In Search of the Truth: Challenges in Merger Simulation Analysis

Christopher R. Knittel and Konstantinos Metaxoglou*

January 3, 2011

Abstract

We illustrate the challenges in Antitrust analysis of merger effects using sophisticated aggregate demand systems, such as the random-coefficient logit of Berry, Levinsohn, and Pakes (1995). Based on data from the cereal and automobile industries, we document variation in hypothetical post-merger prices and its implication for consumer welfare and firm variable profits. We identify two sources of post-merger price variation. The first source is due to variation in demand estimates obtained by different starting values and non-linear search algorithms. The second source is due to variation in the solutions of the non-linear optimization problem associated with the Bertrand equilibrium. More importantly, the solutions to the first-order conditions do not meet the second-order conditions of the assumed oligopoly game. Overall, our results present a cautionary tale not only for the use of random-coefficient logit models in merger analysis, but also the use of the Bertrand first-order conditions in demand estimation for same class of models.

* *Knittel:* Department of Economics, UC Davis; Energy Institute @ Haas; and NBER. *Email: crknittel@ucdavis.edu. Metaxoglou:* Bates White LLC. *Email: konstantinos.metaxoglou@bateswhite.com.*

1 Introduction

We illustrate the challenges in simulating merger effects using sophisticated aggregate demand systems, such as the random-coefficient (RC) Logit of Berry et al. (1995)—BLP, henceforth. The first challenge is the solution of a highly non-linear minimization problem associated with demand estimation, as documented in Knittel and Metaxoglou (2008). Different starting values and non-linear search algorithms yield different demand estimates. These different demand estimates lead to economically significant variation in post-merger market performance. The second challenge is the solution of a complex non-linear optimization problem associated with the Bertrand-Nash equilibrium. Holding the set of demand parameters constant, post-merger market performance varies significantly across different starting points used to find the Bertrand equilibrium.

We quantify the effects of these two sources of variation in post-merger market performance using cereal data from Nevo (2000a) and automobile data from BLP. The hypothetical merger in the cereal industry involves Kellogg's and General Mills. The hypothetical merger in the automobile industry involves GM and Chrysler. The difference between post- and pre-merger cereal prices lies between -0.05 cents per serving to over 50 cents. The cereal post-merger consumer welfare and variable profits also exhibit significant variation: -\$180 million to \$7.6 billion, and \$75 million to \$5.8 billion, respectively. The difference between post- and pre-merger automobile prices is as low as \$28 and as high as \$200. Automobile consumer welfare and variable profits vary significantly: \$219M to \$1.9B, and \$5B to \$187B, respectively.

The variation in merger outcomes holding demand estimates constant, as the one presented here for the cereal industry, is due to nonlinearities in firms' best-response pricing strategies and corroborates, to some extent, recent theoretical work. For example, Allon et al. (2010) have shown that multiple equilibria may exist in sufficiently concentrated markets in the case of Bertrand competition among single-product firms. In addition, we provide an example, where solving the first-order conditions (FOCs) is not sufficient to obtain the equilibrium prices because the second-order conditions (SOCs) are not met. Hence, although the solution of the FOCs provides the Nash equilibrium for the simple Logit case because the SOCs are met (Morrow and Skerlos (2010)), this does not seem to necessarily be the case for the RC-Logit models.

Our finding that the SOCs are not met is problematic for two reasons. First, our read of the literature suggests that SOCs are not typically tested when using a RC-Logit demand system to calculate equilibria; at least, they are not mentioned or reported. Second, researchers have used the Bertrand FOCs as moment restrictions when estimating the demand parameters in a GMM framework under the assumption of profit maximization. If when estimating the demand parameters, the FOCs moments are met, but the SOCs are not, the GMM objective function will be misspecification leading to inconsistent demand estimates.

We believe our work is of immediate interest to Antitrust agencies that routinely review a large number of mergers and frequently employ simulation techniques to assess their potentially anticompetitive effects. However, before providing an overview of the simulating techniques used in antitrust analysis, as well as the details of the issues we have identified in this paper, it is useful for the reader to understand the realities of the merger review process: large number of filings and tight time constraints. We discuss the case of the U.S. because we are the most familiar with.

Since the introduction of the Hart-Scott-Rodino (HSR) Antitrust Improvements Act of 1976, all U.S. mergers valued at more than a threshold are required to file with the Federal Trade Commission (FTC) and the Antitrust Division of the Department of Justice (DOJ) – collectively, the Agencies. Between 1991 and 2009, approximately 45,000 mergers were filed with the Agencies. After various adjustments throughout the years, the filing threshold was set at $65.2 million in fiscal year 2009 (FTC and DOJ (2010a)).

Following an HSR filing, the Agencies have 30 days to conduct their preliminary review and decide whether or not to issue a "second request" for a more thorough investigation of a potentially anticompetitive merger. Once a second request is issued, the merging parties have usually 2 to 3 months to comply. After the parties have complied, the Agencies decide whether to block the transaction, accept some type of remedy, or allow the merger to proceed within 30 days. Almost 97% of the HSR filings were allowed to proceed without a full investigation in the period 1991-2009. The remaining 3% were subject to a second request with two thirds of them being abandoned, blocked, or modified to address the Agencies' concerns.

The Horizontal Merger Guidelines provide the Agencies' analytical framework to determine whether the proposed merger is likely to be anticompetitive (FTC and DOJ (2010b). The to-do list of the Agencies' staff in the course of the investigation, usually, includes product and geographic market definitions, as well as evaluation of (unilateral and coordinated) price-effects theories, entry conditions, and efficiency claims under very tight time constraints. The focus of unilateral-effects theories is the change in the merged firm's incentives to price its products following the merger. In the textbook case, the merged firm will have an incentive to raise prices above the pre-merger levels because it now internalizes some of the substitution due to price increases. Concerns about coordinated effects are raised if the proposed merger increases the likelihood of collusion among competitors.

The simulation of market structures to predict the price effects of mergers has experienced a substantial growth in popularity since the early 1990s. This growth is largely attributed to developments in the industrial organization and the antitrust literatures in two areas: demand calibration/estimation and game-theoretic models of competition Werden and Froeb (2006), Budzinski and Ruhmer (2010), as well as Davis and Garces (2009), provide an in-depth discussion regarding the use of demand and game-theoretic models for antitrust analysis.

Demand calibration, a less demanding exercise compared to estimation, offers a particularly attractive analytical framework for the short-lived merger-review process. The most prominent examples of calibrated demand models are the Antitrust Logit Model (e.g., Werden and Froeb (2002)), and the Proportionately-Calibrated Almost Ideal Demand System (PC-AIDS) of Epstein and Rubinfeld (2001).

Demand models that require more rigorous econometric exercises fall into two broad categories: continuous and discrete. Continuous models include the linear, log-linear, and constant-elasticity models (Werden (1997)), as well as the AIDS models (Hausman et al. (1994)). These models examine the relationship between prices and quantities assuming functional forms and are, usually, suitable for analyzing markets with a small number of products.

Discrete choice models, where demand is derived from utility, other than the RC Logit (Berry et al. (1995), Nevo (2000a)) discussed here, include the simple Logit, the nested Logit, and the GEV models. Weinberg and Hosken (2008), Ivaldi and Verboven (2005), and Peters (2006) illustrate the use of these discrete-choice models in merger simulations. They are particularly suited for markets with many products and increase in their complexity as we move from the simple Logit to the GEV.[1] Of course, the tight time framework of the merger review process very often imposes a reality check on the degree of complexity of the demand model considered for the analysis.

