Conduct Estimation via Ownership Change*

Preliminary and Incomplete Draft -
Please do not cite without permission

Christian Michel†

May 17, 2012

Abstract

This paper proposes a new form of estimating industry conduct. As an identification strategy, I use the structural ownership changes that occur due to a merger. Given both pre-merger and post-merger industry data, I look for the form of conduct that best predicts the market outcome before and after the merger. I provide identification results for a new form of direct industry conduct estimation. Using both pre- and post-merger data also enables me to provide a new evaluation criterion when selecting the form of competition among a discrete "menu" of outcomes. Finally, I estimate industry conduct using data around a merger from the Ready-to Eat (RTE) cereal industry.

Keywords: Conduct Estimation, Identification of Market Structure, Ex-post Merger Evaluation
JEL Classification: L11 (Production, Pricing, and Market Structure), L16 (Industrial Structure and Structural Change), C52 (Model Evaluation, Validation, and Selection)

1 Introduction

One of the core questions in industrial organization is whether prices are driven by costs or by a lack of competition. Modern empirical research in industrial organization often consists of a detailed industry demand model combined with a game-theoretic formalization of supply side behavior. Very often, the form of supply side conduct is however imposed by assumption rather than being tested for. This is potentially problematic: By using a wrong supply-side specification there is a risk of mispredicting the impacts of structural changes in an industry, such as the effects of a merger or the effect of product entry. The main problem is to jointly identify marginal costs and industry conduct.

Previous attempts to estimate both marginal costs and industry conduct have mostly been made using demand side variation. These approaches usually face two

*I would like to thank Philipp Schmidt-Dengler, Volker Nocke, Gautam Gowrisankaran, Andrew Sweeting, and Yuya Takahashi for their helpful comments. I also thank the participants of the University of Mannheim CDSE seminar and of the University of Mannheim IO group.
†Center for Doctoral Studies in Economics, University of Mannheim; Email: christian.felix.michel@gmail.com.

kind of problems. The first problem is the difficulty to find a sufficient number of demand rotators when estimating conduct in differentiated product markets. The second problem relates to the estimation techniques, which only estimate the economic parameters of interest accurately in special cases.

In this paper, I propose a new approach of estimating conduct in differentiated product markets via observing pre-merger and post-merger industry data. In short, I search for the form of supply side competition that best predicts both pre-merger and post-merger industry behavior. The intuition is the following. Given a well-specified demand-side estimation and valid instruments, one can consistently estimate consumer price elasticities irrespective of supply-side behavior. As is common in the literature, I assume that marginal costs in the market are not observed by the researcher. Different game-theoretic assumptions about the competition in the market will then yield different estimates for marginal costs. Given marginal costs estimates and the price-elasticities obtained from a 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) while holding all other things constant, I look for the form of competition that most accurately predicts the effects of an ownership change on prices using only the pre-merger data. As a performance measure, I can then use the observed post-merger prices.

Empirically, using advances in game theory, computational power, as well as econometrics have lead to structural models that are able to recover the empirical counterparts to the theoretical models. This has lead to a plethora of applications with respect to estimating demand or production functions. While estimating industry conduct played a big role in the beginning of the structural evolution, not much progress has been made over the past 15 years. Several reasons contribute to this. Historically, conduct estimation has been closely related to the concept of conjectural variations. Using a Cournot-style setting, one was able to derive equations from which conjectural variation parameters as a form of responsiveness towards industry competitors could be estimated, 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. A conjecture can be seen as a reduced-form game theoretic best response function in quantity setting symmetric games[1]. When accounting for higher-order rationality in the best response correspondences, one can however show that only Cournot competition itself survives iterated deletion of dominated strategies.

Corts (1999) critically discusses the identification of conjectural variation parameters. His critique is twofold. Firstly, he argues that conjectural variation parameters usually differ from the “as-if conduct parameters” one is interested in. This is because a conjectural variation parameter only estimates the marginal responsiveness

[1] In this case, conjectural variation models usually neglect any higher order effects of the best response functions.

of the marginal cost function with respect to changes in a demand shifter. As a researcher, in case of conjectural variation estimation 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, brands which belong to different firms set prices instead of quantities. Furthermore, instead of forming conjectures about other brands' reactions, each brand has an underlying objective function that potentially takes into account preferences for profits of other brands, thus allowing for cooperation between different brands of 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 then allows to estimate both marginal cost parameters and the level conduct parameters, which amount to the "as-if conduct parameters" in Corts (1999).

Corts second critique is related to the static game character of most conduct estimation models. My approach is not exempt from this critique. In this case, it may be that my static approach does not detect certain dynamic collusion patterns. Using a static approach however also has some advantages: Compared to dynamic models, one has a higher degree of tractability. Furthermore, repeated games make identification of conduct even more difficult. With my approach, I am also able to identify patterns of full collusion or partial collusion, i.e. collusion between only a subset of firms.

Menu approach My baseline approach reflects the case in which I estimate the conduct parameters and the marginal cost parameters directly using a moment estimator. The advantage of this approach is that the possible values of the conduct parameters are unconstrained, thus yielding a high degree of flexibility for the researcher. However, many parameter constellations have no real game theoretic foundation and thus it is not clear how one should interpret them. The menu approach on the other hand selects the best fit among a discrete set ("menu") of supply side models, such as for example multi-brand Bertrand-Nash competition or a single profit-maximizing monopolist. One advantage of this approach is that it does not include any conduct parameters directly, but rather pre-imposes them. Therefore one has less parameters to estimate which often also relaxes identification problems. This is similar to Nevo (1998). His menu approach however relies on non-nested tests that only use pre-merger data, such as Vuong (1989). These tests are based on a Kullbach-Leibler measure, and tests how far away a non-nested model is from the true data generating process pre-merger. My test is based on the fit pre-merger and

post-merger, and I exploit industry changes to account for the variation in the data as well as to have a clear structural break

Estimation algorithm I will now outline each step of the estimation algorithm in some detail.

1. **Estimate demand parameters consistently** Using proper instruments, it is possible to consistently estimate the demand parameters without having to specify supply-side competition. Section 2:1 discusses some prominent demand approaches in more detail.

2. **Choose set of potential supply-side specifications** When using the direct approach, one might have to specify some additional assumptions in order to reduce the parameter space. When using the menu approach, one has to predetermine the set of potential supply side models to choose from.

3. **Infer marginal costs and predict post-merger prices for given a supply specification** When estimating the demand side only using the ex-ante data, given proper instruments, one can already infer the marginal costs of production conditional on the form of conduct .

