UNIVERSITY OF CALIFORNIA

Los Angeles

Essays on Merger Simulation in Industrial Organization

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Economics

by

Mee Yeon Kim

2012

ABSTRACT OF THE DISSERTATION

Essays on Merger Simulation in Industrial Organization

by

Mee Yeon Kim

Doctor of Philosophy in Economics

University of California, Los Angeles, 2012

Professor Ackerberg, Daniel A, Co–Chair

Professor Snider, Connan Andrew, Co–Chair

This collection of essays considers three issues regarding the performance of the structural merger simulation. The first chapter addresses the underprediction problem of merger simulation by considering more flexible demand models. To be as flexible as possible in modeling the demand, I estimate the full variance–covariance structure of brand preferences and let the brand intercepts to be different across cities to exploit the multi–market structure of data. Compared to a simple logit demand model, a flexible setting allows the predicted price changes to be closer to the actual price changes. In Chapter 2, because the underprediction problem still exists with the flexible structural demand model, I consider the Almost Ideal Demand System as an alternative demand model to determine how the simulation performance differs from that of Chapter 1. I also address the firm-side modeling issues in Chapter

2. I examine, whether incorporating the retailer side in the pricing game would produce improved merger simulation results. In Chapter 3, I separately estimate the effect of several factors compounded in the actual price changes in the post–merger period. I explain that part of the reasons the merger simulation generally fails to result in accurate predictions is that both demand–side factors and supply–side factors considerably affect the post–merger prices.

The dissertation of Mee Yeon Kim is approved.

HAHN, JINYONG

RILEY, JOHN G

AINSLIE, ANDREW S.

ACKERBERG, DANIEL A, Committee Co–Chair

SNIDER, CONNAN ANDREW, Committee Co–Chair

University of California, Los Angeles

2012

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

It is with much appreciation and gratitude that the author acknowledges the guidance provided by the committee members, Daniel Ackerberg, Connan Snider, John Riley, Jinyong Hahn, and Andrew Ainslie. Especially, the author also acknowledges Daniel Ackerberg for his countless hours of comments, encouragement, and generous guidance during spring 2011 in Michigan. The author also would like to appreciate Connan Snider for generously agreeing to serve as a Co-Chair and providing prompt access and feedback. The author would like to give special thanks to Dr. Minha Hwang, who was extremely helpful and gave constant encouragement and comments during the entire course of research.

The author would also like to thank participants in the industrial organization proseminar at UCLA for their helpful comments. Whatever errors remain are the author's alone.

Finally, the author would like to acknowledge the Ilju Academy and Culture Foundation, UCLA Economics Department, and the UCLA Graduate Division for providing financial support and necessary assistance.

VITA

2004	B.A., Business Administration and Economics, Summa Cum Laude
	Seoul National University
	Seoul, Korea
2006	M.A., Economics
	University of California, Los Angeles
	Los Angeles, California
2005-2006	First-Year Fellowships for Doctoral Studies
	UCLA Graduate Division
2005-2009	Four-Year Fellowships for Graduate Studies
	Ilju Academy and Culture Foundation
	Seoul, Korea
2006-2007	Teaching Assistant
	UCLA Department of Economics
2007-2010	Teaching Associate
	UCLA Department of Economics

1 Chapter 1. Evaluation of the Performance of the Structural Merger Simulation, Revisited: A Case of the Consumer Packaged Goods Industry

1.1 Introduction

Merger simulation based on the structural economic model has been widely employed by many researchers in the industrial organization (IO) and marketing fields to calculate the market power effects of actual or hypothetical mergers in a variety of industries from airlines and banking consumer goods. Not limited only to academics, merger simulation is an important tool used by policy makers and antitrust authorities such as the U.S. Federal Trade Commission and U.S. Department of Justice, who receive thousands of merger proposals every year. Except for a few cases considered to have anticompetitive effects, most proposed mergers are approved without any challenges. Meanwhile, as pointed out in Ashenfelter and Hosken (2008), little evaluation has addressed whether merger simulation is successful in predicting price changes. How successfully does the simulation predict the merger effects? This is essential information for determining the validity of merger analysis in many studies and evaluating the effectiveness of antitrust policy. Few studies have evaluated merger simulation because of the limited availability to researchers of both pre–merger and post–merger data.

Because it is an important policy question, the evaluation of merger simulation is receiving increasing attention. The attempts to assess the performance of the merger simulation include Peters (2006) and Weinberg (2011). They found that the merger simulation failed to predict accurately the price changes in several mergers in the airline and consumer hygiene products industries, respectively. According to their papers, differences between actual price changes and predicted price changes can be attributed to some simplifying assumptions in the merger simulation, for example, fixed marginal cost, fixed product offerings, and so on.

Basically, the evaluation of the structural merger simulations[1] is based on comparing structural merger simulation results to some direct measures of merger effects on prices. Structural merger simulation results are the simulated price changes that the industry would expect from the merger. A simulation is typically conducted by estimating a structural model of demand and firm behavior, and then conducting a what–if analysis—a counterfactual experiment—to simulate how the prices would change if a merger occurred. The direct measures refer to the estimated merger treatment effects using reduced–form analysis of retrospective data.

Two commonly used estimates in the literature are the time–difference estimates of price changes and difference–in–differences estimates of price changes. The difference–in–differences method estimates relative price changes of the merged firms' products to the price changes of a control group considered to have been minimally affected by the merger event. However, a merger is not like an experiment, in which a treatment group and the control group can be clearly defined. For example, especially when national brand manufac-

[1] Standard merger simulation can use either structural or non–structural demand model. I use structural merger simulation to refer to the former case.

turers merge, there should be no markets or competitor brands that are totally immune to the merger. Because of the strict assumptions, a merger simulation is likely to result in an imprecise merger effect compared to an actual merger effect.

In this paper, I analyze a consummated merger between two U.S. peanut butter manufacturers. In October 2001, the JM Smucker Company announced that an agreement to merge the Jif brand of Proctor & Gamble into the company had been settled. Table 1 shows the pre–merger median market share of each of the top five brands of peanut butter across 47 U.S. cities, with other summary statistics from the U.S. peanut butter industry. As shown in the table, Smuckers[2] brand had relatively low market shares compared to the other four brands. Because of the small market share of the Smuckers brand, a merger between Smuckers and Jif was expected to have minimal effect on prices. The motive of the merger did not seem to be to excercise monopoly power.

Given this information, I conducted a structural merger simulation to obtain predicted post–merger changes in the prices of the merged firm's products. Specifically, I estimated a more flexible structural demand model by allowing for a correlation structure between brand preferences. To make the demand model as flexible as possible, I also allowed brand intercepts to be different across cities to fully account for different rivalry structures across cities. As documented in Bronnenberg, Dahr and Dube (2009), brands' market shares show significant heterogeneity across different geographic markets. I find this is also true in the U.S. peanut butter industry in figure 1. By using a flexible demand model, I examine whether

[2] I aggregate three natural and organic line brands of JM Smucker Company, SMUCKERS, LAURA SCUDDER, and ADAMS to a single brand "Smuckers".

a more flexible setting would produce different simulation results.

To evaluate the performance of the merger simulation, the simulation results must be compared to the actual price changes. To get the actual price changes in the post–merger period, I used methods from retrospective merger analysis and directly estimate (1) time–difference and (2) difference–in–differences estimates of the price changes for each brand. These two methods of estimation were used to remove the effect of other factors on prices and to isolate only the merger effects on prices. Thus, these sole merger effects are regarded as the actual price changes attributed to the merger.

In the empirical results, the merger simulation predicts the price increases for Jif and Smuckers to be lower than the direct estimates of price increases for Jif and Smuckers. The flexible random coefficient demand model results are not significantly different from other simple demand models. However, it shows the closest predictions to the actual price changes. In the discussion, I compare them and explain why the simulation returns a lower prediction value than the actual changes. I also explain the factors that may cause the discrepancies between the simulation results and the actual changes.

The remainder of this chapter is organized as follows. The next section reviews related literature. Section 3 describes the industry and the data used in the analysis. Section 4 presents the demand model and the merger counterfactuals. Section 4 also presents the direct estimation method. Section 5 briefly reviews the estimation method. In Section 6, the empirical results are presented and discussions follow. Conclusions and recommendations for future research are presented in Section 7.

1.2 Literature Review

In the empirical industrial organization field, many researchers have used structural merger simulations to study the effects of mergers on product prices. Most studies have addressed the price effects, but some have investigated the effects of mergers on other competition tools, such as product variety (Berry and Waldfogel 2001). Examples of studies using structural merger simulation include Nevo (2000) and Dube (2005). Nevo (2000) examines five mergers in the ready–to–eat cereal industry and simulates the price changes as well as social welfare changes resulting from the mergers. Dube (2005) studies the impact of several mergers in the U.S. carbonated soft drink industry. He uses the merger simulation to predict the equilibrium prices and welfare changes of such mergers as Coke and Dr. Pepper (rejected), Pepsi and 7–Up (withdrawn), and Coke and Pepsi (hypothetical). Nevo (2000) mentions that a natural question would be how well the predictions fit the actual outcomes, admitting that the tests cannot be conducted because of the lack of detailed data for the other period of the merger.

Another strand of literature in merger studies is retrospective merger analysis. While the merger simulation attempts to predict the merger effects by using one period of a merger, either before or after, retrospective merger analysis exploits both periods of data on mergers and examines how the mergers actually affect the prices. By its nature, retrospective merger analysis can be conducted only for consummated mergers. Examples of such studies include Farrell, Pautler, and Vita (2009) investigating hospital mergers, Kim and Singal (1993) examining airline industry mergers, and Ashenfelter and Hosken (2008) investigating mergers in various industries.

To predict mergers' price effects and to determine whether the proposed mergers should be passed or rejected, merger simulation is also used by antitrust authorities. The success of the antitrust policy depends on the quality of the simulation results. The literature on evaluating merger simulations serves this purpose. However, few studies have addressed this topic because of the lack of data. Because the evaluation combines the merger simulation and the retrospective analysis, it requires one to have data from both the pre–merger and the post–merger period.

Recently, some attempts have been made to evaluate how well merger simulation methods perform in estimating the effects of mergers. Using several demand models including a logit, GEV, and linear demand model, Peters (2006) examines several mergers in the airline industry during the 1980s and finds that the merger simulation does not accurately predict the actual price changes. He estimates the effects of some factors that result in the differences between simulated price changes and actual price changes. More recently, Weinberg and Hosken (2012) evaluate the merger simulation in the breakfast syrup and the motor oil industries. In addition, Weinberg (2011) studies Proctor & Gamble's purchase of Tambrands to evaluate the performance of the merger simulation method. Under a simple logit and a nested logit demand models, he finds the simulations significantly underestimate the price effects of the merger considered in the study.

1.3 Industry Overview and Data

1.3.1 U.S. Peanut Butter Industry

The U.S. peanut butter industry is dominated by four major national manufacturers' brands, with annual domestic sales totaling approximately $800 million[3]. As of 2001, Procter & Gamble's Jif brand and Unilever's Skippy brand accounted for approximately 54% of the market share in terms of volume sales across the nation, having 33% and 21%, respectively. ConAgra Foods' Peter Pan brand had 18% of the sales and the J.M. Smucker Company's three major brands—Smuckers, Laura Scudder, and Adams[4]—had about 3%. Other than national manufacturers, private label brands held a relatively strong place in this industry, being third largest at a 25% market share. Considering there are over 50 brands in total including local brands, this industry is a concentrated market with average 2,293 of HHI[5] across 47 U.S. cities without including private labels. The concentration measure, HHI, in the U.S. peanut butter industry shows great variability across cities with a minimum of 1,139 and a maximum of 4,441. Among these, if segmented by quality, private label brands were low end, Jif, Skippy, and Peter Pan were medium, and J.M. Smucker Company's brands were in the high–end segments, producing organic and natural variants of peanut butter. The market shares are summarized in Table 1.

[3] http://barfblog.foodsafety.ksu.edu/barfblog/tag/peanut-butter

[4] Santa Cruz Organic, which is J.M. Smucker Company's high–quality peanut butter product, was introduced to the market in 2006. Because this analysis focuses on 2001 through 2003, Santa Cruz Organic has been excluded from the study.

[5] The Herfindahl-Hirschman Index (HHI) is a commonly accepted measure of market concentration. It is calculated by summing the square of the market share of each firm competing in the market. Under the Horizontal Merger Guidelines, markets in which the HHI is in excess of 1,800 points are considered to be concentrated.

The U.S. peanut butter industry is one of the non–durable consumer packaged goods industries that show heterogeneous rivalry structures. For the U.S. peanut butter industry, I reproduced similar pictures showing distribution of market share levels across U.S. markets as in Bronnenberg, Dahr, and Dube (2009). Figure 1-(1) and 1-(2) show market share asymmetries of the top two brands, Jif and Skippy. As shown, Jif shares were large in the East Coast region, where the Skippy shares were minimal. The market share summary statistics in Table 1 also show market share heterogeneity (S.D.) across cities. In addition, as noted in Nevo (2000), ANOVA analysis shows that variation in prices is largely due to differences across markets after controlling for the variation between brands. Hence, price response could be affected differently across markets. To take into account different brand strengths across cities, I allowed brand intercepts to be different across cities in the demand model.

Table 1: **Nationwide Market Shares (%) Among Inside Goods as of 2001**

Brands	**Median**	**Mean**	**S.D.**	**Min**	**Max**
Jif	30.91	32.75	12.54	14.71	66.63
Skippy	15.63	21.31	18.13	0.40	54.37
Peter Pan	15.12	18.22	15.48	0.15	48.72
Private Label	24.74	25.27	5.56	15.43	43.32
Smuckers	2.48	3.61	3.17	0.56	13.90

Figure 1: Distribution of Market Share

Figure 1-(1). JIF's geographic market share asymmetry. Filled circles indicate the markets where JIF brand takes the highest market share.

Figure 1-(2). SKIPPY's geographic market share asymmetry. Filled circles indicate the markets where SKIPPY brand takes the highest market share.

1.3.2 The Merger

In October 2001, the J.M. Smucker Company announced that an agreement to merge the Jif brand into the company had been settled.[6] The transaction was finally closed in June 2002, according to the company's 2002 annual report. From the first observation of sales of Jif brands through the J.M. Smucker Company's vendor code in the data, I defined the merger date as of June 24, 2002.