As Nevo (2001) illustrates, a demand system combined with an oligopoly model of competition and some additional assumptions suffice to simulate merger effects. Abstracting from technical details, the model of competition offers the FOCs that equate marginal revenue to marginal costs. The demand system identifies the marginal revenue, yielding pre-merger marginal costs through the FOCs. Assuming that a smaller number of competitors is the only

[1] Pinkse and Slade (2004), Chan (2006), and Davis and Ribeiro (2010) borrow ingredients of both continuous and discrete choice demand models.

difference between the pre- and post-merger industry structures, which can be easily handled with an adjustment to the ownership matrix entering the same FOCs, post-merger prices are solutions to a system of non-linear equations. Our work lies in the heart of the challenges of a simulation exercise of this sort.[2]

The remaining of this paper is organized as follows. In Section 2, we discuss a typical RC-Logit demand model and try to shed light in the challenges associated with the underlying optimization problem. Issues associated with the calculation of post-merger prices are discussed in Section 3. We present our findings using the cereal and automobile data in Section 4. We finally conclude.

2 The Demand Model

This section describes the first step of the merger simulation exercise, namely the estimation of the underlying demand system. We start our discussion by presenting the standard BLP-type model of aggregate demand.[3] Following standard notation in the literature, we assume that a consumer i derives utility from a product j in market t that may be written as:

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt} = V_{ijt} + \varepsilon_{ijt}, \tag{1}$$

where p_{jt} is the product's price, x_{jt} is a vector of non-price product characteristics. The vector ξ_{jt} includes the unobserved to the econometrician features of the product, such as after-sale services and image, which are valued by the consumers. Each individual is assumed to choose one of the $1, \ldots, J$ products available in the market, or to no purchase at all. The no-purchase option is usually termed the outside good and its associated utility is $u_{i0t} = \varepsilon_{i0t}$. The Logit error term ε_{ijt} is the first source of consumer heterogeneity in the utility function. The second source of consumer heterogeneity are the random coefficients α_i and β_i, which may be written as follows:

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \Pi D_i + \Sigma v_i, \;\; D_i \backsim P_D\left(D\right), v_i \backsim P_v\left(v\right). \tag{2}$$

The decomposition in (2) leads to terms that are common across consumers, such as α and β, as well as to terms D_i and v_i, which are vectors of observed and unobserved consumer char-

[2]For a different angle on the evaluation of simulation techniques, see Bass et al. (2008), Weinberg and Hosken (2008), Peters (2006), Hausman and Leonard (2005).

[3]There is an extensive literature that uses purchases of a particular product at the customer level that employs random-coefficient logit models in the tradition of Train (2009).

acteristics that affect purchasing decisions and follow the distributions P_D and P_v respectively. The coefficient matrices for D_i and v_i are given by Π and Σ. The main difference between D_i and v_i is that we know something more about P_D compared to P_v. The observed characteristics in D_i may be demographics, such as income, education, and family size. For example, in the case of income, we may use data from Census surveys to either characterize its distribution or use them while constructing D_i *per se.* The v_is capture intrinsic heterogeneity that is not explained by any systematic or observable customer attributes. Often, standard normality is the usual assumption for P_v with Σ containing the associated second moments.

Some preliminary, more descriptive type of analysis in the form of hedonic regressions, is sometimes employed by researchers to identify product characteristics that "matter" and, hence, should enter the utility function. Furthermore, the decomposition of consumer heterogeneity, other than the Logit error term, into observed and unobserved is common.

The assignment of a random coefficient to price is probably the most common among practitioners, due to its desirable property to generate more realistic substitution patterns among products compared to more simple discrete-choice models, such as the Logit. Although attaching random coefficients to the remaining product characteristics is usually justifiable, it increases computational complexity. As a result, it is not uncommon for the matrices Π and Σ to be sparse. Using a more compact notation, and after combining equations (1) and (2), we may use the following expression for the utility:

$$\begin{aligned} u_{ijt} &= \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, D_i, v_i, \theta_2) + \varepsilon_{ijt} \\ \delta_{jt} &= x_{jt}\beta - \alpha p_{jt} + \xi_{jt}, \;\; \mu_{ijt} = [p_{jt}, x_{jt}]'(\Pi D_i + \Sigma v_i) \end{aligned} \tag{3}$$

The δ_{jt}s capture the mean utility associated with the consumption of good j that is common across consumers in market t. Deviations from this mean utility are reflected in μ_{ijt} and ε_{ijt}. The vectors θ_1 and θ_2 differ in that the former contains α and β, while the latter contains the elements of matrices Π and Σ. Under independence of consumer idiosyncrasies for characteristics, the market share of product j is given by:

$$s_{jt}\left(x, p_{\cdot t}, \delta_{\cdot t}; \theta_2\right) = \int_{A_{jt}} dP\left(D, v, \varepsilon\right), = \int_{A_{jt}} dP_{\varepsilon}\left(\varepsilon\right) dP_v\left(v\right) dP_D\left(D\right), \tag{4}$$

$$\text{with } A_{jt}\left(x, p_{\cdot t}, \delta_{\cdot t}; \theta_2\right) = \left\{(D_i, v_i, \varepsilon_{it}) | u_{ijt} \geq u_{ilt}, \forall l = 1, \ldots J\right\}.$$

In the share equation, x, includes the characteristics of the products while $p_{\cdot t} = (p_{1t}, \ldots p_{Jt})'$

and $\delta_{\cdot t} = (\delta_{1t}, \ldots \delta_{Jt})'$. The error term ε can be integrated out analytically in (4) giving rise to the well-known Logit probabilities. Given distributional assumptions for v and D, the integral associated with market shares is commonly evaluated using Monte-Carlo simulation assuming a number ns of individuals:

$$s_{jt}\left(x_{jt}, \delta_{jt}, \theta_2\right) = \frac{1}{ns}\sum_{i=1}^{ns} s_{ijt} = \frac{1}{ns}\sum_{i=1}^{ns} \frac{\exp(\delta_{jt} + \mu_{ijt})}{\sum_{j=1}^{J}\exp(\delta_{jt} + \mu_{ijt})} \tag{5}$$

The aggregate demand shock ξ_{jt}, introduced by Berry (1994), plays the role of the conventional linear-demand shock. In its absence, the market shares given by (5) are deterministic functions of the product characteristics and price. The Logit error term has been washed out and it cannot serve as a source of econometric uncertainty as it does in models employing consumer-level data.

The presence of ξ implies prices are endogenous because both consumers and firm observe ξ and therefore its value enters into the firms' pricing decisions. The standard approach in the literature to address the endogeneity is nonlinear GMM with the identifying assumption $E[\xi_{jt}|x_{jt}, z_{jt}] = 0$ given an appropriate vector of excluded instruments z_{jt}. Excluded instruments may include costs, as well as functions of the observed product characteristics or prices of the same product in different markets. Given a vector of mean utilities δ, a sample analog of the moment condition can be constructed and the researcher may proceed with estimation in the way described below.

The vector of mean utilities δ is retrieved by equating the observed market shares from the data with those implied by the model for a given vector of parameters θ_2:

$$s_{\cdot t}^{obs} = s_{jt}^{pred}\left(x, p_{\cdot t}, \delta_{\cdot t}; \theta_2\right). \tag{6}$$

As opposed to the simple Logit and nested Logit, where analytical solutions for δ are available for the system of equations in (6), the random-coefficient Logit requires a numerical solution of a highly nonlinear system of equations whose dimension equals the number of products in the market. Berry (1994) advocates a contraction mapping to retrieve δ whose kth iteration is given by:

$$\delta_{\cdot t}^{(k+1)} = \delta_{\cdot t}^{(k)} + \ln s_{\cdot t}^{obs} - \ln s_{\cdot t}^{pred}(x, p_t, \delta_{\cdot t}^{(k)}, \theta_2). \tag{7}$$

