GOVERNANCE AND THE EFFICIENCY
OF ECONOMIC SYSTEMS
GESY

Discussion Paper No. 409

Identification and Estimation of Intra-Firm and Industry Competition via Ownership Change

Christian Michel *

* University of Mannheim

May 2013

Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 is gratefully acknowledged.

Sonderforschungsbereich/Transregio 15 · www.sfbtr15.de
Universität Mannheim · Freie Universität Berlin · Humboldt-Universität zu Berlin · Ludwig-Maximilians-Universität München
Rheinische Friedrich-Wilhelms-Universität Bonn · Zentrum für Europäische Wirtschaftsforschung Mannheim

Speaker: Prof. Dr. Klaus M. Schmidt · Department of Economics · University of Munich · D-80539 Munich,
Phone: +49(89)2180 2250 · Fax: +49(89)2180 3510

Identification and Estimation of Intra-Firm and Industry Competition via Ownership Change*

Christian Michel†
University of Mannheim

May 26, 2013

Abstract

This paper proposes and empirically implements a framework for analyzing industry competition and the degree of joint profit maximization of merging firms in differentiated product industries. Using pre- and post-merger industry data, I am able to separate merging firms' intra-organizational pricing considerations from industry pricing considerations. The insights of the paper shed light on a long-standing debate in the theoretical literature about the consequences of organizational integration. Moreover, I propose a novel approach to directly estimate industry conduct that relies on ownership changes and input price variation. I apply my framework using data from the ready-to-eat cereal industry, covering the 1993 Post-Nabisco merger. My results show an increasing degree of joint profit maximization of the merged entities over the first two years after the merger, eventually leading to almost full maximization of joint profits. I find that between 14.3 and 25.6 percent of industry markups can be attributed to cooperative industry behavior, while the remaining markup is due to product differentiation of multi-product firms.

Keywords: Identification of Market Structure, Post-merger Internalization of Profits, Conduct Estimation, Ex-post Merger Evaluation, Estimation of Synergies

*For the latest version, see http://sites.google.com/site/christianfelixmichel/.
I would like to thank my advisors Volker Nocke, Philipp Schmidt-Dengler, and Yuya Takahashi for their guidance and support. I also benefited from conversations with Steve Berry, Pierre Dubois, Georg Duernecker, David Genesove, Alex Shcherbakov, Andre Stenzel, Andrew Sweeting, Naoki Wakamori, and Stefan Weiergraeber, and received helpful comments from seminar participants at Mannheim, Toulouse, Tilburg, Bocconi, Vienna, Compass Lexecon Europe, Copenhagen Business School, Autonoma Barcelona, Pompeu Fabra, ESSEC, the 2012 CEPR Applied IO Summer School, and the EARIE 2012 in Rome.
†Department of Economics, University of Mannheim, L7, 3-5, 68161 Mannheim, Germany.
Email: christian.felix.michel@gmail.com.

1 Introduction

In this paper, I propose a framework to address two open questions in industrial organization and organizational economics, and apply it to a merger from the ready-to-eat cereal industry. First, I examine to what degree merging firms jointly maximize their profits after a horizontal merger. This bridges a gap between empirical industrial organization models and organizational economics models. Second, I provide a way to directly identify and estimate industry conduct in differentiated product models, which has for a long time been a problem in the industrial organization literature.

Existing empirical merger models make several simplifying assumptions with respect to both industry and within firm behavior. On an organizational level, these models assume that merging firms fully internalize their profits immediately after a merger. On an industry level, the form of supply side competition is either assumed to be known, or is chosen from a discrete set of non-nested forms of competition using some selection criterion. These assumptions are usually used together with pre-merger data to predict post-merger industry prices. My framework makes use of both observable pre- and post-merger data. This allows me to relax either the assumption of full profit internalization or the assumption of known industry conduct while keeping the other. I recover pre-merger marginal costs using first-order conditions between merging firms and predict post-merger marginal costs using a cost function estimation. Different forms of industry conduct or post-merger profit internalization of merging firms, respectively, will lead to different markups charged by firms. My identification strategy searches for the form of supply side competition that best predicts post-merger industry prices. I can show that the difference between predicted and observed post-merger prices amounts to a structural cost function error term, which helps to identify the supply-side parameters of interest given proper instruments. Thus, my approach can be seen as a supply side analogue to common demand side models that use structural error terms to identify demand side and cost function parameters, see for example Berry, Levinsohn, and Pakes (1995).

I apply the developed techniques to data from the ready-to-eat cereal industry covering the 1993 Post-Nabisco merger. In January 1993, Philipp Morris corporation's owned Kraft foods with its Post cereal line purchased the Nabisco ready-to-eat cereal branch from RJR Nabisco. The results indicate an increasing degree of joint profit maximization among merging firms after the merger. This suggest the existence of informational or contractual frictions among merging firms shortly after the merger. With respect to industry competitiveness, I find that between 14.3 and 25.6 percent of industry markups can be attributed to cooperative industry behavior, while the remaining markup is due to product differentiation of multi-product firms.

The paper extends the existing literature along several dimensions. To my knowledge, this is the first paper to focus on estimating the degree of joint profit maximization of a merged entity. This links the empirical industrial organization literature to the theoretical organizational economics literature on intra-firm coordination and horizontal integration by allowing for frictions between different divisions of a firm. Conceptually, the approach also differs from existing empirical organizational economics models in that its focus is on behavior within a single (post-merger) organization rather than on correlations across firms and industries. Moreover, I show that using proper supply side variation, it is indeed possible to estimate industry conduct directly in differentiated product industries. Using demand side variation, this can typically not be achieved due to a lack of sufficiently many demand rotators, see for example Nevo (1998).

Following the seminal paper by Berry, Levinsohn, and Pakes (1995), identification of demand and cost parameters relies on orthogonality conditions between structural error terms and appropriate instruments. Unlike the existing literature, I also identify the model's underlying supply side parameters, i.e. the degree of profit internalization and industry conduct, respectively. I show that the difference between observed and predicted post-merger prices represents a structural cost error term. I set up moment conditions that rely on orthogonality conditions between this error term and cost-side instruments to identify the supply side parameters.

Modern empirical industrial organization models assume that a merged entity maximizes the joint profits of all its products, thus abstracting from agency problems within the firm. From an organizational viewpoint, several theories predict that full internalization of joint profits cannot be achieved after a merger. Incentive structures that give managers bonuses based on the performance of their own division rather than the performance of the firm as a whole can cause different horizontal divisions to compete with each other. Fauli-Oller and Giralt (1995) analyze a headquarter's choice of the optimal incentive scheme for division managers. Whenever products of different divisions are substitutes to each other, managers bonuses will be partly based on their own division's performance. There is also a growing literature in organizational economics that focuses on the trade-offs between coordination of decision-making through a headquarter and strategic communication of division managers.[1] Other reasons for no full maximization of joint profits immediately after a merger are delays in post-merger harmonization of firm strategies due to old contractual agreements, or a lack of information concerning revenue potential right after a merger.

I focus on a single merger to assess its consequences on joint maximization of profits. This differs from conventional empirical organizational economics frameworks that focus on correlations between observable firm characteristics across different firms and industries.[2]

[1]Alonso, Dessein, and Matouschek (2008) study the optimal degree of centralization when managers communicate strategically. They show that while centralization can improve horizontal outcomes, it will lead to adverse vertical effects. Dessein, Garicano, and Gertner (2011) show the existence of endogenous incentive conflicts between headquarter managers and division managers within multi-divisional firms.

[2]There is a large empirical literature in organizational economics focusing on the determinants for specific orga-

When estimating industry conduct, I maintain the assumption that merging firms internalize the profits after the merger. Given marginal cost estimates and price-elasticities obtained from demand side-estimation, I can predict the effects of an ownership change on prices ex-post. By varying the form of supply side competition (i.e. industry conduct), and accounting for input price changes on the cost side, I look for the form of competition that most accurately predicts the effects of the merger-induced ownership change on prices. I estimate the predicted post-merger prices and compare them with the observed post-merger prices. The differences of observed prices and predicted prices is used to form moments in order to obtain the model's underlying conduct parameters using a Generalized Methods of Moments estimator.

Previous attempts to estimate both marginal costs and industry conduct have mostly been made using demand side variation. Bresnahan (1982) and Lau (1982) provide identification results for estimating conduct in the homogeneous good case. In differentiated product industries, these approaches usually face two kinds of problems. The first problem is the difficulty to find a sufficient number of demand rotators. Without such rotators, these approaches are not able to identify industry conduct.[3] The second problem relates to the estimation techniques, which only estimate the economic parameters of interest accurately in special cases. Corts (1999) critically discusses the identification of conduct parameters. He argues that the estimated parameters usually differ from the "as-if conduct parameters" and therefore do not reflect the economic parameters of interest. The static conduct estimation models are furthermore not able to detect all dynamic forms of collusion. While I cannot account for the latter point due to the static character of my framework, my estimation technique can overcome the former.

There is a small literature related to the estimation of industry conduct using supply side variation. Ciliberto and Williams (2010) develop an approach that relies on multi-market contact for estimating conduct in the airline industry. Their model includes conduct parameters that can have three different values, accounting for different degrees of cooperation among profit-maximizing firms. Oliveira (2011) uses marginal profit ratios in a dynamic model to distinguish between market competition and efficient "stick-and-carrot" collusion in the airline market. Brito, Pereira, and Ramalho (2012) explore the effects of three Portuguese insurance mergers on coordinated effects and efficiency. They find no indication for an increase in coordinated effects after the mergers.

The remainder of the paper is organized as follows. Section 2 gives an overview over the

nizational structures across firms and industries, see for example Lafontaine and Slade (2012), and Bloom, Genakos, Sadun, and Van Reenen (2012) for an overview over the literature. McElheran (2010) finds a positive correlation between delegation of IT system adoption in multi-divisional firms and local information advantages, but a negative correlation between delegation of system adoption and firm size. Thomas (2011) argues that a reduction in the brand portfolios of firms in the laundry detergent industry across different countries would lead to an increase in their profits.

[3] Nevo (1998) discusses advantages and disadvantages of a direct conduct estimation compared to a non-nested menu approach. He argues that in practice estimating conduct directly will be impossible due to a lack of sufficiently many distinct demand shifters.

industry and the merger. Section 3 introduces the baseline model and discusses the conduct estimation strategy in detail. Section 4 provides identification results for both estimating the degree of joint profit maximization of merging firms and estimating industry conduct, respectively. Section 5 presents estimation results for both techniques. Section 6 introduces several extensions to the baseline model as outlined above. Section 7 concludes with a discussion of the results.

2 Industry Overview and the 1993 Post-Nabisco Merger

2.1 The ready-to-eat cereal industry

There are several factors that make the ready-to-eat (RTE) cereal industry an excellent starting point for oligopoly analysis.[4] Economies of scale in packaging different cereals, as well as in the distribution of products, cause barriers to entry for new firms. There is a frequent introduction of new products by existing firms, which goes in line with large advertising campaigns in the beginning of a product's lifecycle.[5] The cereals differ with respect to their product characteristics, such as sugar content or package design, and target different consumer types. At the start of the period I analyze, the industry consists of 6 main nationwide manufacturers: Kellogg's, General Mills, Post, Nabisco, Quaker Oats, and Ralston Purina. Figure 1 shows the market shares of the different products. It is common to classify the cereals into different groups, such as adult, family, and kids cereals, see also Nevo (2001). Table 1 shows the classification of the different cereal brands in my dataset into different segments. Kellogg's as the firm with the biggest market share has a strong presence in all segments. General Mills is mainly present in the family and kids segments, whereas Post and Nabisco have their main strengths in the adult segments.

On a retail level, cereals are primarily distributed via supermarkets. Supermarket promotions via price reductions or quantity discounts are a further tool used to increase quantities sold for a period of time. Many retailers also own private labels that compete in their stores with the nationwide manufacturers. I use scanner data from the first quarter of 1991 until the fourth quarter of 1995 from the Dominick's Finer Food database. My main dataset for the conduct estimation includes 28 brands from the 6 different nationwide firms. The scanner data involves 35 stores from the Chicago Metropolitan area, see Figure 2 for a geographic map of the stores. In particular, the dataset includes data on product prices, quantities sold, data on promotions, as well as 1990 census data yielding demographic variables for the different store locations. I use additional input price data from the Thomson Reuters Datastream database. Even though I also observe data on Dominick's private label cereal, I do not include it in my conduct estimation. There are two reasons for this. First, I want to

[4]This industry has already been studied extensively, see for example Schmalensee (1978), Nevo (2000), and Corts (1995). Although Corts presents a detailed industry description, to my knowledge the dynamic aspects on the supply side have not been investigated in detail.

[5]See for example Hitsch (2006) for a study of the determinants of successful brand introductions.

focus on the degree of competition between firms that are operating nationwide. Because a private label is only present for one retailer, and in my case a locally operating retailer, it will have different underlying objectives than the nationwide operating manufacturers. Second, a private label firm belongs to its retailer, thus leading to a joint maximization of profits upstream and downstream. This would require additional assumptions to be compatible with estimating "upstream" industry conduct.

In my data, I also observe the retailer's average acquisition costs for each product at a given time. This variable reflects the inventory-weighted average of the fraction of the retail price that was paid to the producer. From this variable I can compute a proxy for the average retail margin for a given period.[6] Figure 3 shows the development of the retail margin proxy over time for the different firms in the dataset. There are several interesting features. The retail margin varies significantly across the different firms. On average retail margins are highest for Ralston, the firm with the smallest market share, followed by Kellogg's, the firm with the highest market share. Thus, there is no clear relationship between retail margin and firm size, suggesting that there is no higher bargaining power for Kellogg's.[7] Another interesting fact is that the retail margin drops significantly around the time of the merger, from over 15% to single digit figures for several firms, including the merging firms. It is not clear whether this drop is due to the merger, which would imply some form of renegotiation between manufacturers and retailer in the period, or whether it is instead a pure coincidence.

2.2 The 1993 Post-Nabisco merger

Between 1990 and 1992, prices steadily increased in the industry, see Figure 4 for the price development per firm. On November 12, 1992, Kraft Foods made an offer to purchase RJR Nabisco's ready-to-eat cereal line. The acquisition was cleared by the FTC on January 4, 1993. On February 10th, 1993, the New York State attorney however sued for a divestiture of the Nabisco assets, which was finally turned down 3 weeks later.[8] Figure 5 shows the merging firms' price development over all stores. Table 2 shows the price development per product for the first four quarters after the merger. Average prices for the merging firms increase over time, which is in line with a unilateral effects merger model. The price development of other firms is more heterogeneous. Kellogg's decreases part of its brands prices while increasing some of its other prices. Prices for Ralston also go up. Prices for most of General Mills products slightly decrease. This can be attributed to both a change in General Mills high management in 1993, in which the company responded to soaring market shares, and to the fact that General Mills was mostly present in the kids and family segment that was not affected as much by the merger. Overall industry behavior remained stable. Between 1993 and March 1995, industry-wide prices for branded RTE cereal increased moderately.

