Essays in Industrial Organization

by

Travis Bruce Triggs

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2021

Doctoral Committee:

Associate Professor Ying Fan, Chair
Assistant Professor Zach Y. Brown
Professor Francine Lafontaine
Professor S. Sriram

Travis Bruce Triggs

ttriggs@umich.edu

ORCID iD: 0000-0002-0161-4720

Acknowledgments

I would like to especially thank Ying Fan, Zach Brown, Francine Lafontaine, and S. Sriram for their invaluable guidance in my research and in my development as an economist. I would also like to thank my coauthor Xuan Teng for her tireless efforts in our joint research. Additionally, my graduate school experience would not have been nearly as rich without the many wonderful friendships I formed in Lorch Hall, including those with Thomas Helgerman, Andrew Simon, Dena Lomonosov, and Elird Haxhiu. Additionally, the research contained in this dissertation benefitted from conversations with Andrew Usher, Kimberly Conlon, Anirudh Jayanti, Colin Watson, Aaron Kaye, Russell Morton, Michael Ricks and many others.

I would also like to express my deepest appreciation to my family, financée, and friends from back home for their loving support of my doctoral pursuits.

Researcher(s)' own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

This dissertation contains three essays in industrial organization economics that use theoretical, reduced-form, and structural analysis to examine various effects of policies and market mechanisms in concentrated industries.

The first chapter, "Equilibrium and Welfare Implications of Store Brands in Vertical Markets," theoretically and empirically studies the effect of store brands on equilibrium outcomes and welfare. I demonstrate the ambiguous theoretical predictions of how store brands can impact equilibrium outcomes using a simplified model, motivating two related empirical exercises that assess the implications of store brands in the U.S. yogurt market. First, I use an event study framework to show that after a store brand introduction, the prices of retailer national brand products increase. Second, I develop a structural model of upstream vertical interaction between manufacturers and retailers, incorporating the strategic use of store brands and downstream consumer demand for differentiated products. I fit the model to observed pricing and quantity data to recover underlying consumer preference parameters and marginal costs. I then conduct counterfactual simulations that remove retailer store brands to show that the benefit of store brands to retailers and consumers, and the harm to manufacturers, is magnified under increased upstream market concentration.

The second chapter, "Patent Licensing and Bias in Estimation and Prediction" (with Xuan Teng), evaluates the consequences of analyzing market outcomes when patent licensing relationships exist but are not observed by the researcher. First, we theoretically illustrate the bias in estimated marginal costs due to not accounting for the alignment of pricing incentives between licensors and licensees and demonstrate the resulting ambiguous direction

of bias in prediction of merger effects. Second, using market simulations, we analyze these biases in two types of mergers in markets with patent licensing: licensor-licensee mergers and licensee-licensee mergers. For both types of mergers, we find that a mis-specified model that does not incorporate existing patent licensing relationships can predict merger effects that are: i) opposite to the true merger effects that account for existing patent licensing relationships; ii) typically over-predicted. Furthermore, we characterize the variation in estimation and prediction biases with respect to royalty rates and the sum of diversion ratios between a licensor and licensees. The simulation results support the importance of incorporating patent licensing relationships when modeling markets with relatively large royalty rates.

The third chapter, "Product Responses to Income-Based Subsidies in the U.S. Infant Formula Industry", develops a structural supply and demand model of the infant formula industry to evaluate the role that endogenous firm product offerings play in determining equilibrium market outcomes and welfare measures. Using the structural model's preference and cost estimates, I evaluate counterfactual scenarios which increase the proportion of infant formula voucher recipients and show i) the policy's negative effect on consumers without vouchers, and, ii) how the magnitude of consumer, producer, and total surplus depend on firms' adjustment margins.

Chapter 1

Equilibrium and Welfare Implications of Store Brands in Vertical Markets

1.1 Introduction

Retailers are intermediaries between the production of goods by manufacturers and the consumption of those goods by consumers. Acting in complementary roles in the vertical supply chain, retailers' and manufacturers' incentives in the vertical supply chain are aligned in producing and selling, respectively, final goods to consumers. If these aligned incentives improve variety, enhance quality, or lower costs, consumers benefit. However, the alignment of incentives may critically divert when retailers additionally compete horizontally with manufacturers in the production of goods. Retailers then face a tradeoff between reselling manufacturers' products and selling their own. In this chapter, I assess the equilibrium implications of retailers' joint role as a vertical intermediary for, and a horizontal competitor to, upstream manufacturers.

This question of what effect does a retailer jointly selling products from arm's-length manufacturers and their own integrated offerings have on a market has recently entered policy discussions. For example, scrutiny over Amazon's use of data on third-party sellers have recently come from antitrust authorities in both the U.S. and European Union.[1] Though, intermediaries that play a dual role of being a retailer as well as integrated with the manu-

[1] For example, see "Amazon to Face Antitrust Charges From EU Over Treatment of Third-Party Sellers", Valentina Pop and Sam Schechner, *Wall Street Journal*, June 11, 2020.

facture of goods predates ecommerce. For example, brick and mortar grocers have, for more than twenty years, sold "store brands"[2] alongside national branded products.

These store brand products are owned exclusively by a retailer, which govern their production and sale. In some settings, retailers are fully integrated with upstream production.[3] Alternatively, the production of some store brand goods is contracted out to third-party producers. However, retailers maintain strict control over the production process, including specifying product features, labeling, and terms of distribution, despite sourcing at arm's-length. In either case, store brands, relative to national brands, empirically exhibit lower prices, higher retail margins, less advertising, and are typically not subject to trade allowances or slotting fees.[4]

In this chapter, I evaluate the effect of intermediaries jointly selling their own products as well as those from national manufacturers by empirically studying the implications of store brands in the U.S. yogurt industry. Though, my findings and the associated policy implications can generalize to other contexts, from other retail product categories to policy-relevant settings such as ecommerce.

This chapter contributes to economics and marketing literatures in three ways. First, I extend the literature that assess the pricing and product mix effects of store brands by exploiting a rich panel dataset of retailers spanning the U.S. in a reduced-form analysis of store brand entry. This adds to the literature assessing retailer pricing responses, for both upstream wholesale prices and downstream retail prices, of store brand introductions, which includes Bonfrer and Chintagunta (2004) and Meza and Sudhir (2010). Relatedly, I am able to additionally use my reduced-form framework to show product responses to store brand introductions, as in Sayman, Hoch and Raju (2002) and Scott-Morton and

[2]Often also referred to as "private label", "own brand", "generic", or "unbranded" products. For uniformity throughout this paper, I refer to these products as store brands.

[3]For example, in their 2013 10K, The Kroger Company states, "Approximately 40% of [store brands] sold are produced in the [Kroger's] manufacturing plants; the remaining...are produced to [Kroger's] strict specifications by outside manufacturers."

[4]Slotting fees are one-time lump sum payments paid by manufacturers to a retailer in exchange for placing a certain product on the retailer's shelf.

Zettelmeyer (2004). Second, I use a structural model that incorporates manufacturers and retailers, drawing from the empirical literature on vertical relationships, as in Villas-Boas (2007), Asker (2016), Miravete, Seim, and Thurk (2017), Hristakeva (2019), and Luco and Marshall (2020). Third, I contribute to the literature of evaluating the welfare effects of new products, specifically store brands. This literature spans from Hausman (1994) and Petrin (2002) that evaluate the welfare effects of new products to two papers on store brands most closely related to this chapter: Ellickson, Kong, and Lovett (2018) and Gross (2019). These latter two papers quantify the welfare effects from store brands by employing an upstream bargaining model requiring the assumption that retailers face zero marginal costs of retailing. I relax this assumption in my structural model, given, in my empirical context, retailing of a perishable good requires noted costs that scale with quantities sold, such as refrigeration and stocking costs. Additionally, I differ from these papers by incorporating retailer competition in my structural modeling. In whole, this chapter combines a theoretical, descriptive, and structural approach that assesses equilibrium effects and quantifies the welfare effects of store brands in oligopoly.

To state concisely, my primary research question is: what role do store brands play in determining equilibrium pricing and associated welfare in successive oligopoly? As a corollary, I also investigate how these effects depend on market concentration.

I first use a simplified theoretical model to show that the addition of a store brand to a retailer's product set has ambiguous welfare effects, motivating an empirical investigation of this question. Using the empirical context of the U.S. industry for yogurt, I conduct both descriptive and structural exercises to evaluate the role store brands play in determining equilibrium and welfare outcomes. First, using an event study framework, I show that, after store brand yogurts are introduced, retailers increase the prices of national brand product lines. I then develop a structural model to run counterfactual simulations showing the existence of yogurt store brands increase total surplus – increasing consumer surplus and (most) retailers' profits, while decreasing manufacturer profits. Finally, I show that under increased upstream concentration, the total welfare benefits of store brands increase.

The remainder of the chapter is organized as such: Section 1.2 develops a theoretical model to evaluate how store brands can impact welfare. Section 1.3 describes the empirical setting and data use in the empirical analysis. Section 1.4 shows the suggestive evidence of the effects of store brands exploiting observed store brand introductions. Sections 1.5 and 1.6 present the structural model and details how the parameters are estimated. Sections 1.7 and 1.8 show the estimation and counterfactual results. Section 1.9 concludes.

1.2 Theory

In this section, I use a theoretical framework to assess whether store brands are unambiguously welfare enhancing or welfare diminishing. I first discuss how a retailer's use of store brands can impact equilibrium outcomes and propose that the equilibrium and welfare effects of store brands depend on the competitive environment. I then formalize a theoretical model of a vertical market that allows me to conduct two sets of simulation-based comparative static exercises: first, what happens to welfare outcomes when a store brand is added? Second, what happens to welfare outcomes when a store brand replaces a national brand product? I show that, depending on model parameters, the strategic use of store brands can increase or decrease total surplus. I use this theoretical evidence as motivation to engage in an empirical investigation of my research question.

1.2.1 Qualitative Discussion

When a store brand is added to the consumer's choice set, product variety increases, which benefits consumers, all else equal. However, manufacturers and retailers may reoptimize respective wholesale and retail prices upon the addition of a store brands – and it is not determinate how these prices would change. On one hand, if the introduction of a store brand, providing the retailer and consumer with a viable substitute to national brands, acted as a constraint on national brand manufacturers markups, theory would predict a decrease in upstream markups. If the decrease in markups are passed through to decreased retail prices, consumers benefit. Alternatively, the retailer, upon introducing a store brand, could increase retail markups, and thus retail prices, on national brand products. This strategy may be profitable for the retailer if this increase in national brand prices, at the detriment of consumers, sufficiently diverts demand to higher-margin store brand products.

Comparing a market with a vertically integrated store brand product in place of a national branded product, theory would predict similar ambiguity in prices, and thus welfare. First, vertical integration eliminates double margins. That is, relative to arms-length transacting which results in successive markups when firms have market power, vertical integration aligns the upstream and downstream firms incentives. In doing so, the integrated firm internalizes the upstream markup on the profits of the downstream firm. Then, optimally, the integrated firm sets no upstream markup, resulting in a total markup that is less than the "double successive markup. In this case, consumers benefit from both the direct effect of a lower store brand retail price, but also, indirectly from resulting downward pricing pressure on retail prices of other products. However, if this coordination between production and sale were instead to result in higher prices, or alternatively, reduced quality or elimination of other national brand products, the welfare effects of replacing a national brand with an identical integrated one would be ambiguous. Higher non-store brand product prices may result if the integrated retailer disadvantaged the sale of these products in favor of store brands through higher markups (as discussed above) or exploitation of access to competitively sensitive information. Additionally, consumers could face a reduction in quality of national brands, for example, if the retailer reduced its promotional activity for these products, or outright foreclosure by the retailer, which would result in the elimination of these products from the consumer choice set.

The mechanisms discussed above suggest an ambiguous welfare effect of store brands, whether in addition to or in lieu of national brands. To formally demonstrate this welfare ambiguity, I next turn to a theoretical model which I describe, solve, and then use to conduct simulation-based comparative static exercises.

1.2.2 Theoretical Model

I develop a theoretical model with sequential profit-maximizing price setting, with, first, manufacturer(s) setting upstream price(s), and, second, a monopolist retailer setting downstream price(s). Based on downstream prices, consumers choose between the product (or outside option) that yields the highest utility. I write down and solve the model backward.

First, consumer demand for product j given by the logit form,

$$s_j = \frac{\exp\{\alpha p_j + \beta x_j + \xi_j\}}{1 + \sum_{k \in J} \exp\{\alpha p_k + \beta x_k + \xi_k\}} \tag{1.1}$$

where demand parameters α and β, which are common across consumers, measure sensitivity to retail price, p_j, and an observable exogenous product attribute, x_j. Finally, ξ_j is a scalar product-specific utility shifter.

Given demand, a monopolist retailer, with product set J, sets retail prices by maximizing variable profits, which are given by,

$$\pi^r = \sum_{j \in J} (p_j - p_j^w - c_j) s_j \tag{1.2}$$

Thus, retail prices follow the retailer's first order condition,

$$s_j + \sum_{j' \in J} \frac{\partial s_{j'}}{\partial p_j} (p_{j'} - c_{j'}) = 0 \tag{1.3}$$

Profits for the single good (non-integrated) manufacturer are given by

$$\pi_j^m = (p_j^w - c_j^w) s_j \text{ with } c_j^w = \gamma x_j + \omega_j \tag{1.4}$$

Then manufacturers set upstream prices to maximize these profits subject to the retailer's first order condition, as specified in (1.3).

In this setup, welfare for suppliers, i.e., producer surplus, is the sum of manufacturer and the monopolist retailer profits. Consumer welfare is given by a measure of consumer surplus, which for a logit demand system is,

$$CS = \frac{1}{|\alpha|} \log \left(\sum_{j \in J}^{N} \exp\{\alpha p_j + \beta x_j + \xi_j\} \right) + C \tag{1.5}$$

Where C is an arbitrary constant.

1.2.3 Comparative Statics

With this simple theoretical model, I analyze equilibrium outcomes under various market structures to assess the impact of store brands on welfare. I consider two baseline scenarios without store brands and one comparative scenario with a store brand present. Specifically, I consider,

A1: One product supplied by an arm's-length manufacturer ($j = 1$)

A2: Two products supplied by two separate arm's-length, single-product manufacturers ($j = 1, 2$)

B: One product supplied by an arm's-length manufacturer ($j = 1$) and one store brand product supplied by an integrated manufacturer ($j = 3$)

The baseline scenarios (Scenarios A1 and A2) aim to understand equilibrium outcomes in markets without store brands, i.e., no integration between manufacturer(s) and retailer. I then compare these outcomes to those under a market with a single store brand good (Scenario B). Moving from scenario A1 and B allows me to assess what happens to market outcomes when a *store brand is added.* Alternatively, the comparison between A2 and B shows the changes in equilibria when a *store brand replaces a national brand.*

Using a simulation approach, for each scenario A1, A2, and B, I calculate optimal prices, both wholesale, p^w_j, and retail, p_j, for manufacturer(s) and the monopolist retailer under a given vector of model parameter values, including consumer preference parameters (α and β), cost parameters (c_j, γ and ω_j), and product attributes (x_j and ξ_j). Solving for equilibrium wholesale and retail prices allows me to calculate product quantites, and, thus in turn, manufacturer profits (from (1.4)), retailer profits (from (1.2)), and consumer surplus (from (1.5)). I calculate total surplus as the sum of manufacturer profits, retailer profits, and consumer surplus. I then repeat this process, solving for the equilibrium and associated welfare measures, for each vector within a specified range of the model's parameter space.

As can be seen from Figure 1.1, when adding a store brand (i.e., moving from Scenario A1 to B), for all considered vectors of parameters, equilibrium retailer profits increase and equilibrium total surplus increases. In this case, for all vectors of model parameters considered, it is both profitable and total surplus enhancing when a retailer adds a store brand product alongside an existing national brand product.[5]

When comparing Scenarios A2 to B, as can be seen from Figure 1.2, replacing one of two existing national brand products with a store brand may either increase or decrease retailer profits. Equilibria in which the retailer's profits decrease upon replacement of a national brand with a store brand are shown on the left-hand side of the plot. Whenever this replacement occurs where a retailer is worse off, total surplus decreases. Alternatively, for equilibria where the retailer gains from a product replacement, total surplus can either increase or decrease. In the upper right quadrant, retailer profits and total surplus increase upon the replacement of a national brand with a store brand. However, in the lower right quadrant, we see a small mass of equilibria where it is profitable for the retailer to replace a national brand with a store brand despite this action resulting in lower total surplus. For these equilibria, the retailer's incentives deviate from the objective of the social planner to maximize total surplus.[6]

Given the demonstrated welfare ambiguity of adding a store brand, either independently, or in replacing an identical national brand product, I turn to an empirical setting in which I investigate my research question. In the next section, I describe the setting and associated

[5]However, the addition of a store brand is not necessarily a Pareto improvement, as, in many simulations, manufacturer profits decrease. See Appendix A.1.1.

[6]Consumer surplus and manufacturer profit comparisons can be found in Appendix A.1.1.

data. Then, I evaluate actual changes in products sets and their impact on observed equilibrium outcomes using an event study framework. Finally, I develop and estimate a vertical structural model that allows for counterfactual simulations to assess the causal effect of store brands on welfare.

1.3 Empirical Setting and Data Description

Given the theoretical ambiguity of the effect of store brands on competition shown in the previous section, I take this research question to an empirical setting. Specifically, I study the role of store brands in the US grocery market for yogurt. Below, I describe this market and the attributes that make it an ideal setting to study this research question. I then detail the data used in the remainder of the chapter.

1.3.1 US Yogurt Industry

I use the US yogurt industry as my empirical context in this chapter. This market exhibits several desirable features to empirically analyze this research question. First, during the sample period, the market exhibited a proliferation of both store brand and national brand products. This allows me to assess responses to product introductions. Relatedly, retailer assortments are heterogeneous amongst both store and national branded yogurt products, providing useful variation when estimating demand parameters. Second, the upstream market is relatively concentrated, with the largest three firms comprising approximately 55% of industry revenues, allowing me to examine the effect of store brands where manufacturers have market power. Third, the final good product – consumer packaged yogurt – due to its inherent properties, has a limited shelf life once produced. This supports modeling the market statically since producers, retailer, and consumers are not able to stockpile yogurt, and thus supply and demand behavior can reasonably abstract from dynamic considerations. Finally, I have detailed data on quantities, expenditure, and product availability of yogurt products at the retail level, which allows me to evaluate observed market outcomes descriptively and fit a model to this data to run counterfactual simulations.

1.3.2 Data Sources

My empirical analysis relies on data from a variety of sources. First, I use point-of-sale data on yogurt products sold in retail stores in the U.S. from 2006–2016 from Nielsen's Retail Scanner Dataset. This scanner data includes quantities sold and consumer expenditures at the weekly-store-universal product code ("UPC") level for yogurt products. For my empirical analysis, I aggregate this data to quarter-retail banner-product line within a designated market area ("DMA").[7] I define a product line as a collection UPCs which are the same brand (including retailer store brands) and share the same combination of observable product characteristics, including whether the yogurt is Greek and its fat content (i.e., low-fat, non-fat, or full fat).[8] For a given product line, I calculate the average price per standard serving by multiplying 6 ounces times the ratio of total consumer expenditure on UPCs within the product line to the total volume sold (in ounces) in a DMA-retailer-quarter.

I also use data on wholesale transactions of yogurt from PromoData PriceTrak. The PriceTrak data records the wholesale transacted prices of products in 41 cities for a single distributor per market between 2005–2012. I can match this data to product line-quarter retail sales from the scanner data described above; however, I am unfortunately not able to match these wholesale transactions to specific retailers. So, for a given observation in the scanner data sample, the associated matched wholesale transaction is not necessarily the price at which that retailer purchased that product in the upstream market. Instead, it represents a contemporaneous transaction price for that same product in the same market for some retailer (which could be the retailer of the matched scanner data observation, albeit coincidentally). Despite the limitations of perfectly matching to the scanner data, this wholesale data still provides a glimpse of the upstream market, which is often unobservable to researchers.

[7] Nielsen defines a DMA as an exclusive collection of counties, "in which the home market television stations hold a dominance of total hours viewed". DMAs roughly correspond to major metropolitan areas and are nearly exhaustive of the entirety of U.S., with only some areas of Alaska not included.

[8] Flavors vary between UPCs within a product line.

I supplement the scanner and wholesale data with data from publicly available sources. I use household counts by county from the U.S. Census Bureau's American Community Survey, which I aggregate to the DMA-level to quantify potential market size for a given DMA-quarter. I also gather data on input costs, including state-level industrial energy prices from the U.S. Energy Information Administration, and state-level industrial food manufacturing wages from the U.S. Census Annual Survey of Manufacturers, which are used to construct the instruments used in estimation (as described in Section 1.3). Finally, I hand collect additional product attributes and manufacturer plant location for use in estimating the structural model, as described in Section 1.6.

