Divestiture Requirements as a Tool for Competition Policy: A Case from the Swedish Beer Market*

Richard Friberg[†], André Romahn[‡]

April 27, 2015

Abstract

We investigate the effect of divestitures on prices and welfare following the Carlsberg-Pripps merger in the Swedish beer market. Both difference-in-difference estimation and simulations using a random coefficients logit model suggest that divestitures are important for dampening price increases. Prices of divested brands fall by around 3 percent and the predicted price increase for Carlsberg falls from 3 to 1.6 percent as a result of the divestitures. To guide practice on divestitures, we investigate the role of the recipient and the number and characteristics of the divested products by simulating post-merger outcomes for all relevant cases. We find that in this setting with large multiproduct firms, the competition authority's most effective means to dampen adverse post-merger outcomes is to aim for a small recipient firm and attain a large number of divested products. Enforcing larger divestitures in terms of market share and raising the average cross-price elasticity between the merging parties' divested and retained products strengthens the dampening effect further.

JEL Classification: K21, L11, L41, L66

Keywords: divestitures, merger simulation, ex-post merger review

1 Introduction

Competition authorities frequently require that merging parties divest a number of brands or operations in order to clear a proposed merger. As we document below, little is known about the impact of such divestitures despite their prominent role in practice. In this paper we use the 2001 takeover of the Swedish Pripps brewery by the Danish brewer Carlsberg to examine the effect of the divestitures on merger outcomes. Our research question is twofold: Firstly, we want to provide a clean case study of the effect of divestitures on prices and welfare in this merger. Secondly, we want to systematically examine how the

*We are grateful to the Swedish Competition Authority, Handelsbankens Forskningsstiftelser and the Jacob Wallenberg foundation for financial support. Thanks to Anna Wikström (Systembolaget) and Lars Näsström (Research International) for providing the data. Valuable comments have been provided by the editor (Julie Holland Mortimer), three anonymous referees, Morten Olsen, Tobias Klein, Philip Schmidt-Dengler, Frode Steen and Michael Waterson as well as seminar audiences at Paris School of Economics, the Swedish Competition Authority, IESE Business School, Uppsala University (national conference), and conference participants at the BECCLE, CRESSE, IIOC and MACCI.

†Corresponding author, Stockholm School of Economics, Sweden. E-mail: nerf@hhs.se

‡KU Leuven, University of Leuven, Belgium; E-Mail: andre.romahn@kuleuven.be

welfare effects of divestitures depend on the acquirer, the number of divested products, their market share and their average cross-price elasticities with the retained products. Our findings are partly specific to this particular merger but we believe that the results and the heuristic algorithm used will be of use for competition authorities in charge of proposing divestitures in other markets and merger cases.

A number of features make this merger in the Swedish beer market an interesting case for examining divestitures. The merging parties account for a substantial share of total sales in the Swedish beer market: Carlsberg's and Pripps's pre-merger market shares by volume were 29 and 17 percent, respectively. The divestitures are substantial: at the time of the merger they account for 6 percent of volume. We have access to barcode level data on prices and quantities, aggregated by month, for the whole market, for a period from January 1996 to January 2003, thus covering two years after the merger. Knowledge of the retailer's (exogenous) markup rule allow us to back out wholesale prices precisely.

We first examine the effects of divestitures on prices in a simple model and highlight that the prices of divested products should fall, *ceteris paribus*. Examining the merger with difference-in-difference methods we indeed find that the prices of divested products fall by about 3 percent. To be able to examine various counterfactual policies we follow the seminal work of Berry, Levinsohn and Pakes (1995, hereafter BLP) and estimate a random coefficients logit model of demand. We find that the effects of the divestitures are sizable: the divestitures lower the predicted price increase from 3 to 1.6 percent for Carlsberg and from 6.5 to 4.9 percent for Pripps. For the average market price increase, the divestitures lower the predicted price hike by two thirds, from 1.6 to 0.5 percent.

Both the difference-in-difference and the BLP demand system are useful for evaluating the merger. The difference-in-difference estimates describe the ex-post developments of prices under transparent assumptions even if, as we discuss below, the merger between two major firms in a national market should lead one to view the control group with some caution. The BLP demand system is useful for ex-ante simulations and for exploring alternative divestitures. We use our structural demand model to pin down which attributes of the divestiture requirements matter most for keeping price rises and consumer welfare losses in check. We find that the recipient firm of the divested beers has a large impact on post-merger outcomes. In the actual merger this firm has a negligible market share and can therefore barely use the divestitures to generate market power for its existing portfolio of brands. Varying the recipient of the divested beers is associated with average price increases ranging from 1.1 percent to 0.5 percent. We also examine the composition of the set of divestitures using a method that can approximate the distribution of all potential post-merger outcomes for a given range of market shares for the divested products. We find that in this market with large multiproduct firms the number of divested beers is an important explanation for low price increases and limited welfare losses post-merger. Raising the number of divestitures by 1 percent reduces the average post-merger price increase by almost 3 percent and reduces the loss in consumer welfare by almost 2.7 percent. The market share of the divestitures and the substitutability between the divestitures and the merging parties' products also have statistically significant and economically important effects on post-merger price increases and consumer welfare losses. Raising both of these parameters dampens adverse outcomes.

Using both a difference-in-difference and a BLP demand system also allow us to compare results. Important qualitative patterns are common to the two methods: prices of divested products fall, prices of legacy products for the acquirer of the divested products rise and despite the merger of the two largest firms in a concentrated market there are small price effects. Nevertheless, as in the previous literature, we find some discrepancies between the ex-post difference-in-difference predictions and the ex-ante structural simulation of the merger when we keep marginal costs at their pre-merger level. Efficiency gains are likely to be one contributor to the discrepancies but lacking separate evidence on product level marginal costs surrounding the merger we cannot rule out other explanations - it may be that the difference-in-difference estimates do not accurately capture the ceteris paribus impact of the merger or there may be concerns with the estimated demand system or firms may not be playing static oligopoly or fail to optimize as we discuss below. Our discussion should prove relevant for future ex-post evaluations of merger simulations.[1]

Let us briefly review the previous literature on divestitures. In both the European Union and the United States a majority of mergers that are subjected to closer scrutiny are cleared subject to remedies in the form of requirements regarding structure (such as divestments) or behavior (such as length of contracts). For mergers that merit closer attention, prohibiting them or permitting as proposed are the exception - clearing subject to remedies is the rule.[2] In many jurisdictions, divestitures are the most prominent form of remedy. Indeed, the European Commission's notice on remedies states that "a general distinction can be made between divestitures, other structural remedies, such as granting access to key infrastructure or inputs on non-discriminatory terms, and commitments relating to the future behavior of the merged entity. *Divestiture commitments are the best way to eliminate competition concerns resulting from horizontal overlaps.*" (European Commission (2008), paragraph 17, emphasis added).

Despite their importance in merger practice, the literature examining the role of divestitures in mergers is scant. Some, largely qualitative, descriptions of divestiture practice can be found in for instance Elzinga (1962), Federal Trade Commission (1999) and DG Competition (2005). The latter two studies establish that in most cases the divested assets are still in operation a few years after the divestiture, and in this sense divestiture policy has been successful. These studies are silent on price reactions surrounding the mergers, however, and there are only a few detailed case studies of the price effects of divestitures. In one such study Tenn and Yun (2011) provide a before-after analysis of the merger between J&J and Pfizer and show that prices of divested brands fell post-divestiture. Pham and Prentice (2013) examine a merger in the Australian cigarette industry that involved divestitures and compare results to counterfactual simulations based on a random coefficients logit model of demand. Data limitations force them to estimate demand for a period several years before the actual merger, but their results nevertheless suggest that divestitures reduced price increases. [3] Apart from the ability to follow a large merger on a market with good data we are also attracted by the beer market having been a prominent testing ground for merger simulations right from the beginning of this literature; Baker and Bresnahan (1985) and Hausman et al. (1994) use simulations to examine prospective mergers in the US beer market, and Pinkse and Slade (2004) apply them to mergers in the

[1] See Peters (2006), Weinberg (2011), Weinberg and Hosken (2013) or Björnerstedt and Verboven (2013) for examples of this, as of yet, small literature.

[2] For instance, among the proposed mergers that were subject to the Phase II procedure by the European Competition Authority between 1990 and 2011, 56 precent were cleared subject to remedies. In comparison, only 13 percent of the proposed mergers were prohibited at this stage and 28 percent were permitted as proposed. Similarly, of 144 mergers challenged by US competition authorities between 2003 and 2007, 64 percent were cleared after remedies had been agreed upon (Tenn and Yun (2011)).

[3] The theoretical work on divestitures is similarly limited. Compte, Jenny and Rey (2002) show that divestitures may facilitate collusion if they lead to a more symmetric industry structure. Cosnita and Tropeano (2009) examine how a competition authority can use policies regarding divestitures to induce the merging parties to reveal private information on the efficiency gains of the merger. Vasconcelos (2010) uses a stylized setting with four ex-ante symmetric Cournot competitors to show that divestitures can increase consumer surplus by creating a more efficient competitor.

UK beer market. Neither of these papers examine divestitures. [4] Ashenfelter et al. (2013) examine the role of efficiencies in the US merger between Coors and Miller and find that a predicted price increase of some two percent was largely offset by declines in marginal costs. Their paper is complementary to ours as both show how substantial concentrations in the beer industry fail to lead to the price hikes that one may ex ante have expected. Efficiencies due to more efficient transport is an important facet in their study while the evidence in our case point to an important role for divestitures.
This paper is structured as follows. Section 2 provides a stylized theoretical example that illustrates the potential for downward pricing pressure when divestiture requirements are imposed on a merger. The following section describes the institutional setting, the data and the merger. Section 4 describes price developments surrounding the merger using the difference-in-difference methodology. Section 5 details our structural model and the estimation results. In Section 6 we use the structural model to examine the impact of divestitures and efficiencies on the merger in question. Section 7 contains our systematic exploration of counterfactual divestitures and in Section 8 we provide concluding remarks.

2 Stylized Example

To provide intuition for our findings, we adapt the concept of upward pricing pressure (UPP) as developed by Farrell and Shapiro (2010) to explore the role of divestitures in price setting. To be concise, we present a stylized example and for the purposes of exposition abstract from efficiency gains. Suppose there are only two firms, where firm MP (*multi-product*) owns products 1 and 2, while firm SP (*single-product*) owns product 3. Let p_j, mc_j and s_j denote price, marginal cost and market share for product j, respectively. Then, the first-order conditions of the firms' maximization problems are as follows.[5]

$$
\begin{aligned}
p_1 &= mc_1 - \left(\frac{\partial s_1}{\partial p_1}\right)^{-1}\left(s_1 + \frac{\partial s_2}{\partial p_1}(p_2 - mc_2)\right) \\
p_3 &= mc_3 - \left(\frac{\partial s_3}{\partial p_3}\right)^{-1} s_3
\end{aligned}
$$

With product prices being strategic complements, the cross-price effect $\partial s_2/\partial p_1$ is (strictly) positive. In response to a price increase of product 1, a fraction of its consumers substitute away to product 2. The multi-product firm MP internalizes this effect and therefore sets higher prices for both products 1 and 2 as compared to the case when each product is owned by a separate firm. We can now think about two cases that are directly related to the Carlsberg-Pripps merger and our counterfactual merger scenarios. First, suppose that firms MP and SP merge and that no divestitures are required. The optimal post-merger prices of products 1 and 3, $\widetilde{p}_1$ and $\widetilde{p}_3$, satisfy the following conditions.

$$
\begin{aligned}
\widetilde{p}_1 &= mc_1 - \left(\frac{\partial s_1}{\partial \widetilde{p}_1}\right)^{-1}\left(s_1 + \frac{\partial s_2}{\partial \widetilde{p}_1}(\widetilde{p}_2 - mc_2) + \frac{\partial s_3}{\partial \widetilde{p}_1}(\widetilde{p}_3 - mc_3)\right) \\
\widetilde{p}_3 &= mc_3 - \left(\frac{\partial s_3}{\partial \widetilde{p}_3}\right)^{-1}\left(s_3 + \frac{\partial s_1}{\partial \widetilde{p}_3}(\widetilde{p}_1 - mc_1) + \frac{\partial s_2}{\partial \widetilde{p}_3}(\widetilde{p}_2 - mc_2)\right)
\end{aligned}
$$

[4] See also Hellerstein (2008) or Rojas (2008), who examine the beer market with similar tools as the merger simulation literature does, but focus on the pass-through of exchange rates and of excise taxes, respectively.

[5] The first-order condition for product 2 is analogous to that for product 1.

We now define the instantaneous pricing pressure at the product level as the difference between the post- and pre-merger first-order conditions, $\Delta p_1 \equiv \widetilde{p}_1 - p_1$. To do so, we evaluate the post-merger first-order conditions at pre-merger prices. This makes clear that UPP is an approximation to a merger simulation. We do not solve for post-merger equilibrium prices, but pin down the direction of price changes following the merger.

$$
\begin{aligned}
\frac{\Delta p_1}{p_1} &= -\widehat{\eta}_1^{-1}\left(\frac{\partial s_3}{\partial p_1}(p_3 - mc_3)\right) > 0 \\
\frac{\Delta p_3}{p_3} &= -\widehat{\eta}_3^{-1}\left(\frac{\partial s_1}{\partial p_3}(p_1 - mc_1) + \frac{\partial s_2}{\partial p_3}(p_2 - mc_2)\right) > 0
\end{aligned}
$$

$\Delta p_1/p_1$ is the net relative price change in response to the merger and $\widehat{\eta}_1 \equiv (\partial s_1/\partial p_1)p_1$ is the semi-elasticity of demand for product 1 and gives the change in market share of product 1 in response to a one percent change in p_1.
The product-level pricing pressure of product 1 is driven by two components: its own-price (semi-)elasticity and the diverted profit margin from its former competitor product 3. Only the diverted margin of the former rival matters for the instantaneous pricing pressure, because we are evaluating each firm's price response to the merger at pre-merger prices. The diversion effect stemming from products that were owned before the merger is netted out. In the absence of any divestiture requirements, we thus obtain the classic result of Deneckere and Davidson (1985) that prices increase post-merger. Note that the price adjustment of product 3 is driven by two profit diversion terms and, conditional on prices and magnitudes of the cross-price effects, the price of product 3 changes more in response to the merger than the price for product 1 does.
To shed some light on the effect of divestitures, consider a second case where firms MP and SP are cleared to merge under the condition that product 2 is divested as an independent rival firm. The pricing pressures for each of the products are then.

$$
\begin{aligned}
\frac{\Delta p_1}{p_1} &= -\widehat{\eta}_1^{-1}\left(\frac{\partial s_3}{\partial p_1}(p_3 - mc_3) - \frac{\partial s_2}{\partial p_1}(p_2 - mc_2)\right) <> 0 \\
\frac{\Delta p_2}{p_2} &= -\widehat{\eta}_2^{-1}\left(-\frac{\partial s_1}{\partial p_2}(p_1 - mc_1)\right) < 0 \\
\frac{\Delta p_3}{p_3} &= -\widehat{\eta}_3^{-1}\left(\frac{\partial s_1}{\partial p_3}(p_1 - mc_1)\right) > 0
\end{aligned}
$$

We first note that the price of the divested product 2 falls unambiguously. Its post-merger pricing does not take into account the diverted profit margin to product 1 and thereby its price falls. The sign of the price change for product 1, however, is ambiguous. If the diverted profits from variety 2 exceed those of the new addition to firm MP's portfolio, product 3, the merged firm finds it optimal to reduce the price of product 1 following the merger. As before, the price of product 3 unambiguously rises post-merger, because it benefits from the diverted profit margin stemming from product 1. Thus, our example shows that properly chosen divestitures can induce falling prices for the divested varieties, and if the divestitures are sufficiently important in terms of their diverted profit margins, it is also possible that the prices of the merging parties' products fall post-merger. Finally, the example illustrates that *ceteris paribus*, the smaller party to a merger raises prices more strongly because it internalizes more diverted profit margins than the bigger party.
How does this carry over to the demand model that we estimate below? The structural model gives us estimates of $\partial s_i/\partial p_j$ for arbitrary i, j. The own- and cross-price elasticities depend on our estimates of heterogeneous consumer preferences and their interplay with product characteristics. The price effects of a merger depend on the estimated elasticities

and on which products merge and on which products are divested. The structural demand model that we use allows for the same ambiguity as the stylized example and our findings are therefore driven by the data and not the specific assumptions placed on demand.