[6] Dominick's uses the following formula for the average acquisition costs (AAC): AAC(t+1) = (Inventory bought in t) Price paid(t) + (Inventory, end of t-l-sales(t)) AAC(t).

[7] Another potential source for bargaining power not modeled here is bargaining power in form of leading to more premium shelf spaces.

[8] See Rubinfeld (2000) for a detailed description.

Starting from the second quarter of 1995, I observe a downward trend in industry prices across different firms. In March 1995, two US congressmen started a public campaign to reduce cereal prices, which received a huge media attention. There is evidence that industry prices decreased after this campaign.[9] For this reason, I only consider the period until the first quarter of 1995 for most estimations.

Exogeneity of merger From an estimation standpoint, it is important to discuss concerns and potential effects of merger endogeneity. After the 1988 leveraged buyout of RJR Nabisco, the ownership group accumulated a substantial pile of debt. There is a popular claim that company divestitures were used to reduce the overall debt level. Merger endogeneity would only bias the results if the merger had led to unknown synergies, or if an anticipation of the merger by firms in the industry had led to a change in the competition between firms. There are no factory closures within the first two years of the merger. Therefore, synergies in factory production are unlikely to be achieved. Within the first two years after the merger, I also do not observe a fundamental change in industry pricing due to the merger, which is backed by anecdotal evidence. Thus, there is a relative steadiness in industry behavior in the short run.

3 Empirical Model

My approach has three basic steps. In the first step, I estimate industry demand using a discrete choice model to back out price elasticities. In the second step, I recover marginal costs using first order conditions, which I subsequently use to estimate a cost function. In the third step, I predict post-merger prices, which I then use to estimate either the degree of profit internalization among merging firms or industry conduct. In the industry conduct case the last two steps are repeated in an iterative process.

3.1 Demand side

My demand specification is closely related to Nevo (2001). There is a total number of J brands in the market. Denote the number of individual consumers in every market by I, and denote the number of markets by T, where a market is defined as a time-store combination. Using a Random Coefficient Logit model, individual i's indirect utility of consuming product j at market t can be written as:

$$u_{ijt} = x_j\tilde{\beta}_i + \tilde{\alpha}_i p_{jt} + \xi_{jt} + \epsilon_{ijt}, \;\; i = 1, .., I; j = 1, .., J; t = 1, .., T. \tag{1}$$

[9] See Cotterill and Franklin (1999) for a detailed analysis. In April 1996, Post decreased the prices for its products nationwide by 20%, thereby permanently increasing its markets share. This was followed by significant price cuts two months later by General Mills and Kellogg's. Overall, margin over production cost fell by 12% in 1996 due to these actions.

x_j denotes a K-dimensional vector of firm j's observable brand characteristics, p_{jt} denotes the price of product j at market t, and ξ_{jt} the brand-specific mean valuation at market t that is unobservable to the researcher but observable to the firms. Finally, ϵ_{ijt} is an idiosyncratic error term. The coefficients $\tilde{\beta}$ and $\tilde{\alpha}$ are indiviual specific coefficients. These coefficients depend on their mean valuations, on demographics in each region, D_i and their associated coefficients Π, as well as on an unobserved vector of shocks, v_i that is interacted with a scaling matrix Σ:

$$\begin{pmatrix} \tilde{\alpha}_i \\ \tilde{\beta}_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i, \quad v_i \sim \; N(0, I_{K+1}). \tag{2}$$

Because not all of the potential consumers purchase a good in each period, I also require an outside good. The indirect utility of not purchasing any product and thus consuming the outside good can be written as:

$$u_{i0t} = \xi_0 + \pi_0 D_i + \sigma_0 v_{i0} + \epsilon_{i0t}.$$

As is common in the literature, I normalize ξ_0 to zero.

Denote the vector of all demand side parameters by γ^D. This vector can be decomposed into a vector of parameters obtained from the linear part of the estimation, $\gamma_1^D = (\alpha, \beta)$, and a vector of parameters obtained from the nonlinear part of the estimation, $\gamma_2^D = (vec(\Pi), vec(\Sigma))$, respectively.

The indirect utility of consuming a product can be decomposed into a mean utility part δ_{jt} and a mean-zero random component $\mu_{ijt} + \epsilon_{ijt}$ that takes into account heterogeneity from demographics and captures other shocks. The decomposed indirect utility can be expressed as

$$u_{ijt} = \delta_{jt}(x_j, p_{jt}, \xi_{jt}, \gamma_1^D) + \mu_{ijt}(x_j, p_{jt}, v_i, D_i; \gamma_2^D)$$

$$\delta_{jt} = x_j \beta - \alpha p_{jt} + \xi_{jt}, \;\; \mu_{ijt} = [p_{jt}, x_j]' * (\Pi D_i + \Sigma v_i), \tag{3}$$

where $[p_{jt}, x_j]$ is a $(K+1) \times 1$ vector.

Consumers either buy one unit of a single product or take the outside good. They will choose the option which yields the highest indirect utility. Using these assumptions, this characterizes the set A_{jt} of unobservables that yield the highest utility for a specific choice j:

$$A_{jt}(x_{.t}, p_{.t}, \delta_{.t}, \gamma_2^D) = \{(D_i, v_i, \epsilon_{it}) | u_{ijt} \geq u_{ilt} \forall l \in \{0, .., J\}\},$$

where dotted indices indicate vectors over all J brands. The market shares predicted by the model can then be obtained via integrating over the different shocks, using population moment functions $P^*(\cdot)$:

$$s_j(x_{.t}, p_{.t}, \delta_{.t}, \gamma_2^D) = \int_{A_{jt}} dP^*_\epsilon(\epsilon) dP^*_v(v) dP^*_D(D). \tag{4}$$

There are several possibilities to estimate the model that depend on different distributional assumptions. The most general case is a Random Coefficients Logit model. Its main advantage is a very flexible form of substitution patterns. This is desirable because it enables a detailed analysis of the substitution patterns between different brands that does not rely on any model structure. To be able to integrate out the market shares, one needs to make distributional assumptions with respect to the unobservable variables $(D_i, v_i, \epsilon_{ijt})$ and then estimate the model using Generalized Methods of Moments.

3.2 Industry technology

The J brands in the industry are produced by $N \leq J$ firms. Each brand can only be produced by one firm. An important part of the model is the representation of marginal cost. As is common in the literature, I assume that marginal costs can be decomposed into cost factors that are observed to the researcher as well as a component that is unobserved to the researcher. I use a linear relationship between marginal costs and the observable cost component. This reflects a relatively weak substitutability of input production factors over the medium- and short-run in the RTE cereal industry. Henceforth, I will omit the market index t in my exposition for notational simplicity. Denote the vector of brand $j's$ observed cost drivers by w_j, and j's unobserved cost component by ω_j. The marginal cost for brand j can be written as:

$$mc_j = w_j \gamma^S + \omega_j \tag{5}$$

where γ^S denotes a vector of marginal cost parameters.[10]

3.3 Pre-merger industry competition

Each firm f owns a portfolio of brands $\mathbb{F}_f$. I further allow a firm's objective function to potentially depend on other firm's profits. Denote by θ_{ij} the degree to which brand i takes into account brand j's profits when setting its optimal price. This implicitly defines a pre-merger ownership matrix Θ with the entries $\Theta_{jr} = \theta_{jr}$. Each of the parameters within Θ are normalized to lie in between 0 and 1, where 0 implies no internalization of profits, and 1 implies full internalization of profits. Note that since only relative weights matter for the first order condition, this is a normalization without loss of generality. Not allowing for negative

[10] The baseline specification implies that marginal costs are constant for different output levels. This is a relatively strict assumption, which can be relaxed by introducing scale effects. Denote q_i firm $i's$ total units sold in a period. If one assumes scale effects, i.e. decreasing marginal costs in total production together with a Cobb-Douglas cost function, then this can be written as: $mc_j = \tau log(q_j) + w_j \gamma^S + \omega_j$, where τ is the scale parameter.

conduct parameters also implies that a firm does not derive a positive utility from "ruining" another firm. This leads to a matrix of the form

$$\Theta = \begin{pmatrix} 1 & \theta_{12} & .. & \theta_{1J} \\ \theta_{21} & 1 & .. & \theta_{2J} \\ .. & .. & .. & .. \\ \theta_{J1} & \theta_{J2} & .. & 1 \end{pmatrix}.$$

Given Θ, the objective function for product j can be written as:

$$\Pi_j = (p_j - mc_j)s_j\overline{M} + \sum_{r \neq j} \theta_{jr}(p_r - mc_r)s_r\overline{M}, \tag{6}$$

where s_r denotes the market share of brand r, and $\overline{M}$ the market size.

The first order condition for product j's objective function with respect to its own price can be written as:

$$s_j(p) + \sum_{r=1}^{J} \theta_{jr}(p_r - mc_r)\frac{\partial s_r}{\partial p_j} = 0. \tag{7}$$

I make the assumption that pre-merger, each firm fully internalizes the profits of all of its brands, which implies $\theta_{ij} = 1$ if $i, j \in \mathbb{F}_f$. The marginal costs of all brands in the industry are unobserved to the researcher but common knowledge among firms.

Define $\Omega_{jr} \equiv -\theta_{jr} * \frac{\partial s_r}{\partial p_j}$. Having estimated the demand parameters γ^D, one can already infer the marginal costs of production mc conditional on the form of the ownership matrix Θ:

$$mc(\gamma^D, \Theta, p, x) = p^{pre} - \Omega^{-1}(\gamma^D, \Theta)s^{pre}. \tag{8}$$

Substituting the recovered pre-merger marginal cost $mc(\gamma^D, \Theta, p, x)$ into equation (5) yields:

$$mc(\gamma^D, \Theta, p, x) = w\gamma^S + \omega \tag{9}$$

I estimate the cost function parameter γ^S in equation (9) using a two staged least squared estimation to account for the unobserved cost component ω. I discuss instrumentation in Section 4. I then use observed post-merger input price drivers, w^{post}, to linearly project the post-merger input price component of marginal cost. Product $j's$ predicted input price component post-merger, $\hat{mc}_j$ can then be written as:

$$\hat{mc}_j^{post}(\gamma^D, \gamma^S, \Theta, p, x, w) = w_j^{post}\gamma^S. \tag{10}$$

It is important to note that this linear projection will not contain the unobserved cost component ω_j. Henceforth, for notational simplicity I will drop the observable factors w, x, and p when referring to marginal costs.

The key differences between estimating the degree of profit internalization of merging firms and estimating industry conduct between firms lies in the treatment of the pre-merger ownership matrix Θ and on the assumptions with respect to the merging firms' post-merger profit internalizations.

3.4 Post-merger prices when estimating the degree of profit internalization between merging firms

When estimating the degree of profit internalization post-merger, I assume that the form of pre-merger industry competition, Θ, is known to the researcher. I assume that a merger will involve a change in the pricing strategies of the merged entity. Denote by $\tilde{\theta}$ the degree of joint profit maximization between the merging firms, which I assume to be unobserved to the researcher, but common knowledge among firms in the industry. I assume that non-merging firms will not change their competitive behavior after a merger, but will adapt to the change in the merging firms' pricing. Under these assumptions, given the degree of joint profit maximization $\tilde{\theta}$, the model's post-merger prices given the parameters γ^D and γ^S can be written as[11]:

$$\hat{p}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) + \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post} + \omega_{\tilde{\theta}}^{post}. \tag{11}$$

Here, $\omega_{\tilde{\theta}}^{post}$ reflects the unobserved cost component in the cost function post-merger. Thus, the unobserved error will not be accounted for in the linear projection of the post-merger marginal cost in equation (10), but remains a separate term in the post merger equation.

3.5 Post-merger prices when estimating industry conduct

When estimating industry conduct, I assume that neither the form of pre-merger conduct, Θ, nor post-merger industry conduct are known to the researcher. A key identifying assumption is that even though the researcher does not know the underlying form of conduct, he knows exactly how the merger will affect industry conduct. There are two channels for this, namely post-merger profit internalization of merging firms, and how competitors consider the merged entity post-merger. I assume that merging firms will fully internalize their profits after a merger. Non-merging firms compete with each other in the same fashion. The change between merging and non-merging firms after the merger is also known to the researcher, which is summed up in the following assumption.

Assumption 1 (Conduct between merging and non-merging firms)**.** *Let f, g be two distinct merging firms, and h a non-merging firm. Let θ_{ik}^{pre} and θ_{ik}^{post} denote the pre- and post-merger*

[11]Even if the degree of joint profit maximization, $\tilde{\theta}$, is part of a post-merger industry conduct matrix, I will explicitly state it in the model. This is to clearly distinguish the case of estimating the joint profit-maximization parameters from the case when estimating the industry conduct matrix Θ.

conduct parameters between firm i and k, respectively. Then, $\forall i \in \mathbb{F}_f$, $\forall j \in \mathbb{F}_g$, $\forall k \in \mathbb{F}_h$, one of the following three cases holds regarding the conduct between a merging and a non-merging firm:
a. $\theta_{ik}^{post} = \theta_{ik}^{pre}$; $\theta_{jk}^{post} = \theta_{jk}^{pre}$ *(no change in conduct);*
b. $\theta_{ik}^{post} = \theta_{ik}^{pre}$; $\theta_{jk}^{post} = \theta_{ik}^{pre}$ *(acquiring firm standard).*
c. $\theta_{ik}^{post} = \theta_{jk}^{pre}$; $\theta_{jk}^{post} = \theta_{jk}^{pre}$ *(target firm standard);*

Under Assumption 1a, the merger does not change how competitors consider the two merging firms after the merger. In this case, the merging firms will fully internalize the profits after the merger. Under Assumption 1b, the fully merged entity is considered and behaves as the acquirer did pre-merger. Assumption 1c implies the reverse, i.e. that the merged entity behaves as the target. I do not have to pre-specify the values of the conduct parameters, but just the way in which the parameters change. Other change patterns can also be accounted for, as long as the change in conduct between merging and non-merging firms is known post-merger. Let the underlying form of pre-merger conduct, Θ, be element of the set B. The post-merger conduct can then be expressed via a transition function $b : B \to B$ which maps the pre-merger conduct matrix Θ into the post-merger conduct matrix $\Theta^{post} = b(\Theta)$. This is because all merger-induced ownership transformations are known to the researcher. Under these assumptions, for a specific form of pre-merger industry conduct Θ and transition function b, the predicted post-merger prices given the estimated parameters γ^D and γ^S can be written as:

$$\hat{p}^{post}(\gamma^D, \gamma^S, \Theta, \Omega^{post}, b(\Theta)) = \hat{mc}(\gamma^D, \gamma^S, \Theta) + \Omega^{-1}(\gamma^D, b(\Theta))s^{post} + \omega_{\Theta}^{post}. \tag{12}$$

As for the case of estimating the degree of post-merger profit internalization, the unobserved cost component ω_{Θ}^{post} is not accounted for in the marginal cost term $\hat{mc}$, but is rather a distinct component in the pricing equation. Appendix A illustrates different cases of pre- and post merger industry conduct. It is worth mentioning the key difference to conjectural variation models. I treat the conduct parameters Θ as part of the firms' underlying objective functions rather than as behavioral responses with respect to the competitors' price setting behaviors. I then form moment conditions to recover the underlying "level" conduct parameters using a Generalized Method of Moments estimator.