Table 1.1 shows descriptive statistics for the sample. The top panel of the table shows the average and dispersion of product line characteristics, including price. The middle three panels describe the competitive environment: the prevalence of products, retailers, and manufacturers in the sample. Finally, the bottom two panels detail the number of geographic markets, time periods, and observations.

The yogurt product lines in the sample have an average retail price of \$1.00 per size ounce serving and standard deviation \$0.46.[9] Approximately one third (34%) of product lines in the sample are Greek, which exhibit an average price of \$1.25 per six ounces (not shown in table). The yogurt product lines are nearly evenly split between the mutually exclusive and exhaustive attributes of non-fat (31%), low-fat (34%), and full-fat (31%).

The average number of retailers per DMA-quarter is 4.69, each of which exhibit an average of 20.96 product lines per DMA-quarter from 5.97 different manufacturers, on average. In total, the sample spans 205 DMAs and 44 quarters, comprising 878,678 product line-retailer-DMA-quarter observations.

1.4 Reduced-Form Evidence:Store Brand Introductions Event Study

We can assess how observed introductions of store brands impact pricing of national brands prior to and following an introduction of a (set of) store brand product line(s). For a

[9]For the 11% of the product lines which are store brands, the average price is \$0.64 (not shown in table).

given retailer-DMA, I define a store brand introduction as a period where the retailer-DMA exhibits positive sales of any store brand product lines and did not exhibit any store brand sales in four prior quarters. For example, if retailer A exhibits non-zero revenues from store brand yogurt in DMA 1 during 2012Q1, and no revenues from store brand yogurt between 2011Q1-2011Q4, I denote a store brand introduction for retailer A in DMA 1 occurred in 2012Q1.

Formally, I define an introduction indicator for retailer r in DMA d at time t as,

$$I_{rdt} = \mathbf{1}\{q_{jrdt} > 0, \sum_{\tau \in \{t-4,..,t-1\}} q_{jr\tau d} = 0, \quad \forall j \in J_{rdt}^{store}\} \tag{1.6}$$

Where J_{rdt}^{store} represents retailer r's set of product lines that are store brands in market d at time t.

In my sample, I record 256 instances of a retailer, within a specific DMA, introduce at least one store brand yogurt product line, according to the definition above. The introductions occur for 15 different retailers, in 147 different DMAs, and in 30 different quarters. The average number of store brand product lines introduced during an introduction event is 1.47 product lines.

1.4.1 National Brand Pricing

As we are interested in how national brand prices change around the retailer's introduction of store brand product lines, we can conduct a two-way fixed effects event study analysis on national brand retail prices.

$$\log(p_{jrdt}) = \zeta + \sum_{\tau \in \{t-T,...,t-2,0,..,t+T\}} \phi_\tau I_{r\tau d} + \zeta_{jrd} + \zeta_t + \epsilon_{jrdt} \tag{1.7}$$

The coefficient ϕ_{-1} is normalized to zero. Then the coefficients ϕ_τ for $\tau \in \{t-T, ..., t-2, 0, .., t+T\}$ indicate how much, on average, retail prices of branded product lines change with respect to an introduction of store brand product line τ quarters from the introduction, relative to prices one quarter prior to the introduction.

The coefficient estimates $\hat{\phi}_\tau$ are plotted in Figure 1.3. We see that, following store brand product line introductions, on average, retail prices of national brands increase and this increase is sustained in subsequent quarters. For example, in the third quarter after an introduction, on average, national brand retail prices are 4.78% higher than in the period prior to introduction $(\hat{\phi}_2) = 0.0477$.[10]

The event study exercise assesses changes in market outcomes around the time of observed retailer store brand introductions to provide suggestive evidence of equilibrium effects. However, retailer-DMAs may condition outcome variables, say national brand retailer prices, on anticipated store brand introductions. Thus, I do not claim the estimated coefficients represent casual effects, but rather descriptive of observed behavior and suggestive of equilibrium effects. To capture causal equilibrium effects, and to quantify the welfare effects of store brands, I turn to a full structural model in the next section.

1.5 Empirical Model

In order to establish casual effects and quantify welfare effects of store brands, I extend the theoretical model in Section 1.2 to develop a full structural model of manufacturers, retailers, and consumers decision making in the US yogurt industry. I solve the model and match it to observed market outcomes, including retailer prices and quantities, and input costs, to recover structural parameters governing demand and supply behavior. I then use the model and estimated parameters to conduct counterfactual simulations in Section 1.8, which allows me to assess the welfare implications of store brands in the US yogurt market.

I model interactions in this market as a repeated static game. In each period t, in each geographic market d, the following subgame is played:

1. Manufacturers and retailers observe demand shocks

[10] The 95% confidence interval for $\hat{\phi}_2$ is $(0.0257, 0.0697)$, which suggests that the positive price difference between national brand prices three quarters after an introduction and right before an introduction is statistically different from zero.

2. Manufacturers set wholesale prices

3. Given wholesale prices and demand shocks, retailers set retail prices

4. Given retail prices, consumers make purchase (or not) decisions

Within each subgame, I solve the model by backwards induction. As such, I will first describe the demand system capturing consumer behavior and then the supply model of retailer competition and manufacturer competition, respectively.

1.5.1 Demand

In each geographic market d, in each period t, consumers face a choice set of available yogurt product lines, J_{dt}. As described in Section 1.3, I define a product line as a unique combination of brand $b \in \{1, ..., B_{dt}\}$ and observable non-price binary attributes. In addition to product lines observed in the data, consumers may choose the outside option, which I index as $j = 0$. Additionally, I assume that consumers only purchase one product line at a time, abstracting from multiple discrete choice.

Household i's latent utility in DMA d at time t from purchasing and consuming product line j from retailer r is given by

$$V_{ijrdt} = \alpha_i p_{jrdt} + x_j \beta_i + \xi_{jrdt} + \epsilon_{ijrdt} \tag{1.8}$$

The retail price of product line j at retailer r in DMA d in time t faced by consumers is given by p_{jrdt}.[11] The observable product attributes of product line j is given by $x_j \equiv (x_{j,1}, .., x_{j,K})$, which is invariant to retailer, market, and time.

These observable product characteristics are scaled by random coefficients, representing tastes that are allowed to differ between households. The random coefficient α_i is the households (dis)taste of price. Similarly, the vector $\beta_i \equiv (\beta_{i,1}, ..., \beta_{i,k})'$ represents the household's taste for each observable product characteristic $k \in \{1, ..., K\}$. These random coefficients

[11] Retail prices for the same manufacturer's product are allowed to vary between two different retailers, a pattern which is observed in the data.

allow the demand system to capture heterogeneity between households. I decompose these random coefficients into a mean and variance term, namely,

$$\begin{pmatrix} \alpha_i \\ \beta_{i,k} \end{pmatrix} = \begin{pmatrix} \alpha + \sigma_\alpha \nu_{\alpha,i} \\ \beta_k + \sigma_{\beta_k} \nu_{\beta_k,i} \end{pmatrix}$$

for $k \in \{1, ..., K\}$. The mean of the random coefficient terms, α and $\beta \equiv (\beta_1, ..., \beta_k)'$, represent the average taste across households for price and product attributes, respectively. The random variables $\nu_{\alpha,i}$ and $\nu_{\beta_k,i}$ follow standard normal distributions and are scaled by the (positive) coefficients σ_α and σ_{β_k} to allow for differences in tastes between households. The joint distribution of (α_i, β_i) is represented by F.

The utility specification also includes a demand shifter, ξ_{jrdt}, which is unobserved by the econometrician but observed by firms. I decompose the demand shifter into retailer, time, and DMA-brand fixed effects and a residual,

$$\xi_{jrdt} \equiv \xi_r + \xi_t + \xi_{d,brand_j} + \Delta\xi_{jrdt} \tag{1.9}$$

This decomposition allows for tastes of the same yogurt products to flexibly vary between retailers and between time periods. Further, by including a DMA-brand fixed effect, I allow for consumers in one geographic market to have different mean preferences over a given brand of yogurt product lines than consumers in other geographic markets. This accounts for regional preferences for specific brands, for example, if customers in West Coast markets have stronger preference for health-oriented brands.

Finally, the demand specification includes an additively separable household-product-retailer-DMA-quarter taste shifter, ϵ_{ijrdt}, which follows a Type I extreme value distribution. As an assumption, I normalize the utility of the outside good to $V_{i0rdt} = \epsilon_{i0rdt}$.

As is standard in the literature, I assume the consumer chooses the single product line or outside good with the highest utility from the set of available product line choices. Product line market shares can then be inverted as,

$$s_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}, \xi) = \int \frac{\exp\{\alpha_i p_{jrdt} + x_j \beta_i + \xi_{jrdt}\}}{1 + \sum_{k \in J_{dt}} \exp\{\alpha_i p_{krdt} + x_k \beta_i + \xi_{krdt}\}} dF(\alpha_i, \beta_i) \tag{1.10}$$

Where retail prices for product lines in DMA-market dt are collected into the vector $\mathbf{p}_{dt}$. Similarly, vectors $\mathbf{x}$ and ξ collect observable and unobservable product characteristics, respectively.

As will be described in Section 1.6, equation (1.10) is matched to observed market shares to estimate the demand parameters that govern the household's utility function.

1.5.2 Supply

I next turn to the supply model, which details how of manufacturers and retailers set upstream and downstream prices, respectively. Since retailers take wholesale prices as given, I first solve the downstream retailer's problem, conditional on upstream prices, and then move to the upstream manufacturer's problem.

Retailer

In geographic market d and time period t, retailer r's variable profits are given by,

$$\pi_{rdt} = \sum_{j \in J_{rdt}} M_{dt} s_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}; \theta)(p_{jrdt} - c_{jrdt} - p^w_{jrdt}) \tag{1.11}$$

Where, for product j at time t in market d, p_{jrdt} is the per-unit retail price paid by consumers to retailer r, c_{jrdt} is the per-unit cost of retailing, and p^w_{jrdt} is the wholesale per-unit price paid by retailer r. The potential size of market d at time t is given by M_{dt}.

For store brands, I assume no wholesale markup and thus set $p^w_{jrdt} = c^w$ for $j \in J^{store}_{rdt}$. Though I do not observe wholesale prices for store brand products to empirically evaluate this assumption, it is reasonable assumption given evidence that large retailers do source store brands through their own production capabilities. With full vertical integration, retailers internalize the upstream markup and thus optimally set $p^w_{jrdt} = c^w$. Alternatively, when retailers opt for arm's-length sourcing of store brands, particularly in the yogurt product category, retailers solicit bids from multiple manufacturers who compete on price to produce a store brand product with given attributes specified by the retailer. This competition in wholesale price between manufacturers, many "private label manufacturers" with limited market power, over a standardized offering drives wholesale prices close to, if not equal to, production costs, c^w, for store brand yogurt products.

Conditional on wholesale prices, retailers compete a la Nash-Bertrand in retail prices. I assume a pure-strategy equilibrium, in which retailers follow the first order condition of the retailer's profit function given in equation (1.11) with respect to prices,

$$s_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}; \theta) + \sum_{j' \in J_{rdt}} \frac{\partial s_{j'rdt}(\mathbf{p}_{dt}, \mathbf{x}; \theta)}{\partial p_{jrdt}} (p_{jrdt} - c_{jrdt} - p^w_{jrdt}) = 0 \tag{1.12}$$

Stacking and rearranging terms,

$$\mathbf{p} = \mathbf{c} + \mathbf{p}^w + (-\mathbf{O}^r * \Delta^r)^{-1} \mathbf{s}(\mathbf{p}, \mathbf{x}; \theta) \tag{1.13}$$

Where the vector of retail prices is given by $\mathbf{p}$, wholesale prices by $\mathbf{p^w}$, retailer marginal costs by $\mathbf{c}$, retailer ownership $\mathbf{O}^r(j, j') \equiv 1\{j, j' \in J_{rdt}\}$, and the partial derivative matrix of market shares with respect to prices is given by $\Delta^r(j, j') \equiv \frac{\partial s_{j'r}}{\partial p_{jr}}$.

Manufacturer

We next solve the manufacturer's problem. Similar to retailers, manufacturers compete Nash-Bertrand in wholesale prices. However, given the timing assumption that manufacturers set wholesale prices prior to retailers, manufacturers take into account the response of retail prices when setting wholesale prices. Specifically, manufacturers know retail price setting will observe equation (1.13). Thus, we can write manufacturer m's problem as maximizing the sum of product-specific profits, given markup over marginal cost, subject to equation (1.13). Formally,

$$\max_{\{p^w_{jrdt}\}_{j\in J_{mdt}}} \sum_{j\in J_{mdt}} s_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}; \theta)(p^w_{jrdt} - c^w_{jrdt}) \tag{1.14}$$

$$\text{subject to} \quad \mathbf{p} = \mathbf{c} + \mathbf{p}^w + (-\mathbf{O}^r * \Delta^r)^{-1}\mathbf{s}(\mathbf{p}, \mathbf{x}; \theta) \tag{1.15}$$

Manufacturer first order conditions with respect to wholesale prices are given by,

$$\mathbf{s}(\mathbf{p}(\mathbf{p}^w)) + (O^m * \Delta^m)(\mathbf{p}^w - \mathbf{c}) = 0 \tag{1.16}$$

where retail prices $\mathbf{p}$ are implicitly a function of wholesale prices $\mathbf{p}^w$ through retailer first order conditions (1.13). The matrix that captures manufacturer product ownership structure, O^m, with elements $O^m(j, j') \equiv 1\{j, j' \in J_{mdt}\}$, is element-wise multiplied ("$*$") by the response matrix Δ^m, which captures how market shares change with respect to to wholesale prices. The elements of the response matrix are defined using the chain rule as,

$$\Delta^m(j, j') = \sum_r \sum_{j''r} \frac{\partial s_{j'r}}{\partial p_{j''r}} \frac{\partial p_{j''r}}{\partial p^w_{jr}} \tag{1.17}$$

Where the partial derivative of retail prices with respect to wholesale prices (i.e., the second partial derivative in the sum)[12] is determined implicitly from (1.13), as in Villas-Boas (2007).

Marginal Costs

Combining the retailer and manufacturer first order conditions (1.13) and (1.16) gives total marginal costs as observed prices less the total vertical markup implied by the model,

$$\mathbf{mc} \equiv \mathbf{c} + \mathbf{c}^w = \mathbf{p} - \mu^r(p, x, \xi; \theta) - \mu^m(p, x, \xi; \theta) \tag{1.18}$$

[12] For $j = j'$, this derivative gives a measure of how a product line's wholesale price is passed through to its retail price.

where retailer and manufacturer markups, respectively, are given by,

$$\mu^r(p, x, \xi; \theta) \equiv (-O^r * \Delta^r)^{-1}\mathbf{s}(\mathbf{p}, \mathbf{x}; \theta) \tag{1.19}$$

$$\mu^m(p, x, \xi; \theta) \equiv (-O^m * \Delta^m)^{-1}\mathbf{s}(\mathbf{p}, \mathbf{x}; \theta) \tag{1.20}$$

From this expression, we see that total marginal costs are a function of observed retail prices, $\mathbf{p}$, and markup terms, which depend on market shares, $\mathbf{s}$, ownership structure, O^r and O^m, and response matrices implied by the demand model and demand parameters, θ. Thus it follows that, once demand parameters are estimated, the vector of total marginal cost can be recovered and taken as a structural object for counterfactual analysis. The next section describes this process in detail.

1.6 Estimation and Identification

Estimation of the empirical model involves first estimating demand and using these demand estimates, in conjunction with the supply model, to recover marginal costs. The demand parameters that govern consumer utility are estimated following the methods of estimating random coefficient logit demand systems due to Berry, Levinsohn, and Pakes (1995). Marginal costs are then determined as the difference between observed prices and optimal markups implied by estimated demand parameters and the observed market structural.

1.6.1 Demand Parameters

Estimation Procedure

Estimation of the parameters of the demand model follows the generalized method of moments ("GMM") procedure of Berry, Levinsohn, and Pakes (1995).

For a guess of σ, I first determine the mean utilities, $\delta(\sigma)$, that match observed (i.e., data) and model-implied shares using the contraction mapping from Berry (1994). Then, following Nevo (2001), I recover estimates of the mean taste and fixed effect parameters, $(\hat{\alpha}, \hat{\beta}, \hat{\xi})$, as the first order condition of regressing the given $\delta(\sigma)$ on these parameters. The estimating equation is,

$$\delta_{jrdt}(\sigma) = \alpha p_{jrdt} + x_j\beta + \xi_r + \xi_t + \xi_{d,brand_j} + \Delta\xi_{jrdt} \tag{1.21}$$

This "concentrating out" of mean taste and fixed effect parameters in the inner loop of the estimation routine allows me to construct demand shocks implied by the candidate value of σ as the residual vector, $\Delta\xi(\sigma)$. In the outer loop of the estimation procedure, I search for the value of the random coefficient variance parameters, $\sigma \equiv (\sigma_\alpha, \sigma_\beta)$, which minimizes the GMM objective function,

$$\hat{\sigma} = \arg\min_{\sigma} \; \Delta\xi(\sigma)' Z W Z' \Delta\xi(\sigma) \tag{1.22}$$

Where Z is a vector of instruments, as detailed below, and W is the optimal weight matrix.[13]

Instrument Construction

Given the assumption that retailers observed demand shocks prior to setting prices, I must address the issue of price endogeneity when estimating demand parameters. To do so, I employ two sets of instrumental variables: input costs and measures of product differentiation.

To construct instruments based on input costs, I first collect publicly available information on the geographic location of production facilities of brands in my data. For brands associated with multiple production facilities, I assume the brand's product lines were produced at the production facility closest to DMA in which the product line was sold.[14] Understanding where a given product line sold in a given market-quarter was produced allows me to proxy for contemporaneous input prices. Namely, I use state-level average wages for food produc-

[13] To construct the optimal weight matrix, I use a two-step estimation procedure that first recovers an estimate of σ, call it $\tilde{\sigma}_1$, using a first step weight matrix, $W_1 = (Z'Z)^{-1}$. I then plug in the demand shocks implied by $\tilde{\sigma}_1$ from this first step estimate to construct the optimal weight matrix, $W_2 = (Z'\Delta\xi(\tilde{\sigma}_1)\Delta\xi(\tilde{\sigma}_1)'Z)^{-1}$. To recover final parameter estimates, I re-run the estimation procedure described above substituting W_2 in the GMM objective function in (1.22).

[14] For store brands, I assign the state of the nearest observed national brand production facility.

tion workers from the US Census Annual Survey of Manufacturers and average industrial electricity prices from the US Energy Information Association. The validity of these input cost instruments to identify the endogenous regressor, α, relies on two conditions: i) input costs are uncorrelated with demand shocks, $\Delta\xi_{jrdt}$; and ii) pass-through of input costs to retail prices is nonzero. These conditions appear reasonable in my empirical context.

Additionally, I construct instruments measuring a product line's degree of differentiation in the product attribute space. Namely, for each non-linear product characteristic, I determine the number of other product lines that share a product line's product characteristic in a given market-period. As suggested by Ghandi and Houde (2019), this instrument measures the degree to which a given product line is isolated in the product space. Given that equilibrium retail prices depend on the competitive environment, specifically the prevalence of close substitutes, these instruments are correlated with observed retail prices. Additionally, it is reasonable to assume that a product's demand shock, particularly after controlling for a rich set of fixed effects, is uncorrelated with the number of "nearby" products in the product attribute space.

1.6.2 Marginal Costs

In order to calculate equilibrium prices under counterfactual scenarios, I must recover unobserved total vertical marginal costs under the factual environment. To do so, I rely on the assumptions in the supply model to "back out" marginal costs as the difference between prices and the total vertical markup implied by the model and demand estimations. From the combined retailer and manufacturer first order conditions given in equation (1.18), marginal costs estimates are given by,

$$\hat{mc}_{jrdt} = p_{jrdt} - \mu^r_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}_{dt}, \xi_{dt}; \hat{\theta}) - \mu^m_{jrdt}(\mathbf{p}_{dt}, \mathbf{x}_{dt}, \xi_{dt}; \hat{\theta}) \tag{1.23}$$

As can be seen in equation (1.19), the retailer markup term, $\mu^r(\cdot)$ consist of three components: retailer product ownership, the response matrix, and product line market shares.

Ownership and market shares are taken directly from the data. The elements of the response matrix, Δ^r, which are derivatives of market shares with respect to prices, are implied by the supply model. In practice, I calculate these derivatives by simulating over household draws over the estimated distribution of demand parameters.