For a given value of θ_2, the contraction mapping in (7) can be initiated with the Logit solution $\delta_{\cdot t}^{(0)} = \ln(s_{\cdot t}) - \ln(s_{0t})$, where s_{0t} is the share of the outside good and continues until some norm of the difference between two consecutive iterates is smaller than some pre-specified tolerance. The contraction mapping has been the most widely used method to retrieve the vectors of mean utilities. Once δ is retrieved, ξ can be inferred from the following equation:

$$\xi_{jt} = \delta_{jt} - x_{jt}\beta - ap_{jt}. \tag{8}$$

The elements of θ_1, namely α and β, in the last equation are retrieved using linear instrumental variables (IVs). Having defined $\theta = (\theta_1, \theta_2)$, with the aggregate demand shock playing the role of a structural error term that is a function of θ, the econometrician faces the following nonlinear-GMM problem:

$$\hat{\theta} = \arg\min_{\theta} \ \xi(\theta)' Z\Phi^{-1}Z'\xi(\theta), \tag{9}$$

where Φ is the covariance matrix of the moment condition in the case of optimal GMM, or some other weighting matrix. Inference is performed using standard results from GMM theory (Pakes and Pollard (1989). The methodology just described allows the econometrician to perform a non-linear search in the parameter space only for θ_2 by concentrating out θ_1. This is feasible because, for a given value of θ_2, we infer δ using (6) and (7) and given δ we obtain θ_1 using linear IVs. Having δ and θ_1 available, the researcher constructs the econometric error that appears in (9). Draws from P_v and P_D required in (5) are made once and are kept constant through the estimation exercise.[4]

Following the publication of computer code by Nevo (2000b), the estimation of BLP-type models has become increasingly popular. Recent studies have identified issues regarding computational aspects of the methodology outlined here. As a background, we should keep in mind that the objective function of a typical BLP-model has not been shown to be globally concave (e.g., Bajari et al. (2007)). Additionally, a non-linear search in the parameter space requires hundreds or even thousands of function evaluations, with each of them involving a call of the contraction mapping.

Knittel and Metaxoglou (2008) illustrate that the underlying GMM problem is a non-trivial one based on data from Nevo (2000a) and Berry et al. (1995). Using more than 10 optimization algorithms from different classes (derivative based, direct searches, random searches) and 50

[4]Recent papers discuss alternative estimation methods for the class of demand models discussed here. Kalouptsidi (2010) uses the MPEC approach for the special case of BLP-type models with a finite number of consumers. Conlon (2010) combines the MPEC approach and Empirical Likelihood on the grounds of computational and statistical efficiency gains, respectively. Petrin and Train (2010) advocate the use of a two-step control-function approach to address price endogeneity. Jiang et al. (2009) implement Maximum Likelihood as opposed to GMM using a Bayesian MCMC method assuming a normal distribution for the vector of demand shocks ξ. Also for the special case of a finite number of consumers, Bajari et al. (2007) offer a non-parametric series estimator via the means of an inequality constrained least-squares approach.

starting values, they show that economic variables of interest such as elasticities and consumer welfare, which are readily available upon estimation, vary widely depending on the choice of optimization routine and starting value.

Dube et al. (2009) show that the temptation to implement loose stopping criteria for the contraction mapping to speed up the estimation process may cause two types of errors in parameter estimates. First, the approximation error of the inner contraction mapping propagates into the outer GMM objective function and its derivatives. Second, even when an optimization run converges, it may falsely stop at a point that is not a local minimum. The authors offer an alternative formulation of the GMM problem as a Mathematical Program with Equilibrium Constraints (MPEC) building on earlier work of Su and Judd (2008). In a nutshell, the unconstrained minimization problem with the contraction mapping is replaced with a constrained minimization problem, with a system of nonlinear constraints requiring the model's predicted market shares to be equal to the observed market shares and nonlinear search over a parameter space that is higher than implied by θ_2.

3 The Bertrand Game

This section describes the second step of the merger simulation exercise assuming a Bertrand oligopoly model. Having retrieved marginal costs using the demand estimates and the first-order conditions of a Bertrand game, we solve for the post-merger prices via simulation. The estimation of the demand model, as well as the merger simulations are often highly nonlinear problems. Knittel and Metaxoglou (2008) contain the details of our methodology to address the highly nonlinear nature of the optimization problem associated with the demand estimation. The details of our approach in the case of merger simulation are provided below.

With demand estimates in hand, we infer marginal costs using the first-order conditions of a static Bertrand model with multi-product firms:

$$p - mc = \Omega\left(p\right)^{-1} s\left(p\right), \tag{10}$$

where p is the price vector, $s(\cdot)$ is the vector of market shares, and mc denotes the corresponding marginal costs. The dimension of these vectors is equal to the number of the products available in the market, say J. The Ω matrix is the Hadamard product of the (transpose) of the matrix of the share price derivatives and an ownership structure matrix. The ownership structure matrix is of dimension $J \times J$, with its (i, j) element equal to 1 if products i and j are produced by the same firm and zero, otherwise. Because prices are observed and demand estimation allows us to retrieve the elements of Ω, estimates of marginal costs, $\widehat{mc}$, are directly obtained using (10).

A simple change of ones and zeros in the ownership structure matrix along with a series of additional assumptions (Nevo (2001)) allows the simulation of a change in the industry's structure, as the one implied by mergers among competitors. Simply put, a merger simulation implies the same Bertrand equilibrium with a smaller number of firms. In what follows, we analyze the range of values for a measure of consumer welfare on the basis of post-merger equilibrium prices. For the automobile data, we assume GM and Chrysler merge. In the case of the cereal data set, we assume Kellogg's and General Mills merge. The vector of post-merger prices p^{post} is the solution to the following system of nonlinear equations:

$$p^{post} = \widehat{mc} + \hat{\Omega}\left(p^{post}\right)^{-1}\hat{s}\left(p^{post}\right). \tag{11}$$

The elements of $\hat{\Omega}\left(p^{post}\right)$ reflect changes in the ownership structure implied by the hypothetical merger. Solving for the post-merger prices is equivalent to solving a system of nonlinear equations of dimension J in the market under consideration. For example, using the cereal data, we have 94 markets with 24 products in each market. As a result, solving (11) requires the solution of 94 systems of nonlinear equations of dimension 24. An approximate solution for the post-merger prices, which avoids the need to solve the systems of nonlinear equations, is given by:

$$p^{approx} = \widehat{mc} + \hat{\Omega}\left(p^{pre}\right)^{-1}\hat{s}(p^{pre}), \tag{12}$$

where $\hat{s}(p^{pre})$ is the pre-merger vector of market shares, readily available from the data, and the elements of $\hat{\Omega}$ associated with share price derivatives are evaluated at the pre-merger prices. In the results discussed below, we solved the system of nonlinear equations in (11) using a dogleg trust-region (DTR) version of Newton's method.[5]

Newton's method for the solution of nonlinear equations of the form $r(x) = 0$ may be described as iterating on the equation $x_{k+1} = x_k - \Delta(x_k)^{-1}r(x_k)$. In terms of notation, x_k denotes the kth iterate and $\Delta(x_k)$ is the Jacobian of $r(x_k)$. Newton's method can be made more robust using trust-region techniques with "dogleg" being a special case. Trust-region methods utilize the notion of a merit function, a scalar-valued function of x, which indicates whether a new candidate iterate is better or worse than the current iterate, in the sense of making progress toward the root of r. A widely used merit function is the sum of squares of the form $||r(x)||^2/2$.

[5]Nocedal and Wright (1999) provide an excellent discussion of the method in their Chapter 11.

We implement the DTR method using the MATLAB *fsolve* function. The termination tolerances for both the merit-function value and the vector of prices we are solving for are set equal to 1e-16. We also impose a maximum number of 1000 iterations. Additionally, the Jacobian is approximated using finite differences.[6] The remaining of the *fsolve* settings are equal to their default values. Finally, we use three sets of starting values for the DTR, namely, the pre-merger prices, the estimated marginal costs, and the vector of approximate solutions for the post-merger prices, p^{approx}. The use of multiple starting values for the DTR tries to address, to some extent, the issues identified in the next paragraph.

In the case of Bertrand competition with multiproduct firms facing random-coefficient Logit demands, there is no result that shows the following: (1) existence of an equilibrium in pure strategies, (2) whether the equilibria are unique solutions to the systems of FOCs of the underlying game. In the papers we are aware of, both existence and uniqueness have been assumed (e.g., footnote 12 in Berry et al. (1995)).

