Beers and Bonds
-
Essays in Structural Empirical Economics

Beers and Bonds - Essays in Structural Empirical Economics

André Romahn

Dissertation for the Degree of Doctor of Philosophy, Ph.D.
Stockholm School of Economics 2012

KEYWORDS: structural demand estimation, merger simulation, pass-through, affine yield curve model

Beers and Bonds - Essays in Structural Empirical Economics

ISBN 978-91-7258-860-8.

PRINTED IN SWEDEN BY:
Ineko, Göteborg 2012

DISTRIBUTED BY:
The Research Secretariat
Stockholm School of Economics
Box 6501, SE-113 83 Stockholm, Sweden
www.hhs.se

Preface

This report is a result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctor's thesis at SSE. The author has been entirely free to conduct and present his research in his own ways as an expression of his own ideas.

SSE is grateful for the financial support which has made it possible to fulfill this project.

Göran Lindqvist
Director of Research
Stockholm School of Economics

Acknowledgement

First and foremost, I would like to thank my supervisor Richard Friberg for his support, encouragement, patience and his dedication to this dissertation. I have learnt a lot from Richard during the last years and am especially grateful for the opportunity to coauthor two papers with him. I would also like to thank Cristian Huse for providing extensive comments on and suggestions for improving several of the papers in this thesis.

In general, I consider myself fortunate for having been a PhD student at the School. I would like to thank the faculties of the departments of economics and finance for creating such a lively and inspiring environment. I am also grateful to Tore Ellingsen, Lars Ljungqvist, Mike Burkart and Roméo Tédongap for having taken the time to discuss my research ideas. Carin Blanksvärd, Ritva Kiviharju, Anneli Sandbladh and Lilian Öberg have kept the department running smoothly and I would like to thank them for helping me many times over the years.

I have made some great friends among my fellow PhD students. I would like to thank Tobias Laun for having been a good neighbor, office mate, lunch buddy and, most importantly, for having been a great friend during the last years. Margherita "Pipita" Bottero has also been a great office companion and I thank her for bringing together our group of PhD students on many memorable occasions. I would also like to thank Johanna Wallenius for making my lunches so much more enjoyable and for suffering through the Sosta coffee ritual, even though she strictly prefers tea. I am also grateful for the financial support of the German-Swedish chamber of commerce in Stockholm.

I owe my gratitude to many other people for making the last years a lot of fun: Johan "Älskling" Almenberg, Patrick Augustin, Anna Dreber, Paul Elger, Sara Formai, Lisa Jönsson, Jan Klingelhöfer, Henrik Lundvall, Arvid Malm, Elena Mattana, Erik Mohlin, Agatha Murgoci, Taneli Mäkinen, Johanna Möllerström, Björn Ohl, Ettore Panetti, Marcus Salomonsson, Rickard Sandberg, Juanna Schrøter Joensen, Nick Sheard, Linus Siming, Ignat Stepanok, Björn Wallace, Claudia Wolff and Christoph Wollersheim.

My family has also supported and encouraged me during my time as a PhD student and I am very grateful to all of them. Last but certainly not least, I would like to thank Kerstin Gies. Her love has been a blessing to me.

Contents

Introduction

This dissertation consists of four papers in structural empirics that can be broadly categorized into two areas. The first three papers revolve around the structural estimation of demand for differentiated products and several applications thereof (Berry (1994), Berry, Levinsohn and Pakes (1995), Nevo (2000)), while the fourth paper examines the U.S. Treasury yield curve by estimating yields as linear functions of observable state variables (Ang and Piazzesi (2003), Ang et al. (2006)).

The central focus of each paper are the underlying economics. Nevertheless, all papers share a common empirical approach. Be it prices of beers in Sweden or yields of U.S. Treasury bonds, it is assumed throughout that the economic variables of interest can be modeled by imposing specific parametric functional forms. The underlying structural parameters are then consistently estimated based on the variation in available data.

Consistent estimation naturally hinges on the assumption that the assumed functional forms are correct. Another way of viewing this is that the imposed functions are flexible enough not to impose restrictive patterns on the data that ultimately lead to biased estimates of the structural parameters and thereby produce misleading conclusions regarding the underlying economics.

In principle, the danger of misspecification could therefore be avoided by adopting sufficiently flexible functional forms. This, however, typically requires the estimation of a growing number of structural parameters that determine the underlying economic relationships. As an example, we can think of the estimation of differentiated product demand. The key object of interest here is the substitution patterns between the products. That is, we are interested in what happens to the demand of good X and all its rival products, as the price of good X increases. With N products in total, we could collect the product-specific changes in demand in a vector with N entries. It is also possible, however, that the price of any other good Y changes and thereby alters the demands for the remaining varieties. Thus, in total, we are interested in N^2 price effects on product-specific demand. With few products, these effects could be estimated directly and the risk of functional misspecification could be excluded

(Goolsbee and Petrin (2004)). With 100 products, however, we are required to estimate 10,000 parameters, which rarely, if ever, is feasible. This is the *curse of dimensionality.*

Each estimation method employed in the four papers breaks this curse by imposing functions that depend on relatively few parameters and thereby tries to strike a balance between the necessity to rely on parsimonious structural frameworks and the risk of misspecification. This is a fundamental feature of empirical research in economics that makes it both interesting and challenging.

In the following, each paper is briefly summarized with a focus on the underlying question of economic interest.

1. Ex-Post Merger Review and Divestitures

Competition authorities face the problem of deciding which proposed mergers to challenge and which to approve. The central question is what happens to prices after the merger. Structural demand estimation can provide valuable guidance to policy makers, but there are still relatively few ex-post merger reviews and some studies stress that the post-merger price effects implied by structural demand estimates can stray from actually observed prices following mergers (Peters (2006) and Weinberg (2011)). This paper contributes to the existing literature by analyzing the merger between the two biggest Swedish breweries (Carlsberg and Pripps) in the beginning of 2001. A random coefficient logit model predicts the actual aggregate price changes following the merger well. Moreover, an interesting aspect specific to this case is that the Swedish competition authority approved the merger only conditional on the requirement that several of the beers owned by the merging parties are divested to a rival firm. In actual antitrust practice, both EU and U.S. competition authorities require divestitures in more than half of all merger cases. Despite this relevance for policy practice, so far, divestitures have been of little importance in ex-post merger reviews. In the merger between Carlsberg and Pripps, the required divestitures reduced the aggregate price increase following the merger by more than half. Divestitures can therefore be an important policy tool to alter the incentives of merging parties to raise prices and structural demand estimation can play an aiding role in determining which products should be divested.

2. The Pass-Through of Cost and Demand Shocks for Multi-Product Oligopolists: A Quantitative Investigation of the Swedish Beer Market

How firms pass on shocks to production cost and demand to final good prices is a central issue in economics (Feenstra (1989), Goldberg and Knetter (1997), Gopinath and Itshoki (2010), Besley and Rosen (1999), Fullerton and Metcalf (2002)). The combination of multiproduct firms and concentrated markets can substantially alter firms' pricing responses to these shocks. The presence of multiproduct firms dampens price reactions to cost shocks, while amplifying the reaction to demand shocks. Both effects stem from a multiproduct firm internalizing the substitutability between the products in its holdings. Raising the price of a product in response to an increase in production cost (or demand) decreases the sales of that product. Some of the foregone sales, however, spill over to the remaining products that are owned by the firm. The empirical relevance of this effect is demonstrated by jointly estimating a structural model of demand and costs for the Swedish beer market (as in for example Berry, Levinsohn and Pakes (1995)) and simulating counterfactual price responses to cost and demand shocks for two cases. In the first case, the actual ownership pattern is imposed, while in the second case, each product is treated as a stand-alone firm. Due to variations in the magnitude of crossprice elasticities and the level of concentration in the different market segments, the size of the dampening effect for cost shocks can range from close to around one percent to six percent. The amplification of reactions to demand shocks is substantial. The high number of beer varieties in the market, however, keeps the overall price impact of demand shocks low.

3. Crowding and Consumer Welfare in the Swedish Light Lager Market

The introduction of new varieties is a central feature of differentiated goods markets. Whether such novel products contribute to an increase in consumer welfare has been of strong interest in the existing literature (Berry and Waldfogel (1991), Petrin (2000), Goolsbee and Petrin (2004)). This paper directly deals with an undesirable feature of the functional form of logit demand. Imposing logit-type preferences on consumers ensures that consumers value goods as such and are thereby willing to pay for having more choice (as in Dixit-Stiglitz preferences, love-for-variety). In the standard logit model, however, consumers' willingness to pay for having the choice among more products drives down consumers' willingness to pay for better quality products, as the number of varieties in the market increases. Imposing this counterintuitive pattern on the data solely by functional form threatens the credibility of welfare estimates. Ackerberg and Rysman (2005) develop a modification of the logit model that allows for a decreasing willingness to pay for more choice as the number of varieties increases (product space congestion). Contrary to alternative approaches (Berry and Pakes (2007), Bajari and Benkard (2005)), the modification is easily implemented for markets with a large number of products. Following the approach of Ackerberg and Rysman (2005), it is found that consumers in the Swedish beer market are not willing to pay for a further widening of their choice set. In other words, the love-for-variety effect is absent. Nevertheless, new beer varieties are on average more attractive to consumers than those beers that exit the market. Moreover, the commonly accepted practice of determining the relevant market outside of the demand estimation can be misleading. A positive correlation between net entry of varieties and total consumer welfare can simply be spurious and does not necessarily imply that entry raises welfare. The identified beneficial effect of product entry in the Swedish beer market, however, is shown to be robust to this concern.

4. Does Treasury Bond Supply Explain the Term Structure of Treasury Yields?

The literature on yield curve estimation has been very successful in fitting observed yields with parsimonious and flexible functional forms (Nelson and Siegel (1987), Svensson (1994), Bliss (1996)). The estimation of such term structure models is often based on unobserved or latent factors such as the level, slope and curvature of the yield curve. It is hard, however, to economically interpret these extracted factors. The ability of the assumed functional forms in producing small pricing errors therefore comes at the cost of a lack of economic interpretability. Several recent papers, however, tie the unobserved factors to observable economic variables or replace them directly (Ang and Piazzesi (2003), Ang et al. (2006)). In line with this approach, the explanatory power of a yield curve model using relative measures of Treasury bond supplies along the term structure is compared to the canonical model that employs the well established yield curve factors (level, slope, curvature). The model based on the relative supply measures captures the average shape of the term structure of U.S. Treasury yields well and the pricing errors for bonds with longer maturities are relatively low. An affine yield curve model that combines the yield curve factors and the observable measures of Treasury supplies produces very low pricing errors, and especially so for long-dated bonds. This demonstrates that Treasury supply factors contain relevant bond pricing information that is not subsumed by the level, slope and curvature of the yield curve. In light of recent interventions by central banks in government bond markets ("Operation Twist"), this finding is both of economic interest by itself and potentially relevant for current policy practice.

References

Ackerberg, Daniel A. and Rysman, Marc, (2005), Unobserved product differentiation in discrete-choice models: Estimating price elasticities and welfare effects, *RAND Journal of Economics*, 36, issue 4, p. 771-88.

Ang, Andrew, Piazzesi, Monika and Wei, Min, (2006), What does the yield curve tell us about GDP growth?, *Journal of Econometrics*, 131, issue 1-2, p. 359-403.

Bajari, P. and Benkard, C. Lanier , (2005), Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach, *Journal of Political Economy*, 113, issue 6, p. 1239-76.

Berry, Steven (1994), Estimating discrete-choice models of product differentiation, *RAND Journal of Economics*, 25, issue 2, p. 242-62.

Berry, Steven, Levinsohn, James, and Pakes, Ariel, (1995), Automobile prices in market equilibrium, *Econometrica*, 63, issue 4, p. 841-90.

Berry, Steven T. and Pakes, Ariel, (2007), The pure characteristics demand model, *International Economic Review*, 48, issue 4, p. 1193-225.

Berry, Steven T. and Waldfogel, Joel, (1999), Free entry and social inefficiency in radio broadcasting, *RAND Journal of Economics*, 30, issue 3, p. 397 - 420.

Besley, Timothy J. and Rosen, Harvey S., (1999), Sales taxes and prices: An empirical analysis, *National Tax Journal*, 52, issue 2, p. 157-78.

Bliss, Robert R., (1996), Testing term structure estimation methods, *Advances in Futures and Options Research*, 9, p. 197-231.

Feenstra, Robert C., (1989), Symmetric pass-through of tariffs and exchange rates under imperfect competition: An empirical test, *Journal of International Economics*, 27, issue 1-2, p. 25-45.

Fullerton, Don and Metcalf, Gilbert E., (2002), Tax incidence, *Handbook of Public Economics*, Eds: A. Auerbach and M. Feldstein, Elsevier Science, The Netherlands.

Goldberg, Penelopi K. and Knetter, Michael M., (1997), Goods prices and exchange rates: What have we learned?, *Journal of Economic Literature*, 35, issue 3, p. 1243-72.

Goolsbee, A. and Petrin, Amil, (2004), The consumer gains from direct broadcast satellites and the competition with cable TV, *Econometrica*, 72, issue 2, p. 351-81.

Gopinath, Gita and Itskhoki, Oleg, (2010), Frequency of price adjustment and pass-through, *Quarterly Journal of Economics*, 125, issue 2, p. 675-727.

Nelson, Charles R. and Siegel, Andrew F., (1987), Parsimonious modeling of yield curves, *The Journal of Business*, 60, issue 4, p. 473-89.

Nevo, Aviv, (2000), Mergers with differentiated products: The case of the ready-to-eat cereal industry, *RAND Journal of Economics*, 31, p. 395-421.

Peters, Craig, (2006), Evaluating the performance of merger simulation: Evidence from the U.S. airline industry, *Journal of Law and Economics*, 49, p. 627-49.

Petrin, A., (2000), Quantifying the benefits of new products: The case of the minivan, *Journal of Political Economy*, 100, issue 4, p. 705-29.

Svensson, Lars E. O., (1994), Estimating and interpreting forward rates: Sweden 1992 - 4, *NBER Working Paper* #4871, National Bureau of Economic Research, Inc.

Weinberg, Matthew C., (2011), More evidence on the performance of merger simulations, *American Economic Review*, 101, issue 3, p. 51-55.

PAPER 1

Ex-Post Merger Review and Divestitures

Richard Friberg
André Romahn

Abstract

We provide an ex-post analysis of the 2001 merger between the two largest brewers on the Swedish beer market. Difference-in-difference estimates suggest low price effects of the merger. This is well matched by a merger simulation, using a random coefficient logit model, which predicts price increases of only 0.4 percent. Knowledge of the retailers markup rules allows us to discard retailer behavior as an explanation for the pricing patterns. We further establish that without the divestitures required by the competition authorities, the price increase would have more than doubled to 1 percent (even though still low in absolute terms).

JEL Classification: K21, L11, L41, L66.
Keywords: Merger simulation, ex-post merger review, demand for beer.

We are grateful to the Swedish Competition Authority, Handelsbankens Forskningsstiftelser and the Jacob Wallenberg foundation for financial support. Thank you to Anna Wikström (Systembolaget) and Lars Näsström (Research International) for providing the data. Valuable comments have been provided by seminar audiences at Uppsala University (national conference) and the Swedish Competition Authority.

1. Introduction

Competition authorities need to take a stand on which proposed mergers they should challenge and which mergers they should allow to proceed. Whether they are making the right choices is clearly an important issue, and many researchers have called for ex-post reviews of mergers[1]. Indeed, in recent years a number of ex-post evaluations of mergers have appeared; see for instance Focarelli and Panetta (2003), Hastings (2004), Taylor and Hosken (2007) and Ashenfelter and Hosken (2011)). A typical finding is that there are modest price increases of a few percent associated with the merger. In a related line of research, a handful of papers have compared the predicted price changes from merger simulations with the actual development of prices after the merger (see Peters (2006) or Weinberg (2011)). These papers point to substantial challenges of merger simulations, as they are commonly applied, in quantitatively matching price developments after the mergers.

In the present paper we aim to contribute to the literatures on merger simulations and ex-post merger reviews. We examine the takeover of Swedish brewery Pripps by the Danish brewer Carlsberg. Carlsberg's pre-merger market share on the Swedish beer market was 33 percent and Pripps' was 17 percent. Despite this substantial concentration there are no important price hikes following the merger. We follow Berry, Levinsohn and Pakes (1995, BLP) and Nevo (2000) and use a random coefficients logit model to estimate demand in this market. We use the model to do counterfactual simulations aimed at understanding the effect of the merger on prices. Our contributions are twofold: Firstly, the merger simulation does well in predicting the low price increases following the merger. If the merging parties were to control all the beers they controlled pre-merger, we predict a quantity weighted price increase for the average beer in the market of 1 percent. While the random coefficient logit model has become the method of choice in published work (see for instance BLP (1995), Nevo (2000) or Petrin (2002)) the evaluations by Peters (2006) and Weinberg (2009) use demand systems that imply

[1] See for instance Whinston (2006). In their critical view of Industrial Organization, Angrist and Pischke (2010, p. 20) highlight the lack of ex-post merger reviews. "An important question at the center of the applied industrial organization agenda is the effect of corporate mergers on prices. One might think, therefore, that studies of the causal effects of mergers on prices would form the core of a vast micro-empirical literature, the way hundreds of studies in labor economics ... But it isn't so." In a recent review, Ashenfelter, Hosken, and Weinberg (2009) found only about 20 empirical studies evaluating the price effects of consummated mergers directly." In the same issue of Journal of Economic Perspectives Einav and Levin (2010) point to the idiosyncracies of mergers and the limits of what can be learnt from ex-post reviews, see also Carlton (2009). In a presentation at the Federal Trade Commision, Benkard (2010) also points to the limits of what can be learnt, based on an evaluation of 75 merger retrospective studies.

more restrictive substitution patterns. Merger simulation, as practiced by competition authorities, frequently uses these restrictive models. Our paper thus points to the potential value of using richer demand models when simulating mergers.[2] Our second contribution is to investigate in detail the role of divestitures in the merger. As a condition for clearing the merger the Swedish Competition Authority required the merging parties to divest a number of brands. Taking account of these divestitures more than halves the predicted price increase, to 0.4 percent. When competition authorities clear a merger they often do so under the condition that the merging parties agree to changes to the structure of the merger or to change their behavior in some way (see for instance Motta (2004)). The parties may for instance agree to change the duration of contracts with suppliers or to divest certain brands. These changes to the proposed mergers are known under various names, the perhaps most common being 'remedies' or 'undertakings'. The European Union also uses the term 'commitment' for changes to the structure of a merger. Such remedies are a central part of merger practice. As an example, the European Competition Authority subjected 165 proposed mergers to deeper analysis during 1990-2011 (the Phase II procedure, no merger was blocked at previous stages). Of these 165 proposed mergers, 56 percent were allowed to proceed after commitments. In comparison, only 13 percent were prohibited and 28 percent were permitted as proposed. Similarly, of 144 mergers challenged by US competition authorities between 2003 and 2007, 64 percent were allowed after remedies had been agreed upon (Tenn and Yun (2009))[3]. Thus, remedies are a crucial part of merger policy. Despite this, they are barely mentioned in most ex-post evaluations. The divesting of brands is one prominent form of merger remedy and in our ex-post evaluation we put the divestitures at center stage. 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 on 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 though merger simulations are accurate it is difficult to bring them into the court room. For divestitures they 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.

Several features of our data, and of the institutional setting, make the Pripps-Carlsberg merger an appealing case for examining the role of divestitures, also for

[2] On a related note, even richer models examine mergers without assuming, as we do, static forms of competition (see for instance Benkard, Bodoh-Creed and Lazarev (2010)).

[3] 51 percent were settled via consent orders and 13 percent were restructured after the Department of Justice communicated its concerns to the merging firms.

those that do not have a strong interest in the Swedish beer market per se. We use package-level data on prices and quantities for the whole market, rather than an estimate based on a selection of stores, as is frequently the case with scanner data. This limits the amount of noise. The data source is the Swedish government-owned retailer, Systembolaget. It has a legal monopoly on retail sales of all alcoholic beverages, including beer (with an alcohol content above 3.5 percent by volume). An additional benefit of the institutional setting is that Systembolaget applies the same transparent markup to all products. We can thus back out the producer price that the profit maximizing producers and importers charge Systembolaget. Gaining access to prices at different stages of the vertical chain is notoriously difficult and many other prospective merger studies, such as Nevo (2000), have been forced to assume that retailer markups are unchanged. We were also attracted by the beer market having been a prominent testing ground for merger simulations right from the start of this literature; see Baker and Bresnahan (1985), Hausman et al. (1994) for examinations of prospective mergers on the US beer market and Pinkse and Slade (2004) for mergers on the UK beer market[4].

In the next Section we describe the data and institutional setting before we turn to a description and introductory analysis of the Pripps-Carlsberg merger. Section 3 presents the model used to estimate the demand side, while section 4 presents the model of the supply side. In the following section, we detail our estimation procedure and the results. Section 6 simulates counterfactual prices for the merger with and without divestitures. Finally, section 7 concludes.

2. Data and Institutional Setting

The data set includes the monthly retail sales of all beers sold in Sweden with January 1996 as the first month and December 2004 as the last month. In the demand estimation that we use for simulating the merger we use data up to November 2000 (the merger was cleared in December 2000, we return to a detailed description of the merger below). Sales volume per month and price are observed at the bar-code level and we use the term a *beer* to denote a product at this level. Samuel Adams Boston Lager in a 33 centiliter (cl) bottle is an example of a beer. In their catalogs Systembolaget classify beers into different categories and we use beers in the light lager, dark lager and ale

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

segments[5]. We include all such beers that are sold in the standard sizes (33 cl and 50 cl bottles and 33 cl, 45 cl or 50 cl cans)[6]. Systembolaget also reports alcohol content by volume as well as measures of bitterness, richness and sweetness for each beer. These latter three measures are reported on a scale from 1-12. All of the above are used as measures of observable product characteristics in the ensuing estimation of demand.

The price for a given beer is the same across all of Sweden and prices can change only when there is a new catalog issued by Systembolaget. There are no temporary sales. The prices that consumers pay are determined by a deterministic markup on the price set by producers or importers. The ingredients in this markup on the producer price are the following: an excise tax on alcohol that is calculated per liter, a percentage markup added by Systembolaget that is common to all products, a markup per container that is also common to all products, value added tax (VAT) of 25 percent and, in the case of some beers, a deposit on the container. There have been a handful of changes to these variables as well as in how the markup is calculated over the period of study, in Section 4 we describe the calculation of the markup in detail. The percentage markup that Systembolaget adds is determined by Swedish parliament and is changed twice during the period covered by the data, in January 2000 and in April 2004. We use the information on these building blocks of the markup to back out the time series of producer prices.

By *brand* we define beer sold under the same name but in different package sizes, or with different alcoholic strengths. We also use monthly advertising expenditure which is observed at the brand level.[7] We sometimes also use the term umbrella brand to denote the case where a number of beers with different characteristics are sold under the same name; Budweiser, Guinness and Warsteiner are examples of *umbrella brands*.

2.1. The Supply of Beer and the Carlsberg-Pripps Merger. The light and dark lager segments are dominated by domestic beers, whereas imports have a high share in the ale segment. A useful way to describe the suppliers to Systembolaget are as follows; major brewers, microbreweries and pure importers. The major brewers are Åbro, Carlsberg, Krönleins, Pripps, Spendrups, and in later periods Kopparberg.

[5] The remaining segments (stout, weiss beer and 'specialty beers', for instance Belgian fruit beers) represent a miniscule share of volume and estimating demand for these marginal beers on the Swedish market would complicate estimations with little apparent benefit.

[6] Other sizes represent a diminutive share of volume.

[7] 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 the observed advertising. Before March 15, 2003 advertising of alcoholic beverages was not legal in Sweden. Beer with an alcohol content below 2.25 percent could be advertised also before this however. In cases where such a low alcohol beer with the same brand name was advertised we use this as a measure of advertising for strong beer.

While Carlsberg is based in Denmark, most of the beer they sold in Sweden at the time was produced locally. Carlsberg's local presence was partly based on having acquired Swedish brewer Falcon in 1995. 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 Kopparberg are independent Swedish brewers. Each of these brewers produces, and sells to Systembolaget, a number of 'their own' beers. They also produce some beers on license agreements with foreign brewers and act as importers and wholesalers for other beers. At the start of the period Carlsberg for instance was the wholesaler for imported beers under the umbrella 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.

In Table 1.1 we present some descriptive statistics on the market, as well as market shares and producer prices for some selected suppliers. Carlsberg, Pripps and Spendrups control roughly a quarter of the market each in the first years of the study. Until 2003 Åbro is the fourth largest supplier - its market share is rather stable around 10 percent. Krönleins' market shares is rather stable around 5 percent. The four firm concentration ratio is close to 85 percent and the Herfindahl-Hirschman index of concentration hovers around 0.2, with a low of 0.195 in the year before the merger and a high of 0.235 in the year after the merger. The highly concentrated supply side is similar to that of brewing in for instance the U.S. (see Tremblay and Tremblay, 2005). The potential for strong effects of the merger on prices is evidenced by the combined pre-merger market share of Carlsberg and Pripps being around 50 percent. Aggregate volume almost doubled over the period. One reason was a sharp drop in the excise tax on January 1, 1997, as reflected in the drop in average consumer price.

The increase in volumes was also spurred by a number of very aggressively priced beers that were introduced during the period. In particular Kopparberg has been central to this development, it advanced from a market share of 3 percent to a market share of 19 percent in 2004.[8] The increasing importance of lower priced beers is seen in the average consumer price that is falling over the period (all prices are in Swedish krona, SEK. In November 2000, 8.62 SEK equalled one Euro and 10.08 SEK equalled one US dollar). The producer price falls as well, with the exception of 1997 when the tax decrease was not fully passed through into consumer prices. Other numbers from

[8] At a fundamental level one may speculate that both developments are affected by European integration, Sweden joins the European Union in 1995 and agrees gradually to ease restrictions on cross-border shopping of alcohol; see Asplund et al. (2007) for an analysis of how crossprice elasticities of demand with respect to foreign price depend on distance to the border around this time.

TABLE 1.1. Market statistics, Swedish market for beer with alcohol content above 3.5 percent (ABV)

		Market Share by Volume (selected groups)								Prices (quantity weighted)				
Year	HHI	Carlsberg	Pripps	Spendrups	Kopparbergs	Galatea	liters sold (millions)	average consumer price	average producer price	Carlsberg	Pripps	Spendrups	Kopparbergs	Galatea
1996	0.210	0.288	0.251	0.230	0.033	0.008	85.47	37.08	9.979	10.13	9.27	9.73	7.15	15.26
1997	0.204	0.260	0.250	0.249	0.037	0.004	95.73	29.52	10.257	10.63	10.20	9.54	8.34	15.62
1998	0.201	0.259	0.256	0.233	0.047	0.003	98.62	28.68	9.667	10.35	9.30	9.44	7.90	15.24
1999	0.210	0.302	0.257	0.193	0.054	0.003	117.47	27.58	9.064	9.13	8.93	9.43	7.31	14.19
2000	0.195	0.329	0.168	0.183	0.099	0.004	130.29	26.58	7.935	8.04	8.69	8.43	5.68	12.15
2001	0.235	0.406	-	0.187	0.064	0.077	143.40	25.92	7.842	8.15	-	8.45	6.66	7.86
2002	0.235	0.396	-	0.214	0.055	0.103	153.51	25.81	7.885	8.03	-	8.38	6.71	7.29
2003	0.194	0.311	-	0.193	0.176	0.118	164.77	24.53	7.131	7.57	-	8.25	4.14	6.59
2004	0.212	0.349	-	0.164	0.191	0.124	163.84	22.24	6.466	6.14	-	8.16	4.30	6.09
mean	0.214	0.325	0.2364	0.208	0.086	0.070	129.13	27.41	8.41	8.63	9.29	8.95	6.45	9.43

Note: Prices in real Swedish krona (base year 1996).

the table that we would like to highlight is the increase in market share (and fall in average price) for Galatea in 2001, which reflects that it took over many of the brands that were divested as a condition for the Carlsberg-Pripps merger. This brings us then to the merger itself. 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 and in turn was owned by Norwegian food and drinks group Orkla. A joint entity was created under the name Carlsberg breweries where Orkla received a 40 percent share. According to reports at the time of the merger, an important motivation was that Carlsberg wanted access to Baltic Beverages Holding Co., that had a strong position in the Russian beer market, and that was owned to 50 percent by Ringnes. Carlsberg and Pripps also sold beer with alcohol content below 3.5 percent that were retailed in supermarkets, bottled water and carbonated soft drinks. By focusing on the market for beer with alcohol content above 3.5 percent in Sweden we thus examine only part of the merger. The part that we examine was viewed as a separate relevant market in the product and the geographic space by the competition authorities.

The first public information about the proposed merger came on May 31, 2000 when Carlsberg announced that it had negotiated a merger with Pripps-Ringnes. During the fall of 2000, the merger is investigated by the Swedish competition authorities and, following a number of divestitures, the Swedish competition authority announces that it will not challenge the merger on December 14, 2000. The merger is finally consummated February 15, 2001. The terms of the merger stipulate that seven domestic and five imported umbrella brands should be divested.[9] Many of the divested brands were transferred to Galatea, an independent wholesaler with no production capacity of its own. As seen in Table 1.1, Galatea expanded from a market share of less than half a percent to one of more than 7 percent following the divestiture.

The first column of Table 1.2 reports market share by volume of the beers that were divested at the time of the merger. The market share of these beers show a decline over the period, that continues after the divestment. The average, quantity weighted, price of the divested beers is somewhat lower than the average price on the market, reflecting that the bulk of the volume for the divested beers stemmed from light lagers. Average price falls slightly after the merger but increases towards the end. This reflects a declining share as well as exit of some of these low priced lagers.[10]

[9] The domestic brands that were divested were Arboga, Bayerbräu, Eagle, Fat, Sailor, Starkbock and TT. The imported brands that were divested were Bass, Caffrey's, Lapin Kulta, Staropramen and Warsteiner.

[10] For instance TT light lager dropped from 1.3 percent market share in 1999 to 0.7 in 2002, and Eagle beer, with a market share of 1 percent in 1999, was discontinued in 2003.

The most important umbrella brand that Carlsberg-Pripps was required to divest was Lapin Kulta, imported from Finland. The transfer of control over Lapin Kulta to Åbro did not take place until November 2002 however. The last column presents the volumes of the divested brands including Lapin Kulta.