3 The Swedish Beer Market and the Merger

3.1 The Retail Setting and the Data

Our data set includes the monthly nationwide retail sales of all beers with a minimum alcoholic content of 3.5 percent of volume and has been provided to us by the Swedish retail monopoly for alcohol, Systembolaget. The data covers the period from January 1996 to January 2003. The merger between Carlsberg and Pripps was consummated in February 2001. We therefore have data for a post-merger period of almost two years. Sales volume per month and price per liter measured in Swedish krona (SEK) are observed at the barcode level.[6] We use the terms product and beer interchangeably to denote a beer with a certain name and certain characteristics - Guinness Draught Beer with 4.2 percent alcohol by volume is a an example of a product. Several of the products are available in different container sizes, this specific beer from Guinness is for instance available in three different containers. We use the term variety to refer to a specific product in a specific container, such as Guinness Draught Beer with 4.2 percent alcohol by volume in a 33 cl bottle.

Systembolaget is wholly owned by the government and its purpose is to supply alcohol without profit motive (see www.systembolaget.se/English/ for an overview). Systembolaget's suppliers are independent profit maximizing firms and Systembolaget is monitored by the Swedish Competition Authority twice yearly on behalf of the EU to ensure that it provides a level playing field for different suppliers. The suppliers of beer to Systembolaget can be categorized as major brewers, microbreweries and pure importers. The major brewers are Åbro, Carlsberg, Kopparbergs, Krönleins, Pripps, and Spendrups. Pripps was at the time of the merger owned by Norwegian food and drinks group Orkla. Krönleins and Spendrups are family controlled domestic brewers and Åbro and Kopparbergs are independent Swedish brewers. Each of these breweries produces and sells to Systembolaget a number of its own beers, they also produce some beers on license agreements with foreign brewers and act as importers and wholesalers for still other beers. For instance, at the start of the period under review Carlsberg was the wholesaler for imported beers under the brands of Budweiser, Caffrey's, Michelob and Staropramen. Micro breweries and independent importers make up a small share of overall volume but control a large number of beers.

The majority of beers and producers sell very little and as we are interested in identifying the effects of a merger between the two biggest beer producers in Sweden, we focus on the largest firms in our study. In particular we include the eight largest firms in our analysis (the six major brewers plus the next two in terms of size, see Table (A.1) for the average market shares). These firms account for more than 95 percent of the total market volume during our sample period.[7]

Apart from prices and liters sold, Systembolaget provides other useful information on

[6] In November 2000, 8.62 SEK equaled one Euro and 10.08 SEK equaled one US Dollar. The average price for a liter of beer was roughly 4 Euros and 3.4 US Dollars, respectively.

[7] Excluding minor brands or competitors is common in structural merger analysis. In the merger analysis of Nevo (2000a), for instance, the data covers between 55 and 60 percent of national market shares.

Table 1: Observable Product Characteristics

	Mean	Std. Dev.		Mean	Std. Dev.
Price per Liter (SEK)	32.89	7.26	Bottle (.5 L)	.22	-
Richness	5.70	1.75	Can (.33 L)	.06	-
Sweetness	2.22	1.33	Can (.5 L)	.37	-
Bitterness	6.08	2.06	Ale	.07	-
Alcohol (% of Vol.)	5.43	1.03	Dark Lager	.04	-
Advertising (mln SEK)	.16	.80	Light Lager	.83	-
Foreign	.42	-	Stout	.03	-
Bottle (.33 L)	.35	-	Wheat Beer	.03	-

Note: Based on 16,867 observations. Means for category or dummy variables are the fraction of beers that fall into this category.

product characteristics in its catalogs. These catalogs are freely available online or in the retail outlets of the monopoly, so that consumers have easy access to them. The observable characteristics are presented in Table (1) along with their means and standard deviations. Beers are classified into the following categories: ale, dark lager, light lager, stout and wheat beer segments.[8] We exclude beers that are not sold in one of the standard sizes, namely bottles and cans of around .33 or .5 liters.[9]

The retail monopoly also provides consumers with taste information. It attributes to each beer values for richness, sweetness and bitterness, which are all measured on a scale from 1 to 12, where higher values indicate a more intense taste. We also observe alcohol content and a measure of advertising expenditure.[10] During the period of study advertising of alcoholic beverages was illegal in Sweden. An exemption was made for beer with an alcoholic content below 2.25 percent of volume. These kinds of beers could be freely sold in regular supermarkets. In cases where such a low-alcoholic beer has the same name as a high-alcoholic beer sold in the outlets of Systembolaget, we expect some spillover from the advertising expenditure on the low-alcoholic version to demand for the high-alcoholic product. We observe positive advertising expenditure, defined in this way, for less than 9 percent of our observations.

3.2 Systembolaget's Pricing Policy

The retail monopoly enforces uniform pricing for each beer variety across all of its stores in Sweden. The retail monopoly applies a fixed formula when determining the prices charged to retail consumers.

$$p_j^r = (p_j^w + x_j^a \tau^a)(1 + \tau^c)(1 + mk^s) + d_j \tag{1}$$

p_j^r and p_j^w are the retail and wholesale price of beer j, respectively. τ^a and τ^c are the alcohol excise tax and value-added tax, while mk^s and d_j are the markup of the retail monopoly and the deposit for the packaging of product j, respectively. x_j^a is the alcohol

[8] The term light refers to the color of the lager not its calorific content.

[9] Cans of .45 liter and .6 liter are included, but kegs of 3 liters are excluded in the demand estimation. The volume and sales share of dropped package sizes is minuscule.

[10] Source: Research International/SIFO. Advertising expenditure is the estimate of the total cost of advertising for a given beer in magazines, newspapers, television and billboards based on observed advertising.

content offered by beer j. (1) is publicly known. Systembolaget does not control the alcohol excise tax, the value added tax or its own markup. Both taxes, as well as the retail markup are set by Swedish parliament. The deposit for the different kinds of packages is small relative to the retail price of the beer. Thus, when firms set wholesale prices, they are in effect directly determining the final retail price charged to consumers. This means that we, in contrast to much other work on the price effects of upstream mergers, do not need to make assumptions about unobservable retailer markups.

3.3 The Merger and the Divestiture Requirements

The takeover of Pripps by Carlsberg had an international dimension and was investigated by competition authorities in Denmark, Finland, Norway and Sweden. Carlsberg merged with Norwegian brewery Ringnes, which owned Pripps. [11] According to reports at the time of the merger, one important motivation was that Carlsberg wanted access to Baltic Beverages Holding Co., which had a strong position in the Russian beer market, and of which Ringnes owned 50 percent. Carlsberg and Pripps also sell beer with alcohol content below 3.5 percent of volume that is retailed in supermarkets as well as bottled water and carbonated soft drinks. By focusing on the market for beer with alcohol content above 3.5 percent of volume, we thereby only examine part of the merger. This part is viewed as a separate relevant market in product and geographic space by the Swedish Competition Authority.

During the fall of 2000, the merger is investigated by the Swedish Competition Authority and, conditional on a number of divestitures, the authority announces that it does not challenge the merger in December 2000. The merger is consummated in February 2001. The agreed upon divestitures are implemented in two steps. In February 2001, when the merger is consummated, the merging parties divest a total of 19 beer varieties, for instance the products associated with the brands of Bass, Caffrey's and Staropramen in addition to a number of purely domestic products such as Arboga and TT. The acquirer of all these beers is the firm Galatea, which at the time of the merger captures a negligible share of the market. [12]

Table (2) presents information about the concentration level as well as the number of beers sold and the share of total volume for the firms of particular interest in the merger. The set of beers that are divested in the consummation of the merger in February 2001 are also shown. We do not strip these beers from the pre-merger holdings of Carlsberg and Pripps. Adding all the market shares during this period across columns therefore double counts these beers.

Until the merger in 2001, the concentration level in the industry falls from .24 to .21. The clearing of the merger, however, raises the Herfindahl-Hirschman Index (HHI) above .25 and thereby, according to the U.S. merger guidelines, changes the industry concentration level from moderate to high. At the time of the merger, Carlsberg and Pripps account for almost half the total volume and their combined portfolios include 50 beer varieties. The beers divested in February 2001 account for 6 percent of the total liters sold and are

[11] Ringnes was itself owned by Orkla, the Norwegian food and drinks group. A joint entity was created under the name Carlsberg Breweries, with Orkla receiving a 40 percent share.

[12] Almost two years later, in November 2002, the merged firm also divests two varieties of Lapin Kulta to Åbro, the by that time third largest producer in terms of market share. The beers divested at the time of the merger account for 63 percent of the volume to be divested. Because of the long period separating the merger from the Lapin-Kulta divestitures our examination focuses on the beers that are divested at the time of the merger.

Table 2: Market Shares and Number of Beers by Firm

Year	HHI	Carlsberg	Pripps	Spendrups	Divestitures	Galatea
1996	.24	.22, 28	.26, 26	.24, 39	.09, 11	.005, 11
1997	.23	.20, 30	.26, 26	.27, 45	.07, 11	.003, 11
1998	.22	.19, 31	.26, 27	.24, 46	.09, 13	.002, 13
1999	.23	.23, 31	.27, 26	.20, 41	.07, 15	.003, 12
2000	.21	.29, 27	.17, 23	.19, 40	.06, 19	.003, 12
2001	.25	.43, 52	-	.20, 37	.06, 19	.071, 31
2002	.25	.41, 48	-	.21, 35	.05, 15	.114, 34

Note: For the firm columns, the first number reported is the firm's average annual market share. The second number is the average number of beer varieties the firm sold during that year rounded to full intergers. We do not report 2003, because we only include January of 2003 in our sample. The divestitures are also included in the pre-merger numbers reported for Carlsberg and Pripps.

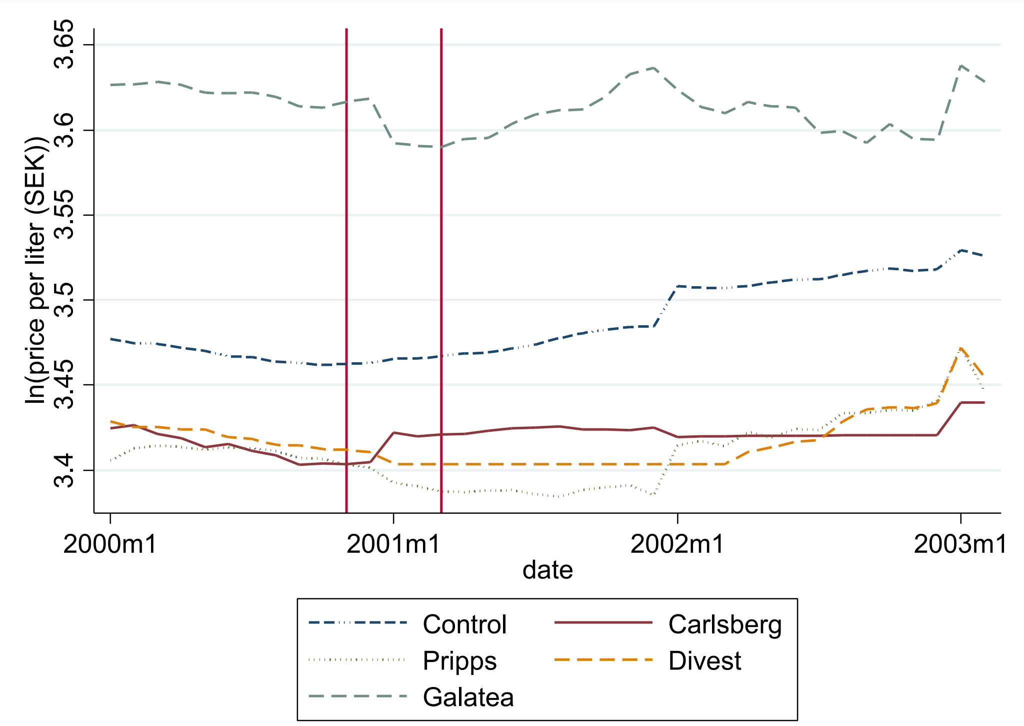
controlled by Galatea in the post-merger period. Before the merger, Galatea accounts for a negligible share of the market, even though it sold 12 varieties in the outlets of the retail monopoly. Spendrups becomes the second largest producer following the merger, and accounts for roughly a fifth of total volume and sells around 40 varieties.

4 A Difference-in-Difference Analysis of the Merger's Effects

As we will see below there was little effect on prices, despite the merger of two firms that jointly control roughly half of the market. This could be the result of a limited market power effect of the merger, or it could reflect cost or demand shocks that counteract the incentives to raise price.[13] Our main interest is to assess if the divestitures associated with the merger are contributing to the muted price response post-merger. As in the ex-post evaluations of mergers in Focarelli and Pannetta (2003), Hastings (2004) and Ashenfelter and Hosken (2010) we adopt a difference-in-difference regression approach. As argued by Angrist and Pischke (2010), the methodology has proved fruitful in many areas of economics. While acknowledging the value of difference-in-difference estimates in describing patterns, we should also note that applying the method to major mergers in a national market is fraught with difficulty - any "control" group that we define is possibly also affected by the merger. In particular we may fear that if a merger triggers the merging parties to raise their prices others in the market will also raise their prices which may lead to a downward bias in the estimated coefficient on the price effect of the merger on the merging parties. As we discuss below, the results and robustness examinations that we perform do not indicate such a bias. Nevertheless we would like to stress that results on a merger between the two largest firms in a national market should be viewed with more

[13]The two most prominent sources of market-wide changes of prices of beer in Systembolaget are changes in excise taxes and changes in Systembolaget's markups. There are no such changes during the window studied after a change in Systembolaget's markup on January 1, 2000. We have also experimented with including various cost (such as labor costs in country of production, price of barley) and demand shocks (such as disposable income) in the difference-in-difference regressions below. As many of these variables are cointegrated their individual significance is low and sensitive to what other controls are included. The merger dummies as defined below were stable across specifications however.

Figure 1: Price Trends Before and After the Merger



Note: The Figure plots 3-month moving averages of the mean log-price for each of the groups. The left vertical line indicates November 2000, the end of our preferred pre-merger window. The right vertical line indicates March 2001, the beginning of our preferred post-merger window.

humility than we view for instance results on the effects of class size on student achievement, where natural experiments can provide us with exogenous variation across a large number of otherwise comparable observations (see Einav and Levin (2010) for a further discussion of the relative merits of demand system estimation vs. difference-in-difference methods in merger cases). Thus, we regress the log-price of variety j in month t on several potential price shifters. Our baseline specification looks as follows.

$$\begin{aligned} \ln(p_{jt}) =& \gamma_j + \gamma_t + \beta_1 post_t * Carlsberg_{jt} + \beta_2 post_t * Pripps_{jt} + \ldots \\ & \beta_3 post_t * divest_{jt} + \beta_4 post_t * Galatea_{jt} + e_{jt} \end{aligned} \tag{2}$$

γ_j and γ_t are product-container and month fixed effects, respectively. We therefore account for time-specific and persistent variety specific shocks to prices. $post_t$ is an indicator variable that equals one if market t is observed after the consummation of the merger and zero otherwise. Similarly, $divest_{jt}$ indicates whether beer j belongs to the set of beers that are divested following the merger. The firm indicators are defined analogously, where the indicators for Carlsberg and Pripps have no overlap with the divested set of beers. e is the residual term.