Allon et al. (2010) provide a sufficient condition under which a Bertrand equilibrium exists and the set of Bertrand equilibria coincides with the solutions of FOCs in the case of *single-product* firms facing random-coefficient Logit demands. This condition precludes a very high degree of market concentration: no firm captures more than 50% of the potential market in any of the consumer segments that it serves. A somewhat stronger version of the same condition, namely, firms shares below 30%, establishes uniqueness. Allon et al. (2010) also provide a sufficient condition for a (unique) equilibrium for markets with an arbitrary degree of concentration in the presence of an exogenous price limit. However, in this case, the equilibrium may not necessarily reside in the interior of the feasible price region and, hence, not be characterized by the FOCs.

The goal of our simulation exercises is to investigate whether the FOCs are met at multiple places for two data sets that have served as example data sets in the literature. Our empirical exercise is not as extensive as in Knittel and Metaxoglou (2008) in the sense that conditional on a set of demand parameters, we use three starting values to search for a Bertrand-Nash equilibrium. Using only these three starting values, we find that the FOCs are indeed met at different prices. We therefore find variation in these merger counterfactuals arising from both variation in the demand estimates and solutions to the non-linear Bertrand-Nash FOCs.

[6] A more detailed description of fsolve is available at http://www.mathworks.com/help/toolbox/optim/ug/fsolve.html. Alternative approaches to infer equilibria assuming a mode of competition are available in some recent work. For example, Miller and Osborne (2010), who examine spatial pricing the cement industry, use of the *fsane* function in R, which implements the non-linear equation solver of Cruz et al. (2006). Morrow and Skerlos (2010) study the automobile industry and present a fixed-point-iteration algorithm as an alternative to Newton's method.

4 Results

Tables 1 through 3 summarize our results of the hypothetical merger using the cereal data. We report post-merger industry profits, the total compensating variation, and the average change in prices.[7] The price changes are weighted by pre-merger market shares. The tables report results based on 13 algorithms used to obtain the demand estimates and 3 sets of starting price vectors utilized in the solution of the Bertrand FOCs. We used 50 starting values for each of the 13 algorithms in the case of demand estimation. We performed the simulation using the demand estimates that gave rise to the lowest GMM objective function value. The three sets of starting price vectors are the pre-merger prices, marginal costs, and approximate post-merger prices. We report results for the market with the smallest share of the outside good.[8]

The entries in each of these three tables exhibit substantial variation both across columns and across rows. The within-column variation is due to different demand estimates obtained for each of the 13 algorithms we employed. The within-row variation is due to different starting price vectors for the solution of the Bertrand FOCs. Using the merger approximation prices as starting values for the FOC search, across algorithms post-merger variable profits range from \$75 million to \$3.8 billion—two order of magnitudes. Using marginal costs as the starting value, post-merger variable profits range from \$456 thousand to \$4.0 billion—a factor of over 8,000. Finally, using pre-merger prices, post-merger variable profits range from \$75 million to \$5.8 billion—a factor of over 70. We note that the demand parameters from SOLVOPT and KNITRO3 yield the lowest GMM objective value found.

Within rows, variable profits vary for the majority of algorithms. The variation often exceeds an order of magnitude. Two points are worth noting. First, the demand parameters that lead to the lowest GMM objective value (SOLVOPT and KNITRO3) show the most variation across FOC starting values. Second, average post-merger variable profits are highest when using the pre-merger prices as starting values. Below, we provide evidence as to why this is the case. To conserve space, we do not go through, in detail, the results with respect to compensating variation and prices, but note that they exhibit similar variation to post-merger variable profits.

A preliminary answer to the more fundamental question of whether the solutions to the

[7]Following McFadden (1974) and Small and Rosen (1981), the compensating variation for individual i in the presence of linear income effects is $CV_i = (1/a_i)\left(\ln\left[\sum_{j=0}^{j=J}\exp\left(V_{ij}^{post}\right)\right] - \ln\left[\sum_{j=0}^{j=J}\exp\left(V_{ij}^{pre}\right)\right]\right)$, We calculate V_{ij}^{pre} and V_{ij}^{post} using the pre- and post-merger prices.

[8]We assume a market size of 250 million as if the single market analyzed is representative of the nation. Although this assumption is somewhat arbitrary, it does not affect the variation in results discussed here. We calculate the total compensating as market size times the average compensating variation. The average compensating variation is given by $\overline{CV} = (1/ns)\sum_{i=1}^{n} CV_i$.

FOCs represent Bertrand-Nash equilibria appears to be no. While we are in the process of calculating the profit hessian for each firm in each of the 94 markets for all three sets of starting points for the FOC solutions in the cereal data, we have checked the hessian for one firm and market when the approximate post-merger prices are used to obtain the FOC solutions. Some of the profit hessian eigenvalues have positive signs indicating that the SOCs are not satisfied.

Additional evidence for the FOC solutions failing to satisfy the SOCs is provided in Figures 1 through 3. Each of these histograms is constructed using the percentage change in market share implied by the 13 demand-estimation algorithms for each of the 24 products in the cereal data.[9] As the spikes at -100 illustrate, for a large fraction of the products the post-merger prices are so high that they imply close-to-zero market shares. This spike pattern is particularly prominent in the case of the simulations that used approximate post-merger prices and marginal costs for the solution of the FOCs.

Figure 4 illustrates the failure of SOCs in an alternative way. Here, we profile the variable profits of the firm whose hessian indicated violation of SOCs around the post-merger price for one of its products with close-to-zero post-merger shares, while keeping the post-merger prices of the remaing products constant. The hump of the profit function is not at the post-merger price but somewhere between the pre- and post-merger prices.

Finally, we provide evidence consistent with recent theoretical results suggesting that multiple equilibria are more likely in markets with larger share of the inside goods. In our case, only the cereal industry exhibits such property, where the share of the inside goods, before the hypothetical merger, often exceeds 50% of the total market size. We can use the variation in the inside-good share across geographic markets to see how variation in the solutions to the FOCs correlated with inside-good share. Based on the theoretical results, we should observe more variation in the Bertrand FOC solutions for those market with higher inside market shares. Such variation in the FOC solutions should also translate into variation of economic variables of interest, such as the average compensative variation (ACV).

We analyze the variation in ACV implied by variation in the share of the inside goods as follows. First, we calculate the ACV and the share of the inside goods for each combination of demand non-linear search algorithm, market, and FOC solutions.[10] Second, we calculate the standard deviation of ACV and the average share of the inside goods by algorithm and

[9]Therefore, these histograms are constructed using $24 \times 13 = 312$ observations for percentage change in market shares.

[10]These calculations give rise to $(3 \times 24 \times 3 = 936)$ values for ACV and the share of the inside goods.

market.[11] Figure 5 provides a scatterplot of these pairs of summary statistics illustrating the positive relationship between inside shares and variation in FOC solutions. This positive relationship appears to be stronger for pairs characterize by inside shares exceeding 30%, which is in the spirit of the results of Allon et al. (2010) with respect to multiple equilibria in the case of single-product firms.

Moving to the automobile data, Tables 4 through 6 indicate substantial variation due post-merger market performance primarily due to variation in demand parameter estimates.[12] Specifically, post-merger variable profits range from \$5 billion dollars to over \$180 billion.[13] Omitting the Mesh Adaptive Direct Search demand estimates reduces the variation considerably, but one could argue that variable profits differences of over \$30 billion is still substantial (\$35B compared to \$5B). Compensating variation shows similar variation ranging from \$219 million to \$1.6 billion. Finally, average changes in prices range from \$28 to \$200. The market share changes in Figures 6 through 8 show a more intuitive pattern for a Bertrand-Nash equilibrium. No product exits the market, and the changes are, on average, negative with left part of the distribution predominantly composed of products from the two merging firms.