   Given the conduct matrix and the estimated demand parameters $_D$ from the demand side estimation in step 1, all parameters necessary to predict marginal costs are identified. This is because in order to estimate the price elasticities, one does not have to include cost parameters if other instruments are available. One can then obtain predicted pre- and post-merger prices.

4. **Compare predicted prices with actual prices** Compare how close the predicted pre- and post-merger prices are to the actual observed prices conditional on the form of conduct , using some distance metric.

5. **Start with step 3 for next supply-side specification**

   Note that when using the direct approach, it is also possible to estimate the "optimal" conduct matrix in a single estimation step without having to compare other specifications. With the menu approach, it is however evident that each potential model has to be estimated individually.

6. **Evaluate the different models** Overall, I am interested in the matrix among all potential supply sides defined in step 2 that minimizes the distance between the predicted and the observed prices both ex-post and ex-ante.

Kraft- Nabisco merger in 1992 In November 1992, Kraft Foods with its Post cereal line purchased the Nabisco ready to eat cereal branch. The antitrust court decision relied heavily on empirical analysis, see for example Rubinfeld (2000) for a detailed

analysis. I use data around the time of the merger to estimate the methods developed in the paper.

Related literature Bresnahan (1982) and Lau (1982) are the firsts to provide identification results for estimating conduct in the homogeneous good case. Lau finds that if the inverse demand function is not separable in the demand shifters, i.e: rotates rather than shifts due to demand shocks, then this is sufficient for an oligopoly solution to be identified. In a differentiated demand model, Bresnahan (1987) tests an hypothesis of a change in supply side competition in the US car market from collusive to competitive against other options. Genesove and Mullin (1998) compare predictions from a homogeneous good conduct estimation in the sugar industry with results from a direct cost estimation. They find that a model with a freely estimated conduct parameter yields more accurate cost estimates than estimates obtained from pre-specified models.

Feenstra and Levinsohn (1995) formalize an oligopoly setting which allows to estimate conduct in differentiated product industries for different non-nested oligopoly models. Nevo (2000) and Nevo (2001) use a random coefficient logit demand estimation in order to estimate marginal costs and market power in the RTE cereal industry, and to predict effects of a merger using only pre-merger data, respectively. Nevo (1998) theoretically 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 not having sufficiently many distinct demand shifters. Therefore, testing different non-nested models and choosing the best fit for the data often seems to be the only viable alternative. For such non-nested tests Vuong (1989) derives asymptotic results using a likelihood ratio test foundation. Shi (2011) develops a refined version of this approach. Gasmi, Laffont, and Vuong (1992) further provide likelihood ratio tests for different oligopoly specifications in the soft drink industry. In this paper, I use a supply driven conduct estimation instead of a demand driven one. Furthermore, this strategy also leads to a new way of testing between different models. Up until now there has not been much work in this field. One exception is Ciliberto and Williams (2010), who develop a conduct approach using multi-market contact in the the airline industry.

Further related questions in Industrial Organization This paper is also related to other open questions in the field of industrial organization. One of these is how well different demand-side models are capable of accurately predicting effects of mergers. Over the last decade, merger simulations have become an regularly-used tool in antitrust cases. However, different demand side models, such as logit, nested logit, or the Almost Ideal Demand Model will lead to different estimates for price-elasticities, and therefore also differences in the inferred marginal costs. Up until

now, there is only a small number of papers exploring these issues, see for example Weinberg and Hosken (2009) and Yoshimoto (2011).

Estimating profit internalization of merging firms Different economic theories would predict different degrees of post-merger internalization of profits. From a macroscopic viewpoint of the firm it makes sense to maximize joint profits of all brands. This is most important if a company has several brands that are relatively close substitutes for consumers. However, there might be delays in post-merger harmonization of firm strategies due to old contractual agreements and incentives structures. Furthermore, from a theory of the firm point of view, it is possible that the different branches remain different profit centers and thus also compete within the firm. With my framework, I am also able to test intra-firm conduct conditional on a specific form of industry competition.

Section 2 introduces the baseline model and discusses the conduct estimation strategy in detail. Section 3 will provide identification results in order to estimate conduct and marginal costs jointly. Section 4 presents the data and estimation results. Section 5 concludes with a discussion of both the results and some open questions.

2 The Model

2.1 Demand side

There is a total number of J brands in the market that is produced by $N \leq J$ firms. Each brand can only be produced by one firm. I assume that the overall revenue is proportional to the market size M. Denote $\mathbb{F}_f$ firm f's portfolio of brands, $f \in \{1, \ldots, N\}$. A firm f faces a fixed cost of production for each brand $j \in \mathbb{F}_f$, C_j, and marginal cost mc_j of producing one unit of brand $j \in \mathbb{F}_f$. Firm f's overall profit Π_f is thus the sum of revenues among all of its brands minus the marginal costs per brand and the firm's fixed costs.

In this section I will focus on two different random utility models, the Logit and the Random Coefficients Logit model, respectively. While the former model has an advantage of a relatively easy implementation, the latter model allows for more flexibility when estimating industry demand.

Logit Demand In the logit model, consumer i's utility of consuming product j in period t can be written as

$$u_{ijt} = x_{jt} \quad - \quad p_{jt} + {}_{jt} + {}_{ijt}; \tag{1}$$

where x_{jt} denotes firm j's observable brand characteristics, p_j denotes the price of product j, and ξ_{jt} denotes brand-specifc fixed effects unobservable to the researcher but observable to the firms, with $\xi_{jt} = \xi_t + \Delta\xi_{jt}$. ϵ_{ijt} is an idiosyncratic error term. Following the standard logit assumption, this error term follows an extreme value distribution, i.e: the cumulative density function has the form $F(\epsilon) = exp(-exp(-\epsilon))$.

Following Berry (1994), we can write the own- and cross price derivatives of product j's market share as

$$\frac{\partial s_j}{\partial p_i} = \begin{cases} -\alpha s_j(1-s_j) & \text{if } i = j; \\ \alpha s_j s_i & \text{if } i \neq j. \end{cases} \tag{2}$$

Thus, given a correctly estimated price-coefficient α, one will be able to estimate the own-price and cross-price elasticities for all products. This however has several unrealistic implications, such as independence of irrelevant alternatives (IIA), see for example Train (2003) for a detailed discussion of these problems.