After the acquisition of Jif brand, the J.M. Smucker Company became the largest player in the market, producing both high–priced products (Smuckers, Laura Scudder, and Adams) and standard–priced products (Jif). The peanut butter category became the largest product line that the J.M. Smucker Company sells (7% in 2001 to 26% in 2003).

As mentioned in Section 1, one of the parties of the merger had a very small market share

[6] "In October 2001, we announced an agreement with The Procter & Gamble Company to merge the Jif and Crisco brands into The J.M. Smucker Company. Our shareholders overwhelmingly approved the merger at a special shareholder meeting held last April, and we thank you for your vote of support. The transaction closed this past June 1." (J.M. Smucker Company, 2002 Annual Report)

relative to other major competitors. Therefore, it was not expected for this merger to have any significant anticompetitive effects in the market. Thus, the merger was not challenged by the antitrust authorities and was passed without any modification order.

1.3.3 Data

The raw data available for this study come from the Information Resources, Inc. (IRI) scanner dataset for consumer packaged goods (CPG) categories in a sample of supermarkets across the 50 U.S. cities from January 2001 through December 2006. Weekly store level prices, volume sales, measures of retailers' marketing activities and product characteristics are available at the UPC level. In addition, the estimated annual sales of each store and the information on which chain—without the specific name of retailers—each store belongs to are also available. I focus on the U.S. peanut butter market from 2001 through 2003 in observation of the consummated merger in mid–2002.

For the estimation, I construct the subsample based on several criteria. First, for pre–merger I use year 2001 (52 weeks) data and for post–merger I use year 2003 (52 weeks) data. This is, first, to eliminate any noise in pricing strategies during the transition period around the merger, and second, to make both pre–merger and post–merger data encompass the same period over a year. Second, because of the frequent close–out of stores or suspension of sales records disclosure, many stores have missing data during the sample period. Hence, I select only stores that provide data consistently[7] for the entire period of 2001 through 2003.

[7] Stores with missing data for less than 5 weeks in a row.

In this way, I remove any noise in pricing that might have resulted from store closing or any other issues such as mergers of stores, store renovations, and so forth. Third, although there are 15 brands, on average, in each market and about 51 brands in total nationwide[8], I limit the analysis to only the top five brands, which have over 90% of the combined market shares. These include four national brands: Jif, Skippy, Peter Pan, and Smuckers, and one aggregated Private Label brand. Purchases of any other national or local brands are regarded as choosing outside alternatives. Using these three criteria, for each pre-merger and post-merger period, I end up with weekly data sets for five brands in each of 924 stores across 47 U.S. cities for 52 weeks. In the empirical model of this paper, the definition of a market will be a city-week combination.

To construct variables for use in the demand estimation, I aggregate all UPCs across sizes and variants (e.g., different flavor, textures, and fat content) into one product if they are under the same brand name. Given that manufacturers and retailers do not price discriminate across the variants, treating them as one product would not have significant effects on the analysis as long as pricing is the main concern.

One caveat of this aggregation is that price discrimination based on the sizes is ignored.[9] However, to study the overall direction of post–merger pricing response at the brand level, price discrimination within individual package sizes would not matter very much. Because of the aggregation, it is hard to define the appropriate product characteristics for the aggregated

[8] This number is only for 2001 though 2003. New brands introduced after 2003 are not included.

[9] There are concerns regarding composition bias because of the non-linear pricing practice in different package sizes (2nd–degree price discrimination).

products. Hence, I use brand dummies in the estimation to capture all intrinsic values of each of the brands. In addition, the demand estimation uses retail prices, market shares, and measures of marketing mix, such as display activity. The price of each brand is constructed as a volume-sales weighted average price across stores in a city-week. Ignoring the retailer side can be too restrictive given the high variance of prices across retail chains. If one wants to take manufacturer-retailer relationship into account, analysis at the chain level would be ideal by defining products as chain-brand combinations in each city-week, as in Villas-Boas (2007a). This alternative supply side model is examined in Chapter 2. Market shares are calculated by dividing total weekly volume–equivalent sales of each brand in a market by total weekly potential consumption in that market. (The most popular standard size, 16oz jar, is defined as one volume–equivalent.) I assume two servings (0.14 volume-equivalent)[10] per capita per week. The InfoScan website has the information on the population size in each IRI market. I construct the appropriate market size for the sample stores using each store's estimated annual sales and the information on the total annual sales of supermarkets in each market.

The share of outside alternatives is obtained by subtracting the sum of the inside goods' market shares from one. In many city-week combinations, outside shares are around 70%, which is normal in CPG industries. The outside option could be regarded as no purchase or purchases of other products that could substitute for peanut butter—for example, jam or other kinds of spreads. Summary statistics for these variables are shown in Table 2.

[10] One serving size is two table-spoons (32g).

Table 2: **Summary statistics**

	Median	Mean	S.D.	Min	Max
Price ($ per 1lb)	1.88	1.97	0.44	0.55	3.21
Share of a brand (%)	3.44	4.19	4.43	0.00	68.81
Combined shares of brands (%)	19.33	20.68	8.37	4.39	78.60
Average display	0.00	0.03	0.05	0	0.63
Average feature	0.00	0.10	0.23	0	3
Average promotion	0.06	0.11	0.15	0	1
Number of SKUs	6.92	6.20	3.27	1.00	13.57

1.4 The Empirical Model

1.4.1 Demand Model

Demand estimation is the most important part of merger simulation studies because the counterfactual price simulation depends on the quality of the demand parameters. The demand parameters determine price elasticities, which determine the market share response matrix (derivative of market shares with respect to prices) in first-order conditions of firms' pricing game. If the model does not reflect the actual demand pattern, then the simulation results should be misleading. In this section, to test a more flexible demand model, I estimate a random coefficient demand model taking into account the correlation structure of brand preferences.

I estimate the demand for peanut butter using both a logit model and a random coefficient discrete choice model. The logit model serves as a benchmark to be compared to the results of the random coefficient model. First, I specify the random coefficient demand model following the literature (McFadden 1984; BLP 1995). I assume the conditional indirect utility of consumer i from choosing brand $j = 1, ..., J$ in city $c = 1, ...C$ in time period $t = 1, ..., T$ as:

$$u_{ijct} = \beta_{ij}^c + \alpha_i p_{jct} + X_{jct}\gamma + d_t^c + \xi_{jct} + \epsilon_{ijct} \tag{1}$$

Hereafter the terms consumers, individuals, and households will be used interchangeably. The definition of a market is the city-time period combination. Model (1) above assumes individuals differ along their brand preferences and price sensitivities. This assumption allows the model to be flexible so that the substitution patterns are not driven solely by the logit errors, ϵ_{ijct}. β_{ij}^c is consumer i's brand-specific fixed effect that captures the consumer's time-persistent intrinsic preference for brand j in the market. The model will also be estimated using common β_{ij} across cities ignoring different brand strengths across cities. I assume brand preferences not only to vary across individuals but also to be correlated between brands. Precisely, I assume the parametric distribution of β_{ij}^c as follows:[11]

$$\beta_{ij}^c = \beta_j^c + \epsilon_{ij}^\beta, \qquad \epsilon_i^\beta \sim N(0, \Sigma) \tag{2}$$

where ϵ_i^β are J by 1 vector of correlated individual heterogeneity on brand preferences. The matrix Σ, which is the variance-covariance of ϵ_i^β, is given by

$$\Sigma = LL' \tag{3}$$

[11]See Jain et al. (1994) for semi-parametric consideration.

Using the Cholesky decomposition, I define L as the lower Cholesky matrix of Σ as follows:

$$L = \begin{bmatrix} \sigma_{11} & 0 & 0 & \ldots & 0 \\ \sigma_{21} & \sigma_{22} & 0 & \ldots & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{J1} & \sigma_{J2} & \sigma_{J3} & \ldots & \sigma_{JJ} \end{bmatrix} \tag{4}$$

Then the full variance-covariance matrix Σ is given by

$$\begin{aligned} \Sigma &= \begin{bmatrix} \sigma_{11} & 0 & 0 & \ldots & 0 \\ \sigma_{21} & \sigma_{22} & 0 & \ldots & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{J1} & \sigma_{J2} & \sigma_{J3} & . & \sigma_{JJ} \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} & \ldots & \sigma_{J1} \\ 0 & \sigma_{22} & \sigma_{32} & \ldots & \sigma_{J2} \\ 0 & 0 & \sigma_{33} & \ldots & \sigma_{J3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & \sigma_{JJ} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{11}^2 & \sigma_{11}\sigma_{21} & \sigma_{11}\sigma_{31} & \ldots & \sigma_{11}\sigma_{J1} \\ \sigma_{21}\sigma_{11} & \sigma_{21}^2+\sigma_{22}^2 & \sigma_{21}\sigma_{31}+\sigma_{22}\sigma_{32} & \ldots & \sigma_{21}\sigma_{J1}+\sigma_{22}\sigma_{J2} \\ \sigma_{31}\sigma_{11} & \sigma_{31}\sigma_{21}+\sigma_{32}\sigma_{22} & \sigma_{31}^2+\sigma_{32}^2+\sigma_{33}^2 & \ldots & \sigma_{31}\sigma_{J1}+\sigma_{32}\sigma_{J2}+\sigma_{33}\sigma_{J3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{J1}\sigma_{11} & \sigma_{J1}\sigma_{21}+\sigma_{J2}\sigma_{22} & \sigma_{J1}\sigma_{31}+\sigma_{J2}\sigma_{32}+\sigma_{J3}\sigma_{33} & \ldots & \sum_{j=1}^{J}\sigma_{j1}^2 \end{bmatrix}. \end{aligned} \tag{5}$$

In the estimation, one only needs to take random draw ν_i^β from $N(0,\, I_J)$ and pre–multiply it by L to obtain correlated ϵ_i^β across brands. The set $\{\sigma_{kl} \mid 1 <= l <= k <= J\}$ is the parameters to be estimated along with β_j^c, the mean brand preference for brand j across individuals in city c. After estimating the lower Cholesky decomposition matrix, I obtain the whole variance–covariance matrix of correlation of choices between brands from (5).[12] The number of parameters in L is $\frac{J(J+1)}{2}$. Hence, with large number of brands, estimation of Σ will suffer from the curse of dimensionality.

In such a case, factor-structure analysis (Goettler and Shachar 2001; Chintagunta 2000, Albuquerque and Bronnenberg 2009) would be useful. p_{jct} is the price of brand j in city c at time period t. The price coefficient α_i is assumed to be heterogeneous across individuals with a distributional assumption:

$$\alpha_i = \alpha + \sigma_p \nu_i^\alpha \qquad\qquad where \;\; \nu_i^\alpha \sim N(0,1) \tag{6}$$

where mean price sensitivity α and price heterogeneity σ_p are parameters to be estimated. X_{jct} includes the average number of SKUs and measure of average displaying activities (across stores) of brand j in city c at time period t. These are included in the specification to account for the effect of the retailers marketing activities, γ, on consumer utility. d_t^c are month-city dummies used to control for any seasonality effect on consumer utility. The variables d_t^c also allow outside option size to be different across both city and time. Finally, ξ_{jct} are time-varying brand-specific demand shocks in city c and ϵ_{ijct} are i.i.d. type I extreme-value distributed idiosyncratic errors.

[12] This matrix can also be estimated as different across cities along the dimension of city characteristics.

Rewriting the utility function using equations (2) and (6) leads to the following:

$$u_{ijct} = \delta_{jct}(p_{jct}, X_{jct}, \xi_{jct}; \beta_j^c, \alpha, \gamma) + \mu_{ijct}(p_{jct}, \nu_i^\alpha, \nu_i^\beta; \sigma_p, L) + \epsilon_{ijct} \tag{7}$$

$$\begin{aligned} where \qquad \delta_{jct} &= \beta_j^c + \alpha p_{jct} + X_{jct}\gamma + d_t^c + \xi_{jct} \\ \mu_{ijct} &= \epsilon_{ij}^\beta + \sigma_p \nu_i^\alpha p_{jct} \end{aligned} \tag{8}$$

$$for \quad i = 1, ..., I_t, \quad j = 1, ..., J, \quad t = 1, ..., T, \quad c = 1, ..., C.$$

The demand model also includes outside alternatives ($j = 0$) so that the consumers may choose not to purchase any of the brands in the choice set. The indirect utility of outside goods is:

$$u_{i0ct} = \delta_{0ct} + \sigma_0 \nu_{i0} + \epsilon_{i0ct}. \tag{9}$$

For identification of inside goods' mean utilities, δ_{0ct} and σ_0 are normalized to be equal to zero. ϵ_{i0ct} are also i.i.d type I extreme-value distributed errors like ϵ_{ijct}.

The set of all demand parameters to be estimated is $\Theta_d \equiv \{\beta_j^c, \alpha, \gamma, L, \sigma_p\}$. I define the set of linear and non-linear demand parameters as θ_1 and θ_2, respectively, where $\theta_1 \equiv \{\beta_j^c, \alpha, \gamma\}$ and $\theta_2 \equiv \{L, \sigma_p\}$. As usual in the single discrete-choice model, I assume that a consumer purchases one equivalent-volume unit (a 16 oz jar) of a brand that maximized her utility among all possible options (including outside alternatives) in each time period t. In the next subsection, I provide the evidence that the peanut butter industry is well fit to the assumption of the discrete choice model. Given this assumption and given the non-linear parameters $\{L, \sigma_p\}$, the logit errors ϵ_{i0ct} and ϵ_{ijct} lead the expression of the market share of brand j in city c at time period t to be:

$$\begin{aligned} s_{jct}(p_{ct}, X_{ct}; \Theta_d) &= \int P_{ijct} dP(\epsilon_i^{\beta}) dP(\nu_i^{\alpha}) \\ &= \int \frac{\exp(\delta_{jct} + \mu_{ijct})}{1 + \sum_k \exp(\delta_{kct} + \mu_{ikct})} dP(\epsilon_i^{\beta}) dP(\nu_i^{\alpha}) \end{aligned} \tag{10}$$

where P_{ijct} represents individual choice probabilities.