5 Conclusions

We illustrate the challenges in merger simulations with sophisticated aggregate demand systems of the BLP tradition and Betrand competition. The first challenge is the solution of a highly non-linear minimization problem associated with demand estimation, as documented in Knittel and Metaxoglou (2008). Different starting values and non-linear search algorithms yield different demand estimates. These different demand estimates lead to economically significant variation in post-merger market performance. The second challenge is the solution of a complex non-linear optimization problem associated with the Bertrand-Nash equilibrium. Holding the set of demand parameters constant, post-merger market performance varies significantly across different starting points used to find the Bertrand equilibrium.

The juxtaposition of the two industries is consistent with recent theoretical work which finds that for single-product firms, multiple Bertrand-Nash equilibria may exist when the inside-good

[11]This new set of calculations gives rise to ($13x24 = 312$) pairs of ACV standard deviations and average inside shares.

[12]The entries of these tables are calculated for year 1990, which corresponds to the market with the largest potential size in the BLP data. We assume a market size of 94 million households.

[13]Using the demand parameters from the Mesh Adaptive Direct Search algorithm and marginal costs as starting values for the FOC search, the FOC search did not converge.

market share is high. While we do not claim to find multiple Nash equilibria, indeed we claim that we do not, this theoretical result suggests that the pricing-best-response functions are more likely to exhibit multiple cross points when this is the case. We find variation in the solutions to the FOCs, across starting values, in the cereal industry, where the inside-good market share for the cereal industry can be as high as 70 percent, but not in the automobile industry where the inside-good market share is roughly 9 percent.

Our results highlight two separate, but related, points. First, while it seems like an obvious point, merger simulations must be sure to check the second order conditions. Given our results, merger simulations may be a time consuming endeavor. Because it appears that in some cases non-linear search algorithms are prone to find local minima for at least some of the products, searching for a Bertrand-Nash equilibrium may entail a large number of searches until the second order conditions are met.

A second implication of our results is that including the FOCs from a Bertrand-Nash equilibrium as additional moments in the GMM objective function associated with demand estimation can lead to inconsistent estimates. Including the FOCs in the GMM objective function is not entirely analogous to our empirical exercise since demand estimation seeks to choose demand elasticities to rationalize observed prices, whereas the merger simulation chooses prices to zero out the FOCs, given elasticities. However, there is no guarantee, within the GMM problem, that the FOC moments are not choosing elasticities to minimize the profits of a subset of the products. Insofar as this occurs, the moments are misspecified leading to inconsistent estimates.

References

ALLON, G., A. FEDERGRUEN, AND M. PIERSON (2010): "Price Competition Under Multinomial Logit Demand Function with Random Coefficients," Tech. rep., Northwester University.

BAJARI, P., J. FOX, AND S. RYAN (2007): "Linear regression estimation of discrete choice models with nonparametric distributions of random coefficients," *The American Economic Review*, 97, 459–463.

BASS, F., D. HUANG, AND C. ROJAS (2008): "What Happens When Demand is Estimated with a Misspecied Model," *Journal of Industrial Economics*, 159–180.

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

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

BUDZINSKI, O. AND I. RUHMER (2010): "Merger simulation in competition policy: a survey," *Journal of Competition Law and Economics*, 6, 277–319.

CHAN, T. Y. (2006): "Estimating a Continuous Hedonic Choice Model with an Application to Demand for Soft Drinks," *The RAND Journal of Economics*, 37, 1027–1057.

CONLON, C. T. (2010): "An MPEC Approach to Empirical Likelihood Demand Estimation," Tech. rep., Yale University.

CRUZ, W. L., J. MARTINEZ, AND M. RAYDAN (2006): "Spectral residual method without gradient information for solving large-scale nonlinear systems of equations," *Mathematics of Computation*, 75, 1429–1448.

DAVIS, P. AND E. GARCES (2009): *Quantitative Techniques for Competition and Antitrust Analysis*, Princeton University Press.

DAVIS, P. AND R. RIBEIRO (2010): "A Simple Globally Consistent Continuous Demand Model for Market Level Data," Tech. rep., London School of Economics and Political Science.

DUBE, J., J. FOX, AND C.-L. SUN (2009): "Improving the Numerical Performance of BLP Static and Dynamic Discrete Choice Random Coefficients Demand Estimation," *NBER Working Paper.*

Epstein, R. and D. Rubinfeld (2001): "Merger Simulation: A Simplified Approach with New Applications," *Antitrust Law Journal*, 69, 883–919.

FTC and DOJ (2010a): *Hart-Scott-Rodino Annual Report*, Federal Trade Commission and Department of Justice.

——— (2010b): *Horizontal Merger Guidelines*, Federal Trade Commission and Department of Justice.

Hausman, J., G. Leonard, and J. Zona (1994): "Competitive Analysis with Differenciated Products," *Annales d'Economie et de Statistique*, 159–180.

Hausman, J. A. and G. K. Leonard (2005): "Using merger simulation models: Testing the underlying assumptions," *International Journal of Industrial Organization*, 23, 693–698.

Ivaldi, M. and F. Verboven (2005): "Quantifying the Effects from Horizontal Mergers in European Competition Policy," *International Journal of Industrial Organization*, 23, 699–702.

Jiang, R., P. Manchanda, and P. E. Rossi (2009): "Bayesian analysis of random coefficient logit models using aggregate data," *Journal of Econometrics*, 149, 136–148.

Kalouptsidi, M. (2010): "From market shares to consumer types: duality in differentiated product demand estimation," *forthcoming, Journal of Applied Econometrics.*

Knittel, C. R. and K. Metaxoglou (2008): "Estimation of Random Coefficient Demand Models: Challenges, Difficulties and Warnings," *NBER Working Paper 14080.*

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by P. Zarembka, New York: Academic Press, 105–142, 6th ed.

Miller, N. H. and M. Osborne (2010): "Competition among Spatially Differentiated Firms: An Empirical Model with an Application to Cement," Tech. rep., Economic Analysis Group Discussion Paper.

Morrow, W. and S. Skerlos (2010): "A Fixed-Point Approach to Equilibrium Pricing in Differentiated Product Markets." *Forthcoming, Operations Research.*

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

——— (2000b): "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

——— (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

Nocedal, J. and S. Wright (1999): *Numerical optimization*, vol. 1, New York: Springer.

Pakes, A. and D. Pollard (1989): "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027–1057.

Peters, C. (2006): "Evaluating the performance of merger simulation: Evidence from the US airline industry," *Journal of Law and Economics*, XLIV, 627–649.

Petrin, A. and K. Train (2010): "A Control Function Approach to Endogeneity in Consumer Choice Models," *Journal of Marketing Research*, 47, 3–13.

Pinkse, J. and M. E. Slade (2004): "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 48, 617–643.

Small, K. and H. S. Rosen (1981): "Applied welfare economics with discrete choice models," *Econometrica*, 49, 105–130.

Su, C.-L. and K. L. Judd (2008): "Constrained Optimization Approaches to Estimation of Structural Models," Tech. rep., Stanford University.

Train, K. (2009): *Discrete Choice Methods with Simulation*, Cambridge University Press.

Weinberg, M. C. and D. Hosken (2008): "Using Mergers to Test a Model of Oligopoly," Tech. rep., University of Georgia.

Werden, G. (1997): "Simulating the Effects of Differentiated Products Mergers: A Practitioners' Guide," *Proceedings: Strategy and Policy in the Food System: Emerging Issues, June 20-21, 1996, Washington, DC*.

Werden, G. J. and L. M. Froeb (2002): "The Antitrust Logit Model for Predicting Unilateral Competitive Effects," *Antitrust Law Journal*, 70, 257–260.

——— (2006): "Unilateral Competitive Effects of Horizontal Mergers," in *Handbook of Antitrust Economics*, ed. by P. Buccirossi, MIT Press, 105–142.