Random Coefficient Logit This demand specification is closely related to Nevo (2001). Denote the number of individual consumers in every market by I, and denote the number of time periods by T. In the Random Coefficient Logit model, individual i's utility of consuming product j at time t can be written as

$$u_{ijt} = x_{jt}\tilde{\beta}_i + \tilde{\alpha}_i p_{jt} + \xi_{jt} + \epsilon_{ijt}; \tag{3}$$

i=1,..,I; j=1,..,J; t=1,..,T. Here, the coefficients $\tilde{\beta}$ and $\tilde{\alpha}$ are random coefficients, and depend on the demographics in each region, D_i, as well as on an unobserved vector of shocks, v_i. Denote the vector of demand side coefficients $(\tilde{\beta}; \tilde{\alpha})$ by θ^D. Then, the predicted population market share for product j can be written as

$$A_{jt}(\mathbf{x_t}; \mathbf{p_t}; \mathbf{xi_t}; \theta^D) = \{(D_i; v_i; \epsilon_{\mathbf{t}}) | u_{ijt} \geq u_{ilt} \forall l \in \{0; ...; J\}\};$$

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

$$s_j(\mathbf{x_t}; \mathbf{p_t}; \mathbf{xi_t}; \theta^D) = \int_{A_{jt}} dP^*_\epsilon(\epsilon) dP^*_v(v) dP^*_D(D). \tag{4}$$

2.2 Supply side

Cost function I use a log-linear form of the cost function, which is similar to Berry, Levinsohn, and Pakes (1995). Product j's marginal costs of production can be

written as

$$mc_j = f(z_j; \omega_j); \tag{5}$$

where z_j is a vector of observable cost characteristics of product j, ω_j is an unobservable brand-fixed effect.

It is common practice in empirical industrial IO models to regress marginal costs on input costs in order to estimate a form of production function and therefore also the cost drivers for each product. This is potentially beneficial when it comes to forecasting marginal costs using input price changes. Naturally, this will also partially hinge on the structure one assumes on the marginal costs. An econometrically convenient way is to assume that the costs are evolving according to a Cobb-Douglas production function:

$$mc_j = \Pi_i e^{z_i \gamma_i^S} e^{\omega_j}; \tag{6}$$

where ω_j is an unobserved brand-specific fixed effect. Rewriting the equation in logarithmic form, one obtains

$$log(mc_j) = \sum_i \gamma_i^S z_i + \omega_j \tag{7}$$

One convenient aspect of this formulation is that the brand-specific fixed effect enters this equation linearly and additively. A further option concerning the production function is to account for 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 production function, then this can be written as

$$mc_j = q_j^\tau \Pi_i e^{z_i \gamma_i^S} e^{\omega_j}; \tag{8}$$

where τ is the scale parameter. The estimable equation then yields

$$log(mc_j) = \tau \, log(q_j) + \sum_i \gamma_i^S z_i + \omega_j. \tag{9}$$

2.3 Industry conduct

The key of this paper is to test for the ownership restrictions rather than to preassume them. This means that I allow for a flexible ownership matrix. Each brand maximizes its objective function with to respect to its price. This can also include profits of other firms' brands. Among all firms, the marginal costs of all firms are common knowledge as well as how each brand takes into account all other brands. Both marginal costs and the conduct parameters are however not observed by the

econometrician. Brand $j's$ objective function can then be written as

$$\Pi_j = (p_j - mc_j)s_j + \sum_{r \neq j} \quad {}_{jr}(p_r - mc_r)s_r; \tag{10}$$

where s_r denotes the market share of brand r. This implicitly includes the assumption that $\quad_{ii} = 1 \forall i = 1; :::; J$, i.e. a brand fully cares about its own profits. Note that since only relative weights matter for the first order condition, this is a normalization without loss of generality. Consequently, I make the restrictions that $0 \leq \quad_{ij} \leq 1 \; \forall i \neq j$. Thus, I also assume that a firm does not derive a positive utility from "ruining" another firm. Furthermore, the profit of all other brands enters the objective function additively. The advantage of such a framework is a linear structure that is easier to estimate than other structures. In any case, the first order condition for brand j with respect to its own price can be written as

$$s_j(p) + \sum_{r=1}^{J} \quad {}_{jr}(p_r - mc_r)\frac{@s_r}{@p_j} = 0: \tag{11}$$

The ownership matrix Ω^* can now be defined as matrix that consists of entries

$$\Omega^*_{jr} = \quad {}_{jr}:$$

This leads to

$$\Omega^* = \quad = \begin{pmatrix} 1 & {}_{12} & :: & {}_{1J} \\ {}_{21} & 1 & :: & {}_{2J} \\ :: & :: & :: & :: \\ {}_{J1} & {}_{J2} & :: & 1 \end{pmatrix}$$

The key of my estimation technique is to find the form of conduct and marginal costs that best rationalize pre-merger and post-merger prices. Thus, when inferring marginal costs conditional on a specific form of conduct using pre-merger data, in a second step I am interested in recovering the matrix of conduct parameters that best predicts these price changes with respect to observed post-merger data. Define $\Omega_{jr} \equiv - \quad {}_{jr} * \frac{\partial s_r}{\partial p_j}$. When estimating the demand side only using the ex-ante data, given proper instruments, one can already infer the marginal costs of production conditional on the form of conduct .

$$\hat{mc}(\hat{\ }^D; \ ^S; \ ; \Omega^{pre}) = p^{pre} - (\Omega^{pre}(\hat{\ }^D; \ ))^{-1}s(p): \tag{12}$$

This also implies $\hat{p}^{pre} = \hat{mc}(\hat{\ }^D; \ ^S; \ ; \Omega^{pre}) + (\Omega^{pre}(\hat{\ }^D; \ ))^{-1}s(p)$ as the predicted pre-merger price vector.

Given the matrix and the parameters $\hat{\ }^D$ from the demand side estimation,

all parameters necessary to recover marginal costs and estimate equation (7)are identified. This is because in order to estimate the price elasticities, one does not have to include cost parameters if other instruments are available. Therefore, the predicted post-merger prices given $\hat{\ }^D$ for a specific   can be written as

$$\hat{p}^{post}(\hat{\ }^D;\ {}^S; \Omega^{pre}(\ ); \Omega^{post}(\ )) = \hat{mc}(\Omega^{pre}(\hat{\ }^D;\ );\ {}^S) - (\Omega^{post}(\hat{\ }^D;\ ))^{-1} s(p). \quad (13)$$

Overall, I am interested in the matrix   among all potential supply sides defined in step 2 that minimizes the distance between the predicted and the observed prices ex-post.

$$\min_{\theta,\gamma^S} ||\hat{p}^{post}(\hat{\ }^D;\ {}^S; \Omega^{pre}(\ ); \Omega^{post}(\ )) - p^{post}; \hat{p}^{pre}(\hat{\ }^D;\ {}^S; \Omega^{pre}(\ )) - p^{pre}||; \quad (14)$$

where $||x||$ denotes a distance metric given a vector x.   is either constrained to belong to a menu of different supply side specifications, or consists of directly estimable parameters. The next section will discuss the differences of these options in more detail, and presents some identification results.