Let us highlight two challenges in applying such a methodology to a merger between two major players on a national market. One challenge is that the mergers that are interesting for competition authorities are typically not exogenous events that come as a surprise to the affected parties. In consequence it may be difficult to make a clear distinction between before and after treatment. In our case the merger was cleared in December 2000 and consummated in February 2001, but the firms had agreed to merge already in May 2000, possibly after long negotiations. Strategic behavior to try to influence the terms of the deal may have affected prices also before May 2000. On the other hand, the earlier one defines the pre-merger period, the more other shocks due to for instance entry and exit of beers are likely to obscure the comparison. The concerns regarding timing are therefore difficult to solve in a perfectly satisfying manner, but as we illustrate below, our qualitative results are not sensitive to how event windows are defined.

The other major challenge in implementing difference-in-difference evaluations of mergers regards defining a control group as already noted. Products that are faced with the same cost and demand shocks as those involved in the merger are likely to be in competition with the merging parties and thus the treatment may have an effect also on the control group.[14] To explore price developments we use Figure (1) which plots price trends for the merging parties, the divestitures, Galatea and our preferred control (all other products). As seen average prices for the merging parties fall relative to the control group - one possible reason for this can be efficiencies due to the merger. We may also use estimates from an estimated demand system to inform our choice of control groups - if cross-price effects are low enough this limits the concerns that the difference-in-difference estimation will yield biased coefficients. In Section 6.2 we return with a discussion of how the estimates from the difference-in-difference estimation match up with the results from the estimation of our demand system.

A requirement for an appropriate control group in a difference-in-difference framework is that the parallel trends assumption is reasonable - that the treatment and control would have followed the same developments absent treatment. One way to examine the reasonableness of the parallel trends assumption is to consider if prices move in tandem prior to the merger. We see that indeed prices of the merging parties and the control follow similar trends prior to the merger. To further examine this we have also run a regression using log-prices, where the groups that are directly affected by the merger are allowed to follow individual trends. The results are presented in Table (A.2) in the Appendix. Overall, we find that even when a pre-merger trend is estimated to be statistically significant, its impact is so small that our results are not sensitive to it. We conclude that pre-merger trends are highly similar for the control group and the treated.

Given the concerns regarding the difference-in-difference setup, we estimate specification (2) with several choices of pre- and post-merger windows as well as different control groups. Columns (1) to (3) in Table (3) vary the pre- and post-merger windows for our preferred choice of control group. Columns (4) and (5) vary the control group.

In column (1), we define the pre-merger period to cover January-November 2000, and the post-merger period to stretch from March 2001 to January 2003. This is the longest post-merger period that our data allow. The control group is all beers that are not directly involved in the merger.

The price of the divested beers are estimated to fall by about 3.2 percent post-merger. This is fully in line with our stylized example above and as can be seen from columns (2) to (5) it is robust across all of the specifications. The divestitures therefore generate downward pricing pressure in this particular merger. Moreover, Carlsberg and Pripps are both estimated to decrease their prices post-merger by 1.6 and 3.2 percent respectively. Following the logic of our stylized example, this finding suggests that the set of divested products generates sufficient pricing externalities to induce the merging parties to lower

[14] One source of identification is to examine price developments in neighboring markets that are affected by the same cost and demand shocks but not by the merger per se. In our case the merger affects the whole country and using prices in neighboring countries as control group is hampered by differences in market structure, that the merger affects those nations as well and finally Sweden's floating exchange rate creates substantial noise in cross-country price comparisons.

Table 3: Difference-in-Difference Estimates

Dep. Var.	(1) $\ln(p_{jt})$	(2) $\ln(p_{jt})$	(3) $\ln(p_{jt})$	(4) $\ln(p_{jt})$	(5) $\ln(p_{jt})$
post * divest	**-.0319**	**-.0294**	**-.0356**	**-.0288**	**-.0528**
	(.0024)	(.0023)	(.0034)	(.0023)	(.0030)
post * Carlsberg	**-.0159**	**-.0127**	**-.0168**	**-.0130**	**-.0369**
	(.0020)	(.0019)	(.0028)	(.0020)	(.0028)
post * Pripps	**-.0316**	**-.0299**	**-.0318**	**-.0292**	**-.0533**
	(.0022)	(.0021)	(.0030)	(.0022)	(.0029)
post * Galatea	**.0261**	**.0116**	**.0316**	**.0294**	**.0074**
	(.0027)	(.0026)	(.0038)	(.0026)	(.0034)
Constant	**1.020**	**1.018**	**1.020**	**1.290**	**1.288**
	(.0046)	(.0046)	(.0051)	(.0051)	(.0055)
Observations	**6,826**	**4,616**	**5,618**	**4,089**	**3,324**
R^2	**.99**	**.99**	**.99**	**.99**	**.99**

Note: Standard errors are reported in parentheses. Market fixed effects and beer variety fixed effects are included in all specifications. In columns (1) and (2) the pre-merger window stretches from January to November 2000. In column (3), the pre-merger window is shortened to May 2000. The post-merger window covers March 2001 to January 2003 in columns (1) and (3), while the post-merger window is shortened to January 2002 in column (2). All other specifications have the same window definitions as column (1). In column (4) Spendrups is the control group, whereas in column (5) Bibendum, the smallest firm in the market, is the control. In all other columns the control group consists of all firms other than the merging parties and Galatea.

their prices post-merger. An alternative explanation is that the merging parties realize efficiency gains in the post-merger period that overwhelm the market power effect of the merger. Further below, we use our structural model to shed light on this question. Finally, Galatea raises prices by 2.6 percent post-merger, which also squares well with the stylized example as it takes into account the diverted profit margins to the newly acquired divestitures post-merger.

In column (2) we shorten the post-merger window to February 2002, thus covering one year following the consummation of the merger. The estimated price effects are qualitatively identical and quantitatively similar. If anything, the effects are lessened marginally. This indicates that to a very limited extent the price effects build up during the post-merger period. Mostly, this applies to Galatea. In column (3), we shorten the pre-merger window to May 2000, the point at which the firms officially agree to merge, and extend the post-merger window again to January 2003. Again, the estimated price effects are qualitatively identical and quantitatively similar. The latter finding applies somewhat less to Galatea, which might simply be the case, because it had a negligible size pre-merger. This is likely to make the prices of the beers offered by Galatea somewhat more volatile. Overall, the results suggest that the specification is quite robust to the choice of pre- and post-merger window.

We next turn to the choice of control group. In column (4), we choose Spendrups, the largest rival to the two merging parties, as the control group. This is the firm for which demand and cost shocks are likely to be the most similar to those of the merging parties. We see that the estimated price effects are similar to the results in columns (1) to (3). In column (5), we choose Bibendum as the control group. This firm is somewhat larger than Galatea in terms of market share and is therefore the second smallest firm in the sample at the time of the merger. The estimated price effects turn out to be much larger in magnitude than with Spendrups as the control. Note however that Bibendum is a wholesaler that only imports and may thus face different shocks to efficiency and demand than the merging parties. Despite these quantitative differences, it is clear that the qualitative patterns are identical across the different choices of pre- and post-merger windows and control groups. In previous versions of this paper we also used various sets of micro breweries as control groups and for these as well the qualitative patterns were the same. We adopt the choices in column (1) as our preferred specification, because it covers the longest post-merger period and defines the broadest control group. We find that the merger has a negative impact on the merging parties' prices and that the prices of divested beers are predicted to fall post-merger. To analyze these findings in more detail, we turn to our structural model.

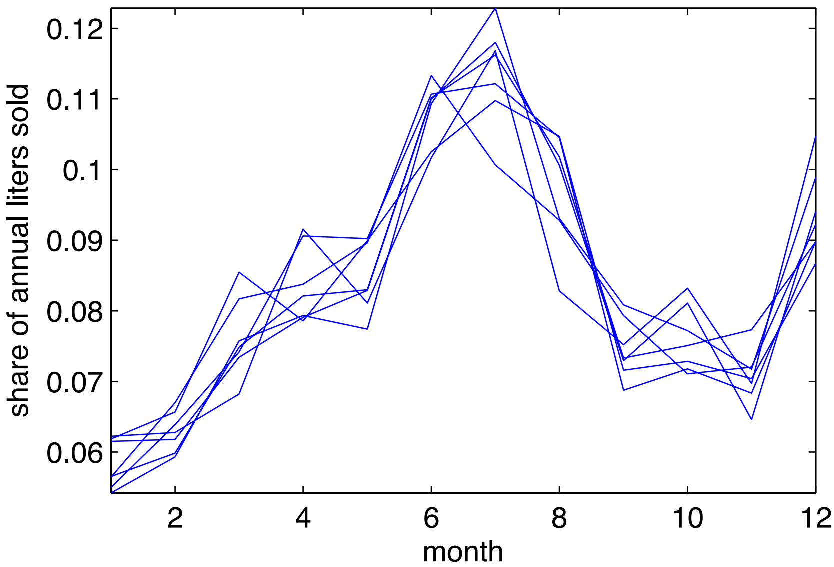
5 A Structural Model of the Swedish Beer Market

As in the merger simulation in Nevo (2000a), we follow the seminal work of BLP, and estimate a random coefficients logit model for demand. The BLP framework allows for very flexible substitution patterns between product varieties and thereby breaks the independence of irrelevant alternatives property that substantially restricts substitution patterns in a standard logit model. As the BLP approach is widely used and described in detail in for example Nevo (2000b), we keep this section deliberately short and focus on our specific choices for estimating the model. Dubé et al. (2012) and Knittel and Metaxoglou (2014) raise several important issues regarding computational aspects of the estimation. We take these issues seriously in our implementation and present the relevant computational details in the Appendix.

5.1 Definition of the Relevant Market

Before estimating our structural model of demand, we need to define the relevant market. After all, if prices of all beers sold by Systembolaget were to rise, we expect demand for beer to fall so that consumers substitute towards the outside good. The volume of liters sold in the Swedish beer market follows two empirical regularities. Firstly, the annual sales cycle is strongly seasonal with peaks in summer and around Christmas time and troughs during autumn and early spring. This regular pattern can be seen in Figure (2). Each month's share of the annual sales volume is plotted for each of the seven years. The plot illustrates that the seasonal pattern is highly stable. Secondly, the volume of beer sold grows on average by roughly eleven percent annually. Part of this change may be driven by changes associated with more liberal attitudes to alcohol following Sweden's entry into the European Union in 1995. Changes in the share of the outside good can drive substitution patterns and we want to avoid that a hard to verify assumption on a decreasing share of the outside good is driving our results. To arrive at a stable share of the outside good, which still allows consumers to opt for not buying beer if prices rise, we follow the approach of Conlon and Holland Mortimer (2010). We initially guess that the potential market size is ten times the actual number of liters sold. We then regress this measure of potential liters sold on a full set of month fixed effects, a linear trend and a constant. This regression explains almost 94 percent of the variation in the relevant market size. Our definition of the relevant market is then the market size predicted by this

Figure 2: Liters of Beer Sold Monthly and Annually



Note: Each of the seven years in our sample is plotted as a separate line. We do not plot January 2003, our only observed market for 2003.

regression. The resulting share of the outside good is quite stable around ninety percent, but the method still allows for substantial substitution between the inside goods and the outside good in response to large aggregate price changes. In the estimation below, we also include a full set of time fixed effects to prevent any remaining errors in our definition of the relevant market to affect the estimated coefficients.

5.2 Demand

We observe $t = 1, \dots, T$ months with J_t varieties in each. The indirect utility that a particular consumer i derives from purchasing variety j at time t is given as follows.

$$u_{ijt} \;=\; x_{jt}\beta - \alpha p_{jt} + \xi_f + \xi_t + \iota_{j\in I}\xi_{b_j} + \xi_{jt} + \textstyle\sum_{k=1}^{K} \sigma^k x_{jt}^k v_i^k + \sigma^p v_i^p p_{jt} + \epsilon_{ijt} \qquad (3)$$

ϵ denotes the *iid* logit errors or idiosyncratic taste shocks. p_{jt} is the price and x contains all of the observable characteristics that are presented in Table (1). We allow for a random coefficient on price and include $K = 3$ additional random coefficients on the attributes sweetness, alcohol content and the constant. The latter captures heterogeneous consumer preferences for the outside good. Heterogeneity is captured by the simulated v_i vector, which contains K draws from a standard normal distribution that is distributed *iid* across consumers. σ contains the standard deviations for the random coefficients. Thus, we have $\alpha_i \sim N(\alpha, \sigma_p)$ and $\beta_{ik} \sim N(\beta_k, \sigma_k)$, where α and β_k are the estimated means for the random coefficients on price and the observable characteristic k.
The structural error term is disaggregated into several components to make the specification more flexible and ensure that the *iid* assumption is justified. Our main concern here is that there are unobserved branding effects in the data.[15] We allow for such effects

[15]Store coverage is another factor that can yield large systematic unobserved differences between products. Note, however, that beers with good brand images are also likely to be sold in most or all of the

along two dimensions. First, consumers can attach specific brand images at the firm level. This is captured by ξ_f, which is a firm-level fixed effect. Second, we allow for product fixed effects. We define a product here as a beer in all its available packages; at most there are four. Advertising expenditure also accrues at this level. ξ_{b_j} is the corresponding product fixed effect. Ideally, we would allow for a full set of these product fixed effects. Due to data limitations, which we describe below in our discussion of endogeneity, we only include product fixed effects for the products that are available in each market in the sample period. These are products for which at least one of its packages is available in all 84 observed markets. There are 52 such products and they constitute on average 63 percent of volume. $\iota_{j \in I}$ is an indicator variable that equals one if j belongs to the set of these beers and zero otherwise. ξ_t is a time fixed effect and ξ_{jt} is a demand shock that is unobserved by the econometrician.

In an ideal setting we would allow for additional random coefficients and estimate a full covariance matrix for these coefficients, as this would give a more flexible demand model. We stick to four independently distributed random coefficients due to data limitations. Relevant consumer survey data, which would have allowed us to append additional micro moments to the estimation as in Petrin (2002), are not available.[16] With the retail monopolist enforcing uniform prices across all of Sweden, regional or store-level data would give us no additional price variation to exploit, but we could have included regional variation in demographics directly in the estimation. Due to the national aggregation of the data, this is not possible. Without these additional sources of observable variation, it is difficult to estimate a large number of random coefficients and their covariances.[17] This level of aggregation is not uncommon in the existing literature. BLP, Björnerstedt and Verboven (2013), and Eizenberg (2014), for example, face similar data limitations. We believe that the demand specification that we adopt is as flexible as the data allows it to be. Moreover, note that the BLP model is typically viewed as allowing for at least as flexible substitution patterns as the nested logit model, which has been implemented in Peters (2006) and Weinberg (2011), and multilevel demand systems as used for example in Weinberg and Hosken (2013). As we discuss in more detail below, our BLP model yields a substantially better fit of the data than a logit model and the observable product characteristics strongly affect product substitutability. Lastly, let us stress that competition authorities are likely to be faced with similar data limitations when they have to make a decision regarding the potential approval of a specific merger. We believe that our frank discussion of the challenges involved and that the results we obtain illustrate that structural demand estimation can still be a useful tool to guide decision-making in such a setting.