4 Identification

In this section, I will present identification results for the different stages of the model, and in particular the two supply side estimation approaches. Section 1 discusses identification restrictions for the demand and cost function estimation. Section 4.2 discusses the identification restrictions when estimating merging firms' profit internalization parameters. Section 4.3 focuses on identification restrictions when estimating industry conduct.

My estimation method requires identification of three sets of parameters: demand-side parameters, γ^D, cost parameters γ^S, and the supply side parameters, which amount either to the degree of profit internalization $\tilde{\theta}$, or to the industry conduct Θ. The correlation between price and both unobserved brand and cost characteristics requires instrumentation for each brand in the demand and pricing equations, respectively.

4.1 Model identification of demand and cost parameters

4.1.1 Identification of demand parameters

Denote by $\xi(\gamma^D, x, p)$ the structural error term vector that consists of the market-specific unobservable brand valuations for all brands. Regarding the demand side, I assume that when being assessed at the true demand parameter values γ_0^D, this error term is uncorrelated with respect to a M_ξ-dimensional set of exogenous demand side instruments, Z_ξ. This leads to the identification restriction:

$$E[Z_\xi' \xi(\gamma_0^D, x, p)] = 0. \tag{13}$$

Note that I implicitly assume that the demand can be estimated independently of the marginal cost and supply side parameters, respectively. The orthogonality conditions would be violated if industry conduct or a change in production costs, not prices, would influence consumer choice through the unobserved brand-specific component.[12] I assume that the observable product characteristics of the different goods are exogenous, and therefore do not respond to changes in industry pricing. Also accounting for potential brand replacement or additional brand introductions would make traction of the full model nearly impossible. Because of the inherent endogeneity between price and unobserved brand characteristics, I need to find adequate instruments for the demand estimation. I use two different sets of instruments to do so.

My first set of instruments relies on production cost shifters. The economic assumption is that input cost variation should be correlated with variation in prices, but not with consumers' preferences for unobservable product characteristics. I use both cost factors that affect all products in similar fashion, such as labor costs, packaging, and transportation, as well as factors that differ among products, such as interactions between product characteristics and input prices for wheat, sugar, and corn. My second set of instruments is the ownership change itself. As argued above, a merger should cause a change in industry prices. Similar to a cost shift, one can assume that the merger affects prices, but not the demand characteristics. This assumption would be violated if the merger caused a change in brand value which would affect the ξ's of the merging firms. Because the actual brand names of the cereals involved did not change after the merger, such an effect seems unlikely.

[12] This assumption would be violated if the merger caused a change in the perceived "brand values" of the merged entities, which would affect the ξ components in the demand equation.

4.1.2 Identification of cost parameters

Conditional on a specific form of pre-merger industry conduct Θ, I can back out the marginal cost via a first order condition and then regress them on observable product characteristics combined with input prices.[13] This allows me to predict the input cost component $\hat{mc}$ of the post merger marginal costs using post-merger input price data and the estimated parameters. I make the implicit assumption that firms cannot substitute between different input goods. The recipes and production processes for a specific product in the ready-to-eat cereal industry remain constant over time, such that this assumption is likely to hold in the medium and short term. My identifying assumption concerning the marginal cost component pre-merger is that the structural error term vector representing unobserved cost characteristics ω^{pre} is uncorrelated to a M_ω-dimensional set of exogenous instruments $Z_{\omega^{pre}}$:

$$E[Z'_{\omega^{pre}}\omega_j^{pre}(\gamma^D, \gamma_0^S, \Theta)] = 0, \tag{14}$$

where γ_0^S reflects the true parameter value for γ^S. Together with the change in ownership, the marginal cost estimates will influence the predicted post-merger prices in the market.

Berry, Levinsohn, and Pakes (1995) argue that the computation of the optimal set of instruments when only conditional moment conditions are available is very difficult and numerically complex. As a less computationally demanding approximation, they use polynomials resulting from first order basis functions of the product characteristics. The validity of these basis functions as instruments relies on exchangeability assumptions of firms' own characteristics with respect to permutations in the order of competitors' product characteristics. Because I allow for the possibilty of collusion among firms, this changes the structure of potential Nash equilibria. The brand specific unobservable marginal cost component ω may be correlated with unobservable product characteristics. Therefore it is essential to look for instruments that are correlated with marginal costs, but not with the structural cost error. To account for the effects of unobserved cost drivers on prices, I use first order basis functions of the own brand characteristics, own firm characteristics, and competitors' characteristics.

4.2 Model identification of post-merger profit internalization parameter

Setting equation (11) equal to the observed post-merger prices p^{post}, and solving for the unobserved post-merger cost-component vector $\omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$, yields:

$$\omega_{\tilde{\theta}}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = p^{post} - \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) - \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post}. \tag{15}$$

As an identification restriction for the degree of joint profit maximization, I use orthogonality conditions between the residual of observed and predicted post-merger prices, which results in the structural error $\omega_{\tilde{\theta},j}^{post}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$ for a product j, and a $M_{\tilde{\theta}}$-dimensional matrix of

[13] Explicitly using input prices to estimate a cost function is a difference to previous studies in the ready-to-eat cereal industry.

instruments $Z_{\tilde{\theta}}$.

The model consists of a system of J equations for the different products whose prices are functions of the profit internalization parameters, $\tilde{\theta}$. The main identification task is to find meaningful moments that allow to identify the parameters. Using the difference between the predicted and observed post-merger prices of all brands would result in only one moment, which would render estimation of more than one parameter infeasible. I instead make use of orthogonality restrictions to generate two additional sets of moments. First, because I treat product characteristics with respect to demand, x, as exogenous with respect to firm behavior in the short run, I can use them as instruments. This is analogous to the identification of the production cost. Second, an increase in consumer income will have a positive demand effect at a given price. If such an income shock does not translate in higher labor costs, then the shock should be uncorrelated with the unobserved post-merger cost component vector ω^{post}. I use regional income data and local consumer price indexes as additional instruments. This leads to the identification restriction

$$E[Z'_{\tilde{\theta}}\omega^{post}_{\tilde{\theta}}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})] = 0. \tag{16}$$

Making use of the structural error term $\omega^{post}_{\tilde{\theta}}$ to construct orthogonality conditions to identify the supply side parameters does not rely on the linear cost function specification and only requires separability between the observed parameters γ^S and the unobserved cost component $\omega^{post}_{\tilde{\theta}}$. Appendix B derives the structural error term under the assumption of a logarithmic cost function.

4.3 Model identification of industry conduct parameters

Comparing to the case when estimating industry conduct, the main difference is that one estimates the industry ownership matrix Θ instead of the internalization parameters $\tilde{\theta}$. Equating equation (12) with the observed post-merger prices p^{post}, and solving for the unobserved post-merger cost error vector $\omega^{post}_{\Theta}(\gamma^D, \gamma^S, \Theta; b())$ yields:

$$\omega^{post}_{\Theta}(\gamma^D, \gamma^S, \Theta; b(\cdot)) = p^{post} - \hat{mc}^{post}(\gamma^D, \gamma^S, \Theta) - \Omega^{-1}(\gamma^D, \Theta, b(\Theta))s^{post}. \tag{17}$$

Besides the conventional first- and second-order polynomials of the observed product characteristics x, I use additional cross-firm polynomials which indicate the proximity between different firms' product portfolios. From a theoretical viewpoint, the proximity of two firms' brand portfolios should at least partly determine the potential profits from collusion between those firms. Furthermore, this should also determine the maximum degree of sustainable collusion between them. Denote x_{sj} as the s^{th} component of product j's observed product characteristic vector x_j. Denote by $h(x_{sj}, f)$ a $J \times 1$ row vector whose entries consists of the average of brand characteristic x_s between firm f and the average brand characteristics of each product's firm. Thus, the entries of this vector are 0 whenever a brand belongs to firm f. In case a brand belongs to firm $g \neq f$, the entry will be $\frac{1}{J_f}\sum_{i \in \mathbb{F}_f} x_{si} - \frac{1}{J_g}\sum_{j \in \mathbb{F}_g} x_{sj}$, where J_f denotes the number of brands of firm f. Under the assumption that cooperation between firms also depends on the proximity of the brand portfolios, given the correct form of conduct, the unobserved cost components of a firm should on average be uncorrelated with the differences in average brand characteristics $h(x_{s.,} .)$. For a given product characteristic s, this yields the additional moment restrictions

$$E[\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))h(x_{s.,}, f)] = 0, \forall f \ \in \{1, .., N\}.$$

Per characteristics used, this will results in $N-1$ additional moment restrictions. As in the profit-internalization case, I will also use data on disposable income as an instrument for the unobserved cost component.

Stacking the different instruments into the M_Θ-dimensional instrument matrix Z_Θ, this yields the following identification restrictions:

$$E[Z'_\Theta \omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))] = 0. \tag{18}$$

One key assumption is that industry conduct is known among firms. Relaxing this assumption would cause two problems. First, this would make the assumption on symmetric behavior between two different firms harder to sustain. Second, I would have to specify beliefs of the different firms regarding other firms' behavior, which could not be identified.

Relationship to Corts' (1999) Critique Previous research has used a conjectural variation approach in order to identify industry conduct, see for example Bresnahan (1989). In these models, a firm forms a "conjecture" about the responses of their competitors towards an increase in its own quantity. In this context, a conjecture can be seen as a reduced-form game theoretic best response function in symmetric quantity setting games. Corts (1999) critically discusses the identification of conjectural variation parameters. He shows that a conjectural variation parameter only estimates the marginal responsiveness of the marginal cost function with respect to changes in a demand shifter. As a researcher, one is however interested in the average slope of the marginal cost function instead of the marginal slope. My approach differs significantly from the conjectural variations approach and is not subject to this critique. In my framework, each firm sets prices for its portfolio of brands instead of quantities. Instead of forming conjectures about other brands' reactions, each firm's underlying objective function includes preferences for profits of other firms, thus allowing for cooperation among different firms. The preference parameters with respect to other firms' profits are essentially the conduct parameters I am interested in. I assume that these conduct parameters, as well as the marginal costs of all brands, are common knowledge in the industry, but not observed by the researcher. Using first order conditions of all brands' objective functions, my identification strategy allows to estimate both marginal cost parameters and the level conduct parameters. These amount to the "as-if conduct parameters" in Corts (1999).

Corts' also criticizes the static game character of conventional conduct estimation models. My approach is not fully exempt from this critique. I partially account for industry dynamics by modeling the merger-induced industry change. Nonetheless, my static approach may not detect certain dynamic collusion patterns. One big advantage of a static approach is a higher degree of tractability. Modeling repeated games makes identification of conduct even more difficult due to a larger set of potential dynamic equilibrium strategies. With my approach, I am also able to identify patterns of full collusion as well as patterns of collusion between only a subset of firms.

4.3.1 Rank conditions for industry conduct

This section provides identification results for different specifications when estimating continuous conduct parameters "directly". This is opposed to the menu approach, which selects among different non-nested models without estimating conduct parameters.

Recall the assumptions made on firms' own-profit maximization. As in standard unilateral merger models, I also assume that a merger does not change the behavior between non-merging firms. There are furthermore some global assumptions that reduce the parameter space which I will discuss in detail.

I only consider cases in which a firm treats all brands of a specific competitor's firm in the same way. This excludes the possibility that single brands of different firms collude while others play against each other competitively. From a pure rank condition perspective the number of parameters I would have to estimate when accounting for brand-specific collusion between firms would easily exceed the number brands in the market. This makes it impossible to identify the parameters.

Bilateral symmetry between firms One way to reduce the number of parameters to be estimated is to restrict the model to cases in which all brands of two firms play against each other in the same way. As a consequence, all brands have the same cross-conduct parameters for all of their brand pairs. This still allows for partial collusion between two firms, but does not allow for more elaborate strategies, such as for example collusion only between some brands of two firms. In terms of the parameter space, this reduces the number of cross-conduct parameters to $\frac{N(N-1)}{2}$.

Proposition 1 (Necessary conditions for identification under bilateral symmetry between firms)**.** *Suppose Assumption* 1 *holds, and that for distinct firms* f, g, $\theta_{ij} = \theta_{ik} = \theta_{ji} = \theta_{ki}$ $\forall i \in \mathbb{F}_f, \forall\ j,k \in \mathbb{F}_g$. *Then industry conduct is identified only if the number of firms is sufficiently small compared to the number of products, i.e. if* $\frac{N(N-1)}{2} \leq J$.

Proof: See Appendix E.

Same responsiveness to all cross-firm brands Another possibility is a case in which each firm behaves in the same way to all of its competitors.

The advantage of this specification is that it reduces the number of parameters to only N different cross-conduct parameters. However, there are also several problems associated with the assumption. First, it is again no longer possible to detect partial collusion between a subset of firms in the industry. Second, there is a consistency problem with respect to a mutual responsiveness: Under this assumption, it can be possible that firm 1 is acting collusively with firm 2, and firm 2 on the other hand acts competitively towards firm 1, something which is hard to justify from an economic perspective.

Proposition 2 (Necessary conditions for identification under same responsiveness to all cross-firm brands)**.** *Suppose Assumption* 1 *holds, and that for distinct firms* f, g, h *,* $\theta_{ij} = \theta_{ik}\ \forall i \in \mathbb{F}_f, \forall\ j \in \mathbb{F}_g, \forall\ k \in \mathbb{F}_h$*. Then rank conditions are met only if* $N \leq J$*.*

Proof: See Appendix E.

It is easy to see that the necessary rank conditions hold trivially. It can still be the case however that there are two or more identical conduct equations, which would violate identification.