3 Identi cation

The key identification question is: Under what circumstances are firms' marginal costs and industry conduct jointly identified? One important point with respect to marginal costs is whether a merger will lead to synergies for the merging firms. In the baseline case I implicitly assume that a merger will either not lead to synergies or involve a known synergy level in form of a certain percentage decrease in marginal costs. The latter case enables to account for synergies that are claimed by merging parties prior to a merger. This is a standard convention in merger simulation models and in many other modern demand models.

I will now present identification results for the direct approach

Identification of conduct The matrix of first-order conditions (12) has five different components: marginal costs, the ownership matrix, prices, market shares and the partial derivatives with respect to prices. All these components affect each equation non-linearly. One can clearly see that if there is no change in post-merger demand or supply, then the system of $2J$ equations will be singular. In the following, I will discuss situations under which the system is clearly identified.

3.1 Direct approach

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. In order to reduce the parameter space, I put certain restrictions on firm behavior. Firstly, each brand takes fully into account its own profits when making the overall pricing decision: $\ _{ii} = 1 \ \forall i \in \{1; :::; J\}$. Secondly, each brand always takes fully into account the profits from all other brands of the same firm. This leads to the following assumption: $\ _{ij} = 1 \ \forall i; j \in \mathbb{F}_f$. 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 might further reduce the parameter space which I will now discuss in detail.

Bilateral brand symmetry The first proposition relates to the case in which two brands of two distinct firms take into account each others' profits in the same way when making their pricing decisions. This reduces the number of parameters to be estimated by a factor $\frac{1}{2}$.

Proposition 1 (Bilateral brand symmetry). *Suppose that for two distinct firms $f; g$, $\ _{ij} = \ _{ji} \ \forall i \in \mathbb{F}_f; \forall \ j \in \mathbb{F}_g$. Then industry conduct and marginal costs are jointly identified only if $dim(\ ^S) + \frac{J(J-1)}{2} + N \leq 2J$.*

Proof. Equation 7 includes $dim(\ ^S)$ different cost function parameters, and J additional brand fixed effects. Because each brand fully takes into account its own profits, i.e. $\ _{ii} = 1 \forall i \in \{1; :::; J\}$, there are $J(J-1)$ remaining conduct parameters. Bilateral brand symmetry reduces the number of conduct parameters to $\frac{J(J-1)}{2}$. Because a brand fully takes into account the brands that belong to the same firm when making its decision, $J - N$ of these parameters are always equal to 1. This results in an overall number of $dim(\ ^S) + J + \frac{J(J-1)}{2} - J + N = dim(\ ^S) + \frac{J(J-1)}{2} + N$ parameters to estimate. There are J pre-merger and J post-merger equations, which leads to an overall of $2J$ equations. Necessary for the model to be identified is that there are at least as many equations as parameters to be estimated, or $dim(\ ^S) + \frac{J(J-1)}{2} + N \leq 2J$. This completes the proof. $\square$

Bilateral symmetry between firms One way to further 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. However, one also needs to specify how conduct will change between merging and non-merging firms after the merger has taken place. Assume for example that there are 3 firms, 1, 2, and 3. Firm 1 and 2 will merge. Before the merger, firm 3 could potentially have had a different relationship towards firm 1 than towards firm 2. For identification reasons, I make 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 $\ ^{pre}$ and $\ ^{post}$ denote pre- and*

post-merger conduct, 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. ${}^{post}_{ik} = {}^{pre}_{ik}$; ${}^{post}_{jk} = {}^{pre}_{jk}$ *(no change in conduct);*
b. ${}^{post}_{ik} = {}^{pre}_{jk}$; ${}^{post}_{jk} = {}^{pre}_{jk}$ *(target firm standard);*
c. Go to "acquiring firm" values: ${}^{post}_{ik} = {}^{pre}_{ik}$; ${}^{post}_{jk} = {}^{pre}_{ik}$ *(acquiring firm standard).*

A stricter specification assumes that for two firms, all brands will 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 2 (Bilateral symmetry between firms). *Suppose Assumption* 1 *holds, and that for distinct firms* $f; g$, ${}_{ij} = {}_{ik} = {}_{ji} = {}_{ki}$ $\forall i \in \mathbb{F}_f; \forall\, j; k \in \mathbb{F}_g$. *Then industry conduct and marginal costs are jointly identified only if* $dim({}^{S}) + \frac{N(N-1)}{2} \leq J$.

Proof. The demand parameters ^{D} can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are $2J$ estimable equations, one equation per brand pre-merger, and one equation per brand post-merger. Since each brand has one brand-specific fixed effect, this requires J parameters. Furthermore, estimation of marginal costs coefficients ^{S} requires to estimate $dim({}^{S})$ more parameters. 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 $2J$ equations with $J + dim({}^{S}) + \frac{N(N-1)}{2}$ parameters. The model is only identified if there are at least as many equations as parameters, i.e: if $dim({}^{S}) + \frac{N(N-1)}{2} \leq J$. This completes the proof. □

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.

Proposition 3 (Same responsiveness to all cross-firm brands). *Suppose Assumption* 1 *holds, and that for distinct firms* $f; g; h$, ${}_{ij} = {}_{ik}$ $\forall i \in \mathbb{F}_f; \forall\, j \in \mathbb{F}_g; \forall\, k \in \mathbb{F}_h$. *Then industry conduct and marginal costs are jointly identified only if* $dim({}^{S}) + N \leq J$.

Proof. The demand parameters ^{D} can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are $2J$ estimable equations, one equation per brand pre-merger, and one equation per brand post-merger. Since each brand has one brand-specific fixed effect, this requires J parameters. Furthermore, estimation of marginal costs coefficients ^{S} requires to estimate $dim({}^{S})$ more parameters. Because each firm has one conduct parameter for all firms, this leads to an overall

number of N conduct parameters. This leads to $2J$ equations with $J + dim(\ ^{S}) + N$ parameters. The model is only identified if there are at least as many equations as parameters, i.e: if $dim(\ ^{S}) + N \leq J$. This completes the proof. $\square$

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. Firstly, it is again no longer possible to detect partial collusion between a subset of firms in the industry. Secondly, 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.