Similarly, the manufacturer markup term, as shown in equation (1.20), is a function of the manufacturer product ownership, response matrix, and market shares. Again, I take factual market shares and observed ownership. Here, since the vertical structure of store brands is assumed to be equivalent to full integration, upstream markups on store brands product lines are zero and are thus omitted from the calculation of manufacturer markups. The response matrix, Δ^m, as shown in equation (1.17), expresses the response of market shares to wholesale prices (as opposed to retail prices). This object, as discussed in Villas-Boas (2007), requires the calculation of the Hessian of market shares with respect to prices. To calculate this term, I again rely simulation based on the empirical distribution implied by demand estimates.

1.7 Estimation Results

1.7.1 Demand Estimates

Demand estimates are included in Table 1.2. The estimate of the mean coefficient on price, $\hat{\alpha} = -8.89$, is the expected sign, indicating consumers' disutility of price. Further, the estimate of the random coefficient on price, $\hat{\sigma}_\alpha = 0.61$ is statistically different than zero, implying consumer heterogeneity in (dis)taste for price. The estimates for non-price observable characteristic imply, on average, consumers do not prefer Greek yogurt product lines and consumers prefer full fat (omitted group) to non-fat and low-fat product lines. However, consumers exihibit heterogeneity is these preferences, such that there is a non-trivial mass of consumers that prefer Greek to non-Greek product lines, all else equal (i.e., for draws of the normally distributed random variable $\nu_{\beta_{\text{Greek}}} > 1.42/4.12 = 0.34$). Similarly, the parameter estimates imply a non-zero mass of consumers prefer non-fat and low-fat product lines.

As a check, I calculate the own price elasticities for each DMA-retailer-product line in my sample at the demand parameter point estimates. The median product line own price elasticity is -8.80, which is slightly less than previous papers that study the same industry. However, these demand estimates do not generate any product line price elasticities which are greater than zero.

As I include DMA-brand coefficients in my demand specification, estimation of demand does not admit a single parameter that captures the taste for store brands. Instead, these DMA-brand coefficients flexibly capture the taste for individual store (and national) brands in each geographic market in my sample.[15] Therefore, to gain an aggregate sense of consumer preferences for store brands, I plot the estimate of each of these fixed effects. Figure 1.4 shows a histogram of these DMA-brand coefficients, grouped by store and national brands.

The average of national brand DMA-brand fixed effects is larger than that of store brands. This conforms to industry knowledge that, all else equal, consumers generally prefer national brands to store brands. However, the dispersion of fixed effects within each group indicate heterogeneity in taste for both different national brands and store brands. Additionally, some store brand-DMA fixed effects are greater than some national brand-DMA fixed effects.

1.7.2 Marginal Cost Estimates

Once demand parameters are estimated, total vertical marginal costs (i.e., combined marginal cost of production and retailing) can be calculated from equation (1.23) (Section 1.6.2) as observed prices less retailer and manufacturer markups. I calculate retailer and manufacturer markups for each product line in my sample by evaluating equations (1.19) and (1.20) at the demand parameter point estimates. The median manufacturer markup is \$0.12 and the median retailer markup is \$0.12.[16]

Most strikingly, the marginal cost estimates differ between Greek and non-Greek yogurt

[15] Within a DMA, I include a different fixed effect for each retailer store brand product lines to allow for different average preferences over between retailers. For example, in DMA 1, non-fat Greek store brand yogurt at retailer A has a different fixed effect than non-fat Greek store brand yogurt at retailer B

[16] By assumption, manufacturer markups on store brand product lines are zero.

products, as plotted in Figure 1.5. Given the production of Greek yogurt requires an additional step of straining liquid whey from regular yogurt, we would expect the marginal cost of production for Greek yogurt product lines to be more than non-Greek. As Figure 1.5 shows, this pattern is borne out in the marginal cost estimates – marginal costs for Greek yogurt product lines are, on average, larger than non-Greek.

1.8 Counterfactual Simulations

I use the structural model, as well as the model's estimated demand and cost parameters, to simulate consumer, retailer, and manufacturer responses to alternative scenarios and the resulting market outcomes to directly assess the equilibrium and welfare implications of store brands in the US yogurt industry.

The first counterfactual assesses the equilibrium and welfare effects of store brands under the observed product line ownership structure. I conduct a counterfactual simulation in which I eliminate all store brand yogurt product lines from each retailer's offerings in my sample. Under this environment, for the remaining set of products, which consist only of national branded products sold by retailers under the factual scenario, I determine equilibrium manufacturer markups that follow first order conditions given in (1.16) and retail markups of remaining products that following equation (1.13). Equilibrium retail prices, which consist of the total vertical markup over estimated marginal costs, along with product characteristics and estimated taste parameters, determine product market shares from equation (1.10). Recovering counterfactual markups and product line shares allows me to calculate manufacturer and retailer profits. Changes in consumer surplus between the counterfactual and observed environment are calculated by compensating variation as,

$$\Delta CS = \sum_{d} \sum_{t} \int \frac{1}{|\alpha_i|} \left[\log \left(\sum_{j'=1}^{J_{dt}^1} \exp(V_{ij'rdt}^1) \right) - \log \left(\sum_{j=1}^{J_{dt}^0} \exp(V_{ijrdt}^0) \right) \right] dF(\alpha_i, \beta_i) \quad (1.24)$$

where V_{ijrdt}^1 equals the utility specified in equation (1.8), evaluated at counterfactual prices, minus the idiosyncratic taste shifter, ϵ_{ijrdt}. Similarly, V_{ijrdt}^0 represents the same utility term evaluated at observed factual prices. The set of product lines available to consumers in DMA d and quarter t is given by J^1_{dt} in the counterfactual and J^0_{dt} in the observed factual.

The second set of counterfactual simulations assess the equilibrium and welfare effects of store brands under a market structure with higher concentration, namely, after a merger between upstream manufacturers. To isolate the effect of store brands under higher upstream market concentration, I first simulate market outcomes under a horizontal merger between the two largest manufacturers in my sample with the factual product set (i.e., with store brands). Second, I simulate market outcomes after this same merger *and* removing store brands.

Table 1.3 shows the resulting changes in welfare for manufacturers, retailers, and consumers from these counterfactual simulations. The first column of results shows how manufacturer, retailer, and consumer surplus measures change when moving from the factual world with both national and store brand product lines to a counterfactual scenario where all retailer store brands are removed. After prices for the remaining national brands are recalculated for the remaining national brands, manufacturer and retailer profits are determined, as well as (change in) consumer surplus. Manufacturers benefit from the removal of store brands, increasing aggregate profits by \$151.42 million, while retailers, in aggregate, are worse off, with profits decreasing by \$203.91 million. However, retailers that do not initially have store brands benefit from the elimination of their competitors' store brands. More so, we see that retailers for whom store brand product lines make up a smaller proportion of their offerings ("low store brand usage retailers") are not as worse off as those retailers with higher proportions of store brands ("high store brand usage retailers").[17]

In this counterfactual without store brands, consumers are worse off than the factual

[17] The three subcategories of retailers are mutually exclusive and mutually exhaustive. "No store brand usage retailers" are observations for which a retailer-DMA-quarter does not exhibit any store brand product lines in their offerings. For the remaining retailer-DMA-quarter observations, I calculate the proportion of a retailer-DMA-quarter's product lines that are store brand and take the conditional (on having one or more store brand) median across my sample, which equals 12.5% of a retailer-DMA-quarter's product lines. "Low store brand usage retailers" are those which have a store brand proportion between zero and the median, 12.5%. "High store brand usage retailers" are the remaining retailer-DMA-quarter observations which have greater than 12.5% of product lines as store brands.

world with store brands. Without store brands, retail prices for national brands increase. Additionally, by removing store brands, consumers mechanically face fewer available product lines to choose from. The reduction in consumer surplus of $935.47 million over the entire sample is a result of the combination of these factors – higher prices and fewer available yogurt products lines. Summing over the entire vertical channel, the change in total surplus due to removing store brands is equal to -$988.47 million.[18]

I run two additional counterfactuals to assess the how the welfare effects of store brands change based on upstream concentration. To provide a baseline, I conduct a merger simulation between the two largest manufacturers in my sample, in terms of yogurt revenue, with factual product sets (i.e., with store and national brands). Under this scenario, the results for which are shown in column (2) of Table 1.3, both merging and non-merging manufacturers benefit to a combined increase of $40.31 million in surplus. This benefit comes at the expense of both retailer and consumer surplus, losing $93.50 million and $444.51 million in surplus, respectively.

I re-run this same merger simulation, now removing store brands. The resulting surplus changes, relative to the factual environment, for this counterfactual exercise are shown in column (3) of Table 1.3. Unsurprisingly, manufacturers, both the merged entity and non-merging firms, realize the largest increase in profits relative to the first two counterfactuals. Consumers and retailers that sell store brands see the largest decrease in respective surplus measures. Retailers that do not sell store brands are slightly worse off post-merger and without store brands relative to the observed factual, but not as worse off than post-merger with store brands.

To assess how the welfare effects of store brands change under increased market concentration, I compare differences in the reported changes in surplus between the factual and the observed upstream structure with store brands (column (1)) and between post-merger with store brands and without store brands (columns (2) and (3)). For manufacturers, post-

[18]This change in total surplus represents approximately 5% of industry revenues in the U.S. over the three year period, 2011-2013, for which this counterfactual simulations are conducted.

merger profits increase more without store brands present for both sets of manufacturers. For example, for the merging manufacturers, moving from the factual to the pre-merger, no store brand counterfactual yielded \$111.58 million of additional profits. Without store brands, post-merger profits for the merging firms would have increased by more, specifically by \$116.05 million (=\$140.92 million − \$24.87 million). Conversely, retailers that initially exhibited store brand product lines would exhibit a larger decrease in profits when removing store brands under higher upstream concentration relative to the observed upstream market structure. Specifically, low usage retailers would exhibit a reduction in profits of \$88.91 million (relative to \$75.55 million without merger) and high usage retailers would realize \$149.78 million less profits (relative to \$132.01 million). Retailers that do not exhibit store brands in the factual realize slightly less gain in surplus under higher upstream market concentration due to the removal of store brands (\$3.36 million) relative to without an upstream merger (\$3.64 million). Finally, consumer surplus, and resultingly, overall welfare, decreases more when removing store brand product lines under higher upstream market concentration.

Alternatively, we can use these counterfactual results to assess how the welfare effects from an upstream merger change when store brands are present versus when they are not. This welfare comparison is then between i) the pre-merger with store brands (factual) scenario and the post-merger with store brands counterfactual (column (2)) and ii) pre-merger without store brands (column (1)) and post-merger without store brands (column (3)). In assessing changes in profits from the merger, both merging and non-merging manufacturers are better off without store brands, collectively gaining \$54.09 million (=\$205.51 million − \$151.42 million) in profit from pre- to post-merger without store brands, relative to \$40.31 million if the merger occurs with store brands in the market. Conversely, the welfare effects of the upstream merger for retailers and consumers are worse without store brands. In aggregate, retailers (consumers) would exhibit a decrease in profits (surplus) of \$124.93 million (\$582.92 million) if the merger occurred without store brands compared to a smaller decrease of \$93.5 million (\$391.32 million) with store brands. Resultingly, total surplus decreases less with store brands ($444.51 million) relative to the decrease in total surplus due to the upstream merger occurring without store brands ($653.76 million).

In summary, the counterfactual exercises showed that i) the existence of store brands shifts surplus downstream, from manufacturers to retailers and consumers; ii) this effect is stronger under increased upstream concentration; and iii) store brands act as a countervailing force from negative welfare effects of increased upstream concentration, such as a horizontal merger.

1.9 Conclusion

In this chapter, I theoretically motivate and conduct two empirical exercises to evaluate the effect of store brands on equilibrium outcomes. Using theoretical simulation, I show how store brands can have ambiguous welfare effects, motiving two empirical exercises. Using an event study framework, I show that retailer store brand introductions of yogurt are associated with an increase in both the number of and the prices of national brand product lines. To evaluate the causal effect of store brand introductions, I develop a structural model and fit it to scanner data to recover structural demand and cost parameters. I then evaluate counterfactual scenarios, removing store brands under both the observed upstream ownership structure and under a simulated merger between two larger manufacturers. The counterfactual exercise suggests the existence of store brands is welfare enhancing, benefiting consumers and (most) retailers, while reducing profits for manufacturers. These effects are shown to be exacerbated if two manufacturers merge, suggesting store brands can act as a countervailing force to anticompetitive merger effects and yield greater total welfare improvement in more concentrated industries.

Table 1.1: Sample Descriptive Statistics

	Mean	Standard Deviation	Minimum	Maximum
Characteristics:				
Price ($ per 6 oz.)	1.00	0.46	0.002	24.99
Binary characteristics:				
Greek	0.34	0.47	0	1
Non-fat	0.35	0.48	0	1
Low-fat	0.34	0.47	0	1
Full-fat	0.31	0.46	0	1
Store brand	0.11	0.31	0	1
Product lines:				
Per DMA-retailer-quarter	20.96	15.11	1	69
Per DMA-quarter	45.50	21.36	1	122
Total brands	42	-	-	-
Total product lines	145	-	-	-
Retailers:				
Per DMA-quarter	4.69	2.46	1	16
Total retailers	146	-	-	-
Manufacturers:				
Per DMA-retailer-quarter	5.97	3.38	1	17
Per DMA-quarter	10.04	3.52	1	21
Total manufacturers	31	-	-	-
Markets:				
Total DMAs	205	-	-	-
Total quarters	44	-	-	-
Total DMA-quarters	8,937	-	-	-
Total observations	878,678	-	-	-

Table 1.2: Demand Estimates

	Mean (α, β)	Random Coefficient (σ)
Constant	0.31	0.18
	(1.54)	(0.38)
Price	-8.89	0.61
	(0.24)	(0.04)
Greek	-1.43	4.12
	(0.08)	(0.04)
Non-fat	-0.15	0.71
	(0.02)	(0.08)
Low-fat	-0.62	1.51
	(0.04)	(0.08)
Median own-elasticity	-8.80	
% own-elasticity > 0	0.00	

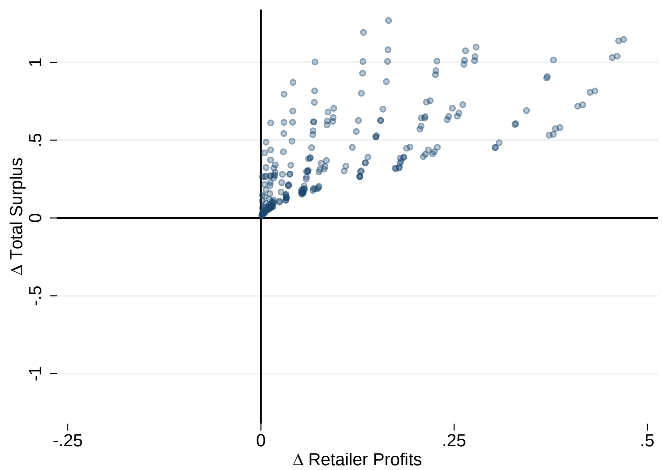
Note: Estimation via 2 step GMM with cost shifters (industrial wages and plant-state electricity prices) and differentiation instruments (number of other products in market sharing characteristics) using aggregated quarterly scanner data for grocery yogurt product line sales in 178 DMAs from 2011Q1-2013Q4 (n = 262,297). 100 households simulated per DMA-period. Specification includes retailer, time, and brand-DMA fixed effects. Standard errors reported below each estimate in parenthesis.

Table 1.3: Counterfactual Simulation Results

	(1)	(2)	(3)
		Upstream Merger	
Δ Surplus ($ million)	No Store Brand	Store Brand	No Store Brand
Δ Manufacturer Profits			
Merging manufacturers	111.58	24.87	140.92
Non-merging manufacturers	39.84	15.44	64.59
All manufacturers	151.42	40.31	205.51
Δ Retailer Profits			
No store brand usage retailers	3.64	-4.43	-1.07
Low store brand usage retailers	-75.55	-49.02	-137.93
High store brand usage retailers	-132.01	-40.06	-189.84
All retailers	-203.91	-93.50	-328.84
Δ Consumer Surplus			
All consumers	-935.98	-391.32	-1,518.90
Δ Total Surplus	-988.47	-444.51	-1,642.23

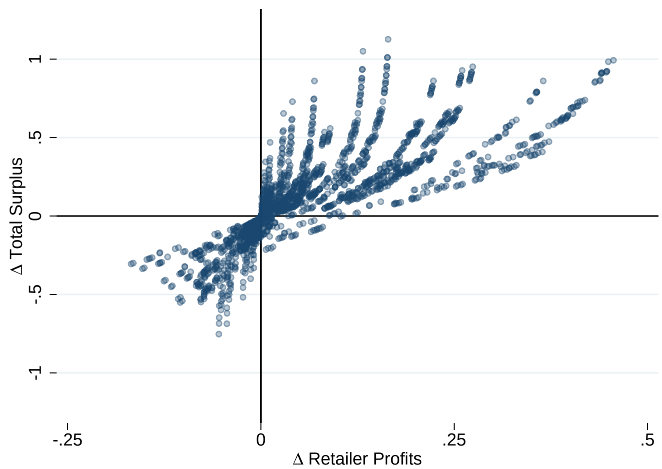
Note: Surplus measures based on simulated counterfactuals using demand parameter and marginal cost estimates for retailer-product lines in 184 DMAs from 2011Q1–2013Q4. Upstream merger simulated between the two largest manufacturers based on observed revenues during the sample period. High store brand usage retailers are those with the proportion of store brand products lines greater than 12.5%, low store brand usage greater than zero and less than or equal to 12.5%, and no store brand usage retailers exhibit 0% store brands in a given retailer-DMA-quarter.

Figure 1.1: A1 to B: adding a store brand



Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A1 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Total surplus is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1,2,3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1, 2, 3$. Equilibria associated with duplicative vectors with variation only in parameters pertaining to product $j = 2$, which do not impact the comparative static exercise between Scenarios A1 and B, are removed for visual clarity.

Figure 1.2: A2 to B: a store brand replacing a national brand



Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A2 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Total surplus is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1, 2, 3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1, 2, 3.$

Figure 1.3: Average Percentage Change in Retail Price of National Brand Product Lines around Store Brand Product Line Introductions

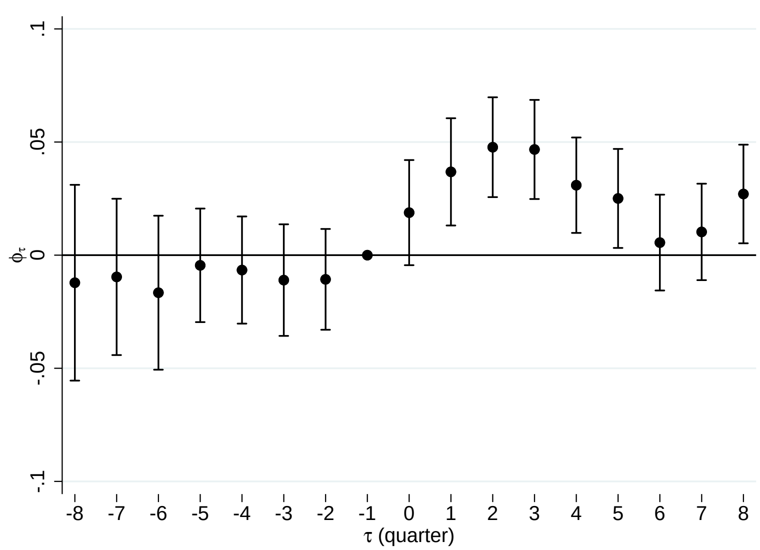


Figure 1.4: DMA-Brand Coefficients by Store and National Brands

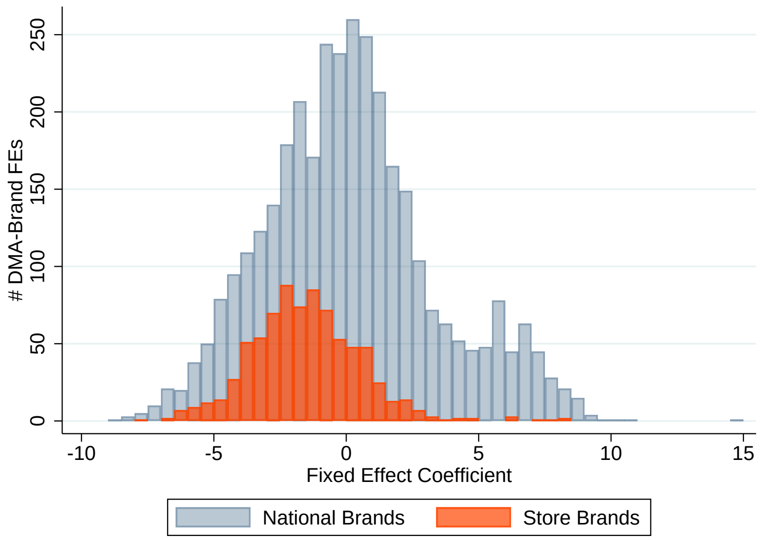
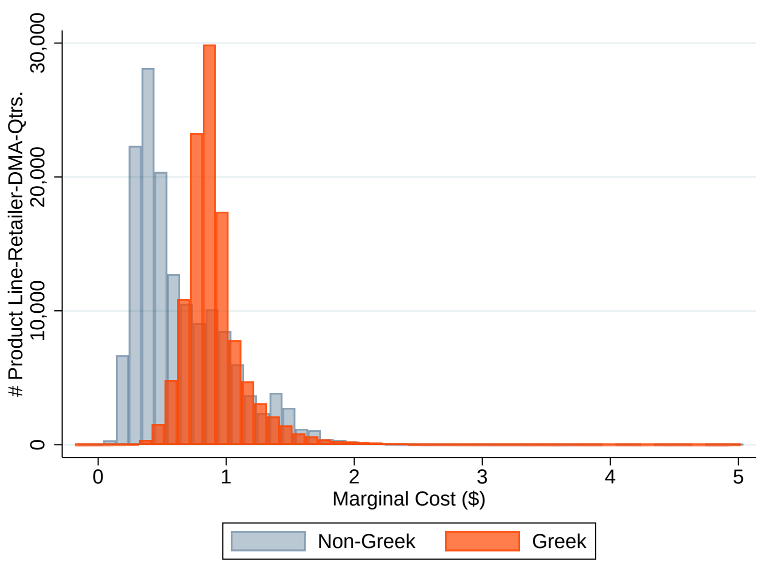


Figure 1.5: Marginal Cost Estimates by Greek and Non-Greek Product Lines



Chapter 2

Patent Licensing and Bias in Estimation and Prediction (with Xuan Teng)

2.1 Introduction

In empirical analyses of markets using structural models of pricing competition, high prices can be explained by high marginal costs, given market concentration and consumer preferences. In this work, we examine a typically unobserved and commonly abstracted-away market feature that additionally contributes to high prices: patent licensing relationships between manufacturers.