5.3 Supply

We assume Nash-Bertrand competition in prices between firms to model the supply side of the Swedish beer market. There are $f = 1, \ldots, F_t$ firms present at date t and each firm

stores. These two aspects are therefore interrelated.

[16]The simple reason is that the statistical office in Sweden does not disaggregate expenditures on alcoholic beverages in sufficient detail.

[17]In a specification where we only allow for firm fixed effects, we also attached a random coefficient to the foreign dummy. When including product dummies, however, this fifth random coefficient becomes insignificant.

maximizes profits for its portfolio of products $\mathcal{F}_f$.

$$\max_{\{p^w_{jt}\}} \Pi_f = \sum_{j \in \mathcal{F}_f} M_t s_{jt}(p^r_{jt})(p^w_{jt} - mc_{jt}) - C_f$$

M_t is the market size at date t, mc_{jt} is the marginal cost of production for beer j and C_f is the fixed cost faced by firm f. We distinguish between retail and wholesale prices, because product shares and elasticities are functions of the prices charged to consumers, while firm margins directly depend on the prices charged to the retail monopoly.
Knowledge of the monopoly's pricing rule, (1), allows us to precisely back out firm margins, $p^w_{jt} - mc_{jt}$. We define $\kappa_t \equiv \partial p^r_{jt}/\partial p^w_{jt} = (1 + \tau^c_t)(1 + mk^s_t)$.[18] The first-order profit maximization condition for product $j \in \mathcal{F}_f$ is given as follows.

$$\sum_{k \in \mathcal{F}_f} \frac{\partial s_{jt}}{\partial p^r_{kt}}(p^w_{kt} - mc_{kt}) = -\frac{s_{jt}}{\kappa_t}$$

Switching to matrix notation, we collect the profit maximization conditions for all firms in the market. Let Ω denote the matrix of market share derivatives where $\Omega_t(j,k) \equiv \partial s_{k,t}/\partial p_{j,t}$ is its entry in the j^{th} row and k^{th} column. Moreover, we define the holding or ownership matrix, $\mathcal{H}_t$, whose entry in row j and column k is 1, if varieties j and k are owned by the same firm and is zero otherwise.

$$(\mathcal{H}_t \odot \Omega_t)(p^w_t - mc_t) = s(p^r_t)\kappa_t^{-1},\ \forall t$$

$\odot$ denotes the element wise product.

$$mc_t = p^w_t + (\mathcal{H}_t \odot \Omega_t)^{-1} s(p^r_t)\kappa_t^{-1} \tag{4}$$

Wholesale prices, market shares, κ_t and the ownership pattern are observed directly, while the elements of Ω_t are functions of the estimated demand parameters. By plugging our estimates into (4), we obtain variety-level estimates of marginal costs.

5.4 Endogeneity

The unobservable product characteristic, ξ_{jt}, has a vertical interpretation in the model. All else equal, a higher realization of ξ_{jt} gives variety j a greater market share. As firms incorporate this into their pricing decisions, realizations of the unobservable and price will tend to be positively correlated, which in turn renders prices endogenous and biases the estimated price coefficient towards zero. Finding instruments that shift prices at the product level while not affecting demand is a challenge in differentiated product markets. One potential instrument is the (log of) the exchange rate which acts as a cost shifter for imported products. We use this as one instrument.
For additional instruments we follow the instrumenting strategy of BLP. These instruments rely on the assumption that each product's location in characteristics space is exogenous. The optimal price of each product depends on its own attributes as well as the attributes

[18] The monopolist's pricing rule has changed over the sample period. Until December 1999, it is given by (1). From January 2000 onwards, the retail price is $p^r_{jt} = (p^w_{jt}(1 + mk^s_t) + c_t + x^a_{jt}\tau^a_t)(1 + \tau^c_t) + d_{jt}$, where c_t is a constant charge per container that is applied to all beers. From January 2000, c_t is 1.5 SEK. For backing out the marginal costs implied by our demand estimates, we need $\partial p^r_{jt}/\partial p^w_{jt}$. It is straightforward to verify that this equals $(1 + \tau^c_t)(1 + mk^s_t)$ for all the pricing functions.

of all other rival products. Given exogeneity, any function of the observed non-price characteristics qualifies as an instrument for price. As advertising expenditure is optimally chosen by each firm, we exclude this variable from the computation of the instruments. Unfortunately, we have been unable to find a good instrument for advertising. We do note, however, that in a previous specification, where we only allow for firm-level fixed effects, the magnitude of the advertising coefficient is three times larger than the one reported below. A substantial part of advertising's effect is therefore absorbed by the product fixed effects. This limits the bias in the estimation to some extent. Moreover, due to legal restrictions at the time, advertising expenditure in the Swedish beer market is only allowed for the low-alcoholic products sold in regular supermarkets. We also note that this form of indirect advertising weakens the role of this endogeneity further. Observed changes in advertising expenditure can be caused by demand shocks for beer in supermarkets (at most 2.25 percent alcohol content) and are not necessarily driven by demand shocks for beer in the stores of the retail monopoly. The endogeneity of advertising expenditure therefore only carries over fully if demand shocks in these two markets are perfectly correlated, something which is unlikely. Anecdotal evidence suggests that consumers strongly distinguish between these two types of beer.[19]

This still leaves the issue of how reasonable it is to assume that the observable characteristics are exogenous. Firms' endogenous choice of these product attributes is an important aspect, especially in industries where firms invest to frequently redesign their products to improve quality over time (see for instance Goettler and Gordon (2011), Blonigen et al. (2013)). Such a pattern, however, does not fit the Swedish beer market well. The characteristics of a beer remain constant once it has entered the market. This leaves the entry of new varieties and the exit of old varieties as a means by which firms can choose characteristics. Estimating an entry game with 8 players, on average 200 existing varieties and up to 9 product attributes is beyond the scope of this paper. Moreover, Figure (A.1) in the Appendix at least indicates that there are no substantial trends in characteristics during our sample period. We therefore treat product characteristics as if they are exogenous in our setting.

We instrument for the price of each product with its own characteristics, the sum of the characteristics of all other products sold by the same firm, the sum of the characteristics of all other products sold by rival firms and the average of the characteristics of all products that are available at that date. To raise the variation of the instruments, we compute them within categories. The same practice has been adopted by Bresnahan et al. (1997). To identify the price coefficient, the instrument should vary at the product-level across markets. Except for advertising and price, however, all product characteristics are fixed once a product has entered the market. To see how this feature of the data constrains the extent to which we can include product fixed effects, consider the following hypothetical data setting. Suppose that in addition to product characteristics remaining fixed for each variety during the sample period we also would not observe any entry or exit. In that case, our BLP-type instruments would not vary either. If we include a full set of product fixed effects in such a setting, the matrix of instrumental variables becomes singular. Another way to look at this is to realize that in this setting the product dummies contain the same information in terms of their correlation with prices as do the BLP-type instruments. The data on breakfast cereals used in Nevo (2000a, 2001) is very close to this setting, which make the use of BLP-type instruments infeasible. Instead, Hausman-type instruments, average prices in other markets, are applied. For our data these alternative instruments

[19] Low-alcoholic beer is still popular as a beverage over lunch, while this is not the case for regular beer, for example.

are of no benefit, because the alcohol retail monopoly enforces uniform pricing across all of its stores. We do, however, observe plenty of entry and exit during our sample period. The number of beers varies between 157 and 213. This by itself does not solve the problem of multicollinearity between the product dummies and the excluded instruments, because the product dummy matrix tracks the entering and exiting of beers. When a new beer enters the market, this changes the values of the excluded instruments but also expands the product dummy matrix by an additional column. The multicollinearity issue remains. We therefore have to compromise to make the use of the excluded instruments feasible. Only products with at least one container size sold in each month are treated with product fixed effects. This has two redeeming features. First, given that these beers are present in all of the markets, their characteristics do not contribute to the variation of the instruments by entering or exiting the market. Second, given that these beers are available in all of the seven years in the sample period, these products are also most likely to have built up significant brand capital. The estimated fixed effects indeed suggest that this is the case: out of a total of 52 fixed effects, 44 are positive and have t-stats that are greater than 2. Only 5 have statistically significant negative values.

If we include additional product fixed effects, the BLP instruments are weakened substantially and our estimator runs into many of the problems that Knittel and Metaxoglou (2014) warn about. We find that this is driven by ill-conditioning of the GMM estimator's weighting matrix. This finding is in line with simulation results reported in Conlon (2013) who notes that the problems identified by Knittel and Metaxoglou (2014) can be viewed as the result of weak instruments.

The data limitations we face are not ideal, but the estimates reported below yield sensible results. The estimated mean of the price coefficient is roughly three times larger than the corresponding coefficient estimate in an uninstrumented BLP model and Table (A.3) shows that our excluded instruments explain roughly half of the observable variation in price, while the full instrumental variable matrix explains 78 percent of the observable variation in prices.[20] Moreover, the implied demand estimates yield no inelastic demands for beer varieties so that there are no economic outliers. Finally, as we discuss in more detail below the marginal cost and markup estimates give reasonable numbers.

5.5 Estimation Results

The estimates are reported in Table (4). Consumers on average value beers that are richer in taste, and dislike both relatively sweet, relatively bitter and high-alcoholic beer varieties. The estimates also point to a home bias. Among the five beer categories, ales and stouts are most preferred, while among the packages, the .33 liter can is most preferred. Except for the constant, all of the random coefficients are statistically significant and as the large value for the Wald statistic shows, the data strongly rejects the standard logit model against the BLP model. We find substantial heterogeneity in consumers' sensitivity to price and consumer taste for alcohol content and sweetness. Moreover, apart from being significant in a statistical sense, the random coefficients also substantially impact substitution patterns. The average cross-price elasticity is more than six times larger than the elasticity implied by the corresponding logit model.[21] The estimated price coefficient is highly statistically significant and implies an average own-price elasticity of roughly -6. This value is close to previously obtained demand estimates for beer reported in Hausman

[20] For the sake of brevity, we do not report the uninstrumented estimates.

[21] By corresponding logit model, we mean the model that has identical point estimates for the means of all the coefficients, but sets the standard deviations of the random coefficients to zero.

Table 4: Estimation Results

	Mean	Std. Dev.		Mean	Std. Dev.
Price per Liter (SEK)	**-.2063**	**.0439**	Dark Lager	**-.6878**	-
	(.0239)	(.0142)		(.0493)	
Richness	**.1141**	-	Stout	**.6650**	-
	(.0075)			(.1139)	
Sweetness	**-.2458**	**.4305**	Wheat Beer	**-.5882**	-
	(.0531)	(.0334)		(.1427)	
Bitterness	**-.0603**	-	Can (.5 L)	**.9668**	-
	(.0053)			(.0237)	
Alcohol (% of Vol.)	**-1.543**	**1.221**	Bottle (.33 L)	**.6014**	-
	(.6057)	(.2649)		(.0311)	
Advertising (mln SEK)	**.1089**	-	Can (.33 L)	**1.645**	-
	(.0084)			(.0823)	
Foreign	**.3125**	-	Constant	**.1115**	**.1765**
	(.0382)			(.9603)	(5.255)
Ale	**1.217**	-			
	(.1118)				
R^2	.42	$\overline{\eta}_{jj}$	-5.95	$\# \ \eta_{jj} > -1$	0
Wald-Stat. $\sim \chi^2(4)$	278.6	$\overline{\eta}_{jk}$	.0202	$\overline{\eta}_{jk,logit}$	.0031

Note: Based on 16,867 observations. Standard errors are reported in parentheses. Fixed effects for products that are available in every month during the sample period, monthly fixed effects and firm fixed effects are included in the specification, but not reported. $\overline{\eta}_{jj}$, $\overline{\eta}_{jk}$ are the average own- and cross-price elasticities implied by the reported point estimates and $\overline{\eta}_{jk,logit}$ is the average cross-price elasticity implied by setting all the estimated standard deviations of the random coefficients equal to zero.

et al. (1996) and Slade (2004). The point estimates imply that all variety demands are elastic.

5.5.1 Marginal Costs and Markups

To examine the implications of our estimates more closely, we turn to the backed-out marginal costs and markups for the firms that act as suppliers to Systembolaget, which are summarized in Table (5). The estimated average markup of around 38 percent is close to the estimates reported by Goldberg and Hellerstein (2013), who estimate a BLP model for beer demand in the greater Chicago metropolitan area.
A potential gauge for how reasonable these implied markups are, is to compare these structural estimates with their accounting equivalents. We could find accounting data for Carlsberg and Spendrups for each year in the sample period. Using sales net of excise taxes and the reported cost of goods sold from annual reports we can calculate markup over average variable cost at the firm level as 48.6 percent for Carlsberg and 46.7 percent for Spendrups. The equivalent numbers using our structural estimates are 41.6 percent and 39 percent, respectively. We should not expect a perfect match as the accounting data are the aggregate for all product lines and both firms also produce non-alcoholic beverages. Furthermore, a long tradition of empirical work in Industrial Organization also stresses that we cannot equate accounting costs to economic costs. With these caveats in mind, it is reassuring that our structural estimates are of similar magnitude as the accounting

Table 5: Implied Marginal Costs of Production and Markups at the Producer Level

Percentile	Marginal Costs (SEK per Liter)	Markups (%)
1^{st}	1.92	17.3
5^{th}	3.24	19.9
10^{th}	3.90	21.8
50^{th}	6.50	38.7
90^{th}	15.4	50.9
95^{th}	17.3	55.7
99^{th}	22.1	67.1
Mean	8.07	37.9
Std. Dev.	4.56	11.6

Note: Based on 16,867 observations. The reported figures are computed by pooling all backed-out marginal costs and markups for the sample period. The markup of product j is defined as $(p_j^w - mc_j)/p_j^w$.

figures.

6 A Structural Analysis of Divestitures on Merger Outcomes

Our estimated structural model allows us to analyze post-merger outcomes in a *ceteris paribus* fashion, which is not possible in the difference-in-difference approach. Due to a lack of observable cost information, we abstract from the possibility of post-merger efficiency gains throughout the counterfactuals. Instead, we hold marginal costs fixed at their pre-merger values. In a first scenario, we simulate the post-merger equilibrium without imposing any divestiture requirements. This allows us to isolate the merger's market power effect. The second counterfactual also abstracts from efficiency gains, but imposes the divestiture requirements of the actual merger. The difference in outcomes between the first and second counterfactual tells us how effective the actual divestitures are in dampening post-merger price increases and consumer welfare losses that are due to the merger's market power effect.
Market equilibrium prices and quantities are determined by the system of firms' profit maximization conditions.

$$p_t^{cf} = \widehat{mc}_t - \left(\mathcal{H}_t^{cf} \odot \widehat{\Omega}_t^{cf}(p_t^{cf}; \widehat{\xi}_t)\right)^{-1} s_t^{cf}(p_t^{cf}; \widehat{\xi}_t) \tag{5}$$

Given the changed ownership pattern, which is reflected in $\mathcal{H}_t^{cf}$, we solve for the counterfactual prices that satisfy firms' profit maximization conditions. As is common in the literature, we fix the unobservables at their pre-merger level and we consider the case where the product offering remains constant; there is no entry and exit.