Same responsiveness between all firms The most restrictive specification assumes that the cross-conduct parameters are identical for all brands in the market. The biggest advantage is that this returns a single cross-conduct parameter instead of a complicated matrix, and thus always meets the rank conditions. One disadvantage is that very often this parameter will severely restricts the set of estimable economic models. For example, one will not be able to test for partial collusion in the market, or for differences in competitive behavior between different firms.

Proposition 3 (Necessary conditions for identification under same responsiveness between all firms)**.** *Suppose for distinct firms* f, g, h *,* $\theta_{ij} = \theta_{ji} = \theta_{ik} = \theta_{jk} = \theta_{kj}\ \forall i \in \mathbb{F}_f, \forall\ j \in \mathbb{F}_g, \forall\ k \in \mathbb{F}_h$*. Then the rank condition for industry conduct is always met.*

Proof: See Appendix E.

Overall, the direct approach requires more structure and a larger parameter space than conventional selection methods. This is because explicitly accounting for conduct parameters requires more degrees of freedom. Therefore, I will provide results specifically tailored for the different assumptions provided in the beginning of this section. Clearly, the most important trade-off is the one between the allowed flexibility of industry conduct and the number of parameters that have to be estimated. Table 3 sums up the necessary rank conditions for all cases.

Identifying industry conduct via product entry or exit Besides using a merger as an identification strategy for estimating industry conduct, one can also think about using other structural changes. Concerning product entry, there is the problem of comparing competition with and without the entrant. While one can still make the assumption that entry does not change how existing brands compete with each other, one has to define how a new product will interact with the existing products.

Unlike product entry, using product exit as an identification strategy is still feasible. However, one has to ask why a product will exit. One reason can be that it is just not profitable, which will then probably imply that its impact on the market is relatively low. Therefore, a reduction of the brand space would not result in a big shift for firms strategies. Another possibility would be that a brand is profitable on its own, but it would be more profitable for a multi-brand firm to exit the product out of the market. This would result in an endogeneity problem when estimating conduct using product exit.

5 Estimation

5.1 Demand estimation

I use the technique of Nevo (2001) to recover the structural demand side parameters γ^D and the unobservable error term $\xi(\gamma^D, x, p)$. Using Nevo's estimation strategy on the demand side allows me to estimate all the demand side parameters independently of the supply side. I solve for the mean utility level across all brands at market t, $\delta_{.t}$, as to match the empirical market shares $s_{jt}(x_{.t}, p_{.t}, \xi_{.t}, \gamma^D)$ from equation (13) with the actual market shares s_{jt} observed in the data. Following equation (13) the objective is to find:

$$\hat{\gamma}^D = \arg\min_{\gamma^D} \xi(\gamma^D, x, p)' Z_\xi \tilde{A}_\xi^{-1} Z_\xi' \xi(\gamma^D, x, p); \tag{19}$$

where $\tilde{A}_\xi^{-1}$ is an estimate of the asymptotically efficient covariance $E[Z_\xi' \xi(\tilde{\gamma}^D, x, p) \xi(\tilde{\gamma}^D, x, p)' Z_\xi]$, given demand parameters $\tilde{\gamma}^D$ obtained from the first-stage GMM estimation.

Defining the market size is an important assumption, for it has implications on the different market shares and also on the differences between markets. I assume that the market size is correlated with store specific characteristics. I compute the market size of a specific store as a function of the average total sales of all supermarket products sold in this store.[14]

Estimates Table 4 shows results for a Random Coefficients Logit demand model. In this specification I include random coefficients for price, a constant, and sogginess of cereal. I furthermore include coefficients for sugar content, content of refined grains, segmentation dummies, a time dummy as well as firm dummies. The inclusion of firm dummies reflects controlling for firm-specific valuations, i.e. accounting for a fixed firm value rather than a brand value. Furthermore, I use demographic data on mean income, income standard deviation, household size and on number of small children to interact them with the random coefficients. The results show a negative relationship between income and price sensitivity,

[14] Because revenue from ready-to-eat cereal only amounts to a very small fraction of the total revenue generated in a store, the endogeneity between the market size variable and the cereal prices is negligible.

which is consistent with higher markups in high income neighborhoods. Price sensitivity also interacts negatively with the number of small children, which might account for their responsiveness to advertising. However, both demographic interaction coefficients are not statistically significant. Appendix C shows details about the estimation routine and other computational issues.

As a robustness check, I also estimate different variants of a multinomial Logit model. Table 5 shows demand side estimation results for several specifications of the multinomial Logit model. I use input prices and prices of other zones together with the ownership change as instruments for the sales price. When also including firm dummies yields a more elastic demand curve than specification (6) without instruments, however, the price coefficients are relatively close to each other. A bigger difference occurs between specifications that include and do not include firm dummies. This can be seen by comparing specifications (1) and (2) to (3) and (4). Overall, all of the price coefficients are lower in absolute magnitude than the mean coefficient of a Random Coefficient Logit estimation. This suggest that the random coefficient model is able to capture some of the consumer heterogeneity through interacting demographic variables which increases the price coefficient in absolute terms.

Demand Elasticities Individual market shares depend on the mean utility as well as on the random and demographic components. Product j's market share for individual i at market t can be written as:

$$s_{ijt} = \frac{exp(\delta_{jt} + \mu_{ijt})}{\sum_{k=0}^{J} exp(\delta_{kt} + \mu_{ikt})} \tag{20}$$

Integrating over the whole distribution of individuals yields the aggregated market shares from the model. The cross-price elasticity between goods j and k at market t, η_{jkt}, can be written as

$$\eta_{jkt} = \begin{cases} \frac{-p_{jt}}{s_{jt}} \int \alpha_i s_{ijt}(1 - s_{ijt}) dP_D(D) dP_v(v) & j = k \\ \frac{p_{kt}}{s_{jt}} \int \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) & j \neq k \end{cases} \tag{21}$$

When using the random coefficients model, one needs to compute the individual market shares using the model structure in equation (20). Table 6 and Table 7 show the mean elasticities over all markets for the baseline random coefficient Logit specification. The own-price elasticities are highly negative for all firms, with the exception of Kellogg's Just Right Fruit, which has an absolute own-price elasticity lower than one. One potential reason for this can be an increased popularity of adult cereals over the period in my dataset, such that both price and demand went up for this cereal at the same time.

There is furthermore significant variation in different brands' substitution patterns, which is related to the type of cereal. Adult cereals, such as Kellogg's Just Right and Kellogg's Nutri Grain, exhibit much lower cross-price elasticities than kids cereals, as for example Kellogg's

Fruit Loops or General Mills Honey Nut Cheerios with coefficients higher than .1. Overall, the substitution patterns are relatively close to previous industry estimates, see for example Nevo (2001).

5.2 Post-merger profit internalization

I first outline each step of the estimation algorithm when estimating the degree of joint profit maximization.

Estimation algorithm

1. **Estimate demand parameters γ^D:** In a first step, I estimate the demand parameters without having to specify the supply side.

2. **Recover pre-merger marginal costs using first order conditions, estimate cost function, and predict post-merger prices** Under the assumption of a known pre-merger industry ownership matrix Θ, I use equations (5) and (8) to back out pre-merger marginal costs. Using variation in the input costs w over time, I then predict the post-merger input price component of marginal costs, $\hat{mc}$.

3. **Pick degree of profit-internalization $\tilde{\theta}$, predict post-merger prices, and compute appropriate moments** I predict the markup firms charge conditional on a specific $\tilde{\theta}$ value. Together with the estimated post-merger input price component of marginal costs from step 2., I then predict the post-merger prices.

4. **Repeat** 3. **until GMM criterion is minimized** I recover the post-merger unobserved cost component $\omega_{\tilde{\theta}}^{post}$ and interact it with the instruments $Z_{\tilde{\theta}}$. I estimate the model using Generalized Method of Moments (GMM) to find the parameters that minimize the weighted moment criterion.

I assume that at a given point in time, marginal costs are constant across all stores. All stores are within the same metropolitan area and are operated by the same retailer. Therefore, the only channels through which marginal costs could differ across stores are either a difference in the retail margin across stores, or a difference in distribution costs. I do not find evidence for structural differences regarding the retail margin in the data. Differences in the distribution costs also do not seem likely because of the relative proximity of the stores.[15]

[15]There may be differences in store-specific fixed costs due to differences in rents or wages between the store locations. Such effect would not translate in marginal cost differences, but may be a channel for cost synergies after a merger.

In the second step of my estimation procedure, I use the marginal costs that were backed out conditional on a specific form of industry conduct and estimate the marginal cost equation (8) via minimizing the following objective function:

$$\hat{\gamma}^S = \arg\min_{\gamma^S} \omega^{pre}(\gamma^D \gamma^S, \Theta)' Z_\omega A^{-1}_{\omega^{pre}} Z'_\omega \omega^{pre}(\gamma^D, \gamma^S, \Theta), \tag{22}$$

where $A^{-1}_\omega = Z'_\omega Z_\omega$, therefore this amounts to a linear GMM estimator.

The brand specific unobserved marginal cost component ω^{pre}_j may be correlated with unobservable product characteristics. Therefore it is essential to look for instruments that are correlated with marginal costs, but not with the error term. To account for the effects of unobserved cost drivers on prices, I use first order basis functions of the own brand characteristics, own firm characteristics, and competitors' characteristics. This relies on an exchangeability argument of product characteristics when facing a unique Nash equilibrium, see for example Berry, Levinsohn, and Pakes (1995).

Having obtained the demand side coefficients γ^D and the cost parameters γ^S for the given form of pre-merger industry conduct Θ, I estimate the degree of profit-internalization $\tilde{\theta}$ by minimizing the GMM objective function:

$$\hat{\tilde{\theta}} = \arg\min_{\tilde{\theta}} \omega^{post}_{\theta}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) Z_{\tilde{\theta}} \tilde{W}^{-1}_{\tilde{\tilde{\theta}}} Z'_{\tilde{\theta}} \omega^{post}_{\tilde{\theta}}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}), \tag{23}$$

where $\tilde{W}_{\tilde{\theta}}(\tilde{\tilde{\theta}}, \tilde{\gamma}^S)$ is a consistent estimate of the covariance matrix $E[Z'_{\tilde{\theta}} \omega^{post}_{\tilde{\theta}}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta, \tilde{\tilde{\theta}}) \omega^{post}_{\tilde{\theta}}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta, \tilde{\tilde{\theta}})' Z_{\tilde{\theta}}]$ for a given first-stage parameter vector $\tilde{\tilde{\theta}}$.

The moments consist of the empirical residuals $\omega^{post}_{\tilde{\theta}}(\gamma^D, \gamma^S, \Theta, \tilde{\theta})$ interacted with the specific instruments $Z_{\tilde{\theta}}$, as described in Section 4.

There are several advantages from using the actual post-merger prices instead of simulating a post-merger price equilibrium. First, when simulating for a new price equilibrium, one needs to make an assumption regarding competition in the market. Already without estimating industry conduct, this is computationally demanding. Furthermore, it does not make use of all the available post-merger data, i.e. market shares and prices. Second, by using post-merger price simulation, one also risks averaging out specific competitive patterns and introduces a simulation error.[16]

Estimates Table 8 shows the cost function estimates for four symmetric forms of industry competition, ranging from multi-brand Nash competition, i.e. $\Theta = 0$, to full collusion, i.e.

[16]One important issue concerns the standard errors. Because of the sequential character of the estimation routine, I have to account for the demand estimation error when estimating standard errors for marginal costs and the supply side parameters, respectively. I account for these effects by using a second estimation routine. After having obtained the parameter estimates of the estimation algorithm, I estimate a sequential model that estimates all of the parameters simultaneously. This has the advantage of increasing efficiency of the estimation, as well as yielding consistent standard errors. Its disadvantage is the computational power required for this estimation. So far, I do not account for the estimation error when computing standard errors, which can potentially cause a bias in these estimates.

$\Theta = 1$. The median marginal costs implied by the model lie between \$.114 per serving for multi-product Nash pricing and \$.072 for full collusion, implying median markups between 40.8 and 63.1 percent. The cost function estimations show that especially the influence of wheat on overall marginal costs is decreasing in the degree of industry competition, Θ. Figure 6 shows the development of profit internalization parameters over time for four different forms of industry competition, ranging from multi-brand Nash competition to partial collusion, i.e. $\Theta = .5$, between all firms pre-merger. The results indicate an increasing degree of profit internalization over the first six quarters for all three industry specifications. The parameter values are the highest for the Nash specification and are decreasing in the degree of industry cooperation. In the last two quarters of 1994, there is a sharp drop in the profit internalization, which is followed by a sharp increase over the last year. However, the parameter estimates are not statistically significant under this specification. Figure 7 accounts for heterogeneity across the merging' entities with respect to the degree of profit internalization for multi-product Nash pricing, i.e. $\Theta = 0$. One can see that Post's profit internalization drops in the last two quarters of 1994 and subsequently increases. Overall, except for the drop in joint profit internalization in 1994, the results for the whole post-merger entity are consistent with a weak increase in profit internalization over time.

5.3 Industry conduct estimation

When estimating industry conduct, I have to iterate the processes of recovering pre-merger marginal costs, predicting post-merger marginal cost using a cost function estimation, and computing the industry conduct moments. This is because my object of interest, i.e. the pre-merger conduct matrix Θ, influences the implied marginal cost in the industry. Overall, the above steps can be decomposed into two parts. I use a nested two-step routine on the supply side. In the first step, I back out marginal cost conditional on a specific form of conduct as the outer loop. In the second step, I recover the supply side parameters by regressing the backed out marginal costs on the observable cost characteristics while controlling for unobserved brand characteristics. First, I will outline the conduct estimation routine.

1. **Estimate demand parameters** Using the instruments discussed above, I estimate the demand parameters, without having to specify supply-side competition.

2. **Pick Θ given the identification restrictions**

3. **Infer marginal costs and predict post-merger prices for given choice of Θ, and compute appropriate moments** Having estimated the demand side parameters, I can infer the marginal costs of production conditional on the form of conduct Θ using proper instruments. Using post-merger input cost data and the estimated cost-parameters, I can predict post-merger marginal costs. Given the conduct matrix Θ and the estimated demand parameters γ_D, I can then predict post-merger prices given Θ.

4. **Repeat steps 2-3 until GMM criterion is minimized** I recover the post-merger unobserved cost component ω_{Θ}^{post} and interact it with the instruments Z_{Θ}. I estimate the model using Generalized Method of Moments (GMM) to find the conduct parameters that minimize the weighted moment criterion.

