations that may explain these findings: a plausible marginal cost increase after the merger because several firms reduced their package size, and the possibility of partial collusion. Both factors help to explain the observed price increases, but we document that some unobserved factors remain. We modelled these unobserved factors as unobserved cost changes in a simple linear cost regression, but one can also interpret these as changes in tacit coordination (by estimating a suitable post-merger coordination parameter in a nonlinear supply framework).

Our analysis was based on a broad range of demand models and simple modifications of a standard multiproduct oligopoly model, without an "elaborate superstructure" to which Angrist and Pischke refer in their discussion. In future research it may nevertheless be interesting to consider various extensions of this model to see whether these can improve the accuracy of the predictions. In particular, more elaborate supply models may be considered with a dynamic component. For example, under brand loyalty due to switching costs (as in e.g. Dubé, Hitsch and Rossi, 2009) price competition and hence the impact of mergers may be different. Furthermore, mergers may affect the extent of tacit coordination, though empirically analyzing this is inherently challenging because of the multiplicity of equilibria in repeated games. Nevertheless, even without these extensions, it would be interesting to see a lot more work that confronts the merger simulations with the actual merger effects.

References

Angrist, Joshua and Steffen Pischke, 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, Orley and Daniel Hosken, 2010, The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies, Journal of Law and Economics, 53(3), 417-466.

Ashenfelter, Orley., Daniel Hosken, and Mathew Weinberg, 2013, The Price Effects of a Large Merger of Manufacturers: a Case Study of Maytag-Whirlpool, American Economic Journal: Economic Policy, 5(1): 239–261.

Ben-Akiva, Moshe and Steven Lerman, 1985, Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, 412pp.

Berry, Steven, 1994, Estimating Discrete Choice Models of Product Differentiation, RAND Journal of Economics, 25, 242-262.

Berry, Steven, James Levinsohn, and Ariel Pakes, 1995, Automobile Prices in Market Equilibrium, Econometrica, 63, 841-890.

Björnerstedt, Jonas and Frank Verboven, 2012, Does Merger Simulation Work? A "Natural Experiment" in the Swedish Analgesics Market, CEPR Discussion paper

Bresnahan, Timothy, Scott Stern, and Manuel Trajtenberg, 1997, Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the late 1980's, RAND Journal of Economics, Special Issue in Honor of Richard Quandt, 1997, 28 (0), pp. S17-S44.

Borenstein, Severin, 1990, Airline Mergers, Airport Dominance, and Market Power, American Economic Review, 80, 400-404.

Chevalier, Judy, Anil Kashyap and Peter Rossi, 2003, Why Don't Prices Rise during Periods of Peak Demand: evidence from Scanner Data, American Economic Review, 93, 15-37.

Chintagunta, Pradeep, 2002, Investigating Category Pricing Behavior at a Retail Chain, Journal of Marketing Research, 39, 141-154.

Crooke, Philip, Luke Froeb, Steven Tschantz and Gregory Werden, The Effects of Assumed Demand Form on Simulated Post-Merger Price Equilibria, Review of Industrial Organization, 15(3), 205-217.

Dubé, Jean-Pierre, 2004, Multiple Discreteness and Product Differentiation: Demand for Carbonated Soft Drinks, Marketing Science, 23(1), 66-81.

Dubé, Jean-Pierre, Gunther Hitch and Peter Rossi, 2009, Do Switching Costs Make Markets Less Competitive?, Journal of Marketing Research, 46, 435-445.

Dubé, Jean-Pierre, Gunther Hitch and Peter Rossi, 2010, State Dependence and Alternative Explanations for Consumer Inertia, RAND Journal of Economics, 41 (3), 417–445.

Epstein, Roy and Daniel Rubinfeld, 2001, Merger Simulation: a Simplified Approach with New Applications, Antitrust Law Journal, 69, 883.

Facarelli, Dario and Fabio Panetta, 2003, Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits, American Economic Review, 93, 1152-1172.

Friberg, Richard and André Romahn, 2015, Ex Post Merger Review and Divestitures, forthcoming International Journal of Industrial Organization.

Goldberg, Pinelopi, 1995, Product Differentiation and Oligopoly in International Market, Econometrica, 891-951.

Gowrisankaran, Gautam and Mark Rysman, 2012, Dynamics of Consumer Demand for New Durable Goods, Journal of Political Economy, 120, 1173-1219.

Grigolon, Laura and Frank Verboven, 2014, Nested Logit or Random Coefficients Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation, Review of Economics and Statistics, 96 (5), 916-935.