TABLE 1.2. Market shares and prices of beers that were divested as a result of the Carlsberg-Pripps merger

Year	Market share by volume	Average price (quantity weighted)	Market share by volume (including Lapin Kulta that was divested in Nov. 2002)
1996	0.093	8.756	0.149
1997	0.078	9.173	0.136
1998	0.091	8.703	0.143
1999	0.073	8.777	0.110
2000	0.061	8.607	0.092
2001	0.056	8.485	0.090
2002	0.047	8.849	0.080
2003	0.035	9.187	0.061
2004	0.031	9.492	0.053
Mean	0.064	8.836	0.103

2.2. The impact of the merger on prices: a first look. The descriptive statistics presented above indicate that, despite the merger of two firms that control roughly a quarter of the market there was little effect on prices. This could be a result of a low price increase of the merger per se, or it could be reflecting cost or demand shocks which counteracted the incentives to raise price. Ex-post evaluations of mergers, such as those by Focarelli and Pannetta (2003) or Hastings (2004) have typically pursued a difference-in-difference methodology. One strives to compare development of prices for the "treated" products, with those of a control group that would be affected by the same demand and cost shocks, but unaffected by the merger. As argued by Angrist and Pischke (2010), the methodology has proved very fruitful in many areas of economics. Applying such a methodology to a merger between two major players on a national market is challenging however. One challenge regards the timing and the difficulty of defining a clear distinction between before and after "treatment". The merger was cleared in February 2001, but the firms had agreed to merge already in May 2000, possibly after long negotiations. The largest umbrella brand to be divested as a result of the merger, Lapin Kulta, was not divested until November 2002, a year and a half after the merger was consummated. Strategic behavior to try to influence the terms of the deal may have affected prices also before May 2000. 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. The other challenge regards defining a control group. Examination of price developments in neighboring countries pointed to that they were affected by other demand shocks and Sweden's floating exchange rate creates substantial noise in cross-country price comparisons. As a rough comparison for the price developments of the merger we use all beers sold by the other main domestic brewers. We thus exclude beers produced by micro breweries, beers that are imported by a purely trading firm rather than by a brewer, and beers sold by Kopparbergs, the producer that is increasing market share rapidly following a very aggressive strategy, as discussed in connection with Table 1.1. The comparison group is likely to be faced by very similar cost and demand shocks as the merging parties. However, these are also producers that are likely to be in close competition with the merging parties. The treatment may thus have an effect also on the "control" group. A standard prediction would be that prices of merging party products increase and that those of substitutes increase, but by lower amounts. A "difference-in-difference" estimation that did not pay attention to this would yield downwardly biased estimates of the price effect of the merger. Below we use regression analysis to describe price developments surrounding the merger, but we would like to stress that, for the reasons just mentioned, a causal interpretation of the results are not warranted.

In Figure 1.1 below we plot the quantity weighted (constant weights given by sales volume for the pre-merger period) average producer price for the merging parties (excluding beers that were divested as a result of the merger review), for the divested beers, and for the other traditional main brewers (Spendrups, Åbro and Krönleins). There is little evidence of dramatic price increases following the merger. Average price of beer sold by the merging firms increased some over the pre-merger period but remained largely constant around the merger. The prices of the other main brewers increased somewhat after the merger. The divested beers fell in price compared to the other two groups of beers. If divestitures are to play a disciplining role in mergers, this is a pattern that we would like to see. The fall in producer prices at the beginning of 2000 is associated with a change in the retailer markup, as described in Section 4.

In the main specifications we define the pre-merger period as the year from November 1999 up to, and including, November 2000. The post-merger period uses April

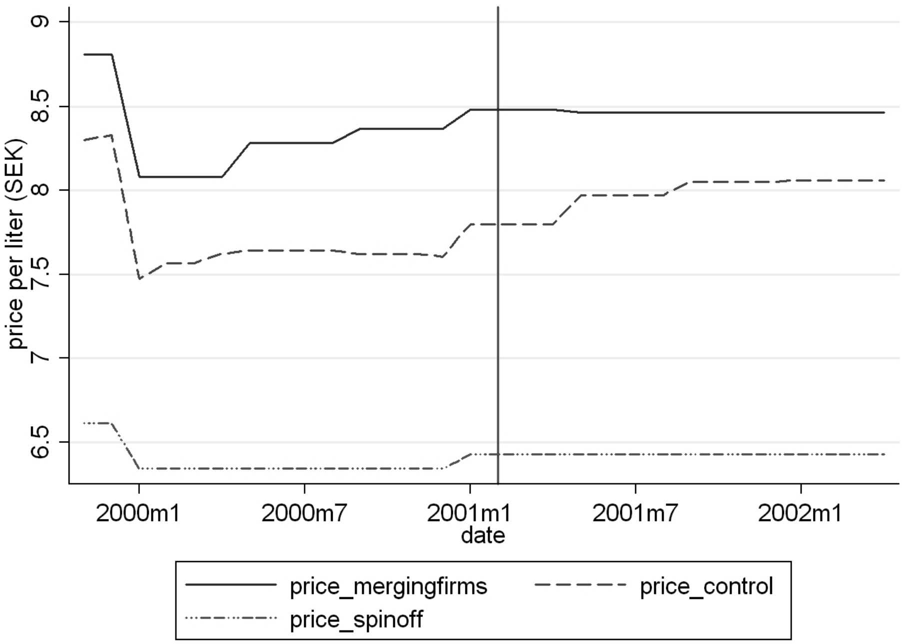


FIGURE 1.1. Quantity weighted price of beer sold by the merging firms, divested beers and a control group before and after the Carlsberg-Pripps merger in February 2001.

2001 as its first month and April 2002 as its last month. We estimate the price of beer i in month t using the following specification.

$$ln(p_{it}) = a_i + \beta_1 * postmerger_t + \beta_2 * postmerger_t * merge_i + \beta_3 * postmerger_t * divest_i + e_{it} \tag{1.1}$$

In this specification, *merge* is a dummy for beers sold by the merging parties after the merger (excluding Lapin Kulta) and *divest* is a dummy for the beers divested at the time of the merger as a condition for the merger. The variable *postmerger* is a dummy for the post-merger period, as defined above, and a_i a fixed effect per beer. We use fixed effects to capture the price of each beer.

In column (1) we use beers from all segments and find that prices of the merging parties fell relative to the comparison group around the time of the merger. Many interpretations are possible: it could for instance be due to cost savings having a rapid

Table 1.3. Difference-in-Difference Estimation

	(1) All in treatment and control	(2) Light lager	(3) Dark lager	(4) Ale
Postmerger	0.036	0.037	-0.020	0.028
	(0.010)***	(0.011)***	(0.029)	(0.013)**
Postmerger*merge	-0.053	-0.057	0.013	0.026
	(0.016)***	(0.018)***	(0.034)	(0.013)*
Postmerger*divest	-0.063	-0.068	-0.209	-0.014
	(0.017)***	(0.018)***	(0.077)***	(0.028)
Constant	2.257	2.180	2.374	2.972
	(0.004)***	(0.004)***	(0.009)***	(0.006)***
Observations	4685	4111	149	425
Number of products	288	244	12	32
Adjusted R^2	0.05	0.06	-0.04	0.07

Note: Regressions include fixed effects at the beer (bar-code) level. Standard errors are clustered by brand. The pre-merger period is November 1999 up to, and including, November 2000. The post-merger period stretches from April 2001 to April 2002. *significant at 10%, ** significant at 5%, *** significant at 1%

effect or that demand shocks affect the merging parties stronger than others. It nevertheless points to that having the merger allowed and consummated was not associated with a price increase of the merging firms relative to a comparison group of similar brewers. In columns (2) to (4) we estimate the same specification as in (1) but do so separately for the light lager, dark lager and ale segments. A comparison of columns (1) to (4) shows that the fall in prices of the merged firms is due to developments of prices of light lager beers; for dark lagers and ales the point estimate of price increases due to the merger are instead 1.3 and 2.6 percent (where only the latter is significant). We note that Åbro's introduction of a highly succesful low priced light lager in March 2001, Kung, may have contributed to these pricing patterns. Keeping all the caveats from above in mind, caveats that we believe are a concern for many merger retrospectives, we do think the regressions are useful in establishing that, despite the merger of two parties that each had a pre-merger market share close to 25 percent, there we no dramatic price increases. The regressions are also useful in establishing that the relative prices of the divested beers fell.

3. Modeling Demand

If we want to deduce the likely effects of the merger on prices, firm profits and consumer welfare before the merger actually takes place, and thereby establish a link between the magnitude of price changes and divestitures, we have to adopt a structural approach. We follow standard practice in merger simulations and estimate a discrete choice model and in particular, following BLP (1995) we allow for random coefficients in the logit estimation. Many have noted that allowing for random coefficients allows for more realistic substitution patterns than those implied by the simple to implement, but restrictive, logit form of demand (see for instance Berry (1994)).

3.1. The Random Coefficients Logit Model. We briefly present the assumptions of the model. Readers interested in a more detailed exposition are referred to Nevo (2000). We observe $t = 1, ..., T$ periods with a total number of J_t beers in each. The indirect utility of consumer i from purchasing beer j at date t is then given by

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \epsilon_{ijt},\ i = 1, ..., I,\ j = 1, ...J_t,\ t = 1, ...T \tag{1.2}$$

x_{jt} is a K-dimensional row vector of observable product characteristics. Here we include the following observables: the taste characteristics richness, sweetness and bitterness, which are all measured on a scale from 1 - 12, percent alcohol by volume, the beer category (dummy variables for ale and dark lager, light lager is the omittted category), packaging dummies with the 33 cl bottle as the base category and brand-level advertising expenditure. p_{jt} is the retail price of product j, ξ_{jt} is the product characteristic that is unobserved by the econometrician and ϵ_{ijt} is a random shock to the consumer's taste for product j. The parameters of interest are the taste coefficients, β_i, and the price sensitivity, α_i. The individual also has the option of not purchasing beer, which is referred to as the outside good and its mean utility is normalized to zero.

The variation of the parameters of interest between the I agents stems from the distribution of observable demographics. Thereby, the individual coefficients are decomposed into two parts: the means across individuals and the deviations from the means for each agent.

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \pi D_i + \Sigma v_i,\ v_i \sim P_v(v),\ D_i \sim P_D(D) \tag{1.3}$$

D_i is a $(d \times 1)$ vector of consumer i's observable demographics. π is a $(K + 1) \times d$ matrix of coefficients to be estimated, allowing the individuals' preferences over product characteristics to vary with observed demographics. v_i is a $(K+1)$-dimensional vector capturing individual taste shocks and Σ is a symmetric matrix of coefficients conformable with v_i. Σ allows for arbitrarily correlated shocks to consumer i's valuation of a product's observable characteristics. It is assumed that D_i and v_i are distributed independently.

We can now collect terms depending on whether they vary with individual demographics and rewrite (1.2) as the sum of the mean utility of consuming product j and individual i's deviation from this mean utility.

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \epsilon_{ijt}, \tag{1.4}$$

where $\mu_{ijt} = [-p_{jt},\, x_{jt}]^T[\pi D_i + \Sigma v_i]$ and $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$. The sum $\mu_{ijt} + \epsilon_{ijt}$ is the individual-specific deviation from the mean and δ_{jt} is the mean utility. It is assumed that ϵ_{ijt} is $i.i.d.$ with a Type I extreme value distribution. The probability that consumer i purchases beer j at date t is then given by

$$s_{ijt} = \frac{exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^{J_t} exp(\delta_{kt} + \mu_{ikt})}. \tag{1.5}$$

We obtain the aggregate market share for product j by integrating over all individuals. The resulting substitution patterns between products are summarized by the elasticities.

$$\eta_{jkt} = \begin{cases} \frac{-p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_D(D) dP_v(v) & ,\ j = k \\ \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) & ,\ j \neq k \end{cases} \tag{1.6}$$

For comparison, we also estimate demand with the logit model. In this approach, all individual-specific variations are set to zero. In other words, observable demographics do not play a role. The market share equation in the logit model is thereby a specialized version of (1.5) with $\mu_{ijt} = 0$.

4. Modeling Supply

To model the supply side of the Swedish beer market, we impose Nash-Bertrand competition between firms. The relevant firms here are the producers/wholesalers that act as suppliers to Systembolaget. There are $f = 1, ..., F_t$ firms present at date t and each firm maximizes profits for its portfolio of products, $\mathcal{F}_f$.

$$\max_{p_{jt}^w} \Pi_f = \sum_{j \in \mathcal{F}_f} M_t s_{jt}(p_{jt}^r)(p_{jt}^w - mc_{jt}) - C_f \tag{1.7}$$

p_{jt}^r and p_{jt}^w are the retail and wholesale prices. 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 the market shares of all products and the elasticities are functions of the prices charged to consumers, while firm margins directly depend on the prices charged to the retail monopoly. As mentioned in section 2, an attractive feature of the Swedish beer market is the deterministic relationship between the retail and wholesale prices of all beers. The retail monopoly, Systembolaget, applies a fixed[11] formula when determining retail prices.

$$p^r_{jt} = (p^w_{jt} + x^a_{jt}\tau^a_t)(1 + \tau^c_t)(1 + mk^s_t) + d_{jt} \tag{1.8}$$

x^a_{jt} is the per liter alcohol content of beer j, 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. As emphasized by the indexing, the tax rates and the retail markup are equal for all products. When setting wholesale prices, firms therefore have certainty about the price charged to consumers.

Knowledge of (1.8) allows us to precisely back out firm margins, $p^w_{jt} - mc_{jt}$. We define $\kappa_t \equiv \partial p^r_{jt}/\partial p^w_{jt} = (1+\tau^c_t)(1+mk^s_t)$. The first-order profit maximization condition for product $j \in \mathcal{F}_f$ is given by

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

Switching to matrix notation, we collect the profit maximization conditions for all firms in the market and define the matrix of market share derivatives, Ω.

$$\Omega(j,k) \equiv \begin{cases} \frac{\partial s_{kt}}{\partial p^r_{jt}} & , \ \exists\{j,k\} \subset \mathcal{F}_f \\ 0 & , \ otherwise \end{cases}$$

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

Ω_t takes into account the actual ownership pattern of beers at date t. We thereby impose a Nash-Bertrand equilibrium, where firms are the players. To illustrate, setting all off-diagonal elements of Ω_t to zero, would define the products as the relevant players and model firms as ignoring the crossprice elasticities between the individual beers in their holdings. By inverting Ω_t, we can now solve explicitly for firms' price-cost margins, given the assumption of Nash-Bertrand competition between firms.

$$p^w_t - mc_t = -\Omega_t^{-1}s(p^r_t)\kappa_t^{-1} \tag{1.9}$$

[11] The formula changes in December 1999. Before this it is given by $p^r_{jt} = (p^w_{jt}+x^a_{jt}\tau^a_t)(1+mk^s_t)(1+\tau^c_t)+d_{jt}$. 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, per container, charge applied to all beers. From January 2000 to April 2004, is c_t 1.5 SEK, and from then on it falls to 0.85 SEK. For backing out the marginal costs implied by our demand estimates, we are interested in $\partial p^r_{jt}/\partial p^w_{jt}$. It is straightforward to verify that this equals $(1+\tau^c_t)(1+mk^s_t)$ for all the pricing functions.

Wholesale prices, market shares and κ_t are observed directly, while the elements of Ω_t are functions of the estimated demand parameters. Marginal costs can be backed out by rearranging and using the demand estimates.

$$\widehat{mc}_t = p_t^w + \widehat{\Omega}_t^{-1} s(p_t^r)\kappa_t^{-1} \tag{1.10}$$

We can allow for different types of strategic conduct by adapting Ω_t. By treating each product as a stand-alone firm, we can back out product margins that are solely due to differentiation, while Nash-Bertrand conduct between firms captures both the product differentiation effect and the additional market power stemming from firms having control over several brands.

5. Estimation

The retail monopoly keeps the prices of all beers and the number of beers on offer fixed in between issues of product catalogs. Ackerberg and Rysman (2004) caution that if we included such periods we would attempt to identify price elasticities without actually observing price changes or the entry and exit of products. We therefore estimate demand using observations only from the periods when prices are permitted to change.

When defining the total market, we keep in mind that Swedish beer sales vary substantially over the seasons with peaks in summer and winter. Defining the market as a fixed number of liters per capita would therefore yield substitution to the outside good that is driven by seasons, and not by prices. Since wine sales follow a similar seasonal cycle, and is the other good available in the Systembolaget stores, we define the total market as the total number of liters sold of both beer and wine in the outlets of Systembolaget.

To estimate the parameters in (1.3), we follow the algorithm of BLP. Before describing the specifics of our estimation, however, we have to address the endogeneity of prices.

5.1. Instruments. ξ_{jt} is unobserved by the econometrician and is typically positively correlated with the price of product j. As the unobserved product attribute increases, consumers' valuation of the beer rises and so does their willingness to pay. The producer of beer j observes ξ_{jt} and incorporates this into the pricing of the product. The resulting positive correlation between prices and the error term biases the estimate of α downwards.

Since we lack comprehensive firm-specific information on cost shifters, and because Systembolaget's price setting does not allow for any regional variation in observed prices, our set of potential instruments is limited to those assuming the location of beers in characteristics space to be exogenous. More specifically, we use the instruments proposed in BLP (1995) for each type of beer. The excluded instruments for each beer's price are obtained by summing the characteristics of all beers of the same type, i.e. ales, dark and light lagers, belonging to the same firm and by summing the characteristics of all beers of the same type belonging to all other firms. We also include the number of beers of the same category held by each firm and the number for beers owned by all other firms. For a motivation of this approach, see for instance BLP (1995) or Nevo (2000).

5.2. Estimation Algorithm. With the level of aggregation of our market data, we do not observe the purchasing decisions of individuals directly. Therefore, we estimate the parameters in (1.3) by drawing 120 observations[12] from the empirical distribution of Swedish total household income (age twenty and above). As we do not have information about the distribution of consumer tastes, we set $\Sigma = 0$[13]. We estimate the vector of parameters $\theta = [\alpha,\ \beta,\ \pi]'$ by efficient GMM and split the problem into a linear and nonlinear part, as in Nevo (2000). Let θ_1 denote the parameters entering linearly and let θ_2 denote the remaining nonlinear parameters.

At each iteration k of the algorithm, we use the Berry (1994) inversion to obtain the vector of mean utilities, δ^k that matches the aggregated simulated market shares, s_{jt}, with their observed counterparts, S_{jt}. BLP show that δ^k can be solved for with a contraction mapping that is guaranteed to converge.

$$s(\delta^k;\theta) = S$$

As Nevo (2000) shows, given δ^k, we can obtain the sample estimate of the unobserved product characteristic, ω, by using a linear instrumental variables estimator.

$$\omega(\theta) = \delta^k - X_1\theta_1$$

X_1 contains the observable characteristics entering linearly. We then form the GMM objective function, $\omega(\theta)'ZWZ'\omega(\theta)$, where W is the weighting matrix and Z is the matrix of instrumental variables. We use the simplex method to determine the parameter values minimizing the GMM objective.

$$\widehat{\theta} = \arg\min_{\theta} \omega(\theta)'ZWZ'\omega(\theta) \tag{1.11}$$

[12] We have simulated up to 700 households, but the estimates did not change substantially.

[13] We initially assumed a multivariate normal distribution for v_i and a diagonal Σ, but found that the estimated coefficients were negligible. Since we use the Simplex method, which is "derivative-free", we chose to drop Σ, even though the inclusion might aid in smoothing the objective function. The payoff in terms of preserved degrees of freedom seemed more relevant to us.

In the first step of the GMM estimation, we assume that the errors are homoscedastic and therefore set $W = (Z'Z)^{-1}$. In the second step[14], we use the estimated errors to form the optimal weighting matrix, W^*, allowing for arbitrary correlation on the product level. Thus, $W^* = (Z'\widehat{\Omega}_c Z)^{-1}$, where we define $\widehat{\Sigma}_j = \omega_{j.}\omega_{j.}'$ and

$$\widehat{\Omega}_c = \begin{pmatrix} \widehat{\Sigma}_1 & & \dots & & 0 \\ \vdots & \ddots & & & \vdots \\ 0 & \dots & \Sigma_j & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & \dots & & \Sigma_{J_T} \end{pmatrix}.$$

5.3. Results. We first discuss the logit and instrumented logit results that we report in Table 1.4. In both specifications, the price coefficient is negative and statistically significant. The instrumented price coefficient is almost four times as large as its uninstrumented counterpart, which tells us that endogeneity of prices in our data is a substantial issue. This is mirrored in the mean ownprice elasticities at the bottom of the table. According to the logit estimation, demand is very inelastic and around sixty percent of all observations are estimated to have elasticities of magnitudes lower than one. In the instrumented specification, however, this fraction of outliers drops to zero and the average ownprice elasticity is close to four, which seems more reasonable.

Most of the other coefficients have the same signs and are of comparable magnitudes. The ale coefficient is positive in the instrumented specification and negative in the logit estimation. As these beers tend to be imported and have prices above the average market price, we view a positive coefficient as more reasonable. Finally, the Sargan test of overidentifying restrictions does not reject the orthogonality of our instruments. Taken together with the finding that the instrumented price coefficient is substantially greater than its logit counterpart, we take this as indication that the instruments are both valid and relevant.

The random coefficient logit estimates indicate that consumers become less price sensitive as their income rises. This is apparent from the positive price-income coefficient. As income increases, however, consumers also attach lower value to beer, because the coefficient on the interaction of income and a constant is large and negative. Given that the outside good is wine, this seems a reasonable outcome.

To allow demographics to matter for product characteristics, we estimate random coefficients for bitterness and alcohol, which we found to be the two taste characteristics with the biggest impact. Wealthier consumers prefer beers that are relatively

[14] We found that updating the weighting matrix repeatedly does not change the estimated parameters significantly after the first update.

TABLE 1.4. Estimation Results

Regressor	Logit	IV-Logit	RC-Logit linear coefficients	RC-Logit interaction with income
Price per Liter	-0.0540 (0.0098)	-0.2188 (0.0319)	-0.8513 (0.4158)	3.1660 (2.4842)
Richness	0.2231 (0.0542)	0.2379 (0.0616)	0.3034 (0.1087)	-
Sweetness	-0.0485 (0.0689)	-0.0568 (0.0726)	-0.1258 (0.1063)	-
Bitterness	-0.1796 (0.0460)	-0.2699 0.0551	-1.8586 (1.1231)	12.7820 (8.2925)
Alcohol as % of Vol.	0.1603 (0.0934)	1.0214 (0.2027)	1.7299 (3.9334)	-1.9440 (26.402)
Ale	-0.8566 (0.2821)	1.3680 (0.4971)	2.7836 (0.8879)	-
Dark Lager	-1.3474 (0.4156)	-0.7178 (0.4251)	-0.2575 (0.5447)	-
Can (33 cl)	0.2663 (0.2854)	1.5284 (0.4044)	2.8412 (0.6958)	-
Can (50 cl)	1.4161 (0.2402)	0.2162 (0.3528)	-0.8331 (0.6615)	-
Bottle (50 cl)	-0.2309 (0.1991)	-0.8958 (0.2674)	-1.4976 (0.4574)	-
Advertising	0.4848 (0.0558)	0.3774 (0.0634)	0.3432 (0.0969)	- -
Constant	-7.0176 (0.4411)	-5.2530 (0.7558)	77.4550 (24.9)	-497.2400 (156.51)
$\bar{\eta}_{jj}$	-1.059	-3.8324	-8.3214	
% of Obs. $\eta_{jj} > -1$	60.18 %	0 %	0.25 %	
Sargan $\sim \chi^2(11)$	-	4.2188	-	
J-Statistic $\sim \chi^2(10)$	-	-	10.3650	

Note: The estimation period covers the pre-merger period from January 1996 to November 2000. Standard errors are given in parentheses and are clustered at the product level.

bitter and dislike beers with a high alcohol content. Given the linear coefficients, exactly the opposite holds for consumers that are located at the lower end of the income distribution.

In comparison to the instrumented logit estimates, the average ownprice elasticity more than doubles. At first brush, this may seem like a high number. Note though that the institutional setting of the Swedish beer market pushes up prices along the demand curve considerably by adding sizable charges to the wholesale price. The average difference between a beer's retail and wholesale price amounts to 23 SEK per liter or roughly 65 percent of the retail price. In light of this, we believe that higher elasticities are a more convincing result for this particular market.

Table 1.5. Estimated Producer Markups

	Logit	IV-Logit		RC-Logit	
		differentiation	Nash-Bertrand	differentiation	Nash-Bertrand
min	0.1489	0.0367	0.0367	-0.0186	-0.0186
1st percentile	0.3438	0.0847	0.0848	0.0500	0.0502
5th percentile	0.5302	0.1289	0.1308	0.0657	0.0678
25th percentile	0.8452	0.2018	0.2085	0.0941	0.1023
median	1.1584	0.2712	0.2858	0.1193	0.1344
mean	1.1867	0.2757	0.2928	0.1237	0.1395
75th percentile	1.4613	0.3344	0.3605	0.1474	0.1691
95th percentile	1.9303	0.4442	0.4762	0.1939	0.2309
99th percentile	2.5788	0.6152	0.6362	0.24551	0.2828
max	9.757	2.3929	2.4069	0.7899	0.8015

Note: The table presents producer markups, $(p_t^w - mc_t^w)/p_t^w$, implied by the estimates in Table 1.4. We have used perfect knowledge of the retail monopoly's pricing rule, (1.8), to back out marginal costs at the producer level.

To further gauge if demand estimates are reasonable we can also consider the producer markups, $(p^w - mc)/p^w$. The implied markups correspond closely to the magnitudes of the estimated ownprice elasticities. As seen in Table 1.5, the logit specification yields unreasonably large markups given the market setup and many estimated markups that are greater than one. For the instrumented specifications, these outliers are negligible and markups move into more reasonable ranges.

We distinguish between the markups stemming from pure differentiation and from multi-product firm pricing, which also takes into account the additional market power that firms derive from selling a portfolio of products. A comparison between the "differentiation" and "multi-product" columns in Table 1.5 points to beers with relatively high markups being more likely to be controlled by firms with wider product portfolios. Beers with below average markups hold on average 19 brands, while above the average markup the mean portfolio size increases to 30.

For the random coefficients specification, the average markup is close to 14 percent, while beers located in the right tails of the distribution boast markups between 23 percent and 29 percent.

As we already argued for the ownprice elasticities, given the market setting, these numbers seem reasonable to us. It would be surprising to see larger markups for the majority of brands with retail prices being raised considerably by the retailer and the government's taxation scheme. Furthermore, several features of the retail environment at Systembolaget point in the direction of making demand more price sensititive. Beers

TABLE 1.6. Own- and Crossprice Elasticities for Selected Beers

	Carl.^{c}	Norr.	Falc.^{c}	Pripps^{p}	TT^{d}	Hein.	Bass^{d}	San M.	Lap.^{d}	Arbo.^{d}	Star.^{g}	Blå^{p}	Mill.^{c}
Carlsberg^{c}	**-7.338**	0.192	0.161	0.018	0.040	0.080	0.004	0.039	0.119	0.006	0.008	0.103	0.361
Norrlands Guld	0.089	**-6.220**	0.174	0.021	0.061	0.040	0.002	0.039	0.230	0.003	0.003	0.100	0.886
Falcon^{c}	0.104	0.243	**-5.871**	0.019	0.047	0.068	0.003	0.040	0.151	0.005	0.006	0.106	0.494
Pripps Blå^{p}	0.103	0.262	0.173	**-6.534**	0.050	0.064	0.003	0.041	0.162	0.005	0.006	0.106	0.547
Three Towns Fat^{d}	0.097	0.317	0.176	0.020	**-5.149**	0.051	0.002	0.040	0.197	0.004	0.004	0.104	0.713
Heineken	0.100	0.108	0.132	0.014	0.027	**-8.083**	0.006	0.033	0.068	0.008	0.011	0.090	0.171
Bass Pale Ale^{d}	0.092	0.078	0.114	0.011	0.021	0.103	**-9.407**	0.029	0.049	0.009	0.012	0.079	0.112
San Miguel	0.105	0.226	0.168	0.019	0.045	0.072	0.004	**-9.610**	0.141	0.006	0.007	0.105	0.450
Lapin Kulta^{d}	0.089	0.370	0.174	0.021	0.061	0.040	0.002	0.039	**-8.835**	0.003	0.003	0.100	0.884
Arboga 7.7^{d}	0.097	0.093	0.124	0.013	0.024	0.101	0.007	0.032	0.058	**-4.434**	0.012	0.085	0.141
Starobrno^{g}	0.091	0.075	0.111	0.011	0.020	0.104	0.007	0.029	0.047	0.009	**-6.007**	0.078	0.106
Blå Gul^{p}	0.105	0.221	0.167	0.018	0.045	0.073	0.004	0.040	0.137	0.006	0.007	**-4.305**	0.435
Millenium^{c}	0.079	0.421	0.168	0.021	0.066	0.030	0.001	0.037	0.261	0.002	0.002	0.094	**-4.729**
IV-Logit η_{jj}	-3.577	-2.379	-2.6527	-2.8282	-2.0603	-4.617	-5.7225	-6.3885	-3.3142	-2.6024	-2.4735	-2.0101	-1.8839
IV-Logit η_{jk}	0.0242	0.0521	0.0374	0.0041	0.0101	0.0193	0.0012	0.0003	0.0324	0.0016	0.0252	0.0236	0.1097

The estimated elasticities are based on market shares in November 2000, the last period before the consummation of the merger. ^{c,p} indicates beers that are owned by Carlsberg and Pripps, respectively, before and after the merger, beers that are divested to obtain merger approval are indexed by ^{d} and ^{g} indicates beers that are owned by Galatea.

are given the same shelf space in a store (rather than dominant brands paying to have larger space or end-of-aisle displays) and are organized according to price in stores and in catalogs. These factors should limit producers' ability to earn higher margins on their beers. At the time of the merger, Spendrups is the only major player that is listed on the Swedish stock market. It's operating margin for 2000 was 9.3 percent. This reflects profits from its other fields of sales as well (low alcohol beer and carbonated soft-drinks), but is indicative of that the relatively low producer markups are in the right ball-park.