Firms can exercise intellectual property rights by implementing technologies in producing goods and services and licensing these rights to other firms, including competitors. By licensing, a firm can recoup private research and development costs, which can incentivize *ex ante* innovation efforts, while publicly disseminating beneficial technologies in the marketplace. In many markets, firms license technologies to competitors[1] in exchange for payments that scale with the quantity sold of the competitors products that implement the licensed technology.[2] Under this payment structure, the incentives of the licensor and its competitors

[1]For example, in the cellphone market, Apple licenses scroll feature to IBM and Nokia. In the electric vehicles market, Toyota licenses hybrid technology to Daimler (Mercedes Benz). In the televisions market, the ATSC patent pool, whose licensors include Panasonic, LG and Samsung, licenses the package of essential patents to not only themselves but also competitors like Sharp (see: https://www.mpegla.com/programs/atsc/licensors/).

[2]Royalty revenue may also be sales-based, i.e., $r_f Q_l p_l$ We follow the theoretical papers in this literature to focus on per-unit royalties (for example, see Reisinger and Tarantino (2019), Layne-Farrar and Lerner (2010), Lerner and Tirole (2005)).

become partially aligned: the licensor does not want to price cut their competitors too much, so as to ensure fruitful royalty revenues. Because prices are strategic complements, the licensors' competitors won't charge low prices neither. Therefore, such alignment effects due to the patent licensing relationships softens pricing competition, resulting in higher prices.

When estimating an empirical model of price competition, if a researcher observes who licenses patents to whom at which royalty rates, the above alignment effect can be directly accounted for and estimates of marginal costs can be accurately obtained. Unfortunately, the terms and even the presence of intellectual licensing agreements are often kept confidential between licensors and licensees and are not publicly observable. If a researcher proceeds to estimate a supply model as if these agreements, and thus the associated alignment of incentives, were not actually present, the resulting estimates will be biased in theory. Therefore, unobserved patent licensing relationships introduce mis-specification errors into estimation of marginal costs.

Further, the bias in the marginal cost estimates will propagate through any analysis of the pricing competition, leading to biased predictions of counterfactual market outcomes. This is because the bias in marginal cost estimates are generally not constant between status-quo and counterfactual scenarios, and therefore cannot be cancelled out when evaluating the effect of market structure changes. The changes in bias of marginal cost estimates due to not accounting for licensing between status-quo and counterfactual scenarios root in the magnitudes of the alignment effect. For example, how much a licensor cares about its competitors product sales depends on the elasticity of its competitors product sales with respect to its own price. If its competitors product sales are unaffected by its own price, the licensor has no need to charge lower prices to secure royalty revenues. Since price elasticities are market outcomes that depend on many aspects of the market structure, so does the alignment effect. Moreover, the structural difference between pricing competition models with and without patent licensing relationships naturally leads to different predictions of market outcomes.

The common existence in market analyses of counterfactual prediction bias due to unobserved existing patent licensing relationships, along with the typical difficulty of collecting patent licensing data, motivates us to ask and answer the following questions: Does the unobserved licensing behavior always lead to positive or negative bias in marginal cost estimates and subsequently the evaluation of counterfactual scenarios? Under what market conditions and which sorts of counterfactual structural changes would we larger prediction biases? When would we expect opposite predicted changes of market outcomes from models with and without patent licensing?

To answer these questions, we first explain the alignment effects and the resulting bias in estimation and predictions in a theoretical framework. Then, we specify a model of pricing competition with patent licensing relationships between competing manufacturers. We calibrate this model and conduct a series of simulation exercises to assess what features of a market, including royalty rates, product substitution patterns, and market concentration, lead to larger biases in estimation and counterfactual predictions when existing patent licensing relationships are unobserved and assumed-away. We consider two types of counterfactuals: a merger between licensor and licensee and a merger between licensees. This enables us to examine which type of counterfactuals are likely to have larger prediction biases. We use two sets of simulation exercises for two complementary purposes: theoretical simulations for examining how biases change with respect to underlying primitives and guidance simulations for showing, for what types of markets and their respective market characteristics would an analyst be most concerned with biases due to unobserved patent licensing.

We find that estimation bias in marginal costs are increasing with royalty rates and the sum of diversion ratios between licensor and licensees. The guidance simulations show that such bias is typically not ignorable: in a sample of reasonably generated markets, the median estimation bias is 28% of the true marginal cost. We show an alarming finding to researchers: not accounting for existing patent licensing relationships can lead to opposite prediction of merger effects. Such opposite predictions show up in both theoretical simulations and guidance simulations for both licensor-licensee mergers and licensee-licensee mergers. Because the guidance simulations cover a richer set of markets (including asymmetric markets) compared to the theoretical simulations that are based on a symmetric benchmark market, opposite predictions show up more frequently in guidance simulations. In particular, in the theoretical simulations, we find 4.3% to 4.4% of cases where the merger effects on prices predicted from a mis-specified model not accounting for existing patent licensing relationships are opposite to the true merger effects predicted from a true model accounting for existing patent licensing relationships. In the guidance simulations, such fraction of opposite prediction cases can be as high as 51.2%. Moreover, both sets of simulations find that, opposite predictions happen when royalty rates are large. In particular, theoretical simulation finds that opposite predictions happen when royalty rates account for about 80% of total marginal costs; while guidance simulation finds that opposite predictions happen at high values of royalty rates in the random sample of markets. We also detail the economic driving forces for opposite predictions in each type of mergers. In licensor-licensee mergers, we find that the ignored saving of royalty payment due to licensor-licensee merger leads to a decrease in prices and an increase of consumer welfare after merger, which causes opposite predictions if one assumes away the patent licensing relationships. In licensee-licensee mergers, we find that ignored decrease in alignment effect due to licensee-licensee merger leads to decrease of licensor's price, which causes opposite predictions if one assumes away the patent licensing relationships. We argue that these two ignored channels when assuming away existing patent licensing relationships are also driving forces for overall prediction biases that are unconditional on opposite predictions.

In both sets of simulations and both types of mergers, we find that market-level merger effects are over-predicted when one assumes away existing patent licensing relationships. We further find that, the prediction biases are smaller in licensee-licensee merger than those in licensor-licensee merger. In particular, the guidance simulation finds that, in licensor-licensee mergers, assuming away existing patent licensing relationships lead to over-predicted increase of share-weighted average price by 2.6% at median, and over-predicted decrease of consumer welfare by 9.3% at median; in licensee-licensee mergers, the two median predictions biases are 0.5% and 1.4% respectively.

We also examine the relationships between prediction biases, royalty rates and sum of diversion ratios between licensor and licensees. Theoretical simulation shows that prediction biases increase with royalty rates; while guidance simulation also shows positive correlations between the two. On the other hand, theoretical simulation shows non-monotonic relationship between prediction biases and the sum of diversion ratios; while guidance simulation finds that the potentials for prediction biases (i.e. maximum prediction biases) are positively correlated with the sum of diversion ratios.

This chapter is related to three strands of the literature. First, we build on the literature studying the role of intellectual property licensing on competition (Reisinger and Tarantino (2019), Layne-Farrar and Lerner (2010), and Lerner and Tirole (2005)). Second, we apply methods from the literature that uses theoretical and Monte Carlo simulations of structural models to evaluate the implications of modeling techniques, merger diagnostic tools, and different market features (Mazzeo, Seim, Varela (2018), Sheu and Taragin (2020), Balan and Brand (2018), Miller, Remer, Ryan, and Sheu (2012, 2013, 2016, and 2017), Miller and Podwol (2017), Dutra and Sabarwal (2020), Crooke (1999), Domnenko and Sibley (2020), Das Varma and De Stefano (2020)). Finally, we contributes to the long literature in industrial organization of estimating parameters in product competition models and predicting counterfactual market outcomes (Berry, Levinsohn, and Pakes (1995), Nevo (2001), and Fan and Yang (2020)). Our contribution lies in examining the consequences of not accounting for existing patent licensing relationship into the pricing competition models and even general supply-side models.

The remainder of this chapter proceeds as follows: section 2.2 illustrates the theoretical framework to analyze estimation bias and prediction bias due to unobserved existing patent licensing relationships; section 2.3 describes the simulation methods; section 2.4 shows simulation results; section 2.5 concludes.

2.2 Theoretical Framework

2.2.1 Alignment Effects on Pricing and Estimation

We begin with explaining the alignment effect on pricing of patent licensing relationships between manufacturers in a Bertrand-pricing competition framework. A set of multiple-product firms compete in one market. Some of the firms are licensors, who license patents to other firms and collect royalty revenues. Some of the firms are licensees, who purchase licenses to legally use the patented technology in production, and pay royalties. Assume that a licensor-manufacturer, f, produces and sells a set of products, $\mathcal{J}_f$, and owns a single patented technology[3]. This firm f chooses product prices to maximize the following profit function:

$$\pi_f = \sum_{j \in \mathcal{J}_f} (p_j - c_j) Q_j + r_f \sum_{l \in \mathcal{E}_f} Q_l \tag{2.1}$$

where c_j is product j's total marginal costs, Q_j is product j's unit sales, $\mathcal{E}_f$ is the set of other firms' products that use firm f's patented technology, and r_f is the exogenous per-unit royalty rate charged by firm f. The patent licensing relationship described by r_f and $\mathcal{E}_f$ partially aligns firms' pricing incentives: firm f's profit contains royalty payments that depend on other firms' sales, Q_l.

Now, we extend the profit function in Equation 2.1 to nest the profit functions for licensee-manufacturers. When $r_f = 0$ and $\mathcal{E}_f$ is empty, it is as if the firm f is a licensee. Moreover, we allow the total marginal cost, c_j, to contain royalty rates whenever necessary. Therefore, $c_j \geq r_f$ when j is a licensee-product. Under the assumption of Bertrand competition, profit-

[3]The model can be easily extended to allow multiple-patents licensor-manufacturers. We focus on single-patent firm for simple exposition.

maximizing prices are characterized by the following first-order conditions[4]:

$$p_j = c_j - \left(\frac{\partial Q_j}{\partial p_j}\right)^{-1}\left(Q_j + \sum_{i \in \mathcal{J}_f \setminus \{j\}} (p_i - c_i)\frac{\partial Q_i}{\partial p_j}\right) - r_f \sum_{l \in \mathcal{E}_f} \frac{\partial Q_l/\partial p_j}{\partial Q_j/\partial p_j}, \qquad \forall j \tag{2.2}$$

Denote $\mu_j := -r_f \sum_{l\in\mathcal{E}_f} \frac{\partial Q_l/\partial p_j}{\partial Q_j/\partial p_j}$. We call μ_j as the *alignment effect* of the existing patent licensing relationship on the price of product j. We note that μ_j is always non-negative, and in particular always positive for licensors whose $r_f > 0$.[5] Therefore, with patent licensing relationships, the licensor has incentive to charge higher prices compared to the case without patent licensing relationships. Because prices are strategic complements, the other firms are incentivized to charge higher prices in response. We also note that Equation 2.2 nests the case of no patent licensing relationships by setting $r_f = 0$ for all f.

Now we consider how the lack of patent licensing data leads to estimation bias. Without patent licensing data, the royalty rates, r_f, the licensed products, $\mathcal{E}_f$, and even the identity of the licensor, f, are unobserved. Therefore, μ_j is unobserved. In such cases, μ_j is typically assumed away, which results in the following commonly used back-out equation for marginal cost:

$$\tilde{c}_j = p_j + \left(\frac{\partial Q_j}{\partial p_j}\right)^{-1}\left(Q_j + \sum_{i \in \mathcal{J}_f \setminus \{j\}} (p_i - \tilde{c}_i)\frac{\partial Q_i}{\partial p_j}\right) \tag{2.3}$$

where the prices are from the data and derivatives are from a correctly specified and unbiased demand model estimation. Comparing Equation 2.3 to Equation 2.2, the resulting bias in marginal cost estimation is driven by the alignment effect μ_j:[6]

$$\Delta c_j := \tilde{c}_j - c_j \approx \mu_j.$$

[4]We think that the existence and uniqueness of the above game does not trivially follow the same argument for Nash-Bertrand pricing game without patent licensing relationships (Caplin and Nalebuff (1991), Vives (2001), and Aksoy-Pierson, Allon, and Federguen (2013)). This is mainly because the licensor-manufacturer does not directly choose the price of licensee-manufacturers' products.

[5]We assume that products are substitutes. Complement goods are beyond the scope of this work.

[6]We note that Equation 2.2 is one equation among a system of equations to back out the vector of marginal costs of firm f. Therefore, the estimation bias is not exactly μ_j.

Because μ_j is always positive, the unobserved alignment effects are likely to result in overestimated marginal costs. In fact, we note that, for a single-product licensor-manufacturer, $\Delta c_j = \mu_j > 0$. Positive estimation bias in marginal costs is intuitive: the licensing relationships create incentives for firms to charge higher prices, which will be explained by higher marginal costs without licensing data. We note that $\mu_j = 0$ for licensees' products. In fact, comparing Equation 2.3 to Equation 2.2 for licensee-manufacturers' products whose $r_f = 0$, the two equations are exactly the same. This implies that licensee's marginal costs are estimated unbiasedly, even when the model ignores the existing patent licensing relationships. Therefore, we focus on estimation bias in the marginal costs of the licensor.

2.2.2 Bias in Counterfactual Predictions

While the bias in estimated marginal costs of single-product firms has a clear sign, the bias in counterfactual predictions does not. We start with counterfactual prices of licensor-manufacturers. There two offsetting effects. First, overestimated marginal costs lead to higher predicted prices. Second, not accounting for the alignment effects lead to lower predicted prices. The net effect on price prediction is determined by the relative importance of the two channels. Since the bias in predicted counterfactual prices is ambiguous, the prediction errors of counterfactual quantities, consumer welfare and profits are also ambiguous.

To formulate the idea, we define the price of a product owned by a licensor-manufacturer that is predicted without licensing data, $\tilde{p}'_j$, as solution to a mis-specified model: Equation2.2 with $r_f = 0$ for all firms, status-quo biased marginal costs $\tilde{c}_j$ and counterfactual market structures. And we define the price of the same product that is predicted with licensing data, p'_j, as solution to a true model: Equation 2.2 with observed r_f and $\mathcal{E}_f$, unbiased marginal costs c_j and the same counterfactual market structures. Now, we consider the difference between $\tilde{p}'_j$ and p'_j, which is characterized by the following equation:

$$\Delta p'_j := \tilde{p}'_j - p'_j = \underbrace{\Delta c_j}_{\approx \mu_j} + (\tilde{\nu}'_j - \nu'_j) + \underbrace{r'_f \sum_{l \in \mathcal{E}'_f} \frac{\partial Q'_l / \partial p'_j}{\partial Q'_j / \partial p'_j}}_{-\mu'_j} \tag{2.4}$$

where $\tilde{\nu}'_j := -\left(\frac{\partial \tilde{Q}'_j}{\partial \tilde{p}'_j}\right)^{-1}\left(\tilde{Q}'_j + \sum_{i \in \mathcal{J}'_f \setminus \{j\}} (\tilde{p}'_i - \tilde{c}_i) \frac{\partial \tilde{Q}'_i}{\partial \tilde{p}'_j}\right)$, and $\nu'_j = -\left(\frac{\partial Q'_j}{\partial p'_j}\right)^{-1}\left(Q'_j + \sum_{i \in \mathcal{J}'_f \setminus \{j\}} (p'_i - c_i) \frac{\partial Q'_i}{\partial p'_j}\right)$. In this chapter, we do not consider counterfactual scenarios where the licensor-manufacturer's marginal costs are changed. We note that $\Delta p'_j$ is determined by three terms: i) estimation bias Δc_j that is likely to be positive; ii) difference between the predicted "mark-ups" if there were no patent licensing relationships $(\tilde{v}'_j - v'_j)$; iii) negative of counterfactual alignment effect, which is negative. While the second term does not have a clear sign, the first and the third terms are offsetting each other. Therefore, $\Delta p'_j$ has an unclear sign.

In general, the alignment effect μ_j itself will change in counterfactual scenarios. This is mainly because diversion ratios$(-\frac{\partial Q_l / \partial p_j}{\partial Q_j / \partial p_j})$ are sensitive to changes of prices and quantities, which are typically changed in the counterfactuals. We focus on mergers among various counterfactual scenarios, because merger is an important studied object in the literature of industrial organization. We consider two types of mergers: licensor-licensee merger and licensee-licensee merger.

The mis-specified model described in Equation 2.3 and the true model described in Equation 2.2 are different in three ways when predicting the effects of horizontal mergers. Firstly, they have different estimated marginal costs as illustrated in section 2.1. Secondly, the mis-specified model ignores a channel of cost saving after licensor-licensee mergers: the merged licensee can save the royalty payment to the licensor after a merger with the licensor. Thirdly, the mis-specified model ignores the royalty revenues as a part of the licensor's pricing incentive, both before and after the merger. This structural difference directly causes difference between the predicted prices from the mis-specified model and the true model. Moreover, the mis-specified model has different inferred pre-merger profits than the true model. In particular, the mis-specified model infer firms' profits as if $r_f = 0$ for all firms in Equation 2.1. We incorporate all these differences during simulations.

Given the same counterfactual, different market structures can lead to different changes of the alignment effect and thus different prediction biases. We note that the magnitude of the alignment effect, μ_j, is determined by two variables: i) the royalty rates (r_f), ii) the sum of diversion ratios between the licensor and licensees ($-\sum_{l \in \mathcal{E}_f} \frac{\partial Q_l/\partial p_j}{\partial Q_j/\partial p_j}$). We will examine how the biases in marginal cost estimations and merger effect predictions change with respect to the two market characteristics.

2.3 Simulations Methods

2.3.1 The Model and Specification

We simulate a model of three single-product firms, A, B and C. A is a licensor-manufacturer, B and C are licensee-manufacturers.

Demand. We follow the demand model in Mazzeo, Seim, and Varela (2018) to allow flexible substitution patterns between "horizontally" equidistant products. A consumer n receives the following utility from choosing product j

$$u_{jn} = \theta_{jn} - \alpha p_j + \epsilon_{jn} \tag{2.5}$$

and utility $u_{0n} = \epsilon_{0n}$ from not purchasing. Here, α is the consumer's price coefficient and $(\theta_{jn}, \epsilon_{jn})$ are two idiosyncratic taste shocks. We assume ϵ_{jn} follows i.i.d. Type 1 Extreme Value (T1EV) distribution with scale parameter σ_ϵ. θ_{jn} is draw from a multivariate normal distribution that allows for correlated shocks across products and non-zero means: $\theta_n \equiv (\theta_{1n}, \theta_{2n}, \cdots, \theta_{Jn})^T \sim \mathcal{N}(\theta|\delta, \Sigma)$[7]. The variance-covariance matrix Σ captures horizontally

[7] As noted by Mazzeo, Seim, and Varela (2018), θ_{jn} is comparable to a linear function of observed product characteristics and random taste shocks: $x_j\beta_n + \xi_{jn}$ where (β_n, ξ_{jn}) are random variables distributed according to some parameterized distribution. We also note that, when products are symmetric, such specification can nests consumer-invariant quality shocks, ξ_j, with $\rho = 1$. When we simulate asymmetric markets, we set $\sigma = 0$ and $\rho = 0$.

equidistant products with equal correlation (ρ) between θ_j and θ_l for each pair of products j and l:

$$\Sigma = \begin{bmatrix} \sigma^2 & \rho & \rho \\ & \sigma^2 & \rho \\ & & \sigma^2 \end{bmatrix}. \tag{2.6}$$

We follow Mazzeo, Seim, and Varela (2018) to call ρ as the *travel parameter*, which captures the distance between products in the preference space, with preferences for close products being highly correlated. We note that ρ is one of the model primitives that determine the diversion ratios.