Hanemann, W. Michael, 1984, Discrete/Continuous Models of Consumer Demand, Econometrica, 52 (3), 541-562.

Hastings, Justin, 2004, Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California, 94(1), 317-328.

Hastings, Justin and Richard Gilbert, 2005, Market Power, Vertical Integration and the Wholesale Price of Gasoline, Journal of Industrial Economics, 53 (4), 469-492.

Hausman, Jerry, Gregory Leonard and Douglas Zona, 1994, Competitive Analysis with Differentiated Products, Annales d'Economie et de Statistique, 34(1), 159-180.

Heiss, Florian and Viktor Winschel, 2008, Likelihood Approximation by Numerical Integration on Sparse Grids, Journal of Econometrics, 144(1), 62-80.

Hosken, Daniel, Louis Silvia, and Christopher Taylor, 2011, Does Concentration Matter? Measurement of Petroleum Merger Price Effects, American Economic Review: Papers & Proceedings, 101(3), 45-50.

Huang, Dongling, Christian Rojas and Frank Bass, 2008, What Happens when Demand is Estimated with a Misspecified Model? The Journal of Industrial Economics, 56 (4), 809–839.

Ivaldi, Marc and Frank Verboven, 2005, Quantifying the Effects of Horizontal Mergers in European Competition Policy, International Journal of Industrial Organization, 23 (9-10), 669-691.

Kim, E. Han and Vijay Singal, 1993, Mergers and Market Power: Evidence from the Airline Industry, American Economic Review, 83, 549-569.

Knittel, Christopher and Kostas Metaxoglou, 2014, Estimation of Random Coefficient Demand Models: Two Empiricists' Perspective, Review of Economics and Statistics, 96(1), 2014, 34-59.

Miller, Nathan and Mathew Weinberg, 2015, Mergers Facilitate Tacit Collusion: Empirical Evidence from the U.S. Brewing Industry, working paper.

Nevo, Aviv, 2000, A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand, Journal of Economics and Management Strategy, 9(4), 513-548.

Nevo, Aviv, 2000, Mergers with Differentiated Products: the Case of the Ready-to-Eat Cereal Industry, RAND Journal of Economics, 31(3), 395-421.

Nevo, Aviv and Michael Whinston, 2011, Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference, Journal of Economic Perspectives, 24(2), 69-82.

Peters, Craig, 2006, Evaluating the Performance of Merger Simulations: Evidence from the U.S. Airline Industry, Journal of Law and Economics, 49, 627-649.

Reynaert, Mathias and Frank Verboven, 2014, Improving the Performance of Random Coefficients Demand Models: the Role of Optimal Instruments, Journal of Econometrics, 179 (1), 83-98.

Slade, Margareth, 2009, Merger Simulations of Unilateral Effects: What can we Learn from the UK Brewing Industry? in: B. Lyons, Cases in European Competition Policy: the Economic Analysis, Cambridge University Press.

Verboven, Frank, 1996, International Price Discrimination in the European car market, RAND Journal of Economics, 27, 240-268.

Werden, Gregory and Luke Froeb, 1994, The Effects of Mergers in Differentiated Products Industries: Logit Demand and Market Policy, 10(2), 407-26.

Weinberg, Mathew, 2011, More Evidence on the Performance of Merger Simulations, American Economic review: Papers & Proceedings, 101 (3): 51-55.

Weinberg, Mathew and Daniel Hosken, 2013, Evidence on the Accuracy of Merger Simulations, Review of Economics and Statistics, 95(5).

Appendix A: Demand model

Extension to random coefficients logit

We start from the following generalization of consumer i's conditional indirect utility (2) of good j:

$$u_{ij} = x_j \beta_i + \xi_j + \alpha_i f(y_i, p_j) + \varepsilon_{ij}, \tag{A1}$$

where β_i and α_i are now individual-specific valuations of the product characteristics, modelled as random coefficients. Following BLP, Nevo (2000) and others, specify the random coefficients β_i and α_i as

$$\begin{pmatrix} \beta_i \\ \alpha_i \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix} + \Sigma \nu_i$$

where β and α are means and Σ is a diagonal matrix with standard deviations of the random coefficients, and ν_i is a vector of standard normal random variables.