Table 1.6 shows an excerpt from the estimated elasticity matrix for the last period before the merger is consummated. We have included the largest brands in terms of sales from the merging parties, Pripps and Carlsberg, Galatea (the acquirer of the divestitures) and some beers with big market shares brewed by firms not directly involved in the merger. The rows of the table are indexed by j and the columns by k. Thus, the entry in the second row and third column, for instance, shows a predicted increase in the market share of Norrlands Guld of 0.174 percent in response to a 1 percent price hike by Falcon. Examining the table in more detail shows that the strengths of our estimates lies with the deviation of the crossprice elasticities from their IV-logit counterparts. The pattern of crossprice elasticities is intuitively appealing. Products that are close in characteristics space are closer substitutes than those which are further apart. As an example, consider Millenium and Norrlands Guld. The observable product characteristics of these two beers are identical, except that Millenium has a sweetness rating of 2, while Norrlands Guld has a rating of 1. The market share of the latter is predicted to rise by almost 0.9 percent in response to a 1 percent price increase by Millenium. The other beers with crossprice elasticities of comparable magnitude are close to Norrlands Guld in product space, as well. These beers are Pripps Blå, Three Towns Fat and Lapin Kulta.

Analogously, Starorbno is also a light lager, but it has a bitterness rating of 9, while the other light lagers in Table 1.6 have a rating of 5. On a scale from 1 - 12, this sets these beers quite far apart. As a consequence, Starobrno is not a very close substitute for the other light lagers. Overall, we find that the estimated substitution patterns are plausible and form a reasonable base for simulating the outcome of the Pripps and Carlsberg merger.

6. Merger Simulation

Having backed out marginal cost estimates under multi-product Nash-Bertrand firm conduct, we can finally perform the merger simulation. To answer this question, we compare two scenarios: a counterfactual merger between the two firms, without any compulsory changes to the joint portfolio of beers, and the actual merger with the divestitures imposed by the Swedish competition authority. We use the set of brands and pre-merger ownership pattern for November 2000 as the basis for this exercise.

We take the estimates of marginal costs and unobserved product characteristics as given in our simulations. The post-merger equilibrium prices solve

$$\widetilde{p}^{w} = \widehat{mc} - \widetilde{\Omega}^{-1}(\widetilde{p}^{r})\widetilde{s}(\widetilde{p}^{r})\kappa^{-1}, \tag{1.12}$$

where the entries of $\widetilde{\Omega}$ reflect the post-merger ownership outcomes in the two scenarios. We arrive at the equilibrium prices by taking pre-merger prices as the initial guess for the solution to (1.12) and then iterate until convergence. This can be thought of as iterating over firms' best responses to price changes by all other firms, until no firm has an incentive to deviate.

Table 1.7 shows the market-share weighted relative price changes resulting from the merger in the two scenarios. A ratio above 1 implies that prices increase and below 1 that they fall. Forcing the merging parties to divest the selected beers generally lowers the overall price increases resulting from the merger of Pripps and Carlsberg. Using the RC-logit specification, prices are predicted to increase by 1 percent without divestitures, whereas they are predicted to increase by only 0.4 percent when divestitures are considered. Thus, the divestitures cut the predicted price increase by more than half. The fact that both equilibria are associated with modest price increases is due to the relatively high ownprice elasticities, which, as we have argued previously, are largely driven by the institutional setup of the Swedish beer market. Focusing on beers produced by Carlsberg after the merger, so this includes beers produced by Pripps before the merger, in row 2 we see that the predicted price increase is 2 percent without divestitures and 1.3 percent with the divesting of beers. The divested beers themselves are predicted to raise prices by 1.5 percent if they were to be kept by Carlsberg-Pripps but lower prices by about 2.4 percent if divested to Galatea.

Turning to the row showing the predicted relative price changes for Galatea products, we can deduce that Galatea was well chosen as a recipient. Recall that we are referring to the post-merger ownership structure here. Thus, the scenario with divestitures moves all the divested beers to Galatea, while the scenario without divestitures leaves Galatea's product portfolio unchanged. Comparing the relative price changes in rows three and four for the scenario with divestitures shows that Galatea derives almost no additional market power from absorbing the divested beers. If the divestitures complemented its existing portfolio of products well, the relative price change of Galatea's grown portfolio would be substantially higher than the predicted relative price change of the spinoffs alone. This is not the case, however, which strongly suggests that the

Table 1.7. Predicted Market-Share Weighted Relative Price Changes

Post-Merger Ownership	Logit		IV-Logit		RC-Logit	
	no divest.	with divest.	no divest.	with divest.	no divest.	with divest.
all beers	1.043	1.021	1.013	1.008	1.010	1.004
Carlsberg	1.086	1.062	1.029	1.026	1.020	1.013
Divestitures (Carlsberg & Pripps)	1.070	0.907	1.024	0.974	1.015	0.977
Galatea	1.000	0.920	1.000	0.977	1.000	0.980
all others	1.002	1.001	0.997	0.997	1.001	1.000

choice of Galatea as a recipient for the divested beers aided in limiting price increases resulting from the Carlsberg-Pripps merger.

In the last row, we note that prices of beers not directly involved in the merger are little affected. As seen in Table 1.6, the cross-price elasticities between brands are non-trivial. The small price increases for beers not directly affected by the merger are thus largely due to the small predicted price increases of beers sold by the the merging parties.

Let us now compare our simulated effects to the time series evidence that we considered in Section 2. While there are concerns about our ability to capture the causal effect of the merger on prices using the methods of Section 2, we note that all specifications pointed to small price effects of the Carlsberg-Pripps merger. If the merger was associated with large price increases we would have needed large drops in marginal cost or large negative demand shocks to counteract an incentive to raise price. We have not found any plausible candidate to such large shocks. Thus, we are confident that the time series points to very moderate price effects of the merger, even if we do not want to put to much faith into any one of the specifications.

The RC-logit estimates clearly match the low price effects of the merger. In comparison, the logit results, reported in the first columns of Table 1.7, point to much greater price effects. Time and computational constraints have implied that logit and nested logit have been seen as the main alternatives for merger simulations (see for instance Peters (2006) or Weinberg (2011)). However, falling computing costs are likely to make RC logit easier to implement and we take the ability of the RC logit to match the limited price increases observed in this market as encouraging. We also note that in this particular case the predicted price increases from the IV-logit are rather close to those of the RC logit. This result stresses the role of finding valid instruments in the case where one opts for using a logit estimation.

For reasons explained above we do not put much trust in the logit estimates reported in the first two columns. Even so, one might view them as a robustness check on what the price effects would be if demand were less elastic. We then note that the divestitures have an important effect on the estimated price increases also in this case - divestitures lower predicted price increases from 4.3 to 2.1 percent.

7. Conclusion

There are important limits to how much can be learnt from one single case study of a consummated merger. Let us nevertheless highlight a few findings from our study which we believe are of more general interest. Firstly, divestitures and other remedies are a crucial part of merger control. If ex-post merger reviews are used to analyze whether merger policy is effective or not, they need to be careful in how they deal with remedies. For mergers that are seen as problematic, remedies are the rule, rather than the exception. Despite this, remedies have been conspicuously absent from ex-post merger reviews. We put them center stage and show that they had an important impact on the predicted price changes of the Carlsberg-Pripps merger. Secondly, while RC-logit is seen as superior to logit and nested logit by economists, previous ex-post evaluations have typically used these more restrictive demand specifications. In our case RC-logit provides a much better match with actual price changes than logit. With falling complexity of implementing demand systems, that allow for richer substitution patterns, we should not discard merger simulations based on the criticism that too restrictive methods are being used.[15] Finally, the case study also points to that a reliance on simple measures of concentration to decide what mergers to block, can lead astray. Despite the merger of two parties, where each had close to a quarter of the market, the price effects were small. In this case, high demand elasticities served to keep the price effect of the merger muted. The high elasticities were plausibly generated by taxes pushing prices far up along the demand curve and a retail setting with an explicit aim to provide a level playing field. These institutional features are specific to our data. The benefit of merger simulations, and its simplified cousin in the form of upward pricing pressure (UPP, see for instance Farrell and Shapiro (2010)), is that it forces us to be explicit about these institutional features.

[15] For instance access to consumer level data on purchases of differentiated consumer products is now common. The tools to estimate a random coefficient logit demand system on such data are now available as a canned routine in STATA.

8. References

Angrist, Joshua D., and Pischke, Jörn-Steffen, (2010), The credibility revolution in empirical economics: How better research design Is taking the con out of econometrics, *Journal of Economic Perspectives*, 24, issue 2, p. 3-30.

Ashenfelter, Orley C. and Hosken, Daniel S., (2011), The effect of mergers on consumer prices: Evidence from five selected case studies, *forthcoming in Journal of Law and Economics.*

Ashenfelter, Orley C.,Hosken, Daniel S. and Weinberg, Matthew, (2009), Generating evidence to guide merger enforcement, *CPI Journal*, Competition Policy International, vol. 5.

Asplund, Marcus, Friberg, Richard and Wilander, Fredrik, (2007), Demand and distance: Evidence on cross-border shopping, *Journal of Public Economics*, 91, issue 1-2, p. 141-57.

Baker, Jonathan B. and Bresnahan, Timothy F., (1985), The gains from merger or collusion in product-differentiated industries, *Journal of Industrial Economics*, 33, issue 4, p. 427-44.

Benkard, Lanier, Bodoh-Creed, Aaron and Lazarev, John, (2010), Simulating the dynamic effects of horizontal mergers: U.S. airlines, *Manuscript*, Yale University.

Benkard, Lanier (2010), "Merger retrospectives, panel discussion", slides available at http://www.ftc.gov/be/workshops/microeconomics/2010/docs/benkard_slide.pdf, downloaded November 17, 2011.

Berry, Steven (1994), Estimating discrete-choice models of product differentiation, *RAND Journal of Economics*, 25, issue 2, p. 242-62.

Berry, Steven, Levinsohn, James, and Pakes, Ariel, (1995), Automobile prices in market equilibrium, *Econometrica*, 63, issue 4, p. 841-90.

Carlton, Dennis, (2009), Why we need to measure the effect of merger policy and how to do it, *CPI Journal*, Competition Policy International, vol. 5.

Einav, Liran, and Levin, Jonathan, (2010), Empirical industrial organization: A progress report, *Journal of Economic Perspectives*, 24, issue 2, p. 145-62.

Farrell, Joseph and Shapiro, Carl, (2010), Antitrust evaluation of horizontal mergers: An economic alternative to market definition, *B.E. Journal of Theoretical Economics*, 10, issue 1.

Focarelli, Dario and Panetta, Fabio, (2003), Are mergers beneficial to consumers? Evidence from the market for bank deposits, *American Economic Review*, 93, p. 1152-72.

Hastings, Justine S., (2004), Vertical relationships and competition in retail gasoline markets: Empirical evidence from contract changes in Southern California, *American Economic Review*, 94, issue 1, p. 317-28.

Hausman, Jerry, Leonard, Gregory and Zona, J. Douglas, (1994), Competitive analysis with differentiated products, *Annales D'Economie Et De Statistique*, 34, p. 160-80.

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, issue 1, p. 14-32

Motta, Massimo, (2004), Competition policy. Theory and practice, Cambridge: Cambridge University Press.

Nevo, Aviv, (2000), Mergers with differentiated products: The case of the ready-to-eat cereal industry, *RAND Journal of Economics*, 31, p. 395-421.

Pesendorfer, Martin, (2003), Horizontal mergers in the paper industry, *RAND Journal of Economics*, 34, p. 495-515.

Peters, Craig, (2006), Evaluating the performance of merger simulation: Evidence from the U.S. airline industry, *Journal of Law and Economics*, 49, p. 627-49.

Petrin, Amil, (2002), Quantifying the benefits of new products: The case of the minivan, *Journal of Political Economy*, 110, issue 4,p. 705-29.

Pinkse, Joris and Slade, Margaret E., (2004), Mergers, brand competition, and the price of a pint, *European Economic Review*, 48, p. 617-43.

Rojas, Christian, (2008), Price competition in U.S. Brewing, *Journal of Industrial Economics*, 61, issue 1, p. 1-31.

Taylor, Cristopher T. and Hosken, Daniel S., (2007), The economic effects of the Marathon-Ashland joint venture: The importance of industry supply shocks and vertical market structure, *Journal of Industrial Economics*, 60, issue 3, p. 419-51.

Tenn, Steven and Yun, John M., (2009), The success of divestitures in merger enforcement: Evidence from the J&J Pfizer transaction, Bureau of Economics, Federal Trade Commission, Working Paper #296.

Tremblay, Victor J. and Tremblay, Carol Horton, (2005), The U.S. brewing industry: Data and economic analysis, MIT Press, Cambridge (MA).

Weinberg, Matthew C., (2011), More evidence on the performance of merger simulations, *American Economic Review*, 101, issue 3, p. 51-55.

Whinston, Michael, (2006), Lectures on antitrust economics, MIT Press, Cambridge (MA).

PAPER 2

The Pass-Through of Cost and Demand Shocks for Multi-Product Oligopolists: A Quantitative Investigation of the Swedish Beer Market.

Richard Friberg
André Romahn

Abstract

We show, in a simple stylized setting, that the combination of multiproduct firms and concentrated markets yield dampened price responses to marginal cost shocks but amplified price responses to demand shocks. To quantitatively assess the empirical relevance of this effect, we estimate a structural demand model for the Swedish retail market for beer and use our model and estimates to simulate counterfactual prices in response to cost and demand shocks. We find that when modeling each product as a firm, aggregate price responses to cost (demand) shocks are higher (lower) than for the actual ownership pattern. Due to variations in the magnitude of crossprice elasticities and the level of concentration in the different market segments, the size of the effect for cost shocks can range from close to around one percent to six percent. The effect amplifies reactions to demand shocks substantially. The high number of products in the market, however, keeps the overall price impact of demand shocks low.

JEL Classification: L11, L13, L66.
Keywords: Price rigidity, multi-product firms, demand estimation.

We are grateful to the Swedish Competition Authority, Handelsbankens Forskningsstiftelser and the Jacob Wallenberg foundation for financial support. Thank you to Anna Wikström (Systembolaget) and Lars Näsström (Research International) for providing the data. Valuable comments have been provided by seminar audiences at Stockholm School of Economics.

1. Introduction

The determinants of price responses to cost and demand shocks are of fundamental importance to many questions in economics. Two prominent areas of investigation relate to how exchange rate changes affect prices of imported goods (Feenstra (1989), Goldberg and Knetter (1997), Gopinath and Itshoki (2010)) and to how tax changes feed into prices (Besley and Rosen (1999), Fullerton and Metcalf (2002)). In Industrial Organization, recent work stresses that the pass through of costs is crucial for many comparative statics results and serves as a building block for market delineation (Farrell and Shapiro (2010), Weyl and Fabinger (2011)). In Macroeconomics, the links between (lack of) price adjustment and market structure has been of interest at least since early empirical work by Means (1929), Hall and Hitch (1939) and Sweezy (1939). By now it is well established that for a monopolistic firm that receives an idiosyncratic shock to marginal costs, its price response will be greater the more convex the demand curve (Bulow and Pflederer (1983), Feenstra (1989)). For instance, a constant elasticity demand curve will feature a magnified response of prices to a marginal cost shock whereas a linear demand curve implies that prices change by half the size of the marginal cost shock. In oligopolistic competition the extent of pass-through will depend on assumptions on for instance the extent of product differentiaton and the number of firms. Pass-through under oligopoly has been explored in applications covering both exchange rates (Dornbusch (1987)) and taxes (Delipalla and Keen (1992), Reny, Wilkie and Williams (2011)).

One aspect of pass-through that remains largely unexplored however is the impact of multiproduct firms on pass-through. An exception is Hamilton (2009) who examines how product breadth and prices respond to tax changes in a multi-product extension of Salop's (1979) competition on the circle model. His results point to a more than proportional pass-through of costs occurs *except* under the case when demand is highly convex - a result at sharp odds with the results under monopoly and which points to the need for further analysis of pass-through by multiproduct firms. [1] We argue that the gap in the literature is especially important since casual observation suffices to establish that many markets are dominated by oligopolistic firms that each control a large number of products.

In the present paper we investigate how the breadth of products that a firm controls affects how prices change in response to cost and demand shocks. While the details of pass-through are likely to be sensitive to what form of competition we assume,

[1] The key mechanism in Hamilton (2009) is that higher excise taxes lead firms to narrow the range of varieties on offer, which in his setting softens price competition, creating an additional upward pressure on prices.

we note that Bertrand competition with differentiated products is seen as a workable description of many consumer goods markets (see for instance Davis and Garces (2010), Nevo (2011)). In particular we note that a rich previous literature has examined the quantitative impact on prices of a firm gaining control over more products via mergers in such a setting (Hausman et al (1994), Nevo (2000), Ivaldi and Verboven (2005)). That literature examines how prices are affected by changes in the market structure - we use similar tools to systematically explore how prices respond to cost and demand shocks. We first show in a simplified example (logit demand, disregarding competitive responses) that a firm who controls more products will pass-through an idiosyncratic marginal cost shock to a lesser extent but an idiocyncratic demand shock to a greater extent than a single product firm would. We proceed to investigate the quantitative implications of the pass-through incentives using data from the Swedish beer market.

Empirical studies of pass-through are faced with differentiated product prices being rigid and only adjusting at irregular intervals (Blinder et al (1998), Nakamura and Steinson (2008)). One way to generate such rigidity in models is by introducing a fixed cost of adjusting prices (menu costs). Rigidity of prices will be determined by the *costs* of adjusting prices on the one hand, and the *gains* from adjusting prices on the other hand (which in turn have a close relation to the desired price change, the pass-through rate). The overwhelming part of work aimed at understanding the sources of rigidity has assumed that products are supplied by a monopolist or by monopolistically competing firms that each sell one product (Barro (1972), Sheshinski and Weiss (1977), Akerlof and Yellen (1985)). However, a set of recent papers examine price adjustment with multi-product firms with a focus on the *costs* of price adjustment. They are complementary to the present paper in that they study the price setting behavior of firms and back out the nature and size of price adjustment costs that are consistent with observed price adjustments - whereas the present paper focuses on studying pass-through of multi-product firms. Midrigan (2011) considers the case where a firm sells two products and by paying a single menu cost can change both prices - he shows how these "economies of scope" in price adjustment allow for a better match of some stylized facts on price adjustment, in particular the existence of many small price changes.[2] Stella (2012) allows for one menu cost component that is common to all products sold by a firm as well as an additional menu cost that is product specific. Using data on cheeese sold by a US retail chain he establishes that both types of menu costs are present. Closely related are also two papers that examine the reasons

[2] Midrigan (2011) embeds his model in a general equilibrium model. A comparison between Midrigan (2011) and Golosov and Lucas (2007) - who use single product firms - shows that the seemingly minor addition of economies of scope in menu costs, coupled with a fat-tailed distribution of shocks, drastically raises the impact of monetary policy on real variables.

for incomplete pass-through of exchange rates and costs into prices: Nakamura and Zerom's (2010) study of coffee prices and Goldberg and Hellerstein's (2012) study of beer prices. Nakamura and Zerom (2010) assume that menu costs are random but equal in a given period for all products sold by a firm, whereas Goldberg and Hellerstein (2012) estimate bounds on the menu costs, assuming that they are product specific. [3]

Our focus on examining the links between the number of products controlled by a firm, and pass-through rates, is largely unexplored in the literature just mentioned.[4] Midrigan (2011) compares pricing of just one or two products and Bhattarai and Schoenle (2012) pricing of one, two or three products. [5] In Goldberg and Hellerstein's (2012) data there are only two cases of multiproduct firms, each of whom sells two products.

In contrast, our data set features substantial variation in the number of products controlled by firms - across both time and market segments. A number of institutional details of the Swedish beer market make it appealing for an empirically grounded examination of how pass-through rates vary with the breadth of product portfolios. State owned Systembolaget is the monopolist retailer for alcoholic beverages in Sweden. This is our source for nine years of market-encompassing monthly observations on sales, prices, taxes, retailer markup and product characteristics: all observed at the bar-code level. Beer is supplied to Systembolaget by a number of independent, profit maximizing wholesalers and producers. The retailer follows a deterministic and publicly known markup rule - we can thus focus squarely on the price setting of upstream firms without being forced to make assumptions on retailer markup rules.[6] Prices can change only a few times every year, in connection with catalogs by the retailer and there are no temporary sales. Temporary sales generate substantial noise in the data, but as economists we are typically more interested in the pass-through of regular prices (see for instance Nakamura and Steinson (2008) for an analysis of the price rigididity that pays attention to the role of temporary sales). Finally, the addition of a new data set that can be used to study price adjustment with a structural model is potentially

[3] Another related paper is Noton (2011) who examines exchange rate pass-though on European car markets using the estimation methodology of Bajari, Benkard and Levin (2007). He assumes that there is no fixed cost of adjusting a price, but instead estimates a price adjustment cost that depends on the size of the price adjustment.

[4] A number of empirical papers examine pass-through of multi-product firms (see for instance Goldberg (1995)). The multi-product aspect is not in focus however.

[5] In related theoretical work Alvarez and Lippi (2012) characterize optimal price adjustment patterns for a firm that sells an arbitrary number of goods and where paying a single fixed cost allows it to change all prices. The model is quite stylized however and the loss of not adjusting a price is assumed to be quadratic in the deviation of the actual price from the frictionless price.

[6] Recent work by for instance Gopinath et al (2012), and Goldberg and Hellerstein (2012), point to that must of the interesting price dynamics are at the upstream level, where acquiring accurate price data is notoriously hard.

valuable in itself: Data sets that feature the desired level of detail are often hard to access, and indeed a very large share of the papers above, and related work, use data from Chicago retailer Dominick's Finer Foods: (various sub-sets from this data set are used by Besanko et al (2005), Kano (2006), Burstein and Hellwig (2007), Midrigan (2011), Goldberg and Hellerstein (2012) and Stella (2012)). While using the same data set can be conducive to methodological contributions, concerns can also be raised that specificities in the behavior of one retailer is used as a foundation of stylized facts.

To explore the links between multi-product pricing and pass-through we estimate a discrete choice model of demand, following Berry (1994) and Berry, Levinsohn and Pakes (1995, BLP). The estimated parameters can be used to inform us about own- and cross-price elasticities for all products, as well as giving us an estimate of product level demand shocks in each period. We estimate the demand and supply sides jointly, which allows us to extract the comovement of demand and supply shocks from the data. This allows us to simulate price responses to demand and cost shocks for different counterfactual ownership patterns and correlations between demand and cost shocks in the industry. We compare the response of prices to cost and demand shocks under two scenarios - one with the actual multi-product firm pricing and one where the price of each product is optimized as if it were sold by a stand-alone firm. As in Goldberg and Hellerstein (2012) we assume Nash-Bertrand conduct when firms set optimal prices. For a sophisticated forward looking firm, the price that it sets may be affected by future expectations about the drift and volatility of cost and demand factors, as well as the response of competitors to such factors and the future costs of adjusting prices. To make such a problem computationally feasible we would need a number of simplifying assumptions which would partly cloud our interest in how pass-through depends on the set of products that a wholesaler/producer controls.

The market segments in the beer market (ales, dark lagers, light lagers, stouts and weissbeers) are characterized by different magnitudes of crossprice elasticities and levels of concentration. In the dark lager segment, having both the highest concentration and greatest crossprice elasticities, the aggregated (wholesale) price response of single product firms to a common shock to marginal costs is six percent higher than the response when taking the actual ownership pattern of products into account. In the ale segment, on the other hand, which is characterized by similarly sized crossprice elasticities and very low levels of concentration, the effect is only around one percent.

2. Price Setting in the Logit Model

To provide a simple exposition of links between pass-trough and shocks to costs and demand we consider a "logit", discrete choice model of demand. This demand model, and refinements thereoff, is a workhorse in empirical studies of demand (see for instance Anderson, de Palma and Thisse (1992), Berry (1994), BLP (1995), Besanko, Gupta and Jain (1998), Sudhir (2001), Nevo (2011)). Let n differentiated products be offered in a market and let the utility of a consumer from purchasing good i at date t be given by

$$U_{it} = \delta_{it} + \epsilon_{it}, \ \ i = 1, ..., n, \ \ t = 1, ..., T, \tag{2.1}$$

where $\delta_{it} = x_i - \alpha p_{it} + \xi_{it}$ is product i's mean utility and ϵ_i is a product specific taste shock. x_i summarizes product i's observable attributes and p_{it} is the product's price. ξ_{it} represents the product attributes that are observed by market participants, but unobserved by the econometrician. We can therefore think of variation in ξ_{it} as capturing temporary demand shocks.

Imposing an extreme value distribution on the taste shocks yields the familiar logit expression for product i's market share.

$$s_i = \frac{e^{\delta_i}}{1 + \sum_{j=1}^{n} e^{\delta_j}} \tag{2.2}$$

Here, we have normalized the utility of the outside good, namely not buying any of the differentiated goods on offer, to zero, $\delta_o \equiv 0$.

Suppose initially that a monopolist offers n product varieties. When maximizing profits, the monopolist internalizes the substitutability of all products on offer and sets prices accordingly. The n first-order conditions are

$$\begin{aligned}
\frac{\partial s_1}{\partial p_1} + s_1 + \sum_{j \neq 1} \frac{\partial s_j}{\partial p_1}(p_j - c_j) &= 0 \\
&\vdots \\
\frac{\partial s_i}{\partial p_i} + s_i + \sum_{j \neq i} \frac{\partial s_j}{\partial p_i}(p_j - c_j) &= 0 \\
&\vdots \\
\frac{\partial s_n}{\partial p_n} + s_n + \sum_{j \neq n} \frac{\partial s_j}{\partial p_n}(p_j - c_j) &= 0
\end{aligned}$$

We rewrite this system of equations in matrix notation by collecting the market share derivatives in the $n \times n$ matrix Ω with entries

$$\Omega_{ij} = \partial s_i / \partial p_j = \begin{cases} -\alpha s_i (1 - s_i) & , i = j \\ \alpha s_i s_j & , i \neq j \end{cases} \tag{2.3}$$

Ω captures the substitutability of products and because market shares are strictly positive, a price increase of good i reduces its own market share while raising that of all rival products. The price-cost margins are collected in the $(n \times 1)$ vector $p - c$ and the $(n \times 1)$ vector s keeps track of market shares. Optimal prices satisfy

$$p - c = -\Omega^{-1} s. \tag{2.4}$$

The functional form of the logit model can be exploited to analytically invert Ω for an arbitrary number of products, n. The adjoint of Ω is symmetric,

$$adj(\Omega) = \alpha^2 \begin{pmatrix} (\Pi_{j \neq i} s_j)(1 - \sum_i s_i) & \cdots & \Pi_i s_i \\ \vdots & \ddots & \vdots \\ \Pi_i s_i & \cdots & (\Pi_{j \neq i} s_j)(1 - \sum_i s_i) \end{pmatrix},$$

and its determinant is given by

$$|\Omega| = \alpha^3 (-\Pi_i s_i)(1 - \sum_i s_i).$$

It follows that (4) can be rewritten as

$$p - c = \left(\alpha (1 - \sum_i^n s_i) \right)^{-1} \iota, \tag{2.5}$$

where ι is the $n \times 1$ unit vector.

Thus, in the logit model, the monopolist optimally chooses one price-cost margin for all products.[7] Equipped with (2.5), we can derive the monopolist's price response to marginal cost shocks and temporary shifts in demand. Defining $f(s;\alpha) \equiv \left(\alpha(1 - \sum_i^n s_i)\right)^{-1}$, totally differentiate (2.5) and let the price, marginal cost and demand for product k change while keeping marginal costs and prices for other products fixed. We can then solve for the change in the price of product k, dp_k.

$$dp_k = \left(1 - \frac{\partial f(s;\alpha)}{\partial p_k}\right)^{-1} dc_k + \left(1 - \frac{\partial f(s;\alpha)}{\partial p_k}\right)^{-1} \frac{\partial f(s;\alpha)}{\partial \xi_k} d\xi_k$$

Using the fact that $\partial s_i / \partial \xi_j = -\partial s_i / \partial p_j$, we arrive at the following relationship between price, cost and demand changes for product k.

$$dp_k = \frac{1 - \sum_i^n s_i}{1 - \sum_{j \neq k}^n s_j} dc_k + \frac{s_k}{\alpha(1 - \sum_{j \neq k}^n s_j)} d\xi_k \tag{2.6}$$

[7] The first statement of this result that we have been able to find is in Besanko, Gupta and Jain (1998, equation (5)).

The reaction of price to cost and demand shocks depends on the coefficients $A \equiv (1 - \sum_i^n s_i)/(1 - \sum_{j \neq k}^n s_j)$ and $B \equiv s_k/(\alpha(1 - \sum_{j \neq k} s_j))$. The monopolist reacts differently to demand and cost shocks. As seen from the numerator of the first term in (2.6), the higher the market share captured by all the varieties sold by the firm, the more muted is the price response to cost shocks. The denominator in (2.6) shows that a firm's response to cost shocks is lower for products with a higher market share. This implies that firms react relatively more (less) to cost shocks in markets with a relatively low (high) level of concentration.

Faced with demand shocks, however, the firm reacts differently. The greater the market share of a specific variety, and the greater the market share captured by all products in the firm's holdings, the larger is the resulting price response caused by a temporary shift in demand. Moreover, consumer's marginal utility of income, α, affects the magnitude of price responses to demand shocks. As demand becomes perfectly elastic, $\alpha \to \infty$, the price reaction is driven towards zero. Lowering the level of concentration in the market, on the other hand, dampens the optimal price response to demand shocks.

Moreover, from (2.6) it is obvious that the correlation of demand and cost shocks is another important determinant of both the size and frequency with which price changes are observed in the data. Note that in examining the pass-through in (2.6) above we only considered the change in the first order conditions for the firm, treating all other prices as fixed. This, what Weyl and Fabinger (2011) term "short-run own" pass-through, is obviously not the full story. It nevertheless forms some guidance for what to expect in the empirical work that follows, where we compare different Nash equilibria in response to demand and cost shocks. Before presenting the structural estimation framework, we turn to a description of the Swedish beer market.

3. The Swedish Beer Market

To assess the quantitative impact of multiproduct firm pricing for the observed pattern of price changes, we use data from the Swedish beer market. The data set includes the monthly retail sales of all beers sold in Sweden with January 1996 as the first month and December 2004 as the last month. Sales are aggregated at the national level. Each product's sales volume per month and price are observed at the bar-code level and we use the term a beer to denote a product. Leffe Blonde in a 33 centiliter (cl) bottle is an example of a beer.

As mentioned in the Introduction, state-owned Systembolaget is the sole retail outlet for alcoholic beverages in Sweden. This gives us complete market coverage for beers with an alcohol content of at least 3.5 percent by volume (beers with alcohol contents below this threshold can be bought in supermarkets).