For the simple logit demand model, I assume the conditional indirect utility of consumer i from choosing brand $j = 1, ..., J$, in city $c = 1, ..., C$, in time period $t = 1, ..., T$ as follows:

$$u_{ijct} = \beta_j + \alpha p_{jct} + X_{jct}\gamma + d_t + d_c + \xi_{jct} + \epsilon_{ijct} \tag{11}$$

where β_j are brand intercepts assumed to be common across individuals and cities, and d_t and d_c are month and city fixed effects, respectively. All other variables are defined as in (1). The simple logit specification will also be estimated using β_j^c and d_t^c instead of β_j, and d_t and d_c to account for different brand strengths across geographic markets.

1.4.2 Single–Unit Purchase Assumption

The single–unit purchase assumption has been criticized as unrealistic in some industries. Especially, in the carbonated soft drink (CSD) industry, Dube (2005) reports that consumers usually purchase multiple units or multiple brands during a single shopping trip. However, the peanut butter category shows single–unit purchase behavior most of the time. In Table 3, I provide the evidence of this finding by analyzing consumer purchasing pattern on the number of units/brands at each shopping trip in two behavioral cities in the IRI dataset. Though these two cities are excluded from the construction of the sample data I use because of the many missing records in the aggregated–level data, the panel data in these cities provide the justification for single-unit purchase assumption when analyzing peanut butter industry demand.

For comparison purposes I also present a multiple-unit purchasing pattern in the CSD category for those two cities. As shown in Tables 3 and 4, the peanut butter category and the CSD category show a clear difference in purchasing behavior. In the peanut butter industry, about 93% of shopping trips are single–brand purchases, and among these purchases, 85% are single–unit purchases. In total, over 79% of shopping trips show single–brand, single–unit purchases. On the other hand, in the soft drink industry, 57% of shopping trips are single–brand purchases, and among these, 61% are single–unit purchases. In total, only 35% of shopping trips show single–brand, single–unit purchases. The discrete choice model with single–unit purchase assumption, therefore, is reasonable to use for estimating the demand in the peanut butter industry because the real demand pattern supports the assumption.

Table 3: **Distribution of number of alternatives and number of units purchased on a shopping trip in the peanut butter category**

	number of units								
number of alternatives	1	2	3	4	5	6	7	8+	Total
1	14,508	2,162	147	117	122	14	3	11	17,084
2	0	926	112	74	51	13	5	5	1,186
3	0	0	21	7	11	6	0	3	48
4	0	0	0	5	3	0	0	2	10
Total	14,508	3,088	280	203	187	33	8	21	18,328

Table 4: **Distribution of number of alternatives and number of units purchased on a shopping trip in the CSD category**

	number of units								
number of alternatives	1	2	3	4	5	6	7	8+	Total
1	33,285	10,681	4,677	2,653	984	1,059	62	847	54,248
2	0	11,894	4,983	2,931	1,319	1,083	371	1013	23,594
3	0	0	4,195	2,327	1,193	891	373	1005	9,984
4	0	0	0	1,543	826	582	312	778	4,041
5	0	0	0	0	446	367	225	550	1,588
6	0	0	0	0	0	167	121	380	668
7	0	0	0	0	0	0	42	239	281
8+	0	0	0	0	0	0	0	182	182
Total	33,285	22,575	13,855	9,454	4,768	4,149	1,506	4,994	94,586

1.4.3 Firm Conduct and Price Counterfactuals

In this section, I use notation m—instead of ct—to denote a city-time period combination, that is equivalent to the definition of a market. Assuming that there are $f = 1, ..., F$ firms and each firm produces subset J_{fm} in market m, the profit of firm f in market m is:

$$\Pi_{fm} = \sum_{j \in J_{fm}} (p_{jm} - mc_{jm}) Y_m s_{jm}(p_m) \tag{12}$$

where mc_{jm} is the constant marginal cost of brand j in market m; $s_{jm}(p_m)$ is the market share of brand j in market m, which is a function of all brands' prices; and Y_m is the total market size. Assuming the static Bertrand–Nash oligopoly model of pricing competition between firms with the pure-strategy static Nash equilibrium assumptions, the first–order conditions of profit maximizing firms are as follows:

$$s_{jm}(p_m) + \sum_{k \in J_{fm}} (p_{km} - mc_{km}) \frac{\partial s_{km}(p_m)}{\partial p_{jm}} = 0 \qquad j = 1, ...J_{fm}, \quad f = 1, ..., F, \quad m = 1, ...M \tag{13}$$

The total number of brands is J. Let us define Ω^{pre} as a J by J pre-merger ownership matrix with element $\Omega^{pre}(k, r)$ equal to one if brand k and r are produced by same firm and zero otherwise. By rearranging the first-order conditions, the pre-merger implied marginal cost in vector expression is:

$$mc_m = p_m - (\Omega^{pre} * \Delta(p_m))^{-1} s_m(p_m) \tag{14}$$

where the J by J matrix of market share response with respect to prices is:

$$\Delta(p_m)_{k,r} = \frac{\partial s_{km}(p_m)}{\partial p_{rm}} \qquad k = 1, ..., J \qquad r = 1, ..., J \tag{15}$$

To perform the merger counterfactuals, it is generally assumed that the marginal costs are fixed at the pre–merger level and that we have the same Bertrand–Nash price equilibrium holds in the post–merger period. Then simply changing the ownership matrix allows simulated prices for the post–merger period to be:

$$\widehat{p_m} = mc_m + (\Omega^{post} * \Delta(\widehat{p_m}))^{-1} s_m(\widehat{p_m}) \tag{16}$$

where mc_m is obtained from (14). The simulated price changes for each brand in each city-time can be constructed by calculating:

$$\Delta p_{jm}^{simulated} = \widehat{p_{jm}} - p_{jm} \tag{17}$$

This measure will be compared to the direct measures of merger effect from reduced form analysis described in the next subsection.

1.4.4 Direct Estimates of Merger Effects

The actual price effects of a merger can be directly estimated using pre–merger and post–merger data. Because the direct estimation needs data from both period of time, this method is primarily used in the retrospective literature on merger studies. For consummated mergers, retrospective merger studies examine how mergers actually affect prices using a reduced form analysis.

To estimate the merger effect, the simplest way would be calculating the difference in prices before and after the merger. However, this simple before–and–after analysis cannot identify what factor has caused the price changes in the post–merger period because there might be other factors besides the merger that have affected the prices, including cost changes, consumer preference changes, or time trends, for example. If input cost changes when a merger occurs, purely comparing before and after prices cannot determine what portion of price changes should be attributed to the merger event.

In this section, I consider two direct estimates of the merger effect. The first one is time-difference estimates that explicitly control for other factors, such as cost changes and time trends in the regression model. To obtain accurate estimates of the merger effect and reasonable coefficients for other covariates, one needs to be cautious when specifying the reduced form regression model. The other type of direct estimate I consider in this section is difference-in-difference estimates that employ a certain group of products to control for other factors.

Time–Difference Estimates The following equation estimates time–difference merger effects:

$$\ln(p_{jct}) = \gamma^0_{jc} + \rho^0_j Cost_t + \theta^0_j X_{jct} + \delta^0_j PostMerger_t + \sum_{n=1}^{11} \tau^0_{jn} T^n_t + \varepsilon^0_{jct} \quad (18)$$

$\ln(p_{jct})$ is the log of brand j's price in city c at time t. γ^0_{jc} are brand-city specific fixed effects. $Cost_t$ represents input costs that affect prices in time t. X_{jct} is the vector of marketing activity measures that affect brands' prices. I include measures of display activity, feature activity, and promotion activity. θ^0_j is a row vector of coefficient that measures each of these marketing mix covariates on brand j's price. $PostMerger_t$ is a dummy variable equal to one if t is the post–merger period. We are interested in the coefficient δ^0_j, and $100 * \delta^0_j$ represents each brand j's percent change in price after the merger in city c. Finally, T^n_t represents month dummy variables to control for seasonal time fixed effects. The regression of model (18) is done separately by brands.

Table 5 shows the time–difference regression results with percent changes of prices due to the merger, the estimates of cost coefficients, and the estimates of marketing mix coefficients for each brand. For the input cost measure, I use peanut prices from two quarters before of time t. I fit the regression using various periods of peanut prices from the spot price to four quarters before time t. All of them except the peanut prices from two quarters before time t give negative coefficients on the input cost term. Given that peanuts are the major ingredient of peanut butter, the price of peanut butter should be positively correlated with peanut prices. The two–quarter prior peanut prices give the most reasonable estimate of cost coefficient, probably because of the pre–purchase contracts between peanut butter manufacturers and peanut suppliers.

Table 5: **Time difference estimates of merger effects**

Variable	Jif	Skippy	Peter Pan	Private Label	Smuckers
Post Merger[1]	0.96	1.50	2.79**	1.45	3.33***
	(0.009)	(0.013)	(0.012)	(0.013)	(0.005)
Cost	0.216***	0.080	0.245***	0.631***	0.031
	(0.079)	(0.128)	(0.085)	(0.119)	(0.041)
Display	−0.254***	−0.424***	−0.217***	−0.339***	−0.178*
	(0.068)	(0.128)	(0.066)	(0.051)	(0.093)
Feature	−0.088***	−0.086***	−0.072***	−0.110***	−0.011
	(0.017)	(0.014)	(0.013)	(0.012)	(0.007)
Promotion	−0.192***	−0.305***	−0.256***	−0.208***	−0.142***
	(0.017)	(0.021)	(0.024)	(0.027)	(0.021)

[1] 100 times of coefficient estimates are reported to express in percent effects. ***, **, and * denote 1%, 5%, and 10% significance levels, respectively. Standard errors in parentheses are clustered by city.

The merger between Jif and Smuckers has increased prices of all brands. The price of Jif and Smuckers, the two brands of the merging firm, increased by 0.96% and 3.33%, respectively. The test that the merger effect on Jif price would be zero cannot be rejected at a 10% significance level. This result can be interpreted that the actual merger effect on Jif may be minimal or lower than 0.96%. The coefficients of the marketing activities all have expected signs in the sense that products put on marketing promotions are usually priced lower during the promotion period. The effect of input cost on brand prices have expected signs, as well, in the sense that, when peanut prices increase (decrease), the price of peanut butter should also increase (decrease).

Difference–in–Differences Estimates Another direct estimate used widely in the literature is difference–in–differences estimate. The difference–in–differences estimator uses a control group to eliminate the effects of other factors on prices, assuming that all factors except the merger event have the same effect on both interested groups and control groups. In this way, we control for the effects of other exogenous changes that may occur at the same time as the merger event. Another critical assumption in difference–in–differences estimation is that the ideal control group should be naive to the merger event or should be minimally affected by the merger. Given those assumptions, I obtain the relative price changes of interested groups relative to the control group.

For example, let us assume that the pre–merger price and post–merger price of the specific product are a and b, respectively. Let us also assume that the pre–merger price and post–merger price of the control group are c and d, respectively. Then the simple before and after changes for the specific product and the control group are $(b-a)$ and $(d-c)$, respectively. If some portion of these changes have been caused by other factors, then a more accurate merger effect would be $(b-a)-(d-c)$, which is the difference in the pre–post price changes between the two groups. For $(b-a)-(d-c)$ to reflect the accurate merger effect on the group of interest, the change $(d-c)$ should be caused only by other factors and not by the merger.

For the difference-in-difference estimates to remove any exogenous changes caused by other factors than the merger event, I follow Weinberg's (2011) style of control group, which is the Private Label brand. All other brands except Private Label are national brands that are distributed nationwide and available in all markets in the sample. The equation for the estimation of difference–in–differences merger effects is given by:

$$\begin{aligned}\ln(p_{jct}) \quad = \quad & \gamma_{jc}^{1} + \rho_{j}^{1} Cost_t + \rho_{j}^{1,B} Cost_t \cdot B_j + \theta_{j}^{1} X_{jct} + \theta_{j}^{1,B} X_{jct} \cdot B_j \\ & + \delta_{j}^{1} PostMerger_t + \delta_{j}^{1,B} PostMerger_t \cdot B_j + \sum_{n=1}^{11} \tau_{jn} T_t^n + \varepsilon_{jct}\end{aligned} \tag{19}$$

The equation is separately estimated with OLS for each pair of national brand and control group. γ_{jc} are brand-city specific fixed effects; B_j is a dummy variable equal to one if product j is not a control group; the term $Cost_t \cdot B_j$ allows for different effects of input cost between a branded product price and the private label product price; the term $X_{jct} \cdot B_j$ allows for different effects of marketing activities between the branded product price and private label product price. Term $\sum_{n=1}^{11} \tau_{jn} T_t^n$ controls for seasonal time effects and I assume these are common to both brand j and the control group. The coefficient of interest, $\delta_j^{1,B}$, measures the relative effect of the merger on brand j's price. Specifically, $100 * \delta_j^{1,B}$ represents brand j's percent changes in prices relative to the control group.

Table 6 shows time–difference and difference–in–differences estimates of price changes for each brand. In the difference–in–differences estimation, the prices of Jif and Smuckers go up by 1.63% and 7.48%, respectively. Compared to the time–difference estimates, the difference–in–difference estimates result in higher estimates of price changes for all national brands. In the results section, these results will be compared to the simulation results.

Table 6: **Direct estimates of merger effects**

	Model	
Brand	**Time-Diff**	**Diff-in-Diff**
Jif	0.96	1.63
	(0.009)	(0.013)
Skippy	1.50	4.45**
	(0.013)	(0.018)
Peter Pan	2.79**	4.22***
	(0.012)	(0.015)
Private Label	1.45	–
	(0.013)	
Smuckers	3.33***	7.48***
	(0.005)	(0.014)

100 times of coefficient estimates are reported to express in percent effects. ***, **, and * denote 1%, 5%, and 10% significance levels, respectively. Standard errors in parentheses are clustered by city.

1.5 Identification and GMM Estimation

1.5.1 Identification

To address potential price endogeneity issues in the demand estimation, which stem from the unobserved (by econometricians) demand shocks (ξ_{jct}) being correlated with prices (p_{jct}), we need a valid set of instruments. The instruments should be correlated with prices but not with the demand shocks. I follow an approach similar to Nevo (2000) and use other regions' average prices and manufacturer input prices (peanut and sugar prices) as instruments. The identification comes from the assumption that city-specific demand shocks are independent across cities. Prices across cities are correlated because the input cost shocks have common effects on prices in all cities. If demand shocks are not correlated among cities, the set of other cities' prices serve as valid instruments. As described in the previous Section 2, because brand strength—the order ranking of each brand's market share— and rivalry structures differ across cities, it is probable that the demand shocks are independent across cities. Using the input prices as instruments follows the logic that changes in costs correlate with prices but not necessarily with demand shocks of a product.