For the two conditional demand specifications (3), we can again write utility (2) more compactly:

$$u_{ij} = K_i + \delta_j + \mu_j(\nu_i) + \varepsilon_{ij}, \tag{A2}$$

where δ_j is the mean valuation for product j as before, and $\mu_j(\nu_i)$ is an individual-specific valuation for product j, with $\mu_j(\nu_i) = \begin{pmatrix} x_j & p_j \end{pmatrix} \Sigma \nu_i$ in the unit demand specification and $\mu_j(\nu_i) = \begin{pmatrix} x_j & \ln p_j \end{pmatrix} \Sigma \nu_i$ in the constant expenditures specification. The unit demand and constant expenditures specification essentially differ in the fact that price enters linearly or logarithmically in both δ_j and $\mu_j(\nu_i)$.

Given random utility maximization and an extreme value (logit) distribution for ε_{ij}, the conditional probability that consumer i chooses product j is:

$$\pi_j(\boldsymbol{\delta}, \sigma, \nu_i) \equiv \frac{\exp\left(\delta_j + \mu_j(\nu_i)\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_k + \mu_k(\nu_i)\right)},$$

where σ is the vector of standard deviations in the diagonal matrix Σ.

Aggregate demand for product j is the probability that a consumer buys product j multiplied by the quantity purchased, $d_j(y_i)$, aggregated over all L consumers according to income distribution P_y and the distribution of taste parameters P_ν, assumed to be independent of income

$$\begin{aligned} q_j &= \int \pi_j(\boldsymbol{\delta}, \sigma, \nu)\, d_j(y)\, \mathrm{d}P_\nu(\nu)\, \mathrm{d}P_y(y)\, L \\ &= \int \pi_j(\boldsymbol{\delta}, \sigma, \nu)\, \mathrm{d}P_\nu(\nu) \int d_j(y)\, \mathrm{d}P_y(y)\, L \\ &= s_j(\boldsymbol{\delta}, \sigma) \int d_j(y)\, \mathrm{d}P_y(y)\, L. \end{aligned} \tag{A3}$$

The second equality follows from the fact that the choice probability $\pi_j(\boldsymbol{\delta}, \sigma, \nu_i)$ is independent of income. The third equality substitutes the usual unconditional choice probability of BLP's aggregate random coefficients model:

$$s_j(\boldsymbol{\delta}, \sigma) \equiv \int \frac{\exp\left(\delta_j + \mu_j(\nu)\right)}{1 + \sum_{k=1}^{J} \exp\left(\delta_k + \mu_k(\nu)\right)} \mathrm{d}P_\nu(\nu). \tag{A4}$$

Similar to the nested logit model, the integral in (A3) is simply $\int d_j(y) \mathrm{d}P_y(y)\, L = L$ in the unit demand specification, and $\int d_j(y) \mathrm{d}P_y(y)\, L = \gamma Y / p_j$ in the constant expenditures specification. This results in the same expressions for the choice probabilities in terms of observables derived in the text (6), where $s_j(\boldsymbol{\delta}, \sigma)$ is now given by the market share integral (A4).

This shows that the constant expenditures specification is a straightforward variant of BLP's unit demand specification, where the unconditional choice probability should be set equal to the market share in value terms instead of volume terms, and price enters logarithmically instead of linearly. Estimation is otherwise similar as in BLP, i.e. the market share system can be solved numerically for the mean utility δ_j using BLP's contraction mapping and simulated GMM can be applied.

Details on the nested logit model

In the text, we showed that

$$
\begin{array}{ll}
\text{Unit demand} & \frac{q_j}{L} = s_j\left(\boldsymbol{\delta}, \sigma\right) \\
\text{Constant expenditures} & \frac{p_j q_j}{B} = s_j\left(\boldsymbol{\delta}, \sigma\right)
\end{array}
\tag{A5}
$$

where $s_j\left(\boldsymbol{\delta}, \sigma\right)$ is given by (5). We now provide several details not provided in the text.

1. The inclusive values or "log sum" formulas I_{hg}, I_g, and I are defined by:

$$
\begin{aligned}
I_{hg} &\equiv (1-\sigma_1)\ln\sum_{k=1}^{J_{hg}} \exp((\delta_k)/(1-\sigma_1)) && \text{(A6)} \\
I_g &\equiv (1-\sigma_2)\ln\sum_{h=1}^{H_g} \exp(I_{hg}/(1-\sigma_2)) \\
I &\equiv \ln\left(1+\sum_{g=1}^{G}\exp(I_g)\right).
\end{aligned}
$$