Systembolaget enforces uniform pricing for each of the beers on sale across the whole of Sweden. In regular intervals, the retail monopoly publishes a catalog that is free of charge, containing all the products that are available in its retail outlets. Prices are only allowed to change when a new catalog issue is published. In between issues, Systembolaget holds prices fixed. Moreover, there are no temporary sales during the year. Ackerberg and Rysman (2004) caution that including market observations in which prices are fixed in the demand estimation implies that we are attempting to identify price elasticities in these periods without actually observing prices changing. This can yield biased coefficients in the estimation. To take account of this fact, we exclude all periods in between catalog issues, reducing the number of months used in the sample from 108 to 50.

The prices that consumers pay, p_r, are set as a deterministic markup on the wholesale price, p_w. The components of this markup are the following: an excise tax on alcohol that is calculated per liter, τ_a, a percentage markup added by Systembolaget that is common to all products, m_s, a charge per container that is also common to all products, c, a value added tax (VAT) of 25 percent, τ_c, and a deposit charge on the container, f. Systembolaget is fully transparent when setting retail prices and applies a publicly known formula to arrive at the retail price of beer j at date t.[8]

$$p_{rjt} = (p_{wjt} + \tau_{ajt})(1 + \tau_{ct})(1 + m_{st}) + f_t, \quad t = 1, \ldots, T \tag{2.7}$$

This institutional setup allows us to abstract from jointly modeling the price setting behavior of wholesalers and retailers. Instead, wholesalers directly determine the final retail prices for their products given the mechanic markup imposed by the retail monopoly. We can thereby back out price-cost margins and elasticities at the wholesale level exactly. The wholesalers are profit maximizing firms and they can be described as being in one of two categories. The supply of beer is dominated by domestic brewers that act as wholesalers for the beers that they themselves brew, while also acting as wholesalers for some imported beers (Danish brewer Carlsberg has a strong market presence in Sweden and is included in this category). The other type of wholesaler is

[8] The formula changes in December 1999. Before this date, it is given by $p_{rjt} = (p_{wjt} + \tau_{ajt})(1 + m_{st})(1+\tau_{ct})+f_{jt}$. From January 2000 onwards, the retail price is $p_{rjt} = (p_{wjt}(1+m_{st})+c_t+\tau_{ajt})(1+\tau_{ct}) + f_{jt}$, where c_t is a constant per package charge applied to all beers. From January 2000 to April 2004, c_t is 1.5 SEK, and from then on it falls to 0.85 SEK. When we turn to firm price setting, we are interested in $\partial p_{rjt}/\partial p_{wjt}$, i.e. how the retail price changes with the wholesale price that firms charge to Systembolaget. It is straightforward to verify that it is equal to $(1+\tau_{ct})(1+m_{st})$ for all the pricing rules.

Table 2.1. Market Descriptives

	Liters Sold (in '000s)	Retail Price (SEK)	Wholesale Price (SEK)
Ale	167	49.38	23.55
Dark Lager	183	36.34	14.65
Light Lager	10,303	31.71	11.76
Stout	56	41.43	19.29
Weissbeer	19	38.63	17.29
All	10,727	35.14	14.09

Note: Liters sold and prices are monthly averages over the sample period from January 1996 to December 2004.

	Number of Products	Number of Firms	HHI
Ale	45	17	.15
Dark Lager	13	8	.75
Light Lager	232	29	.21
Stout	11	10	.44
Weissbeer	7	6	.46
All	308	33	.21

Note: The numbers of products and firms are monthly averages over the sample period from January 1996 to December 2004. The Herfindahl-Hirschman Index (HHI) is computed by using the actual number of liters sold in a given month for each type of beer and all beers, respectively.

a pure importer. Typically these are rather small firms that have imports of alcohol as their main business.

In the catalogs of Systembolaget, beers are classified into different categories. We use beers in the ale, dark lager, light lager, stout and weissbeer segments.[9] Table 2.1 presents descriptive statistics for the market as a whole and each of the beer categories separately. The top panel compares liters sold, retail and wholesale prices. Light lagers make up by far the biggest category in terms of sales, which on average account for 96 percent of total beer sales. Next in line are the dark lager and ale categories, which sell on average around 170,000 liters every month. The remaining categories of stouts and weissbeers are much smaller in terms of sales.

Apart from being the category with the highest volumes, light lagers on average also sell at the lowest price, while ales are the most expensive category in the Swedish beer market. All prices are in Swedish kronor (SEK). In January 2000 the SEK/Euro exchange rate was 8.60 and the SEK/US Dollar 8.47. The large differences between

[9] The remaining beers in the specialty and spontaneously fermented categories represent a negligible share of total market volume and there are few observations of these beers in the data. We therefore decided to drop these two categories.

wholesale and retail prices highlights the impact of the alcohol tax, Systembolaget's markup and a 25 percent VAT.

The bottom panel of the table highlights the variation in the data that is most relevant for assessing the impact of multiproduct firm price setting on the pattern of price changes. The ale and dark lager categories, for example, are similar in terms of liters sold and revenue generated, but they differ strongly in terms of concentration. With a Herfindahl-Hirschman Index (HHI) of .15 at the wholesale level, the ale category exhibits the lowest degree of concentration, while the dark lager segment is characterized by the highest level of concentration with an HHI of .75. This pattern carries over in terms of the number of products and firms in each of these categories. There are roughly three times as many products and roughly twice as many firms in the ale segment as in the dark lager category. The remaining beer types lie between those two in terms of the level of concentration measured by the HHI.

Before proceeding to the demand estimation, we describe the pattern of price changes in the data in some detail, to dispel concerns that the strictly regulated retail setting could yield patterns that are very different from the well established stylized facts in the literature on price adjustment.

3.1. Observed Price Changes. Table 2.2 compares the pattern of price changes in the Swedish retail beer market with that of the United States' and the euro area's processed food sectors, as presented in Dhyne et al. (2006) and Nakamura and Steinsson (2008). We chose this comparison, because the processed food sector seems to be the closest match for the beer market. Moreover, we include all time periods in-between catalog issues to ensure that our observed price descriptives are comparable to these two papers. As the reported price durations in Nakamura and Steinsson (2008) strip out the effects of temporary sales, they are a good comparison with the patterns we report. In terms of the median price duration, the two sets of data seem quite close. The median life span of a price in Systembolaget's outlets is a bit more than eight months, while the equivalent measure for the US processed food sector comes in at nine months. Also, the fraction of prices changing in each month is quite similar: 7.18 percent of all beers change price while 10.6 percent of all processed food items in the US adjust prices in the same time frame.

An obvious difference between the patterns reported in the US and the euro area is the net change in prices: in the Swedish beer data retail prices drop by almost two percent, while prices are reported to increase over time in the comparison data. This difference can be explained by exogenous changes in alcohol excise taxes and the markup of Systembolaget. In January 1997, the average alcohol tax per liter was lowered by almost forty percent. This translated into a fall in the average retail price

Table 2.2. Price Changes

All Price Changes					
	up	down	Δ (%)	% of all prices	duration (months)
Retail Prices					
mean	3.80%	-8.48%	-1.93%	7.18%	11.11
median	3%	-6%	1%	-	8.17
share of changes	.53	.47	-	1	-
Wholesale Prices					
mean	8.75%	-8.62%	2.25%	9.35%	9.31
median	6%	-7%	2%	-	8.11
share of changes	.61	.39	-	1	-
Dhyne et al. (2006) - Processed Food (Euro Area)					
	up	down	Δ (%)	% of all prices	duration (months)
mean	6.9 %	-8.1 %	7.5 %	13.7 %	≈ 13
share of changes	.54	.46	-	1	-
Nakamura and Steinsson (2008) - Processed Food, No Sales (USA)					
	up	down	Δ (%)	% of all prices	duration (months)
median	11.5 %	-17.6 %	13.2 %	10.6 %	9
share of changes	.72	.28	-	1	-

of nearly fifteen percent. When looking at the wholesale prices of beer, the average and median net price change are about two percent. Also, the average rise in wholesale prices is close to nine percent while the equivalent change in the retail price is only roughly four percent.

Thus, the patterns in the Swedish beer data are comparable to the stylized facts established in the price rigidity literature regarding the duration of prices and the magnitude of price changes. The observed downward trend in prices is driven by changes in taxes and retailer markups, which are both decided upon by the government and can therefore be viewed as exogenous to firms' price setting decisions.

4. Structural Estimation of Demand and Supply

The different beer segments exhibit interesting variations in terms of concentration levels. Given that product prices in more concentrated markets should react relatively less to cost shocks and relatively more to demand shocks, we want to exploit this variation to investigate the role of multiproduct firms for the magnitude of price responses to these two types of shocks. As light lagers dominate the Swedish beer market in terms of both market shares and the number of products, most of the differences between the beer categories would be swamped in a simple logit model. To allow different patterns of product substitutability within and across the beer categories, while keeping the computational burden low, we estimate a nested logit model, where each beer category is defined as a separate nest. As we are also interested in the comovement of demand and cost shocks, we want to be able to extract the correlation between these two shocks from the data. We therefore estimate the demand and supply sides of the Swedish beer market jointly. This gives us a consistent estimate of the covariance between demand and cost shocks.

4.1. Demand Side. We first describe the demand side estimation. Let the utility of consumer i from purchasing beer j at date t be given by

$$u_{ijt} = \delta_{jt} + \zeta_{ig} + (1-\sigma)\epsilon_{ijt}. \tag{2.8}$$

$\delta_{jt} = x_{jt}\beta - \alpha p_{rjt} + \xi_{jt}$ is product j's mean utility level that is determined by its observable product characteristics, x_{jt}, its retail price, p_{rjt} and its unobservable product characteristics, ξ_{jt}. As observable characteristics, we include the three taste characteristics, richness, sweetness and bitterness, which are all measured by Systembolaget on a scale from one to twelve and published in its product catalogs. The beer's alcohol content enters the observables as well as the advertising expenditure for the beer's particular brand.[10] We also include dummy variables for the different packagings.[11] Table 2.3 presents descriptive statistics for the observable product characteristics. As in Section 2, the unobservable product characteristics capture temporary shocks to demand.

ζ_{ig} is consumer i's valuation of nest g. Its distribution depends on the parameter $\sigma \in (0,1)$. As the nesting parameter tends to zero, the correlation of consumer taste shocks within nest g goes to zero; in other words, the nesting of products becomes meaningless.

Regarding the definition of the outside good, we note that Swedish beer sales exhibit a seasonal pattern with peaks in summer and winter. Wine sales follow a very similar cycle. By defining the relevant market as the sum of beer and wine sales, we avoid seasonal substitution towards the outside good.

Given (2.8), we can rewrite the market share of product j at date t as $s_{jt} = s_{j|g,t}s_{gt}$, where $s_{j|g,t}$ is beer j's share of the liters sold by the category or nest it is located in and s_{gt} is category g's share of the overall market. As is well known, we have that

$$s_{j|g,t} = \frac{e^{\delta_{jt}/(1-\sigma)}}{\sum_{j\in\mathcal{G}_g} e^{\delta_{jt}/(1-\sigma)}}, \tag{2.9}$$

[10] By brand we define beers sold under the same name but in different package sizes, or with different alcoholic strengths. A brand can therefore include several individual beers.

[11] The .33 liter bottle is the base category.

TABLE 2.3. Observable Product Characteristics

Column	Variable	Mean	[Min,Max]	Std. Dev.
X_1	Price Per Liter (SEK)	35.25	$[15.6, 183]$	10.78
X_2	Richness	5.78	$[1, 11]$	1.83
X_3	Sweetness	2.23	$[1, 11]$	1.38
X_4	Bitterness	6.34	$[1, 12]$	2.12
X_5	Alcohol (% of Vol.)	5.42	$[4, 17]$	1.12
X_6	Advertising (mln SEK)	.10	$[0, 13.85]$	.66
X_7	.5 Liter Bottle	.30	$[0, 1]$	.46
X_8	.33 Liter Can	.05	$[0, 1]$	.22
X_9	.5 Liter Can	.28	$[0, 1]$	.45
X_{10}	Constant	1	$[1, 1]$	0

and

$$s_{gt} = \frac{\left(\sum_{j\in\mathcal{G}_g} e^{\delta_{jt}/(1-\sigma)}\right)^{(1-\sigma)}}{\sum_g \left(\sum_{j\in\mathcal{G}_g} e^{\delta_{jt}/(1-\sigma)}\right)^{(1-\sigma)}}, \tag{2.10}$$

where $j \in \mathcal{G}_g$ indicates that the particular product j belongs to category g. Differentiating (2.9) and (2.10) with respect to the retail price of beer j and the retail price of a rival product k, we obtain the substitution patterns implied by our nested logit assumption.

$$\Omega_{jk} = \frac{\partial s_{jt}}{\partial p_{rkt}} = \begin{cases} -\frac{\alpha s_{jt}}{1-\sigma}\left(1 - \sigma s_{j|g,t} - (1-\sigma)s_{jt}\right), & j,k \in G_g \\ \frac{\alpha s_{kt}}{1-\sigma}\left(\sigma s_{j|g,t} - (1-\sigma)s_{jt}\right), & j,k \in G_g \\ \alpha s_{jt}s_{kt}, & j \in G_g, k \notin G_g \end{cases} \tag{2.11}$$

These market share derivatives are determined by the structural parameters of the demand side and enter in the modeling of the supply side. As in Section 2, these derivatives determine how each firm optimally sets the prices for the products in its holdings.

Berry (1994) shows, that the system of market shares, which are functions of product mean utilities, δ_{jt}, can be inverted analytically to yield the estimating equation for the demand side.

$$ln(s_{jt}/s_{ot}) = x_{jt}\beta - \alpha p_{rjt} + \sigma ln(s_{j|g,t}) + \xi_{jt},\ t = 1, \ldots, T,\ j = 1, \ldots, J_t \tag{2.12}$$

We next turn to modeling the supply side.

4.2. Supply Side. Let F denote the total number of firms active in the market and let $\mathcal{F}_f$ denote the set of all beers owned by a particular firm f. Then, profit-maximizing prices satisfy the following system of equations.

$$p_{rt} = c_t - \mathcal{H} \circ \Omega_t^{-1} s_t, \quad t = 1, \dots, T \tag{2.13}$$

$\circ$ denotes the Hadamard or pointwise product and $\mathcal{H}$ accounts for the actual ownership pattern in the market. Its entries are

$$\mathcal{H}_{ij} = \begin{cases} 1, & j, k \in \mathcal{F}_f \\ 0, & j \in \mathcal{F}_f, k \notin \mathcal{F}_f. \end{cases} \tag{2.14}$$

Recall that the entries in Ω are determined by the structural demand parameters and actual market shares. Prices are directly observed in the data as well, so that (2.13) can be used to back out each product's marginal cost, as is done in Nevo (2000), for example. We are interested in the comovement of demand and cost shocks, however. To obtain a consistent estimate of the correlation between demand and cost shocks in the Swedish beer market, we have to combine (2.13) with an assumption about marginal costs. We follow the modeling strategy of Berry, Levinsohn and Pakes (1995, BLP) and assume that analogous to the demand side, the marginal cost of beer j, c_j, is composed of directly observable factors, W_j, and an unobserved component, ω_j.

$$\ln(c_{jt}) = W_{jt}\gamma + \omega_{jt} \tag{2.15}$$

As observable cost factors, we include the taste characteristics richness, sweetness, bitterness, as well as the alcohol content, the prevailing exchange rate between the beer's brewing origin and Sweden, measured in Swedish krona (SEK) per unit of foreign currency and a constant. ω_{jt} thereby captures temporary shocks to marginal costs.

Using (2.13) to substitute for marginal costs in (2.15) yields the estimation equation for the supply side.

$$\ln(p_t + \mathcal{H} \circ \Omega_t^{-1} s_t) = W_{jt}\gamma + \omega_{jt}, \; t = 1, \dots, T, \; j = 1, \dots, J_t \tag{2.16}$$

The, joint estimation of demand and supply is then given by (2.12) and (2.16). Before presenting the results of the estimation, we discuss the instruments we use for the endogenous regressors in the demand and supply regressions, and the identification strategy.

4.3. Instruments. On the demand side, ξ_{jt} is unobserved by the econometrician. A higher value of the unobserved product attribute is associated with a higher mean utility of consuming that beer and thereby a higher willingness to pay for that particular beer. The producer of beer j observes ξ_{jt} and incorporates this into the optimal pricing of the beer. The resulting positive correlation between prices and the error term biases the estimate of the price coefficient, α, downwards.

Similarly, on the supply side, ω_{jt} is unobserved by the econometrician, but incorporated by the firm in its price-setting decisions. All else equal, a higher realization of ω_{jt} raises the price of beer j and thereby generates a dependency between prices and ω_{jt}, as well as between product price-cost margins, $\mathcal{H} \circ \Omega_t^{-1} s_t$, which are also a function of prices, and ω_{jt}. Finally, the nest shares, $s_{j|g,t}$, that are included in the demand side regression are also functions of prices and are thereby also endogenous to the regression.

Similar to Brenkers and Verboven (2006), we adapt the widely used instrumenting strategy of BLP to the nested logit model to construct instruments for our joint estimation. Cost shifters at the firm or product level would be ideal, but as is often the case in the existing literature, cost-side data on such a detailed level is not available to us. Looking back at firms' first-order profit maximizing conditions in (2.13), any variable shifting the right-hand side of the equation qualifies as a potentially relevant and valid instrument. For each individual beer, j, the right-hand side of the equation pinning down its optimal price is a function of j's own market share and that of all its rival products' market shares. Market shares in turn depend on observable product characteristics.

As instruments for the prices, market shares and markups of beers owned by a specific firm, we therefore use functions of the characteristics of products owned by that firm and functions of the characteristics of products owned by all other firms. Moreover, to take account of the nested market structure, we compute these functions within each nest. In addition to these instruments, we also use the number of products owned by the firm and the number of products owned by all other firms. As included instruments, we employ the observed product characteristics excluding prices, nest shares and packaging dummies. Table 2.9 in the Appendix lists the excluded instruments and their correlations with the endogenous regressors. Typically, the correlations have a value of at least .2 and frequently above .3. Moreover, an instrument that is weakly correlated to one of the endogenous regressors, tends to be more strongly correlated to the other endogenous variable. Overall, our set of excluded instruments looks relevant at first sight.

This impression is confirmed when looking at the results of the first-stage regressions of the excluded instruments on prices and category shares, as reported in Table

2.10. In both cases, the excluded instruments are jointly highly statistically significant and explain a sizable share of the variation in prices and nest shares. In the first-stage regression on prices, more than forty percent of the total variation of prices is explained by the excluded instruments. This share falls in the regression on nest shares. Nevertheless, the excluded instruments still explain a third of the variation in nest shares. We therefore conclude that our selected instruments are relevant. To address the validity of our instruments, we move on to the identification strategy.

4.4. Identification. As is common practice in the existing literature, we assume that our instruments are orthogonal to the error terms in (2.12) and (2.16) and estimate the system of supply and demand by efficient GMM instrumental variables estimation.

$$E\begin{bmatrix} z_{jt}\xi_{jt} \\ z_{jt}\omega_{jt} \end{bmatrix} = 0,\ t = 1, \ldots, T,\ j = 1, \ldots, J_t$$

Given this orthogonality assumption and defining $\theta \equiv [\alpha, \beta, \sigma, \gamma]'$, we can write the objective function of the GMM estimator as

$$J(\theta) = g(\theta)'\widetilde{W}g(\theta), \tag{2.17}$$

where

$$g(\theta) \equiv \sum_t \sum_j \begin{bmatrix} z_{jt}\xi_{jt}(\theta) \\ z_{jt}\omega_{jt}(\theta) \end{bmatrix}$$

are the stacked moment conditions of the demand and supply equations and $\widetilde{W}$ is the optimal weighting matrix. With L instruments in total, $g(\theta)$ is $(2L \times 1)$ and $\widetilde{W}$ is $(2L \times 2L)$. We use an updating procedure to obtain a consistent estimate of the optimal weighting matrix. In the first stage, we assume homoscedastic errors, so that we set $\widehat{W} = I_2 \otimes (Z)'^{-1}$. This stage yields the error estimates $[\widehat{\xi}, \widehat{\omega}]$, which we use to update the weighting matrix. Here we allow the demand and supply shocks to be clustered at the firm level. Hoxby and Paserman (1998) show that tests of instrument orthogonality assuming homoscedasticity tend to overreject the null of orthogonality if the data is characterized by strong intra-cluster correlation and little within-group variation of the instruments. Given the large number of products owned by several firms and in the market overall, the construction of the BLP instruments tends to yield limited variation at the firm level. We found that even though the computation of the instruments for each category raises the variation, accounting for clustering of the errors at the firm level still aids in reducing the GMM estimator's objective function value.

Then, our estimate of the optimal weighting matrix is given by [12]

$$\widetilde{W} = I_2 \otimes \begin{bmatrix} (Z'\widehat{\Omega}_d Z)^{-1} \\ (Z'\widehat{\Omega}_s Z)^{-1} \end{bmatrix},$$

where $\widehat{\Omega}_d$ and $\widehat{\Omega}_s$ are the covariance matrices for the demand and supply shocks, respectively. The true structural demand and supply parameters minimize the objective function value, (2.17). We verified that our estimation procedure yields the same coefficient vectors for different starting values. Our estimate of the covariance matrix between the demand and supply shocks, $\Sigma_{\xi,\omega}$, is then simply

$$\Sigma_{\xi,\omega} = T^{-1} \sum_t [\xi_t, \omega_t]'[\xi_t, \omega_t].$$

We turn to our results next.

4.5. Results. Table 2.6 presents the results of jointly estimating the demand and supply relations, as well as the estimates of two comparison models. In the column labeled (I) we estimate only the demand-side regression equation, (2.12), without instrumenting for price and the nest market shares. In specification (II), we repeat the estimation with our set of instruments and employ the identification strategy of the previous section. That is, we use efficient instrumental variables GMM, where we iterate the estimation procedure until the estimate of the optimal weighting matrix converges. Compared with the uninstrumented regression, the coefficients of the endogenous regressors change substantially, indicating that the endogeneity of prices and nest shares is an important feature of the data. The estimated price coefficient increases by more than forty percent, while the nesting parameter, σ, is reduced by more than half.

Apart from the coefficients for the taste characteristics richness and bitterness, there are no coefficients in specifications (I) and (II) with opposite signs. The magnitudes of the remaining coefficients are quite similar with the alcohol content, advertising and the .33 liter can packaging dummy benefiting product market shares more in the instrumented specification.

Specification (III) jointly estimates the demand and supply sides by efficient instrumental variables GMM as does specification (II). This is our preferred specification, and we are going to use the estimated parameters in our counterfactual price simulations below. We would like to emphasize two points regarding the estimates. First, when estimating demand and supply jointly, we have to make an assumption about firm conduct on the supply side. We have assumed that firms behave according to

[12] We iterate the updating procedure for the weighting matrix until convergence.

Table 2.4. Estimation Results

	Demand		
Regressor	(*I*)	(*II*)	(*III*)
Price Per Liter (SEK)	**0.0926**	**0.1323**	**0.1242**
	(.0013)	**(.0069)**	**(.0054)**
$\ln(s_{j\|g})$	**0.7229**	**0.2330**	**0.2563**
	(.004)	**(.0236)**	**(.0187)**
Richness	**-0.2517**	**0.0441**	**-0.1863**
	(.0075)	**(.01824)**	**(.0126)**
Sweetness	**-0.0796**	**-0.1378**	**-0.0893**
	(.0092)	**(.0137)**	**(.0142)**
Bitterness	**0.006**	**-0.0220**	**0.1089**
	(.0059)	**(.0151)**	**(.0099)**
Alcohol (% of Vol.)	**0.5105**	**0.7804**	**0.7698**
	(.0124)	**(.0292)**	**(.0255)**
Advertising (mln SEK)	**0.1857**	**0.3697**	**0.3668**
	(.0153)	**(.0190)**	**(.0155)**
.5 Liter Bottle	**-0.7426**	**-0.5507**	**-0.0364**
	(.0253)	**(.0745)**	**(.0719)**
.33 Liter Can	**0.642**	**1.0680**	**1.6236**
	(.0481)	**(.0805)**	**(.0524)**
.5 Liter Can	**0.0216**	**0.2352**	**0.5709**
	(.0294)	**(.1007)**	**(.1024)**
Constant	**-0.6433**	**-5.6646**	**-5.7611**
	(.07)	**(.3550)**	**(.2991)**
	Marginal Cost		
Richness	**-**	**-**	**-0.0052**
			(.0068)
Sweetness	**-**	**-**	**-0.0122**
			(.0044)
Bitterness	**-**	**-**	**0.0275**
			(.0043)
Alcohol (% of Vol.)	**-**	**-**	**0.0534**
			(.0034)
Exchange Rate (SEK/FOR)	**-**	**-**	**0.0440**
			(.0019)
Constant	**-**	**-**	**3.6281**
			(.0291)
Observations	**15,937**	**15,937**	**15,937**
R^2	**.75**	**.54**	
$\chi^2\left(J(\widetilde{\theta}), \#df\right)$	**-**	**.89**	**.86**
$\Sigma_{\xi,\omega} = \begin{bmatrix} 3.4066 & -0.1338 \\ -0.1338 & 1.0757 \end{bmatrix}$			

Note: The χ^2 statistic of specification (II) is based on 24 degrees of freedom, while that of specification (III) is based on 53.

Nash-Bertrand competition in prices. Making a specific assumption about conduct introduces a potential source of misspecification. We can dispel these concerns, however, when comparing the estimated structural demand parameters in specifications (II) and (III). Except for the richness and bitterness coefficients, as well as the half liter bottle dummy, all of the coefficients are very close to each other, suggesting that the estimation of the cost side does not affect the demand-side parameters in a substantial way.

Second, the estimated cost-side parameters are quite intuitive. Except for richness and sweetness, each of the included taste characteristics raises marginal costs, indicating that attaining higher values for the taste characteristics is costly for producers. Moreover, a beer that is imported from a brewing destination with a high exchange rate also has a higher marginal cost than a rival product that is imported from a country with a lower exchange rate. This is exactly what we would expect, given that importers of foreign beer varieties should pass through (at least part) of a rise in the exchange rate.

The obtained values of the J-statistics for both specifications (II) and (III) show that we cannot reject the orthogonality of the instruments at the ten percent significance level. Finally, the estimated covariance matrix of the demand and cost shocks indicate that the shocks exhibit little correlation and that demand shocks have a substantially higher variance than cost shocks.

Having determined the structural demand and cost parameters, we next use our knowledge of the retail monopolist's pricing rule to back out marginal costs, markups and price elasticities at the firm-level.

5. Firm-Level Marginal Costs, Markups and Price Elasticities

With public knowledge of the retail monopolist's pricing function, (7),we can back out firm-level marginal costs and markups. We simply have to adjust the optimal pricing formula to take into account how a change in wholesale prices impacts the final retail prices, $\partial p_{rjt}/\partial p_{wjt} = \kappa_{jt} \equiv (1+\tau_{ct})(1+m_{st})$.

$$c_{wt} = p_{wt} + \mathcal{H} \circ \Omega_{wt}^{-1} s_t, \; t = 1, \ldots, T \tag{2.18}$$

The wholesale index, w, indicates that the entries of the matrix of market share derivatives, Ω, have been multiplied by κ_t. To avoid confusion with the terms margin and markup, we define the markup of beer j over its marginal cost of production as $(p_{wjt} - c_{wjt})/p_{wjt}$, while the (price-cost) margin is $(p_{wjt} - c_{wjt})$.

The backed-out patterns of markups in Table 2.5 relate closely to the market descriptives presented in Table 2.1. The ale segment of the market is characterized by the lowest level of concentration with an HHI value of only .15, whereas the dark lager segment exhibits the highest level of concentration with an HHI value of .75. This carries over to the average markups attained for the two categories. Dark lagers on average boast a markup of almost forty percent, which is double that of ales. Only light lagers attain a higher markup on average. At first sight, this is somewhat surprising, given

TABLE 2.5. Firm-Level Marginal Costs and Markups

	Markup over Wholesale Price (%)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	20	38	42	28	27
median	19	32	41	25	26
1^{st} percentile	7.6	11.6	17.6	13.1	17.1
99^{th} percentile	36.4	78.9	87.8	70.3	41.2
	Wholesale Marginal Cost (SEK)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	18.96	9.78	7.29	14.25	12.49
median	17.52	8.49	6.27	14.88	12.89
1^{st} percentile	6.76	1.79	.61	2.47	6.98
99^{th} percentile	51.96	29.89	19.25	27.53	21.05

Note: There are four outliers with a markup greater than 1. All marginal costs are below their corresponding wholesale prices.

that the level of concentration is not nearly as high as in the dark lager segment. The light lager segment has by far the highest number of products on offer, which naturally limits the HHI value in this category, given that there are many small firms with only a single beer on offer. On average, however, each firm that is active in the light lager segment owns eight beers, while the corresponding numbers in the dark lager and ale segments are only 1.63 and 2.65, respectively. Similarly, in the stout and weissbeer segments, the average firm only owns one beer. Thus, firms in the light lager segment capture market power by selling large portfolios of products instead of only a single or relatively few beers. This raises attained markups for the big light lager producers. The stout and weissbeer segments settle in-between the ale and dark lager segments in terms of markups, which mirrors the observed levels of concentration in these two segments.