Same responsiveness between all firms

Proposition 4 (Same responsiveness between all firms). *Suppose for distinct firms* $f;g;h$, $\ _{ij} = \ _{ji} = \ _{ik} = \ _{jk} = \ _{kj}$ $\forall i \in \mathbb{F}_f; \forall j \in \mathbb{F}_g; \forall k \in \mathbb{F}_h$. *Then industry conduct and marginal costs are jointly identified only if* $dim(\ ^{S}) \leq J - 1$.

Proof. The demand parameters $\ ^{D}$ can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are $2J$ estimable equations, one equation per brand pre-merger, and one equation per brand post-merger. Since each brand has one brand-specific fixed effect, this requires J parameters. Furthermore, estimation of marginal costs coefficients $\ ^{S}$ requires to estimate $dim(\ ^{S})$ more parameters. Because each firm has one conduct parameter for all firms, this leads only one conduct parameter. This leads to $2J$ equations with $J + dim(\ ^{S}) + 1$ parameters. The model is only identified if there are at least as many equations as parameters, i.e: if $dim(\ ^{S}) \leq J - 1$. This completes the proof. $\square$

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 is always identified. The big disadvantages are that very often this parameter will not have a clear economic interpretation and severely restricts the number of economic models. Furthermore, one will not be able to test for partial collusion in the market.

The direct approach requires more structure and a larger parameter space. This is because the free-floating conduct parameters require 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.

Multi-brand pricing imposes restrictions on intra-firm behavior.

One problem with removing this assumption is that identification of these parameters requires new additional assumptions. The only exception to this is when one makes the assumption that a merger between two firms does not change the interaction between the different brands. This is the next lemma.

Lemma 1 (No strict multi-brand pricing). *Suppose Assumption* 1 *does not hold. Then the only case in which one can identify industry conduct without further assumptions is if conduct between merging firms' brands does not change at all.*

The next subsection will focus on identification of the menu approach.

3.2 Menu approach

The menu approach selects among a subset of pre-specified models. A big advantage of the menu approach is that it severely reduces the parameters to be estimated. Furthermore, it allows the researcher to restrict its search to models that also have a clear economic interpretation, such as multi-brand competition or partial collusion between firms. However, there is the risk of forgetting a relevant economic model when relying on the non-nested approach. As an intuition, one can think of the potential combinations of tacit collusion. Most importantly, I will now present results for which selecting among different non-nested models is identified. In particular, it will be of interest to see when it is possible to also estimate the synergies that merging firms experience after a merger.

Proposition 5 (Nonnested approach with no or known synergies). *When choosing among a discrete set of non-nested models, industry conduct and brands' marginal costs are jointly identified only if* $dim(\ ^S) \leq J$.

Proof. The demand parameters $\ ^D$ can be estimated from equations 10 and 13, respectively. Regarding the supply side, there are $2J$ estimable equations, one equation per brand pre-merger, and one equation per brand post-merger. Since each brand has one brand-specific fixed effect, this requires J parameters. Furthermore, estimation of marginal costs coefficients $\ ^S$ requires to estimate $dim(\ ^S)$ more parameters. Because the non-nested approach does not require the estimation of any conduct paramter, this leads to $2J$ equations with $J + dim(\ ^S)$ parameters. The model is only identified if there are at least as many equations as parameters, i.e: if $dim(\ ^S) \leq J$. This completes the proof. □

3.3 Direct estimation of synergies

There are various ways in order to estimate synergies directly, and no method works best all the time. Recall the marginal cost formulation (7): $log(mc_j) = \ ^S z_j + !_j$: Note that the supply-side vector $\ ^S$ is equal among all brands, i.e: firms have an

identical production function with respect to input, and only differ in a brand-specific cost shock. With this restriction, one easy way to estimate synergies among firms would be to estimate the post-merger brand-specific fixed effects $!_j$ for the brands of merging firms. This would then amount to synergies for factors that are not accounted for in the production function, such as distribution network, or advertising. In addition, one can also account for direct impacts on the production function, either by estimating the $_{ij}$ directly post-merger, or by accounting for differences while using a dummy estimation. A different question is whether one should always estimate $!_j$ when trying to estimate synergies directly. One argument against omitting these fixed effects is that it might lead to biased results for the production function parameters due to an omission bias. Accounting for firm-specific fixed effects can be another way of estimating synergies directly. The advantage is that depending on the number of brands per firm, the researcher has to estimate fewer parameters than when estimating brand-fixed effects. However, for two merging firms, the number of firm-specific fixed effects reduces from two parameters pre-merger to one single parameter post-merger. Changes in the values of the firm-specific fixed effects are then hard to interpret. When including both firm and brand specific fixed effects, however, only estimating firm-specific synergies will result in accounting for effects that do not involve specific brand characteristics or the production function, which could again amount to distribution effects. Proposition 2 sums up the identification requirements for the different variants.

Proposition 6 (Estimation of unknown synergies). *Let J_M and N_M denote the number of brands and firms involved in a merger, respectively. When estimating the synergies resulting from a merger directly together with industry conduct and marginal costs, both under the direct approach and the menu approach, this will lead to following number of additional parameters to estimate:*
a. $J_M \leq J$ when estimating brand synergy effects
b. $2J_M$ when estimating brand and production function synergy effects;
c. N_M when estimating firm synergy effects.

This shows that it is also possible to estimate synergies with the direct approach. However, this requires even more parameters to be estimated. Using the menu approach, it will often be feasible to estimate synergies and test for industry conduct jointly.

3.4 Related issues

Identifying 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

Conduct specification	No or known synergies	brand synergies
Bilateral brand symmetry	$dim(\gamma^S) + \frac{J(J-1)}{2} + N \leq 2J$	$\frac{J(J-1)}{2} + N + J_M \leq 2J$
Bilateral symm btw. firms	$dim(\gamma^S) + \frac{N(N-1)}{2} + \leq J$	$dim(\gamma^S) + \frac{N(N-1)}{2} + J_M \leq J$
Same resp. to cross-firm br.	$dim(\gamma^S) + N + \leq J$	$dim(\gamma^S) + N + J_M \leq J$
Same resp. btw. all firms	$dim(\gamma^S) \leq J - 1$	$dim(\gamma^S) + J_M \leq J - 1$
Menu approach	$dim(\gamma^S) \leq J$	$dim(\gamma^S) + J_M \leq J$

Table 1: Identification conditions for different specifications

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. The menu approach is theoretically feasible under entry.

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 also 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 then result in an endogeneity problem when estimating conduct using product exit.