In the second step of my estimation procedure, I use the marginal costs that were backed out conditional on a specific form of industry conduct and estimate the cost function from equation (8) via minimizing the following GMM objective function:

$$\hat{\gamma}^S = \arg\min_{\gamma^S} \omega^{pre}(\gamma^D, \gamma^S, \Theta)' Z_{\omega} \tilde{A}^{-1}_{\omega^{pre}} Z'_{\omega} \omega^{pre}(\gamma^D, \gamma^S, \Theta). \tag{24}$$

$\tilde{A}^{-1}_{\omega}$ is a consistent estimate of the covariance $E[Z'_{\omega}\omega^{pre}(\hat{\gamma}^D, \tilde{\gamma}^S, \Theta)\omega^{pre}(\hat{\gamma}^D, \tilde{\gamma}^S, \Theta)' Z_{\omega}]$ for a given first-stage parameter vector $\tilde{\gamma}^S$.

Having obtained the demand side coefficients γ^D and the cost parameters γ^S for any form of industry conduct, I estimate the conduct parameters Θ by minimizing the GMM objective function. The moments consist of the empirical residuals $\omega_{\Theta}^{post}(\gamma^D, \gamma^S, \Theta; b(\cdot))$ interacted with the specific instruments Z_{Θ}, as described in section 3.

Then the GMM objective in can be written as:

$$\hat{\Theta} = \arg\min_{\Theta} \omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta; b(\cdot)) Z'_{\Theta} \tilde{W}_{\Theta}^{-1} Z'_{\Theta} \omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \Theta; b(\cdot)), \tag{25}$$

where $\tilde{W}(\tilde{\Theta}, \gamma^S)$ is a consistent estimate of the covariance $E[Z'_{\Theta}\omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \tilde{\Theta}; b(\cdot))\omega_{\Theta}^{post}(\hat{\gamma}^D, \hat{\gamma}^S, \tilde{\Theta}; b(\cdot))' Z_{\Theta}]$ for a given first stage conduct matrix $\tilde{\Theta}$. Here, $\hat{\gamma}^S$ denotes the cost estimates from the second step that are conditional on a specific form of conduct.

Estimates Table 10 shows estimation results when estimating a single industry conduct parameter. The obtained parameter value is 0.708. It is interesting to compare the implied price cost-margins to those from a multi-product Nash pricing supply side model. Under multi-product Nash pricing, all of the markup can be attributed to product differentiation, and not to cooperative effects. When estimating a single conduct parameter, 25.6 percent of the price-cost margin is attributable to cooperative behavior between firms.

Table 11 shows the results when imposing symmetry in a firm's behavior towards all of its rivals. One can see that the two biggest players, General Mills and Kellogg's, act in the most cooperative behavior, while smaller companies act more competitively. According to this specification, General Mills acts fully cooperatively, with a parameter value of 0.98. Under this specification, 17.4 percent of the markups are attributable to cooperative behavior.

To account for even more heterogeneity with respect to behavior across firms, Table 12 shows the conduct estimation results under the assumption of bilateral brand symmetry. The parameter estimates show a lot of heterogeneity in the parameter values, however, under this specification, none of the parameters are statistically significant. The implied median price-cost margins from the estimation are 14.3 percent higher than the median multi-brand Nash price cost margins. This is lower than under a single conduct parameter specification, reflecting the heterogeneity across different firm pairs.

6 Extensions

In this section I present extensions to my basic framework that address several merger related issues.

6.1 Supply side selection methods

The menu approach selects the best fit among a discrete set ("menu") of supply side models, for example multi-brand Bertrand-Nash competition or full collusion among all firms. This approach does not include any explicit conduct parameters, but rather fully pre-imposes the form of competition, which often relaxes identification problems. In practice, there are two popular ways to select among different non-nested industry specifications. However, both have significant weaknesses.

The first method compares the marginal cost estimates of the different supply side specifications with cost estimates from other sources, such as accounting data, see for example Nevo (2001). At first sight this seems to be an intuitive way to select the most appropriate specification from the data. This approach, however, has several disadvantages. First, outside cost estimates are not always available, or do not have a reliable economic interpretation. Second, such data is often only available on a very aggregate level, which makes a detailed industry introspective nearly impossible.[17] Third, if different specifications yield similar cost estimates, it is not clear how one can use these results for a reliable model selection.

The second method uses forms of non-nested selection test in combination with pre-merger data to look for the supply specification that is closest to the true data generating process, see for example Vuong (1989) or Rivers and Vuong (2002). In vertical relations frameworks, such tests are relatively successful for selecting among different non-nested models, see for example Bonnet and Dubois (2010) for a detailed exposition. When using only pre-merger data in a horizontal framework, however, in practice there is the problem that such tests have relatively low predictive power, due to only very limited variation between the different pre-merger model specifications.

I exploit changes in ownership as well as variation in the input cost data to select among different non-nested horizontal models.Using pre-and post merger data in combination with non-nested supply side models provides a tractable in-sample test. Testing can be done using a J-Test or using a variant of the Rivers and Vuong (2002) test. Appendix D shows the estimation details for applying the Rivers and Vuong (2002) approach.

Table 13 shows the results of using the Rivers and Vuong (2002) approach for testing different non-nested hypotheses against each other. I use five non-nested specifications in

[17] Nevo (2001) states that his data is not sufficiently detailed to test for collusion among a subset of firms.

which each firm play symmetrically against each other, with values 0, .25, .5, .75, and 1. The results show that the non-nested test clearly rejects hypotheses of low industry cooperation against the hypotheses of high cooperation among firms. Overall, this approach would select a fully collusive model.

6.2 Direct estimation of synergies

From an antitrust viewpoint, the magnitude of synergies plays a key role for the welfare and consumer welfare effects of horizontal mergers, see for example Farrell and Shapiro (1990) and Nocke and Whinston (2010).[18] To my knowledge there is no approach that uses a differentiated goods framework to estimate the magnitude of merger related marginal cost synergies directly. I propose the following estimation method. Assume that industry conduct is known in an industry pre-merger and post-merger, and merging firms fully internalize their profits. When accounting for the change in price elasticities and the change in conduct after the merger, I can back out marginal costs both pre-merger and post merger via the vector of first-order conditions. When conduct and demand is known, the only systematic change can occur with respect to marginal costs. I will use information on the timing of the merger to assess the impact of the merger on marginal costs of the merging firms, which in economic terms reflects cost synergies.

Denote Θ^{pre} and Θ^{post} the known pre- and post-merger industry conduct, respectively. Then, using equation 8, I can back out the pre-merger and post-merger marginal cost from the model:

$$mc^{pre} = p^{pre} - \Omega^{-1}(\Theta^{pre})s^{pre}$$

$$mc^{post} = p^{post} - \Omega^{-1}(\Theta^{post})s^{post}$$

Define mc^{all} as the marginal cost vector both pre- and post-merger: $mc^{all} \equiv [mc^{pre}; mc^{post}]$.

I will now propose three different specifications one can use to estimate for synergies between merging firms given pre- and post-merger data.

1. Synergies in observable cost characteristics If one assumes that synergies affect all observable brand characteristics in the same way, but do not affect unobserved brand characteristics, then one can estimate the following equation:

$$mc_{jt} = (1 + \kappa \mathbb{1}_{merge,j})(\gamma^S w_j) + \omega_j + \epsilon_{jt}, \tag{26}$$

where κ represents the change in the observable brand characteristics on input prices, and $\mathbb{1}_{merge}$ is an indicator function equal to one if the brand belongs to one of the merging firms in the post-merger periods.

[18] One example for an industry with significant synergies is the beer industry. After the 2005 Coors-Molson merger, the company stated that it made $ 66 million worth merger related synergies in its first year as joined entity.

2. Synergies in unobservable cost characteristics If one assumes that synergies will affect only the unobserved brand specific component, then one can estimate the following equation:

$$mc_{jt} = \gamma^S w_j + \mathbb{1}^{pre}\omega_j^{pre} + (1 - \mathbb{1}^{pre})\omega_j^{post} + \epsilon_{jt}, \tag{27}$$

where $\mathbb{1}^{pre}$ is an indicator function equal to one in pre-merger periods.

3. Synergies in output A third possibility to account for synergies is to test for returns to scale in total firm output. This can account for increasing returns to scale in distribution cost or advertising. In this case, the marginal cost for brand j of firm f can be written as

$$mc_{jt} = \tau \sum_{i \in \mathbb{F}_f} log(q_i) + \gamma^S w_j + \omega_j + \epsilon_{jt}, \tag{28}$$

where q_i denotes the total quantity sold of brand i. Since I only observe output in one metropolitan area and brand, I have to assume that my data is representative for the average output over all retailers in the industry.

7 Conclusion

This paper proposes a framework to estimate the degree of joint profit maximization between merging firms and the form of industry conduct in the ready-to-eat cereal industry. The merger-induced ownership change serves as an important variation to identify firm behavior in the industry.

The availability of pre- and post-merger industry data allows me to estimate the degree of joint profit maximization rather than to assume it. The results shed light on the question of cooperation within a firm after a merger. The empirical descriptive findings show a partial pricing adjustment by the merging firms immediately after the merger. Furthermore, the structural estimation suggest an overall increase in the joint profit maximization over the first 10 quarters. The results are in line with informational and contractual frictions in a post-merger integration period.

The merged firm's pricing potential also crucially depends on the form of industry conduct. The biggest difference of my conduct estimation approach compared to other approaches lies in exploiting supply-side shifts using both pre- and post-merger industry data and thereby inferring the underlying degree of competitiveness. I do not have to rely on aggregate outside data or on relatively weak selection methods to determine the industry conduct. The estimation results suggest that markups in the industry are above those predicted under multi-product Nash pricing. I find that between 14.3 and 25.6 percent of the estimated markups can be attributed to cooperative industry behavior.

The proposed methods require sufficient variation in the price movement across different products or a variation in input prices. This might not be achieved in all horizontal merger cases. However, in the growing literature on ex-post merger evaluations there are already examples with seemingly sufficient variation, for example painkillers (Bjoernerstedt and Verboven 2012), motor oil and syrup (Weinberg and Hosken 2012), and cars (Yoshimoto 2011). Up until now, this literature mainly compares the predictions of different demand models under the assumption of multi-product Nash competition. Both methods can be applied to all of these mergers. Besides additional information about competition and merging firms' behavior within these industries, using such data would also give interesting information about behavior across industries. From an organizational perspective, this can yield insights about the effects of different managerial firm structures on post-merger behavior, and about differences in the potential to maximize joint profits. From an industry perspective, this can also provide information about the relationship between competition and market power across different industries.

References

Alonso, R., W. Dessein, and N. Matouschek (2008): "When Does Coordination Require Centralization?," *The American Economic Review*, 98(1), 145–179.

Appel, M., R. Labarre, and D. Radulovic (2004): "On accelerated random search," *SIAM Journal on Optimization*, 14(3), 708–731.

Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile prices in market equilibrium," *Econometrica*, 63(4), 841–890.

Bjoernerstedt, J., and F. Verboven (2012): "Does Merger Simulation Work? A "Natural Experiment" in the Swedish Analgesics Market," *mimeo.*

Bloom, N., C. Genakos, R. Sadun, and J. Van Reenen (2012): "Management practices across firms and countries," *The Academy of Management Perspectives*, 26(1), 12–33.

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

Bresnahan, T. (1982): "The oligopoly solution concept is identified," *Economics Letters*, 10(1-2), 87–92.

——— (1989): "Empirical studies of industries with market power," *Handbook of industrial organization*, 2, 1011–1057.

Brito, D., P. Pereira, and J. Ramalho (2012): "Mergers, Coordinated Effects and Efficiency in the Portuguese Non-Life Insurance Industry," *mimeo.*

Ciliberto, F., and J. Williams (2010): "Does Multimarket Contact Facilitate Tacit Collusion? Inference on Conjectural Parameters in the Airline Industry," *mimeo.*

Corts, K. (1995): *The ready-to-eat breakfast cereal industry in 1994 (A)*. Harvard Business School.

——— (1999): "Conduct parameters and the measurement of market power," *Journal of Econometrics*, 88(2), 227–250.

Cotterill, R., and A. Franklin (1999): "An estimation of consumer benefits from the public campaign to lower cereal prices," *Agribusiness*, 15(2), 273–287.

Dessein, W., L. Garicano, and R. Gertner (2011): "Organizing for Synergies," *American Economic Journal: Microeconomics*, 2(4), 77–114.

Farrell, J., and C. Shapiro (1990): "Horizontal mergers: An equilibrium analysis," *The American Economic Review*, 80(1), 107–126.

Fauli-Oller, R., and M. Giralt (1995): "Competition and cooperation within a multidivisional firm," *The Journal of Industrial Economics*, 43(1), 77–99.

Hitsch, G. (2006): "An empirical model of optimal dynamic product launch and exit under demand uncertainty," *Marketing Science*, 25(1), 25–50.

Knittel, C., and K. Metaxoglou (2011): "In Search of the Truth: Merger Simulations using Random Coefficient Logit Models," *UC Davis Department of Economics working paper*.

Lafontaine, F., and M. Slade (2012): "Inter-Firm Contracts: Evidence," in *The Handbook of Organizational Economics*, ed. by R. Gibbons, and J. Roberts. Princeton University Press.

Lau, L. (1982): "On identifying the degree of competitiveness from industry price and output data," *Economics Letters*, 10(1-2), 93–99.

McElheran, K. (2010): "Delegation in Multi-Establishment Firms: The Organizational Structure of IT Purchasing Authority," *US Census Bureau Center for Economic Studies Paper No. CES-WP-10-35*.

Nevo, A. (1998): "Identification of the oligopoly solution concept in a differentiated-products industry," *Economics Letters*, 59(3), 391–395.

——— (2000): "Mergers with differentiated products: The case of the ready-to-eat cereal industry," *The RAND Journal of Economics*, 31(3), 395–421.

——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2), 307–342.

Nocke, V., and M. Whinston (2010): "Dynamic Merger Review," *Journal of Political Economy*, 118(6), 1201–1251.

Oliveira, A. (2011): "Estimating Market Power with a Generalized Supply Relation: Application to an Airline Antitrust Case," *mimeo.*

Rivers, D., and Q. Vuong (2002): "Model selection tests for nonlinear dynamic models," *The Econometrics Journal*, 5(1), 1–39.

Rubinfeld, D. (2000): "Market definition with differentiated products: the Post / Nabisco cereal Merger," *Antitrust Law Journal*, 68, 163–185.

Schmalensee, R. (1978): "Entry deterrence in the ready-to-eat breakfast cereal industry," *The Bell Journal of Economics*, 9(2), 305–327.

Thomas, C. (2011): "Too Many Products: Decentralized Decision Making in Multinational Firms," *American Economic Journal: Microeconomics*, 3(1), 280–306.