6.1 The Observed Divestitures' Impact on Post-Merger Outcomes

Table (6) presents the results regarding the impact of actual divestitures on merger outcomes in detail. The price effects are reported in the top panel, while the middle and bottom panel report profit and consumer welfare effects. We first focus on columns (I) and (II), the counterfactual scenarios that do not allow for the merging parties' post-merger efficiency gains. The 95 percent confidence bounds show that all the relevant pricing effects are statistically significant. In column (I) we present the case where a merger is cleared without divestitures and without efficiency gains. The market-wide average price increase post-merger is 1.6 percent, while this figure drops by two thirds if the divestitures are implemented as shown in column (II). The divested beers themselves are predicted to raise their prices by on average almost 3.6 percent if they are owned by the merging parties post-merger. If control of these beers goes to Galatea, on the other hand, the divestitures on average lower prices by almost 4 percent. This effect is fully in line with our stylized example above and mirrors our findings in all specifications of the difference-in-difference regression of prices. The loss of the diverted profit margins makes it optimal for Galatea to charge lower prices on the divested beers.

With regard to the merging parties, we can see two asymmetries. As implied by the stylized model, Pripps, the smaller party to the merger, raises prices more strongly than Carlsberg, irrespective of whether the divestitures go to Galatea or not. The second asymmetry is that the divested beers are much more effective at reducing the price increases of Carlsberg's beers. This is driven by the fact that most of the divestitures stem from the pre-merger holdings of Carlsberg. Thereby, the loss of the diverted profit margins impacts Carlsberg's pricing decisions more strongly. Finally, Galatea gains some market power by controlling the divestitures post-merger and can thereby raise the prices of the beers it owned pre-merger by on average 1 percent.

The middle panel reports profit changes for each of the merging parties. Without the divestiture requirements the merger would raise profits for the merging parties. We also see in column (I) that the outsider firms benefit more strongly from the pricing externalities generated by the merger than the merging parties themselves. This is in line with the theoretical results of Deneckere and Davidson (1985). In column (II) we note that the merger is unprofitable if the divestiture requirements are imposed. It is important however to note that this drop in profits is entirely due to the loss of revenue from the divested beers. If we only consider the beers that are sold by Carlsberg and Pripps before and after the merger, the combination of the two product portfolios is profitable. The gain in market power, however, is not sufficient to compensate for the lost profit from the divested beers.

Why then do Carlsberg and Pripps merge? First, we do not know how much Galatea paid for the acquisition of the divestitures and the acquisition fee may have more than compensated for the lost profit. Second, and probably more importantly, the acquisition of Pripps' operations in Sweden is only a fraction of the merger's total volume. As noted above, Carlsberg's aim is to gain control of Baltic Beverages Holdings, which is one of the biggest producers of beer in Eastern Europe. Subsequent developments point to the importance of this motivation: during the three years from 2011 to 2013 Carlsberg's profits in the Russian market alone are roughly ten times as large as its Swedish market profits[22].

[22] Based on accounting data for Baltic Beverages Holding and Carlsberg Sweden from www.largestcompanies.com

Table 6: Change in Average Prices Relative to Blocking the Merger (in percent)

	(I) No Divestitures No Efficiency Gains	(II) Actual Merger No Efficiency Gains
Carlsberg	**3.04** [2.45,4.19]	**1.57** [1.29,2.03]]
Pripps	**6.46** [5.07,9.82]	**4.94** [3.87,7.66]
Divestitures	**3.58** [2.88,5.01]	**-3.96** [-6.10,-3.19]
Spendrups	**.32** [.22,.64]	**.14** [.10,.26]
Galatea	**-.04** [-.25,.00]	**1.05** [.85,1.48]
All Beers	**1.60** [1.26,2.36]	**.53** [.41,.76]

Change in Firm Profits Relative to Blocking the Merger (in percent)

	(I)	(II)
Carlsberg	**3.53** [2.71,6.03]	**-15.5** [-15.8,-14.5]
Pripps	**.17** [-.65,.22]	**-8.60** [-10.8,-7.88]
Joint	**2.51** [1.90,4.22]	**-13.4** [-13.6,-13.3]
All Firms	**5.46** [4.26,8.40]	**2.21** [1.75,3.27]

Change in Consumer Welfare Relative to Blocking the Merger (in mln SEK)

	(I)	(II)
$CW_i - CW_{blocked}$	**-4.83** [-8.10,-4.42]	**-2.08** [-4.94,-1.26]

Note: In the actually observed merger, the divested beers are controlled by Galatea post-merger. The simulations are computed using the observed market outcomes in the month prior to the consummation of the merger, January 2001. 95 percent confidence bounds are reported in brackets. These confidence bounds are computed with a parametric bootstrap procedure. 1,000 draws from the estimated asymptotic covariance matrix are used.

It therefore seems reasonable that Carlsberg is willing to make some sacrifices to get the Swedish competition authority's approval for the merger.

The bottom panel reports the merger's welfare effects. Approving the merger without divestitures yields almost 5 million SEK of welfare losses. As with prices, the divestiture requirements prove effective in dampening the adverse effects of the merger. The loss in welfare is more than halved to 2.1 million SEK as a result of divestitures.

6.2 Comparing the Predictions of the Difference-in-Difference Regression with the Structural Simulation of the Merger's Price Effects

A first point regarding a comparison of the two ways of gauging price effects is that structural demand estimates may be used to inform the choice of control group in difference-in-difference estimation for ex-post evaluations of mergers. For instance, while Spendrups is likely to be the most similar to Carlsberg and Pripps in terms of cost structure and demand shocks, one may worry that its prices would be strongly affected by the merger. The results in Table (6) suggest that Spendrup's price response, at 0.14 percent, is low enough not to have a major qualitative impact on the difference-in-difference estimates, even if it is significantly different from 0. This supports the notion that biased results because of the treatment having an effect on the control group is not a major concern in the difference-in-difference estimates.

Let us now turn to a closer comparison of the results from simulations in Table (6) with the difference-in-difference estimates of the merger's price effects in Table (3). Qualitatively the results rhyme well: Prices of divested products fall, prices for the acquirer's legacy products rise and the price effects of the merging parties are relatively minor. There are also some divergences however - quantitatively the estimates do not match fully. For instance the simulations indicate that *ceteris paribus* Carlsberg prices increase by 1.6 percent whereas the difference-in-difference estimates point to a fall in the price of Carlsberg products of 1.6 percent. The difference is most marked for Pripps' products.

A first set of possible reasons for the divergence is that shocks after the merger affect different products in different ways. While the treated and the control group are similar, and we cannot reject parallel trends pre-merger, the possibility remains that idiosyncratic shocks may drive post-merger developments. For difference-in-difference applications in education or labor economics, where we often have a large number of treated individuals, such idiosyncratic shocks may be expected to wash out. In contrast we are here analysing results for no more than four firms. Key potential shocks are due to i) costs, ii) demand shocks that affect a particular product or iii) entry and exit whose effect will differ across products. As regards costs it is clear that efficiencies are a key part of the motivation for a typical merger. Assuming zero efficiencies, as we do inTable (6), is therefore likely to overstate the price effect of the merger. One way to explore this is to back out the cost savings that are consistent with the estimated demand system and post-merger prices, as we do in Table (7). Such an exercise suggests substantial efficiencies - around 5 percent for Carlsberg and around 10 percent for Pripps. We lack exogenous information on marginal costs in these firms so we cannot draw conclusions about the extent to which efficiencies drive observed price developments. What we can say, however, is that efficiencies will drive down observed differences between the simulations and the difference-in-difference estimates and that marginal cost savings of 10 percent for Pripps and 5 percent for Carlsberg would be sufficient to bring them more or less fully in line. One way to explore if idiosyncracies are driving results is to use the structural model to make ex-post comparisons under different scenarios. We have also performed counterfactual simulations where we include the observed post-merger estimates of demand shocks, ξ's, in the simulation, but found these to have a minor impact on the simulated price effects. Moreover, we have investigated the role of entry and exit by simulating post-merger outcomes where we correctly anticipate observed changes to product selection. As with demand shocks, observable entry and exit has a minor impact.

Another potential reason for divergences is of course that the structure we impose - a logit model with four random coefficients coupled with static Bertrand competition is at odds

Table 7: Backed Out Marginal Cost Changes in the Post-Merger Period

Dependent Variable	$\ln(\widehat{mc})$	$\ln(\widehat{mc})$	
post * divest	**-.0004**	post * Åbro	**-.0090**
	(.0040)		(.0035)
post * Carlsberg	**-.0500**	post * Kopparbergs	**.0097**
	(.0037)		(.0043)
post * Pripps	**-.1052**	post * Krönleins	**-.0297**
	(.0038)		(.0041)
post * Galatea	**.0086**	Constant	**.7653**
	(.0044)		(.0058)
post * Spendrups	**-.0157**	R^2	.99
	(.0036)	Observations	6,826

Note: The pre-and post-merger windows are defined identically to column (1) Table (3). The control group is made up of the beers sold by Bibendum. Choosing the same control group as in column (1) of Table (3) yields somewhat lower post-merger efficiency gains for Carlsberg and Pripps of respectively roughly 4 and 10 percent. Standard errors are reported in parentheses. Market fixed effects and beer variety fixed effects are included.

with reality. If we had even richer data and a setting with regional price variation we could have estimated additional random coefficients and perhaps captured substitution patterns even better. We do note, however, that the market level prices and quantities that we use are a common type of data when mergers are evaluated. While any demand system imposes some structure on cross-price effects, also note that the random coefficients logit model that we use is typically seen as allowing for at least as rich substitution patterns as is common in applications and in published work evaluating merger review (nested logit in Peters (2006), Weinberg (2011) or multilevel demand systems (AID system) in Weinberg and Hosken (2013)). As to firm behavior it may be that firms are more sophisticated (playing a dynamic game) or less sophisticated (using crude estimates of demand sensitivity or being subject to behavioral biases) than what we assume. We argue that the reasonable size of coefficients and of the backed out markups, as well as the relatively close match with the difference-in-difference estimates point to that the structural model performs sufficiently well for counterfactual analysis to be of interest. We also have to concede that any evaluation of "close enough" is likely to partly be in the eye of the beholder.

7 A Structural Analysis of Divestitures - The Role of Recipients and of Products

We have now seen that in this case divestitures were of a sufficient magnitude to have a substantial effect on prices. This is supportive for the use of divestitures by competition authorities. Let us in the following use our structural model to examine if we can offer some advice to competition authorities on how to select recipients and which products to require for divestment. We consider different recipients first before turning to the role of products and an overall discussion.

7.1 Varying the Recipient for a Given Set of Products

In the actual merger the control of the divestitures goes to Galatea that accounts only for .3 percent of the volume of liters sold pre-merger. To pin down to what extent the particular choice of recipient affects post-merger prices and consumer welfare changes, we simulate the remaining 6 counterfactual merger cases. We only consider "pure" cases: all of the divestitures go to one recipient firm. Table (A.5) presents the simulated post-merger changes to prices and consumer welfare including 95 percent confidence bounds. All price changes that are of interest and all consumer welfare changes are statistically significant. Taking into account the confidence bounds, Galatea or Bibendum, the two smallest firms in the market are ideal candidates to receive control of the divestitures following the merger between Pripps and Carlsberg. The worst possible choice, excluding the case of no divestiture requirements, is Spendrups, the second largest firm following the merger. By either picking Galatea or Bibendum instead of Spendrups, the average post-merger price increase falls by almost 51 percent, while the loss in consumer welfare is reduced by 40 percent. These effects are economically substantial and underline that the choice of recipient firm can have an important impact on post-merger outcomes.
Table (8) boils down the large amount of information to the noteworthy patterns. Our aim is to arrive at results that are also likely to be applicable in other merger cases. We therefore abstract from the specific composition of each firm's portfolio and instead focus on the following attributes of each counterfactual scenario: the inside market share accounted for by the recipient firm, $S_r = \sum_{j \in \mathcal{F}_r} s_j$, the number of varieties in the recipient firm's holdings, J_r, and the average cross-price elasticity between the divested beers and the beers owned by the recipient, $\overline{\eta}_{DIV,r}$. These attributes can either be observed directly or are easy to compute in a demand model. Inspection of Table (8) points to that post-merger price hikes and consumer welfare losses are greater, the bigger the recipient firm in terms of market share and the larger the number of varieties in its product portfolio. Moreover, the better substitutes the divestitures are for the beers owned by the recipient, the higher are post-merger price increases and welfare losses. All of these observations are fully in line with our stylized example above. The diverted profit margins between the divestitures and the recipient firms' beers are greater if these two sets of beers are more substitutable, account for a larger share of the market and if the number of varieties controlled by the recipient firm is higher. We cannot infer a causation in this setting, because these three attributes move in lockstep for the counterfactual scenarios we consider. Apart from the fact that we only have 7 observations, this strong comovement makes it impossible to identify each attribute's effect. Our observations here, however, do carry over to the following section, where we consider a very large number of different sets of divestitures and can thereby estimate the effect of each attribute.

7.2 Varying the Set of Divested Products

We next consider how the composition of the divestitures themselves affects post-merger outcomes - for this analysis we keep Galatea as the recipient. To arrive at realistic counterfactual sets of divestitures, we require that all container sizes of a product are divested to the same recipient. It is easy to show that we cannot follow the same approach here as we do when varying the recipient of the set of divested beers. Enumerating all possible sets of divestitures and then simulating the post-merger equilibrium for each of these sets leads to a prohibitive computational burden. In the actual merger, $K_D = 12$ products are

Table 8: Recipient-Firm Attributes and Post-Merger Outcomes

Recipient	$\Delta\overline{p}$	ΔCW (mln SEK)	S_r	J_r	$\overline{\eta}_{DIV,r}$
Carlsberg-Pripps	**1.601**	**-4.829**	**.4432**	**50**	**.0308**
Spendrups	**1.062**	**-3.521**	**.1948**	**38**	**.0186**
Åbro	**.8940**	**-3.023**	**.1090**	**32**	**.0138**
Kopparberg	**.7627**	**-2.905**	**.0917**	**16**	**.0239**
Krönleins	**.7280**	**-2.746**	**.0768**	**16**	**.0194**
Galatea	**.5293**	**-2.075**	**.0056**	**15**	**.0016**
Bibendum	**.5213**	**-2.110**	**.0121**	**14**	**.0034**

Note: $\Delta\overline{p}$ is the average post-merger price change in percent, ΔCW is the change in consumer welfare, S_r is the recipient firm's share of market volume measured in liters sold, J_r is the number of varieties sold by the recipient firm and $\overline{\eta}_{DIV,r}$ is the average cross-price elasticity between the divestitures and the recipient firm's beers.

divested, while the merging parties jointly control $N_D = 44$ products pre-merger.[23] Moreover, there are $N_R = 6$ potential recipient firms. Let N_S denote the number of different sets of divestitures with K_D members each that have been sampled without replacement from the merging parties' pre-merger holdings. Then, the total number of unique sets of divestitures is given as follows.

$$N_S = \binom{N_D}{K_D} N_R^{K_D} \tag{6}$$

Setting $N_R = 1$, as in the actual merger, this gives roughly 21 billion unique sets of divestitures. This vast number illustrates that negotiating remedies is not an easy task for a competition authority if the merging parties hold large portfolios of products. This is also the case, when a realistic structural demand model is available.