3.5 Theoretical examples

In this section I present further examples that highlight the effects of the assumptions made above. Again, 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. $_{ij}$ 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

$$\max_{p_1} (p_1 - mc_1)s_1(p) + \ _{12}(p_2 - mc_2)s_2(p) + \ _{13}(p_3 - mc_3)s_3(p) + \ _{14}(p_4 - mc_4)s_4(p)$$

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

$$(p_1 - mc_1)\frac{@s_1}{@p_1} + s_1 + \ _{12}(p_2 - mc_2)\frac{@s_2}{@p_1} + \ _{13}(p_3 - mc_3)\frac{@s_3}{@p_1} + \ _{14}(p_4 - mc_4)\frac{@s_4}{@p_1} = 0$$

The fully flexible conduct matrix for all firms can be written as

$$\Omega^* = \quad = \begin{pmatrix} {}_{11} & {}_{12} & {}_{13} & {}_{14} \\ {}_{21} & {}_{22} & {}_{23} & {}_{24} \\ {}_{31} & {}_{32} & {}_{33} & {}_{34} \\ {}_{41} & {}_{42} & {}_{43} & {}_{44} \end{pmatrix}$$

Assuming that each brand maximizes its own profits, the pre-merger conduct matrix yields

$$\Omega^* = \quad = \begin{pmatrix} 1 & {}_{12} & {}_{13} & {}_{14} \\ {}_{21} & 1 & {}_{23} & {}_{24} \\ {}_{31} & {}_{32} & 1 & {}_{34} \\ {}_{41} & {}_{42} & {}_{43} & 1 \end{pmatrix}$$

When making the additional assumption that a firm maximizes the profits of all of its brands, the pre-merger conduct matrix is

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {}_{13} & {}_{14} \\ 1 & 1 & {}_{23} & {}_{24} \\ {}_{31} & {}_{32} & 1 & {}_{34} \\ {}_{41} & {}_{42} & {}_{43} & 1 \end{pmatrix}$$

Using these assumptions on intra-firm maximization, one can already see a change in the ownership matrix pre- and post-merger if firms 2 and 3 merge. In such a case, the associated post-merger matrix yields

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {}_{13} & {}_{14} \\ 1 & 1 & {}_{23} & {}_{24} \\ {}_{31} & {}_{32} & 1 & 1 \\ {}_{41} & {}_{42} & 1 & 1 \end{pmatrix}$$

As one can already see 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, however, there are 12 parameters (8 conduct parameters, 4 marginal cost parameters) to estimate, with only 8 equations, such that identification is not feasible in this case. I will now briefly introduce different assumption on firm supply in order to reduce the number of parameters to be estimated.

Bilateral brand symmetry Under bilateral brand symmetry, for two distinct brands, each brand takes the other brands profits into account in the same way when making

its pricing decision. This leads to the following pre-merger conduct matrix:

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {\scriptstyle B} & {\scriptstyle D} \\ 1 & 1 & {\scriptstyle C} & {\scriptstyle E} \\ {\scriptstyle B} & {\scriptstyle C} & 1 & {\scriptstyle A} \\ {\scriptstyle D} & {\scriptstyle E} & {\scriptstyle A} & 1 \end{pmatrix}$$

The associated post-merger conduct matrix results in

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {\scriptstyle B} & {\scriptstyle D} \\ 1 & 1 & {\scriptstyle C} & {\scriptstyle E} \\ {\scriptstyle B} & {\scriptstyle C} & 1 & 1 \\ {\scriptstyle D} & {\scriptstyle E} & 1 & 1 \end{pmatrix}$$

Under the restriction, there are 9 parameters to estimate, with only 8 equations, such that this system is still not identified.

Bilateral symmetry between firms Instead of bilateral brand symmetry, as 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 conduct can then be written as

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {\scriptstyle a} & {\scriptstyle b} \\ 1 & 1 & {\scriptstyle a} & {\scriptstyle b} \\ {\scriptstyle a} & {\scriptstyle a} & 1 & {\scriptstyle c} \\ {\scriptstyle b} & {\scriptstyle b} & {\scriptstyle c} & 1 \end{pmatrix}$$

Post-merger conduct in this case is

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {\scriptstyle a} & {\scriptstyle b} \\ 1 & 1 & {\scriptstyle a} & {\scriptstyle b} \\ {\scriptstyle a} & {\scriptstyle a} & 1 & 1 \\ {\scriptstyle b} & {\scriptstyle b} & 1 & 1 \end{pmatrix}$$

This leads to a number of 7 parameters to estimate, with 8 available equations, such that the system is identified in absence of multi-collinearity.

Symmetry among all cross-firm brands I now assume that all brands take the brands of all other firms into account in the same way. This results in the following pre-

merger conduct:

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {}^{a} & {}^{a} \\ 1 & 1 & {}^{a} & {}^{a} \\ {}^{a} & {}^{a} & 1 & {}^{b} \\ {}^{a} & {}^{a} & {}^{b} & 1 \end{pmatrix}$$

.

Post-merger conduct is then

$$\Omega^* = \quad = \begin{pmatrix} 1 & 1 & {}^{a} & {}^{a} \\ 1 & 1 & {}^{a} & {}^{a} \\ {}^{a} & {}^{a} & 1 & 1 \\ {}^{a} & {}^{a} & 1 & 1 \end{pmatrix}$$

There are now 6 parameters and 8 equations. This potentially allows for estimation of synergies due to a merger directly, or by inclusion of other cost variables.

Differences of menu approach My identification strategy relies on $2J$ pricing equations. When using the menu approach, even with only 4 brands in the market, one has already some flexibility to estimate a more detailed cost function or to incorporate brand synergies for merging firms.

4 Data and Estimation

4.1 Estimation technique

I have two different estimation steps. In the first step, I estimate the demand elasticities using a discrete choice model. Using these elasticities, I then estimate marginal costs and industry conduct using a second GMM estimation routine.

4.1.1 Demand side

I use the technique of Nevo (2001) to recover the structural demand side parameters and unobservable error term . Using Nevo's estimation strategy on the demand side allows me to estimate all the structural demand side parameters independently of the supply side. This has major advantages when it comes to estimating industry conduct with respect to computational complexity. For the most flexible specification, I use a random coefficient Logit model estimated via a GMM routine.