Let M be the total number of consumers. The model-predicted unit sales of product j is given by

$$Q_j(p) = M \int \frac{e^{\frac{1}{\sigma_\epsilon}(\theta_{jn} - \alpha p_j)}}{1 + \sum_{l \in \mathcal{J}} e^{\frac{1}{\sigma_\epsilon}(\theta_{ln} - \alpha p_l)}} d\Phi(\theta_n | \delta, \Sigma) \tag{2.7}$$

where $\mathcal{J} = \{A, B, C\}$ is the set of products. We normalize market size to 1. This normalization is without loss of generality because market size is cancelled out everywhere in the first-order condition Equation2.2.

Pricing Game. We maintain the Nash-Bertrand Pricing competition assumption. Denote r as the royalty rate charged by firm A. We allow the licensees to have the same marginal costs as the licensor, as long as the licensees' marginal costs are no smaller than the royalty rate: $c_j \geq r, \forall j \in \{B, C\}$. Applying the above set-up of three single-product firms and patent licensing relationships to the first-order condition, Equation 2.2, the equilibrium prices are characterized by the following equations:

$$\begin{aligned} p_A &= c_A - \left(\frac{\partial Q_A}{\partial p_A}\right)^{-1} Q_A - r \sum_{l \in \{B,C\}} \frac{\partial Q_l / \partial p_A}{\partial Q_A / \partial p_A}, \\ p_j &= c_j - \left(\frac{\partial Q_j}{\partial p_j}\right)^{-1} Q_j, \qquad c_j \geq r, \forall j \in \{B, C\}. \end{aligned} \tag{2.8}$$

We apply the measurement of the equivalent variation in McFadden (1973) to quantify consumer surplus as blow

$$CS(p) = \int M\sigma_\epsilon \ln \left[1 + \sum_{j\in\mathcal{J}} e^{\frac{1}{\sigma}(\theta_{jn}-\alpha p_j)} \right] d\Phi(\theta_n|\delta,\Sigma). \tag{2.9}$$

Merger Effects. We consider two types of merger: i) merger between the licensor, A, and one of the licensees; ii) merger between the two licensees, B and C. We measure merger effects on a market outcome y in percentage change before and after the merger: $\%\Delta y := (y' - y)/y$. We consider market outcomes (y) including prices, quantities, consumer welfare and producer surplus. Lastly, we measure prediction bias of merger effects with difference between the merger effects ($\%\Delta\tilde{y}$) predicted from the mis-specified model that does not account for patent licensing relationships and biased marginal cost estimates and the merger effects ($\%\Delta y$) predicted from the true model that accounts for patent licensing relationships and true marginal costs: denoted as $\Delta\Delta y := \%\Delta\tilde{y} - \%\Delta y$.

2.3.2 The Simulation Algorithm

We have two goals of simulation: i) examine the theoretical ambiguity between model primitives and the biases in estimation and prediction; ii) provide guidance on correlations between typically observed market structures and the biases in estimation and prediction. For these two different goals, we design two different simulation algorithms. For the first goal of theoretical relationships, the designed simulation algorithm changes one model primitive over a continuum of values, holding all the other primitives fixed at a benchmark level. We call this set of simulations as theoretical simulation. For the second goal of guidance, the designed simulation algorithm randomly draws a rich set of various markets, backs out model primitives from the drawn market outcomes, then simulates the biases in estimation and predictions in each market. We call this set of simulations as guidance simulation. We explain the details of each simulation algorithm as below.

Theoretical Simulation. In the theoretical simulation, we specify a benchmark market, and then change one single model primitive in each simulation, holding all other model primitives fixed at the benchmark value. The benchmark market is summarized in Table 2.1. In the benchmark market, we shut down all random coefficients by setting σ and ρ to be zero. Then, the demand model in Equation 2.7 is reduced to a logit-demand model. Next, we set the scale parameter of the logit error, σ_ε, to be 0.1, so that changes in demand parameters have more bites on quantities. Also, we shut down patent licensing relationship in the benchmark market, by setting royalty rate, r, to be zero. We normalize market size, M, to be 1. Lastly, we set the benchmark market as a symmetric one: all products have the same marginal cost value, c, equal to 1; market share of each product is 25%; own price elasticity of demand of each product is -2.5. Based on such a set of model primitives and market outcomes, we back out the value of the common mean utilities (δ) of the three symmetric products; the value of the price coefficient (α); and calculate the price that are consistent with the market shares and price elasticities.

In theoretical simulations, we set $\sigma = 1, \sigma_\varepsilon = 0.1$[8]. We simulate biases in estimation and predictions with respect to different values of royalty rate (r) within $[0, 1]$ and the travel parameter (ρ) within $[0, 1]$. The simulation step is 0.01 for each parameter. To be aligned with the literature, we report the bias with respect to the resulting sum of diversion ratios between licensor and licensees from ρ,i.e. $-\left(\frac{\partial Q_B/\partial p_A}{\partial Q_A/\partial p_A} + \frac{\partial Q_C/\partial p_A}{\partial Q_A/\partial p_A}\right)$ [9]. Holding all the other model primitives fixed at the benchmark value in Table 2.1, for each pair of (r, ρ), we take the following steps to simulate the biases:

1. Simulate the pricing game described in Equation 2.8, which accounts for existing patent licensing relationships.

2. Back out the mis-specified marginal costs, $\tilde{c}_j$, as described in Equation 2.3. Calculate

[8] We follow Mazzeo, Seim, and Varela (2018) to set $\sigma_\varepsilon = 0.1$ so that variation in horizontal differentiation, i.e., in the travel parameter ρ, may have a large impact on market outcomes.

[9] In the Appendix Figure B.1 we show that the travel parameter ρ determines the sum of diversion ratios between licensor and licensees.

the estimation bias, $\Delta c_j := \tilde{c}_j - c_j$.

3. Simulate post-merger equilibrium with biased marginal costs, $\tilde{c}_j$, and the mis-specified model where $r = 0$ in Equation 2.8 with $r = 0$. Predict biased merger effects. In licensor-licensee mergers, we merge firm A and firm B[10]. We note that, in the mis-specified model, product B's total marginal cost is not changed after the merger, since r is unobserved and assumed away. In licensee-licensee mergers, we merger firm B and firm C.

4. Merger simulation with true marginal costs, c_j, and the true model as described in 2.8. Predict true merger effects. We note that, in the true model, after the licensor-licensee merger, product B's total marginal cost is decreased by r.

5. Compute the prediction bias in merger effects: $\Delta\Delta y = \%\Delta\tilde{y} - \%\Delta y$, i.e. the difference between predicted percentage changes in the market outcome y after the merger from the mis-specified model and the true model.

Guidance Simulation. In guidance simulation, we set $\sigma = 0, \sigma_\varepsilon = 1$. We generate and analyze a given theoretical market by the procedure outlined below. To construct our sample of markets, we repeat this process sufficiently as to generate 1,000 markets. For each market, we normalize initial prices, $p_j^0 = 1$ for $j = 1, ..., |J|$, where $|J|$ is the number of single-product firms in the market.

1. Obtain market shares by first drawing random variables $\check{s}_j \sim U[0, 1]$ for $j = A, B, C$. Calculate firm j's market share by, $s_j = (1 - s_0)\frac{\check{s}_j}{\sum_{j'=A,B,C}\check{s}_{j'}}$, where $s_0 = 0.25$ represents the initial share of the outside good. Back out mean utilities, $\delta_j = \log(s_j) - \log(s_0)$.

2. Draw the licensor's product margins as, $m_A \sim U[0.25, 0.75]$.

3. Draw the royalty rate uniformly as, $r \sim U[0, m_A]$.[11]

[10] Because the two licensees are symmetric during the simulation, mergers between firm A and firm B are the equivalent to mergers between the licensor, A, and licensee C.

[11] Setting the upper bound of the support on r to m_1 ensures the model is well-defined, specifically, that $\alpha < 0$.

4. From the licensor's first order condition taking into account the licensing relationship, back out the price sensitivity parameter as,

$$\alpha = \frac{-1}{(1 - s_A)m_A p_A - r(\sum_{j=B,C} s_j)} \tag{2.10}$$

5. From the licensees' first order conditions, back out remaining margins as,

$$m_j = \frac{1}{p_j}\left(\frac{-1}{\alpha(1 - s_j)} + r\right) \text{ for } j = B, C \tag{2.11}$$

6. Determine implied "true" costs for licensor and licensees as,[12]

$$c_j = p_j(1 - m_j) \text{ for } j = A, B, C \tag{2.12}$$

7. Given prices, p, using first order conditions that do not account for the licensing relationships back out the mis-specified marginal costs, $\tilde{c}_j$, as in Equation (2.3).

8. Conduct counterfactual simulation with the mis-specified model and $\tilde{c}_j$.

9. Conduct counterfactual simulation with the true model and true c_j.

10. Compute the prediction bias in merger effects, $\Delta\Delta y$.

11. Repeat Steps 1–10 until a total of 1,000 sample markets are obtained.

Guidance Simulation Sample. In order to understand the scope of markets comprising the sample generated from the above outlined procedure, we detail the key market characteristics in the sample. Table 2.2 documents the empirical distributions of these market characteristics for the generated sample. While royalty rates, r, and the licensor's margin, m_A, are drawn uniformly (see Guidance Simulation Section), we drop markets for which model primitives imply negative costs, inducing a non-uniform observed distribution in the

[12] Markets with any negative costs (i.e., $c_j < 0$ for any j) are discarded. This can occur when r is large.

sample over these primitives. Under Nash-Bertrand competition with a logit demand system, the licensees' costs are decreasing in both royalty rates and the licensor's margin. Thus, each respective empirical distribution is shifted left relative to parameter's drawn upon uniform distribution. For example, royalty rates are drawn uniformly from zero to one, however the empirical distribution of markets with non-negative implied costs is centered around 0.225 with a right tail. Similarly, the empirical distribution of the licensor margin is slightly left-shifted relative to the uniform distribution over [0.25, 0.75], from which this parameter is initially drawn.

The resulting sample exhibits variation in concentration, as quantified by market shares and HHI, margins, costs, and market elasticity. Note that prices deviate from the initially normalized prices due to resolving the pricing game upon recovering costs, and the price sensitivity parameter, α (see Step 7) and exhibit some heterogeneity, particularly for the licensor. However, margins, both the drawn licensor margin, m_A, and the implied licensee margins, m_B and m_C, exhibit larger variation than prices, mechanically giving rise to heterogeneity in costs. Additionally, looking ahead to the counterfactuals that we will evaluate under this framework – a merger between licensor and licensee (i.e., AB) and a merger between the two licensors (i.e., BC) – the resulting changes in market concentration vary substantially.

2.4 Simulation Results

2.4.1 The Simulated Bias in Marginal Cost Estimation

Theoretical Simulation Results. Figure 2.1 shows the theoretical simulation results on marginal cost estimation bias in a heat-map. The horizontal axis is the sum of diversion ratios between the licensor A and licensees B and C, i.e. $-\left(\frac{\partial Q_B/\partial p_A}{\partial Q_A/\partial p_A} + \frac{\partial Q_C/\partial p_A}{\partial Q_A/\partial p_A}\right)$. The vertical axis is the royalty rate r charged by the licensor, and also the fraction of r over product-level total marginal cost, since total marginal costs are normalized to 1. The color shows the exact value of the estimation bias in marginal costs of the licensor, $\Delta c_A := (\tilde{c}_A - c_A)/c_A$. It is not surprising that the features of Δc_A are consistent with the features of the alignment effect, because in the pre-merger market, the licensor is a single-product firm, the estimation bias is equal to the alignment effect. In particular, consistent with the fact that the alignment effect is always non-negative, the estimation bias in the licensor's marginal costs is always positive. This implies that ignoring existing patent licensing relationships will lead to over-estimated marginal costs of the licensor. Moreover, consistent with the increasing monotonicity of the alignment effect with respect to royalty rates and the sum of diversion ratios between the licensor and the licensees, the estimation bias is also increasing with respect to these two model features. The average bias in estimated marginal costs from the mis-specified model is 25.28% of the true marginal costs. [13]

Guidance Simulation Results. As described in Step 8 of the Guidance/Monte Carlo Simulation Procedure, costs are backed out for the mis-specified model for each firm. We compare these costs, $\tilde{c}$, to the true costs implied by a model correctly accounting for the licensing relationships between firms. Given licensee first order conditions are identical between true and mis-specified models (see Equation 2.8), the implied licensee costs are the same between the two models, i.e., $c_j = \tilde{c}_j$ for $j = B$, C. However, for the licensor, the alignment effect term, μ_A, is not accounted for in the mis-specified model, and thus implied licensor costs diverge between the two models. Table 2.3 and Figure 2.2 show the distribution of the bias in estimation of licensor marginal costs from the simulation sample. In all simulated markets, licensor marginal costs are overestimated in the mis-specified model relative to those consistent with the true model.

As detailed in Table 2.3, estimates of the licensor's costs under a mis-specified model can lead to significant differences from the true costs. At the median, estimates of the licensor's costs under the mis-specified model differ by 28% from those implied by the true model. Even at the 5th percentile, costs estimates differ by a non-negligible amount (i.e., 2.5%).

[13] We report estimation bias of licensee's marginal costs in Appendix Figures B.4 and B.5. Consistent with the theoretical framework, the bias are extremely small, and there is no pattern with respect to royalty rates and diversion ratios. Thus, we interpret these "estimation bias" as computation errors.

In order to understand when we would expect this bias in cost estimation to be larger in real world markets, we can assess how the bias varies by underlying model primitives – specifically, the licensing royalty rate, r, and the sum of the diversion ratios between the licensor and each respective licensee. Given the assumed exogeneity of the royalty rate, r, from Equation 2.8, we would expect size of the cost estimation bias to scale with the royalty rate. Indeed, in Figure 2.4, we see that, for a given sum of diversion ratios between the licensor and each respective licensee, the greater the royalty rate, the larger the bias in estimation of licensor costs.[14]

Given this set of guidance simulations is aimed at understanding in what real world markets would we expect larger differences between the estimates of a mis-specified model and the true model, we analyze these cost estimates relative to a common measure of market concentration – HHI. In Figure 2.3, we see that in less concentrated markets (i.e., smaller HHI), there is larger variance in the percentage difference in estimated licensor marginal costs between mis-specified and true models. While in more concentrated markets, there is smaller variance in cost estimate differences, the magnitude is still relatively large. For example, on the right-hand side of Figure 2.3, one relatively concentrated simulated market is observed with a pre-merger HHI of 4523 while exhibiting a relatively large estimated licensor cost difference of 50.62%.

Understanding that, in these simulated markets, implied licensor costs can be significantly overestimated in a model that does not account for licensing, in Section 2.4.2 below, we look to understand the implications of this overestimation on counterfactual equilibrium objects of interest to researchers, policy makers, and antitrust authorities.

2.4.2 Prediction Bias in Licensor-Licensee Mergers

Theoretical Simulation Results. Figures 2.5 to 2.8 shows the results on the prediction biases in the effects of mergers between the licensor, A, and licensee B. The horizontal axis

[14] This correlation between royalty rate levels and the bias in licensor cost estimation is additionally demonstrated in Appendix Figure B.6 where cost biases are plotted against royalty rates.

is the sum of the diversion ratios between the licensor A and the two licensees. The vertical axis is the royalty rate r. The colors show the prediction biases. Figure 2.5 shows the prediction biases in consumer welfare effects of the mergers. Figure 2.6 shows the prediction biases in the merger effects on sales-weighted average price. Figure 2.7 shows the prediction biases in the merger effects on total quantities. Figure 2.8 shows the prediction biases in the merger effects on producer surplus.

Overall, the results in Figures 2.5 to 2.8 show that i) the higher the royalty rates, the larger the magnitude of the prediction bias, and ii) the relationship between diversion ratios and the prediction bias is non-monotonic. Given a level of the sum of diversion ratios, figure 2.5 shows that the magnitude of the prediction bias in the consumer welfare effects of the merger increases with royalty rate. Similar patterns exist for Figures 2.6, 2.7, and 2.8, including merger effects on sales-weighted average price, total quantities and producer surplus. Given a royalty rate, Figure 2.8 shows that the prediction bias on the merger effects on producer surplus increases with the sum of diversion ratios. However, Figure 2.5 shows that, given a low level of royalty rate, the magnitude of the prediction bias in the consumer welfare effects typically decrease as the sum of diversion ratios increases. But at high levels of royalty rates, such as 80% of the licensees' marginal costs, the prediction bias might reach the largest magnitude when the sum of diversion ratios is at middle range such as 0.5. Similar patterns show up in the Figure 2.7. Figure 2.6 shows an even more obvious non-monotonic relationship between the prediction bias and the sum of diversion ratios, when it comes to the prediction of merger effects on sales-weighted average price. In particular, given a level of royalty rate, the prediction bias in Figure 2.6 firstly increases then decreases as the sum of diversion ratios increases.

The above results have two implications. First, when predicting merger effects on consumer surplus, sales-weighted average price and total quantities, it's better to be careful if the royalty rates are high and diversion ratios are in the middle range. Second, estimation bias in marginal costs is not the deterministic factor for prediction biases. As estimation bias in marginal costs always increase with the sum of diversion ratios as shown in Figure 2.1, we see non-monotonic relationship between the prediction biases and the sum of diversion ratios. This is consistent with the theoretical analysis of $\Delta p_{j'}$ in Equation 2.4.

Table 2.4 provides summary statistics of the theoretical simulation results on the prediction bias, conditional on positive royalty rates. Each row is a market outcome of general interests. The first column reports the average prediction bias. The second column reports the average true merger effects, measured in the percentage change of a market outcome relative to its pre-merger levels. The third column reports the percentage of cases that generate over-predicted merger effects in magnitudes during the simulation. Recall that each case is a pair of royalty rates and the travel parameter, i.e. (r, ρ). The last column reports the percentage of cases where the mis-specified model predicts opposite direction of merger effects during the simulation. The results are conditional on positive royalty rates, so that ignoring patent licensing relationship or not is non-trivial.

Column (2) of Table 2.4 shows that the simulated merger effects are economically reasonable. On average, the true model with patent licensing relationships predicts decrease in consumer surplus, increase in all prices, decrease in the quantities of the merged products and increase in the quantity of the non-merging product, and increase in all profits after the licensor-licensee merger.

At market level, Column (1) of Table 2.4 shows that, on average, the mis-specified model predicts more loss in consumer welfare by 2.4%, more increase in sales-weighted average price by 2.7%, more decrease in total quantities by 1.8% and more increase in producer surplus by 2.8% after the licensor-licensee merger. At product level, Column(1) of Table 2.4 shows that, the market-level overly predicted post-merger price increase is driven by the overly predicted post-merger price increase of the merged licensee's product, B. As prices are strategic complements, it's not surprising to see that post-merger prices of the licensor's good, A, and the non-merging licensee's good, C, are also overly predicted by the mis-specified model. The largely over-predicted post-merger price B also leads to largely over-predicted post-merger decrease in quantity B, which leads to the overall over-predicted post-merger decrease in quantities at the market level[15]. At product-level, we see over-predicted post-merger profit increase as what we see at market-level. These results are consistent with column (3) of Table 2.4, which shows that in almost all of the cases, merger effects are over-predicted in the mis-specified model[16].

The last column of Table 2.4 shows the existence of opposite prediction when the prediction is based on mis-specified model ignoring existing patent licensing relationship. This is alarming to researchers. It shows 3% of simulation cases see opposite predictions on merger effect on sales-weighted average price. It further shows that the opposite prediction in market-level average price roots in opposite prediction in price B; and the opposite prediction in price B also leads to opposite prediction in Quantity B.

We provide more details on the opposition predictions in Table 2.5. Column (1) shows the average royalty rates where the opposition prediction on each market outcome happens. Column (2) shows the corresponding average sum of diversion ratios. Column (3) shows the corresponding average travel parameter. Column (4) shows the corresponding average true merger effects. Column (5) shows the corresponding average mis-specified merger effects. Column (6) shows the corresponding average prediction bias.