A Tables and Figures

Demand Estimation			**Merger Simulation Starting Values**		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	19.553	3,649	3,768	3,768
Quasi-Newton 2	Derivative based	4.562	1,440	3,538	265
Conjugate Gradient	Derivative based	18.561	3,040	3,798	3,798
SOLVOPT	Derivative based	4.562	1,441	3,539	266
KNITRO 1	Derivative based	15.456	3,802	3,837	3,837
KNITRO 2	Derivative based	15.465	3,802	3,837	3,837
KNITRO 3	Derivative based	4.562	1,442	3,539	266
Simplex	Direct search	17.224	3,360	3,769	3,769
Mesh Adaptive Direct Search	Direct search	17.095	3,220	4,024	4,024
Generalized Pattern Search	Direct search	50.993	2,747	4,070	4,070
Simulated Annealing 1	Stochastic search	131.321	2,307	5,775	2,691
Genetic Algorithm	Stochastic search	34.236	75	75	0
Simulated Annealing 2	Stochastic search	108.128	2,346	2,862	2,767

Note: Post-merger annual industry profits in millions of dollars.

Table 1: Post-merger annual industry profits in the cereal industry

Demand Estimation			**Merger Simulation Starting Values**		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	19.553	31	-97	-97
Quasi-Newton 2	Derivative based	4.562	3,698	953	5,859
Conjugate Gradient	Derivative based	18.561	1,111	-122	-122
SOLVOPT	Derivative based	4.562	3,699	953	5,862
KNITRO 1	Derivative based	15.456	-48	-63	-63
KNITRO 2	Derivative based	15.465	-47	-62	-62
KNITRO 3	Derivative based	4.562	3,700	953	5,863
Simplex	Direct search	17.224	116	-87	-87
Mesh Adaptive Direct Search	Direct search	17.095	1,158	-129	-129
Generalized Pattern Search	Direct search	50.993	1,502	-182	-182
Simulated Annealing 1	Stochastic search	131.321	5,972	817	7,602
Genetic Algorithm	Stochastic search	34.236	44	44	98
Simulated Annealing 2	Stochastic search	108.128	1,679	-25	-1,024

Note: Change in annual consumer welfare in millions of dollars.

Table 2: Change in annual consumer welfare in the cereal industry

Demand Estimation			**Merger Simulation Starting Values**		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	19.553	3.01	0.06	0.06
Quasi-Newton 2	Derivative based	4.562	27.42	8.38	36.22
Conjugate Gradient	Derivative based	18.561	23.00	0.03	0.03
SOLVOPT	Derivative based	4.562	27.41	8.38	36.22
KNITRO 1	Derivative based	15.456	2.67	0.12	0.12
KNITRO 2	Derivative based	15.465	2.53	0.12	0.12
KNITRO 3	Derivative based	4.562	27.41	8.38	36.22
Simplex	Direct search	17.224	11.81	0.08	0.08
Mesh Adaptive Direct Search	Direct search	17.095	28.94	0.09	0.09
Generalized Pattern Search	Direct search	50.993	27.59	-0.05	-0.05
Simulated Annealing 1	Stochastic search	131.321	51.74	18.96	56.06
Genetic Algorithm	Stochastic search	34.236	1.12	1.12	11.55
Simulated Annealing 2	Stochastic search	108.128	19.03	7.22	-1.06

Note: Change in average price, weighted by pre-merger market share, in cents per serving.

Table 3: Average change in prices in the cereal industry

Demand Estimation			**Merger Simulation Starting Values**		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	215.08	27	27	27
Quasi-Newton 2	Derivative based	207.49	27	27	27
Conjugate Gradient	Derivative based	215.09	27	27	27
SOLVOPT	Derivative based	178.06	29	29	29
KNITRO 1	Derivative based	277.58	35	35	35
KNITRO 2	Derivative based	277.63	35	35	35
KNITRO 3	Derivative based	277.73	35	35	35
Simplex	Direct search	215.09	27	27	27
Mesh Adaptive Direct Search	Direct search	215.07	27	27	27
Generalized Pattern Search	Direct search	196.70	27	27	27
Simulated Annealing 1	Stochastic search	193.25	23	23	23
Genetic Algorithm	Stochastic search	215.60	180	187	NA
Simulated Annealing 2	Stochastic search	180.06	5	5	5

Note: Post-merger annual industry profits in billions of dollars.

Table 4: Post-merger annual industry profits in the automobile industry

Demand Estimation			Merger Simulation Starting Values		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	215.08	711	711	711
Quasi-Newton 2	Derivative based	207.49	1,235	1,235	1,193
Conjugate Gradient	Derivative based	215.09	712	712	712
SOLVOPT	Derivative based	178.06	1,550	1,550	1,550
KNITRO 1	Derivative based	277.58	1,141	1,141	1,141
KNITRO 2	Derivative based	277.63	1,134	1,134	1,134
KNITRO 3	Derivative based	277.73	1,151	1,151	1,151
Simplex	Direct search	215.09	716	716	716
Mesh Adaptive Direct Search	Direct search	215.07	728	728	728
Generalized Pattern Search	Direct search	196.70	1,394	1,394	1,394
Simulated Annealing 1	Stochastic search	193.25	1,197	1,197	1,197
Genetic Algorithm	Stochastic search	215.60	1,610	382	NA
Simulated Annealing 2	Stochastic search	180.06	219	219	219

Note: Reduction in annual consumer welfare in millions of dollars.

Table 5: Change in annual consumer welfare in the automobile industry

Demand Estimation			Merger Simulation Starting Values		
Algorithm Description	**Algorithm Class**	**GMM Value**	**App. Post-Merger Prices**	**Pre-Merger Prices**	**Marginal Costs**
Quasi-Newton 1	Derivative based	215.08	87	87	87
Quasi-Newton 2	Derivative based	207.49	151	151	149
Conjugate Gradient	Derivative based	215.09	87	87	87
SOLVOPT	Derivative based	178.06	187	187	187
KNITRO 1	Derivative based	277.58	140	140	140
KNITRO 2	Derivative based	277.63	139	139	139
KNITRO 3	Derivative based	277.73	142	142	142
Simplex	Direct search	215.09	87	87	87
Mesh Adaptive Direct Search	Direct search	215.07	89	89	89
Generalized Pattern Search	Direct search	196.70	172	172	172
Simulated Annealing 1	Stochastic search	193.25	146	146	146
Genetic Algorithm	Stochastic search	215.60	200	45	NA
Simulated Annealing 2	Stochastic search	180.06	28	28	28

Note: Change in average prices in dollars.

Table 6: Average change in prices in the automobile industry

Figure 1: Percentage change in market shares resulting from merger, in the cereal industry using approximated post-merger prices as starting values for the FOC search