Denote the vector of the mean utility level at time t by ${}_{\mathbf{t}}$. I solve for ${}_{\mathbf{t}}$ as to match the empirical market shares $s_{jt}(\mathbf{x_t}; \mathbf{p_t}; {}_{\mathbf{t}}; {}^{D})$ from equation (4) with the actual market shares s_{jt} observed in the data. Denote by $!({}^{D})$ an error term

depending on the demand side parameters, and denote by Z a matrix of demand side instruments. Then using a GMM estimator, the objective is to find

$$\hat{\gamma}^D = \arg\min_{\gamma^D} \omega(\gamma^D)' Z A^{-1} Z' \omega(\gamma^D); \tag{15}$$

where A^{-1} is an estimate of the asymptotically efficient covariance function $E[Z'\omega\omega'Z]$.

4.1.2 Supply side

Having obtained the demand side coefficients $\hat{\gamma}^D$, I estimate the conduct parameters θ and marginal cost parameters γ^S by matching the predicted post-merger prices $\hat{p}^{post}(\hat{\gamma}^D; \gamma^S; \theta)$ to the empirical post-merger prices p^{post}. Denote the error term of these moments by $G(\theta; \gamma^S)$. Then the second stage GMM objective can be written as

$$(\hat{\theta}; \hat{\gamma}^S) = \arg\min_{\theta, \gamma^S} G(\theta; \gamma^S)' \tilde{W} G(\theta; \gamma^S); \tag{16}$$

where $\tilde{W}(\tilde{\theta}; \tilde{\gamma}^S)$ is a positive definite, asymptotically efficient weighting matrix given first-stage estimates $\tilde{\theta}$ and $\tilde{\gamma}^S$.

4.2 Data

On November 12, 1992, Kraft Foods purchased 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.[2] I use weekly scanner data from January 1991 until December 1995 from the Dominick's Finer Food database, with an overall 21 brands from 5 different firms. The scanner data involves 68 stores from the Chicago Metropolitan area, see Fig. 1 for a geographic map of the stores. In particular, the dataset includes data with respect to product prices, quantities sold, data on promotions, as well as 1990 census data yielding demographic variables for the different store locations.

Furthermore, I use input price data from the Bloomberg Terminal database over the same time.

Exogeneity of merger When it comes to the cause of the merger, one has to adress whether the reason to merge is endogenous with respect to market structure. After the 1988 leveraged buyout of RJR Nabisco, the ownership group accumulated a relatively high pile of debt. There is a popular claim that company divestitures were used in order to reduce the overall debt level, and not in order to increase industry

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

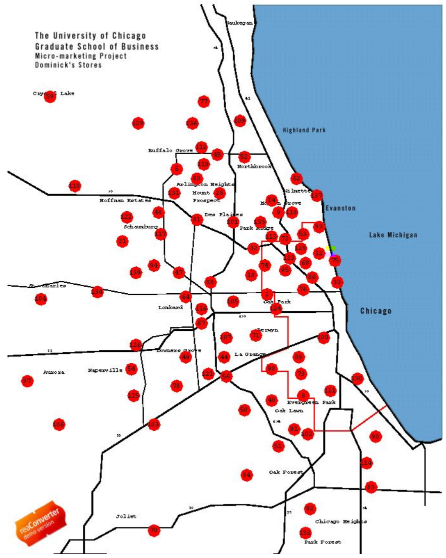

Figure 1: Geographical location of stores in dataset

profits. Even if this claim was not true, this would only bias the results if the merger had lead to unknown synergies, or if an anticipation of the merger by firms in the industry had lead to a change in behavior. Using my data, I can test for the former case. Concerning the latter case, such change in behavior pre-merger seems unlikely when looking at the data.

4.3 Instruments

There are several potential sources of endogeneity in the model. Firstly, from the demand side, prices may be correlated with unobserved product characteristics. Secondly, unobserved cost components may also be correlated with price. I will explain the use of my instruments.

4.3.1 Demand side instruments

Input prices To control for endogeneity of unobserved product characteristics, I use input prices as an instrument for sales prices in the demand estimation. Input costs variation should be correlated with with variation in prices, but not with consumers' preferences for unobservable product characteristics. Table 2 shows summary statistics for several input factors of the RTE cereal industry. One can see that for most factors, there are relatively high standard deviations. Table 3 shows the percentage price changes in input factors post-merger relative to the pre-merger prices. Input price changes in absolute values range from 4.5% for marketing expenses up to a 25% decrease in prices for gasoline. Table 4 shows aggregated price statistics pre- and post-merger. One can see that overall sales prices in the dataset are on average 11:1% higher in the post-merger sample than in the pre-merger sample.

	mean	sd	min	max
Corn	254.4	59.6	189.0	507.5
Wheat	354.0	77.4	250.5	634.5
Raw Sugar	10.4	1.7	7.8	14.9
Gasoline	63.7	10.7	41.9	108.2
Milk	128.0	4.1	121	136.5
CPI Cereals/ Bakery	155.6	9.1	141.6	173.6

Table 2: Summary statistics for input prices

,

	change (in %)
Transport costs	7.1
Milk	4.6
Wages food sector	8.0
Marketing	4.5
Cereal CPI	10.4
Milk CPI	4.5
Soybeans	9.4
Corn	12.4
Wheat	19.4
Sugar	13.1
Gasoline	-25.4

Table 3: Average percentage input price increases post-merger

Zone prices A second instrumentation possibility in this case is to use demand side instruments, see for example Hausman (1996) and Nevo (2001). In my dataset, stores are located in thirteen different clusters within the same metropolitan area. They rely on different zone pricing. Under the assumption that unobserved demand shocks are independent across the different clusters, prices of other zones are valid instruments. This seems to be a very strong assumption in reality, for example due to common advertising campaigns within the same metropolitan area.

4.3.2 Supply-side instruments

The brand specific marginal cost fixed effect $!_j$ 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 exchangability argument of product characteristics when facing a unique Nash equilibrium, see for example Berry et al. (1995).