1.5.2 GMM Estimation

Using the set of instruments with all other exogenous covariates in the demand model, I construct moment conditions that the instruments and exogenous covariates are independent of demand shocks. I estimate demand parameters using the generalized method of moments (GMM) procedure. First, the population moment condition is:

$$G(\theta) = E[\xi_{jct} \otimes Z_{jct}] = 0 \quad \text{at the true } \theta$$

where Z_{jct} is the vector of all instruments and exogenous covariates of brand j in city c in time period t.

The sample analog to this population moment is:

$$\begin{aligned} G_N(\theta) &= \frac{1}{N}\sum_j\sum_c\sum_t [\xi_{jct} \otimes Z_{jct}] \\ &= \frac{1}{N}\sum_j\sum_c\sum_t [(s_{jct} - s_{jct}(p_{ct}, X_{ct}; \Theta_d)) \otimes Z_{jct}] \end{aligned} \tag{20}$$

where s_{jct} is the observed market share. The objective function to be minimized has the quadratic form:

$$G_N(\theta)' W G_N(\theta) \tag{21}$$

where W is the weight matrix. Any matrix can be used as the weight matrix. For the most efficient estimates, the inverse of the approximation of variance matrix of the moments is used. The detailed procedure is as follows:

1) Take an initial guess of the non-linear parameters θ_2. Take an initial guess of δ_{jct}.

2) Take $J + 1$ number of normal draws for each of NS number of simulated consumers in market m. I choose $NS = 100$.

3) Given the initial θ_2 and initial δ_{jct}, calculate the individual choice probability for each simulated consumer i using the analytic form:

$$P_{ijct} = \frac{\exp(\delta_{jct} + \mu_{ijct}(\theta_2))}{1 + \sum_k \exp(\delta_{kct} + \mu_{ikct}(\theta_2))}$$

4) Calculate the average value of P_{ijct} across simulated consumers:

$$s_{jct}(p_{ct}, X_{ct}; \Theta_d) = \frac{1}{NS} \sum_{i=1}^{NS} P_{ijct}$$

5) Find the fixed point of δ_{jct} using contraction mapping $\widetilde{\delta_{jct}} = \delta_{jct} + s_{jct} - s_{jct}(p_{ct}, X_{ct}; \Theta_d)$. For each estimation round of θ_2, I obtain $\delta_{jct}(\theta_2)$.

6) Use the instruments and regress $\delta_{jct}(\theta_2)$ on covariates as in (8). The residual from this regression is $\xi_{jct}(\theta_2)$, which will be used to form the sample moment condition (20). The set of linear parameters θ_1 is acquired in this step.

7) Search over θ_2 that minimize the objective function (21). θ_1 is re–estimated every round.

1.5.3 Technical Issue in the Simulation

After the demand estimation, we perform the merger simulation. One technical issue regarding simulation of counterfactual prices is whether to use demand shocks that are inverted out from the pre–merger demand estimation or to let them be equal to zero. If we use the pre–merger demand shocks in the simulation we implicitely assume that unobserved preferences that might be included in the demand shock terms are fixed at the pre–merger level. If preferences not captured by demand parameters change, then post-merger simulated prices would have biases because the post-merger demand would not be reflected appropriately.

Theoretically, GMM residuals are mean zero at the true parameter value. For the sample analog, the GMM procedure forces the residuals to be sample mean zero. Therefore, given that the global minimum is found, whether using residuals from the demand estimation or setting them at zero in the simulation should not change the results much. A more flexible demand specification is expected to produce less variable counterfactual results for the selection of demand shocks. I find that, in the demand specification where brand intercepts are common across cities, setting residuals to zero in the simulation often produces different results from the simulation in which residuals are set to the pre–merger level. In case where more restriction is imposed on the demand model, the simulation results seem to be influenced quite strongly by the choice of demand shocks.

In section 1.6, I report results on the predicted prices for each case of unobserved demand shocks: at the pre–merger level and at zero. The more flexible model shows less vulnerability to the choice of unobserved demand shocks in the simulation.

1.6 Results and Discussion

1.6.1 Demand Estimation Results

The estimates of the demand parameters are presented in Table 7. I estimate models of OLS—ordinary least squares—, simple logit, simple logit with different brand intercepts across cities, nested logit, random coefficient (mixed logit), and random coefficient with different brand intercepts across cities. The last one is the most flexible specification. Across specifications, all the estimated coefficients on price, display, and the number of skus—stock–keeping unit—are significant and have expected signs. With price coefficients in Column (2) through Column (6) all larger than that of OLS, price endogeneity seems to have been appropriately addressed by the instruments.

The own–price and cross–price elasticities are presented in Table 8. All values of own–price and cross–price elasticities have expected signs. Smuckers brand has the highest own–price elasticity in all specifications, meaning that the brand has the most elastic demand among the competitors. Considering the cross–price elasticities, Jif is the closest substitute for Smuckers. For the two random coefficient specifications in Table 8, element (j, k) is $\epsilon_{jk} = \frac{\partial s_j}{\partial p_k}\frac{p_k}{s_j}$.

The first column of the elasticity table of the random coefficient with common brand intercept model (Table 8-(2)) represents what percent of market share of each brand would change if Jif's price increases by one percent. With respect to Jif's price change, Skippy's elasticity is 0.023, Peter Pan's is 0.083, Private Label's is 0.226, and Smuckers' is 0.871. Smuckers brand's demand increases most among competitor brands when Jif's price increases, giving the merger incentive for Jif and Smuckers. In the fifth column of Table 8-(2), Jif's elasticity with respect to Smuckers' price change is 0.100, which is the highest among competitor brands. Jif's and Smuckers' cross–price elasticities are also nearly the highest in the random coefficient with the different brand intercept model (Table 8-(3)).

Considering the demand parameter estimates in Table 7, display activities make people buy more of the product being displayed. People prefer brands with more skus being carried. Caution is advised, however, when interpreting the coefficients for the number of skus. The decisions on how many variants to carry depend on the retailers as well as manufacturers.

Demand for peanut butter is fairly stable over time and may have been built to be persistent over a long period of time. Hence, high demand might have caused retailers and manufacturers carry more variants to satisfy the high demand. Conversely, a large shelf presence (i.e., large number of skus) might have led people to buy more. However, an estimation without the number of skus as a covariate did not appreciably change other demand parameters. Each brand intercept is interpreted as a mean intrinsic value attached to the brand. The brand intercepts in Columns (3) and (6), where I set the brand intercepts to be different across cities, represent median values across cities. At the market–by–market level, the order of brand intercepts agrees with the order of pre–merger market shares, except for Smuckers. Considering that Smuckers brand represents higher quality segments, a high value of brand intercept for Smuckers seems reasonable.

The estimated Cholesky parameters for model (5) in Table 7 are presented in Table 9. The variance–covariance matrix of brand correlation is calculated as LL' and is presented in Table 10. Jif and Smuckers' own variances are estimated at 6.60 and 15.79, respectively. These values are relatively larger than those of other brands.

1.6.2 Comparing Direct Estimates of Price Changes and Predicted Price Changes

Table 11 shows a comparison of directly estimated price changes and simulated price changes of the merged firm's two brands, Jif and Smuckers. The Difference–in–Difference price changes are higher than the Time–Difference price changes for both Jif and Smuckers. For

Table 7: **Parameter Estimates of Demand Models**

	(1) OLS	(2) Logit	(3) Logit	(4) Nested Logit	(5) RC	(6) RC
Price	-1.368***	-3.248***	-2.283***	-1.675***	-2.913***	-1.509***
	(0.029)	(0.274)	(0.132)	(0.086)	(0.267)	(0.045)
S.D. of Price	–	–	–	–	0.208**	0.009
					(0.103)	(0.028)
Nest parameter	–	–	–	0.775***	–	–
				(0.081)		
Jif	-2.458***	2.108***	0.069	-1.662***	-4.860***	-1.771***
	(0.080)	(0.668)	(0.310)	(0.207)	(0.650)	(0.152)
Skippy	-2.840***	1.620***	0.353	-2.044***	-2.593***	-2.153***
	(0.078)	(0.652)	(0.294)	(0.201)	(0.636)	(0.184)
Peter Pan	-2.894***	1.491**	-0.626**	-2.101***	-1.064*	-2.358***
	(0.077)	(0.641)	(0.265)	(0.197)	(0.625)	(0.115)
Private Label	-2.905***	0.950**	-0.810***	-2.194***	-2.658***	-2.349***
	(0.070)	(0.564)	(0.262)	(0.173)	(0.550)	(0.137)
Smuckers	-1.896***	3.815***	0.668	-0.902***	-6.221***	-1.518***
	(0.095)	(0.834)	(0.367)	(0.258)	(0.811)	(0.141)
Display	2.273***	0.411*	0.681***	1.512***	2.436***	1.517***
	(0.091)	(0.289)	(0.152)	(0.107)	(0.310)	(0.102)
Number of SKUs	0.315***	0.281***	0.203***	0.276***	0.398***	0.239***
	(0.002)	(0.005)	(0.011)	(0.003)	(0.006)	(0.011)
Different brand intercepts across cities			Yes			Yes

1. Standard errors in parenthesis
2. *, **, and *** are 10%, 5%, and 1% significance, respectively.

Table 8: **Own- and Cross-Price Elasticities**

Table 8-(1): Logit

	Common Brand Intercepts		Different Brand Intercepts	
	Own-price Elasticity	Cross-price Elasticity	Own-price Elasticity	Cross-price Elasticity
Jif	-5.777	0.349	-4.061	0.245
Skippy	-5.902	0.209	-4.149	0.147
Peter Pan	-5.731	0.135	-4.028	0.095
Private Label	-4.653	0.223	-3.271	0.157
Smuckers	-8.567	0.042	-6.022	0.029

Table 8-(2): Random Coefficient (Common Brand Intercepts)

	Jif	Skippy	Peter Pan	Private Label	Smuckers
Jif	-2.316	0.011	0.028	0.142	0.100
Skippy	0.023	-4.090	0.024	0.099	0.001
Peter Pan	0.083	0.023	-4.750	0.076	0.024
Private Label	0.226	0.080	0.042	-3.147	0.013
Smuckers	0.871	0.002	0.070	0.068	-6.068

Table 8-(3): Random Coefficient (Different Brand Intercepts)

	Jif	Skippy	Peter Pan	Private Label	Smuckers
Jif	-2.682	0.097	0.057	0.103	0.019
Skippy	0.161	-2.739	0.056	0.104	0.020
Peter Pan	0.156	0.045	-2.624	0.096	0.016
Private Label	0.162	0.098	0.056	-2.159	0.019
Smuckers	0.163	0.099	0.052	0.102	-3.977

Table 9: **Lower Cholesky Estimates**

	coefficient	s.e.
σ_{11}	2.570^{**}	(0.294)
σ_{22}	-0.144	(2.140)
σ_{33}	-0.457	(0.967)
σ_{44}	-0.131	(1.622)
σ_{55}	-1.783^{*}	(0.971)
σ_{21}	-1.380^{**}	(0.543)
σ_{31}	0.988^{*}	(0.558)
σ_{32}	1.566^{**}	(0.583)
σ_{41}	0.265	(0.651)
σ_{42}	-0.792	(1.160)
σ_{43}	0.489	(2.171)
σ_{51}	1.888^{**}	(0.510)
σ_{52}	2.269^{**}	(0.807)
σ_{53}	-1.9730^{**}	(0.915)
σ_{54}	0.047	(0.756)

* and ** denote 10% and 5% significance, respectively.

Table 10: **Variance-Covariance of Brand Preference**

	Jif	Skippy	Peter Pan	Private Label	Smuckers
Jif	6.605	-3.545	2.540	0.680	4.853
Skippy	-3.545	1.924	-1.588	-0.251	-2.931
Peter Pan	2.540	-1.588	3.639	-1.203	6.321
Private Label	0.680	-0.251	-1.203	0.954	-2.268
Smuckers	4.853	-2.931	6.321	-2.268	15.788

Jif, time–difference and difference–in–differences merger effects are 0.96% and 1.63%, respectively. For Smuckers, time–difference and difference–in–difference merger effects are 3.33% and 7.48%, respectively. They show, however, the same asymmetric pattern[13] of price changes between Jif and Smuckers. This indicates the actual merger effects depend on policy makers. Compared to both direct estimates, the simulated price changes of Jif and Smuckers are lower for all specifications except Column (6), where I let brand intercepts be common across cities in random coefficient model. In the sense that ignoring different brand strengths across different regions may mislead the demand parameters in a flexible setting, Column (6) can be removed from consideration.

Comparing Column (1), (3), (5), and (7), the demand model gets more flexible. Column (7) is the most flexible demand model in which brand intercepts are set to be different across cities and the full variance–covariance structure between brand preferences are included. Column (7) model predicts Jif's price change to be 0.25% and Smuckers' price change to be 1.99%. Although the differences in the simulation predictions across specifications are not significantly large, moving from the simple model to the flexible model, the simulated price changes are larger and closer to the actual price changes.

Basically, in the consumer packaged goods industry, when we estimate a structural demand model using logit errors, the share of outside alternatives is usually very high (around 70–80%). Thus, substitution patterns within inside goods are likely to be underestimated. Underestimation of substitution patterns affects the simulation, causing it to underpredict

[13] Jif's actual price changes are estimated to be much lower than Smuckers' price changes.

the post–merger price changes.

To resolve the underprediction problem of the logit model, I estimated more flexible settings from the nested logit model to the random coefficient model with the brand preference correlation structure. The results in Table 11 offer evidence that the logit model typically underpredicts the merger effects and that a more flexible demand model could help to resolve this problem. In the future, it would be interesting to investigate whether stimating a full covariance structure using consumer–level data would result in closer predictions to the actual price changes.