Table 2.6 shows the own- and cross-price elasticities for the different beer categories. The computation of elasticities at the retail- and wholesale-levels is analogous to the computation of marginal costs. Thus, at the retailer level, we use observed retail prices and the market share derivatives contained in Ω. At the wholesaler- or firm-level, we use observed wholesale prices and the elements of Ω_w. As can be expected, the pattern of elasticities ties in with the pattern of markups. Overall, ales tend to be more price-elastic than the other beer categories, which are similar in terms of own-price elasticities. The difference between wholesale and retail own-price elasticities is sizable and driven by the big differences between retail and wholesale prices, as shown in Table 2.1. The within-category substitutability of products is highest in the dark lager segment and lowest in the light lager segment. Given the high number of light

Table 2.6. Own- and Cross-Price Elasticities

	Retail Own-Price Elasticities				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	-8.18	-6.09	-5.34	-6.56	-6.04
median	-7.41	-5.52	-5.18	-6.59	-6.06
1^{st} percentile	-19.75	-14.23	-8.53	-9.95	-8.02
99^{th} percentile	-4.78	-3.02	-3.30	-3.52	-4.53
	Wholesale Own-Price Elasticities				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	-5.77	-3.63	-2.93	-4.56	-3.99
median	-5.35	-3.23	-2.69	-4.80	-3.98
1^{st} percentile	-14.43	-8.68	-5.79	-8.08	-6.18
99^{th} percentile	-2.75	-1.60	-1.30	-1.62	-2.82
	Average Retail Within-Segment Cross-Price Elasticities				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	.04	.12	.01	.16	.26
median	.01	.03	.01	.07	.12
1^{st} percentile	0	0	0	0	0
99^{th} percentile	.31	1.04	.13	.84	.84
	Average Retail Cross-Segment Cross-Price Elasticities				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
Ales	.04	.0033	.0080	.0013	.0007
Dark Lagers	.0033	.12	.0080	.0012	.0006
Light Lagers	.0080	.0080	.01	.0011	.0006
Stouts	.0013	.0012	.0011	.16	.0006
Weissbeers	.0007	.0006	.0006	.0006	.26

Note: There are no outliers with retail own-price elasticities greater than -1.

lager varieties, this finding is not too surprising. Across categories, light lagers and dark lagers, as well as light lagers and ales are most substitutable.

We now move on to simulating firms' price responses to demand and cost shocks.

6. Counterfactual Price Simulations

To simulate counterfactual prices, we determine the retail prices, $\widetilde{p}_r$, that solve the system[13]

$$\widetilde{p}_r = \Delta \circ \widehat{c} + \widetilde{\Omega}^{-1}\widetilde{s}, \tag{2.19}$$

where $\circ$ is the pointwise or Hadamard product and Δ is a vector of perturbations. By setting the entries in Δ, we can simulate the general equilibrium prices that firms choose in response to the particular shock to marginal costs. To be clear, raising all

[13] Alternatively, we can base our simulations on the system $\widetilde{p}_w = \Delta \circ \widehat{c} + \widetilde{\Omega}_w^{-1}\widetilde{s}$. Both systems yield the same prices.

TABLE 2.7. Three Standard Deviation Shock to All ξ's

	Actual Ownership Pattern				
	Retail Price Change (%)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	.33	1.72	1.43	1.26	.21
median	.05	.85	.98	.56	.14
1^{st} percentile	0	0	11	0	0
99^{th} percentile	2.77	6.88	5.57	5.08	.98
	Single-Product Firms				
	Retail Price Change (%)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	0	.03	.07	.01	0
median	0	0	.01	0	0
1^{st} percentile	0	0	0	0	0
99^{th} percentile	.05	.42	.71	.08	.02

Note: There are four outliers with a markup greater than 1. All marginal costs are below their corresponding wholesale prices.

marginal cost by ten percent, Δ's entries are 1.1. Shocks to demand, $\Delta\xi$, directly enter the matrix of market share derivatives and the vector of market shares.

The tilde superscript indicates all the terms adjusting during the counterfactual simulation. Retail prices change in response to shocks to marginal costs and demand, which in turn reallocates market shares. The elements of $\widetilde{\Omega}$ are functions of the estimated demand parameters, which we keep fixed, and the reallocated market shares, $\widetilde{s}$. We obtain the solution to (2.19) by iterating over firms' best responses until convergence to the perturbed Nash-Bertrand pricing equilibrium.

Tables 2.7 and 2.8 present the simulated price responses caused by an aggregate positive three standard deviation shock to demand and an aggregate ten percent rise in marginal costs, respectively. We consider two scenarios here. In the top panels, we show the changes in retail prices imposing the actually observed pattern of product ownership. In the bottom panels, we impose that each product is treated as a standalone firm.

Turning to Table 2.6, there are large differences between the two scenarios. This is in line with the results from Section 2. For dark and light lagers, the mean price response imposing the actual ownership pattern is roughly sixty and twenty times larger compared with imposing single-product firm price setting, respectively. For the other categories the effect is also substantial. In absolute levels, the average price changes are biggest for the dark and light lager segments. This is in line with the estimated high elasticities in the dark lager segment and the large portfolios of beers that firms own in the light lager segment.

Table 2.8. Ten Percent Increase in All Marginal Costs

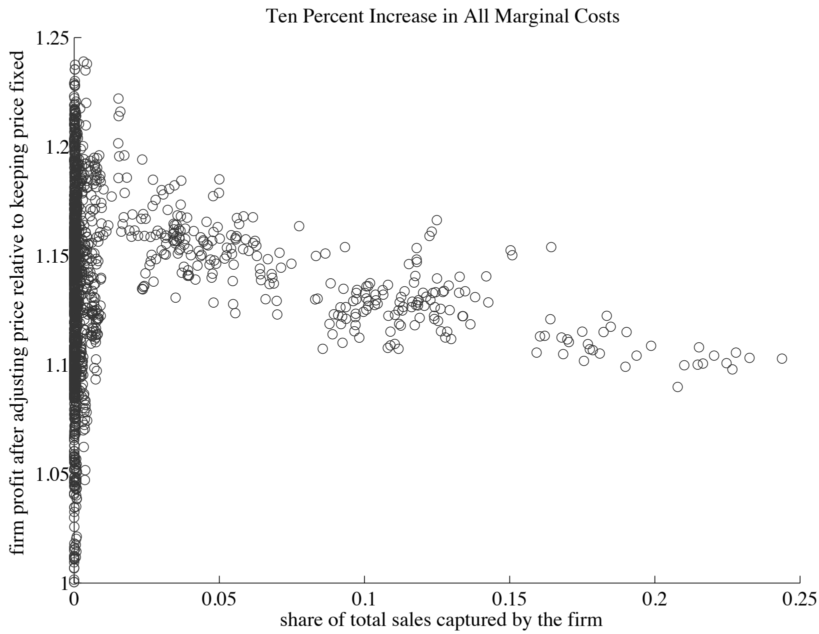
	Actual Ownership Pattern				
	Retail Price Change (%)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	8.58	7.75	7.69	8.09	8.21
median	8.57	7.95	7.72	8.15	8.23
1^{st} percentile	7.72	5.36	6.02	6.58	7.81
99^{th} percentile	9.49	9.28	8.76	8.91	8.64
	Single-Product Firms				
	Retail Price Change (%)				
	Ales	Dark Lagers	Light Lagers	Stouts	Weissbeers
mean	8.68	8.24	8.04	8.41	8.32
median	8.64	8.19	8.06	8.41	8.32
1^{st} percentile	7.98	7.05	6.98	7.68	7.97
99^{th} percentile	9.49	9.30	8.82	8.99	8.74

Note: There are four outliers with a markup greater than 1. All marginal costs are below their corresponding wholesale prices.

The simulated price responses in Table 2.8 show that for shocks to marginal costs the price increase with the actual ownership matrix is lower than in the counterfactual simulation where each product is treated as a stand-alone firm. Again, this is in line with expected results from Section 2. The magnitude of the price changes is lowest in the segment with the highest concentration (dark lagers) and in the segment with the biggest firm portfolios (light lagers). The response to marginal cost shocks is dampened by firms capturing larger shares of total sales and a higher level of own-price and cross-price elasticities. This is exactly what characterizes the dark and light lager segments, compared to the other beer categories.

As noted in the Introduction some recent papers have examined the links between multi-product firms and menu costs. Price setting for a portfolio of products will affect not only the pass-through rates as examined above, but also the curvature of the profit function around the optimal prices. To investigate this latter issue we examine the slope of each firm's profit function around the optimal prices of all the beers in each firm's holdings. This is what we do in Figures 2.1 and 2.2. On the y-axis, we plot the relative change in firm profits, that can be attained by each firm adjusting the prices of its beers in response to the particular shock. To be clear, we obtain the pre-adjustment profits by letting the shock hit firms, holding prices fixed and computing the equilibrium market shares for this scenario. The post-shock profits, are the result of allowing all prices to adjust optimally and computing the new Nash-Bertrand pricing equilibrium. On the x-axis, we plot the total share of sales captured by the firm. We conduct these counterfactual simulations for each of the time periods in the sample and pool the

Figure 2.1. Slope of Firms' Profit Functions

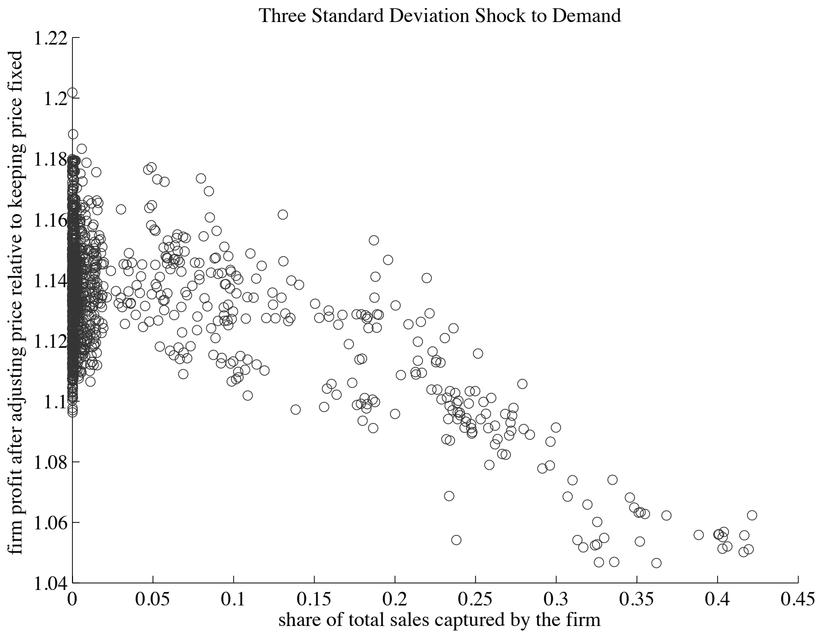


resulting outcomes in the figures.[14] For both cases we can see that there is a wide distribution of profitability levels for firms that only account for a very small share of total sales. For firms with a non-negligible share of the market, the relationship between the relative gain in firm profits and firm shares is clear. As firms become bigger in terms of their products' combined market shares, they gain relatively less by adjusting prices in response to shocks. In other words, the larger the firm's product portfolio in terms of sales, the flatter is its profit function around its portfolio's optimal price vector.

This result speaks directly to size of menu costs necessary to keep prices fixed in response to cost and demand shocks. Each firm would be willing to pay at most the relative gain in productivity to adjust its prices. For firms with few products and small market shares, this relative amount is larger than the amount a firm with many products and a larger market share is willing to pay. Thus, in a concentrated market, lower price adjustment costs are needed to generate the same level of price stickiness

[14] That is why summing the market shares captured by each firm yields a sum larger than 1.

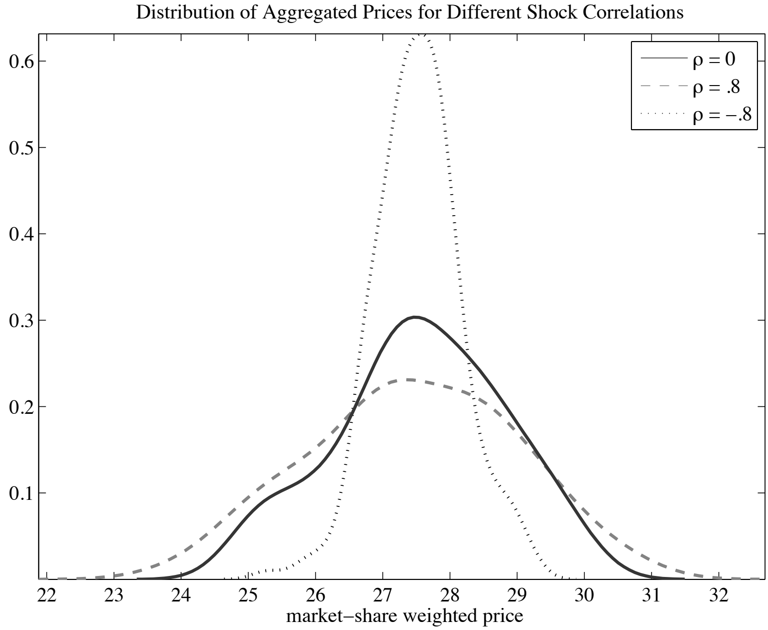
FIGURE 2.2. Slope of Firms' Profit Functions



as in an unconcentrated market, irrespective of if it is shocks to cost or to demand that is affecting prices.

Finally, we turn to the comovement of demand and cost shocks. We estimated the actual covariance matrix of demand and cost shocks to be $\Sigma_{\xi,\omega} = \begin{bmatrix} 3.4066 & -0.1338 \\ -0.1338 & 1.0757 \end{bmatrix}$, which yields a correlation of $-.07$. In Figure 2.3, we show the distribution of aggregated prices for different correlation patterns between ξ and ω. The diagonal elements of $\Sigma_{\xi,\omega}$ are as estimated. The off-diagonal elements are modified to yield a zero, a .8 and a $-.8$ correlation for the three cases, respectively. A positive demand shock requires higher prices to restore optimal profits. The same holds for higher costs. Thus, if demand and cost shocks are correlated, both shocks pull prices in the same direction. With a negative correlation, the effects point in opposite directions. To generate the actual sample of cost and demand shocks in each case, we draw two hundred times from a multivariate normal distribution with mean zero and the modified $\Sigma_{\xi,\omega}$ as covariance matrix. We then simulate the optimal general equilibrium prices and aggregate them

FIGURE 2.3.



with the altered allocation of market shares. As should be the case, all the distributions have roughly the same means. The distribution for the case with a strong negative correlation is most concentrated around the mean, while the distribution for the case of a positive correlation is flattest.

Analogous to the impacts of product substitutability and concentration levels, the correlation of demand and cost shocks can substantially alter firms' incentives to change prices.

7. Conclusion

Both from a theoretical and empirical perspective multiproduct firms and high product substitutability are important determinants for the size of price adjustments in reaction to marginal cost and demand shocks. In the presence of strictly positive crossprice elasticities, multiproduct firms react less to cost shocks and more to demand shocks than their single product rivals. This effect strengthens with the degree of concentration in the market, but importantly it is not simply driven by single products capturing higher market shares. Single product firms' price reactions increase with the market share of their product. What matters for multiproduct firms' price setting is not only the total market share captured by their portfolio of products, but the substitutability of the products within its holdings.

Using data from the Swedish retail market for beer, we estimate a nested logit model of demand, back out firms' marginal costs using the demand parameters and simulate counterfactual price responses for aggregate cost and demand shocks. The interesting aspect here is the variation in both concentration levels, as measured by the HHI, and the size of crossprice elasticities for the different beer types. While both ales and dark lagers are characterized by high crossprice elasticities, the former segment is unconcentrated and the latter is highly concentrated. This translates into virtually no effect for the magnitude of price changes when moving from single product to multiproduct firms in the ale segment. In the dark lager segment, however, firms' price reactions to demand and cost shocks is affected substantially.

These sizable effects point to the dangers of overestimating price adjustment costs (convex adjustment or menu costs) when abstracting from variations in product substitutability and multi-product firms.

8. References

Akerlof, George A. and Yellen, Janet L ., (1985), 'A Near-rational model of the business cycle, with wage and price inertia, *Quarterly Journal of Economics*, 100, issue 5, p. 823-38.

Alvarez, Fernando and Lippi, Francisco, (2012), Price setting with menu costs for multi-product firms", *Manuscript*, University of Chicago.

Anderson, Simon P., De Palma, André and Thisse, Jacques-François, (1992), Discrete choice theory of product differentiation, MIT Press, Cambridge, MA.

Bajari, Patrick, Benkard, C. Lanier and Levin, Jonathan, (2007), Estimating dynamic models of imperfect competition," *Econometrica*, 75, issue 5, 1331–370.

Barro, Robert J., (1972), A theory of monopolistic price adjustment, *Review of Economic Studies*, 39, 17-26.

Berry, Steven T., (1994), Estimating discrete-choice models of product differentiation, *RAND Journal of Economics*, 25, issue 2, p. 242-62.

Berry, Steven T., Levinsohn, James and Pakes, Ariel, (1995), Automobile prices in equilibrium, *Econometrica*, 63, issue 4, 841-90.

Besanko, David, Gupta, Sachin and Jain, Dipak, (1998), Logit demand estimation under competitive pricing behavior: An equilibrium framework, *Management Science*, 44, issue 11, p. 1533-47.

Besanko, David, Dubé, Jean-Pierre and Gupta, Sachin, (2005), Own-brand and cross-brand retail pass-through, *Marketing Science*, 24, issue 1,p. 123-37.

Besley, Timothy J. and Rosen, Harvey S., (1999), Sales taxes and prices: An empirical analysis, *National Tax Journal*, 52, issue 2, p. 157-78.

Bhattarai, Saroj and Schoenle, Raphael, (2012), Multi-product firms and price-setting: Theory and evidence from U.S. producer prices, *Manuscript*, Pennsylvania State University.

Blinder, Alan S., Canetti, E. R. D. , Lebow, D. E. and Rudd, J. B., (1998), Asking about prices: A new approach to understanding price stickiness, New York, Russell Sage Foundation.

Brenkers, Randy and Verboven, Frank, (2006), Liberalizing a distribution system: The European car market, *Journal of the European Economic Association*, 4, issue 1, 216 - 51.

Bulow, Jeremy I. and Pfleiderer, Paul, (1983), Note on the effect of cost changes on prices, *Journal of Political Economy*, 91, issue 1, p. 182-85.

Burstein, Ariel and Hellwig, Christian, (2007), Prices and market shares in a menu cost model, *Mansucript*, UCLA.

Davis, Peter and Garcés, Eliana, (2010), Quantitative techniques for competition and antitrust analysis, Princeton University Press.

Delipalla, Sofia and Keen, Michael, (1992), The comparison between ad valorem and specific taxation under imperfect competition, *Journal of Public Economics*, 49, p. 351-67.

Dhyne, Emmanuel, Álvarez, Luis J., Hervé, Le Bihan, Veronese, Giovanni , Dias, Daniel , Hoffmann, Johannes , Jonker, Nicole, Lünnemann, Patrick, Rumler, Fabio and Vilmunen, Jouko, (2006), Price changes in the Euro area and the United States: Some facts from individual consumer price data, *Journal of Economic Perspectives*, 20, issue 2, p. 171-92.

Dornbusch, Rudiger (1987), Exchange rates and prices, *American Economic Review*, 77, issue 1, 93-106.

Farrell, Joseph and Shapiro, Carl, (2010), Antitrust evaluation of horizontal mergers: An economic alternative to market definition, *The B.E. Journal of Theoretical Economics*, 10, issue 1.

Feenstra, Robert C., (1989), Symmetric pass-through of tariffs and exchange rates under imperfect competition: An empirical test, *Journal of International Economics*, 27, issue 1-2, p. 25-45.

Fullerton, Don and Metcalf, Gilbert E., (2002), Tax incidence, *Handbook of Public Economics*, Eds: A. Auerbach and M. Feldstein, Elsevier Science, The Netherlands.

Goldberg, Penelopi K., (1995), Product differentiation and oligopoly in international markets: The case of the U.S. automobile industry, *Econometrica*, 63, issue 4, p. 891-951.

Goldberg, Penelopi K. and Knetter, Michael M., (1997), Goods prices and exchange rates: What have we learned?, *Journal of Economic Literature*, 35, issue 3, p. 1243-72.

Goldberg, Penelopi K. and Hellerstein, Rebecca, (2011), A structural approach to identifying the sources of local currency price stability, *forthcoming in Review of Economic Studies*.

Golosov, Mikhail and Lucas, Robert E., (2007), Menu costs and the Phillips curve, *Journal of Political Economy*, 115, p. 171-99.

Gopinath, Gita and Itskhoki, Oleg, (2010), Frequency of price adjustment and pass-through, *Quarterly Journal of Economics*, 125, issue 2, p. 675-727.

Gopinath, Gita , Gourinchas, Pierre-Olivier, Hsieh, Chang-Tai , and Li, Nicholas, (2011), International prices, costs, and markup differences", *American Economic Review*, 101, p. 2450-86.

Hall, R. L. and Hitch, C. J., (1939), Price theory and business behavior," *Oxford Economic Papers*, 2, 12-45.

Hamilton, Stephen F., (2009), Excise taxes with multiproduct transactions, *American Economic Review*, 99, issue 1, p. 458-71.

Hausman, Jerry, Leonard, Gregory and Zona, J. Douglas , (1994), Competitive analysis with differentiated products, *Annales D'Economie Et De Statistique*, 34, p. 160-80.

Hoxby, Caroline and Paserman, M. Daniele , (1998), Overidentification tests with grouped data, *NBER Technical Working Paper* #223.

Ivaldi, Marc and Verboven, Frank, (2005), Quantifying the effects from horizontal mergers in European competition policy, *International Journal of Industrial Organization*, 23, issue 9-10, p. 669-91.

Kano, Kazuko, (2006), Menu costs, strategic interactions, and retail price movements", *Manuscript*, Queen's University.

Karp, L.S., and Perloff, J.M., (1989), Estimating market structure and tax incidence: The Japanese television market", *Journal of Industrial Economics*, 37, p. 225-39.

McFadden, Daniel, (1978), Modelling the choice of residential location, in A. Karlgvist, et al., eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland.

Means, Gardiner, (1935), Industrial prices and their relative inflexibility, *U.S. Senate Document*, #13, 74th Congress, 1st Session, Washington, 1935.

Midrigan, Virgiliu, (2011), Menu costs, multiproduct firms, and aggregate fluctuations, *Econometrica*,79, issue 4, 1139-80.

Nakamura, Emi and Steinsson, Jón , (2008), Five facts about prices: A reevaluation of menu cost models, *The Quarterly Journal of Economics*, 123, issue 4, p. 1415-64.

Nakamura Emi and Zerom, Dawit, (2010), Accounting for incomplete pass-through, *Review of Economic Studies*, 77, issue 3, p. 1192-1230.

Nevo, Aviv, (2000), Mergers with differentiated products: The case of the ready-to-eat cereal industry, *The RAND Journal of Economics*, 31, p. 395-421.

Nevo, Aviv, (2011), 'Empirical models of consumer behavior, *forthcoming in Annual Review of Economics.*

Noton, Carlos, (2011), Home bias, structural demands and the European car market", *Manuscript*, University of Warwick.

Reny, Philip, Wilkie, Simon and Williams, Michael A., (2011), Tax incidence under imperfect competition," available at *SSRN*, http://ssrn.com/abstract=1874346.

Sheshinski, Eytan, and Weiss, Yoram, (1977), Inflation and costs of price adjustment, *Review of Economic Studies*, 44, p. 287-303.

Stella, Andrea, (2012), "The magnitude of menu costs: A structural estimation", *Manuscript*, Harvard University.

Sudhir, K., (2001), Structural analysis of manufacturer pricing in the presence of a strategic retailer, *Marketing Science*, 20, issue 3, p. 244-264

Sweezy, Paul M., (1939), Demand under conditions of oligopoly, *Journal of Political Economy*, 47, p. 568-73.

Weyl, E. Glen and Fabinger, Michal, (2011), Pass-through as an economic tool, *Manuscript*, Harvard University.

PAPER 3

Crowding and Consumer Welfare in the Swedish Light Lager Market

André Romahn

Abstract

I estimate consumer welfare gains from the entry of new products in the Swedish market for light lager beers by estimating a structural demand model using bar-code level data provided by the Swedish alcohol retail monopoly. Following Ackerberg and Rysman (2005), I explicitly allow for crowding in the product space as new varieties enter. I find strong evidence for crowding effects. My results suggest that consumers do not value new goods *per se*; the love-for-variety effect is absent. Further inspection of the entering and exiting products suggests that entrants on average provide consumers with higher utility than exiting lagers. In a methodological contribution, I show that following the commonly accepted practice of determining the relevant market outside of the demand estimation can be misleading. A positive correlation between net entry and total consumer welfare can simply be spurious and does not necessarily imply that entry raises welfare. My findings regarding the benefits of new varieties are robust to this concern.

JEL Classification: D6, L10, L66.

Keywords: structural demand estimation, crowding, consumer welfare, entry and exit, beer.

I would like to thank Richard Friberg and Cristian Huse for detailed comments on this paper and seminar participants at the Research Institute of Industrial Economics, the Stockholm School of Economics and the University of Mannheim for their suggestions regarding an earlier draft of this paper.

1. Introduction

The introduction of new varieties is a central feature of differentiated goods markets. Whether such novel products contribute to an increase in consumer welfare has been of strong interest in the existing literature (Berry and Waldfogel (1991), Petrin (2000), Goolsbee and Petrin (2004)). In many cases, previous findings are based on logit-type structural models of demand that allow to estimate changes in consumer welfare as new product varieties enter the market.

This paper directly deals with an undesirable feature of the functional form of logit demand. Imposing logit-type preferences on consumers ensures that consumers value goods as such and are thereby willing to pay for having more choice (as in Dixit-Stiglitz preferences, love-for-variety effect). In the standard logit model, however, consumers' willingness to pay for having the choice among more products drives down consumers' willingness to pay for better quality products, as the number of varieties in the market increases. Imposing this counterintuitive pattern on the data solely by functional form threatens the credibility of welfare estimates.

Ackerberg and Rysman (2005) develop a modification of the logit model that allows for a decreasing willingness to pay for more choice as the number of varieties increases (product space congestion or crowding). Contrary to alternative approaches (Berry and Pakes (2007), Bajari and Benkard (2005)), the modification is easily implemented for markets with a large number of products.

Methodologically, I analyze how predictions of changes in consumer welfare in response to entry differ between a demand model adjusting for crowding effects and a demand model that does not. At one extreme, a pure congestion model eliminates the love-for-variety effect, while a logit model maximizes the impact of changes in the number of products on welfare. The difference between the two models is how the change in welfare is allocated between two components: consumers' monetary valuation of new varieties entering the market and consumers' willingness to pay for the average mean utilities of the available products (quality). I show that in-sample both types of models make exactly the same predictions about the rate of change of total consumer welfare. This is based on the commonly accepted practice of defining the market's potential outside of the demand model and the fact that consumer welfare is simply a transformation of the outside good's market share.

In a novel decomposition of consumer welfare in the framework of Ackerberg and Rysman (2005), I relate the share of the outside good and the number of products to the two components of consumer welfare. Then, in the logit model, all else equal, a fall of the outside good's market share and an increase in the number of products raises consumer welfare. In the pure congestion model, a reduction in the share of the outside good is sufficient to yield the same outcome. Thus, if the entry of new products is associated with a decrease in the share of the outside good, both models predict that consumers gain. This finding is robust across the logit and pure congestion models.

Even though economically intuitive, I caution that this reasoning can be misleading. As is common practice in the literature, I define the market's potential size outside of the demand estimation. Put differently, the share of the outside alternative is not jointly estimated with the structural demand parameters. A positive correlation between net entry and consumer welfare can therefore be spurious.

To address this concern, I use my analytical welfare decomposition to derive deterministic relationships between the estimated and directly observed components of each model. Changes in the number of varieties and the share of the outside good are observed directly, while consumers' willingness to pay for the size and quality of their choice set are estimated. In the pure congestion model, the share of the outside good and consumers' willingness to pay for quality are perfectly negatively correlated. While in the standard logit model, consumers' willingness to pay for having more choice is perfectly positively correlated with the number of varieties in the market. These deterministic relationships hold for the aggregate market and not necessarily for individual or groups of products. I therefore disaggregate the market into groups of incumbents, entrants and exiting varieties and compare how closely the estimated welfare components for each of these groups conforms to the aggregate relationship with the share of the outside good and observed net entry.

The methodological upshot of the paper is that it is desirable to jointly estimate the market share of the outside alternative with the structural demand parameters. This could alleviate concerns that entry and consumer welfare are spuriously correlated by the exogenous delineation of the relevant market.

Empirically, I add to the few papers (Mariuzzo et al. (2010)) that use the framework of Ackerberg and Rysman (2005) to assess the empirical importance of crowding effects in a market with many differentiated products. The Swedish market for light lagers provides a good testing ground, because a liberalization of market access in the wake of Sweden entering the European Union's common market in 1995 has resulted in a rapid increase of the number of lagers available to Swedish retail customers.

Indeed, I find strong evidence for substantial crowding effects in the Swedish light lager market. The estimation results suggest that the market is best described by a pure congestion model, where consumers do not value additional goods *per se*. In other words, the love-for-variety effect is absent.

Nevertheless, consumers are still willing to pay for higher quality beers. I find that exiting lagers provide the representative consumer with less utility than newly introduced lagers. This effect is strongly driven by entering lagers having lower prices than the beers that exit the market, while offering similar non-price characteristics. Net entry is therefore beneficial to consumer welfare. Moreover, I can dispel concerns that this finding is the artifact of a spurious correlation between changes in the number of varieties and changes in the share of the outside good.

For both the logit and pure congestion models, the welfare estimates for the group of entering beers are least influenced by movements in the observable changes in the number of lagers and the share of the outside good. It is therefore unlikely that the positive effect of entry on consumer welfare is a product of the functional form of the estimator.

The remainder of the paper is structured as follows. Section 2 presents some stylized facts about the Swedish light lager market. Section 3 outlines the advantages of the approach by Ackerberg and Rysman (2005) over alternative demand models allowing for crowding for the data at hand. The demand model and the relationship between its observable and estimated components are analyzed in detail. Section 4 presents the outcomes of the demand estimation and its implications for consumer welfare. Section 5 analyzes the relation between the entry and exit of lagers and consumer welfare. Section 6 concludes.

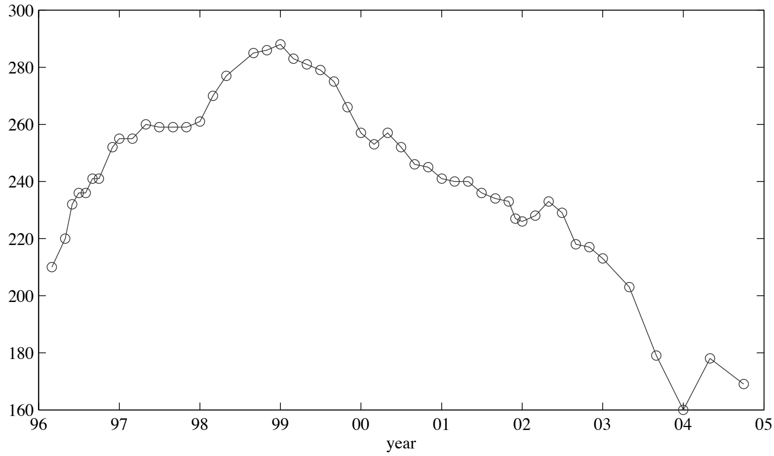
2. The Swedish Market for Light Lagers

I use nationwide monthly retail sales data at the bar-code level provided by Systembolaget, the Swedish retail monopoly for alcohol. The sample covers the period from January 1996 to December 2004. I only observe sales of beers (liters and revenue in Swedish krona) with a minimum alcohol content of 3.5 percent of volume. Beers with a lower alcohol content can be purchased in regular supermarkets and these sales are not covered by the data. I use the terms a lager and a product interchangeably. Heineken in a .33 liter bottle is an example of a lager.