Weinberg, M. C., and D. Hosken (2012): "Evidence on the Accuracy of Merger Simulations," *Review of Economics and Statistics, forthcoming.*

Yoshimoto, H. (2011): "Reliability Examination in Horizontal-Merger Price Simulations: An Ex-Post Evaluation of the Gap between Predicted and Observed Prices in the 1998 Hyundai–Kia Merger," *mimeo.*

8 Rank conditions examples

In this section I present further examples that highlight the effects of the assumptions made above. The main question will be under which circumstances marginal costs and industry conduct will be jointly identified in a model.

3 firms, brands 1 and 2 belong to same firm Consider an industry that consists of 4 brands, where brands 1 and 2 belong to the same firm. For simplicity, assume in this example that marginal costs are constant for each firm. Furthermore, denote by p_i, mc_i, s_i the price, marginal costs and market share of firm i, respectively. θ_{ij} describes the degree to which brand i takes into account the profits of brand j when making its decision. In the example, the maximization problem of brand 1 thus yields

$$\max_{p_1}(p_1 - mc_1)s_1(p) + (p_2 - mc_2)s_2(p) + \theta_{13}(p_3 - mc_3)s_3(p) + \theta_{14}(p_4 - mc_4)s_4(p)$$

The first-order condition for brand 1 with respect to its price then yields

$$(p_1 - mc_1)\frac{\partial s_1}{\partial p_1} + s_1 + (p_2 - mc_2)\frac{\partial s_2}{\partial p_1} + \theta_{13}(p_3 - mc_3)\frac{\partial s_3}{\partial p_1} + \theta_{14}(p_4 - mc_4)\frac{\partial s_4}{\partial p_1} = 0$$

There is a change in the ownership matrix pre- and post-merger if firms 2 and 3 merge. When making the additional assumption that each firm maximizes the profits of all of its brands, and merging firms fully internalize their profits, the associated pre- and post-merger conduct matrices can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & \theta_{13} & \theta_{14} \\ 1 & 1 & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & 1 & \theta_{34} \\ \theta_{41} & \theta_{42} & \theta_{43} & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta_{13} & \theta_{14} \\ 1 & 1 & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & 1 & 1 \\ \theta_{41} & \theta_{42} & 1 & 1 \end{pmatrix}$$

From firm 1's first order condition, conditional on the form of industry conduct, firms will adapt their prices after an ownership change. In the above example, without symmetry, there are 10 parameters to estimate, with only 4 equations, such that the rank conditions are never met for identification. I introduce different assumption on firm supply to reduce the number of parameters to be estimated.

Bilateral symmetry between firms Instead of bilateral brand symmetry, a stricter assumption is that for all brands of two distinct firms, each brand will take the other firms' brands into account in the same fashion when making its pricing decision. Pre-merger and post-merger conduct can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & \theta^a & \theta^b \\ 1 & 1 & \theta^a & \theta^b \\ \theta^a & \theta^a & 1 & \theta^c \\ \theta^b & \theta^b & \theta^c & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta^a & \theta^b \\ 1 & 1 & \theta^a & \theta^b \\ \theta^a & \theta^a & 1 & 1 \\ \theta^b & \theta^b & 1 & 1 \end{pmatrix}$$

This leads to a number of 3 parameters to estimate, with 4 available equations, such that the system is identified in absence of multi-collinearity.

Symmetry among all cross-firm brands When assuming that all brands take the brands of all other firms into account in the same way, this results in the following pre- and post-merger conduct:

$$\Theta = \begin{pmatrix} 1 & 1 & \theta^a & \theta^a \\ 1 & 1 & \theta^a & \theta^a \\ \theta^a & \theta^a & 1 & \theta^a \\ \theta^a & \theta^a & \theta^a & 1 \end{pmatrix}; \qquad \Theta^{post} = b(\Theta) = \begin{pmatrix} 1 & 1 & \theta^a & \theta^a \\ 1 & 1 & \theta^a & \theta^a \\ \theta^a & \theta^a & 1 & 1 \\ \theta^a & \theta^a & 1 & 1 \end{pmatrix}$$

There is only one conduct parameter to estimate and in 4 equations.

Post-merger internalization of profits When estimating the internalization of post merger profits among merging firms, the form of industry competition. If one assumes multi-product Nash pricing in the 3 firm, 4 brand case, then pre- and post-merger industry conduct can be written as

$$\Theta = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}; \qquad \Theta^{post} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & \tilde{\theta} \\ 0 & 0 & \tilde{\theta} & 1 \end{pmatrix}$$

Here, $\tilde{\theta}$ represents the degree of profit-internalization between firms 2 and 3.

9 Derivation of structural production error term using logarithmic cost function

Instead of the linear cost specification from equation (9), I now assume a logarithmic cost specification.

$$log(mc(\gamma^D, \Theta, p, x)) = w\gamma^S + \omega. \tag{29}$$

Combining the above equation with the recovered marginal cost estimates from equation (5), and solving for the pre-merger structural error term ω^{pre} yields

$$\omega_j^{pre}(\hat{\gamma}^D, \gamma^S, \Theta) = log\left(\frac{mc_j(\gamma^D, \gamma^S, \Theta, p, x, w)}{e^{\gamma^S w_j}}\right). \tag{30}$$

The post-merger structural error when estimating post-merger profit internalization can be derived as

$$\omega_{\tilde{\theta}^{post}}(\gamma^D, \gamma^S, \Theta, \tilde{\theta}) = log\left(\frac{p^{post} - \Omega^{-1}(\gamma^D, \Theta, \tilde{\theta})s^{post}}{\hat{mc}^{post}(\gamma^D, \gamma^S, \Theta, p, x, w)}\right). \tag{31}$$

The structural error term can be obtained in an analogous fashion. The computed error terms can then again be combined with instruments discussed in section 3.

10 Computational details

For the Random Coefficient Logit model, I use the derivative-based SOLVOPT algorithm. In practice, this algorithm has shown to provide more accurate estimation results in terms of a lower GMM objective function, see for example Knittel and Metaxoglou (2011). I also compute theoretical and numerical gradients at the equilibrium, which both are very close to zero, the highest gradient being of magnitude 0.003. In my estimation routine, I draw 50 individuals per store, and combine the draws with demographic store characteristics.

For my GMM supply-side routine, I use a basic finite-descent accelerated random search (ARS) algorithm, as proposed by Appel, Labarre, and Radulovic (2004). For each estimation step, I use 1500 starting points.

11 Rivers and Vuong approach

Formal Rivers and Vuong approach Consider a non-nested conduct specification h, with conduct matrix Θ^h at a time. Note that this can include differences in the conduct matrix between pre- and post-merger.[19] Under this hypothesis, one can recover the marginal cost using equation (5): $mc = p - \Omega^{-1}(\Theta^h)s$. When estimating a cost function, I decompose the structural unobserved error term into a fixed brand-specific unobserved component and a random shock component: $\omega_{jt}^h = \omega_j^h + \epsilon_{jt}^h$. Using my marginal cost specification 9 leads to

$$mc_{jt}^h = w_{jt}'\gamma^S + \omega_j^h + \epsilon_{jt}^h. \tag{32}$$

I assume that the random shock component uncorrelated to the observable input characteristics and to the unobserved brand-specific component:$E[\epsilon_{jt}^h|\omega_j^h, w_{jt}] = 0$.

The test looks for the cost equation with best statistical fit using the observable input

[19]This Appendix uses similar structure and notation as Bonnet and Dubois (2010).

characteristics from the cost function. Note that these characteristics are brand-specific, but do not change with the different conduct specifications. If one tests between two different models, h, and h', one obtains the following equations for the industry prices.

$$p_{jt}^{h} = \Omega^{-1}(\Theta^{h})s + w_{jt}'\gamma_{h}^{S} + \omega_{j}^{h} + \epsilon_{jt}^{h}$$

$$p_{jt}^{h'} = \Omega^{-1}(\Theta^{h'})s + w_{jt}'\gamma_{h'}^{S} + \omega_{j}^{h'} + \epsilon_{jt}^{h'}$$

For a given specification h, denote $Q_n^h(\gamma^S, \omega)$ the lack-of-fit criterion given the parameters γ^S and ω. For any specification h, one now aims to find the parameter values that minimize the lack of fit criterion Q^h:

$$\min_{\gamma_h^S, \omega_j^h} Q_n^h(\gamma_h^S, \omega_j^h) = \min_{\gamma_h^S, \omega_j^h} \frac{1}{n}\sum_{j,t}(\epsilon_{jt}^h)^2 = \frac{1}{n}\sum_{j,t}[p_{jt} - \Omega^{-1}(\Theta^h)s - \omega_j^h - w_{jt}'\gamma_h^S]^2$$

This does not require any of the specifications to be the correct model. Denote by $\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h)$ the expected lack-of-fit criteria for specification h.

Given specifications h and h', there are three different hypothesis one has to test against each other, which have the following asymptotic properties.

H_0 : h and $h^{'}$ are asymptotically equivalent if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} = 0$.

H_1 : h asymptotically better fit than h' if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} < 0$.

H_2 : h' asymptotically better fit than h if $lim_{n\to\infty}\{\overline{Q}_n^h(\overline{\gamma}_h^S, \overline{\omega}_j^h) - \overline{Q}_n^{h'}(\overline{\gamma}_{h'}^S, \overline{\omega}_j^{h'})\} > 0$.

Denote by T_n the test statistic that accounts for the variation in the lack-of-fit criteria for the different hypotheses.

$$T_n = \frac{\sqrt{n}}{\hat{\sigma}_n^{hh'}}\{Q_n^h(\hat{\gamma}_h^S, \hat{\omega}_j^h) - Q_n^{h'}(\hat{\gamma}_{h'}^S, \hat{\omega}_j^{h'})\}, \tag{33}$$

where $\sigma_n^{hh'}$ is the variance of the difference of the estimated lack-of-fit criteria. Rivers and Vuong show that if two models are strictly non-nested, the asymptotic distribution of T_n is standard normal. Thus, one has to compare sample values of T_n with the critical values of a standard normal distribution.

12 Proofs

Proof of Proposition 1

Proof. The demand parameters γ^D can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are J estimable equations, one equation per brand post-merger. Because each firm has one conduct parameter for each competitor, this leads to an overall number of $N(N-1)$ parameters. The bilateral symmetry assumption reduces this number to $\frac{N(N-1)}{2}$. This leads to J equations with $\frac{N(N-1)}{2}$ parameters. The model is only identified if there are at least as many equations as parameters, i.e. if $\frac{N(N-1)}{2} \leq J$. □

Proof of Proposition 2

Proof. The demand parameters γ^D can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are J estimable equations, one equation per brand pre-merger, and one equation per brand post-merger.Because each firm has one conduct parameter for all firms, this leads to an overall number of N conduct parameters. This leads to J equations with N parameters. The model is only identified if there are at least as many equations as parameters, i.e. if $N \leq J$. □

Proof of Proposition 3

Proof. Using the same reasoning as in the proof for Proposition 2, there are J equations and one parameter to estimate, so that the result trivially holds. □

13 Graphs and tables

Table 1: Product segmentation

Adult enhanced	Adult simple	Family	Kids
PO Raisin Bran	NAB Orig. Shrd. Wheat	GM Wheaties	PO Honeycomb
GM Raisin Nut Bran	NAB Spoon Size Shrd.	GM Cheerios	GM Apple-Cinn. Cheerios
KE Raisin Bran	PO Grape Nuts Cereal	KE Corn Flakes	GM Honey Nut Cheerios
KE Nutri Grain	GM Total Corn Flakes	KE Crispix	GM Lucky Charms
QO 100% Natural	KE Special K	RA Corn Chex	GM Trix
	KE Just Right Fruit Nut	RA Wheat Chex	KE Froot Loops
		RA Rice Chex	KE Frosted Flakes
			KE Corn Pops
			KE Smacks
			QO Cap'n Crunch

Table 2: Product specific price development

Brand Name	% 92Q3-93Q1	% 92Q3-93Q3	% 92Q3-94Q1
NAB Orig Shred Wheat	0.03	0.12	0.11
NAB Sp.Size Shrd Wheat	0.01	0.03	0.09
PO Grape Nuts Cereal	0.04	0.01	0.04
PO Raisin Bran	0.07	0.04	0.23
PO Honeycomb	0.08	0.08	0.10
GM Raisin Nut Bran	-0.02	0.03	0.12
GM Apple-Cin Cheerios	-0.02	-0.10	-0.05
GM Wheaties	-0.01	-0.12	-0.17
GM Cheerios	-0.09	-0.14	-0.09
GM Honey Nut Cheerios	-0.02	-0.07	-0.17
GM Lucky Charms	-0.07	-0.11	-0.05
GM TOTAL Corn Flakes	0.00	0.06	-0.09
GM Trix	-0.22	-0.12	0.07
KE Froot Loops	0.00	-0.14	-0.12
KE Special K	0.04	0.02	-0.09
KE Frosted Flakes	0.01	0.03	-0.13
KE Corn Pops	0.05	-0.23	-0.28
KE Raisin Bran	0.06	-0.06	-0.04
KE Corn Flakes	-0.02	-0.03	-0.24
KE Smacks	0.06	-0.01	0.13
KE Crispix	0.04	0.08	0.15
KE Just Right FruitNut	-0.24	0.02	0.18
KE Nutri Grain	0.03	0.01	0.06
RA Corn Chex	0.03	0.06	0.05
RA Wheat Chex	0.03	0.05	0.05
RA Rice Chex	0.02	0.05	0.06
QO 100% Natural Cereal	0.03	0.05	-0.14
QO Cap'n Crunch	-0.27	0.04	-0.12

Note: Column 1 shows of quantity-weighted average percentage deflated price change between quarter 3, 1992, and quarter 1, 1993. Columns 2 and 3 show the price developments between 1992, quarter 2 and 1993, quarter 3, and 1994, quarter 2, respectively.

Table 3: Identification conditions for different specifications

Conduct specification	Necessary condition
Bilateral symm btw. firms	$\frac{N(N-1)}{2} \leq J$
Same resp. to cross-firm br.	$N \leq J$
Same resp. btw. all firms	$J \geq 1$ (always met)
Menu approach	$J \geq 0$ (always met)

Table 4: Demand side estimates γ^S for Random Coefficient Logit model

Variable	Mean	Std. Dev.	Interaction Small Child	Interaction Income	Interaction Household Size	Interaction St.Dev. Income
price	-34.97	1.69	-387.90	-93.38	–	44.37
	(7.69)	(3.80)	(333.60)	(191.10)	–	(32.35)
const	-1.42	-.14	–	–	12.30	–
	(.55)	(.60)	–	–	(7.87)	–
mushy	.12	.01	6.33	1.36	–	–
	(.40)	(.23)	(19.14)	(12.14)	–	–
sugar	-.03					
	(.00)					
refined grains	.46					
	(.03)					
quarter trend	.02					
	(.00)					
adult segment	1.25					
	(.06)					
kids segment	.47					
	(.04)					

Note: Num Obs: 19600. Interactions with demographics from US 1990 Census data around each store. Omitted category: Firm dummy variables to account for firm-specific brand valuations.