To answer where the opposite predictions happen, Column (1) of Table 2.5 shows that, the opposite prediction cases mainly happen in relatively extreme cases: royalty rates accounts for more than 85% of total marginal costs. This is good news. Column(2) and (3) show

[15] As for the other merged product, A, while its post-merger price are over predicted (i.e. higher post-merger price A), it's quantity decrease is not over predicted and practically always under-predicted. This might be due to the relative over predictions in prices of A, B and C. Note that price of B is the most over-predicted price. Therefore, good A will be predicted more attractive to consumers than good B in the mis-specified model, which leads to the less decrease in quantity A. Similar reasons explain why the non-merging product, C, also see over-predicted post-merger price increase, and higher post-merger downloads predicted from the mis-specified model.

[16] We additionally note that, given no opposite prediction on the merger effects on price of A, p_A, over-prediction is the same as $\Delta p'_A > 0$. Therefore, the 98.8% of over-prediction cases imply that in most of the cases, the ignored alignment effect in Equation 2.4 is not driving the prediction biases in post-merger price of the licensor-manufacturer. However, given the averagely negative true merger effect and positive prediction bias on quantity of A, $\tilde{Q}'_A > Q'_A$, and thus $\tilde{v}'_A > v'_A$. Therefore, both the first and the second terms in Equation 2.4 are positive, which disenables us to conclude that marginal cost estimation bias is driving the prediction bias.

that, to have opposite predictions on market-level sales-weighted average price, we also need extremely high substitutability between products. However, only middle substitutability between products are needed to generate opposite predictions on product-level price and quantity, conditional on extremely high royalty rates.

To answer how the opposite prediction happens, column(4) and (5) of Table 2.5 shows that the reason is that the mis-specified model ignores cost-saving of product B after its merger with the licensor. In particular, column (4) shows that the opposite prediction of the merger effect on sales-weighted average price happens because the true merger effects actually lead to decrease in the sales-weighted average price by 0.16% on average. Such lower price after merger happens because the merged licensee, B, does not need to pay royalty rates to the licensor, A, after the merger. With the cost saving, product B can enjoy a 0.78% lower price on average. However, column (5) shows that, the mis-specified model cannot account for such cost reduction channels, and thus, predicts that the price of product B will increase by 10.8% after the merger, on average. Such discrepancy in predicted post-merger price of B also leads to opposite predictions on post-merger quantities of B between the true model and the mis-specified model: on average, the true model predicts 2.7% increase of the quantity of product B, while the mis-specified model predicts 10.3% decrease of the quantity of product B. Now, the close relationship between the opposite prediction and high royalty rates are intuitive: a higher royalty rate implies larger cost reduction un-captured by the mis-specified model.

To answer how important are those opposite-prediction cases, we compare column (4) to (6) in Table 2.5 to column (1) and (2) in Table 2.4. Comparing column (1) of Table 2.4 to column (6) of Table 2.5, it shows that in the cases of opposite predictions, the prediction bias are larger in magnitudes. Comparing column (2) of Table 2.4 to column (4) of Table 2.5, it shows that in the cases of opposite predictions, the true merger effects are smaller in magnitudes. More importantly, we note that the average prediction bias on price of B is 11.61% conditional on opposite predictions, while the unconditional average prediction bias is 6.23% as reported in Table 2.4. Since the average prediction bias of merger effects on price B is driving the market-level prediction biases on prices, we conclude that the ignored royalty payment saving due to licensor-licensee merger is driving the prediction bias in licensor-licensee merger.

Guidance Simulation Results. We can analyze the bias in counterfactual predictions between the true and mis-specified models under a horizontal merger between the licensor ($j = A$) and one of the licensees (without loss, $j = B$). From these simulations, we find that, in general, due to the overestimation of licensor costs from not accounting for the licensing relationship, as detailed above in Section 2.4.1, the mis-specified model predicts inflated counterfactual welfare measures. Particularly, relative to the true model, the mis-specified model predicts larger welfare losses for consumers and greater welfare gains for producers following a merger between the licensor and a licensee.

To assess for what characteristics of real-world markets we would expect to find larger biases in counterfactual prediction, we compare the biases in counterfactual measures across a rich set of simulated markets that span the space of reasonable underlying model primitives. This allows examine where the bias in counterfactual predictions is largest, and thus cause for concern for market analysts, across generated markets with different distributions of market shares, margins, and royalty rates.

To explore when prediction biases are largest and provide guidance to analysts evaluating markets with potentially unobserved licensing behavior, we plot the bias in a given counterfactual measure derived for each market in our sample against the distribution of an underlying model parameter.

First, we show how the bias in key counterfactual measures vary with the licensing royalty rate, r. Figure 2.9 plots the relationship between a market's royalty rate, r, and the bias in reported counterfactual-factual change in share-weighted average prices between the mis-specified and true model. Each circle represents a generated market with its associated values plotted in r-$\bar{p}$ space. It can be seen from the strictly positive biases (i.e., only positive values on the y-axis) that the mis-specified model exhaustively overestimates the percent change in share-weighted average prices, relative to the true model (see also the top panel of Table 2.6).[17] Further, moving from left to right within Figure 2.9, we see that markets that exhibit a larger royalty rate are correlated with a larger bias in predicted change in share-weighted average price, with heterogeneity in the prediction bias at larger values of r. Therefore, in real world markets, we would expect markets with larger royalty rates to have the potential to exhibit larger errors in reported prices from a licensor-licensee merger under a mis-specified model and thus more reason for concern for market analysts.

Given that mis-specified model exhaustively predicts a larger change in share-weighted prices than the true model and that larger magnitudes of this overprediction (particularly the maximum bias) are associated with royalty rates, we would expected the mis-specified model to predict larger changes in the shares of the inside goods (i.e., $\sum_{j=A,B,C} s_j$) and that the magnitude of this overprediction increases in r. Indeed, this relationship is exhibited in Figure 2.10 – when royalty rates are larger, we see that the mis-specified model generally predicts a larger decrease in total inside shares relative to the true model.[18] Given the overprediction of price increases and the associated overprediction of inside share decreases, we would generally expect an overprediction of consumer welfare losses as the royalty rate increases. In our sample, this relationship holds, as shown in Figure 2.11, which plots the bias in predicted changes in consumer surplus between the mis-specified and true models against a market's royalty rate, r. Finally, predictions for change in producer profits due to the licensor-licensee merger seem to follow a less clear trend, in terms of the correlation with royalty rates, as exhibited in Figure 2.12, however are largely overpredicted. Overall, these results with respect to the market's royalty rate suggest that for markets with a larger royalty rate, the analysis of a merger between the licensor and a licensee under a model that

[17]In all simulated markets, the mis-specified model predicts an increase in share-weighted average prices due to the licensor-licensee merger, while the true model does so for only 63.7% of markets.

[18]The mis-specified model underpredicts post-merger shares for Firm B (i.e., the merging licensee), while it overpredicts shares for Firm A (i.e., the merging licensor) and Firm C (i.e., the non-merging licensee). See Table 2.6

does not account for licensing behavior can critically overpredict pre- to post-merger changes in key welfare measures.

We can additionally assess the mis-specified model's counterfactual predictions in the simulated markets based on the sum of the diversion ratios between the licensor and each licensee. Figures 2.13 to 2.16 shows the relationship between the same biases in predicted changes in welfare measures as Figures 2.9 to 2.12 – share-weighted average prices, total inside market shares, consumer surplus, and producer surplus – now against the model's implied sum of diversion ratios between the licensor and the two respective licensees. The trend in prediction bias is less clear, however we can observe that maximum (i.e., frontier) over- or underprediction does exhibit an association with the sum of diversion ratios measure.[19] For example, Figure 2.13 shows that, for larger sum of diversion ratios, the maximum overprediction of share-weighted average prices is generally increasing. Similarly, the maximum overprediction in pre- to post-merger decreases in inside shares and the maximum overprediction in pre- to post-merger decreases in consumer surplus are generally increasing in the sum of the royalty rate. The relationship between the maximum prediction error in changes in producer surplus is non-monotonic in the sum of diversion ratios. Specifically, at smaller values of the sum of diversion ratios, the maximum prediction bias in producer surplus is increasing; however, at larger values, this frontier is decreasing. For market analysts, this non-monotonicity suggests concern for larger overpredictions in producer surplus lie in the interior of feasible sum of diversion ratios, while larger prediction biases for share-weighted average prices, total inside shares, and consumer surplus occur at high values of the sum of diversion ratio term.

Not only are we interested where in the model primitive space do the magnitude of predictions diverge between the mis-specified and true models, but also when they predict changes in welfare measures of the opposite sign. If, for example, a mis-specified model is used to evaluate a merger that (incorrectly) predicts losses in consumer surplus, whereas a

[19]Given the simulated outside market shares are calibrated at $s_0 = 0.25$, the maximum possible sum of diversion ratios is $1 - s_0 = 0.75$.

correctly-specified model would predict consumer surplus gains, an analyst may critically arrive at an opposite policy recommendation against allowing the merger occur. In Table 2.7, we summarize the predictions in welfare measures between the true and mis-specified models across the generated sample of markets. In the top panel, first column, we see that, on average, share-weighted prices increase by 2% following the merger between licensor (Firm A) and licensee (Firm B). This increase in share-weighted average prices is larger driven by increases in Firm A's product, which is 9.9% higher than pre-merger levels, on average. The mis-specified model overpredicts share-weighted average prices by 3.5% on average, largely due the overprediction of post-merger price of Firm B (i.e., the merging licensee). This follows from the intuition that Firm B, by merging with the licensor, no longer needs to pay the royalty rate, which reflects a cost savings that induces downward pressure on Firm B's post-merger price. Given the mis-specified model does not account for the pre-merger licensing relationship, no cost savings are exhibited and, thus, the mis-specified model overpredicts Firm B's price increase, on average by 8.9 percentage points.

In the generated sample, the true model predicts increase in share-weighted prices due to the licensor-licensee merger 63.7% of the time, while the mis-specified model always predicts higher post-merger prices. In the remaining 36.3% of sample markets, the true model predicts a decrease in share-weighted average prices, with the mis-specified model predicting the opposite effect in these markets, as shown in the third column of Table 2.7.

We see corresponding trends in the average effects, biases, and opposite predictions in terms of market shares. The true model predicts, on average, a decrease in inside shares, driven by large decreases Firm A's market share. On average, the mis-specified model overpredicts this decrease in total inside market shares by 4.7% and predicts an opposite change in the inside total share in 32.7% of markets.

The true model that accounts for licensing predicts, on average, a 5.0% decrease in consumer surplus due to the licensor-licensee merger. Largely driven by the inability to account for the cost savings from the merging licensee, and the associated prediction of a higher post-merger price for Firm B, the mis-specified model overpredicts the loss in consumer surplus by 15 percentage points. Critically, the true model predicts increases in consumer surplus in 32.7% of markets, while the mis-specified model exclusively predicts consumer surplus losses. In other words, the mis-specified model predicts an opposite effect on consumer for approximately one third of markets in our simulated sample. Given modern merger enforcement's focus on consumer welfare to drive policy recommendations, in these markets with opposite consumer surplus predictions, a mis-specified model that does not take into account licensing behavior would lead to an errant recommendation.

What are the characteristics of the markets in which the true and mis-specified model predict opposite effects? In Table 2.8, we detail the conditional distribution of market characteristics for markets that exhibit opposite predictions between the true and mis-specified model for at least one welfare measure (i.e., prices, shares, profits, or consumer surplus). Relative to the full distribution of the sample given in Table 2.2, markets that exhibit an opposite prediction in at least one welfare measures by the mis-specified model generally exhibit: larger royalty rates, larger licensee shares and margins, higher concentration both pre- and post-merger, larger costs, particularly for licensees, and more elastic demand.

2.4.3 Prediction Bias in Licensee-Licensee Mergers

Theoretical Simulation Results. Figures 2.17 to 2.20 shows the theoretical simulation results on the prediction biases in the effect of merger between the two licensees, B and C. Firstly, Figure 2.17, Figure 2.18 and Figure 2.19 show that i) the prediction biases in the merger's effect on consumer welfare, sales-weighted average prices and total quantities increase with royalty rate; and ii) the prediction biases firstly increase then decrease with the sum of diversion ratios in two rounds. Secondly, Figure 2.20 shows that the prediction biases in the merger effects producer surplus increase with both royalty rate and the sum of diversion ratios.

We interpret the above results with three implications similar to those from Figures 2.5 to 2.8. First, when a researcher predicts post-merger quantities and consumer surplus without data on existing patent licensing relationships, the researcher is better to be careful if royalty rates are large and the diversion ratios are at middle range. Second, when a researcher predicts post-merger prices and profits without data on existing patent licensing relationships, the research is better to be careful if royalty rates and the sum of diversion ratios are large [20]. Third, estimation bias in marginal costs is not the deterministic factor for prediction biases: As alignment effect always increase with the sum of diversion ratios, we see up and downs in the prediction bias with respect to the sum of diversion ratios.

We summarize the prediction bias in Table 2.9, in the same format as Table 2.4. In particular, each row is a market outcome of interests. Column (1) reports the average prediction bias. Column (2) reports the average true merger effects. Column (3) reports the percentage of cases that generate over-predicted merger effects in magnitudes during the simulation. Column (4) reports the percentage of cases where the mis-specified model predicts opposite direction of merger effects during the simulation. The results are conditional on positive royalty rates, so that ignoring patent licensing relationship or not is non-trivial. Compared to the results in Table 2.4 for licensor-licensee merger, Table 2.9 shows that, for licensee-licensee merger, prediction bias on market-level outcomes have the same signs and smaller magnitudes; while prediction bias on product-level outcomes are quite different even in signs. We firstly discuss each group of outcomes, then focus on the opposite prediction cases.

With respect to licensee-licensee merger's effect on consumer surplus, the first row of Table 2.9 shows that, the average prediction bias is -0.2% with an average true merger effects of -10.9%. This implies that the mis-specified model typically over-predict the decrease in consumer surplus due to the licensee-licensee merger. This is also seen in the first row of Table 2.4. The last two columns in the first-row shows that, in 99.9% cases during simulation,

[20] We note that non-monotonicity with respect to the sum of diversion ratios is more obvious in prediction bias on post-merger quantities and consumer welfare than in prediction bias on post-merger prices. We think the reason is that, apart from indirectly affecting the demand non-linearly through affecting prices with alignment effects, the travel parameter (determinant of diversion ratios) directly affect demand through consumer preferences.

the mis-specified model predicts larger consumer welfare effect than the true model, and in no case does the mis-specified model predict opposite consumer welfare effect compared to the true model.

With respect to licensee-licensee merger's effect on prices, the rows two to five of Table 2.9 show that, the average prediction biases and the average true merger effect on sales-weighted average price and product-level prices are all positive. This implies that the mis-specified model typically over-predicts the increase in prices due to the licensee-licensee merger, which also happens in licensor-licensee merger as shown in Table 2.4. The high ratios of over-prediction cases are consistent with this implication. Lastly, there are 4.4% cases where the mis-specified model predicts opposite licensee-licensee merger effect on price of the licensor's product, A. We discuss the underlying reasoning later.

With respect to licensee-licensee merger's effect on quantities, the rows six to nine of Table 2.9 show that, the average prediction bias and the average true merger effect on total quantities are negative. This implies that the mis-specified model typically over-predicts the decrease in quantities due to the licensee-licensee merger, which also happens in licensor-licensee merger as shown in Table 2.4. This implication is consistent with the 99.9% cases that see over-predictions. At product-level, we see different merger effects and prediction biases compared to Table 2.4. This is reasonable since the merged firms and different. In particular, with the merged firms being the licensees, the licensor's product, A, sees increase in quantities due to the merger. Then, given the average negative prediction bias on merger effect on quantity A (-0.69%), the result implies that the mis-specified model typically under-predicts the increase in the non-merging firm's quantity due to the licensee-licensee's merger. Notice that this is qualitatively different from the result in Table 2.4, where the averagely positive effect on the non-merging firm's quantity, C, is typically over-predicted. We will explain this result together with the opposite predictions on merger effect on price A later. As for merged firms' quantities, Table 2.9 shows that, the licensee-licensee merger averagely leads to decrease in merged firms' quantities (by 15.0% to 15.6%), and the mis-specified model averagely weakly under-predicts the decrease their quantities. This result is consistent with the close-to-zero ratio of over-prediction[21]. There is no opposite-prediction case with respect to quantity effects of the licensee-licensee merger.

With respect to licensee-licensee merger's effect on profits, the last three rows of Table 2.9 shows that, the mis-specified model averagely over-predicts the increase in firms' profits due to the licensee-licensee merger. This is consistent with the high ratio of over-prediction cases. Lastly, given the opposite prediction on price effects on product A, there also exists opposite prediction on profit effects on product A.

We summarize the opposite-prediction cases in Table 2.10, in the similar format as Table 2.5. In particular, Column (1) shows the average royalty rates where the opposite prediction on each market outcome happens. Column (2) shows the corresponding average sum of diversion ratios. Column (3) shows the corresponding average travel parameter. Column(4) shows the corresponding average true merger effects. Column (5) shows the corresponding average mis-specified merger effects. Column (6) shows the corresponding average prediction bias.

Column(4) of Table 2.10 show that the opposite predictions of merger effects on price A happen with negative true merger effects on price A. Why would the true model predict lower price of A, the non-merging licensor, after licensee-licensee merger? This is due to decrease in alignment effect after the merger. In particular, after the licensee-licensee merger, the prices of the licensees' products increase, and their quantities decrease, which decreases how much the licensor cares about the licensees' sales. In fact, the alignment effects, in the case of opposite prediction of price of the licensor, are always smaller after the merger, and on average, decrease by 8.73%. While prices are strategic complements, the higher prices of B and C lead to smaller positive effect on the price A than the negative effect due to lower alignment effect. Such decrease in price of A after the licensee-licensee merger, also leads to

[21]For clarification, in Table 2.9, we note that the product-level over-prediction ratios for quantity effects are small while the market-level over-prediction ratio is large. This is because the over-prediction at market-level quantity effect is due to the under-prediction of quantity effect on product A: the increase in quantity of A is under-predicted, thus the decrease in market-level quantities is over-predicted.

decrease of profit of A after the merger, by about 0.23%.

To answer when the opposite predictions happen. Column (1) of Table 2.10 shows that such opposite predictions typically happen at relatively extreme cases: royalty rates account for 80% to 95% of total marginal costs on average. This is intuitive, because the larger the royalty rates, the more important is the alignment effect. Column (2) and (3) of Table 2.10 shows that such opposite predictions happen with relatively low values of diversion ratios and small travel parameter.

To answer how important are these opposite-prediction cases, we compare column (4) to (6) in Table 2.10 to column (1) and (2) in Table 2.9. Comparing column (1) of Table 2.9 to column (6) of Table 2.10, it shows that in the cases of opposite predictions, the prediction biases are not quite different in magnitudes, if not smaller. Comparing column (2) of Table 2.9 to column (4) of Table 2.10, it shows that in the cases of opposite predictions, the true merger effects are smaller in magnitudes. More importantly, we note that the average prediction bias in merger effects on price A conditional on opposite prediction is 0.39%, while the unconditional average prediction bias is 0.44% as reported in column (1) of Table 2.9. The smaller conditional average prediction bias implies that while the channel of ignored decrease in alignment effect after merger is driving the opposite predictions, there are other larger prediction biases happening when the true merger effects on price A are positive[22]. However, the conditional average prediction bias in merger effects on price A is larger than unconditional average prediction biases in merger effects on price B and C as reported in column (1) of Table 2.9. Therefore, we conclude that, ignored decrease in alignment effect due to licensee-licensee merger is a driving force of prediction biases.

[22]Based on Equation 2.4 and the results in Table 2.9 and Table 2.10, we argue that estimation bias in marginal cost of the licensor-manufacturer, Δc_A, is also a driving force for overall prediction biases. In particular, given no opposite predictions on Q_A, the over-prediction ratio for merger effects on Q_A implies that, in almost all cases, $\tilde{s}'_A < s'_A$. Since firm A is a single-product firm, $v_A = -1/(\alpha(1 - s_A))$. Therefore, $\tilde{v}'_A < v'_A$ in almost all cases. Then, both the second and the third terms in Equation 2.4 are negative, the first term – marginal cost estimation bias – is the only positive term. Moreover, Table 2.9 shows that on average, the prediction bias on merger effects on price A is 0.446%, which is positive. Therefore, on average, the estimation bias on marginal costs is driving the prediction bias on post-merger prices of the licensor-manufacturer, p'_A.