I also conduct experiments concerning whether setting the unobserved demand shock terms at the pre–merger level or at zero affects the simulation results. This is the technical issue in the simulation pointed out in subsection 1.5.3. Columns (2), (4), and (8) in Table 11 show the results. Except for Jif's predicted price change in Column (2), it seems that choice of unobserved demand shocks does not greatly affect the simulation results. Especially when brand intercepts are set to be different across cities, there is almost no significant difference between the results with zero demand shocks and the results with pre–merger demand shocks, whether a logit model or a more flexible random coefficient model is used. When more restriction is imposed on the demand model in terms of the fixed effects of brand preference, the simulation results seem to be influenced quite strongly by the choice of demand shocks. On the other hand, the more flexible model shows less vulnerability to the choice of unobserved demand shocks in the simulation.

In all specifications in Table 11, the asymmetric pattern of price changes between two brands are well predicted. They indicate that Jif's price change is predicted to be small while Smucker's is predicted to be larger in the market. Because this result agrees with the pattern of direct estimates, some confidence exists concerning the structural demand models considered in this chapter explaining well the substitution patterns and the demand in the U.S. peanut butter industry, given that I have successfully estimated the actual merger effects from the direct estimation. Another interesting pattern is found in Figure 2. It depicts Smuckers' predicted price changes and Jif's pre–merger market shares across cities. Smuckers' price changes are predicted to be closely and positively correlated (correlation coefficient 0.73) to Jif's pre–merger market shares. In other words, in markets where Jif's pre–merger market shares are higher, Smuckers' price is predicted to increase more. It is intuitive that, as an acquiring firm, the J.M Smucker Company would want to increase Smuckers' price more in markets where buying–in brands are stronger to exploit the combined market power.

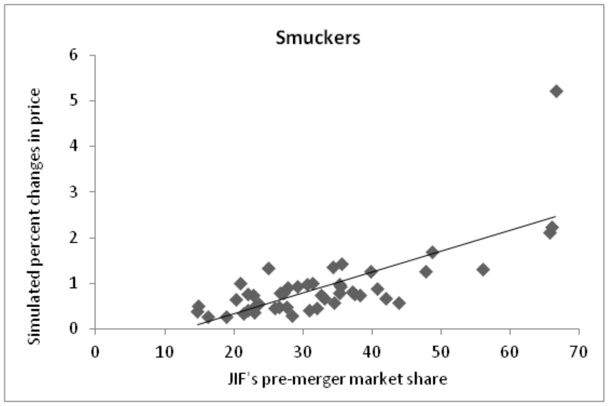
1.7 Conclusion

Chapter 1 examines how the results of structural merger simulation are compared to direct measures of merger effects on a consummated merger between two U.S. national peanut butter manufacturers. In the literature, structural merger simulation has been reported to give inaccurate and usually lower prediction of price changes than the actual changes. The

Table 11: **Actual and Predicted Percent Price Changes for Merging Firm's Brands**

			Logit				Nested Logit	Random Coefficient with full covariance		
Brand intercepts			Common	Common	Different	Different	Common	Common	Different	Different
Demand Shock			Pre-merger	Zero	Pre-merger	Zero	Pre-merger	Pre-merger	Pre-merger	Zero
Brands	Time-diff	Diff-in-diff	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Jif	0.96	1.63	0.12	0.43	0.17	0.17	0.23	1.83	0.25	0.21
Smuckers	3.33	7.48	0.92	1.12	1.31	1.31	1.79	8.64	1.99	1.97

Figure 2: Smuckers' simulated percent price changes vs. Jif's pre-merger market shares (correlation: 0.73)



problem is partially due to the logit demand model in which substitution patterns within inside goods are not captured well, especially in the consumer packaged goods industry.

After Peters (2006) and Weinberg (2011) noted the failure of the structural model in predicting accurate merger effects, I built on Weinberg and extended the literature by estimating a more flexible demand model. To determine if this more flexible demand model can resolve the underprediction problem, I estimated a random coefficient demand model allowing brand intercepts to be different across cities and also allowing the full variance-covariance structure of brand preferences. To my knowledge, this model which considers demand with a full variance–covariance structure is the first attempt in the literature to evaluate merger simulations. I find that a more flexible demand model slightly improves the simulation results, though the difference is not significantly large than other simple demand models. Within each demand model, addressing different brand strengths across cities slightly improves the simulated results. Moving from a simple demand model to a flexible model, the predicted price changes become closer to the actual merger effects. Overall, the most flexible setting estimated in this chapter gives the closest prediction to the actual price changes. Although improved somewhat, the underprediction problem still exists with the flexible demand model.

There are three main possibilities explaining why merger simulation fails to result in an accurate prediction of post–merger price changes. First, a demand model might be wrong, so it might not capture the real substitution patterns. In Chapter 2, I consider the Almost Ideal Demand System as an alternative demand model and examine how it compare to the structural model in simulation.

Second, the equilibrium assumption of the pricing game might be wrong. One extension to address this issue would be to consider retailer pricing in the model and determine how it affects counterfactual results, as in Villas-Boas (2007b). In fact, scanner data for consumer packaged goods industries are retail–level data, so the prices involve a double margin structure, that is, a manufacturer margin and a retailer margin. Thus, pricing decisions not only depend on manufacturers' decisions but vary according to retailers' decisions as well. In Chapter 2, I also consider this fact in an alternative supply model and determine how the merger simulation results differ.

Last, several assumptions in the merger simulation can lead to inaccurate predictions. In Chapter 3, I run counterfactuals to determine the effect of each assumption on the simulated price changes, as in Peters (2006). Future research could address the underlying motive of the merger by estimating marginal costs by directly exploiting market structure heterogeneity.

The economies of scale effects would be greater in markets where buying-in products are strong. Finally, considering the results shown in Figure 2, it would be interesting to examine the effect of initial advantage on post–merger pricing power in a vertical differentiation context.

2 Chapter 2. Evaluation of Merger Simulation with an Alternative Demand and Supply Model

2.1 Introduction

In Chapter 1, I examined whether a flexible demand setting could resolve the typical underprediction problem in merger simulation which uses a structural demand model with logit errors. To address a real substitution pattern, I explicitly considered the correlation structure of brand preferences and addressed different brand strengths across cities. Given the results in Chapter 1 that the underprediction problem still exists with the flexible structural demand model, in this chapter, I estimate an alternative demand model and investigate whether another functional form of the demand model gives more accurate predictions of post–merger price changes in the U.S. peanut butter industry.

As an alternative demand model to the structural one, I consider the Almost Ideal Demand System (AIDS). The AIDS model has been widely used by researchers as an alternative model of structural demand since Deaton and Muellbauer (1980). The model is based on a reduced form approach. Hadeishi and Schmidt (2004) use a three–stage AIDS model in analyzing the merger between Jif and Smuckers. They include jelly data in the analysis to study the merger from the perspective of a merger between complementary goods companies. Peters (2008) estimates two– and three–stage AIDS models for the mergers in the motor oil industry and the breakfast syrup industry, respectively. More studies that use the AIDS model based on a multi–stage budgeting approach include Hausman and Leonard (2002) and Chaudhuri, Goldberg, and Jia (2006).

In this chapter, I estimate a two–stage AIDS model using the same data I analyzed in Chapter 1. For the firm model, I maintain the Bertrand–Nash pricing game between manufacturers. The brief result is that the AIDS model gives very different simulation results from those of the structural model. Jif's predicted price change is almost zero while Smuckers' predicted price change is much higher. The prediction on Smuckers brand closely resembles the difference-in-difference estimates of the merger effect.

Another attempt studied in this chapter is the inclusion of retailer–side pricing in the model together with manufacturer pricing. Because scanner data in the consumer packaged goods industries are at the retail level, the price involves retailer margins. Depending on retailer–manufacturer interactions on pricing, post-merger prices could be affected. If this feature is not addressed appropriately, the merger simulation could result in wrong predictions.

To include the retailer pricing in the model, I follow Villas–Boas (2007a) and Villas–Boas (2007b). Villas–Boas (2007b) reports that ignoring the retailer side in the model underestimates the merger effects in the German coffee industry. Villas–Boas (2007b) does not evaluate the performance of the simulation, though.

I examine whether considering retailer–side pricing in the model would produce more accurate predictions closer to the actual merger effects of the merger between Jif and Smuckers in the U.S. peanut butter industry. In the results section, explicitly accounting for the retailer pricing in the model gives higher predicted price changes for both Jif and Smuckers, suggesting that the inclusion of retailer–side pricing partially resolves the underprediction problem of logit or mixed–logit structural demand model.

In the next section I describe the AIDS model in detail. Section 3 presents the model with retailer pricing. Section 4 reports empirical results. Conclusions are presented in Section 5.

2.2 Alternative Demand Model

2.2.1 AIDS Demand Model

To estimate the peanut butter industry demand, I use a two–level AIDS demand system, based on the approach that consumers make decisions at two different levels. At the top level, they first decide how much of their budget to spend on the peanut butter category. Then at the bottom level, consumers choose which brand or product to purchase within the budget set at the top level.

Because I define the product in this study as an aggregated one across package sizes, fat content, textures, and process—for example, Jif reduced fat creamy 16oz jar is the same product as Jif regular fat crunchy 28oz jar—a two–stage budgeting approach is appropriate for the analysis. If the data were disaggregated further at the sub–segment level, such as fat content or texture, then a multiple–stage demand beyond two levels should be used. For simplicity and for the purpose of comparing the results to the structural demand model in Chapter 1, I use a two–level demand system and the same data structure as in Chapter 1.

Examples using the three–level AIDS demand model can be found in Hausman, Leonard, and Zona (1994) and Hausman and Leonard (2002).

Top level In the top–level model, it is assumed that consumers first assign their budget to the category—the total category demand—depends on the Stone price index. They choose how much to spend on the peanut butter before they make choices about which brand to buy. For each city c and week t, the top–level demand specification is as follows:

$$\log(\frac{Y_{ct}}{P_{ct}}) = \mu_c + \theta \log(P_{ct}) + \sum_{m=1}^{11} M_m D_t^m + e_{ct} \tag{22}$$

where Y_{ct} is the total category sales in city c in week t. P_{ct} is the Stone price index, which is the sum of the weighted average of each brand's price in the category. I use fixed weights in the price index, which are the average shares of each brand's revenue across the time period in the category sales. Specifically, the weights $w_{jc} = \frac{1}{T}\sum_{t=1}^{T} s_{jct}$ and $s_{jct} = \frac{p_{jct}q_{jct}}{Y_{ct}}$. Then, the Stone price index is given by:

$$P_{ct} = \prod_{j=1}^{J} p_{jct}^{w_{jc}} \tag{23}$$

On the right hand side of (22), μ_c are city-specific fixed effects to control for the regional effects and D_t^m are month dummies to control for the seasonal effects on the real expenditure. The coefficient of interest at the top–level demand is θ, which measures the effect of the category price index on the category-wide real spending.

Bottom level Bottom–level demand is modeled as follows:

$$s_{jct} = \alpha_{jc} + \beta_j \log(\frac{Y_{ct}}{P_{ct}}) + \sum_{k=1}^{J} \gamma_{jk} \log(p_{kct}) + \sum_{m=1}^{11} M_{jm} D_t^m + \eta_{jct} \tag{24}$$

At the bottom level, each product's demand depends on total category demand and its own price and the price of all other products competing in the same market. s_{jct} is brand j's revenue share of category sales in city c in week t; α_{jc} controls for city-brand specific fixed effect; Y_{ct} is the total category sales; P_{ct} is the Stone price index; p_{kct} is product k's price in city c in week t. D_t^m represents month dummies. We are interested in the coefficients β_j and $\gamma_{jk.}$ They measure the effect of category demand on product i's budget share and the effect of each product's price on product j's demand, respectively.

In each market (city-week combination), the revenue share of each product must add up to 1 by its definition. Hence, when estimating the bottom-level demand system, one brand's share equation should be dropped to make the system just–identifiable. Without loss of generality, product J's equation is dropped in the estimation. β_J and $(\gamma_{J1,}\gamma_{J2}, ..., \gamma_{JJ})$ are then calculated from the restrictions:

$$\sum_{j=1}^{J} \beta_j = 0 \tag{25}$$

and

$$\sum_{j=1}^{J} \gamma_{ji} = 0 \quad \text{for each } i = 1, ...J \tag{26}$$

2.2.2 Identification

If prices are set by manufacturers[14] after they observe demand, the prices should be correlated with the demand shock at both the top and bottom levels, causing the endogeneity problem—OLS estimators of θ, β_j,and γ_{jk} are going to be inconsistent—in estimating the AIDS model above. Each market needs $J+1$ instruments—J for the price of each product and 1 for the segment–wide real expenditure.

Typically, researchers use other regions' average prices as instruments under the assumptions that (1) marginal costs are correlated across regions because of the common input price shocks, and (2) demand shocks are independent across regions. If advertising or social networking, which may cause different regions' demands to be correlated, can be ignored, then these instruments are reasonable.

The validity of this kind of instrument is supported further by the fact that, in the consumer packaged goods industry, brands' market shares show high levels of heterogeneity across different regions (Bronnenberg, Dahr, and Dube 2009). This phenomenon partially stems from the long history of consumer preferences evolving differently in each region. The other approach besides using instruments is based on the assumption that retailers[15] set prices before demand is realized, in wich case, prices are regarded as exogenous, and OLS estimators are consistent. I take the latter approach and use OLS estimates of demand parameters for AIDS in this chapter.

[14] Here we assume omission of retailer's pricing.

[15] Although I do not model retailers' pricing in the first part of Chapter 2, they are the prices I observe.