Table 5: Demand estimation results for different Logit specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Price	-11.5485***	-11.6427***	-7.3154***	-10.2634***	-13.8593***	-11.2802***
	(0.2287)	(0.4481)	(0.2736)	(0.494)	(0.2376)	(0.1254)
Constant	-2.2791***	-3.2689***	-2.3519***	-1.8492***	-3.1581***	-2.147***
	(0.0487)	(0.1515)	(0.0602)	(0.1051)	(0.0878)	(0.0358)
Sugar	-0.0344***	-0.051***	-0.0143***	-0.0189***	-0.0528***	-0.0414***
	(0.0013)	(0.0022)	(0.0014)	(0.0021)	(0.0013)	(0.0007)
Fat	-0.0369***	0.0346***	0.0155***	0.0153***	0.0184***	0.0073***
	(0.004)	(0.0067)	(0.003)	(0.0052)	(0.0038)	(0.0025)
Refined grains	0.4682***	0.3481***	0.2636***	0.2679***	0.4377***	0.4383***
	(0.0207)	(0.0349)	(0.0277)	(0.0395)	(0.0198)	(0.0134)
Sales	0.7569***	3.4081***	2.3924***	5.0058***	1.5109***	0.1584***
	(0.067)	(0.2007)	(0.0918)	(0.215)	(0.073)	(0.0102)
Quarter dummy	-0.0001***	-0.0492***	-0.0304***	-0.0647***		
	(0.0016)	(0.0041)	(0.0023)	(0.0049)		
Kids Cereal	1.0309***	1.698***	0.7072***	0.9768***	2.7253***	1.581***
	(0.054)	(0.0816)	(0.0528)	(0.0784)	(0.1024)	(0.0338)
Adult Cereal	0.3237***	0.6644***	0.4202***	0.6491***	1.6531***	0.2053***
	(0.0301)	(0.0488)	(0.0327)	(0.0502)	(0.1019)	(0.0385)
Input price IV	Yes	No	Yes	No	No	No
Zone IV	No	Yes	No	Yes	Yes	No
Ownership IV	Yes	Yes	Yes	Yes	Yes	No
Firm dummies	Yes	Yes	No	No	Yes	No

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 19600.

Table 6: Mean Elasticities Random Coefficient Logit Model Part 1

	NAB ShW	NAB SSW	PO GNu	PO RBr	PO RNB	GM Hon	GM ACh	GM Whe	GM Che	GM HNC	GM LCh	GM TCF	GM Tri	KE FLo
NAB Shred Wheat	-3.835	0.038	0.043	0.056	0.071	0.022	0.037	0.036	0.113	0.122	0.094	0.055	0.056	0.123
NAB Sp Shr Whea	0.012	-3.366	0.045	0.057	0.069	0.022	0.039	0.035	0.117	0.125	0.090	0.049	0.054	0.115
PO Grape Nuts	0.010	0.035	-2.630	0.059	0.063	0.023	0.042	0.032	0.121	0.129	0.083	0.039	0.048	0.101
PO Raisin Bran	0.011	0.035	0.046	-2.910	0.063	0.023	0.042	0.032	0.123	0.130	0.084	0.040	0.049	0.102
PO Honeycomb	0.012	0.037	0.044	0.057	-3.713	0.022	0.039	0.035	0.116	0.125	0.092	0.051	0.055	0.118
GM RaisNut Bran	0.011	0.035	0.046	0.060	0.064	-4.232	0.044	0.033	0.129	0.137	0.087	0.041	0.051	0.105
GM AppCin Cheer	0.010	0.033	0.046	0.060	0.060	0.024	-2.899	0.031	0.126	0.133	0.080	0.036	0.046	0.095
GM Wheaties	0.012	0.037	0.044	0.057	0.070	0.023	0.039	-3.956	0.118	0.127	0.093	0.051	0.055	0.118
GM Cheerios	0.010	0.034	0.046	0.060	0.063	0.024	0.044	0.032	-3.495	0.134	0.084	0.039	0.049	0.102
GM HonNut Cheer	0.011	0.035	0.046	0.060	0.064	0.024	0.043	0.033	0.126	-3.613	0.086	0.040	0.050	0.104
GM Luck Charms	0.012	0.037	0.044	0.057	0.070	0.023	0.039	0.035	0.118	0.127	-4.167	0.052	0.056	0.119
GM TCorn Flakes	0.014	0.039	0.040	0.052	0.075	0.021	0.033	0.038	0.106	0.116	0.100	-4.580	0.061	0.134
GM Trix	0.012	0.038	0.044	0.056	0.071	0.023	0.038	0.036	0.117	0.126	0.095	0.053	-4.329	0.122
KE Froot Loops	0.013	0.038	0.043	0.056	0.072	0.022	0.037	0.036	0.114	0.124	0.096	0.056	0.057	-4.203
KE Special K	0.013	0.038	0.042	0.055	0.073	0.022	0.036	0.037	0.112	0.121	0.096	0.058	0.058	0.127
KE Frost Flakes	0.012	0.038	0.044	0.056	0.070	0.022	0.038	0.035	0.115	0.124	0.093	0.053	0.055	0.119
KE Corn Pops	0.013	0.038	0.043	0.055	0.072	0.021	0.036	0.036	0.112	0.121	0.095	0.057	0.057	0.125
KE Raisin Bran	0.011	0.036	0.046	0.059	0.066	0.023	0.041	0.033	0.121	0.129	0.087	0.044	0.051	0.108
KE Corn Flakes	0.011	0.035	0.046	0.059	0.063	0.023	0.042	0.032	0.120	0.128	0.083	0.040	0.048	0.101
KE Smacks	0.012	0.038	0.044	0.056	0.071	0.022	0.038	0.036	0.115	0.124	0.093	0.053	0.056	0.120
KE Crispix	0.013	0.039	0.041	0.053	0.074	0.021	0.034	0.037	0.107	0.117	0.098	0.062	0.059	0.131
KE JRight Fruit	0.008	0.028	0.045	0.057	0.049	0.022	0.045	0.025	0.119	0.123	0.063	0.023	0.036	0.071
KE Nutri Grain	0.011	0.036	0.045	0.059	0.067	0.025	0.042	0.034	0.127	0.136	0.091	0.046	0.054	0.113
RA Corn Chex	0.013	0.038	0.042	0.055	0.072	0.021	0.036	0.036	0.111	0.120	0.096	0.058	0.058	0.126
RA Wheat Chex	0.012	0.037	0.045	0.058	0.067	0.023	0.040	0.034	0.119	0.127	0.089	0.047	0.053	0.112
RA Rice Chex	0.013	0.039	0.042	0.054	0.073	0.021	0.035	0.037	0.109	0.119	0.097	0.060	0.059	0.129
QO 100 Natural	0.007	0.028	0.046	0.061	0.051	0.029	0.051	0.027	0.143	0.149	0.072	0.024	0.040	0.077
QO Capn Crunch	0.011	0.035	0.046	0.059	0.064	0.024	0.043	0.033	0.126	0.134	0.086	0.041	0.050	0.105

Table 7: Mean Elasticities Random Coefficient Logit Model Part 2

	KE	KE	KE	KE	KE	KE	KE	KE	KE	RA	RA	RA	QO	QO
	SpK	FFl	CPo	RBr	CFl	Sma	Cri	JRi	NGr	CCh	WCh	RCh	QO1	CCr
NAB Shred Wheat	0.105	0.142	0.097	0.079	0.066	0.014	0.077	0.000	0.005	0.028	0.016	0.038	0.023	0.065
NAB Sp Shr Whea	0.097	0.136	0.090	0.079	0.068	0.013	0.069	0.000	0.005	0.026	0.016	0.035	0.026	0.067
PO Grape Nuts	0.083	0.123	0.078	0.078	0.070	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.034	0.068
PO Raisin Bran	0.084	0.123	0.079	0.078	0.070	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.035	0.069
PO Honeycomb	0.100	0.138	0.093	0.079	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.026	0.067
GM RaisNut Bran	0.085	0.124	0.079	0.080	0.069	0.012	0.058	0.001	0.005	0.022	0.016	0.030	0.042	0.073
GM AppCin Cheer	0.077	0.117	0.073	0.078	0.070	0.011	0.052	0.001	0.005	0.020	0.015	0.027	0.041	0.070
GM Wheaties	0.099	0.138	0.092	0.079	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.027	0.067
GM Cheerios	0.083	0.122	0.078	0.079	0.070	0.012	0.056	0.001	0.005	0.022	0.016	0.029	0.040	0.071
GM HonNut Cheer	0.085	0.124	0.079	0.079	0.069	0.012	0.058	0.001	0.005	0.022	0.016	0.030	0.039	0.071
GM Luck Charms	0.100	0.138	0.093	0.080	0.067	0.014	0.072	0.000	0.005	0.026	0.016	0.036	0.028	0.068
GM TCorn Flakes	0.117	0.151	0.107	0.077	0.062	0.015	0.088	0.000	0.005	0.031	0.017	0.043	0.018	0.062
GM Trix	0.103	0.140	0.095	0.080	0.066	0.014	0.074	0.000	0.005	0.027	0.016	0.037	0.027	0.067
KE Froot Loops	0.106	0.143	0.098	0.079	0.065	0.014	0.077	0.000	0.005	0.028	0.016	0.038	0.024	0.066
KE Special K	-4.148	0.145	0.100	0.079	0.065	0.014	0.080	0.000	0.005	0.029	0.017	0.039	0.022	0.065
KE Frost Flakes	0.101	-3.519	0.094	0.079	0.067	0.014	0.073	0.000	0.005	0.027	0.016	0.036	0.024	0.066
KE Corn Pops	0.107	0.144	-3.871	0.079	0.065	0.014	0.079	0.000	0.005	0.028	0.016	0.039	0.022	0.065
KE Raisin Bran	0.090	0.129	0.084	-3.314	0.069	0.013	0.063	0.000	0.005	0.024	0.016	0.032	0.032	0.068
KE Corn Flakes	0.083	0.123	0.079	0.078	-2.401	0.012	0.057	0.001	0.005	0.022	0.016	0.029	0.032	0.067
KE Smacks	0.102	0.140	0.095	0.079	0.067	-3.904	0.074	0.000	0.005	0.027	0.016	0.037	0.025	0.066
KE Crispix	0.113	0.149	0.105	0.078	0.063	0.015	-4.208	0.000	0.005	0.030	0.017	0.042	0.019	0.063
KE JRight Fruit	0.056	0.093	0.054	0.070	0.069	0.009	0.035	-0.252	0.004	0.015	0.013	0.019	0.047	0.064
KE Nutri Grain	0.093	0.131	0.086	0.081	0.068	0.013	0.065	0.000	-4.748	0.024	0.016	0.033	0.038	0.072
RA Corn Chex	0.108	0.145	0.100	0.078	0.065	0.014	0.080	0.000	0.005	-3.992	0.017	0.039	0.021	0.065
RA Wheat Chex	0.093	0.133	0.087	0.079	0.069	0.013	0.066	0.000	0.005	0.025	-3.399	0.033	0.029	0.068
RA Rice Chex	0.111	0.147	0.102	0.078	0.064	0.015	0.083	0.000	0.005	0.029	0.017	-4.123	0.020	0.064
QO 100 Natural	0.058	0.093	0.054	0.075	0.067	0.009	0.035	0.001	0.005	0.015	0.014	0.019	-3.968	0.077
QO Capn Crunch	0.086	0.125	0.080	0.079	0.069	0.012	0.059	0.001	0.005	0.023	0.016	0.030	0.038	-3.761

Table 8: Cost function estimates γ^S

Industry competition	Parameter Value $\Theta = 0$	Parameter Value $\Theta = 0.25$	Parameter Value $\Theta = 0.5$	Parameter Value $\Theta = 1$
Wheat	0.123**	0.112***	0.091**	0.046**
Rice	0.057**	0.058***	0.061***	0.066***
Oat	0.019***	0.017***	0.014***	0.007***
Corn	0.092***	0.085***	0.077***	0.057***
Sugar	0.001***	0.001***	0.000***	0.000***
Vitamin	0.028***	0.027***	0.024***	0.02***
Firm Dummies	Yes	Yes	Yes	Yes
Median mc^{pre}	0.1036	0.0961	0.0864	0.0606
Mean mc^{pre}	0.1005	0.0930	0.0837	0.0586
Std. mc^{pre}	0.0669	0.0616	0.0564	0.0481
Median mc^{post}	0.1138	0.1062	0.0971	0.0721
Mean mc^{post}	0.1161	0.1087	0.0994	0.074
Std. mc^{post}	0.0807	0.0709	0.0608	0.0489

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Num Obs: 7840. Columns 2-5 show estimates for different forms of symmetric industry conduct, raging from full competition between firms to full collusion. Wheat, Oat, Corn, Sugar, Rice, Vitamins reflect coefficients for interaction between input prices and relative ingredient content in product.

Table 9: Joint profit maximization estimates $\tilde{\theta}$

Industry competition	$\Theta = 0$ (Nash) Estimated $\tilde{\theta}$	$\Theta = .25$ Estimated $\tilde{\theta}$	$\Theta = .5$ Estimated $\tilde{\theta}$
1-2 quarters	.30 (1.03)	.11 (.69)	.02 (.51)
3-4 quarters	.36 (1.05)	.19 (.71)	.02 (.49)
5-6 quarters	.43 (1.41)	.22 (.87)	.07 (.94)
7-8 quarters	.02 (.92)	.02 (.79)	.02 (.65)
9-10 quarters	.98 (1.88)	.98 (1.83)	.98 (1.95)

Note: Standard errors in parentheses. Num Obs: 9800. Degree of merging firms' joint profit maximization over time, $\tilde{\theta}$, for different degrees of symmetric industry competition, Θ.