As a summary of main findings from theoretical simulations on licensor-licensee merger and licensee-licensee merger, we show a result that should be alarming to researchers: unobserved patent licensing relationships can lead to opposite prediction of merger effects. And the opposite predictions are fundamentally due to ignored mechanisms: i) royalty payment saving due to licensor-licensee merger; ii) smaller alignment effect due to licensee-licensee merger. The good news is that, in the theoretical simulations, which is based on a symmetric benchmark market, opposite predictions happen in relatively extreme cases. In particular, opposite predictions happen when royalty rates account for about 80% of total marginal costs.

Theoretical simulations also illustrate on the patterns between prediction bias and two model primitives: i) royalty rates; ii) the sum of diversion ratios between the licensor and licensees (controlled by the *travel parameter*). We find that larger royalty rates, larger bias in marginal cost estimation and merger evaluations. We also find that with larger sum of diversion ratios (i.e. substitutability), the bias in marginal cost estimation increases, but the prediction biases in merger effects typically firstly increase then decrease.

Lastly, theoretical simulations find that, on average, merger effects are over-predicted. In particular, for both licensor-licensee and licensee-licensee mergers, by ignoring patent-licensing relationships, consumer surplus decrease is over-predicted(2.4% for licensor-licensee merger, 0.2% for licensee-licensee merger), increase of sales-weighted average price is over-predicted(2.65% for licensor-licensor merger, 0.25% for licensee-licensee merger), decrease in total quantities is over-predicted (1.76% for licensor-licensee merger, 0.15% for licensee-licensee merger), increase in producer-surplus is over-predicted (2.7% for licensor-licensee merger, 2.95% for licensee-licensee merger). Moreover, on average, the prediction bias in licensor-licensee merger is larger than that in licensee-licensee merger.

Guidance Simulation Results. Across the generated sample of simulated markets, we analyze a horizontal merger between the two licensees, i.e., Firm B and Firm C, and assess the predictions of counterfactual measures from the mis-specified model relative to those from the model accounting for the licensing in the market. We similarly evaluate how these predictions change with respect to underlying model parameters which comprise the alignment effect term, namely the royalty rate and the sum of diversion ratios.

In Figures 2.21 to 2.24, which plot simulated markets' biases in predicted welfare measures form a licensee-licensee merger against its royalty rate, we see similar trends to predictions in the guidance simulations of the licensor-licensee merger case. The mis-specified overpredicts increases in share-weighted prices (Figure 2.21), overpredicts decreases in inside shares (Figure 2.22), and overpredicts decreases in consumer surplus (Figure 2.23), relative to the true model, and the magnitudes of each of these overpredictions generally increases in the market's royalty rate, r. Additionally, for licensee-licensee mergers, a trend is present for the prediction bias is producer surplus with respect to the royalty rate – the mis-specified model overpredicts changes in producer surplus and the magnitude of this overprediction is correlated with r (see Figure 2.24).

We also analyze how predictions differ based on the other component of the alignment effect – the sum of the diversion ratios. In Figure 2.13, we see that the maximum overprediction in share-weighted average prices is first increasing in the sum of diversion ratios and then sharply drops at extreme values of the sum of diversion ratios. A similar pattern in overprediction of decreases in total inside shares is shown in Figure 2.14. Correspondingly, the overprediction in decrease in consumer surplus is largest at interior, but relatively large values of the diversion ratios (Figure 2.15). Finally, the maximum bias in producer surplus due to a licensee-licensee merger in increasing in the sum of the diversion ratios, as is shown in Figure 2.16.

For licensee-licensee mergers, we want to understand when a mis-specified model predicts the opposite sign of a key welfare measure, relatively to the effect under a true model. In Table 2.12, we document, across the generated sample of markets, the average true effect, average prediction bias, and the frequency in which an opposite prediction occurs for key counterfactual objects – prices, shares, profits, and consumer surplus. We see, on average, due to the licensee-licensee merger, the true model predicts an increase in share-weighted average prices (by 4.8%) and a decrease in inside shares (by 5.4%), which is associated with an average decrease in consumer surplus (by 17.2%). In these three measures, prices, shares, and consumer surplus, the mis-specified model overpredicts the respective effects, however to a relatively small magnitude. For example, the mis-specified model overpredicts the percentage decrease in consumer surplus by 2.2 percentage points. Further, we see that in these three measures, the mis-specified model predicts the correct sign of the change in welfare measure in all markets (column (3)). However, in terms of producer surplus, the mis-specified model predicts an opposite effect in an overwhelming majority of markets in the sample (87.6%).

What are the characteristics of the markets in which the true and mis-specified model predict opposite effects? In Table 2.8, we detail the conditional distribution of market characteristics for markets that exhibit opposite predictions between the true and mis-specified model for at least one welfare measure. We see that relative to the distributions in the full sample detailed in Table 2.2, the markets with opposite predictions have larger royalty rates. However, in many other key measures, the conditional distributions of these market characteristics are relatively close to those of the full sample.

Guidance Summary. In the guidance simulations, we assess differences in cost estimates and counterfactual merger predictions between a mis-specified model that does not take into account licensing behavior and a true model that does. In terms of costs, the bias in estimation of a licensors cost from mis-specified model can be significant. Further, we show that this estimation bias increases in the licensors royalty rate and exhibited larger potential magnitudes in less concentrated markets.

This difference in estimated costs plays through to model predictions of counterfactual scenarios. Using this sample of generated markets, we evaluate two mergers one between the licensor and a licensee, and one between the two licensees. Under the licensor-licensee merger, we show the mis-specified model overpredicts decreases in consumer surplus (from overpredicting increases share-weighted average prices and thus overpredicting decreases inside shares) and, on average, overpredicts increases in producer surplus. The magnitude of the overpredictions is shown to positive correlated in the licensors royalty rate. Further, we show that the mis-specified model, under a licensor-licensee merger, can lead to policy recommendation opposite of that suggested by the true model. For example, the mis-specified model always predicts a loss of consumer surplus due to the merger, while in approximately one third of the simulated markets the true model predicts increases in consumer surplus.

Additionally, we use the generated sample to evaluate a licensee-licensee merger. Similarly, under this type of merger, the mis-specified model overpredicts changes in prices, shares, consumer and producer surplus, and these overpredictions increasing in underlying royalty rate. However, opposite implied policy recommendations under the mis-specified model are less frequent.

2.5 Conclusion

We use simulations to examine estimation bias in marginal costs and prediction bias in merger evaluations due to not accounting for existing patent licensing relationships between manufacturers. We examine two types of mergers: mergers between a licensor and a licensee and mergers between licensees. We use two sets of simulation approaches for two complementary purposes: theoretical simulations for examining the relationships between model primitives and biases; and guidance simulations to show the potential magnitude of the biases and for what types of observed market features could be associated with larger biases. We find that estimation bias in marginal costs are increasing with royalty rates and the sum of diversion ratios between licensor and licensees. The guidance simulation shows that such bias is typically not ignorable: the median estimation bias is 28% of true marginal cost values.

We show an alarming finding to researchers: not accounting for existing patent licensing relationships can lead to opposite prediction of merger effects. Such opposite predictions show up in both theoretical simulations and guidance simulations for both licensor-licensee mergers and licensee-licensee mergers. Because the guidance simulations cover a richer set of markets (including asymmetric markets) compared to the theoretical simulations that are based on a symmetric benchmark market, opposite predictions show up more frequently in guidance simulations. In particular, in the theoretical simulations, we find 4.3% to 4.4% of cases where the merger effects on prices predicted from a mis-specified model not accounting for existing patent licensing relationships are opposite to the true merger effects predicted from a true model accounting for existing patent licensing relationships. In the guidance simulations, such fraction of opposite prediction cases can be as high as 51.2%. Moreover, both sets of simulations find that, opposite predictions happen when royalty rates are large. In particular, theoretical simulation finds that opposite predictions happen when royalty rates account for about 80% of total marginal costs; while guidance simulation finds that opposite predictions happen at high values of royalty rates in the random sample of markets. We also find the economic driving forces for opposite predictions in each type of mergers. In licensor-licensee mergers, we find that ignored saving of royalty payment due to licensor-licensee merger leads to decrease of prices and increase of consumer welfare after merger, which causes opposite predictions if one assumes away the patent licensing relationships. In licensee-licensee mergers, we find that ignored decrease in alignment effect due to licensee-licensee merger leads to decrease of licensor's price, which causes opposite predictions if one assumes away the patent licensing relationships. We argue that these two ignored channels when assuming away existing patent licensing relationships are also driving forces for overall prediction biases that are unconditional on opposite predictions.

In both sets of simulations and both types of mergers, we find that market-level merger effects are over-predicted when one assumes away existing patent licensing relationships. We further find that, the prediction biases are smaller in licensee-licensee merger than those in licensor-licensee merger. In particular, the guidance simulation finds that, in licensor-licensee mergers, assuming away existing patent licensing relationships lead to over-predicted increase of share-weighted average price by 2.6% at the median, and over-predicted decrease of consumer welfare by 9.3% at the median; in licensee-licensee mergers, the two median predictions biases are 0.5% and 1.4% respectively.

We also examine the relationships between prediction biases, royalty rates, and sum of diversion ratios between licensor and licensees. Theoretical simulations show that prediction biases increase with royalty rates; while the guidance simulations also shows positive correlations between the two. On the other hand, the theoretical simulation exhibits a non-monotonic relationship between prediction biases and the sum of diversion ratios; while the guidance simulations show that the potential for prediction biases (i.e., the maximum prediction biases) is positively correlated with the sum of diversion ratios.

There are two important caveats on the analysis and findings in this chapter. First, we assume an exogenous royalty rate that does not change upon merger. Additionally, we impose that merging parties must merger, which may not be realistic if a given merger is not profitable. In the future, we will exclude mergers that are not profitable from our simulation analysis to allay this concern.

Additionally, as we aim to extend this theoretical exercise to understand the extent of estimation and prediction bias in a real-life context, we hope to calibrate our model to an empirical setting in future research.

Table 2.1: The Benchmark Market in Theoretical Simulations

(a) Product Specific Values	A	B	C	(b) Common Values	
Mean Utility ($\boldsymbol{\delta}$)	1/3	1/3	1/3	Price Coefficient (α)	0.2
Marginal Cost (c)	1	1	1	T1EV Scale (σ_ϵ)	0.1
Royalty Rate (r)	0	0	0	Market Size(M)	1.0
Market Share (%)*	25%	25%	25%	Taste Heterogeneity (σ)	0
Price	5/3	5/3	5/3	Travel Parameter (ρ)	0
Elasticity*	-2.5	-2.5	-2.5		

Notes: Equilibrium market outcomes that are used to back out the benchmark model primitives $(\boldsymbol{\delta}, \alpha)$, given $(\boldsymbol{c}, \boldsymbol{r}, \sigma_\varepsilon, \sigma, \rho)$.

Table 2.2: Sample Statistics

Percentile	5%	25%	50%	75%	95%
Royalty Rate (r)	0.019	0.11	0.225	0.353	0.538
Shares:					
s_A	0.046	0.172	0.267	0.342	0.483
s_B	0.035	0.145	0.246	0.33	0.469
s_C	0.034	0.15	0.246	0.333	0.444
Total Inside Share	0.75	0.75	0.75	0.75	0.75
HHI:					
Pre-Merger	1895	2025	2256	2617	3274
Post-AB Merger	2636	2799	3026	3553	4721
Δ from AB Merger	145	505	760	1167	2117
Post-BC Merger	2558	2633	2737	2924	3534
Δ from BC Merger	72	356	527	615	692
Margins (Pre-Merger):					
m_A	0.275	0.366	0.498	0.606	0.708
m_B	0.302	0.414	0.567	0.714	0.891
m_C	0.293	0.429	0.57	0.712	0.898
Share-Weighted Avg.	0.286	0.401	0.539	0.673	0.862
Costs:					
c_A	0.292	0.394	0.502	0.634	0.725
c_B	0.333	0.577	0.701	0.79	0.864
c_C	0.361	0.578	0.704	0.791	0.869
Share-Weighted Avg.	0.326	0.486	0.602	0.701	0.789
Market Elasticity	-7.187	-4.625	-3.369	-2.494	-1.707

Note: Market elasticity under logit demand is given by $\alpha\bar{p}(1 - s_0)$, where $\bar{p}$ represents the share-weighed average price.

Table 2.3: Bias in Cost Estimation

	5%	25%	50%	75%	95%
$\%\Delta c_A$	0.025	0.13	0.28	0.493	0.987

Note: Bias in cost estimation is defined as the percentage difference between costs implied by the mis-specified model and the true costs, i.e., $\%\Delta c_j = (\tilde{c}_j - c_j)/c_j$. Given licensees' first order conditions are identical between the mis-specified and true models, $\%\Delta c_B = \%\Delta c_C = 0$.

Table 2.4: Theoretical Simulation: Predict the Effects of Licensor-Licensee Merger

	(1) avg. prediction bias (%)	(2) avg. true effect (%)	(3) over-prediction (%)	(4) opposite prediction (%)
Consumer Surplus	-2.414	-10.668	100	0
Sales-weighted Avg. Price	2.655	8.283	100	0.03
Price: A	0.65	16.334	98.812	0
Price: B	6.237	8.302	100	4.307
Price: C	0.887	2.612	100	0
Total Quantities	-1.761	-5.85	100	0
Quantity: A	1.218	-18.916	0	0
Quantity: B	-7.664	-4.987	99.594	25.228
Quantity: C	1.249	5.577	100	0
Producer Surplus	2.773	4.715	100	0
Profit: A+B	2.206	2.516	100	0
Profit: C	2.635	9.65	100	0

Note: results are conditional on positive royalty rates.

Table 2.5: Theoretical Simulation: Mis-specified Model Predicts Opposite Licensor-Licensee Merger Effects

	(1) royalty rates $\overline{r}$	(2) diversion ratios $\overline{d}$	(3) travel parameters $\overline{\rho}$	(4) true effects $\overline{\%\Delta y}$	(5) mis-specified effects $\overline{\%\Delta \tilde{y}}$	(6) prediction biases $\overline{\Delta\Delta y}$
Sales-weighted Avg. Price	0.99	0.907	1	-0.159	3.931	4.090
consumer surplus				-1.067	-2.819	-1.752
total quantities				-0.712	-1.898	-1.186
producer surplus				2.363	12.970	10.606
Price: B	0.947	0.402	0.313	-0.778	10.835	11.613
consumer surplus				-6.100	-11.101	-5.002
sales-weighted avg. price				4.671	9.596	4.925
total quantities				-4.272	-7.863	-3.591
producer surplus				1.970	4.538	2.568
Quantity: B	0.859	0.464	0.422	2.733	-10.336	-13.069
consumer surplus				-6.491	-10.797	-4.307
sales-weighted avg. price				5.493	10.051	4.558
total quantities				-4.557	-7.666	-3.109
producer surplus				2.629	6.235	3.606

Note: results are conditional on opposite predictions.

Table 2.6: Bias in Licensor-Licensee Merger Prediction

	5%	25%	50%	75%	95%
Prices:					
p_A	-0.009	0	0.002	0.005	0.013
p_B	0.007	0.041	0.08	0.126	0.2
p_C	0	0.002	0.005	0.009	0.018
Share-Weighted Avg.	0.001	0.009	0.026	0.053	0.098
Shares:					
s_A	0.001	0.013	0.044	0.12	0.254
s_B	-1.037	-0.486	-0.228	-0.092	-0.011
s_C	0.002	0.018	0.058	0.152	0.357
Inside Shares	-0.139	-0.072	-0.032	-0.011	-0.001
Profits:					
π_A	0.016	0.131	0.289	0.454	0.633
π_B	-4.415	-1.972	-0.813	-0.307	-0.035
π_C	0.003	0.027	0.083	0.197	0.403
Producer Surplus	0.003	0.024	0.048	0.075	0.1
Consumer Surplus	-0.509	-0.235	-0.093	-0.033	-0.004

Note: Post-merger licensing revenues allocated to π_A.

Table 2.7: Average Bias and Opposite Predictions in Licensor-Licensee Mergers

	(1) Avg. True Effect	(2) Avg. Prediction Bias	(3) Opposite Prediction
Prices:			
p_A	0.099	0.002	0
p_B	0.014	0.089	0.499
p_C	0.003	0.006	0.326
Share-Weighted Avg.	0.02	0.035	0.363
Shares:			
s_A	-0.269	0.077	0.011
s_B	0.126	-0.342	0.609
s_C	0.049	0.103	0.326
Inside Shares	-0.02	-0.047	0.327
Profits:			
π_A	-0.272	0.3	0.702
π_B	1.384	-1.367	0.245
π_C	0.057	0.13	0.326
Producer Surplus	0.029	0.05	0.091
Consumer Surplus	-0.05	-0.157	0.327

Note: Average true effect is given by percent change. Average prediction bias is given by the difference in percent changes between the mis-specified model. Opposite prediction represents a different sign in counterfactual prediction of a given measure between the true and mis-specified model.

Table 2.8: Opposite Prediction Market Characteristics

Percentile	5%	25%	50%	75%	95%
Royalty Rate (r)	0.168	0.27	0.372	0.486	0.603
Shares:					
s_A	0.017	0.098	0.204	0.297	0.429
s_B	0.041	0.163	0.271	0.354	0.509
s_C	0.063	0.188	0.278	0.354	0.518
Total Inside Share	0.75	0.75	0.75	0.75	0.75
HHI:					
Pre-Merger	1894	2029	2324	2700	3360
Post Merger	2858	3058	3395	3956	5263
Δ from Merger	133	603	1118	1651	2536
Margins (Pre-Merger):					
m_A	0.267	0.362	0.483	0.576	0.68
m_B	0.351	0.479	0.627	0.784	0.92
m_C	0.359	0.481	0.645	0.771	0.919
Share-Weighted Avg.	0.339	0.472	0.618	0.749	0.87
Costs:					
c_A	0.32	0.424	0.517	0.638	0.733
c_B	0.516	0.689	0.772	0.835	0.889
c_C	0.457	0.694	0.775	0.826	0.892
Share-Weighted Avg.	0.447	0.582	0.664	0.75	0.823
Market Elasticity	-8.546	-5.797	-4.531	-3.5	-2.474

Note: Market elasticity under logit demand is given by $\alpha\bar{p}(1 - s_0)$, where $\bar{p}$ represents the share-weighed average price

Table 2.9: Theoretical Simulation: Predict the Effects of Licensee-Licensee Merger

	(1) avg. prediction bias (%)	(2) avg. true effect (%)	(3) over-prediction (%)	(4) opposite prediction (%)
Consumer Surplus	-0.203	-10.888	99.921	0
Sales-weighted Avg. Price	0.253	11.27	99.95	0
Price:A	0.446	2.221	97.911	4.446
Price:B	0.103	16.819	97.97	0
Price:C	0.093	16.722	97.089	0
Total Quantities	-0.152	-7.089	99.921	0
Quantity: A	-0.687	10.599	0.01	0
Quantity: B	0.079	-15.631	0.386	0
Quantity: C	0.099	-15.048	0.198	0
Producer Surplus	2.959	4.901	100	0
Profit:A	9.571	5.581	100	4.119
Profit:B+C	0.237	4.678	99.98	0

Note: results are conditional on positive royalty rates.

Table 2.10: Theoretical Simulation: Mis-specified Model Predicts Opposite Licensee-Licensee Merger Effects

	(1) royalty rates	(2) diversion ratios	(3) travel parameters	(4) true effects	(5) mis-specified effects	(6) prediction biases
	$\overline{r}$	$\overline{d}$	$\overline{\rho}$	$\overline{\%\Delta y}$	$\overline{\%\Delta \tilde{y}}$	$\overline{\Delta\Delta y}$
Price: A	0.801	0.348	0.085	-0.191	0.199	0.391
consumer surplus				-9.757	-9.967	-0.210
sales-weighted avg. price				7.500	7.691	0.190
total quantities				-6.018	-6.167	-0.149
producer surplus				1.046	2.958	1.912
Profit: A	0.953	0.370	0.177	-0.225	6.458	6.683
consumer surplus				-9.838	-10.138	-0.300
sales-weighted avg. price				7.955	8.238	0.283
total quantities				-6.278	-6.496	-0.219
producer surplus				0.909	3.345	2.435

Note: results are conditional on opposite predictions.

Table 2.11: Bias in Licensee-Licensee Merger Prediction

	5%	25%	50%	75%	95%
Prices:					
p_A	0	0.003	0.009	0.018	0.038
p_B	0	0.001	0.002	0.003	0.006
p_C	0	0.001	0.002	0.003	0.006
Share-Weighted Avg.	0	0.002	0.005	0.009	0.017
Shares:					
s_A	-0.237	-0.07	-0.026	-0.008	-0.001
s_B	0	0.001	0.004	0.01	0.022
s_C	0	0.001	0.004	0.01	0.022
Inside Shares	-0.026	-0.011	-0.005	-0.002	0
Profits:					
π_A	0.004	0.044	0.112	0.238	0.564
π_B	0.011	0.135	0.549	1.803	12.361
π_C	0.012	0.13	0.523	1.848	13.756
Producer Surplus	0.024	0.14	0.282	0.423	0.591
Consumer Surplus	-0.07	-0.034	-0.014	-0.005	-0.001

Note: Post-merger licensee profits allocated to each respective pre-merger licensee product.