On average almost 9.5 million liters of light lager beer are sold monthly during the sample period. This corresponds to average monthly sales of 274 million Swedish krona. In terms of sales of all types of beer[1], light lagers capture between 93 and 97 percent of total revenue and liters sold. Moreover, the number of light lagers tracks the total number of beers sold by Systembolaget very well. I therefore focus on the light lager segment and drop the other types of beer.

[1] There are also ales, dark lagers, spontaneously fermented beers, special beers such as Christmas beers, stouts and weissbeers on sale.

FIGURE 3.1. The Number of Light Lagers for Sale During the Sample Period



Before the beginning of 1995, the Swedish market for alcohol was characterized by two monopolies. Vin & Sprit[2] owned the monopoly on the production, import, export and wholesaling of alcohol, while Systembolaget owned the monopoly for sales of alcohol to restaurants and the retail monopoly.

On the first of January 1995, however, Sweden joined the European Union (EU) and its common market. The European Commission ruled that the monopolies owned by Vin & Sprit and Systembolaget are not compatible with the common market and have to be abandoned upon Sweden's entry to the EU. As the Swedish government views the consumption of alcohol as a potential hazard to the public's health, however, it was allowed to let Systembolaget retain its retail monopoly for alcohol. All other monopolies had to be abandoned. One of the consequences was that Vin & Sprit no longer had control over which product is allowed to enter the market. For the Swedish light lager market this led to a rapid rise in the number of products available for sale in the retail outlets of Systembolaget until the beginning of 1999. Figure 3.1 plots the number of available varieties over the sample period.

Initially, the number of light lagers on offer increases by almost forty percent from January 1996 to January 1999. A period of consolidation follows until May 2002 in

[2] Outside of Sweden, Vin & Sprit is probably best known for its Absolut Vodka brand. On the 31st of March, 2008, the Swedish government eventually sold the company to Pernod Ricard for 55 billion Swedish krona

which the number of products is reduced by roughly twenty percent. During the remaining months of the sample, the pace of product reductions quickens substantially as the lagers on offer are reduced by almost another thirty percent. Given the substantial variation in the number of products and the rapid initial increase at the beginning of the sample, that is likely due to the liberalization of market entry, the Swedish light lager data should provide a good testing ground for the extent and relevance of crowding effects in a market with a very large number of differentiated products.

Before I move on to the description of the demand model, I have to mention that Systembolaget publishes the range of products that can be bought in its retail outlets in regular intervals in freely available catalogs. It enforces the rule that prices can only be changed and entry and exit of products can only take place when a new catalog is published. Thus, in the periods in-between catalog issues, prices and the number of products are held fixed. Ackerberg and Rysman (2005) caution that when including such periods, I would effectively be attempting to identify price elasticities and crowding intensities in these periods without ever observing a price change or entry and exit of products. This can yield biased estimates of the structural demand parameters. I therefore drop the periods in-between catalog issues, which reduces the months in the sample period from 108 to 50. When estimating a logit model with the full and reduced sample, the differences in the point estimates are small, which suggests that the periods in-between issues do not contain much useful variation.

3. Demand

The inability of standard discrete choice models to capture crowding effects as the number of products increases, is based on the assumption of each individual consumer, i, having a logit-type taste shock for each product j, ϵ_{ij}. Crucially, these shocks have unbounded support. Bajari and Benkard (2003) investigate the implications of this assumption in a discrete choice model nesting the logit, nested logit, generalized extreme value, and random coefficients logit models. The authors derive two major implications of the logit taste shock assumption as the number of products tends to infinity. First, for any given share of the outside good, consumer welfare rises without bound. Second, individual i's taste shock for product j, ϵ_{ij}, fully accounts for i's utility derived from purchasing the product, u_{ij}.[3]

The first finding implies that using standard random utility models to analyze the welfare gains from new product introductions is faced with serious limitations. It is unappealing to impose welfare gains from enlarging consumers' choice set, if there

[3] Bajari and Benkard (2003) go on to identify more unpleasant properties of the logit taste shock assumptio as the number of products becomes very large. For my purposes, however, the above two findings are the most relevant.

already are many products in the market. Given a limited space in which products differentiate, any two closely neighboring products eventually become indistinguishable and consumer utility is unaffected when removing one of the products. This mechanism, however, is absent in standard logit type models of demand.

The second finding implies that estimates of the observable taste coefficients are biased downwards in a market with many products. As the number of products expands, the estimated mapping from product characteristics to consumer utility is becoming less and less relevant for explaining purchasing decisions, until the distributional assumption alone fits the observed pattern of market shares.

In response to these undesirable features of standard discrete choice models, several alternatives have been proposed. Ackerberg and Rysman (2005) allow the mean of the random taste shocks, ϵ_{ij}, to be a decreasing function of the number of products in the market.[4] This limits both the welfare gains from new product introductions and the role of random taste shocks in accounting for consumer utility. Thereby, the model addresses both of the undesirable features identified by Bajari and Benkard (2003). Moreover, implementing the approach for a market with many products is straightforward.

Berry and Pakes (2007) and Bajari and Benkard (2005) drop the random error terms ϵ_{ij} altogether and estimate hedonic models of demand, where consumers have preferences over a finite set of product characteristics. These approaches are attractive, as their theoretical predictions are "well behaved" as the number of products increases: closely neighboring products eventually become perfect substitutes and welfare gains from enlarging consumers' choice set only stem from the structurally estimated mean utilities of the products. Thereby, the love-of-variety effect is excluded *a priori*.

Given the high number of light lagers in the Swedish beer market, however, these two approaches are difficult to implement. Berry and Pakes (2007) caution that their algorithm used to minimize the distance between the simulated and observed market shares is not guaranteed to converge and equating the model's predicted shares with the actual ones becomes harder as the number of products increases.

The model of Bajari and Benkard (2005) is unattractive for the data at hand, because it is likely to generate counterintuitively large elasticities. Moreover, with few observable product attributes the model tends to yield few strictly positive cross-price elasticities for any product j.

[4] In an appendix to the paper, Ackerberg and Rysman (2005) illustrate that one can also modify the logit-type estimating equations to allow the number of products to alter the variance of the random errors. The implications are very similar.

TABLE 3.1. Observable Product Characteristics, X

Column	Variable	Mean	[Min,Max]	Std. Dev.
X_1	price per liter	32.09	$[15.6, 64.24]$	6.46
X_2	richness	5.28	$[1, 9]$	1.57
X_3	sweetness	2.10	$[1, 7]$	1.18
X_4	bitterness	6.00	$[1, 12]$	1.88
X_5	alcohol (% of vol.)	5.38	$[4, 10.2]$	.95
X_6	advertising (mln SEK)	.13	$[0, 13.85]$	.76
X_7	.5 liter bottle	.26	$[0, 1]$	.44
X_8	.33 liter bottle	.36	$[0, 1]$	.48
X_9	.5 liter can	.34	$[0, 1]$	.47
X_{10}	entrant	.07	$[0, 1]$	.26
X_{11}	exiter	.13	$[0, 1]$	.33
X_{12}	foreign	.38	$[0, 1]$	.49
X_{13}	constant	1	$[1, 1]$	0

3.1. Demand Model. Following Ackerberg and Rysman (2005), I modify the standard logit framework to allow the mean of consumers' taste shocks, ϵ_{ij}, to decrease with the number of products in the market. Let consumer i's indirect utility from purchasing product j be

$$u_{ijt} = \delta_{jt} + \epsilon_{irt}, \quad i = 1, \ldots, I, \quad j = 1, \ldots, J_t, \quad r = 1, \ldots, R_t. \tag{3.1}$$

$\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ is product j's mean utility, mapping the observable attributes, x_{jt}, the price, p_{jt}, and the unobservable characteristic, ξ_{jt}, to consumer utility. Table 3.1 presents the observable product characteristics included in x_j and their descriptive statistics for the sample period.

Integrating over consumer taste shocks yields the market share of product j.

$$s_{jt} = \frac{R_{jt}\exp(\delta_{jt})}{R_{ot}\exp(\delta_{ot}) + \sum_k R_k \exp(\delta_{kt})} \tag{3.2}$$

δ_{ot} and R_{ot} denote the mean utility and the crowding term for the outside good. Without detailed information about the outside good, it is common practice to hold its mean utility constant. I follow this practice here and make the normalizations $\delta_{ot} = R_{ot} \equiv 0, \ \forall t$. Moreover, I follow Ackerberg and Rysman (2005) and parameterize R_{jt} as

$$R_{jt} = \gamma/\widetilde{J}_t + 1 - \gamma. \tag{3.3}$$

$\widetilde{J}_t \equiv J_t + 1$ is the number of inside products plus the outside good. This choice ensures that R_{jt} is decreasing in the number of products for positive values of γ, the crowding parameter. Applying the results of Small and Rosen (1981) and McFadden (1981), the monetary value of consumer welfare is given by

$$CW_t = (1/\alpha)\ln\left[R_{ot}\exp(\delta_{ot}) + \sum_k R_{kt}\exp(\delta_{kt})\right] + K,$$

where K is an arbitrary constant. Using $R_{jt} = R_t \; \forall t$, $R_{ot} = 1 \; \forall t$, and $\delta_{ot} = 0 \; \forall t$ the expression can be rearranged as follows.

$$CW_t = (1/\alpha)\ln(R_t \widetilde{J}_t \overline{\delta}_t) + K,$$

where $\overline{\delta}_t = \widetilde{J}_t^{-1}\left(1/R_t + \sum_{j=1}^{J_t}\exp(\delta_{jt})\right)$ is the average of the exponential mean utilities of both the outside and inside goods. For simplicity, I simply refer to $\overline{\delta}_t$ as the quality of the choice set. Then, consumer welfare can be decomposed into two relevant terms, the monetary values of the size and the quality of consumers' choice set.

$$CW_t = \frac{\ln(R_t \widetilde{J}_t)}{\alpha} + \frac{\ln(\overline{\delta}_t)}{\alpha} + K \tag{3.4}$$

It is now easy to see that (3.3) has a structural interpretation. For $\gamma = 0$, the crowding term equals one and the expression for consumer welfare specializes to that of the standard logit model. For $\gamma = 1$, the crowding term equals the reciprocal of the total number of products and the first term on the left-hand side vanishes. Thereby, changes in the number of products have no effect on consumer welfare, and (3.4) specializes to the pure congestion model.

Due to the arbitrary constant, interpreting the level of consumer welfare is meaningless. Welfare changes over time, however, are independent of K. I define $\Delta X_{t+1} \equiv X_{t+1}/X_t$ as the (gross) rate of growth of a variable X. The change in consumer welfare between periods t and $t+1$ is then given by

$$CW_{t+1} - CW_t = \frac{\ln(\Delta R_{t+1}\Delta \widetilde{J}_{t+1})}{\alpha} + \frac{\ln(\Delta \overline{\delta}_{t+1})}{\alpha}.$$

Using (3.2), changes in consumer welfare can be related to changes in the share of the outside good.

$$\ln(s_{ot}) = -\ln(R_t \widetilde{J}_t \overline{\delta}_t)$$

It follows immediately that

$$CW_{t+1} - CW_t = -\ln(\Delta s_{ot+1})/\alpha.$$

Combining the two expressions for welfare changes between dates $t+1$ and t yields an identity for the rates of changes of terms that are directly observable and those terms that derive from the structural demand estimation.

$$(\Delta s_{ot+1})^{-1} = \Delta\overline{\delta}_{t+1}\Delta\widetilde{J}_{t+1}\Delta R_{t+1} \tag{3.5}$$

ΔR_{t+1} and $\Delta\overline{\delta}_{t+1}$ are determined by the structural demand parameters, while the remaining terms are observed directly. This identity illustrates that in-sample welfare changes predicted by the logit and pure congestion models are perfectly correlated. If both models yield the same estimate of consumers' price sensitivity, even the level of consumer welfare estimates are identical. Given that substantial congestion in product space implies that taste coefficients are biased downwards in the logit model, this outcome is unlikely, however.

In-sample, the difference between the models is how the change in welfare is allocated between the two components of consumer welfare, $\Delta\overline{\delta}_{t+1}$ and $\Delta\widetilde{J}_{t+1}\Delta R_{t+1}$. Posing the question "Did the entry of good A raise total consumer welfare?" can therefore yield misleading conclusions, because the answer completely depends on the definition of the relevant market. Typically, the delineation of the relevant market lies outside of the demand model and is taken as given during the estimation of the structural demand parameters. Without certainty that the relevant market has been correctly determined, this creates the possibility that entry of products and fluctuations in the share of the outside good are spuriously related.

To further illustrate the dependence of the two welfare components on the outside good's market share, for small gross rates of growth, (3.5) can be approximated in terms of the net rates of growth, $g_X = X_{t+1}/X_t - 1$, by noting that $\ln(\Delta X_{t+1}) \approx g_X$.

$$-g_o = g_{\widetilde{J}} + g_R + g_\delta$$

With the structural crowding term (3.3), this approximation can be specialized to the cases of the logit model and the pure congestion model.

$$g_\delta = \begin{cases} -(g_o + g_{\widetilde{J}}) & ,\gamma = 0 \\ -g_o & ,\gamma = 1 \end{cases}$$

Thus, in the pure congestion model, periods in which the share of the outside good increases (falls) are periods in which the quality of the choice set falls (increases). As $\overline{\delta}$ contains the unobservable product characteristics, ξ, the demand estimation's error term is included in the identity (3.5) and the above approximation. Thus, this relationship between the quality of the choice set and the share of the outside good is deterministic.

For the logit model, a similar relationship holds. The quality of the choice set is predicted to increase (fall), when the sum of the rates of growth of the outside good and the number of products is negative (positive). Holding the share of the outside good constant, it is clear that a rise in the number of products is associated with a fall in quality. The logit model allows a fall in the number of products to compensate for a rise in the share of the outside good. Even as consumers switch to the outside alternative, quality is unaffected, as long as a sufficient number of products exits. In contrast, in the pure congestion model, the number of products is irrelevant. Only the changes in the share of the outside good matter.

When determining the relevant market outside of the demand estimation, one should be aware of the deterministic relationships between the observable and unobservable variables in the demand model. Gauging the impact of a new product on consumer welfare at a time when the outside good's market share is falling will inevitably lead to the conclusion that total welfare is increasing. This result, however, is not necessarily driven by entry but by the definition of the relevant market's size.

I present the estimation framework and results next.

4. Estimation

To arrive at an estimating equation for the crowding model of the previous section, I use the market share equation for product j, (3.2), and apply the Berry (1994) inversion. This yields the regression specification for an arbitrary crowding term, R_{jt}.

$$\ln(s_{jt}) - \ln(s_{ot}) = X_t\beta - \alpha p_t + \xi_t + \ln(R_{jt})$$

I adopt the structural crowding term from Ackerberg and Rysman (2005) and estimate R_{jt} parametrically, as shown in (3.3). Then, my final estimating equation is given by

$$\ln(s_{jt}) - \ln(s_{ot}) = x_{jt}\beta - \alpha p_{jt} + \ln(\gamma/J_t + 1 - \gamma) + \xi_{jt}. \tag{3.6}$$

4.1. Instruments. The unobservable product characteristic ξ_{jt} has a vertical interpretation in the model. All else equal, a higher realization of ξ_{jt} gives product j a greater market share. In other words, consumers' willingness to pay is increasing in the unobservable product characteristic. As firms incorporate this into their pricing decisions, realizations of the unobservable and prices will tend to be positively correlated, which in turn renders prices endogenous. This is a well-known problem in the existing literature and I follow the instrumenting strategy of Berry, Levinsohn and Pakes

TABLE 3.2. Excluded Instruments

Column	Variable	$\rho_{z_i,p}$	Column	Variable	$\rho_{z_i,p}$
$z_{t,1}$	$J_t^{-1}\sum_{j=1}^{J_t} x_{t,2}$	$-.2523$	$z_{t,9}$	$\sum_{j\in\mathcal{F}} x_{t,5}$	$-.2078$
$z_{t,2}$	$J_t^{-1}\sum_{j=1}^{J_t} x_{t,3}$	$.2131$	$z_{t,10}$	$\sum_{j\in\mathcal{F}} x_{t,6}$	$-.1094$
$z_{t,3}$	$J_t^{-1}\sum_{j=1}^{J_t} x_{t,4}$	$-.3362$	$z_{t,11}$	$\sum_{j\notin\mathcal{F}} x_{t,2}$	$.0314$
$z_{t,4}$	$J_t^{-1}\sum_{j=1}^{J_t} x_{t,5}$	$-.2917$	$z_{t,12}$	$\sum_{j\notin\mathcal{F}} x_{t,3}$	$.0895$
$z_{t,5}$	$J_t^{-1}\sum_{j=1}^{J_t} x_{t,6}$	$-.1201$	$z_{t,13}$	$\sum_{j\notin\mathcal{F}} x_{t,4}$	$.0097$
$z_{t,6}$	$\sum_{j\in\mathcal{F}} x_{t,2}$	$-.1907$	$z_{t,14}$	$\sum_{j\notin\mathcal{F}} x_{t,5}$	$.0405$
$z_{t,7}$	$\sum_{j\in\mathcal{F}} x_{t,3}$	$-.2110$	$z_{t,15}$	$\sum_{j\notin\mathcal{F}} x_{t,6}$	$-.0701$
$z_{t,8}$	$\sum_{j\in\mathcal{F}} x_{t,4}$	$-.2044$	$z_{t,16}$	$J_f \equiv \sum_{j\in\mathcal{F}}(1)$	$-.1940$

(1995). Table 3.2 lists all the excluded instruments and their correlation with price. The remaining columns of the instrument matrix Z are the observable characteristics listed in Table 3.1.

To ensure sufficient variation of the instruments across observations, I only include the first five columns of the observable characteristics matrix, because the remaining columns contain dummy variables and the constant. For the majority of the excluded instruments the magnitude of their correlation with price is at least .2, indicating that they can qualify as relevant. To examine whether the instruments fulfill this requirement, the two instrumental variable tables in the Appendix present the results of regressing the excluded instruments only and all instruments on price, respectively. The former regression explains roughly twenty percent of the variation in price and the excluded instruments are jointly significant as implied by the value of the F-statistic. When utilizing the included instruments as well, the regressors explain nearly sixty percent of the variation in price, the instruments are jointly significant and, importantly, the included instruments do not drive out the excluded ones.

To address the question of validity, I test the overidentifying restrictions imposed by the instruments. The bottom panel of Table 3.3 shows the values of the Sargan statistic for the linear instrumental variables regression, specification (II), and the value of the J-statistic for the efficient GMM-instrumental variables estimation, specification (IV). Both statistics are distributed Chi squared with degrees of freedom given by the difference between the number of instruments and the number of regressors. For both instrumental variables specifications, the null of the instruments being orthogonal to the residuals cannot be rejected at the five percent significance level. For the computation of both statistics, I cluster the errors at the firm level.[5] I conclude that the instrumenting strategy of Berry, Levinsohn and Pakes (1995) yields relevant and valid instruments for the data.

4.2. Estimation Results and the Implications for Consumer Welfare. Table 3.3 presents the results of the demand estimation. The first two columns estimate the standard logit model without any adjustment for potential crowding effects. The estimation in column (II) instruments for the endogeneity of prices, as discussed in the previous section, while the specification in column (I), does not use instruments. The fact that the price coefficient in column (II) is almost three times as large as that in column (I) illustrates that the endogeneity of prices is an important feature of the data.

Across all the specifications the estimated coefficients on lagers' taste parameters, richness, sweetness, bitterness and alcohol content are quite similar. Lagers that are very rich in taste and have a high alcohol content tend to have higher market shares, while relatively high scores for sweetness and bitterness yield all else equal smaller market shares. Finally, marketing expenditures tend to raise sales.

The large and highly significant coefficients on the entrant and exiter dummies show that these types of products behave quite differently from incumbent lagers. In contrast to the descriptive statistics section, I assign the entrant dummy to all new products that have been in the market for at most three periods. Similarly, a lager is an exiter during the last three periods of its life.[6] Shortening these entrant end exiter periods yields coefficients of greater magnitude. This outcome is intuitive, as these effects should eventually vanish. As with the taste coefficients, the estimated effects of being an entering or exiting product are very similar across all specifications.

Given this result, the two major implications of the logit taste shock assumption derived by Bajari and Benkard (2003) should be quantitatively important for the sample data. First, I examine the implication that the taste coefficients tend to be biased towards zero in markets with a large number of products. Looking at Table 3.3, a pairwise comparison of the estimated coefficients between both the instrumented

[5] As Hoxby and Paserman (1998) show, not doing so in the presence of intra-cluster correlation tends to yield too frequent rejections of the overidentifying restrictions. For the data at hand, this effect is important.

[6] Naturally, I exclude the first and the last three periods of the sample, when setting up the dummies. This definition of exiter introduces a forward-looking variable. When dropping exiter from the regression, the coefficients are nearly unchanged. The R^2 of the estimation drops, however. This is because even though exiting products tend to lower price, the fall in market share cannot be fully explained by the price adjustment. The remaining product characteristics are fixed. A substantial drop in sales can therefore only be explained by large negative unobservables, ξ.

Table 3.3. Estimation Results

Regressor	**(I)**	**(II)**	**(III)**	**(IV)**
Price per liter	**-.0432**	**-.1009**	**-.0445**	**-.1006**
	(.0136)	**(.0069)**	**(.0130)**	**(.0110)**
Richness	**.0931**	**.0988**	**.0908**	**.0898**
	(.0368)	**(.0168)**	**(.0363)**	**(.0221)**
Sweetness	**-.1023**	**-.0992**	**-.0964**	**-.0453**
	(.1010)	**(.0407)**	**(.0998)**	**(.0432)**
Bitterness	**-.1172**	**-.1118**	**-.1157**	**-.1127**
	(.0361)	**(.0209)**	**(.0357)**	**(.0231)**
Alcohol Vol. %	**.4251**	**.5259**	**.4202**	**.4657**
	(.1219)	**(.0581)**	**(.1220)**	**(.0575)**
.5 Liter Bottle	**-.9455**	**-1.5453**	**-.9570**	**-1.5053**
	(.2727)	**(.1375)**	**(.2618)**	**(.1394)**
.33 Liter Bottle	**-.3375**	**-.8323**	**-.3502**	**-.9357**
	(.2476)	**(.1139)**	**(.2474)**	**(.1161)**
.5 Liter Can	**.6892**	**-.1526**	**.6446**	**-.2221**
	(.2713)	**(.1591)**	**(.2676)**	**(.1899)**
Entrant	**-1.2880**	**-1.3212**	**-1.2711**	**-1.2580**
	(.1728)	**(.0627)**	**(.1721)**	**(.0658)**
Exiter	**-3.6303**	**-3.6612**	**-3.6571**	**-3.6466**
	(.1354)	**(.0520)**	**(.1367)**	**(.0590)**
Foreign	**-.1455**	**.1254**	**-.1538**	**.0632**
	(.2222)	**(.0981)**	**(.2191)**	**(.1171)**
Advertising	**.3512**	**.3342**	**.3487**	**.3316**
	(.0571)	**(.0063)**	**(.0595)**	**(.0071)**
Constant	**-7.0773**	**-5.3170**	**-1.4951**	**.4926**
	(.8440)	**(.2912)**	**(1.0107)**	**(.5766)**
γ	**-**	**-**	**1.0000**	**1.0000**
	-	**-**	**(.0016)**	**(.0009)**
R^2	**.43**	**.41**	**.43**	**.42**
Sargan Stat.		**9.21**		
J-Stat.				**21.72**
$\chi^2(15)$		**.56**		
$\chi^2(14)$				**.92**

and uninstrumented specifications shows that more than half of the coefficients are of greater magnitude in the crowding specifications. The by far biggest difference, however, is in the value of the constant. When adding the structural crowding term to the uninstrumented specification, the constant increases from -7 to -1.5. In the instrumented regressions, the constant even switches sign and moves from about -5 to .5.

To see how this impacts the role of product characteristics in explaining market shares, recall (3.2) and the fact that the constant enters the observable characteristics matrix in the computation of mean utility, $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$. Each product's market share is positively related to its mean utility, $\partial s_{jt}/\partial\delta_{jt} = s_{jt}(1 - s_{jt}) > 0$. Thus, all else equal, an increase in the constant term is equivalent to a rise in all product market shares. Given that the observed shares of each lager and the outside good are fixed, a rise in the constant increases the proportion of market shares and consumer utility that is explained by product characteristics. Looking back at the closed-form decomposition of consumer welfare, (3.4), it follows immediately that the proportion of consumer welfare explained by mean utilities rises, too.

Thus, my estimates indeed suggest that not accounting for crowding effects reduces the role of the estimated taste coefficients in explaining market shares and thereby consumer welfare.

I reinforce this point by computing consumer welfare using the decomposition in (3.4) for each date in the sample period. The top row of Figure 3.2 plots consumers' valuations of their choice set's size (left plot) and quality (right plot) using the estimated structural parameters from specification (II), which corresponds to a logit model. The bottom row repeats this exercise using the results from specification (IV), which corresponds to a pure congestion model.

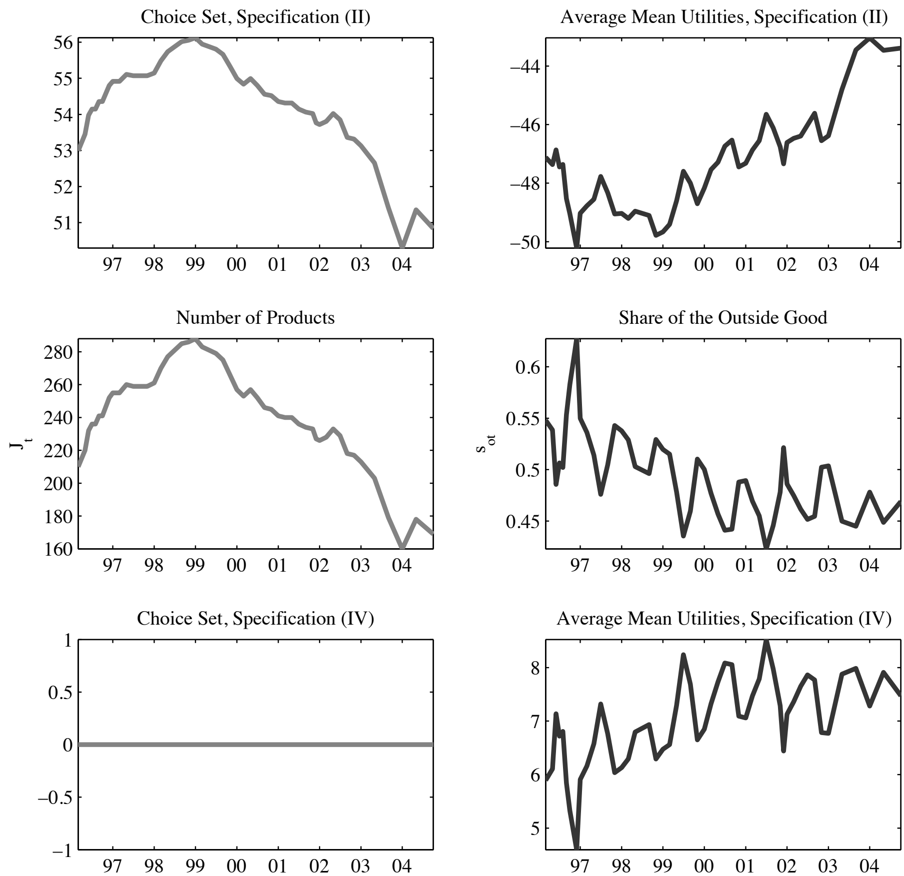
In the plots, I have set the arbitrary constant K to zero. As I mentioned before, it is not meaningful to interpret the level of each welfare component. Nevertheless, it is still informative to compare the role of the two components across the different specifications and to compare the time series of each component with that of the number of products and the share of the outside good.

As $\gamma = 1$ in the bottom row, changes in the number of products have no effect on welfare. In the top panel, however, we have the polar opposite case of $\gamma = 0$. Changes in the number of light lagers are fully passed on to consumer welfare. Accordingly, the plot of consumers' valuation of the choice set in the upper left panel is perfectly correlated with the number of products.

Looking at the valuation of average mean utility, we can see that for the pure congestion model this component of welfare is perfectly negatively correlated with the share of the outside good. For the standard logit model, this relationship seems to be weaker. Instead, changes in quality appear to be much more driven by changes in the number of products. In fact, the correlation between average quality and the number of products and the share of the outside good is $-.88$ and $-.67$. For the pure congestion model, corresponding numbers are $-.25$ and -1.

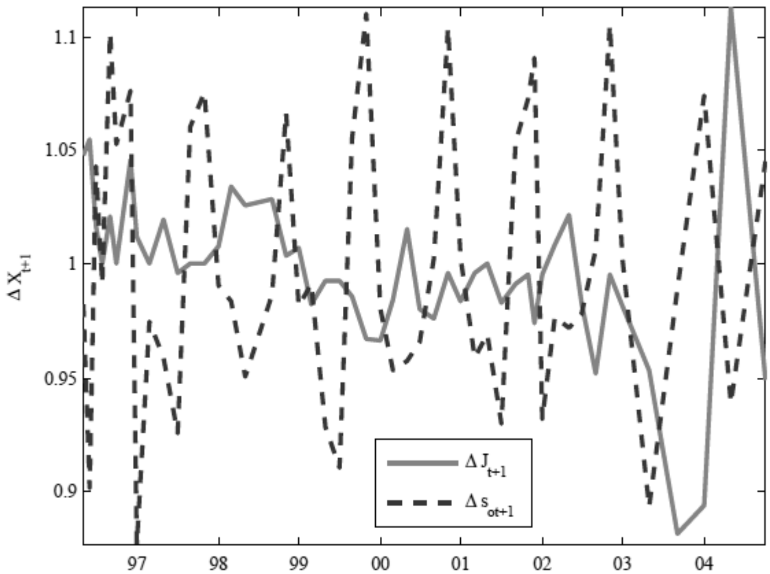
These findings are fully in line with the derivation of (3.5) in section 3.1. Equipped with this identity, we can now easily determine the correlation between net entry and consumer welfare in the logit and pure congestion model. Figure 3.3 plots the gross

Figure 3.2. The Number of Products, the Outside Good and the Components of Welfare During the Sample Period



rate of changes of the total number of inside products and the share of the outside good. As can be gauged from the plot, the two series tend to be negatively correlated, especially towards the end of the sample. The actual correlation between the series is −.19. We can thereby conclude immediately that for both specifications, net entry is positively correlated with total consumer welfare. For the instrumented crowding model, specification (IV), the actual correlation is .17, whereas for specification (II) the actual correlation is .18.