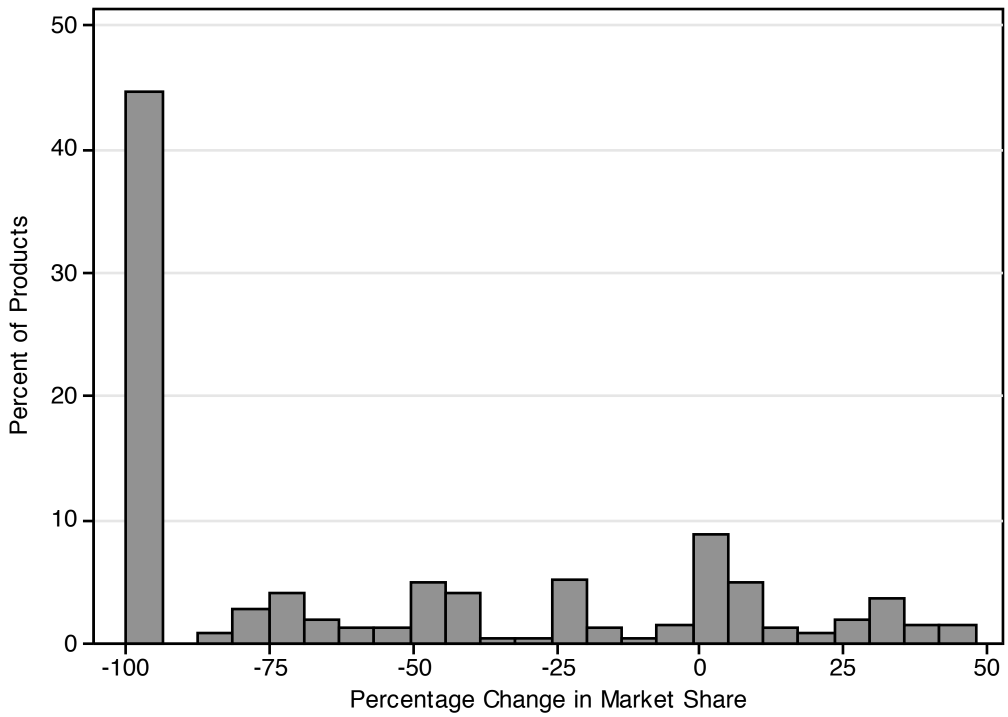


Figure 2: Percentage change in market shares resulting from merger, in the cereal industry using pre-merger prices as starting values for the FOC search

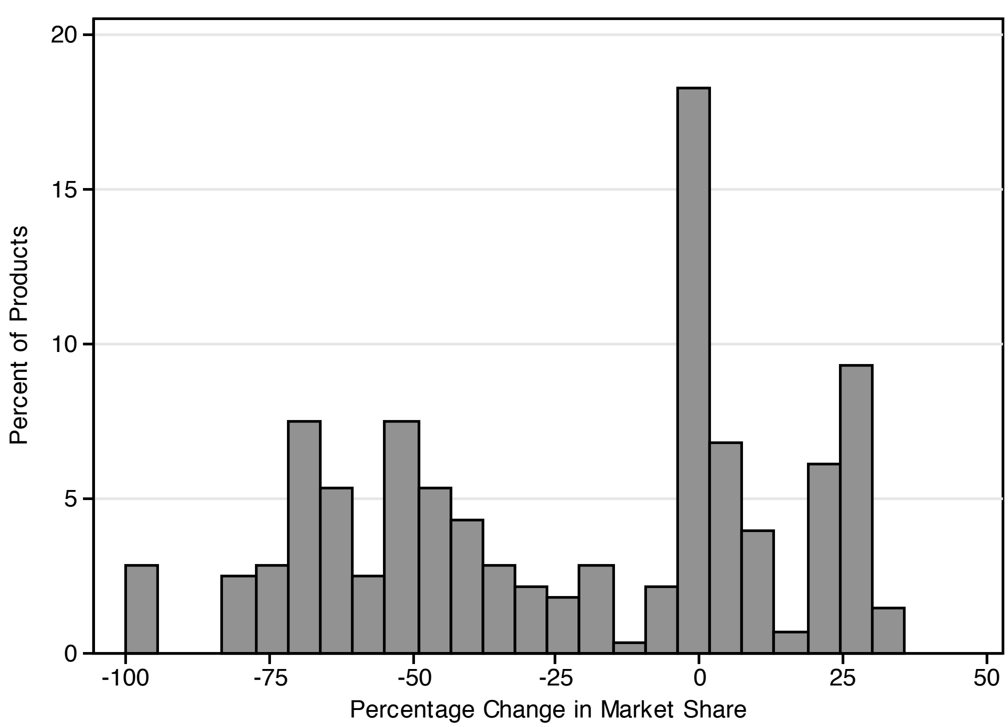


Figure 3: Percentage change in market shares resulting from merger, in the cereal industry using MC as starting values for the FOC search

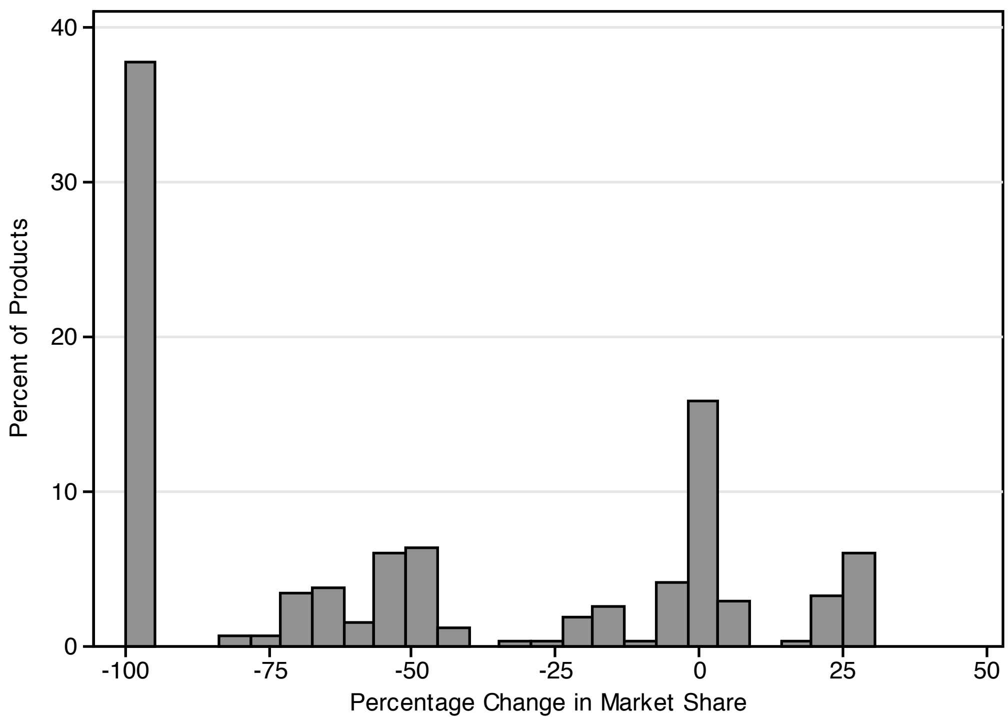
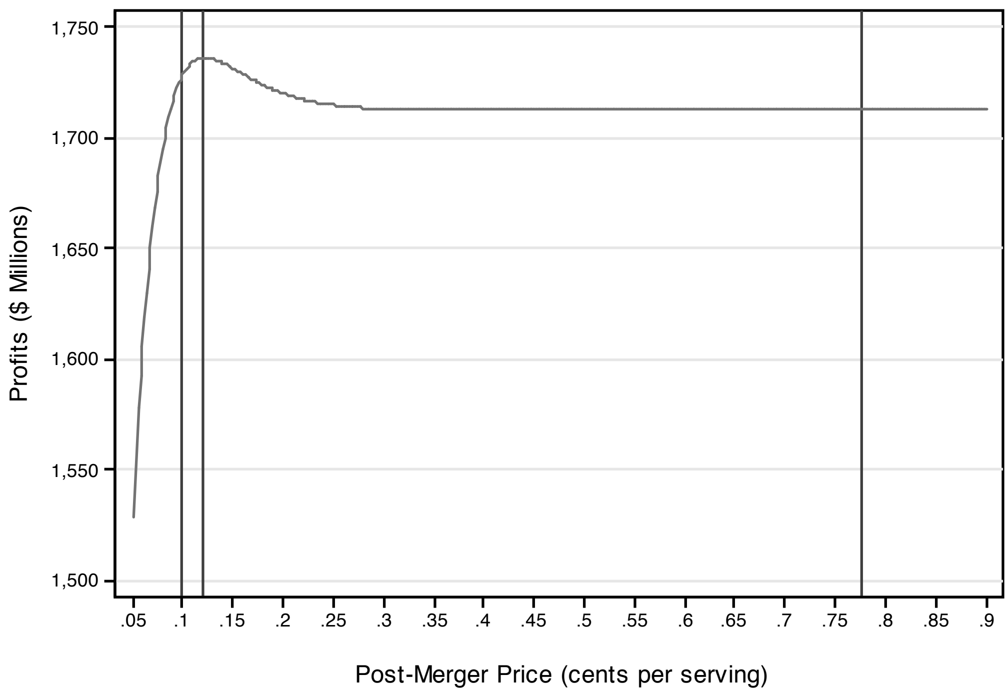


Figure 4: Percentage change in market shares resulting from merger, in the cereal industry



Note: First vertical line represents pre-merger price, the second the post-merger profit maximizing prices (holding other prices constant), and the third, the solution to the FOC.