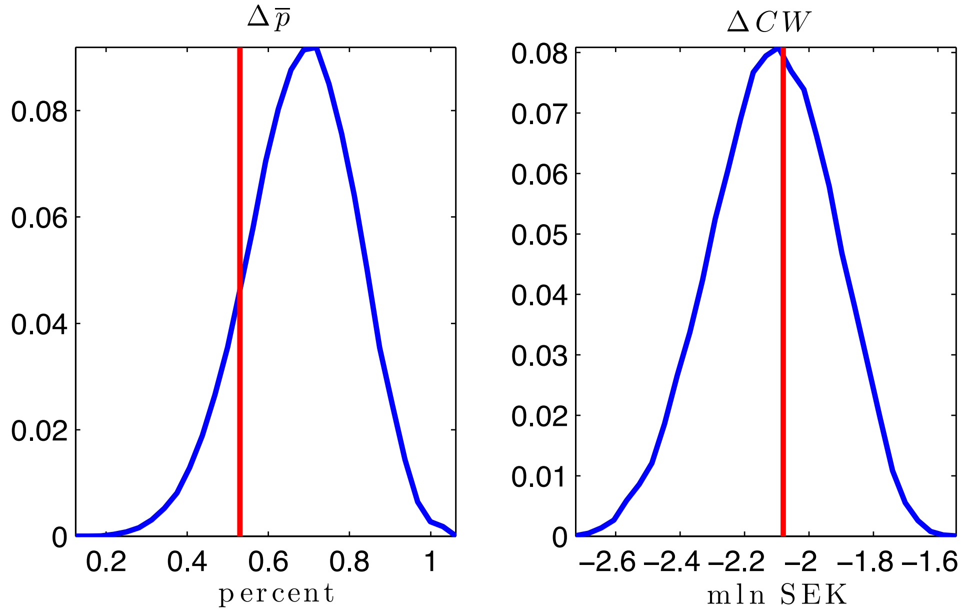
To tackle this problem, we use a simple approach that approximates the entire distribution of post-merger outcomes with divestiture requirements. We want to modify the composition of the divestitures and still be able to compare the simulated outcomes to the actual outcomes in a meaningful way. To do so, we impose a market share constraint on each counterfactual set of divestitures. Let $\mathcal{S}_i$ denote a specific set of beers to be divested. We impose the constraint that the total market share of the divested beers should lie within a symmetric 20 percent window of the market share that the actually divested beers account for: $s_{DIV,i} = \sum_{j \in \mathcal{S}_i} s_j \in [.9, 1.1] * s_{DIV,actual}$.

Instead of enumerating all possible sets of divestitures and selecting those that conform to our market share criterion, we compose the counterfactual sets by randomly selecting beers from the merging parties' pre-merger holdings subject to our market share target. The details of our heuristic approach are presented in the Appendix, where we also describe our use of the two-sample Kolmogorov-Smirnov (KS) test to infer if the sample size is sufficiently large. For each of these constructed sets, we simulate the post-merger equilibrium. If we draw a sample of $\{\mathcal{S}_i\}$ that is sufficiently large, we cover all relevant cases and can thereby approximate the entire distribution or population of potential sets of divestitures subject to our market share target. Having computed the counterfactual equilibria for each of these relevant sets, we obtain the distributions of all post-merger outcomes of interest. We focus on welfare losses and average price changes.

Figure (3) plots the post-merger distributions for consumer welfare and average price

[23] In terms of varieties, 19 beers are divested and jointly the merging parties control 69 varieties pre-merger.

Figure 3: Post-Merger Distributions of Price and Consumer Welfare Changes



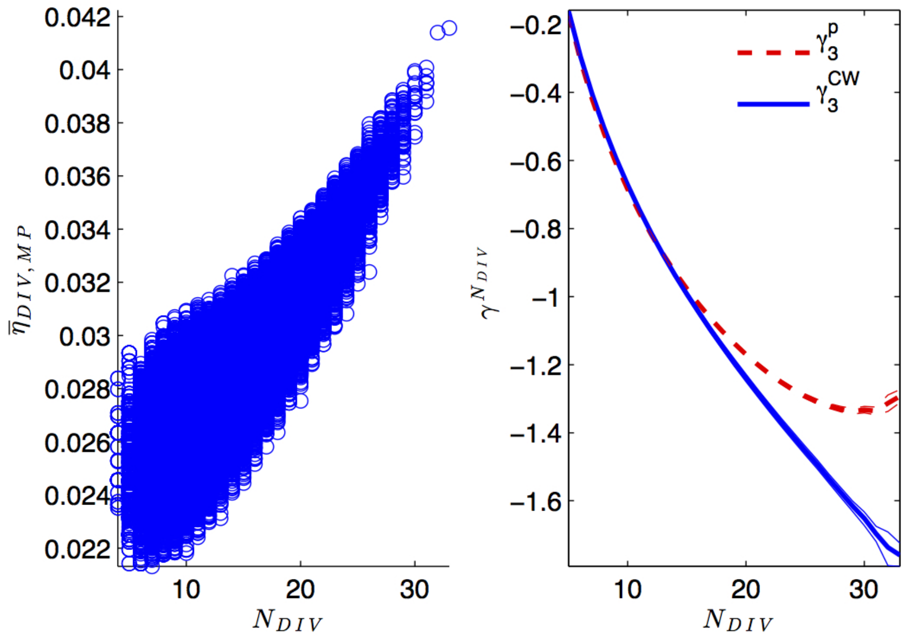
Note: Based on 200,000 simulated mergers with divestiture requirements. Each panel plots the empirical distribution for the selected outcome. The vertical line indicates the realized outcome in the actual merger.

changes that we obtain. The vertical line indicates the simulated outcome of the actual merger without efficiency gains. The large range of attainable outcomes for both price and welfare changes underlines that the composition of divested beers has an economically important impact on post-merger outcomes. Relative to the divestiture requirements of the actual merger the post-merger price increase can be reduced to only .03 percent. The loss in consumer welfare can be reduced to roughly 1.6 million SEK from more than 2.6 million SEK. Given that these effects are sizable, we next determine which attributes make the counterfactual sets of divestitures successful in dampening post-merger price increases and consumer welfare losses.

7.3 Guidelines for Structuring Divestiture Requirements

The counterfactual exercise allows us to pinpoint the set of divestitures that minimizes total consumer welfare losses. Given that we find substantial heterogeneity in consumer preferences for beer attributes, it is not surprising that there are several sets of divestitures that yield consumer welfare losses very close to this minimum, but which show limited overlap with the minimizing set. To illustrate this, we can look at the divestiture sets yielding the ten lowest welfare losses. The greatest relative difference among these sets is roughly 2.8 percent in terms of welfare loss. These cover cases from 11 to 18 divested products. This tells us that welfare losses among consumers with specific preferences can be compensated by welfare gains among other consumers with different tastes for beer attributes. Analyzing this issue in detail will not give us results that are likely to carry over to other merger cases, because it is specific to our estimates of taste heterogeneity. Instead, we focus on our findings that should prove relevant to other merger cases.
We estimate a log-log specification below to describe patterns in the simulated divestitures. Let i index the counterfactual set of divestitures and let y stand in for one of the two

Figure 4: The Elasticity of Post-Merger Outcomes w.r.t. N_{DIV} and $\overline{\eta}_{DIV,MP}$



Note: In the left panel for each of the 200,000 simulated mergers with divestiture requirements, the number of divested beer varieties is plotted against the average substitutability between the divested beers and the beers that remain in the holdings of the merged firm. The right panel plots the estimated coefficients $\gamma_3^{N_{DIV}}$, while the 99 percent confidence intervals for these coefficients are plotted using thin lines. Given the large number of observations and the clear patterns in the simulated data, the confidence intervals are difficult to distinguish from the point estimates.

outcome measures, post-merger increases in average prices, $\Delta\overline{p}$, and losses in consumer welfare, ΔCW. We relate these outcomes to the number of divested beer varieties (N_{DIV}), the market share of the divestitures (s_{DIV}) and the average cross-price elasticity between the divested beers and the merging parties' remaining holdings ($\overline{\eta}_{DIV,MP}$).

$$\ln(y_i) = constant + \gamma_1 s_{DIV} + \gamma_2 \ln(N_{DIV}) + \gamma_3^{N_{DIV}} \ln(\overline{\eta}_{DIV,MP}) + error_i \qquad (7)$$

Our measure of the divestitures' market share is already expressed as a fraction, so we include it in levels instead of in logs. To consistently estimate the elasticity of the outcome variable with respect to $\overline{\eta}_{DIV,MP}$, we allow γ_3 to vary for each observed value of N_{DIV}. The left panel of Figure (4) illustrates why this is necessary. The number of divested beers and their substitutability vis-á-vis the merging parties' remaining products strongly depend on each other. Each circle in the scatter plot represents one counterfactual divestiture scenario. Thus, for cases between 10 and 20 divested beer varieties, there are no sets that attain values for $\overline{\eta}_{DIV,MP}$ greater than .034. Similarly, for values of N_{DIV} between 20 and 30, there are no observations of $\overline{\eta}_{DIV,MP}$ below .026. The strong association between these two factors is driven by the underlying discreteness of the problem. There are only 44 products to choose from and each divestiture set has to satisfy our market share target. It is simply not possible to compose sets that fill the space in the scatter plot. This introduces a selection effect. We can still consistently estimate the coefficients by allowing γ_3 to vary with each observed number of divestitures. This conditions on N_{DIV} and uses the remaining observable variation in $\overline{\eta}_{DIV,MP}$ to identify the elasticity. Finally, there is no strong comovement between s_{DIV} and N_{DIV}.

Table 9: Estimation Results

	Dependent Variable			
	ln($\Delta\overline{p}$)		ln($-\Delta CW$)	
Coefficient	(1)	(2)	(3)	(4)
$\gamma_1 : s_{DIV}$	**-8.703**	**-7.685**	**-16.47**	**-16.99**
	(.0337)	(.0171)	(.0342)	(.0332)
$\gamma_2 : ln(N_{DIV})$	**-.3387**	**-2.945**	**.1412**	**-2.673**
	(.0007)	(.0052)	(.0007)	(.0100)
$\gamma_3 : ln(\overline{\eta}_{DIV,MP})$	**-1.254**	-	**-.6843**	-
	(.0028)		(.0028)	
$\overline{\gamma}_3^{N_{DIV}}$	-	**-1.013**	-	**-1.126**
$min(\gamma_3^{N_{DIV}})$	-	**-1.338**	-	**-1.761**
		(.0032)		(.0139)
$max(\gamma_3^{N_{DIV}})$	-	**-.1676**	-	**-.1615**
		(.0006)		(.0011)
Constant	**-3.412**	**4.591**	**-.9599**	**5.592**
	(.0113)	(.0083)	(.0114)	(.0162)
R^2	**.93**	**.98**	**.57**	**.60**

Note: Based on 200,000 simulated mergers with divestiture requirements. Standard errors are reported in parentheses.

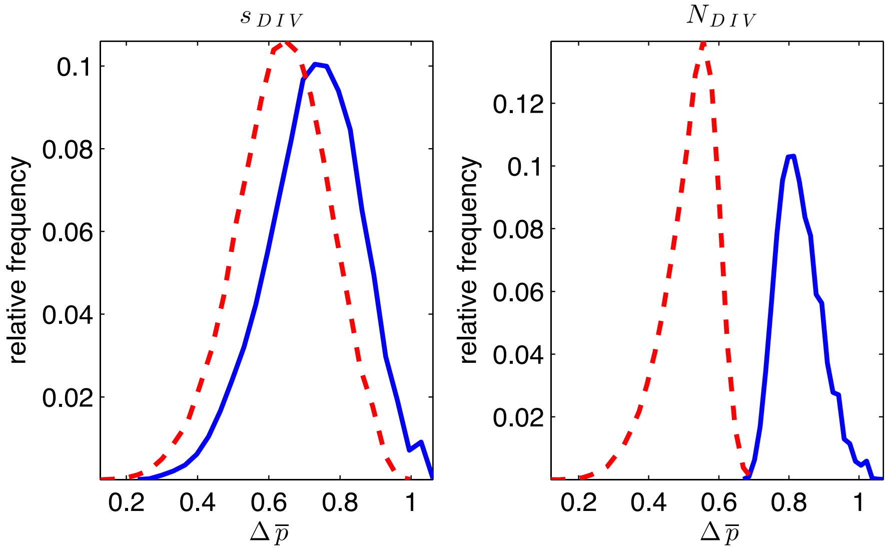
Table (9) presents our results in detail. We focus on columns (2) and (4), which present the consistent estimates. γ_1 gives us the semi-elasticity of the average price change and the consumer welfare loss with respect to these shares. In case of welfare loss, the unit of measurement for γ_1 is millions of SEK. Raising s_{DIV} has a large negative and highly statistically significant impact on both $\Delta\overline{p}$ and $-\Delta CW$. Raising the number of varieties by 1 percent reduces the average post-merger price increase by almost 3 percent and lowers the aggregate loss of consumer welfare by almost 2.7 percent. The share of liters sold that the counterfactual sets of divestitures account for lies between 6 and 7.4 percent. The right panel of Figure (4) plots the point estimates of $\gamma_3^{N_{DIV}}$ along with 99 percent confidence bounds for the price and welfare regressions. Each point estimate is highly statistically significant and shows that raising the substitutability between the divested beers and the remaining holdings of the merging parties dampens both post-merger price hikes and consumer welfare losses.

All of these effects are economically important and reiterate our findings thus far: divestiture requirements prove effective in reducing both price increases and welfare losses following the consummation of a merger. Summing up we find that *ceteris paribus* post-merger increases in average prices, $\Delta\overline{p}$, and losses in consumer welfare, ΔCW, are lower

- if the market share of the divestitures, s_{DIV}, is higher, and
- if the number of divested beer varieties, N_{DIV}, is higher,
- if the average cross-price elasticity between the divested beers and the merging parties' remaining holdings, $\overline{\eta}_{DIV,MP}$, is higher.

Qualitatively these results are hardly surprising but a comparison of their relative effects and the algorithm used to simulate the effects may be of interest to competition authori-

Figure 5: The Elasticity of Post-Merger Outcomes w.r.t. s_{DIV} and N_{DIV}



Note: In the left-hand plot, the dashed line shows the relative frequency plots of $\Delta\overline{p}$ for the hypothetical divestiture sets that lie above the 75^{th} percentile of the distribution of s_{DIV}. The solid line shows the same plot for all the divestiture sets that lie below the 25^{th} percentile. The plot on the right-hand side redraws the figure using N_{DIV}, instead.

ties. The estimated elasticities in Table (9) are informative for a marginal change in the dependent variables but let us use Figure (5) to illustrate the overall impact on the average post-merger price increases. Let s^{25}_{DIV} and s^{75}_{DIV} denote the 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of the distribution of s_{DIV}. The left panel shows the relative frequency histogram for $\Delta\overline{p}$ for all divestiture sets i for which $s_{DIV,i} \leq s^{25}_{DIV}$ (solid line) and for which $s_{DIV,i} \geq s^{75}_{DIV}$ (dashed line). We can see that the histograms are close, which illustrates that varying the market share of the divested products over the examined range has a minor impact on post-merger pricing. The right panel shows the same plot for the divestiture sets i that satisfy $N_{DIV,i} \leq N^{25}_{DIV}$ (solid line) and $N_{DIV,i} \geq N^{75}_{DIV}$ (dashed line). This plot tells a very different story. In fact, the histograms have almost no overlap and a larger number of divestitures yields substantially lower post-merger price increases. Due to the strong association between N_{DIV} and $\overline{\eta}_{DIV,MP}$, repeating the exercise for the latter yields a very similar figure.