4.4 Demand estimation

Table 5 shows demand side estimation results for several specifications of the logit model. Using both input prices and zone prices as instruments for the sales price

	mean	sd
p_pre	3.11	.55
p_post	3.48	.49
p_perc	.11	.06

Table 4: Price statistics pre- and post-merger

	IV Logit (costs) (1)	IV Logit (costs) (2)	IV Logit (zone prices) (3)	Logit (no IV) (4)
price (α)	-1.04***	-1.04***	-.99***	-.56***
advertising	.04**	.13**	.05**	.11***
sugar	.02***	.01***	.02***	.01***
fat	.19***	.06***	.15 ***	.13 ***
mushy	1.80***	.88***	2.06***	1.74 ***
income	.48***	.48***	.48 ***	.43 ***
household size	.08	.08	.08	.09
Brand FE	Yes	No	Yes	No
Cost IV	Yes	Yes	No	No
Zone IV	No	No	Yes	No

Table 5: Demand side estimates for Logit Model

Variable	Mean	Std. Dev.	Interaction income	Interaction std. income	Interaction household size	Interaction children
price	-16.47	.79	8.2	-.91	–	67.34
	(2.49)	(3.05)	(41.55)	(7.31)	–	(30.64)
const	-12.03	-.02422	-1.19	–	1.49	–
	(.21)	(.83)	(10.52)	–	(3.94)	–
sugar	2.16	-.01	-.34	–	-.03	–
	(.21)	(3.30)	(3.73)	–	(6.08)	–
mushy	-.01	.03	-.01	–	-.01	–
	(.10)	(.02)	(.08)	–	(.14)	–
MD weighted R^2	.20					
Numb. Obs.	21840					

Table 6: Demand side estimates for Random Coefficient Logit model

in specifications (1)- (3) yields a more elastic demand curve than specification (4) without instruments. Table 6 shows results for a random coefficients logit demand model. In this specification I include random coefficients for price, constant, sugar content and sogginess of cereal ("mushy"). Furthermore, I use demographic data on mean income, income standard deviation, household size and on number of small children to estimate these random coefficients. The results show a positive relationship between income and price, which is consistent with more aggressive pricing in high income neighborhoods. Price also interacts positively with the number of small children, which might account for their responsiveness to advertising. Furthermore, income interacts negatively with sugar, which might be attributed to more health-consciousness of higher income cohorts.

4.5 Conduct estimation (preliminary results)

There are several ways to predict the price changes caused by a merger. The main differences lie in how to account for post merger market shares. One option is to use only pre-merger data and then to simulate for each potential conduct assumption a new post-merger equilibrium consisting of simulated post-merger prices and simulated post-merger market shares. The main advantage of this approach is that it checks for consistency of prices and market shares in the post merger equilibrium. However, this approach neglects the available post-merger market shares and is computationally very intense.

	Firm A	Firm B	Firm C	Firm D	Firm E
Firm A	1	.39*	.70*	.49	.65
Firm B		1	.33	.27	.36
Firm C			1	.60	.40
Firm D				1	.68
Firm E					1
Median Price-Cost Margin	.28				

Table 7: Conduct estimates under bilateral firm symmetry

	Conduct Parameter
Firm A	.64***
Firm B	.72***
Firm C	.17
Firm D	.39**
Firm E	.39***
Median Price-Cost Margin	.35

Table 8: Conduct estimates under symmetry to all firms

As an alternative, I can also use the observed post-merger market shares directly as equilibrium market shares in my estimation. In this case I do not have to simulate

Inter-Firm Conduct	. 58**
Median Price-Cost Margin	.42

Table 9: Estimation of a single industry conduct parameter

for a new post-merger equilibrium, but can rather estimate the equilibrium parameters that most accurately predict post-merger prices. I will use this second approach for the remainder of the paper.

Table 7 shows the conduct estimation results under the assumption of bilateral brand symmetry, with an input price IV logit demand side being used. Using this approach, no parameter is significant on a 95% level. Table 8 shows the conduct estimation results under the assumption that each firm plays the same way to all of its rivals. Here we see that firms A and B play relatively cooperatively, while the other firms are more aggressive in the market. All but firm C's conduct parameter are significant at a 95% level. Table 9 indicates that if one only estimates a single conduct parameter for the industry, the estimation will yield a relatively cooperative market outcome with a relatively high median price-cost margin.

5 Discussion

In this paper I have proposed a new approach of estimating industry conduct. The difference of this paper to other price setting approaches lies in exploiting supply-side shifts using also post-merger industry data and thereby inferring the underlying industry conduct.

Another point worthy of discussion concerns a direct estimation of firm synergies. Using post-merger data and sufficiently many degrees of freedom, one can estimate firm synergies for the merging firms. In this context, one important question is how one accounts for economies of scale after the merger. It is unclear whether or how fast a merged entity can adjust its production line to reduce costs. This is especially important in energy-intense industries like the RTE cereal industry, and should also affect the firms' pricing decisions.

The underlying identification assumption that overall industry conduct does not change after the merger is arguably relatively strong. However, this assumption is also made in conventional merger simulation models. Furthermore, there is further supporting evidence for the validity of this assumption. Between 1993 and March 1996, prices for branded RTE cereal increased very moderately. In April 1996, Post started a price war and decreased the prices for all of its products by 20%, thereby permanently increasing its markets share. This was followed by significant price cuts two months later by General Mills and Kelloggs. Overall margin over production cost

fell by 12% in 1996 due to these actions.[3] These findings have several implications. Firstly, the results suggest that all of the main players priced significantly over marginal costs before April 1996. One explanation consistent with Post's price cuts would be stronger synergies of production after the merger that made a defection from a relatively cooperative equilibrium profitable. Secondly, the results seem to be fairly consistent with a change in equilibrium only occurring after April 1996. Thirdly, this can potentially be considered as another example for the differences between short and long term implications of merger policies.

[3]See for example Food Review (2000), Volume 23, Issue 2, pages 21-28.

References

Berry, S. (1994): "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 25(2), 242–262.

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

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

——— (1987): "Competition and collusion in the American automobile industry: The 1955 price war," *The Journal of Industrial Economics*, pp. 457–482.

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

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

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

Feenstra, R., and J. Levinsohn (1995): "Estimating markups and market conduct with multidimensional product attributes," *The Review of Economic Studies*, 62(1), 19.

Gasmi, F., J. Laffont, and Q. Vuong (1992): "Econometric Analysisof Collusive Behaviorin a Soft-Drink Market," *Journal of Economics & Management Strategy*, 1(2), 277–311.

Genesove, D., and W. Mullin (1998): "Testing static oligopoly models: conduct and cost in the sugar industry, 1890-1914," *The Rand Journal of Economics*, 29(2), 355–377.

Hausman, J. (1996): "Valuation of new goods under perfect and imperfect competition," *Chicago: National Bureau of Economic Research.*

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

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*, pp. 395–421.

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

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

Shi, X. (2011): "Size Distortion and Modification of Classical Vuong Tests," *Unpublished working paper, University of Wisconsin-Madison.*

Train, K. (2003): *Discrete choice methods with simulation.* Cambridge Univ Pr.

Vuong, Q. (1989): "Likelihood ratio tests for model selection and non-nested hypotheses," *Econometrica*, 57(2), 307–333.

Weinberg, M., and D. Hosken (2009): "Using Mergers to Test a Model of Oligopoly," *mimeo.*

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," .