2. The estimating equation (7) can be written out as follows for the unit demand specification:

$$
\ln\frac{q_j}{L-\sum_{j=1}^{J} q_j} = x_j\beta - \alpha p_j + \sigma_1 \ln\frac{q_j}{\sum_{j\in H_{hg}} q_j} + \sigma_2 \ln\frac{\sum_{j\in H_{hg}} q_j}{\sum_{h=1}^{H_{hg}}\sum_{j\in H_{hg}} q_j} + \xi_j \tag{A7}
$$

and for the constant expenditures model

$$
\ln\frac{p_j q_j}{B-\sum_{j=1}^{J} q_j p_j} = x_j\beta - \alpha \ln p_j + \sigma_1 \ln\frac{p_j q_j}{\sum_{j\in H_{hg}} p_j q_j} + \sigma_2 \ln\frac{\sum_{j\in H_{hg}} p_j q_j}{\sum_{h=1}^{H_{hg}}\sum_{j\in H_{hg}} p_j q_j} + \xi_j. \tag{A8}
$$

Note that the unit demand specification can immediately be interpreted as an inverse demand system (by writing price on the left hand side). This is not the case for the constant expenditures specification.

3. The price elasticities can be computed as follows. First, the derivatives of the choice probability $s_j(\boldsymbol{\delta}, \sigma)$, as given by (5), with respect to the mean utility δ_k can be shown to be

$$\frac{\partial s_k}{\partial \delta_j} = s_k \left(\frac{1}{1-\sigma_1} D^1_{jk} - \left(\frac{1}{1-\sigma_1} - \frac{1}{1-\sigma_2} \right) s_{j|hg} D^2_{jk} - \frac{\sigma_2}{1-\sigma_2} s_{j|g} D^3_{jk} - s_j \right) \tag{A9}$$

where we define three dummy variables: $D^1_{jk} = 1$ if $j = k$, $D^2_{jk} = 1$ if j and k are in same subgroup, $D^3_{jk} = 1$ if j and k are in same group (and zero otherwise). Second, using (A5), the aggregate demand derivatives are

$$\begin{array}{lcl} \text{Unit demand} & \frac{\partial q_k}{\partial p_j} = & -\alpha \frac{\partial s_k}{\partial \delta_j} L \\ \text{Constant expenditures} & \frac{\partial q_k}{\partial p_j} = & -\alpha \frac{\partial s_k}{\partial \delta_j} \frac{B}{p_j p_k} - s_j \frac{B}{p_j^2} D^1_{jk}, \end{array} \tag{A10}$$

Substituting (A9) into (A10), one can obtain the following expressions for the aggregate price elasticities. In the unit demand specification, we have

$$\frac{\partial q_k}{\partial p_j} \frac{p_j}{q_k} = -\alpha \left(\frac{1}{1-\sigma_1} D^1_{jk} - \left(\frac{1}{1-\sigma_1} - \frac{1}{1-\sigma_2} \right) s_{j|hg} D^2_{jk} - \frac{\sigma_2}{1-\sigma_2} s_{j|g} D^3_{jk} - s_j \right) p_j,$$

while in the constant expenditures specification we have

$$\frac{\partial q_k}{\partial p_j} \frac{p_j}{q_k} = -\alpha \left(\frac{1}{1-\sigma_1} D^1_{jk} - \left(\frac{1}{1-\sigma_1} - \frac{1}{1-\sigma_2} \right) s_{j|hg} D^2_{jk} - \frac{\sigma_2}{1-\sigma_2} s_{j|g} D^3_{jk} - s_j \right) - D^1_{jk}.$$

This shows that the price elasticities are increasing quasi-linearly in prices across products in the typical unit demand specification, whereas they are quasi-independent of prices in the constant expenditures demand specification. In both cases, we write "quasi", since there is indirect dependence on the prices through the market shares.

Appendix B: Further empirical results

First stage demand estimates

Table B1: Definition and summary statistics of instruments

Variable name	Definition	Mean	St Dev	Min	Max
# prod	# of products per market	43.35	3.34	37	52
# prod (group)	# of products per market/group	23.07	5.83	15	34
# prod (firm)	# of products per market/firm	8.54	4.26	1	15
# prod (firm group)	# of products per market/firm/group	4.74	1.99	1	8
# prod (subgroup)	# of products per market/subgroup	10.78	4.13	1	17
# prod (firm subgroup)	# of products per market/firm/subgroup	3.95	1.78	1	7

Notes: 7240 observations for 1995-2008.