2.2.3 Firm Behavior: Bertrand Nash

Because the main purpose of the first part of Chapter 2 is to compare the merger simulation results of the alternative demand model, I maintain the Bertrand–Nash assumption of firms' pricing game as in Chapter 1. Assuming a strictly positive pure Nash price equilibrium, the first–order condition for each product i is:

$$s_{ic}(p_c) + \sum_{j \in J_f} (p_{jc} - mc_{jc}) \frac{\partial s_{jc}(p_c)}{\partial p_{ic}} = 0 \qquad i = 1, ...J_f, \quad f = 1, ..., F \tag{27}$$

Each term in (27) is averaged across time periods - for example, s_{ic} is product i's average revenue share across time periods in city c. I use the uncompensated demand elasticity formula for the two–level AIDS model to obtain an analytic expression of the term $\frac{\partial s_c(p_{ct})}{\partial p_c}$ in (27). The elasticity of product j's demand with respect to product i's price in city c is given by:

$$\varepsilon_{jic} = -1(j = i) + (1 + \theta)s_{ic} + \frac{1}{s_{jc}}(\beta_j \theta s_{ic} + \gamma_{ji}) \tag{28}$$

Equation (27) can be rewritten as:

$$s_{ic}(p_c) + \sum_{j \in J_f} \frac{(p_{jc} - mc_{jc})}{p_{ic}} \epsilon_{jic} s_{jc}(p_c) = 0 \qquad i = 1, ...J_f, \quad f = 1, ..., F \tag{29}$$

If I express the first–order condition in vector notation, J by 1 vector of marginal costs in city c is:

$$mc_c = p_c - (\Omega^{pre} * \Delta(p_c))^{-1} s_c(p_c) \tag{30}$$

where $\Delta(p_c)_{j,i} = \frac{\epsilon_{jic} s_{jc}(p_c)}{p_{ic}}$. Assume that marginal costs are fixed at the pre–merger level and that we have the same Bertrand–Nash price equilibrium holds in the post–merger period. Then, counterfactual prices for the post–merger period are:

$$\widehat{p_c} = mc_c + (\Omega^{post} * \Delta(\widehat{p_c}))^{-1} s_c(\widehat{p_c}) \tag{31}$$

Solving for J first-order equations to obtain marginal costs and running the simulation are done separately for each city c. The demand parameter estimates, the own–price and cross–price elasticities, and the results of simulated price changes are reported in Section 4.

2.3 Alternative Supply Model

Because the supermarket scanner dataset is available only at the retailer level, researchers assume that demand curves estimated using retailer data should be the same as those for manufacturers. In this section I consider the retailer–side pricing and the relationship between retailers and manufacturers in pricing.

2.3.1 Logit Demand Model

Following the discrete choice model, I assume that a consumer chooses one product among J_m number of products available in market m. Market m is defined by a city (denoted by c)–and–time (denoted by t) combination. And the product is defined by manufacturer's brand–and–retailer combination. Hence, the same brand sold by a different retailer chains is regarded as a different product. In the model when retailer pricing is explicitly considered, the number of products available in a market is typically large because there are multiple retail chains that carry the same manufacturers' brands. The large number of products causes a dimensionality problem, especially when one tries to estimate the correlation structure of consumer preferences among products.

For this reason, I estimate a simple logit demand model here. Given that (1) I am interested in how inclusion of retailer pricing in a supply model gives different post–merger price predictions and, (2) as indicated in Chapter 1, considering correlated preferences structure does not result in very different post–merger predictions compared to the logit demand model, I continue to use the logit demand model in this section. The indirect utility of consumer i purchasing product j in city c in time period t is

$$u_{ijct} = \beta_{jc} + \delta T_t + \gamma X_{jct} + \alpha p_{jct} + \xi_{jct} + \epsilon_{ijct} \tag{32}$$

where β_{jc} represents product–city specific fixed effect that captures a time invariant consumer preference for product j in city c. Variable T_t is a time trend, and δ measures time effects on the consumer utility. X_{jct} is observed product characteristics of product j in city c in time t. I include the average display level of product j in city c in time t. Because a product has been aggregated from the UPC–level to brand–level, the average value across UPCs for the same brand is calculated and included in the model to obtain a proxy measure of the marketing activity, such as display. p_{jct} is product j's price in city c in time t. ξ_{jct} represents unobserved product characteristics, and ϵ_{ijct} indicates idiosyncratic errors.

Under the assumptions that (1) prices are correlated among different regions because of common input cost shocks and (2) demand shocks across regions are uncorrelated, I instrument the price p_{jct} using peanut prices and the average prices for the same manufacturer's brand in other regions during the same time period.

2.3.2 Considering Retailer Pricing in a Supply Model

I estimate several models of retailer–manufacturer relationships in this section. Among the supply models in Villas-Boas (2007a), I focus on three of them and investigate how these models predict price changes differently and whether they render more accurate merger simulation results compared to the conventional firm model. In the following, I present the double marginalization model, the model of zero retail margins, and the model of retailer collusion.

Double Marginalization Model (Linear Pricing or Stackelberg) M number of markets are denoted as $m = 1, ..., M$. Each market m has R_m number of retailers competing with each other. There are $f = 1, ..., F$ manufacturers, including the aggregated Private Label brand. For ease of notation, I denote a market as m instead of (c, t). Let S_m^f and S_m^r be the set of products that each manufacturer produces or distributes and the set of products that each retailer carries in market m, respectively.

It is assumed, in the linear pricing model, that manufacturers first set their product prices and then retailers follow and set their own prices. The product is defined by the brand–retailer combination. Jif creamy at retailer A, for example, is regarded as a different product from Jif creamy at retailer B. Retailer r's profit in market m is given by:

$$\Pi_m^r = \sum_{j \in S_m^r} (p_{jm} - p_{jm}^f - mc_{jm}^r) s_{jm}(p_m) \qquad \text{for } r = 1, ..., R_m \text{ and } m = 1, ..., M \tag{33}$$

where p_{jm} is the retail price of the product (brand-retail combination) j in market m, p_{jm}^f is the manufacturer's (wholesale) price of product j in market m, mc_{jm}^r is the retailer r's marginal cost of product j in market m, and $s_{jm}(p_m)$ is the market share of product j in market m, which is the function of all product prices in the market. Assuming a pure-strategy Nash equilibrium in retail prices, the retailers' implied price-cost margins can be acquired from the first-order conditions as follows:

$$s_{jm}(p_m) + \sum_{k \in S_m^r} (p_{km} - p_{km}^f - mc_{km}^r) \frac{\partial s_{km}}{\partial p_{jm}} = 0 \qquad \text{for } \forall j \in S_m^r, \quad r = 1, ..., R_m \text{ and } m = 1, ..., M \tag{34}$$

The matrix notation of implied price-cost margins of retailers is:

$$p_m - p_m^f - mc_m^r = -(\Omega_m^r * \Delta_m^r)^{-1} s_m(p_m) \tag{35}$$

Each vector in (35) is J by 1 dimension, where J is the total number of products (brand-retailer) sold in market m. Ω_m^r is the retailer owenership matrix whose element $\Omega_m^r(x, y)$ is equal to 1 if both products x and y are sold by the same retailer and zero otherwise in market m. Δ_m^r is the retail price–response matrix, whose elements are the derivatives of market shares of all products in market m with respect to all retail prices in market m. Specifically, the element $\Delta_m^r(x,y) = \frac{\partial s_{ym}}{\partial p_{xm}}$. Hereafter, the notation $*$ denotes element by element multiplication of matrices.

Given the optimal pricing of retailers, manufacturers set their prices to maximize the profits[16] as follows:

$$\Pi_m^f = \sum_{j \in S_m^f} (p_{jm}^f - mc_{jm}^f) s_{jm}(p_m(p_m^f)) \qquad \text{for } f = 1, ..., F \text{ and } m = 1, ..., M \tag{36}$$

S_m^f is the set of products sold by manufacturer f in market m. I am assuming that manufacturers make decisions on the optimal prices separately by markets. This assumption should be violated if manufacturers coordinate their price decisions nationally. However, given the industry convention that different marketing strategies are used for different regions and prices are re-optimized as time passes, the assumption seems to be reasonable. Thus, the implied price–cost margins of manufacturers are written as:

$$p_m^f - mc_m^f = -(\Omega_m^f * \Delta_m^f)^{-1} s_m(p_m) \tag{37}$$

where Ω_m^f is the ownership matrix of manufacturers analogous to the retailers' ownership

[16] Although, in the data, I observe that some retail chains operate in multiple cities, I assume manufacturers regard them as different retailers and set different wholesale prices. Otherwise, solving the first–order conditions of manufacturer profit function would be very complicated because of the cross–city effect of optimal pricing of products distributed to the same chain across cities. It would be interesting in future research to see how the results of a merger simulation could be affected by taking into account this feature.

matrix. Δ_m^f is the matrix of derivatives of market shares of all products in market m with respect to all manufacturers' wholesale prices in market m, with element $\Delta_m^f(x,y) = \frac{\partial s_{ym}}{\partial p_{xm}^f}$. I do not observe p^f, the wholesale prices.

Hence, to compute the wholesale price response matrix Δ_m^f, it must be rewritten in other terms that can be observed or computed. As in Villas-Boas (2007a), the wholesale price response matrix is the product of Δ_m^p and Δ_m^r. Δ_m^p is the pass–through matrix—the matrix of derivatives of all retail prices with respect to all wholesale prices—, and Δ_m^r is the retail price response matrix:

$$\Delta_m^f = \Delta_m^r \Delta_m^p \tag{38}$$

The element $\Delta_m^p(x,y) = \frac{\partial p_{ym}}{\partial p_{xm}^f}$.

Now we only need to get the expression of the pass–through matrix Δ_m^p. First we totally–differentiate the retailers' first–order condition (equation (34)) with respect to all retail prices $(p_{1m}, p_{2m}, ..., p_{J_m,m})$ and a wholesale price p_{wm}^f (for $w = 1, ..., J_m$).

$$\begin{aligned} 0 \;=\; & \sum_{l=1}^{J_m} \left[\frac{\partial s_{jm}}{\partial p_{lm}} + \sum_{k=1}^{J_m} \left(\Omega_m^r(k,j) \frac{\partial^2 s_{km}}{\partial p_{jm} \partial p_{lm}} (p_{km} - p_{km}^f - mc_{km}^r) \right) + \Omega_m^r(l,j) \frac{\partial s_{lm}}{\partial p_{jm}} \right] dp_{lm} \\ & - \Omega_m^r(w,j) \frac{\partial s_{wm}}{\partial p_{jm}} dp_{wm}^f \end{aligned} \tag{39}$$

Let the expression in the bracket in the first part of the equation (39) be the (j,l)-th element of matrix G.

$$G(j,l) = \frac{\partial s_{jm}}{\partial p_{lm}} + \sum_{k=1}^{J_{mt}} \left(\Omega_m^r(k,j) \frac{\partial^2 s_{km}}{\partial p_{jm} \partial p_{lm}} (p_{km} - p_{km}^f - mc_{km}^r) \right) + \Omega_m^r(l,j) \frac{\partial s_{lm}}{\partial p_{jm}} \tag{40}$$

And let H_w be the J_m-dimensional vector with j-th element $\Omega_m^r(w,j)\frac{\partial s_{wm}}{\partial p_{jm}}$. Then $Gdp_m - H_w dp_{wm}^f = 0$.

$$\frac{dp_m}{dp_{wm}^f} = G^{-1} H_w \tag{41}$$

Putting J_m columns of matrix H together, the pass-through matrix is

$$\Delta_m^p = \frac{dp_m}{dp_m^f} = (G^{-1}H)' \tag{42}$$

By combining the two equations (35) and (37), the sum of the retailer's and the manufacturer's marginal costs is given by:

$$\begin{aligned} mc_m &= mc_m^r + mc_m^f \qquad (43) \\ &= p_m + (\Omega_m^r * \Delta_m^r)^{-1} s_m(p_m) + (\Omega_m^f * \Delta_m^f)^{-1} s_m(p_m) \\ &= p_m + (\Omega_m^r * \Delta_m^r)^{-1} s_m(p_m) + (\Omega_m^f * \Delta_m^p \Delta_m^r)^{-1} s_m(p_m) \end{aligned}$$

Model of Zero Retail Margins (Passive Retailer Model) This is one kind of two–part tariff model in which one party prices at marginal cost and the other party claims the residuals and pays fees to the marginal–cost–pricing party. If retailers set their prices only to cover the retail costs, they add only retail marginal costs to the wholesale price. In this case, retailers receive franchise fees from manufacturers and can recover part or all of the manufacturer's profits, depending on the bargaining power between them. The retail margins in this case are given by:

$$p_m - p_m^f - mc_m^r = 0 \tag{44}$$

and the manufacturer margins are:

$$\begin{aligned} p_m^f - mc_m^f &= -(\Omega_m^f * \Delta_m^f)^{-1} s_m(p_m) \\ &= -(\Omega_m^f * \Delta_m^r)^{-1} s_m(p_m) \end{aligned} \tag{45}$$

The second equality in equation (45) comes from the fact that Δ_m^p, the pass-through matrix, is the identity matrix, which is implied from (44). By summing (44) and (45), the sum of retailer's and manufacturer's marginal costs is given by:

$$mc_m = p_m + (\Omega_m^f * \Delta_m^r)^{-1} s_m(p_m) \tag{46}$$

The other case of a two–part tariff model occurs when manufacturers set the wholesale prices equal to the marginal costs and let retailers claim the residuals. Retailers set their profit–maximizing retail prices given the wholesale prices, and the manufacturers take part or all of the industry profits in the form of fixed fees from retailers. The retailer price–cost margins are given by (35) where $p_m^f = mc_m^f$. Hence, the implied marginal costs of retailers and manufacturers are:

$$mc_m = p_m + (\Omega_m^r * \Delta_m^r)^{-1} s_m(p_m) \tag{47}$$

I am not estimating this model because the merger occurred only at the upstream level. In equation (47), there is nothing to change when performing merger simulation, so the simulated prices will be the same as the pre–merger prices.

Model of Retailer Collusion When retailers follow collusive pricing, retail prices are set to maximize joint profits of all products sold by all retailers. In this case, the retail margins are given by:

$$p_m - p_m^f - mc_m^r = -(\Delta_m^r)^{-1} s_m(p_m) \tag{48}$$

All elements of the retailer ownership matrix are one because all products are regarded to be owned by one firm in the joint–profit maximization. The manufacturer margins are the same as in equation (37). By summing equations (37) and (48), I obtain

$$mc_m = p_m + (\Delta_m^r)^{-1} s_m(p_m) + (\Omega_m^f * \Delta_m^p \Delta_m^r)^{-1} s_m(p_m) \tag{49}$$

In all three cases considered above, with access to the wholesale price data, I would be able to separately estimate retailer and manufacturer marginal costs. Furthermore, I would be able to estimate how each changes after the merger using both pre–merger and post–merger data in each case. Table 12 summarizes the formula for implied marginal costs in three cases of inclusion of retailers.