Table 10: Estimation of single conduct parameter

	Conduct Parameter	Std. Dev
Inter-Firm Conduct	.708***	.023
Type of competition	Median PCM	St. Dev PCM
Estimated Conduct	.548	.216
Multi-brand Nash	.408	.285
Full collusion	.631	.217

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 7840. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

Table 11: Conduct estimates under symmetry to all firms

	Conduct Parameter	Std. Dev.
General Mills	.981***	.318
Ralston	.070	.206
Kellogg	.402***	.013
Post	.102	.106
Nabisco	.314***	.025
Quaker Oats	.195***	.070
Type of competition	Med. PCM	Std.PCM
Estimated Conduct	.494	.244
Multi-brand Nash	.408	.285
Full collusion	.631	.217

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 11760. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

Table 12: Conduct estimates under bilateral firm symmetry

	GM	RA	KE	PO	NA	QO
General Mills	1	.886	.046	.555	.799	.982
Ralston		1	.932	.856	.070	.982
Kellogg			1	.841	.086	.022
Post				1	.456	.977
Nabisco					1	.018
Quaker Oats						1
Type of competition	Med. PCM	Std.PCM				
Estimated Conduct	.476	.193				
Multi-brand Nash	.408	.285				
Full collusion	.631	.217				

Note: P-values: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; Num Obs: 11760. Recovered parameters from multi-product Nash pricing reflects the case of $\theta = 0$ for all cross-firm conduct parameters. Recovered parameters from full collusion reflects the case of of $\theta = 1$ for all cross-firm conduct parameters.

Table 13: Selection method results using Rivers and Vuong test

Cross-firm Conduct	$\Theta_{ij} = 0.25$	$\Theta_{ij} = 0.5$	$\Theta_{ij} = .75$	$\Theta_{ij} = 1$
$\Theta_{ij} = 0$	14.10	12.64	11.17	9.75
$\Theta_{ij} = .25$		11.35	9.84	8.35
$\Theta_{ij} = .5$			8.26	6.67
$\Theta_{ij} = .75$				4.90

Note: Test statistic follows a standard normal distribution. Num obs: 11760.

Figure 1: Pre-merger market shares

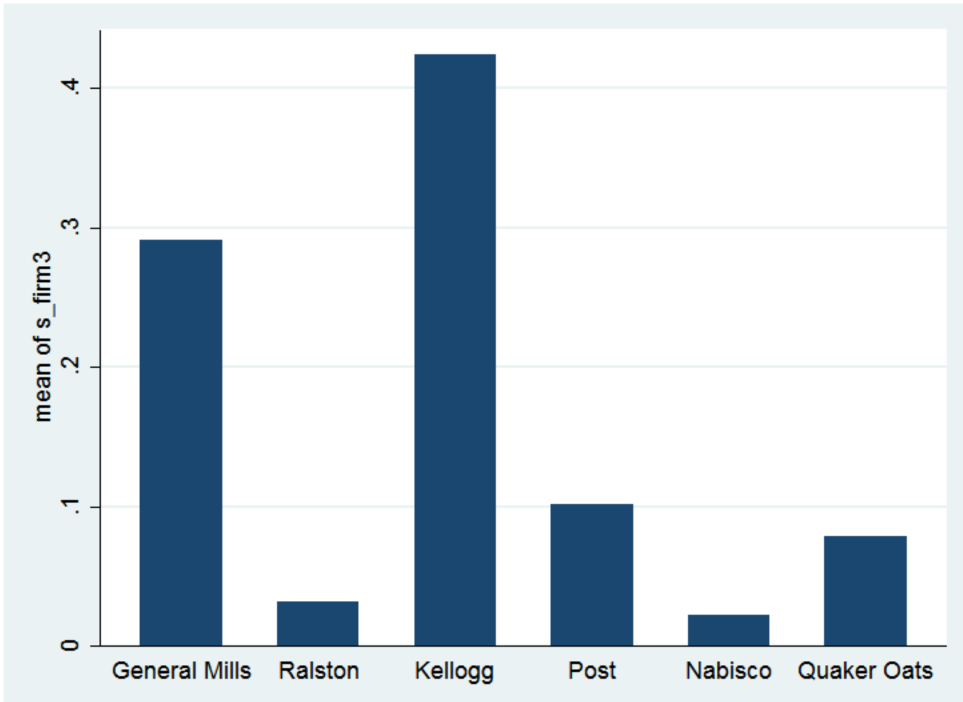


Figure 2: Geographical location of stores in dataset

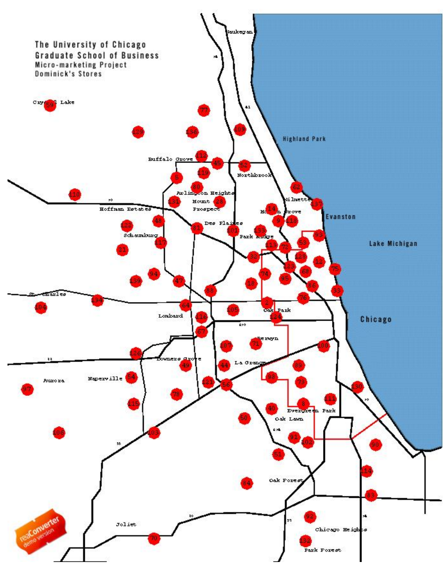
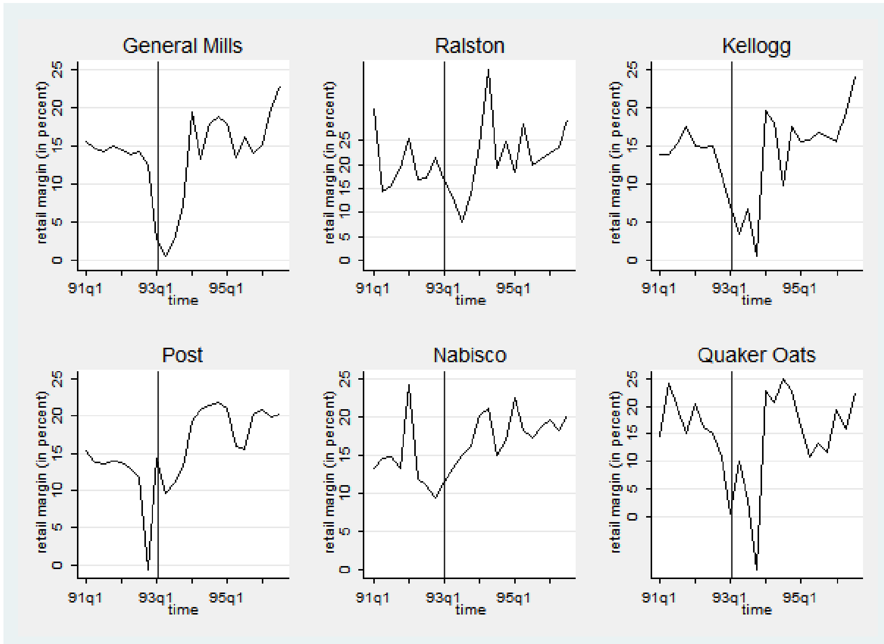
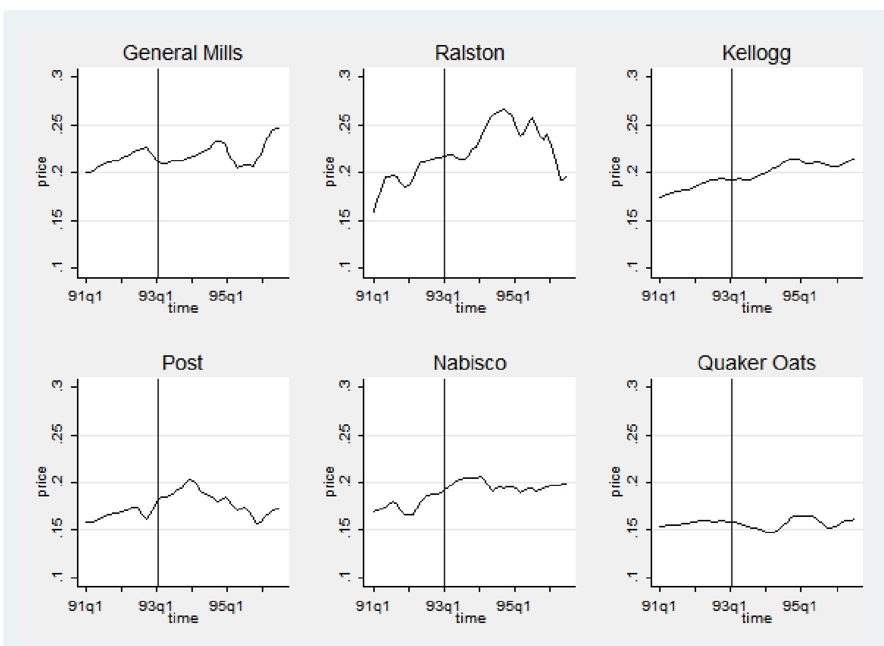


Figure 3: Retail margin development per firm



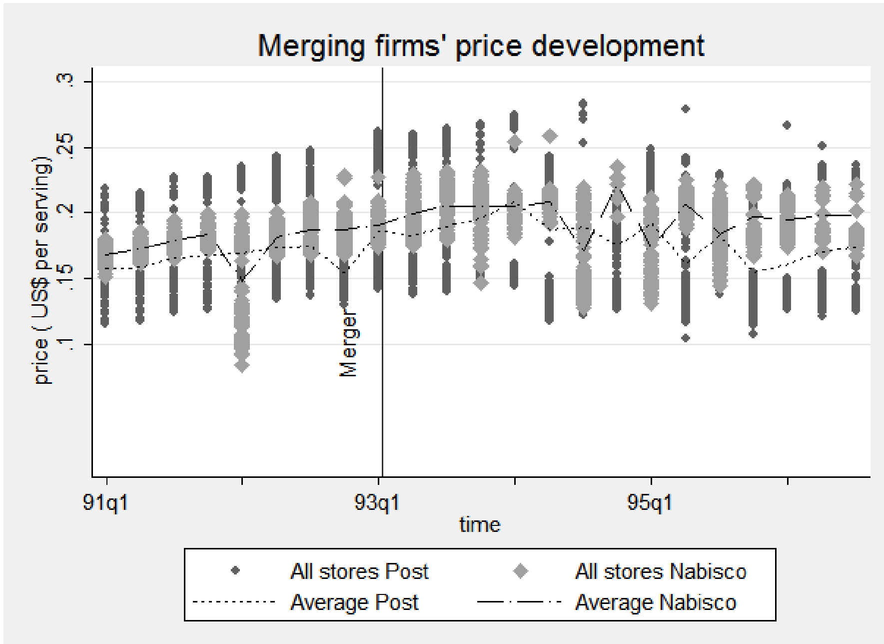
Note: Retail margin proxy computed as $1 - AAC$, where AAC are the average acquisition costs observed in the data for a firm per quarter. AAC reflects the percentage of the retail price paid to a producer.

Figure 4: Average price development per firm



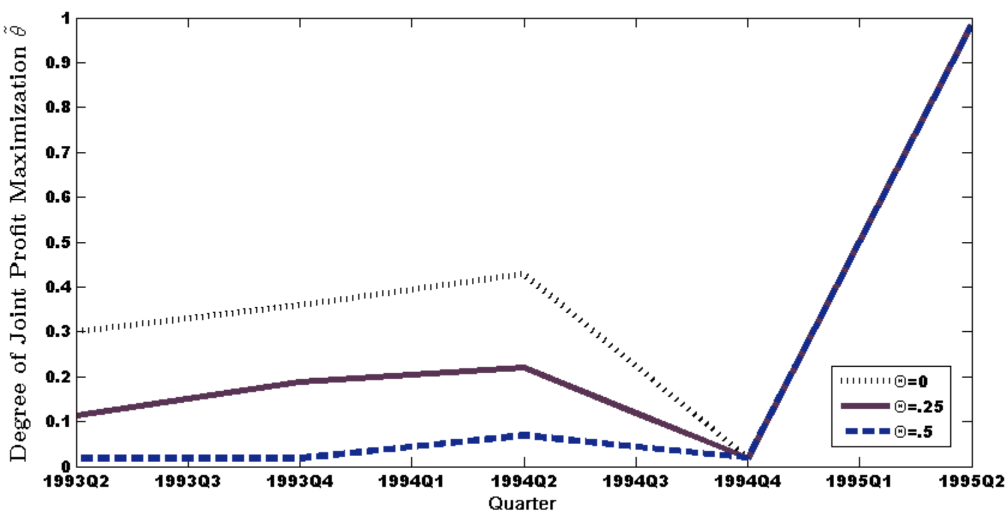
Note: Firm prices computed as deflated quantity-weighted average prices over all products sold of same firm per quarter.

Figure 5: Price development of merging firms across stores



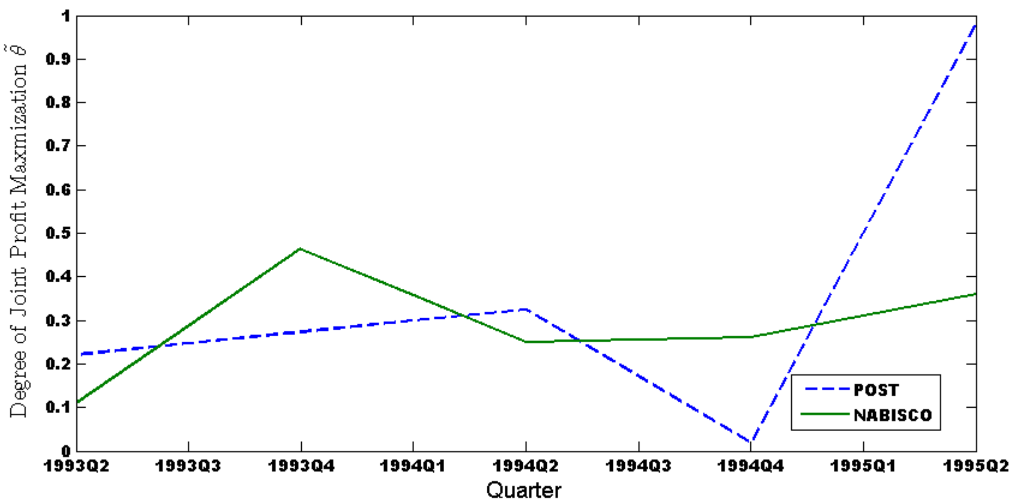
Note: Squares indicate quantity-weighted average deflated firm prices per store per quarter. Lines indicate average firm prices across all stores.

Figure 6: Degree of joint profit maximization $\tilde{\theta}$ over time



Note: Lines reflect degree of merging firms' joint profit maximization over time, $\tilde{\theta}$, for different degrees of symmetric industry competition, Θ. Degree of joint profit maximization is estimated separately for two-month intervals, and then harmonized between neighboring intervals.

Figure 7: Heterogeneous joint profit maximization under multi-product Nash pricing



Note: Lines reflect degree of merging firms' joint profit maximization over time, $\tilde{\theta}$, for multi-product Nash pricing, $\Theta = 0$. Degree of joint profit maximization is estimated separately for two-month intervals, and then harmonized between neighboring intervals.

Figure 8: Distribution of random price coefficient $\tilde{\alpha}$