Table B2: Nested logit first-stage estimates

	Const Exp			Unit demand		
	$\log(p_j)$	$\log(s_{j\|hg})$	$\log(s_{h\|g})$	p_j	$\log(s_{j\|hg})$	$\log(s_{h\|g})$
marketing1	0.923	75.830	2.498	0.907	73.956	4.388
	(0.926)	(9.372)	(1.766)	(0.392)	(9.216)	(2.023)
# products	-0.008	-0.036	-0.014	-0.004	-0.034	-0.010
	(0.001)	(0.008)	(0.003)	(0.000)	(0.008)	(0.003)
# products (group)	0.024	0.069	-0.049	0.009	0.058	-0.051
	(0.001)	(0.012)	(0.003)	(0.000)	(0.012)	(0.003)
# products (firm)	-0.001	-0.152	0.016	-0.003	-0.147	0.012
	(0.002)	(0.022)	(0.003)	(0.001)	(0.022)	(0.004)
# products (firm group)	-0.046	0.234	-0.014	-0.012	0.276	-0.010
	(0.003)	(0.033)	(0.005)	(0.001)	(0.033)	(0.006)
# products (subgroup)	-0.036	-0.149	0.186	-0.006	-0.115	0.180
	(0.001)	(0.017)	(0.002)	(0.000)	(0.017)	(0.002)
# products (firm subgroup)	0.071	0.001	-0.020	0.017	-0.049	-0.034
	(0.004)	(0.052)	(0.006)	(0.001)	(0.051)	(0.007)
F-test excluded instruments	227.650	39.290	2459.730	135.340	35.400	2099.390

Notes: 7240 observations for 1995-2008. First-stage regressions also include the variables sickwomen and sickmen, monthly fixed effects and 56 product fixed effects. Robust standard errors are reported. Market shares are defined in values for the constant expenditure specification and in units for the unit demand specification, as discussed in the text.

Cost parameter estimates

Table B3: Cost parameter estimates

	Nested logit		Random coefficient logit	
	Const Exp	Unit	Const Exp	Unit
constant	-3.827	-4.562	-3.595	-5.234
	(0.885)	(0.821)	(0.718)	(1.623)
log(dosage)	0.323	0.363	0.281	0.239
	(0.139)	(0.129)	(0.113)	(0.254)
log(package size)	-0.328	-0.429	-0.337	-0.674
	(0.058)	(0.054)	(0.047)	(0.107)
fizzy	0.273	0.426	0.356	0.643
	(0.068)	(0.063)	(0.055)	(0.125)
paracetamol	0.019	0.093	-0.004	-0.064
	(0.091)	(0.084)	(0.074)	(0.167)
ibuprofen	0.632	0.751	0.557	0.768
	(0.120)	(0.111)	(0.097)	(0.220)

Notes: Based on second stage regression of 56 product fixed effects. The first stage (not reported) regresses marginal cost (under Bertrand-Nash) on monthly fixed effects and 56 product fixed effects (7240 observations for 1995-2008). Robust standard errors are reported.

Price and unobserved cost effects: alternative comparison period

Table B4: Price and unobserved cost effects - four year window

	Price (% change)		Unobserved cost (% change)			
			Nested logit		RC logit	
	Coeff	St Err	Coeff	St Err	Coeff	St Err
Regressions at the level of the substance						
subst. f.e.	yes		yes		yes	
Paracetamol×merger	42.1	(1.0)	1.8	(2.1)	26.2	(1.7)
Ibuprofen×merger	0.1	(1.2)	2.2	(1.1)	3.6	(1.4)
ASA×merger	14.7	(3.3)	11.4	(2.8)	11.3	(3.6)
Regressions at the level of the firm × substance						
firm×subst f.e.	yes		yes		yes	
Paracetamol						
AZT×merger	42.9	(0.8)	3.9	(2.6)	26.4	(2.3)
GSK×merger	45.8	(1.1)	-8.0	(2.5)	26.8	(2.2)
Ibuprofen						
McNeil×merger	2.5	(0.9)	3.9	(1.7)	5.8	(2.0)
Meda×merger	0.0	(0.1)	-2.6	(1.1)	-0.7	(1.7)
Nycomed×merger	1.5	(0.9)	0.5	(1.0)	2.7	(1.7)
ASA						
McNeil×merger	21.4	(2.7)	14.5	(3.5)	17.2	(3.9)
Meda×merger	8.5	(3.6)	-1.0	(3.8)	2.4	(4.1)
Bayer×merger	10.7	(0.9)	5.8	(1.5)	14.3	(1.8)

Notes: This table compares the price and unobserved marginal cost effects, based on (12), using a four-year instead of two-year comparison window. This contains confounding deregulation effects because of sales shift to department stores. The percentage effects are obtained from a transformation of the parameters using $\exp(\beta_i)-1$ (and a corresponding adjustment of the standard errors using the delta method). Robust standard errors are reported.