Table 12: **Comparison of Different Supply Models**

Model	Retailer margins	Manufacturer margins	$mc = mc^f + mc^r$
1	$(\Omega^r * \Delta^r)^{-1}s$	$(\Omega^f * \Delta^f)^{-1}s$	$(\Omega^r * \Delta^r)^{-1}s + (\Omega^f * \Delta^f)^{-1}s$
2	0	$(\Omega^f * \Delta^r)^{-1}s$	$(\Omega^f * \Delta^r)^{-1}s$
3	$(\Delta^r)^{-1}s$	$(\Omega^f * \Delta^f)^{-1}s$	$(\Delta^r)^{-1}s + (\Omega^f * \Delta^f)^{-1}s$

Model 1: Double Marginalization
Model 2: Zero Retail Margins
Model 3: Retailer Collusion

2.3.3 Merger Simulation

For each of the three supply models above, I simulate post–merger prices under the assumption that marginal costs and demand are fixed at the pre–merger levels. I use the demand parameters estimated in the first part of Section 3. The simulated post–merger prices in market m are as follows:

1. Double marginalization model:

$$p_m^* = mc_m - (\Omega_m^r * \Delta_m^r)^{-1} s_m(p_m^*) - (\Omega_m^{f,\text{ Postmerger}} * \Delta_m^p \Delta_m^r)^{-1} s_m(p_m^*) \tag{50}$$

2. Zero retail margins model:

$$p^*_m = mc_m - (\Omega^{f,\text{ Postmerger}}_m * \Delta^r_m)^{-1} s_m(p^*_m) \tag{51}$$

3. Retailer collusion model:

$$p^*_m = mc_m - (\Delta^r_m)^{-1} s_m(p^*_m) - (\Omega^{f,\text{ Postmerger}}_m * \Delta^p_m \Delta^r_m)^{-1} s_m(p^*_m) \tag{52}$$

2.4 Empirical Results

In this section, I present the demand estimates and the simulation results for Sections 2 and 3. To evaluate the performance of all models, I compare the simulated price changes of each model to the actual price changes.

2.4.1 AIDS Demand Estimation Results

The parameter estimates of the AIDS demand model are shown in Table 13. θ is the effects of the price index on the category real expenditure. Each β_i is the effect of category real expenditure on each brand i's demand. γ_{ij} is the effect of brand j's price on brand i's demand.

The own–price and cross–price elasticities are shown in Table 14. Because I did not restrict γ_{ij} to be equal to γ_{ji} in the bottom–demand model, the cross–price elasticities between two products are not symmetric. Some cross–elasticity values are negative, possibly resulting in unreasonable simulation results.

I report price–cost margins of each brand in Table 15. Smuckers brand's margin is the highest. The average margin in this industry is 38%.

Table 13: **AIDS Demand Parameters**

θ		β_i	γ_{ij}				
			Jif	Skippy	Peter Pan	Private Label	Smuckers
-2.3209	Jif	-0.0692	-0.5058	0.0823	0.1043	0.0506	0.0329
(0.0591)		(0.0058)	(0.0142)	(0.0081)	(0.0097)	(0.0107)	(0.0366)
	Skippy	0.1247	0.2485	-0.2273	0.1441	0.0961	-0.0573
		(0.0058)	(0.0142)	(0.0081)	(0.0097)	(0.0107)	(0.0366)
	Peter Pan	0.0153	0.1207	0.0772	-0.3315	0.0535	0.0480
		(0.0058)	(0.0142)	(0.0081)	(0.0097)	(0.0107)	(0.0366)
	Private Label	-0.0555	0.1332	0.0654	0.0827	-0.2074	0.0160
		(0.0058)	(0.0142)	(0.0081)	(0.0097)	(0.0107)	(0.0366)
	Smuckers	-0.0153	0.0033	0.0025	0.0003	0.0072	-0.0397

Standard errors in parentheses

Table 14: **AIDS model: Own-Price and Cross-Price Elasticity**

	Jif	Skippy	Peter Pan	Private Label	Smuckers
Jif	-2.7724	0.1042	0.1326	-0.0103	0.0669
Skippy	0.6063	-3.2744	0.3114	-0.0426	-0.6832
Peter Pan	0.1888	0.3400	-3.5549	0.0412	0.2413
Private Label	0.3839	0.2371	0.3006	-2.1232	0.0477
Smuckers	-0.0265	0.0363	-0.0144	0.1037	-2.0812

Median values across cities are reported.

Table 15: **AIDS model: Price-Cost Margins**

Brands	Margins
Jif	36.07%
Skippy	30.59%
Peter Pan	28.14%
Private Label	47.10%
Smuckers	48.05%

Median values across cities are reported.

2.4.2 Retailer-Level Model Demand Estimation Results

The number of products defined as retailer–brand combinations in the sample is 880. Data on mass merchandisers, such as Walmart, are not included in the IRI dataset, so the retailers are all local chains, with each city having a different set of retailers. In Table 16, I report the logit IV demand parameters. For convenience of reporting, I present only the median value of manufacturers' brand intercepts and the median value of their standard errors.

Own–price and cross–price elasticities are shown in Tables 17 and 18. Table 17 includes median values of price elasticities at the manufacturer brand level. Table 18 shows the actual price elasticities among retailer-brand combinations in one market. Because the prices in the data are retailer level, defining a product as a retailer–brand should reflect more realistic substitution patterns between products.

Table 16: **Inclusion of Retailer: Logit IV Demand Parameters**

Variable	
Price	-1.2065
	(0.3473)
Average display	0.8904
	(0.2455)
Time trend	0.0273
	(0.0043)
Jif	-1.9627
	(0.6605)
Skippy	-2.9386
	(0.6381)
Peter Pan	-3.0490
	(0.6143)
Private Label	-2.8433
	(0.5377)
Smuckers	-3.4580
	(0.9315)

Estimates are all statistically significant at 1%.

Table 17: **Inclusion of Retailer: Own-Price and Cross-Price Elasticities**

Brand	Own-Price Elasticity	Cross-Price Elasticity
Jif	-2.2541	0.0213
Skippy	-2.1976	0.0107
Peter Pan	-2.0883	0.0110
Private Label	-1.8165	0.0120
Smuckers	-3.1243	0.0036

Table 18: **Inclusion of Retailer: Own-Price and Cross-Price Elasticities in One Market**

Product	Own-Price Elasticity	Cross-Price Elasticity
Jif in Retailer 1	-2.0514	0.0047
Skippy in Retailer 1	-2.0549	0.0003
Peter Pan in Retailer 1	-1.9800	0.0033
Private Label in Retailer 1	-1.5207	0.0023
Jif in Retailer 2	-2.1660	0.0684
Skippy in Retailer 2	-2.1688	0.0075
Peter Pan in Retailer 2	-1.8331	0.0494
Private Label in Retailer 2	-1.6319	0.0505
Smuckers in Retailer 2	-3.3166	0.0017
Jif in Retailer 3	-2.1548	0.0702
Skippy in Retailer 3	-2.2155	0.0129
Peter Pan in Retailer 3	-2.4720	0.0244
Private Label in Retailer 3	-1.9813	0.0184
Smuckers in Retailer 3	-2.8794	0.0043
Jif in Retailer 4	-2.2909	0.0129
Skippy in Retailer 4	-2.0262	0.0008
Peter Pan in Retailer 4	-1.9974	0.0074
Private Label in Retailer 4	-1.6854	0.0065
Smuckers in Retailer 4	-3.3658	0.0005

Retailer 1 in this market does not carry Smuckers.

2.4.3 Performance of the Merger Simulation

For each model of retailer–manufacturer relationship on pricing, I compare the simulated price changes of the merging firm's two brands, Jif and Smuckers to the direct measures. I also compare them to the simulated price changes from the structural demand model. Tables 19 and 20 show the simulation results for AIDS and three models of retailer inclusion.

All models with retailer–side pricing incorporated produce higher predictions on price changes than the model without retailer–side pricing. Because of the assumed monopolistic behavior at the retailer level, the model of retailer collusion shows the highest predicted price changes, with 0.49% for Jif and 4.11% for Smuckers. Differences in the predicted price changes are relatively large across different models of retailer–manufacturer pricing relationships.

This result indicates that addressing precisely the upstream and downstream relationships of firms in merger simulations is important. Wrong models of firm conduct could result in misleading recommendation concerning the antitrust policy. Given the empirical result that incorporating retailer pricing in the model produces higher predicted price changes than the conventional oligopoly models at the manufacturer level, analysis at the retailer level could partially resolve the underprediction problem of merger simulation in the consumer packaged goods industry.

As mentioned in Chapter 1, underprediction is partly caused by the logit error structure combined with an outside option which has high market share. The logit assumption causes the substitution patterns to depend on the brands' shares. Because the ouside option has high market share, which is typical in the consumer packaged goods industry, shares of inside options are relatively low. Hence, the presence of a large outside option necessarily makes the substitution patterns of the logit model understated. Compared to the logit IV result in Table 20 (0.17% increase for Jif and 1.31% increase for Smuckers), the results of models with retailer inclusion show considerable differences. Given that the directly estimated price changes are correct, I conclude that incorporating the retailer side in the model gives a closer prediction of post–merger prices when researchers use scanner data in the merger analysis.

2.5 Conclusion

In Chapter 2, I considered an alternative demand model and supply models in the merger simulation. In the first part, I estimated an AIDS instead of a structural demand model. The AIDS demand model combined with the conventional Bertrand–Nash oligopoly model of manufacturer pricing game produces very different predictions of post–merger price changes from those of the structural demand model. Crook et al. (1999) report that different assumptions on the functional form of the demand model cause significant differences in predictions of post–merger price increase. Their finding has been supported in this chapter. Because there is no rule of thumb that best explains demand in every industry, empirical research on the characteristics of demand curves is required before one chooses a demand model in a merger simulation.

Table 19: **Alternative Demand: Actual and Predicted Percent Price Changes for Merging Firm's Brands**

	Actual Merger Effect		Simulated Merger Effect		
			Structural Demand Model		Alternative Demand Model
Brand	Time-diff	Diff-in-diff	Logit IV	RC IV	AIDS
Jif	0.96	1.63	0.17	0.25	-0.01
Smuckers	3.33	7.48	1.31	1.99	7.44

Table 20: **Inclusion of Retailer Pricing: Actual and Predicted Percent Price Changes for Merging Firm's Brands**

	Actual Merger Effect		Simulated Merger Effect				
			Without Retailer Pricing		With Retailer Pricing		
Brand	Time-diff	Diff-in-diff	Logit IV	RC IV	Double Marginalization	Passive Retailers	Retailer Collusion
Jif	0.96	1.63	0.17	0.25	0.35	0.31	0.49
Smuckers	3.33	7.48	1.31	1.99	2.62	2.17	4.11

Note: Brand intercepts are set to be different across cities in the Logit IV demand model. RC IV refers to the random coefficient IV demand model with full covariance structure of brand preferences. Brand intercepts are also set to be different across cities. All price changes are mean values across markets.

In the second part of this chapter, I incorporated retailer–side pricing into the model and noted the merger simulation performed differently. The conventional merger simulation assumes omission of retailer–side pricing and uses oligopoly models of the pricing game at the manufacturer level. Doing so has been shown to be problematic if the available data are at the retailer–level. To incorporate retailer side into the model, I considered three cases of retailer–manufacturer relationships. In all three models, inclusion of retailer pricing in the model produced closer predictions to the directly estimated price changes. Hence, the underprediction problem could be partially resolved by incorporating retailer side when one analyzes mergers using scanner data in the consumer packaged goods industry.

3 Chapter 3. Explaining the Difference Between the Predicted Price Changes and the Actual Price Changes

3.1 Introduction

In Chapter 3, I examine one of the possibilities that explains why merger simulation generally does not give accurate predictions of the actual effects of mergers. As mentioned in the conclusion of Chapter 1, the first possibility was the validity of the demand model. I addressed this issue and considered a more flexible demand model in Chapter 1 and an alternative demand model in Chapter 2. To address the second possibility, that the equilibrium assumption of the firms' pricing game might be wrong, I included the retailer–side pricing in the supply model and considered various types of retailer–manufacturer relationships in pricing. The third possibility lies with the strict assumptions of merger simulation itself. In the periods before and after a merger, many things can change besides the merger event. Among them, I estimate the effect of four factors that lead to the differences between simulated price changes and actual price changes. The actual price changes refer to the price changes that are directly estimated using retrospective analysis. The analysis in this chapter follows Peters (2006).

There are four factors that change during the post–merger period. First, market power changes. With fewer competitors, the merging firm sets the prices to maximize joint profits coming from existing products and newly purchased products. Merger simulation assumes that demand and cost are all fixed at the pre–merger level. The only thing changed in the simulation is the ownership structure. Hence, what the merger simulation measures is the pure market power effect. If the assumptions that demand and cost are all fixed at the pre–merger level is true, the pure market power effect would reflect the actual price changes. However, it is generally expected that the simulated price changes differ from the actual changes if the factors that are fixed in the merger simulation change significantly in the post–merger periods. A second factor that changes in the post–merger period is observed consumer preferences. Observed consumer preferences are captured by the structural demand parameters. This second factor is one of the things that is assumed not to change in the merger simulation. By running counterfactuals using post–merger observed consumer preferences instead of pre–merger observed consumer preferences, I estimate the effect of changes in observed preferences on predicted prices.

Third, along with the observed preferences, unobserved demand shocks can also change in the post–merger period. In the structural demand model, the unobserved demand shocks are all captured in the demand residual term, ξ. Using the post–merger data, this term is inverted out in the same way as the pre–merger values were recovered. Using the post–merger demand parameters and post–merger demand residuals, I estimate the effect of the changes in unobserved demand shocks on predicted prices.

Finally, changes in firm conducts or unobserved supply side changes such as regulations and input cost shocks can also affect the post–merger prices. After addressing the effect of changes in observed and unobserved demand separately, all the remaining differences are explained by this final factor.

The result, using the logit IV demand model, indicates that, for the merger between Jif and Smuckers consummated in mid–2002 in the U.S. peanut butter industry, the market power effect accounts for 0.17% and 1.31% of the price increase for Jif and Smuckers in the post–merger period. The changes in demand explain about 9% and 7% of the increase in prices for Jif and Smuckers, respectively. The effect of unobserved demand shocks on the post–merger prices is minimal. The reduction of prices caused by changes in unobserved supply side factors accounts for 7% and 4% price decreases for each product, respectively.