FIGURE 3.3. Net Entry and the Share of the Outside Good



It is tempting to take this story at face value, because intuitively, entry should raise competition and thereby increase the quality of all goods. As I have argued before, however, with the share of the outside good not having been determined in the demand estimation, it is potentially misleading to end the investigation at this point.

5. A Closer Look at Entry and Exit

To obtain a more reliable assessment of the welfare benefits of entry and exit, I compute the quality terms for the groups of entering, exiting and incumbent products separately. These terms are determined by the structural part of the demand model and not by the distributional assumption placed on consumers' taste shocks. Quality is therefore a more reliable assessment of each group of products' value to consumers. As in the previous section, I use the structural parameters from specification (II), the instrumented logit model and specification (IV), the instrumented crowding model.

Looking back at the estimation results in Table 3.3, it is clear that including the exiter and entrant dummies favors new product introductions over those lagers that exit the market in the following period.

TABLE 3.4. Average Mean Utility by Group

Including Unobservables, ξ

	Specification (II)			Specification (IV)		
	Incumbents	Entrants	Exiters	Incumbents	Entrants	Exiters
sample mean	.0199	.0076	.0043	4.0521	1.7605	1.0280
sample median	.0095	.0054	.0042	2.3517	1.3390	1.0117
correlation with J_t	-.3186	-.0523	-.1978	-.2991	.1662	.1664
correlation with s_{ot}	-.1904	-.2693	-.3749	-.1925	-.2167	-.3087

Excluding Unobservables, ξ

	Specification (II)			Specification (IV)		
	Incumbents	Entrants	Exiters	Incumbents	Entrants	Exiters
sample mean	.0021	.0001	.0001	.4796	.0150	.0119
sample median	.0020	.0000	.0000	.4686	.0099	.0096
correlation with J_t	-.4835	-.1363	-.3381	-.4730	-.1360	-.3343
correlation with s_{ot}	-.6053	-.2120	-.3810	-.6103	-.2033	-.3894

As I am interested in obtaining robust results regarding the benefits of entry and exit, I exclude both dummies. If entering lagers have a higher quality than exiting lagers without the dummies, this difference is only going to widen when including them. Moreover, to investigate the importance of the estimated unobservable product characteristics, I compute two sets of results, one including the error term and the other excluding it.

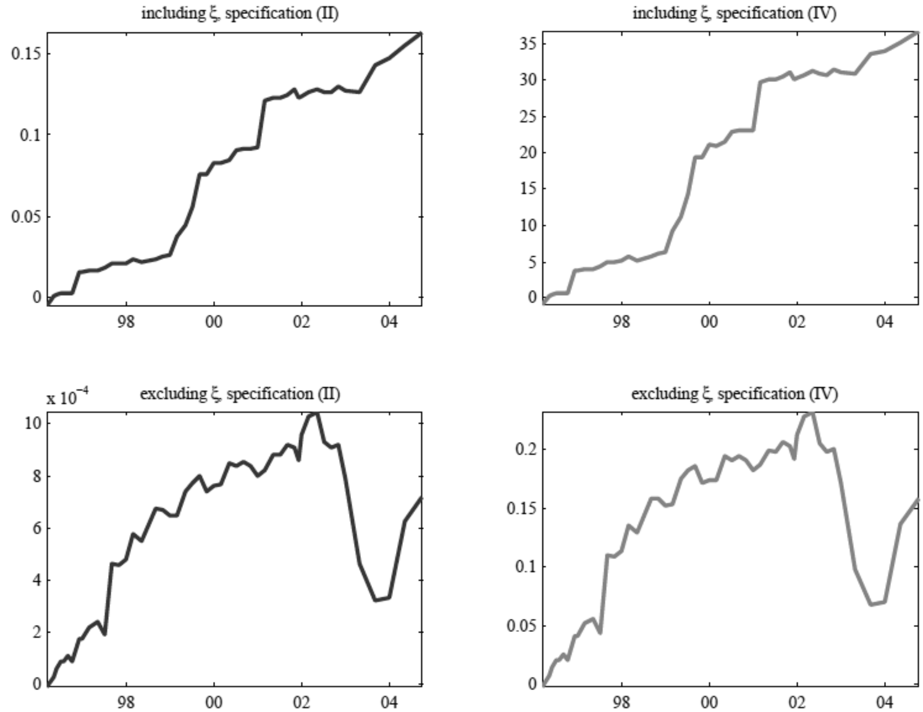
For the comparisons I define the group quality terms, $\widehat{\delta}_{gt}$ as follows.

$$\widehat{\delta}_{gt} = (J_{gt})^{-1} \sum_{j=1}^{J_{gt}} \exp(\widehat{\delta}_{jt})$$

J_{gt} is the total number of products in the specific group and the hat above the mean utilities is a reminder that $\widehat{\delta}_{jt} \neq \delta_{jt}$, because the entrant and exiter dummies and depending on the set of results the unobservable product characteristics, have been stripped out. I also ignore the outside good here to ensure that differences between the groups are driven by different qualities and not by the entrant and exiter groups being much smaller than the incumbent lagers group. Summing the $\widehat{\delta}_{gt}$ therefore does not yield $\overline{\delta}_t$.

Table 3.4 and Figure 3.4 tell the same story about the impact of entry and exit on welfare. On average, lagers entering the market have a higher quality than those exiting the market. This holds for both specifications. Moreover, the fact that the two specifications draw such a similar picture of the quality differences between entering

FIGURE 3.4. Cumulative Average Quality Differences Between Entering and Exiting Products During the Sample Period



and exiting lagers is particularly reassuring, because there are substantial differences between the logit and pure crowding estimates regarding the aggregate market's quality.

This impression is strengthened when looking at the correlations between the quality measures and the number of products and the outside good's market share. For the logit specification, the correlation between the number of products and average quality is weakest for the group of entering products. This holds both for the case where product unobservables are included and excluded and suggests that the estimated quality of entrants is not mainly driven by the observed changes in the number of products.

For the pure congestion estimates, the correlation between average quality and the share of the outside good is lowest for the group of entrants. By the same logic, this finding builds confidence in the result that entry raises consumer welfare by providing better lagers and is not simply the result of a spurious correlation between changes in the number of varieties and changes in the share of the outside good.

6. Conclusion

Using a structural demand model within the framework of Ackerberg and Rysman (2005), I find strong evidence for substantial crowding effects in the Swedish market for light lagers. The estimation results suggest that market shares are generated by a pure congestion model. In other words, the available choice set to consumers has become so wide that additional products are not valued *per se*. There is no longer a love-for-variety effect. Instead, consumer welfare can only be raised by improving the quality of the lagers in the market.

During the sample period, I find entry to be positively correlated with consumer welfare. I consider the possibility that this finding stems from the exogenous definition of the relevant market. As decreases in the share of the outside good imply an increase in the average quality and thereby a rise in consumer welfare in the pure congestion model, entry can simply be correlated with but not the cause of consumers switching from the outside alternative to one of the lagers in the market.

To address this concern, I assess the difference between the average mean utilities of entering and exiting products. Here, I use the structural part of the demand model, because it is less likely to be driven by the distributional assumption imposed on the logit taste shocks. I find that entering products tend to be more attractive to the representative consumer than exiting products and that this finding is stable throughout the sample period. Moreover, changes in the average mean utilities of entering products appear much less correlated with the share of the outside good, than this is the case for incumbent products and exiting lagers. This strengthens the robustness of the finding that entry and exit raise consumer welfare.

From a methodological point of view, the upshot of the paper is that it is desirable to jointly estimate the market share of the outside alternative with the structural demand parameters. This could alleviate concerns that entry and consumer welfare are spuriously correlated by the exogenous delineation of the relevant market. Huang and Rojas (2010) emphasize the potentially biased estimates of consumer demand resulting from exogenously delineating the share of the outside. Their suggested estimation methods for jointly estimating the share of the outside good with the structural demand parameters could complement my approach for assessing to which extent welfare gains from product introductions are driven by the functional form of the estimator.

7. References

Ackerberg, Daniel A. and Rysman, Marc, (2005), Unobserved product differentiation in discrete-choice models: Estimating price elasticities and welfare effects," *RAND Journal of Economics*, 36, issue 4, p. 771-88.

Bajari, P. and Benkard, C. Lanier, (2003), Discrete choice models as structural models of demand: Some economic implications of common approaches, *Manuscript*, Yale University.

Bajari, P. and Benkard, C. Lanier , (2005), Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach," *Journal of Political Economy*, 113, issue 6, p. 1239-76.

Berry, Steven T., (1994), Estimating discrete-choice models of product differentiation, *RAND Journal of Economics*, 25, issue 2, p. 242-62.

Berry, Steven T., Levinsohn, J. A., and Pakes, Ariel , (1995), Automobile prices in market equilibrium, *Econometrica*, 63, issue 4, p. 841-90.

Berry, Steven T. and Pakes, Ariel, (2007), The pure characteristics demand model, *International Economic Review*, 48, issue 4, p. 1193-225.

Berry, Steven T. and Waldfogel, Joel, (1999), Free entry and social inefficiency in radio broadcasting, *RAND Journal of Economics*, 30, issue 3, p. 397 - 420.

Goldberg, Penelopi K. and Hellerstein, Rebecca, (2011), A structural approach to identifying the sources of local currency price stability, *forthcoming in Review of Economic Studies*.

Goolsbee, A. and Petrin, Amil, (2004), The consumer gains from direct broadcast satellites and the competition with cable TV, *Econometrica*, 72, issue 2, p. 351-81.

Hoxby, Caroline and Paserman, M. Daniele , (1998), Overidentification tests with grouped data, *NBER Technical Working Paper* #223.

Huang, Dongling and Rojas, Cristian, (2010), Eliminating the Outside Good Bias in Logit Models of Demand with Aggregate Data, Manuscript, University of Massachusetts Amherst.

Mariuzzo, F., Walsh, P. P., and Whelan, Ciara , (2010), Coverage of retail stores and discrete choice models of demand: Estimating price elasticities and welfare Effects, *International Journal of Industrial Organization*, 28, issue 5, p. 555 - 78.

McFadden , D., (1981), Econometric models of probabilistic choice, in C.F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press.

Nevo, A., (2000), A practitioner's guide to estimation of random coefficients logit models of demand, *Journal of Economics & Management Strategy*, 9, issue 4, p. 513-48.

Petrin, A., (2000), Quantifying the benefits of new products: The case of the minivan, *Journal of Political Economy*, 100, issue 4, p. 705-29.

Rysman, M., (2004), Competition between networks: A study of the market for yellow pages, *Review of Economic Studies*, 71, issue 2, p. 483-512.

Slade, M., (2004), Market power and joint dominance in U.K. brewing, *Journal of Industrial Economics*, 52, issue .1, p. 133-63.

Small, K.A. and Rosen, H.S., (1981), Applied welfare analysis with discrete choice models, *Econometrica*, 49, issue 1, p. 105-30.

8. Appendix

TABLE 3.5. Regressing the Excluded Instruments on Price

Variable	Coefficient	$P > \|t\|$	Variable	Coefficient	$P > \|t\|$
z_1	44.58	.071	z_{10}	-.02	.321
z_2	60.20	.000	z_{11}	-.20	.068
z_3	-54.79	.000	z_{12}	-.25	.045
z_4	-96.06	.001	z_{13}	.17	.005
z_5	3.66	.700	z_{14}	.43	.001
z_6	-.11	.317	z_{15}	-.04	.115
z_7	-.24	.000	z_{16}	-1.80	.000
z_8	.08	.193	constant	522.19	.000
z_9	.09	.481			
Observations	12,080				
F(16,12063)	204.53				
$P > F$	.000				
R^2	.213				
adjusted R^2	.212				

TABLE 3.6. Regressing the Excluded Instruments on Price

Variable	Coefficient	$P > \|t\|$	Variable	Coefficient	$P > \|t\|$
z_1	47.22	.009	z_{15}	-.04	.011
z_2	82.61	.000	z_{16}	-.84	.000
z_3	-53.20	.000	x_2	.14	.000
z_4	-67.20	.002	x_3	-.11	.006
z_5	5.37	.158	x_4	.03	.318
z_6	-.16	.047	x_5	2.03	.000
z_7	-.36	.000	x_6	-.06	.242
z_8	.13	.004	x_7	-10.06	.000
z_9	.13	.149	x_8	-7.96	.000
z_{10}	-.02	.217	x_9	-13.10	.000
z_{11}	-.20	.011	x_{10}	-.58	.000
z_{12}	-.35	.000	x_{11}	0	.990
z_{13}	-.17	.000	x_{12}	4.28	.000
z_{14}	.30	.001	constant	292.47	.000
Observations	12,080				
F(24,12553)	619.97				
$P > F$	.000				
R^2	.581				
adjusted R^2	.581				

PAPER 4

Does Treasury Bond Supply Explain the Term Structure of Treasury Yields?

André Romahn

Abstract

The relative supplies of Treasury securities along the term structure matter for the dynamics of the yield curve. A no-arbitrage affine term structure model based solely on these supply measures captures the average shape of the Treasury yield curve well. The pricing errors decrease with the maturity of the bonds. A combination of yield and supply factors attains lower pricing errors for bonds with maturities beyond twenty years than does the canonical affine model with the observable level, slope and curvature of the yield curve. Treasury supply factors therefore contain relevant bond pricing information that is not subsumed by these three yield curve factors.

JEL Classification: E43, G12.

Keywords: term structure, affine yield curve model, Treasury securities.

I wrote part of this paper during a stay at the ECB's Division for Monetary Policy Strategy. I gratefully acknowledge the Division's hospitality and emphasize that the views expressed in this paper are mine and do not necessarily reflect the views of the ECB.

1. Introduction

In efficient financial markets with rational market participants, asset quantities do not impact asset prices. A bond's price, for example, is simply a reflection of the present value of all future coupon payments during the remaining life of the bond. Both recently and in the past, however, actual policy making has attempted to affect asset prices by changing the quantities of asets.

Several central banks have intervened openly in bond markets with the intent of lowering yields by creating additional demand for government bonds. The Federal Reserve has rediscovered a policy experiment from the Kennedy era. The Treasury attempted to twist the yield curve ("Operation Twist") by purchasing long-term Treasury bonds and selling short-term Treasuries. By making long-term bonds relatively scarce, the Fed aims at raising the prices of such bonds and thereby lowering their yields. This in turn decreases the long-term financing costs of the private sector and thereby attains an expansionary effect. Modigliani and Sutch (1967) analyze the Kennedy administration's "Operation Twist" and in line with efficient markets find no significant effects.

The findings in several recent papers, however, indicate that this is not the whole story. Greenwood et al. (2010) uncover evidence that the private sector creates substitute liquidity if the public sector retrenches from specific maturity segments in the bond market. Thereby, the maturity of corporate bond issues is significantly negatively correlated with the maturity of outstanding government debt in the United States (maturity gap-filling). As the government tilts its funding towards the short end, for example, corporates issue relatively more bonds at the long end of the yield curve. If the outstanding volume of long-dated Treasury securities impacts their price, firms are incentivized to provide substitute liquidity, because they can borrow at lower interest rates. Krishnamurthy and Vissing-Jorgensen (2010) present similar findings. They show that Treasuries have money-like features in that they are very liquid and safe. When their supply falls, Treasury yields are reduced and the supply of bank-issued money rises. This is exactly the mechanism that "Operation Twist" attempts to exploit to lower long-term rates.

Taking these findings on the relationship between bond prices and quantities at face value, the question is what the possible underpinnings for these effects might be. Modigliani and Sutch (1967) argued that investors have preferred habitats along the term structure. That is, some investors have strong preferences for long-dated bonds while others prefer short-dated bonds. This yields a partial segmentation of the yield curve and allows the outstanding volume of Treasuries in each habitat to influence the prices of these bonds.

Insurance companies and pension funds are prime candidates for having strong preferences for long-dated bonds. Faced with very long-term liabilities, long-dated Treasuries are the ideal hedging asset to close the maturity gap between the asset and liability sides of such a financial institution's balance sheet. In contrast, commercial banks typically have very short term liabilities as bank customers can withdraw their deposits at any time.

To the best of my knowledge, however, there is no generally accepted framework that explicitly models the interaction of these two types of financial institutions on the bond markets in an attempt to rationalize the impact of asset quantities on asset prices. Greenwood et al. (2010) and Krishnamurthy and Vissing-Jorgensen (2010), instead, follow the tradition of the money-in-the-utility approach and make the assumption that the representative household gains utility from holding Treasuries. Vayanos and Vila (2009) exogenously assume that bond market clienteles have different preferences for maturities and model their interaction with risk-averse arbitrageurs.

The lack of a widely accepted general equilibrium model, however, is no obstacle to investigating the relationship between Treasury quantities and prices more closely. I use the framework of Ang et al. (2006) to model the dynamics of the U.S. Treasury yield curve in a vector autoregression (VAR) framework. This approach has the advantage that the VAR can parsimoniously summarize the joint dynamics of the state variables, which in turn explain bond yields at different maturities. Moreover, the affine yield curve model imposes no-arbitrage conditions on bond yields and thereby enforces economically meaningful outcomes. This is not necessarily the case in the approaches of Greenwood et al. (2010) and Krishnamurthy and Vissing-Jorgensen (2010) who do not make these restrictions. Their results could therefore be partly driven by misspecification of their estimating frameworks.

In this sense, this paper complements their efforts by testing the relevance of Treasury security quantities for the yield curve in a more rigorous setting. I use the share of each Treasury maturity offering of the total market value of all outstanding Treasury securities to measure the relative scarcity of bonds with different maturities. An increase in the share of bonds with maturities of up to one year, for instance, necessarily brings about a fall in the shares of bonds with longer lives. More than 90 percent of the total variation of these shares can be summarized by three principal components.

I then compare the ability of an affine yield curve model that employs the Treasury supply factors to price bonds at all available maturities with that of the canonical affine model that uses the three observable yield curve factors level, slope and curvature. The latter factors are well-known to fit the yield curve with very low pricing errors (Nelson and Siegel (1982), Svensson (1994), Ang and Piazzesi (2003), Ang et al. (2006)). Even though the relative measures of Treasury supplies can capture the average shape of the Treasury yield curve well, the implied pricing errors are larger than those of the canonical model. The interesting aspect, however, is that the pricing errors generated by the supply factor model decrease with maturity. This suggests that the supply factors can contain relevant economic information that is not contained in the level, slope and curvature of the yield curve. I confirm this finding by estimating a hybrid model that uses a combination of the supply and yield curve factors. For long-dated bonds, the implied pricing errors are smaller than those of the canonical model and for the remaining maturities, the deviations of fitted yields from observed yields are similar.

In the next section, I describe the data sources and the sample period in detail, before presenting the no-arbitrage affine modeling framework in Section 3. The estimation procedure and its results are presented in Section 4. Section 5 concludes.

2. Data and Descriptives

I use data from two sources: the zero-coupon yield curve data estimated in Gürkaynak et al. (2006) for bonds with maturities ranging from one to thirty years and the outstanding face value data for all issued U.S. Treasury bonds from the CRSP Monthly Treasury Master File.

The sample covers the period from November 1985 to December 2007 at a monthly frequency. As I want to assess the relevance of bond supply factors for Treasury bond yields of all available maturities, I cannot extend the coverage of the sample period further into the past. Gürkaynak et al. (2006) note that the reported yields for thirty-year bonds before November 1985 seem unreliable and do not report estimated zero-coupon yields for these long-term bonds before that date. Another factor to take into account is that the issuance policy of the Treasury should remain stable over the sample period. Garbade (2007) demonstrates that before 1982 the Treasury attempted to lower the interest costs of market borrowing by "tactical" offerings of Treasury securities that were not fully anticipated by market participants. Experience showed, however, that such attempts of timing bond markets raised borrowing costs by introducing an additional source of uncertainty. From 1982 onwards, the Treasury therefore relied on a "regular and predictable" publicly announced schedule of Treasury security offerings. This should ensure that the estimated issuance policy of Treasuries is stable over the sample period.

Similarly, I set the end of the sample period to avoid including the Financial Crisis. During the Crisis, the comovement between returns on assets increased substantially and the maturity profile of Treasury bond issues underwent some abrupt changes.

Including such periods with structural breaks in the relationships between yields and Treasury supply factors could substantially bias the estimated joint dynamics of the state variables away from "normal" market conditions.

Another relevant point regarding the yield curve data is that the reported zero-coupon yields by Gürkaynak et al. (2006) are computed using the flexible functional form of Svensson (1994), which is an extension of the original model by Nelson and Siegel (1987). The yields, therefore, are the outcomes of a smoothing process and are not identical to the unsmoothed Fama-Bliss zero-coupon yields, that are most commonly used in the existing literature. The drawback of the Fama-Bliss yields is that they are only available for maturities of up to five years. Moreover, Bliss (1996) compares several methods of extracting zero-coupon yields from raw yield data and he finds that even though the unsmoothed Fama-Bliss yields generally perform best, the generalized Nelson-Siegel method still performs well. Similarly, Cochrane and Piazzesi (2009) note that the differences between the Gürkaynak et al. (2006) and unsmoothed Fama-Bliss yields are "small on most dates".

As a reliable measure for the total volume of Treasury bonds, notes and bills outstanding at each point in time, I use the CRSP Treasury database, which contains information on virtually every Treasury security that has been issued since 1925. I only include the volumes of noncallable notes, bonds and bills to match the volume data closely to the yields provided in Gürkaynak et al. (2006). Along with the face value of each treasury security, the database also contains information on the prices of each bond during its life. This allows me to measure available market volumes at market prices.

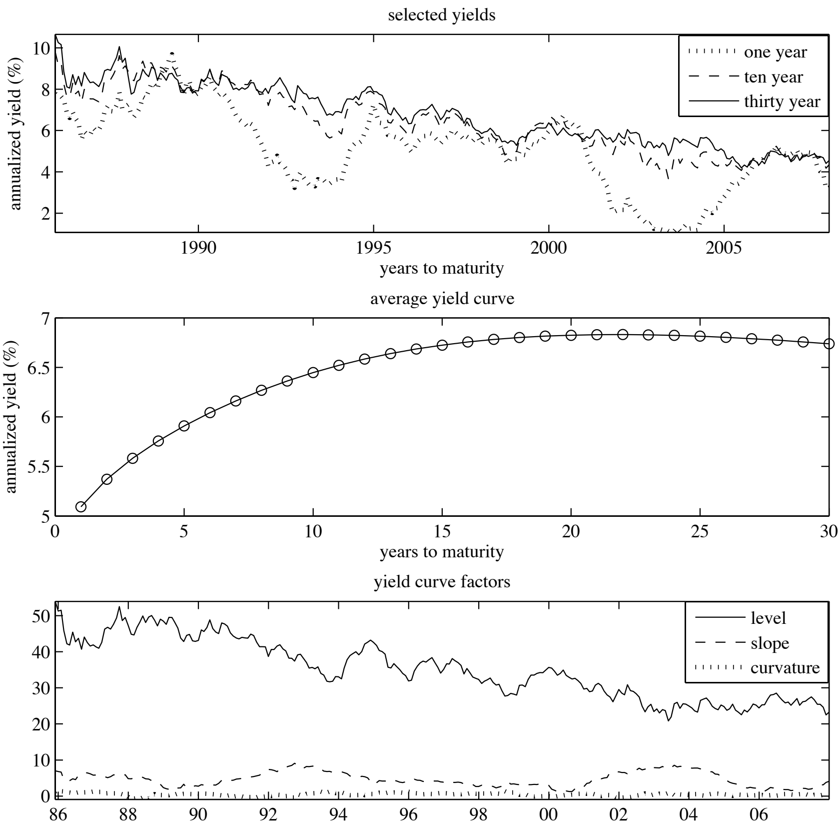
Figures 4.1 and 4.2 give an impression of the two data sources. The top panel of Figure 4.1 plots the one-, ten- and thirty-year yields over the sample period. The general downward trend in yields reflects the Fed's success in reducing inflation ("Great Moderation"). The volatility of yields decreases with the time to maturity; the thirty-year yield being the least volatile. Moreover, the rapid drops in the one-year yield in 1991 and in 2001 are due to the Fed's policy response to recessions following these dates.[1] The middle panel plots the average yield curve. As is well established in the stylized facts of the existing literature, the yield curve is on average upward sloping.

To give an impression of the supply of Treasury bonds over the maturity spectrum, I plot the volume-weighted maturity of all non-matured Treasury bonds along with the share of all active Treasuries for several maturity ranges.[2] The initial increase in

[1] The NBER dates the peak of economic activity at July 1990 and the corresponding trough at March 1991. For 2001, the peak is dated in March, while the trough is located in November.

[2] I use the terms active and non-matured interchangeably to denote all Treasury securities at a given date t that have been issued before that date and mature after date t.

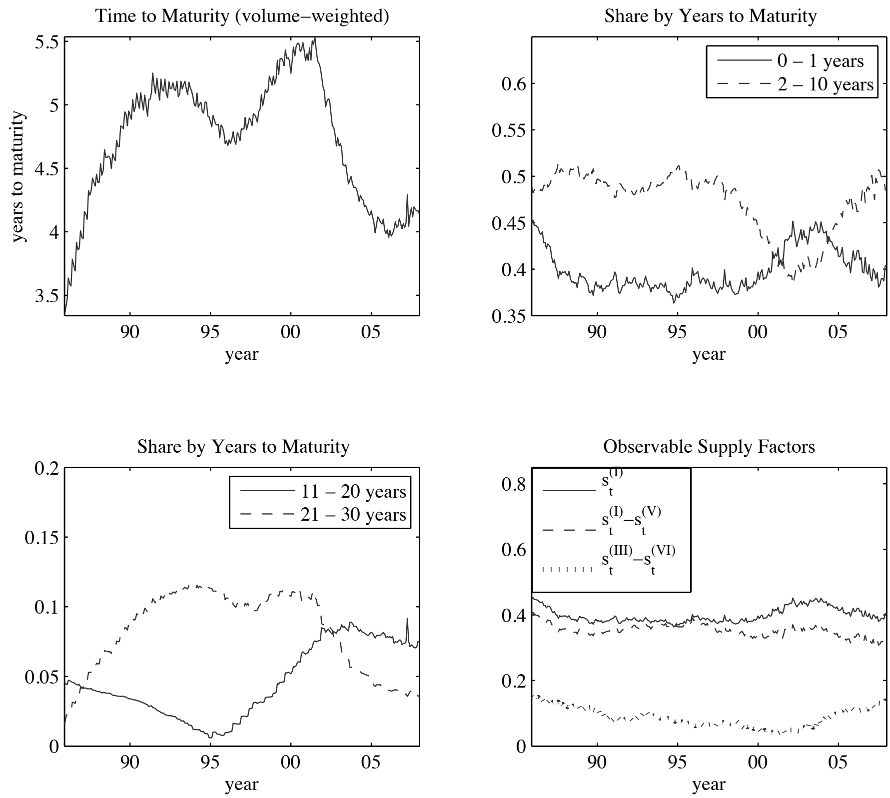
FIGURE 4.1. The Treasury Yield Curve from November 1985 to December 2007



the maturity stems from a rise in the share of bonds with maturities of at least twenty years. The sharp drop in the volume-weighted maturity around 2001 reflects the public surpluses of the Clinton administration. During that time, the Treasury did not fully replace maturing longer term bonds, but instead shifted issuance towards the very short end of the maturity spectrum.

For fitting the yield curve with an affine model, the information that is contained in yields and outstanding volumes of Treasury securities must be summarized in a few meaningful factors. Simply including thirty yields and estimating a VAR to summarize their joint dynamics requires the estimation of more than nine hundred parameters. The same holds for running a VAR for the outstanding volumes of all thirty maturities.

FIGURE 4.2. Characteristics of Treasury Securities from November 1985 to December 2007



The most widely accepted way of solving this problem is to summarize the information contained in yields by extracting their principal components (Nelson and Siegel (1982), Svensson (1994), Ang and Piazzesi (2003), Ang et al. (2004)). The top panel of Table 4.1 replicates the well-established fact that three factors can explain almost all of the variation in bond yields. These extracted factors correspond well with directly observable yields or combinations of yields. The first factor is almost perfectly negatively correlated to the thirty-year yield[3], while the second and third factors comove

[3] Most papers report a correlation of this magnitude for the yield corresponding to the shortest maturity included in the sample. For my sample the correlation between the first factor and the one-year yield is -.82, which is still substantial but less than that for the thirty-year yield. The difference is most likely due to the fact that I include all maturities, while Ang et al. (2006) for instance only look at Treasuries with a maturity of up to five years.

TABLE 4.1. Extracted Factors and Their Observable Counterparts

Yield Curve Factors

	Factor				
	1	2	3	4	5
Cumulative Share of Total Variation	.9391	.9964	.9990	.9998	1.0000

	Factor		
Correlation of Factors with Observables	1	2	3
thirty-year yield, $y^{(30)}$	-.9667	.1369	.4518
term spread, $y^{(30)} - y^{(1)}$	.2290	.9824	-.1616
curvature, $y^{(1)} - 1.9y^{(5)} + y^{(30)}$	.1054	.0117	.8080

Treasury Securities Supply Factors

	Factor				
	1	2	3	4	5
Cumulative Share of Total Variation	.6573	.8693	.9328	.9906	1.0000

	Factor		
Correlation of Factors with Observables	1	2	3
$-s_t^{(I)}$	.8704	.1918	.2486
$s_t^{(I)} - s_t^{(V)}$	-.1888	.9114	-.0595
$s_t^{(VI)} - s_t^{(III)}$	-.4237	-.2255	.8432

tightly with the term spread and the curvature of the yield curve. The ability of only a few factors to explain nearly all of the joint variation of yields is driven by the fact that yields in general and especially yields of neighboring maturities are strongly correlated.