Figure 5: Percentage change in market shares resulting from merger, in the cereal industry

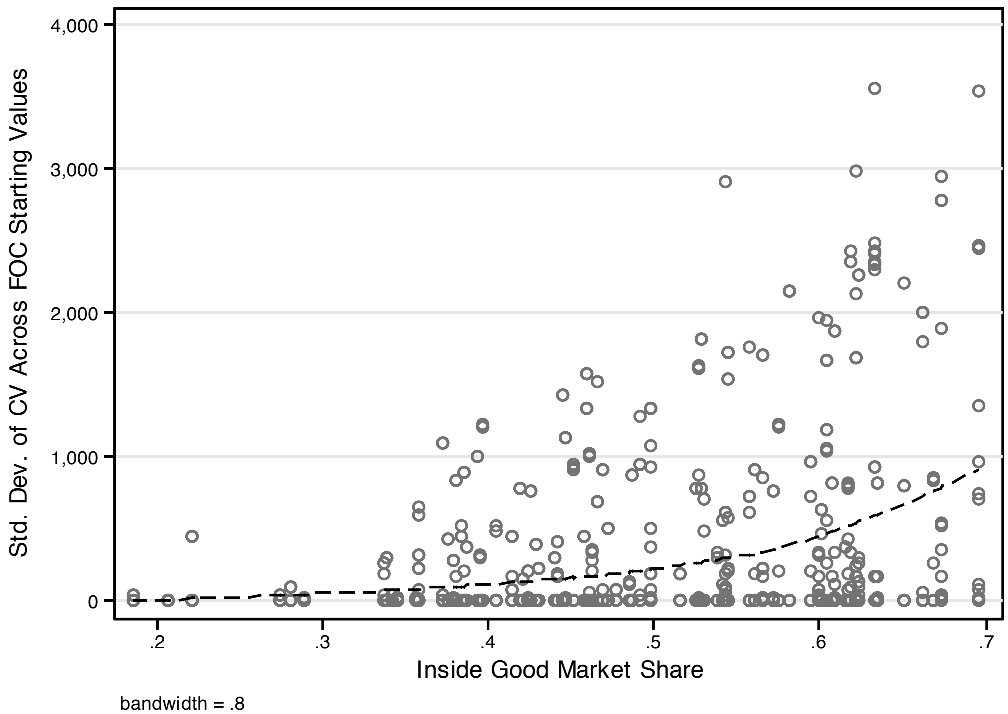


Figure 6: Percentage change in market shares resulting from merger, in the automobile industry using approximated post-merger prices as starting values for the FOC search

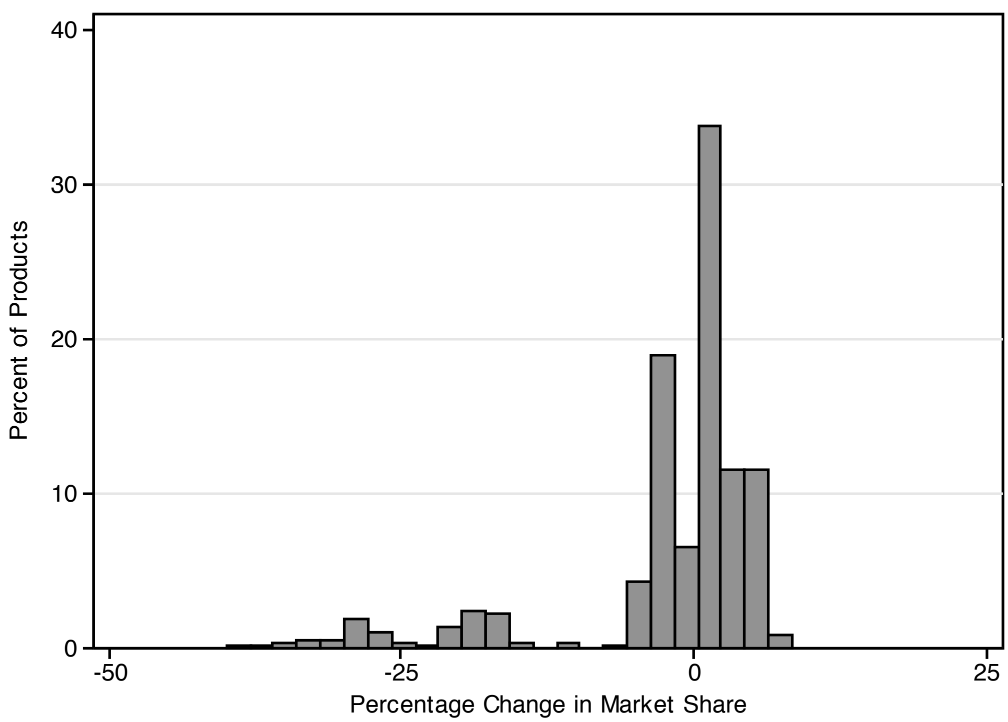


Figure 7: Percentage change in market shares resulting from merger, in the automobile industry using pre-merger prices as starting values for the FOC search

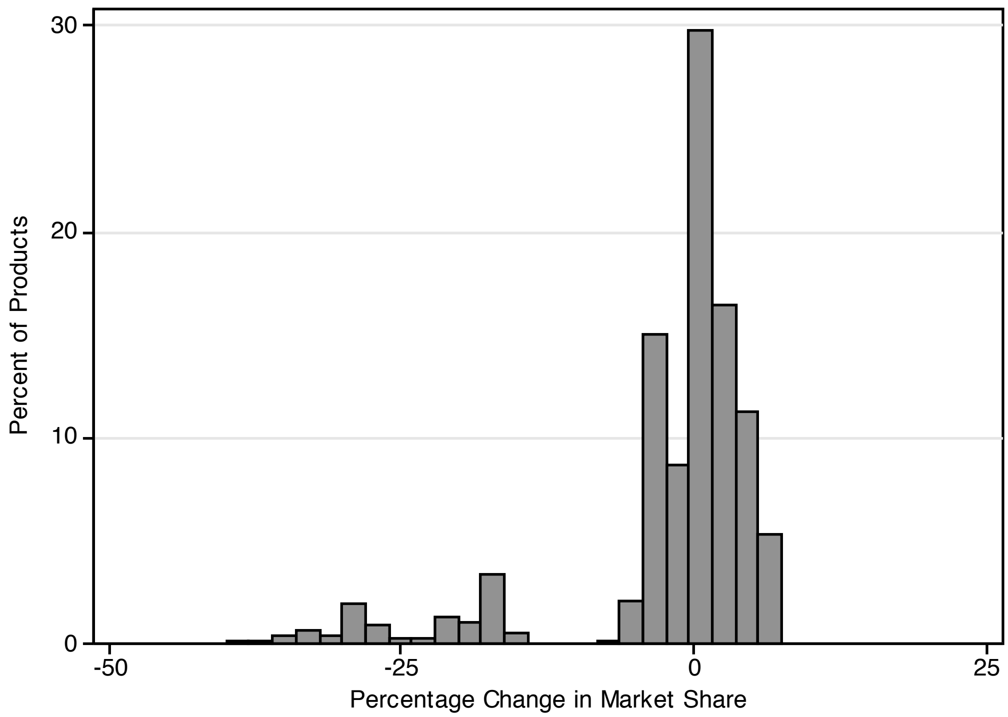


Figure 8: Percentage change in market shares resulting from merger, in the automobile industry using MC as starting values for the FOC search