3.2 What Factors Cause the Differences Between Simulated and Actual Price Changes?

Merger analysis using simulation is not a simple task and often produces inaccurate predictions compared to the actual merger effects. Because a merger is not like a lab experiment in which one can control for all other factors that might affect the results. When a merger occurs, the market power of the merging firm definitely increases due to the loss of competition in the market. If only the number of competitors changed then the post–merger prices would only be affected by the changes in market power. However, there must be other changes along with the merger event. For example, consumer preferences may change. Failing to reflect such changes appropriately in the simulation could cause the discrepancies between predictions and the actual price changes.

Other than the demand–side changes, there could also be changes in the supply side. For example, different input costs in the post–merger period may affect the post–merger prices. Even if the input cost is stable over time, firms' conduct between competitors or firms' interactions with the suppliers of raw material could also change, affecting the cost of production in the post–merger period.

Another possible change is the potential efficiency gain effect of the merger on production cost. A merger can achieve large cost savings in the production process of the merging firm. More efficient use of the production facility or increased bargaining power may allow the merging firm to price its products at a lower level.

In sum, the pure market power effect, demand–side changes, and supply–side changes are all compounded in the actual price changes in the post–merger period. Among these, the merger simulation can address only the pure market power effect, assuming all other potential changes do not happen. Then the difference between the price changes predicted by merger simulation and the actual price changes is attributed to the effect of those other factors besides the market power. In this section, I run several counterfactuals to estimate the separate effect of each of these factors on the actual merger effect.

3.2.1 Counterfactuals

Having the post–merger period data allows estimating separately each factor's contribution to the actual price changes in the post–merger period. I explain how I run the counterfactuals to do so in the following subsections.

Market Power First, changes in market power affect post–merger prices. These changes are what the merger simulation actually measures. In performing a merger simulation, three values at fixed at the pre–merger level. One is the set of demand parameters, which are based on the the assumption that observed consumer preferences do not change in the post–merger period. Another value that is fixed at the pre–merger level is unobserved demand shock. The last is the marginal cost. With these values fixed, merger simulation changes only the ownership structure to address the market power effect on the new equilibrium prices.

Changes in Observed Demand Before discussing how the changes in observed consumer preference affect the simulated prices, one first must determine whether the demands are really different between the pre–merger and post–merger periods. If the post–merger demand is not significantly different from the pre–merger demand, then the effect of this factor would not explain much about the actual price changes. The simplest econometric test to determine whether the demands during the two periods are different is the Chow test (Chow 1960). It is conducted by estimating three different demand models. First one uses the whole data, from both the pre–merger and post–merger periods. Second and third are estimated separately for each period. The statistic of the Chow test is:

$$\frac{(S_c - (S_1 + S_2)/k}{(S_1 + S_2)/(N_1 + N_2 - 2k)} \sim F(k, N_1 + N_2 - 2k)$$

where S_c is the sum of squared residuals from whole data, S_1 and S_2 are sum of squared residuals from pre–merger data and post–merger data, respectively. k is the number of parameters in the demand models. N_1 and N_2 are the number of observations in the pre–merger and post–merger data, respectively. The statistic is distributed as F distribution with degrees of freedom k and $N_1 + N_2 - 2k$.

I estimate the OLS logit demand model using the sample data on the merger between Jif and Smuckers in the U.S. peanut butter industry to check whether demand in the two periods differ. The test statistic I derived is 15.609 and the distribution is F(752,22588). The null hypothesis that there is no change in demand across periods is rejected 99.99% of the time.

Given that we know there are statistically significant changes in the demand, the effect of these changes on the actual price changes can be estimated by running the simulation as follows. In the simulation, I use the post–merger demand parameters, which are estimated from the separate demand estimation. Unobserved demand shocks and marginal cost are fixed at the pre–merger level as in the original merger simulation. Then I change the ownership matrix and simulate the predicted price changes. The difference between the predicted price changes acquired here and the predicted price changes from the original merger simulation are attributed to the observed changes in demand.

Changes in Unobserved Demand Next, I consider the effect of changes in unobserved demand shocks on the actual price changes. In the structural demand model, all the demand shocks that are not captured by the model are thrown into the unobserved demand shock term $\dot{\xi}$. I invert out the post–merger values of ξ from the separate demand estimation for the post–merger period and then use this term in the simulation. Thus, the counterfactuals are run using post–merger demand parameters, post–merger unobserved demand shocks, and pre–merger marginal costs. The difference between the predicted price changes obtained here and the predicted price changes obtained in the previous subsection will be the effect of the changes in the observed demand shocks.

Changes in Cost and Supply Side After changes in demand side are accounted for, all the remaining discrepancies between the predicted price changes and the actual price changes are caused by the supply–side changes. Depending on which one is chosen as the actual price changes between the time-difference estimates and the difference-in-difference estimates, the effect of the supply–side factors could be different.

3.2.2 Results

Table 21 shows each factor's effect that consists of the actual post–merger price changes. In each counterfactual, I use a simple logit IV demand model with different brand intercepts across cities to address the multi–market structure of the data.

Table 21: **Impact of Different Factors on Actual Price Changes in the Post-merger Period**

	Components				
Brand	Market Power Effect	Changes in Observed Demand	Changes in Unobserved Demand	Supply-side Effect	Actual Changes (Time-Diff)
Jif	0.17%	9.11%	0.02%	-7.64%	0.96%
Smuckers	1.31%	6.53%	0.01%	-4.27%	3.33%

Note: All the price effects are the mean value across markets.

Initially, the sole market power effect underpredicts the actual price changes for both Jif and Smuckers. Then, the changes in observed demand mainly lead the predicted prices to increase greatly. Without the supply–side effect on price reduction, the market power effect along with the changes in demand would have led the prices in the market to increase overmuch.

For Smuckers, market power has caused the price to increase 1.31%, while changes in demand account for another 6.53% increase in prices. Unobserved demand shocks do not have significant effect on the prices. Reduction of prices by 4.27% is accounted for by the changes in supply–side factors.

The reduction of prices explained by the unobserved supply side changes might be because of the reduction in price of peanuts. The peanut prices dropped by 37% in the post–merger period compared to the pre–merger period. Table 22 shows the results of the regression of the log of prices of peanuts on the post–merger dummy variable after controlling for monthly seasonal effects. Hence, the lower input prices may have partially caused the prices to decrease. Alternatively, there might have been entry of more products into the market, which have been omitted in the model. New products may have caused the firms to compete more intensely and to drop their prices. Another possible explanation is the potential efficiency gain, which is the synergy effect, resulting from the merger. Savings in production and distribution costs may have enabled the merged firm to drop its prices.

Table 22: **Regression of Peanut Prices**

Variable	Coefficient (Std. Err.)
Post-Merger	-0.372***
	(0.000)
Intercept	-1.552***
	(0.001)
Monthly fixed effects suppressed	
N	24077
R^2	0.9690
$F_{(2,197)}$	62765.22

3.3 Conclusion

In Chapter 3, I estimated the separate effect of many factors on actual post–merger prices. Merger simulation inherently accounts only for a pure market power effect among these factors. Besides the merger, there are many things going on in the market. All are compounded in the post–merger prices, making it difficult to separate each factor's contribution to the prices changes. The merger simulation generally fails to predict an actual price change accurately because the demand–side factors and supply–side factors also change during the post–merger period.

By exploiting both periods of merger data, I have successfully estimated each factor's effect on the actual price changes. In the merger between Jif and Smuckers in the U.S. peanut butter industry, the changes in demand have positively affected the post–merger price changes. In addition, the supply–side factors have negatively affected the prices. This last can be partly explained by the reduction of peanut prices. Otherwise, potential efficiency gain or changes in firm conduct during the post–merger period could offer explanations for the reduction of prices attributed to the supply–side changes.

References

[1] Albuquerque, P. and B. J. Bronnenberg (2009), "Estimating Demand Heterogeneity Using Aggregated Data: An Application to the Frozen Pizza Category," *Marketing Science*, 28(2): 356-372.

[2] Ashenfelter, O. C. and D. S. Hosken (2008), "The Effect of Mergers on Consumer Prices: Evidence from Five Selected Case Studies" (NBER Working Paper 13859), National Bureau of Economic Research, Cambridge, MA.

[3] Ashenfelter, O. C., D. S. Hosken, and M. C. Weinberg (2011), "The Price Effects of a Large Merger of Manufacturers: A Case Study of Maytag-Whirlpool" (NBER Working Paper 17476), National Bureau of Economic Research, Cambridge, MA.

[4] Anderson, S., A. De Palma, and J. F. Thisse (1992), *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press.

[5] Baker, J. B. (1999), "Econometric Analysis in FTC v. Staples," *Journal of Public Policy & Marketing,* 18(1): 11-21.

[6] Berry, S. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25: 242-262.

[7] Berry, S., J. Levinsohn, and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4): 841-890.

[8] Berger, A. N., S. D. Bonime, L. G. Goldberg, and L. J. White (2004), "The Dynamics of Market Entry: The Effects of Mergers and Acquisitions on Entry in the Banking Industry," *Journal of Business*, 77(4): 797-834

[9] Bronnenberg, B. J., S. K. Dhar, and J. P. Dube (2009), "Brand History, Geography, and the Persistence of Brand Shares," *Journal of Political Economy*, 117(1): 87-115.

[10] Berry, S. and J. Waldfogel (2001), "Do Mergers Increase Product Variety? Evidence From Radio Broadcasting," *Quarterly Journal of Economics*, 116 (3): 1009-1025.

[11] Bonfrer, A. and J. S. Raju (2003), "The Demand and Supply-side Impact of the Kimberly-Clark, Scott Paper Products Merger in the Facial Tissues Category" (Working paper, Wharton-SMU Research Center, Singapore, November 2003).

[12] Carlton, D. W. (2009), "Why We Need to Measure the Effect of Merger Policy and How To Do It" (NBER Working Paper 14719), National Bureau of Economic Research, Cambridge, MA.

[13] Chintagunta P. K., A. Bonfrer, and I. Song (2002), "Investigating the Effects of Store-Brand Introduction on Retailer Demand and Pricing Behavior," *Management Science,* 48(10): 1242-1267.

[14] Chintagunta, P. K., J. P. Dube, and V. Singh (2002), "Market Structure across Stores: An Application of a Random Coefficients Logit Model with Store Level Data," *Advances in Econometrics,* 16: 191-221.

[15] Chintagunta, P. K. (2000), "A Flexible Aggregate Logit Demand Model" (Working Paper, University of Chicago, July 2000).

[16] Chow, G. C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," Econometrica, 28(3): 591–605.

[17] Crooke, P., L. Froeb, S. Tschantz, and G. J. Werden (1999), "Effects of Assumed Demand Form on Simulated Postmerger Equilibria," *Journal of Industrial Organization,* 15: 205-217.

[18] Dube, J. P. (2005), "Product Differentiation and Mergers in the Carbonated Soft Drink Industry," *Journal of Economics and Management Strategy*, 14(4): 879-904.

[19] Einav, L. and J. Levin (2010), "Empirical Industrial Organization: A Progress Report," *Journal of Economic Perspectives,* 24(2): 145-162.

[20] Feenstra, R. and J. Levinsohn (1995), "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies*, 62: 19-52.

[21] Farrell J., P. A. Pautler, and M. G. Vita (2009), "Economics at the FTC: Retrospective Merger Analysis with a Focus on Hospitals," *Review of Industrial Organization,* 35: 369-385.

[22] Genakos, C. (2004), "Differential Merger Effects: The Case of the Personal Computer Industry," mimeo, London Business School.

[23] Goeree, M. S. (2009), "Was Mr. Hewlett Right? Mergers, Advertising and the PC Industry" (Working paper, University of Southern California, Economics Department, Los Angeles, February 2009).

[24] Goettler, R. L. and R. Shachar (2001), "Spatial Competition in the Network Television Industry," *RAND Journal of Economics*, 32(4): 624-656.

[25] Kim, D. (2004), "Estimation of the Effects of New Brands on Incumbents' Profits and Consumer Welfare: The U.S. Processed Cheese Market Case", *Review of Industrial Organization,* 25: 275-293

[26] Kim, E. H, and V. Singal (1993), "Mergers and Market Power: Evidence from the Airline Industry," *The American Economic Review,* 83(3): 549-569.

[27] Knittel, C. R. and K. Metaxoglou (2011), "Challenges in Merger Simulation Analysis," *American Economic Review: Papers & Proceedings*, 101(3): 56-59.

[28] Langer, A. (2011), "Demographic Preferences and Price Discrimination in New Vehicle Sales" (Working paper, University of Michigan, Ann Arbor, January 2011).

[29] Nevo, A. (2000), "Mergers with differentiated products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 31(3): 395-421.

[30] Nevo, A. (2001), "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69(2): 307-342.

[31] Nevo, A. and M. D. Whinston (2010), "Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference," *Journal of Economic Perspectives,* 24(2): 69-82.

[32] Peters, C. (2006), "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law and Economics,* 49(2): 627-649.

[33] Shapiro, C. (1996), "Mergers with Differentiated Products," *Antitrust*, 10(2): 23-30.

[34] Singh, V. and T. Zhu (2008), "Pricing and Market Concentration in Oligopoly Markets," *Marketing Science,* 27(6): 1020-1035.

[35] Thomadsen, R. (2005), "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries," *RAND Journal of Economics*, 36(4): 908-929.

[36] Villas-Boas, S. B. (2007a), "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *Review of Economic Studies*, 74: 625-652.

[37] Villas-Boas, S. B. (2007b), "Using Retail Data for Upstream Merger Analysis," *Journal of Competition Law and Economics,* 3(4): 689–715.

[38] Weinberg, M. C. (2011), "More Evidence on the Performance of Merger Simulations," *American Economic Review: Papers & Proceedings*, 101(3): 51-55.

[39] Weinberg, M. C. and D. Hosken (2012), "Evidence on the Accuracy of Merger Simulations" (Working paper).

[40] Werden, G. J. (1997), "Simulating the Effects of Differentiated Products Mergers: A Practitioners' Guide," Department of Resource Economics Regional Research Project 967, University of Massachusetts.

[41] Werden, G.J. and L. M. Froeb (1994), "The Effects of Mergers in Differentiated Product Industries: Logit Demand and Merger Policy," *Journal of Law, Economics, and Organization*, 194: 407-26.