The fact that N_{DIV} has the greatest impact on both outcome measures has a simple arithmetic reason. If the merging parties control N varieties pre-merger, there are $N(N-1)$ diverted profit margins that yield market power. If one of the varieties is divested, this number falls to $(N-1)(N-2)$ or for N_{DIV} divested varieties it falls to $(N-N_{DIV})(N-N_{DIV}-1)$. This gives a higher than linear rate at which the number of diverted profit margins that remain in the merging parties' holdings falls with N_{DIV}. At the extreme, where the merging parties keep only one product post-merger, the last divested beer still eliminates two diverted profit margins. As can be seen in Table (A.1) the Swedish beer market is populated by multiproduct firms with a minimum number of 10 products and 13 varieties in their holdings. This explains why we find N_{DIV} to be so important in our setting. In markets where firms' portfolios of products are smaller, it is likely the case that the two remaining attributes become relatively more important.

Unless diverted profit margins depend to a large extent on a single product, competition authorities should therefore resist the temptation to require the divestiture of the bestselling product or few large products. This becomes more relevant, the larger the number of products offered by the merging parties. Divesting a larger number of varieties with smaller market shares will then tend to be more effective in reducing post-merger price increases and welfare losses.

8 Conclusion

In the Carlsberg and Pripps case the two largest firms in a concentrated market merge and the divestitures required by the competition authority are substantial, accounting for roughly 6 percent of market volume. In addition institutional factors allow access to barcode level data on wholesale and consumer prices for the entire market.
We first explore how divestitures affect prices using a UPP framework along the lines of Farrell and Shapiro (2010). One prediction is that prices of divested products fall and that divestitures have a disciplining effect on the price increases that we expect on products controlled by the merging parties throughout. These effects are consistent with what we observe in a difference-in-difference analysis. Our simulations, building on a BLP demand model, show that the effects of divestitures on prices are substantial.
Varying the recipient and the set of divested brands, we find that the recipient and the number of divested products are important aspects for a competition authority that wants to use divestitures to limit the price increase associated with a merger. Overall, we believe that the evidence in this paper stresses the role of merger simulations in guiding what divestitures competition authorities should pursue. The implicit focus in much discussion of the pros and cons of merger simulations is that they should be used to determine which mergers to block and which to allow (see for instance Angrist and Pischke (2010) and Einav and Levin (2010)). It has also been noted that even carefully executed merger simulations are difficult to bring into the court-room. We suggest that the greatest value for merger simulations may lie in the ability to evaluate different divestiture arrangements. For such purposes, merger simulation can form the basis for a discussion with the merging parties and explaining the intricacies of demand estimation to lawyers is less of an issue.

9 References

Angrist, Joshua D., and Jörn Steffen Pischke, (2010), "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics," Journal of Economic Perspectives, Vol. 24, No. 2, pp. 3-30.

Ashenfelter, Orley C. and Daniel S. Hosken, (2010), "The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies," Journal of Law and Economics, 53(3), pp. 417-466.

Ashenfelter, Orley C., Daniel S. Hosken, and Mathew Weinberg, (2013), "Efficiencies Brewed: Pricing and Consolidation in the U.S. Beer Industry," NBER Working Paper No. 19353.

Baker, Jonathan B. and Timothy F. Bresnahan, (1985), "The Gains from Merger or Collusion in Product-Differentiated Industries," Journal of Industrial Economics, 33(4), p. 427-444.

Berry, Steven, James Levinsohn and Ariel Pakes, (1995), "Automobile Prices in Market Equilibrium," Econometrica, 63(4), pp. 841-890.

Björnerstedt, Jonas and Frank Verboven, (2013), "Does Merger Simulation Work? Evidence from the Swedish Analgesics Market," Working Paper, KU Leuven, University of Leuven.

Blonigen, Bruce A., Cristopher R. Knittel and Anson Soderbery, (2013), "Keeping it Fresh: Strategic Product Redesigns and Welfare," NBER Working Paper No. 18997.

Bresnahan, Timothy F., Scott Stern and Manuel Trajtenberg, (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," RAND Journal of Economics, pp. 17-44.

Carlton, Dennis, (2009), "Why We Need to Measure the Effect of Merger Policy and How to Do It," CPI Journal, Competition Policy International, 5(1).

Compte, Olivier, Frédéric Jenny and Patrick Rey (2002), "Capacity constraints, Mergers and Collusion," European Economic Review pp. 1-29.

Conlon, Christopher T. and Julie Holland Mortimer, (2010), "Effects of Product Availability: Experimental Evidence," NBER Working Paper 16506.

Conlon, Christopher T., (2013), "The Empirical Likelihood MPEC Approach to Demand Estimation," mimeo, Columbia University.

Cosnita, Andreea and Jean-Philippe Tropeano, (2009), "Negotiating Remedies: Revealing the Merger Efficiency Gains," International Journal of Industrial Organization, 27(2), pp. 188-196.

Deneckere, Raymond and Carl Davidson, (1985), "Incentives to Form Coalitions With Bertrand Competition," RAND Journal of Economics, Vol. 16, No. 4, pp. 473-486.

DG Competition, (2005), "Merger Remedies Study," European Commission.

Dubé, Jean-Pierre, Jeremy T. Fox, and Che-Lin Su, (2012), "Improving the Numerical Performance Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation," Econometrica, 80(5), pp. 2231-2267.

Einav, Liran, and Levin, Jonathan, (2010), “Empirical Industrial Organization: A Progress Report,” Journal of Economic Perspectives, 24(2), pp. 145-62.

Elzinga, Kenneth G., (1969), “The Antimerger Law: Pyrrhic Victories?,” Journal of Law and Economics, 12(1), pp. 43-78.

European Commission (2008), ”Notice on Remedies Acceptable under Council Regulation (EC) No 139/2004 and under Commission Regulation (EC) No 802/2004”.

Farrell, Joseph and Carl Shapiro, (2010), “Antitrust Evaluation of Horizontal Mergers: An Economic Alternative to Market Definition,” The B.E. Journal of Theoretical Economics, 10(1), DOI: 10.2202/1935-1704.1563.

Federal Trade Commission (1999), “A Study of the Commission’s Divestiture Process,” Staff Report, Federal Trade Commission.

Focarelli, Dario and Fabio Panetta, (2003), “Are Mergers Beneficial to Consumers? Evidence from the Market for Bank Deposits,” American Economic Review, 93(4), p. 1152-1172.

Goettler, Ronald L. and Brett R Gordon, (2011), “Does AMD Spur Intel to Innovate More?,” Journal of Political Economy 119(6), 1141-1200.

Goldberg, Pinelopi K. and Rebecca Hellerstein, (2013), “A Structural Approach to Identifying the Sources of Local Currency Price Stability,” Review of Economic Studies, 80(1), pp. 175 - 210.

Hastings, Justine S., (2004), “Vertical Relationships and Competition in Retail Gasoline Markets: Empirical Evidence from Contract Changes in Southern California,” American Economic Review, 94(1), pp. 317-328.

Hausman, Jerry, Gregory Leonard, and J. Douglas Zona, (1994), “Competitive Analysis with Differenciated Products,” Annales D’Economie Et De Statistique, 34, pp. 160-180.

Hellerstein, Rebecca, (2008), “Who Bears the Cost of a Change in the Exchange Rate? Pass-Through Accounting for the Case of Beer,” Journal of International Economics, 76(1), pp. 14-32

Hess, Stephane, John Polak and Kenneth Train, (2006), “Use of a Modified Latin Hypercube Sampling (MLHS) Method in the Estimation of a Mixed Logit Model for Vehicle Choice”, Transportation Research Part B, Vol. 40, No. 2, pp. 147-167.

Knittel, Christopher R. and Konstantinos Metaxoglou, (2014), “Estimation of Random-Coefficient Demand Models: Two Empiricists’ Perspective,” Review of Economics and

Statistics, 96(1), pp. 3459.

Konkurrensverket (2000) “Beslut: Fretagskoncentration Dnr 615/2000,” (In Swedish, Swedish Competition Authority, Decision: Concentration Decision nr 615/2000).

Motta, Massimo, (2004), “Competition policy. Theory and Practice,” Cambridge: Cambridge University Press.

Nevo, Aviv, (2000a), “Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry,” RAND Journal of Economics, Vo. 31, No. 3, pp. 395-421.

Nevo, Aviv, (2000b), “A Practitioner’s Guide to Estimation of Random-Coefficients Logit Demand Models,” Journal of Economics and Management Strategy, pp. 513-548.

Nevo, Aviv, (2001), “Measuring Market Power in the Ready-to-Eat Cereal Industry,” Econometrica, Vo. 69, No. 2, pp. 307-342.

Peters, Craig, (2006), “Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry,” Journal of Law and Economics, 49(2), p. 627-649.

Petrin, Amil, (2002), “Quantifying the Benefits of New Products: The Case of the Minivan,” Journal of Political Economy, 110(4), pp. 705-829.

Pham, Vivienne and David Prentice, (2013), “A Random Coefficients Logit Analysis of the Counterfactual: A Merger and Divestiture in the Australian Cigarette Industry,” Working Paper No. 2013.04 LaTrobe University.

Pinkse, Joris and Margaret E. Slade, (2004), “Mergers, Brand Competition, and the Price of a Pint,” European Economic Review, 48(3), pp. 617-643.

Rojas, Christian, (2008), “Price Competition in U.S. Brewing,” Journal of Industrial Economics, 56(1), issue 1, pp. 1-31.

Slade, Margaret E., (2004), “Market Power and Joint Dominance in U.K. Brewing,” Journal of Industrial Economics 52(1), 133-163.

Tenn, Steven and John M. Yun, (2011), “The Success of Divestitures in Merger Enforcement: Evidence from the J&J Pfizer Transaction,” International Journal of Industrial Organization, 29(2), pp. 273-282.

Vasconselos, Helder, (2010), “Efficiency Gains and Structural Remedies in Merger Control,” Journal of Industrial Economics, 58(4), pp. 742-766.

Weinberg, Matthew C., (2011), "More Evidence on the Performance of Merger Simulations," American Economic Review, Papers and Proceedings, 101(3), pp. 51-55.

Weinberg, Matthew C., and Daniel Hosken (2013), "Evidence on the Accuracy of Merger Simulations," Review of Economics and Statistics, 95(5), pp. 1584-1600.

10 Appendix

10.1 Firms in the Sample

As we are interested in the impact of the merger of the two biggest beer producers in Sweden, we include the eight largest firms in our sample and leave out smaller firms such as micro breweries. Table (A.1) presents each firm's average share of the volume of liters sold for the pre- and post-merger periods. It also shows how many beer varieties and beer products each firm offers on average during the entire sample period. Recall, that we define a variety as a specific package of a specific beer, whereas a product can be available in at most four different packages. The substantial impact of the divestitures on Galatea's market share is apparent. We also note that each firm offers more than 10 products and at last 13 varieties during our sample. The Swedish beer market is therefore made up of multiproduct firms with large product portfolios. We also report each firms share of inside sales for the full sample containing all firms in brackets. You can see that the shares for the restricted and full sample are very close. On average we drop less than five percent of total sales volume by only including the eight largest firms in our sample.

Table A.1: The Firms in the Sample

	Number of Varieties	Number of Products	Share of Liters Sold (1996 - 2000)	Share of Liters Sold (2001 - 2002)
Carlsberg	**43.5**	**29.0**	**.298** (.285)	**.417** (.402)
Pripps	**20.0***	**11.2***	**.251** (.240)	-
Spendrups	**40.6**	**22.1**	**.228** (.218)	**.202** (.194)
Åbro	**31.3**	**19.3**	**.098** (.093)	**.136** (.131)
Kopparberg	**15.3**	**10.2**	**.057** (.054)	**.063** (.061)
Krönleins	**15.8**	**10.1**	**.051** (.049)	**.069** (.067)
Bibendum	**13.8**	**10.1**	**.015** (.014)	**.015** (.014)
Galatea	**18.2**	**15.3**	**.003** (.003)	**.092** (.089)

Note: Unless otherwise indicated all figures are averages over the period from January 1996 up to and including December 2002. The bracketed figures reported for the inside market shares (share of liters sold) are the firms' inside share if we keep all the firms in our sample, instead of retaining only the eight largest firms. * The number of varieties and products for Pripps refers to the pre-merger period from January 1996 to December 2000.

10.2 Pricing Trends Pre- and Post-Merger

We explicitly test for statistically significant differences in pre-merger trends. Table (A.2) presents the results for two windows. Column (2) corresponds to the pre-merger window of our preferred difference-in-difference specification. We can see that except for the divestitures, all pre-merger trends are statistically significant but close to zero. Taken at face value this would be cause for concern. Column (1), however, demonstrates that this pattern is not stable. Lengthening the window slightly to November 1999, changes the results. The trend for Pripps becomes insignificant and the trend for Carlsberg flips sign. Similarly, the statistical significance of the trends for Galatea and the divestitures weakens. This instability also occurs when lengthening the window further. Different trends pop in and out of statistical significance. As another robustness check, we have repeated the regression in Column (1) of Table (3) using the pre-merger window that is lengthened to November 1999. The estimated coefficients barely move. This underlines our finding that even if some of the trends are statistically significant, their quantitative impact is small and does thereby not substantially bias the difference-in-difference estimates reported in Table (3).

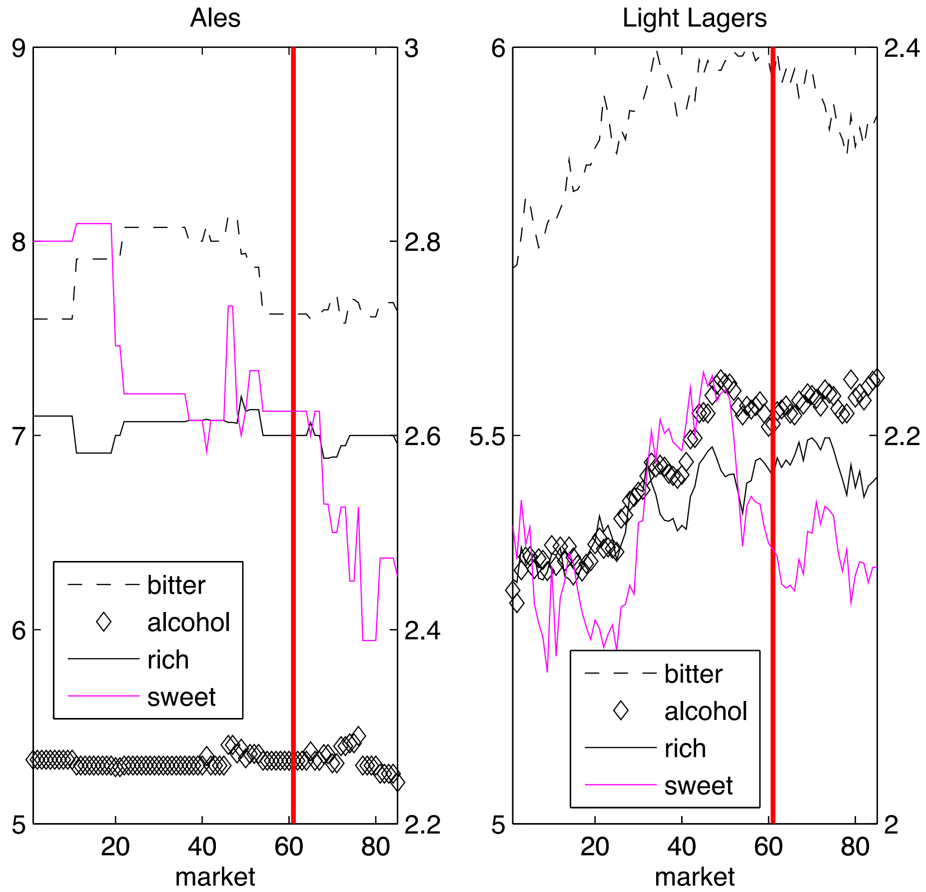
Table A.2: Testing for the Significance of Pre-Merger Pricing Trends

	Window (1) 11/1999 - 11/2000	Window (2) 01/2000 - 11/2000
Carlsberg * trend	**-.0007**	**.0007**
	(.0003)	(.0001)
Pripps * trend	**-.0001**	**.0007**
	(.0003)	(.0001)
Divestitures * trend	**-.0011**	**-.0003**
	(.0003)	(.0001)
Galatea * trend	**.0009**	**.0011**
	(.0003)	(.0001)
Constant	**1.020**	**1.190**
	(.0038)	(.0068)
Observations	2,609	2,206
R^2	.991	.996

10.3 The Observable Characteristics over Time

As we are assuming that the observable characteristics are exogenous, we also check whether there seem to be important trends in these beer attributes. We focus here on the taste information richness, bitterness and sweetness as well as the alcohol content. There are systematic differences in taste characteristics between the five beer types: ales, dark and light lagers, stouts and wheat beers. Figure (A.1) plots the evolution of the observable characteristics within the ale and light lager types. These two types account on average for almost 91 percent of all observations. We can see in both plots that overall the characteristics are quite stable over time. For ales, sweetness is trending downwards over the sample period, but stays within a rather tight range of 2.4 to 2.8. Recall that the scale for all taste characteristics ranges from 1 to 12. Similarly for light lagers the alcohol content is trending slightly upwards. Over the seven years in our sample, it increases by less than 6 percent. The vertical line in both plots indicates the month immediately preceding the consummation of the merger. The merger itself does not seem to affect the evolution of the characteristics in a substantial manner.