For the outstanding volumes of Treasury securities, it is also the case that the volume of neighboring maturities are highly correlated. Nevertheless, I cannot simply extract factors from the nominal volumes, because they share a common trend over time. As output grows, so typically does the nominal value of public debt. In other words, bond volumes are not stationary time series. Greenwood et al. (2010) and Krishnamurthy and Vissing-Jorgensen (2010) tackle this problem by normalizing bond volumes by GDP. Data on output, however, is not available at a monthly frequency. Instead of interpolating GDP data in-between quarters, I simply compute each maturity's share of the total volume of outstanding Treasury securities. As the plots in Figure 4.2 suggest, the resulting series are stationary and the shares can be interpreted as a measure of the relative scarcity of specific maturities.[4] An increase in the share of bonds with maturities of up to one year, necessarily brings about a fall in the shares of bonds with longer lives.

Moreover, I also define six maturity brackets before extracting the principal components from the series: 0 to 1, 1 to 3, 3 to 7, 7 to 10, 10 to 20, and 20 to 30 years. This classification of maturity brackets mirrors the Treasuries' issuance policy.[5] Looking at the bottom panel of Table 4.1, the first three Treasury securities supply factors explain more than 93 percent of the total variation in the maturity bracket shares. Similar to the case of yields, the extracted factors are highly correlated with directly observable shares or linear combinations of shares. To avoid confusion with the notation for bond maturity, I refer to the specific maturity brackets by Roman numerals. The first factor, for example, exhibits a -.87 correlation with the share of Treasury securities with a maturity of at most one year, $s_t^{(I)}$.

The most relevant question is of course how well the observable supply factors explain the zero-coupon yield curve data. As a first pass at this issue, I compute the R^2's of simple predictive OLS regressions for each of the yields. Thus, I estimate the regression equation

$$y_{t,t+k}^{(n)} = \alpha + X_t^i \beta_i + \epsilon_t, \quad i = \{s, y\}, \tag{4.1}$$

where k is measured in months and X_t^i contains the observable factors corresponding to yields ($i = y$) or Treasury security supplies ($i = s$).

$$X_t^y = \begin{bmatrix} y_t^{(30)} \\ y_t^{(30)} - y_t^{(1)} \\ y_t^{(1)} + y_t^{(30)} - 1.9 y_t^{(5)} \end{bmatrix} \tag{4.2}$$

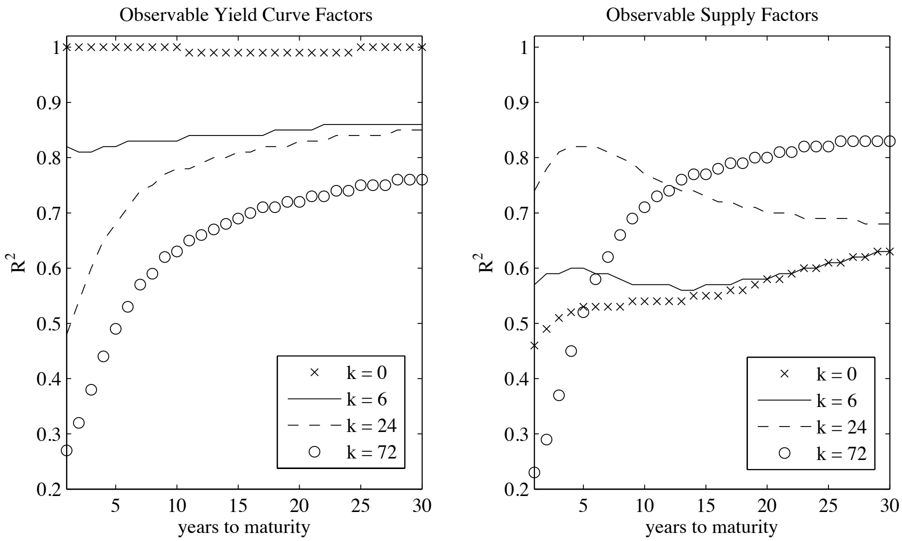
$$X_t^s = \begin{bmatrix} -s_t^{(I)} \\ s_t^{(I)} - s_t^{(V)} \\ s_t^{(VI)} - s_t^{(III)} \end{bmatrix} \tag{4.3}$$

Figure 4.3 plots the resulting R^2's for the two sets of observable factors. In the left panel, we can gauge how well the observable yield curve factors can explain each of the yields along the maturity spectrum. Recalling that the three factors explain more than 90 percent of the total variation of yields, it is not surprising that the contemporaneous

[4] I confirmed this in a more formal sense by estimating a VAR describing the joint dynamics of the states. The estimated coefficients and the implied impulse response functions show that the system returns to the stationary state after it has been perturbed.

[5] Treasury notes and bonds are at this time offered with maturities of 2, 3, 5, 7, 10 and 30 years. There are also occasional offerings of Treasury bonds with maturities of 15 and 20 years. Treasury bills have a maximum life of 1 year.

FIGURE 4.3. OLS Predictive Yield Regressions



fit ($k = 0$) produces R^2's of close to 1 for all the maturities. Increasing the horizon produces lower R^2's, especially for bond yields with shorter maturities. Given that these yields are more volatile than yields of bonds with longer lives, this result is intuitive. This could also be a driver for the outcome that the fit generally improves with the maturity of the zero-coupon bond. The Treasuries supply factors produce a worse contemporaneous fit than the observable yield curve factors, but still explain more than 50 percent of the variation across all maturities. Given the disconnect between asset prices and quantities in frictionless financial markets with rational agents, this fit is surprisingly good. Moreover, it turns out that for intermediate values of k, the observable supply factors contain more relevant information for the yields of bonds with shorter maturities. This suggests that the supply factors contain information that is not captured by the yield curve factors. Similarly, for the longest horizon ($k = 72$), the supply factors explain the yields of bonds at the very long end of the term structure better than do the observable yield factors.

3. Modeling the Yield Curve

I follow Ang et al. (2006) and model yields at any date t as affine functions of K observable states, X_t. The joint dynamics of the states are summarized by a first-order vector autoregression.

$$X_t = \mu + \Phi X_{t-1} + \Sigma \epsilon_t \tag{4.4}$$

I assume that the innovations to (4.4) are distributed standard normal, $\epsilon_t \sim N(0, I)$. Σ is a $(K \times K)$ lower triangular matrix that allows for arbitrary contemporary correlations between the innovations to the states. μ is a $(K \times 1)$ vector of constants and Φ is a $(K \times K)$ matrix that collects the autoregressive coefficients.

The stochastic discount factor is affine in the state variables and given by

$$m_{t+1} = \exp\left(-y_t^{(1)} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}\right), \tag{4.5}$$

where λ_t contains the market prices of risk associated with each of the states and is also a linear function of X_t.

$$\lambda_t = \lambda_0 + \lambda_1 X_t \tag{4.6}$$

λ_0 is a $(K \times 1)$ vector and λ_1 is a $(K \times K)$ matrix of coefficients. Estimating the prices of risk in this form allows bond risk premia to be time-varying.

The price of an n-year to maturity bond can then be obtained by using the fact that a bond paying off instantaneously must satisfy $p_t^{(0)} = 1$ and the following no-arbitrage pricing relationship.

$$p_t^{(n)} = E_t(m_{t+1} p_{t+1}^{(n-1)}) \tag{4.7}$$

Given that risk premia are exponentially affine in the state vector, bond prices are conditionally log-normal functions of the observable states.

$$p_t^{(n)} = \exp(A_n + B_n' X_t) \tag{4.8}$$

Detailed derivations of the coefficients A_n and B_n that solve (4.8) can be found in the appendices of Ang and Piazzesi (2003) and Cochrane and Piazzesi (2005). Given that the solution satisfies (4.7), the resulting bond prices satisfy no-arbitrage.

$$A_{n+1} = A_n + B_n'(\mu - \Sigma\lambda_0) + \frac{1}{2}B_n'\Sigma\Sigma' B_n$$
$$B_{n+1} = (\Phi - \Sigma\lambda_1)' B_n - e_1$$

A_n is a $(K \times 1)$ vector, B_n is a $(K \times K)$ matrix and e_1 is the $(K \times 1)$ vector with its first element being equal to one and all other elements set to zero. Applying logs to (4.5) and annualizing yields gives the yield-fitting functions for each of the maturities.

$$y_t^{(n)} = -\frac{A_n}{n} - \frac{B_n' X_t}{n} \tag{4.9}$$

This modeling strategy implicitly assumes that there are no structural breaks in the joint dynamics of the observable state variables.[6]

As state variables, I am going to use the two sets of observable factors, X_t^y and X_t^s, corresponding to the observable yield curve and Treasury supply factors described in the previous section, respectively.

Looking back at (4.9) it is clear that some of the yields that are fitted by the model are directly observed. To ensure that (4.9) holds for all yields, I have to impose that the yields that are directly included in the set of observable yield curve factors are fitted without error. Thus, when using the observable yield curve factors, I impose the following coefficients: $A_1 = 0$, $B_1' = [1, 1, 0]$, $A_5 = 0$, $B_5' = [-10/1.9, 5/1.9, 5/1.9]$, $A_{30} = 0$ and $B_{30}' = [-30, 0, 0]$. For the second set of observable states, I obviously do not impose any restrictions on the estimated coefficients. Nevertheless, as bond prices are solved recursively, I have to estimate the first or last coefficients of A and B directly to obtain the remaining coefficients by forward or backward induction. I choose to estimate A_1 and B_1 directly.

I next present the two-step estimation procedure.

4. Estimation

I follow the two-step estimation strategy of Ang et al. (2006). In the first stage, I estimate the parameters determining the joint dynamics of the state variables, which are μ, Φ and Σ, with a VAR. In the second stage, the parameters determining the time-varying market prices of risk are estimated. This step differs, depending on whether observable yield curve factors or observable Treasury supply factors are used. For the latter case, I estimate the additional parameters contained in A_1 and B_1 directly, while in the former restrictions on A_1, B_1, A_5, B_5, A_{30}, B_{30} are imposed *a priori*. Thus, for the model using observable yield curve factors, the parameters to be estimated are $\Theta^y = (\mu, \Phi, \Sigma, \lambda_0, \lambda_1)$, while for the model employing observable Treasury supply factors, the estimated structural parameters are $\Theta^s = (\mu, \Phi, \Sigma, \lambda_0, \lambda_1, A_1, B_1)$.

The parameters obtained in the second stage minimize the total sum of squared yield-fitting errors.

$$\min \sum_t \sum_n (\widehat{y}_t^{(n)} - y_t^{(n)})^2$$

[6] To avoid including time periods that are characterized by such shifts in the relationship between the state variables, I focus on the period of November 1985 to December 2007. This sample period excludes the disinflationary policy shift of the Federal Reserve at the beginning of the 80s and it also excludes the Financial Crisis, that erupted in September 2008 with the bankruptcy of Lehman brothers.

Statistically, this approach is not efficient, but consistent. The efficient alternative is to jointly estimate all of the parameters by maximum likelihood, as is done in for example Ang and Piazzesi (2003). Efficiency in this setting comes at a price, however, because the large number of parameters to be estimated yields a nonlinear likelihood surface with "flat patches" that make the identification of the global maximum difficult. This problem has been noted in the existing literature (Ang and Piazzesi (2003), Kim (2008)) and Hamilton and Wu (2012) show that several popular representations of the canonical affine model are unidentified. Estimating λ_0 and λ_1 by minimizing the sum of squared fitting errors produces more reliable and robust parameter estimates by circumventing the need to find the global maximum of a highly nonlinear likelihood surface.

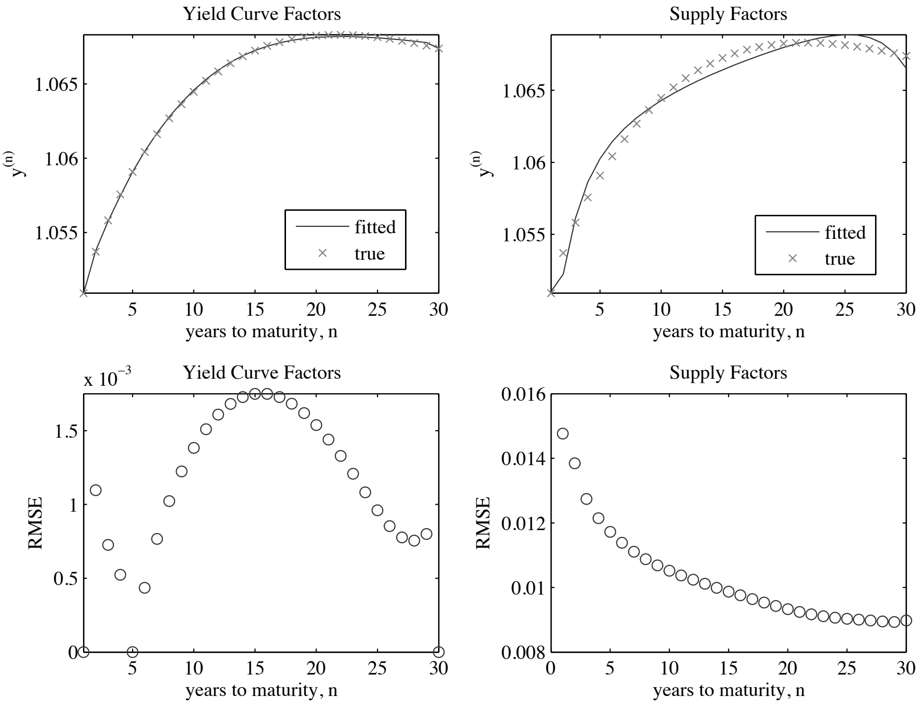
The standard errors for the consistent parameter estimates can be obtained by stacking the moment equations from the two steps of the estimation procedure and using GMM. Detailed derivations can be found in the Appendix of Ang et al. (2006).

4.1. Results. Figure 4.4 illustrates how well the two models perform in terms of fitting the yield curve data. All tables presenting the parameter estimates are relegated to the Appendix. The left-hand side panels show the fit for the specification using the observable yield curve factors, X_t^y. I refer to this model as the benchmark. The right-hand side panels illustrate the results for the model employing the observable supply factors, X_t^s. The average fitted yield curve implied by the benchmark model is practically indistinguishable from the actually observed average term structure. The supply factor model captures the shape of the average yield curve reasonably well, but differences between the true and fitted curve are apparent. This can also be seen from the scales of the plots in the bottom panels. The root mean squared errors (RMSEs) for the supply factor models are about an order of magnitude larger than the benchmark model. The RMSEs for the supply factor model are clearly decreasing with maturity. The average fitting error for the one-year yield is about 64 percent larger than that of the thirty-year yield.

For the benchmark model, the pattern is not as obvious. It is apparent that restricting the one-, five-, and thirty-year yields to be fitted without error also reduces the RMSEs of the neighboring yields. The way the fitting errors line up along the term structure could therefore be driven to some extent by the selection of yields that are included in the state variables.

Establishing that the benchmark model delivers a considerably better fit of the term structure during the sample period, however, is not surprising, because the plots in Figure 4.3 and the factor analysis of Table 4.1 already established that the explanatory power of yield factors for the term structure at a given date, is greater than that of

FIGURE 4.4. Comparing the Models

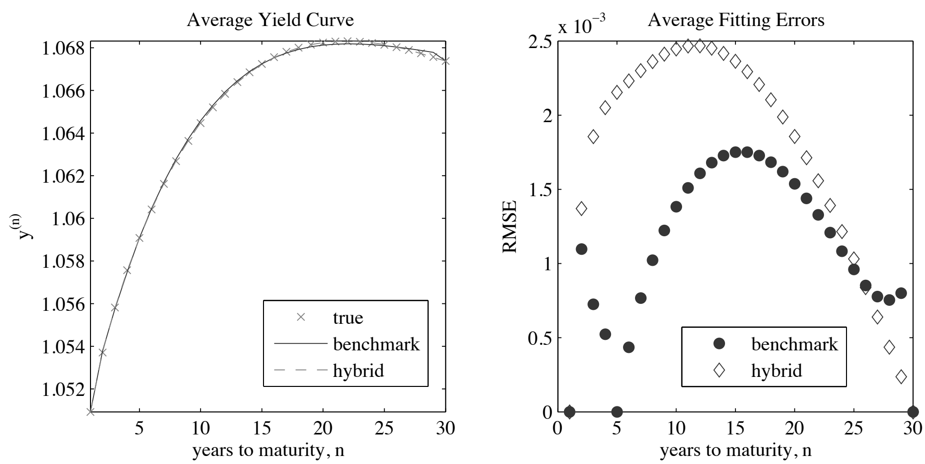


the supply factors. The relevant question is, whether the information contained in the yield factors subsumes the informational content of the observable supply factors. Putting it differently, do the supply factors add economically relevant information for understanding the dynamics of the yield curve?

I can only answer this question by estimating a "hybrid" model that contains both observable yield and supply factors. I collect the first two yield and supply factors in the hybrid state vector, X_t^h.

$$X_t^h = \begin{bmatrix} y_t^{(30)} \\ y_t^{(30)} - y_t^{(1)} \\ -s_t^{(I)} \\ s_t^{(I)} - s_t^{(V)} \end{bmatrix} \tag{4.10}$$

FIGURE 4.5. Comparing the Benchmark and Hybrid Models



Taken together, the observable states should account for more than 90 percent and more than 80 percent of the total variation in yields and Treasury supplies, respectively.[7]

Figure 4.5 compares how the hybrid model fares in comparison to the benchmark. In terms of the average fitted yield curve, the two specifications yield almost identical plots that match the actually observed average yield curve well. The right-hand side plot of the average fitting errors reveals differences between the two models. The hybrid model seems to price Treasury securities of maturities between five and twenty years with less precision than the benchmark model. As I argued before, however, this could at least partly be driven by the fact that in the benchmark model the five-year yield is measured without fitting error, which also tends to reduce the fitting errors of yields with similar maturities. Both specifications impose that the thirty-year yield is fitted without error, making the RMSEs for long-term bonds more readily comparable. As in the supply factor model, the RMSEs of the hybrid specification are decreasing with maturity. At the long end of the term structure, the hybrid model even attains lower pricing errors than the benchmark.

[7] I do not include all the observable factors here, because of the number of parameters to be estimated. I have a total of 266 monthly observations. If I include six state variables, I am estimating a total of 84 parameters. Only including the four state variables reduces this number to 40.

This indicates that the observable Treasury supply factors contain useful information for bonds with maturities beyond twenty years that is not subsumed in the three observable yield curve factors.

5. Conclusion

Using the Treasury zero-coupon yields estimated by Gürkaynak et al. (2006) and data on the outstanding volumes of Treasury securities for the entire term structure, I investigate if the relative supplies of Treasuries with different maturities can explain the dynamics of the term structure. A no-arbitrage affine term structure model based solely on these supply measures does not fit the yield curve as well as the popular benchmark model with the observable level, slope and curvature of the yield curve. Nevertheless, the supply factor model manages to capture the average shape of the term structure well and the bond pricing errors decrease with the maturity of the bonds. This indicates that the supply measures contain relevant economic information that is not subsumed by the observable yield curve factors. The fact that a hybrid model attains lower pricing errors for bonds with maturities beyond twenty years supports this finding.

This result complements and strengthens the findings of recent papers that find a significant impact of the supply of Treasuries on the level of bond yields (Greenwood et al. (2010), Krishnamurthy and Vissing-Jorgenson (2010)). The authors of these papers do not impose no-arbitrage on bond yields and thereby risk that their findings are driven by functional misspecification.

The results I present are tentative in the sense that it is unclear what underlying informational content of Treasury supplies is useful for pricing government bonds. This lack of interpretability is common in the literature using (no-arbitrage) affine term structure models, however. In this regard, it seems worthwhile for future research to put more structure on the joint dynamics of yield factors and Treasury supplies to allow for more direct economic interpretability.

6. References

Ang, Andrew, Piazzesi and Monika, (2003), A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables, Journal of Monetary Economics, 50, p. 745 - 787.

Ang, Andrew, Piazzesi, Monika and Wei, Min, (2006), What does the yield curve tell us about GDP growth?, *Journal of Econometrics*, 131, issue 1-2, p. 359-403.

Bliss, Robert R., (1996), Testing term structure estimation methods, *Advances in Futures and Options Research*, 9, p. 197-231.

Center for Research in Security Prices, (2011), *US Treasury database guide*, Chicago Booth.

Cochrane, John and Piazzesi, Monika, (2005), Bond risk premia, *American Economic Review*, 95, issue 1, p. 138-60.

Cochrane, John and Piazzesi, Monika, (2008), Decomposing the yield curve, *Manuscript.*

Fama, E. F. and Bliss, Robert R., (1987), The information in long-maturity forward rates, *American Economic Review*, 77, issue 4, p. 680-92.

Garbade, Kenneth D., (2007), The emergence of "regular and predictable" as a Treasury debt management strategy, *Federal Reserve Bank of New York Economic Policy Review*, March 2007, p. 53-71.

Greenwood, Robin, Hanson, Samuel and Stein, Jeremy C., (2010), A gap-filling theory of corporate debt maturity choice, *The Journal of Finance*, 65, issue 3, p. 993-1028.

Gürkaynak, Refet S., Sack, Brian and Wright, Jonathan H., (2006), The U.S. Treasury yield curve: 1961 to the present, *Finance and Economics Discussion Series (FEDS) Working Paper* #2006-28, Board of Governors of the Federal Reserve System (U.S.).

Hamilton, James D. and Wu, Cynthia, (2012), Identification and estimation of gaussian affine term structure models, *forthcoming in Journal of Econometrics.*

Kim, Don H., (2008), Challenges in macro-finance modeling, BIS Working Paper #240, *FEDS Working Paper* #2008-06.

Krishnamurthy, Arvind and Vissing-Jorgenson, Annette, (2010), The aggregate demand for Treasury debt, *Manuscript*, Northwestern University.

Modigliani, Franco and Sutch, Richard, (1967), Debt Management and the Term Structure of Interest Rates: An Empirical Analysis of Recent Experience, *Journal of Political Economy*, 75, issue 4, p. 569-89.

Nelson, Charles R. and Siegel, Andrew F., (1987), Parsimonious modeling of yield curves, *The Journal of Business*, 60, issue 4, p. 473-89.

Svensson, Lars E. O., (1994), Estimating and interpreting forward rates: Sweden 1992 - 4, *NBER Working Paper* #4871, National Bureau of Economic Research, Inc.

Vayanos, Dimitri and Vila, Jean-Luc, (2009), A preferred-habitat model of the term structure of interest rates, *NBER Working Paper* #15487, National Bureau of Economic Research, Inc.

7. Appendix

Table 4.2. Estimation Results for the Three Observable Yield Curve Factor Model

	μ	Φ			Σ		
$y_t^{(30)}$	.0309	.9736	-.0042	-.0269	.0026	0	0
	(.0131)	(.0381)	(.0108)	(.0097)			
$y_t^{(30)} - y_t^{(1)}$	.0179	-0.0151	.9845	-.0152	.0011	.0023	0
	(.0111)	(.0126)	(.0366)	(.0098)			
$y_t^{(1)} + y_t^{(30)} - 1.9y_t^{(5)}$	.02203	-.0045	-.0131	.8401	-.0005	.0006	.0022
	(.0112)	(.0107)	(.0115)	(.0333)			
	λ_0	λ_1					
$y_t^{(30)}$	13.17	-17.047	45.429	48.24			
	(.27)	(.12)	(1.11)	(2.61)			
$y_t^{(30)} - y_t^{(1)}$	4.2381	29.034	-256.54	-331.33			
	(.50)	(1.62)	(9.53)	(20.18)			
$y_t^{(1)} + y_t^{(30)} - 1.9y_t^{(5)}$	24.249	-62.383	393.945	408.22			
	(1.35)	(1.61)	(9.44)	(30.44)			

Table 4.3. Estimation Results for the Three Observable Treasury Supply Factor Model

	μ	Φ			Σ		
$-s_t^{(I)}$	-.04208	.9612	-.0185	.07123	.0062	0	0
	(.009)	(.022)	(.009)	(.019)			
$s_t^{(I)} - s_t^{(V)}$	.00514	.0295	.9846	.0151	-0.0008	.0043	0
	(.018)	(.007)	(.015)	(.015)			
$s_t^{(VI)} - s_t^{(II)}$	.0430	.0399	.0036	.9233	-0.0065	.0003	.0021
	(.013)	(.012)	(.010)	(.024)			
	λ_0	λ_1			A_1	B_1	
$-s_t^{(I)}$	172.39	83.679	-7.032	-63.106	-1.1462	-.5141	
	(3.03)	(1.27)	(.74)	(1.13)	(.003)	(.003)	
$s_t^{(I)} - s_t^{(V)}$	1354.99	-61.824	-117.15	1.906		.3637	
	(31.86)	(.83)	(.02)	(3.62)		(.001)	
$s_t^{(VI)} - s_t^{(II)}$	964.96	105.41	-70.564	-120.54		-.2209	
	(33.53)	(1.07)	(.15)	(2.19)		(.012	

TABLE 4.4. Estimation Results for the Hybrid Model

	μ	Φ				Σ			
$y_t^{(30)}$	.0401	.9666	.0069	.0127	-.0027	.0026	0	0	0
	(.0155)	(.0063)	(.0084)	(.0318)	(.0220)				
$y_t^{(30)} - y_t^{(1)}$	0	-.0012	.97589	-.0069	.0133	.0011	.0022	0	
	(.0131)	(.0148)	(.0060)	(.0214)	(.0219)				
$-s_t^{(I)}$	-.0819	.0504	-.0350	.9297	-0.0091	.0004	-0.0004	.0063	0
	(.0131)	(.0125)	(.0377)	(.0153)	(.0148)				
$s_t^{(I)} - s_t^{(V)}$	-.0296	.0329	-0.0217	.0145	.9919	.0003	0	-0.0008	.0043
	(.0088)	(.0125)	(.0319)	(.0260)	(.0106)				
	λ_0	λ_1							
$y_t^{(30)}$	-68.97	65.45	-77.61	7.01	2.61				
	(8.12)	(7.50)	(11.64)	(.0039)	(.6638)				
$y_t^{(30)} - y_t^{(1)}$	-113.41	109.01	-371.38	4.456	24.043				
	(4.52)	(3.97)	(7.64)	(.2573)	(1.17)				
$-s_t^{(I)}$	-9408.5	8776.9	-11763	311.28	554.48				
	(217.34)	(191.97)	(752.98)	(25.49)	(47.40)				
$s_t^{(I)} - s_t^{(V)}$	6410.5	-5897.6	3429.5	-84.518	-158.07				
	(100.11)	(101.05)	(631.80)	(21.15)	(21.65)				

The Stockholm School of Economics

A complete publication list can be found at *www.hhs.se/research/publications*. Books and dissertations are published in the language indicated by the title and can be ordered via e-mail: *publications@hhs.se*.

A selection of our recent publications

Books

Alexandersson, Gunnar. (2011). *Den svenska buss- och tågtrafiken: 20 år av avregleringar.* Forskning i fickformat.

Barinaga, Ester (2010). *Powerful dichotomies.*

Benson, Ilinca Lind, Johnny, Sjögren, Ebba & Wijkström, Filip (red.) (2011). *Morgondagens industri: att sätta spelregler och flytta gränser.*

Einarsson, Torbjörm (2011). *Medlemsorganisationen: Individen, organisationen och samhället.*

Engwall, Lars (2009). *Mercury meets Minerva: business studies and higher education: the Swedish case.*

Ericsson, Daniel (2010). *Den odöda musiken.*

Ericsson, Daniel (2010). *Scripting Creativity.*

Holmquist, Carin (2011). *Kvinnors företagande – kan och bör det öka?*

Lundeberg, Mats (2011). *Improving business performance: a first introduction.*

Melén, Sara (2010). *Globala från start: småföretag med världen som marknad.* Forskning i Fickformat.

Modig, Niklas & Åhlström, Pär. (2011). *Vad är lean?: en guide till kundfokus och flödeseffektivitet.*

Mårtensson, Pär & Mähring, Magnus (2010). *Mönster som ger avtryck: perspektiv på verksamhetsutveckling.*

Ottosson, Mikael (2011). *Skogsindustrin och energiomställningen.* Forskning i fickformat.

Sjöström, Emma (2010). *Ansiktslösa men ansvarsfulla.* Forskning i fickformat.

Wijkström, Filip (2010). *Civilsamhällets många ansikten.*

Dissertations

Bjerhammar, Lena (2011). *Produktutvecklingssamarbete mellan detaljhandelsföretag och deras varuleverantörer.*

Bohman, Claes (2010). *Attraction: a new driver of learning and innovation.*

Bottero, Margherita (2011). *Slave trades, credit records and strategic reasoning: four essays in microeconomics.*

Bång, Joakim (2011). *T Essays in empirical corporate finance.*

Carlsson-Wall, Martin (2011). *Targeting target costing: cost management and inter-organizational product development of multi-technology products.*

Formai, Sara (2012). *Heterogeneous firms, international trade and institutions.*

Freier, Ronny (2011). *Incumbency, divided government, partisan politics and council size: essays in local political economics.*

Fries, Liv (2011). *Att organisera tjänstesektorns röst: en teori om hur organisationer formar aktörer.*

Hemrit, Maetinee (2011). *Beyond the Bamboo network: the internationalization process of Thai family business groups.*

Kaisajuntti, Linus (2011). *Multidimensional Markov-functional and stochastic volatility interest rate modelling.*

Khachatryan, Karen (2011). *Biased beliefs and heterogeneous preferences: essays in behavioral economics.*

Lundmark, Martin (2011). *Transatlantic defence industry integration: discourse and action in the organizational field of the defence market.*

Lundvall, Henrik (2010). *Poverty and the dynamics of equilibrium unemployment: [essays on the economics of job search, skills, and savings].*

Migueles Chazarreta, Damián (2011). *The implications of trade and offshoring for growth and wages.*

Monsenego, Jérôme (2011). *Taxation of foreign business income within the European internal market: an analysis of the conflict between the objective of achievement of the European internal market and the principles of territoriality and worldwide taxation*.

Ranehill, Eva (2011). *Essays on gender, competition and status*.

Runsten, Philip (2011). *Kollektiv förmåga: en avhandling om grupper och kunskapsintegration*.

Setterberg, Hanna (2011). *The pricing of earnings: essays on the post-earnings announcement drift and earnings quality risk*.

Stepanok, Ignat (2011). *Essays on international trade and foreign direct investment*.

Söderblom, Anna (2011). *Private equity fund investing: investment strategies, entry order and performance*.

Wallace, Björn (2011). *Genes, history and economics*.