Figure A.1: Ale and Light Lager Characteristics



Note: The plots show the average value of the characteristics within the selected beer type for each of the 85 markets. Ales and light lagers together account for roughly 91 percent of all observations. Richness, bitterness and alcohol content in percent of volume are plotted against the left-hand y-axes. Sweetness is plotted agains the right-hand y-axis. The vertical line indicates the pre-merger month, January 2001.

10.4 Estimation Details

Our estimation algorithm follows the nested fixed point approach (NFP). The NFP algorithm consists of an outer and inner loop. In the outer loop, an optimization algorithm determines the minimum of the GMM estimator's objective function.

$$Q(\xi(\sigma)) = \min_{\sigma}(\xi^T Z)W(Z^T \xi) \tag{8}$$

W is the weighting matrix and Z is the matrix of instrumental variables. The GMM objective function can be expressed solely in terms of the standard deviations of the random coefficients, σ, because the means of the random coefficients are a function of σ (see Nevo (2000b)). For each candidate value of σ, the inner loop solves for the mean utilities, δ, that equate the observed market shares, S, with those implied by the model, $s(\delta;\sigma)$, by iterating over the following contraction mapping.

$$\delta_t^{h+1} = \delta_t^h + \ln(S_t) - \ln(s_t(\delta_t^h;\sigma)),\ t = 1, ..., T,\ h = 0, ..., H \tag{9}$$

In our case, mean utilities are given as follows.

$$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_f + \xi_t + \iota_{j\in I}\xi_{b_j} + \xi_{jt} \tag{10}$$

H is the number of iterations required to attain convergence of successive iterates within the chosen tolerance level of the inner loop, ϵ_{tol}.

$$||\delta_t^{h+1} - \delta_t^h|| \leq \epsilon_{tol} \tag{11}$$

Dubé et al. (2012) and Knittel and Metaxoglou (2014) demonstrate potential pitfalls in the NFP approach. We take lessons from both papers in the setup of our estimation algorithm.

Dubé et al. (2012) show that a loose tolerance for the inner loop can cause the optimization in the outer loop to fail to converge or converge to a point that is not a minimum. We address this issue by setting a tight tolerance level of $\epsilon_{tol} = 10^{-14}$.[24] In a similar vein, Knittel and Metaxoglou (2014) emphasize that different initial guesses for σ can converge to different candidate minima. We therefore use 100 randomly generated starting values[25] for σ to arrive at a set of candidate minima. Following the recommendations in McCullough and Vinod (2003) and Dubé et al. (2012), we verify that each of the candidate minima is indeed a local minimum. We discard a candidate minimum if the Hessian is not positive definite or if the Hessian is ill-conditioned. Given the set of candidate local minima, we adopt the estimate with the lowest value as the global minimum and base all our computations on this value for σ.

Finally, to simulate the heterogeneous tastes for product characteristics, we use 350 random draws that we generate using modified latin hypercube sampling. Hess et al. (2006) find this method to perform somewhat better than Halton sampling.

10.5 First-Stage Regression of the Instruments on Price

In total, we use 12 excluded instruments. 11 of these are computed using the method of BLP and Bresnahan (1997). The remaining instrument, z_{12}, is the log of the exchange rate. More than 42 percent of all observations cover beer varieties that are brewed abroad. It is therefore intuitive to see that exchange rate changes are quite strongly correlated with prices. To avoid that our results could partly be driven by weak instruments, we require that each of the instruments attains a correlation with price that is at least .15 in magnitude. Moreover, we also require that the matrix of instruments is well conditioned, which simply means that we drop nearly collinear instruments. In the right-most column of Table (A.3), we report each instrument's correlation with price. These range from .19 to .54 in magnitude.

The excluded instruments jointly explain 49 percent of the observed variation in prices and are jointly highly statistically significant as the F-statistic indicates. Moreover, 11 of the 12 excluded instruments are individually statistically significant at the 99 percent confidence level. The first-stage regression therefore demonstrates that our excluded instruments qualify as relevant. We also repeat this regression with the full instrumental variable matrix. 78 percent of the observable variation in price is explained. The included instruments do not drive out the excluded ones, as 9 excluded instruments are still

[24] We actually set $\epsilon_{tol} = 10^{-16}$, but we lose roughly two digits of precision by applying the exponential function to the contraction mapping, (9). This speeds up the contraction considerably, however (see Nevo 2000b).

[25] We use modified latin hypercube sampling to generate the different starting values.

Table A.3: First-Stage Regression of the Instruments on Price

	Dependent Variable		
Instrument	Price (SEK per Liter)	Price (SEK per Liter)	$\rho_{z_i,p}$
z_1	$-.0012^*$	.0006	$-.38$
z_2	$.0054^*$	$-.0293^*$	$-.22$
z_3	$-.0031^*$	.0006	$-.51$
z_4	$-.0262^*$	$.0092^*$	.22
z_5	$-.0077^*$	$-.0092^*$	$-.41$
z_6	$-.0011^*$	$-.0008^*$	$-.48$
z_7	$-.0214^*$	$-.0320^*$	.28
z_8	$-.0170^*$	$-.0171^*$	$-.19$
z_9	$.0018^*$	$.0052^*$	$-.42$
z_{10}	$-.0026^*$	$-.0004^*$	$-.54$
z_{11}	$.0146^*$	$-.0135^*$	$-.34$
z_{12}	.0049	.0049	$-.43$
Constant	4.610^*	.0050	-
Included Instruments	No	Yes	
R^2	.49	.78	
F-Stat.	1,322	355.8	

Note: Based on 16,867 observations. Coefficients with an asterisk are statistically significant at the 99 percent confidence level. $\rho_{z_i,p}$ is the correlation coefficient between the i^{th} column of the excluded instrument matrix and prices.

statistically significant at the 99 percent level. Again, the F-statistic indicates that the instruments are jointly highly statistically significant. We note that the full instrumental variable matrix consists of 169 columns, 143 of which are fixed effects. The latter can only help explain the level of prices, but not changes over time. This could explain the drop in the F-statistic.

10.6 Approximating the Distribution of Post-Merger Outcomes With Divestiture Requirements

Our approach for approximating the distributions of post-merger outcomes with divestiture requirements consists of two steps. First, we construct eligible counterfactual sets of divestitures by randomly selecting from the merging parties' pre-merger holdings. We implement the following algorithm to construct counterfactual set i. Let $\mathcal{S}_D$ denote the set of candidate divestitures. Initially this set is identical to the pre-merger beer holdings of the merging parties. Let $s_{DIV,actual}$ denote the market share that the actually observed divestitures account for. This is $s_{DIV,actual} = .0668$ of the inside market volume. Thus, we consider divestitures that account for between 6 and 7.4 percent of the total liters of beer sold.

1. Draw beer j from S_D and add it to the candidate set S_i.
2. If $\sum_{j \in \mathcal{S}_i} s_j < .9 * s_{DIV,actual}$, remove j from $\mathcal{S}_D$ and return to Step 1.

Table A.4: Outcomes of the Two-Sample Kolmogorov-Smirnov Test

Subsample Size	$\Delta\overline{p}$ # Rejections	Δp_j Average # of Rejections
1,000	**10**	**6.45**
10,000	**5**	**3.65**
50,000	**4**	**4**
100,000	**2**	**4.8**

Note: For each subsample size 100 random splits are used to implement the KS-test. Thus, the number of rejections is out of 100. For the variety-level price change, Δp_j, we compute the average number of rejections for each of the 100 randomly generated splits.

3. If $\sum_{j \in \mathcal{S}_i} s_j > 1.1 * s_{actual}$, discard draw j from S_i and return to Step 1. Here S_D is not reduced by an element.
4. If $\sum_{j \in \mathcal{S}_i} s_j \in [.9, 1.1] * s_{actual}$, the counterfactual set of divestitures $\mathcal{S}_i$ is complete.

Given that we consider a wide range of market shares, we split the 20 percent window into 4 subintervals and draw an equal number of counterfactual sets in each subinterval. This prevents us from drawing a sample that is skewed towards the low end of the market share target range. The above algorithm is then simply applied to each subinterval.

The second step of the approach infers whether the number of counterfactual sets that we draw is sufficient to approximate the whole distribution of post-merger outcomes with divestiture requirements that conform to our market share constraint. For a given number of counterfactual sets of divestitures, we randomly split this sample into two equally sized subsamples. We then use the two-sample Kolmogorov-Smirnov test to assess if the two subsamples are in fact generated by the same empirical distribution function. If the KS-tests fail to reject this null hypothesis, we conclude that our sample of counterfactual divestiture sets is sufficiently large. If the KS-test rejects the null, it must be the case that one of the samples contains information that the other does not. We then return to Step 1 and raise the number of counterfactual sets of divestitures. This procedure is continued until Step 2 is passed. To address the possibility that we get a lucky random split of the sample that gives us a KS-test result that fails to reject the null hypothesis, while the alternative hypothesis is true, we repeat Step 2 for each given number of counterfactual sets 100 times. If the number of rejections is below 5, we pass Step 2.

Table (A.4) illustrates this procedure for several numbers of counterfactual sets of divestitures. With only 2,000 sets, which gives us a subsample size of 1,000, the KS-test rejects the null in 10 cases, while this number drops to 5 if we raise the number of counterfactual sets to 10,000. As should be the case, the number of rejections drops with the number of counterfactual sets. With a subsample size of 100,000 sets, the null hypothesis is rejected only twice out of a 100 random splits of the sample. In the main text, we combine both subsamples and use the full 200,000 sets to analyze the effect of the divestitures' composition. Finally, the right-most column shows the results of the KS-test procedure if we test the equality of the empirical distribution functions at the level of individual beer varieties. For each of the 100 random splits of the sample, we therefore conduct 200 KS-tests, because there are 200 varieties available in January 2001. The average number of rejections over the 100 splits tends to fall with the number of counterfactual sets, but this pattern is somewhat less clear cut than for the average post-merger price increase in the market. Some of the tails of the variety-level price change distributions are volatile, which produces this pattern. We find, however, that this does not impact our results reported in Table (9) in a material way. To drive the average number of rejections at the variety level close to zero, up to a million counterfactual sets are needed. With such a large number of counterfactual scenarios, we found a full-blown merger simulation no longer to be feasible, because of the high computational burden. For such a large number of counterfactual scenarios, we recommend switching to an approximation to merger simulation, such as UPP. As this aspect does not impact our results in a material way, we stick to the more precise merger simulation with only 200,000 counterfactual sets.

Table A.5: Change in Average Prices Relative to Blocking the Merger (in percent), No Efficiency Gains

	Recipient of the Divestitures						
	Carlsberg-Pripps	Spendrups	Åbro	Kopparberg	Krönleins	Galatea	Bibendum
Carlsberg	**3.04**	**1.78**	**1.71**	**1.67**	**1.67**	**1.57**	**1.58**
	[2.45,4.19]	[1.37,2.48]	[1.33,2.28]	[1.31,2.24]	[1.31,2.20]	[1.26,2.05]	[1.26,2.07]
Pripps	**6.46**	**5.16**	**5.08**	**5.04**	**5.04**	**4.94**	**4.95**
	[5.07,9.82]	[3.86,8.14]	[3.81,7.95]	[3.80,7.90]	[3.79,7.87]	[3.74,7.73]	[3.74,7.75]
Divestitures	**3.58**	**-.76**	**-2.13**	**-2.43**	**-2.72**	**-3.96**	**-3.88**
	[2.88,5.01]	[-1.73,-.51]	[-3.75,1.68]	[-3.96,-1.93]	[-4.66,-2.11]	[-6.19,-3.07]	[-6.04,-3.00]
Spendrups	**.32**	**1.41**	**.19**	**.18**	**.18**	**.14**	**.14**
	[.22,.64]	[1.08,2.21]	[.11,.35]	[.11,.34]	[.11,.33]	[.09,.26]	[.09,.27]
Åbro	**.19**	**.13**	**1.48**	**.11**	**.10**	**.07**	**.08**
	[.08,.34]	[.05,.23]	[1.10,2.17]	[.04,.20]	[.04,.18]	[.03,.13]	[.03,.13]
Kopparberg	**.14**	**.10**	**.09**	**1.65**	**.08**	**.05**	**.05**
	[.06,.27]	[.05,.19]	[.03,.17]	[1.19,2.47]	[.03,.15]	[.03,.09]	[.03,.10]
Krönleins	**.13**	**.09**	**.08**	**.08**	**1.60**	**.05**	**.05**
	[.00,.21]	[.01,.16]	[.00,.13]	[.00,.14]	[1.16,2.28]	[.01,.08]	[.01,.08]
Galatea	**-.04**	**-.03**	**-.03**	**-.03**	**-.02**	**1.05**	**-.01**
	[-.25,.00]	[-.21,.00]	[-.18,.00]	[-.17,.00]	[-.16,.00]	[.83,1.52]	[-.11,.10]
Bibendum	**-.03**	**-.02**	**-.02**	**-.02**	**-.02**	**-.01**	**.87**
	[-.36,.02]	[-.26,.01]	[-.22,.01]	[-.20,.01]	[-.19,.01]	[-.11,.00]	[.72,1.37]
All Beers	**1.60**	**1.06**	**.89**	**.76**	**.73**	**.53**	**.52**
	[1.26,2.36]	[.80,1.58]	[.66,1.29]	[.55,1.12]	[.53,1.04]	[.40,.77]	[.39,.77]
ΔCW (mln SEK)	**-4.83**	**-3.52**	**-3.02**	**-2.91**	**-2.75**	**-2.08**	**-2.11**
	[-6.86,-3.78]	[-5.00,-2.75]	[-4.11,-2.33]	[-4.01,-2.23]	[-3.71,-2.13]	[-2.80,-1.65]	[-2.87,-1.68]

Note: The simulations are computed using the observed market outcomes in the month prior to the consummation of the merger, January 2001. 95 percent confidence bounds on the price and welfare changes are reported in brackets. These confidence bounds are computed with a parametric bootstrap procedure. 1,000 draws from the estimated asymptotic covariance matrix are used.