Table 2.12: Average Bias and Opposite Predictions in Licensee-Licensee Mergers

	(1) Avg. True Effect	(2) Avg. Prediction Bias	(3) Opposite Prediction
Prices:			
p_A	-0.003	0.012	0.512
p_B	0.098	0.002	0
p_C	0.099	0.002	0
Share-Weighted Avg.	0.048	0.006	0
Shares:			
s_A	0.2	-0.061	0
s_B	-0.203	0.007	0.006
s_C	-0.205	0.007	0.002
Inside Shares	-0.054	-0.008	0
Profits:			
π_A	0.002	0.171	0.513
π_B	-17.953	17.976	0.623
π_C	-4.885	4.907	0.624
Producer Surplus	-0.214	0.29	0.876
Consumer Surplus	-0.172	-0.022	0

Note: Post-merger licensee profits allocated to each respective pre-merger licensee product.

Table 2.13: Opposite Prediction Market Characteristics

Percentile	5%	25%	50%	75%	95%
Royalty Rate (r)	0.069	0.16	0.249	0.378	0.551
Shares:					
s_A	0.05	0.175	0.269	0.345	0.491
s_B	0.03	0.139	0.242	0.326	0.473
s_C	0.029	0.145	0.244	0.333	0.452
Total Inside Share	0.75	0.75	0.75	0.75	0.75
HHI:					
Pre-Merger	1895	2029	2269	2642	3345
Post Merger	2564	2641	2744	2926	3538
Δ from Merger	68	339	522	616	696
Margins (Pre-Merger):					
m_A	0.276	0.365	0.498	0.606	0.707
m_B	0.31	0.419	0.576	0.715	0.891
m_C	0.298	0.438	0.575	0.721	0.903
Share-Weighted Avg.	0.315	0.439	0.577	0.709	0.846
Costs:					
c_A	0.293	0.394	0.502	0.635	0.724
c_B	0.437	0.619	0.719	0.801	0.867
c_C	0.411	0.611	0.726	0.8	0.873
Share-Weighted Avg.	0.371	0.513	0.617	0.709	0.791
Market Elasticity	-7.382	-4.805	-3.63	-2.675	-1.923

Note: Market elasticity under logit demand is given by $\alpha\bar{p}(1-s_0)$, where $\bar{p}$ represents the share-weighed average price

Figure 2.1: Theoretical Simulation: marginal cost estimation bias and royalty rates and the sum of diversion ratios between the licensor and the two licensees.

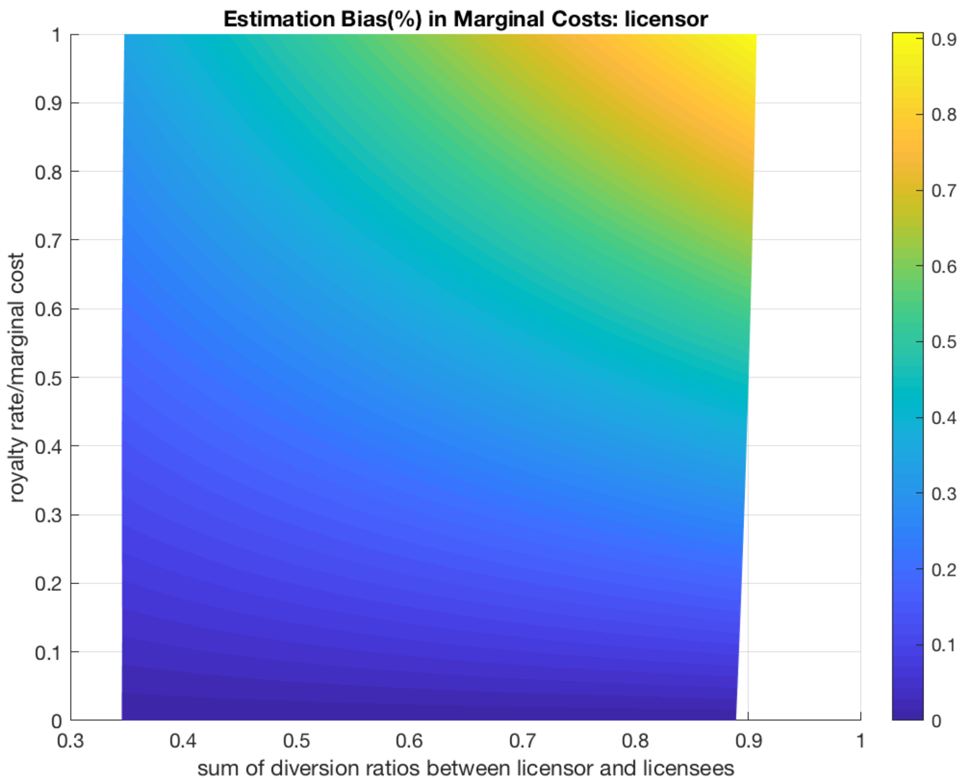


Figure 2.2: Guidance Simulations: estimation bias in costs between mis-specified and true models for the licensor (firm A)

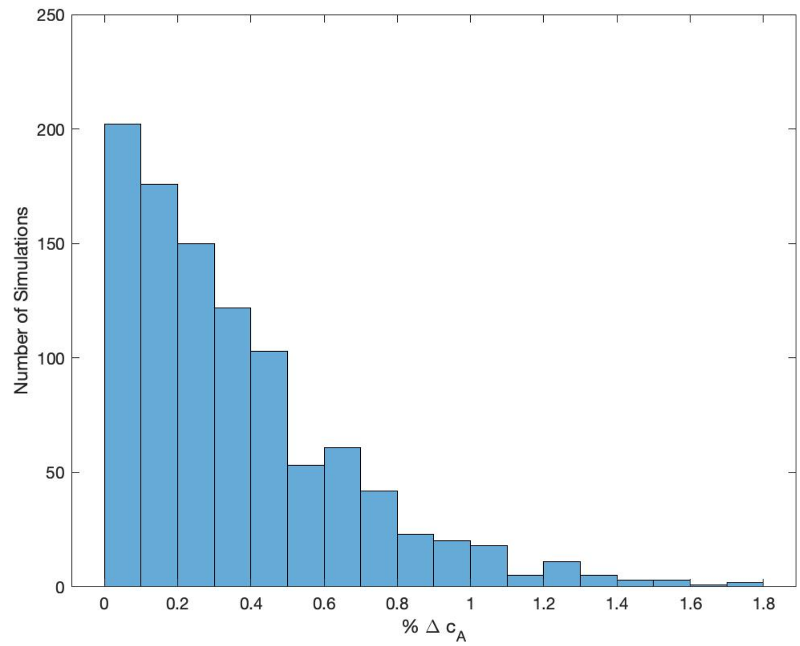


Figure 2.3: Guidance Simulations: estimation bias in costs between mis-specified and true models for firm A (licensor) by HHI

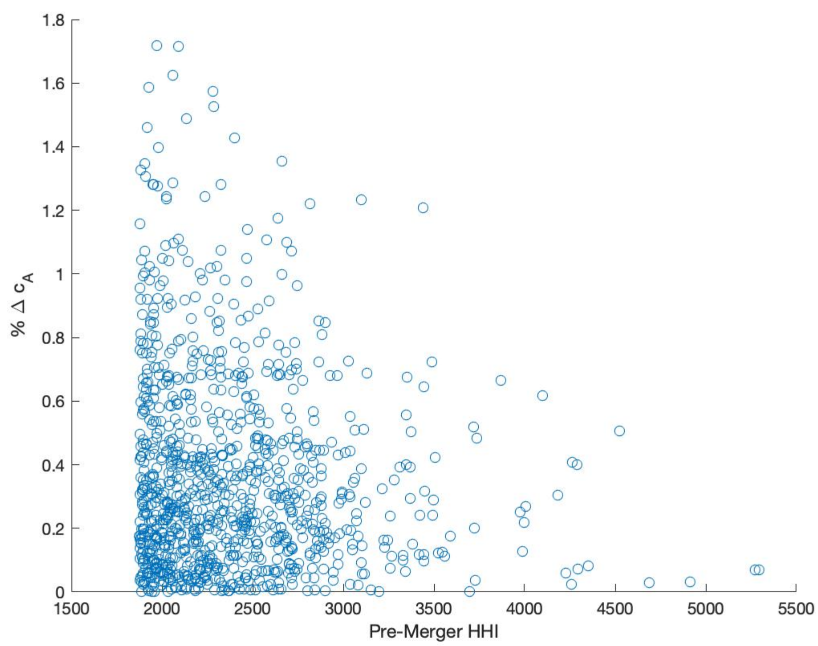


Figure 2.4: Guidance Simulations: estimation bias in costs between mis-specified and true models for firm A (licensor) by model parameters

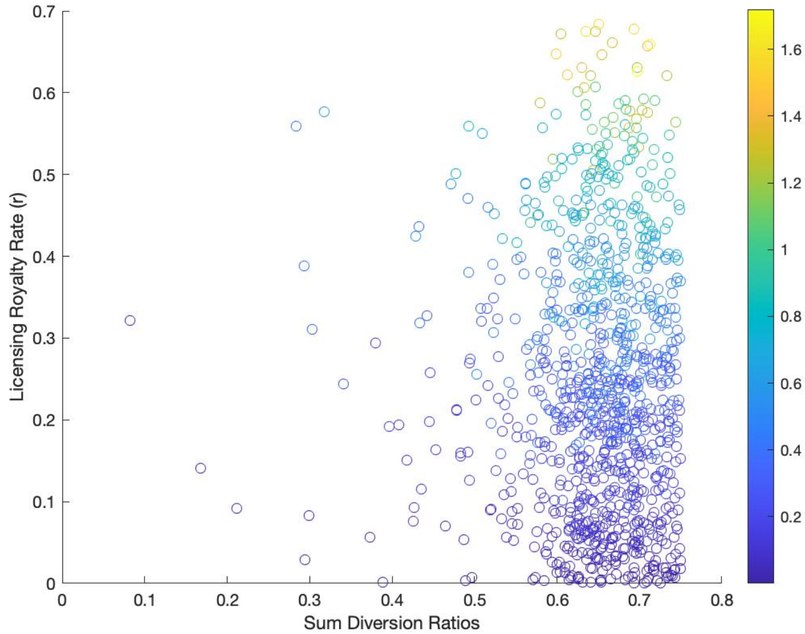


Figure 2.5: Theoretical Simulation: prediction bias in consumer welfare merger effect from the merger between licensor and licensee.

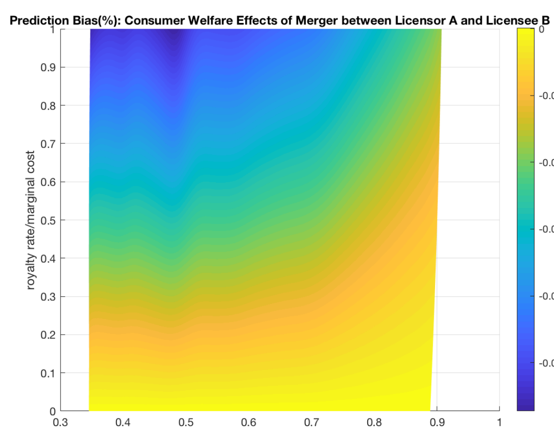


Figure 2.6: Theoretical Simulation: prediction bias in sales-weighted average price merger effect from the merger between licensor and licensee.

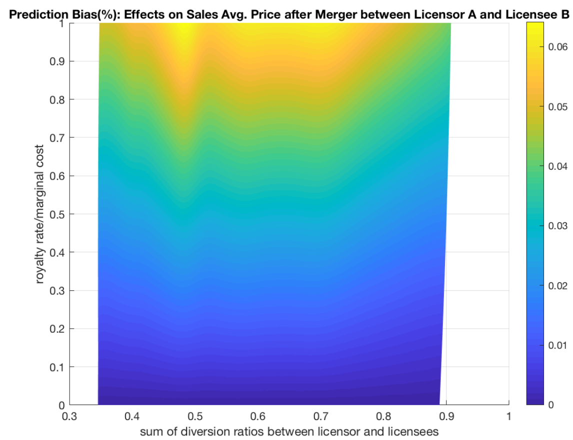


Figure 2.7: Theoretical Simulation: prediction bias in total quantity merger effect from the merger between licensor and licensee.

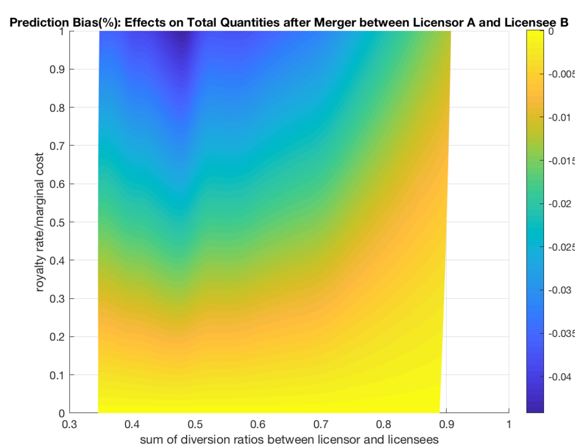


Figure 2.8: Theoretical Simulation: prediction bias in producer surplus merger effect from the merger between licensor and licensee.

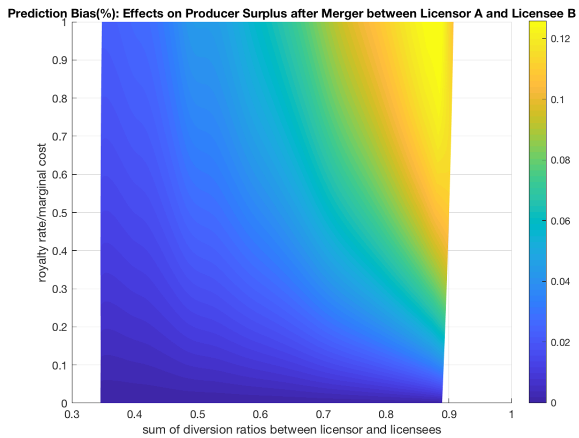


Figure 2.9: Guidance Simulations: prediction bias in the share-weighted average price merger effect from the merger between licensor and licensee by licensing royalty rate.

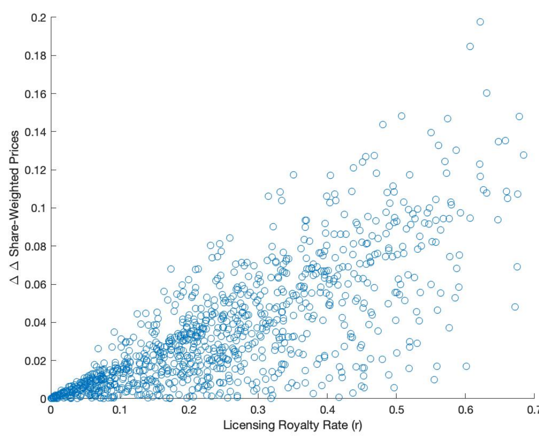


Figure 2.10: Guidance Simulations: prediction bias in the total quantities merger effect from the merger between licensor and licensee by licensing royalty rate.

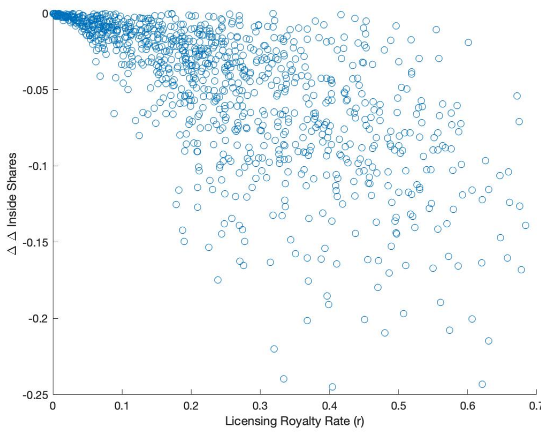


Figure 2.11: Guidance Simulations: prediction bias in the consumer surplus merger effect from the merger between licensor and licensee by licensing royalty rate.

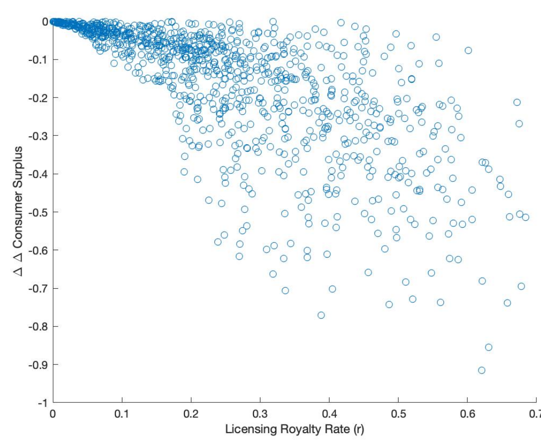


Figure 2.12: Guidance Simulations: prediction bias in the producer surplus merger effect from the merger between licensor and licensee by licensing royalty rate.

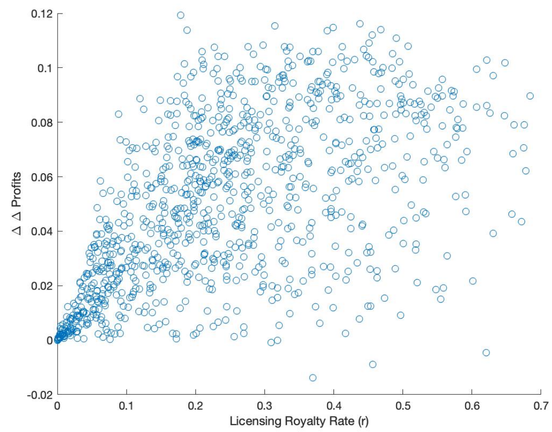


Figure 2.13: Guidance Simulations: prediction bias in the share-weighted average price merger effect from the merger between licensor and licensee by sum of diversion ratios.

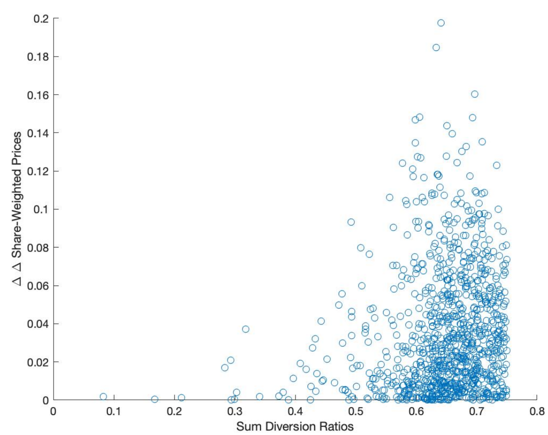


Figure 2.14: Guidance Simulations: prediction bias in the total quantities merger effect from the merger between licensor and licensee by sum of diversion ratios.

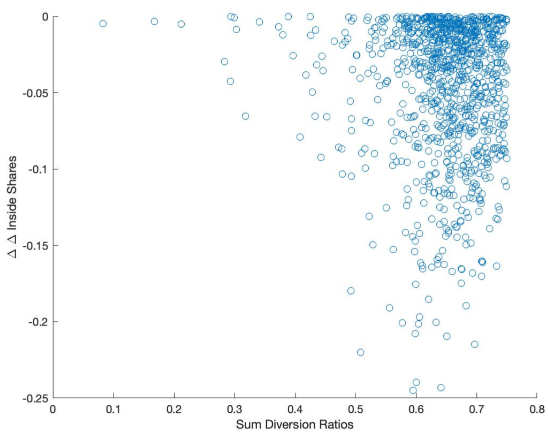


Figure 2.15: Guidance Simulations: prediction bias in the consumer surplus merger effect from the merger between licensor and licensee by sum of diversion ratios.

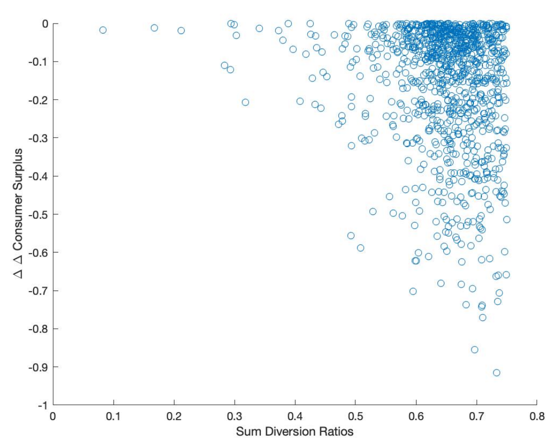


Figure 2.16: Guidance Simulations: prediction bias in the producer surplus merger effect from the merger between licensor and licensee by sum of diversion ratios.

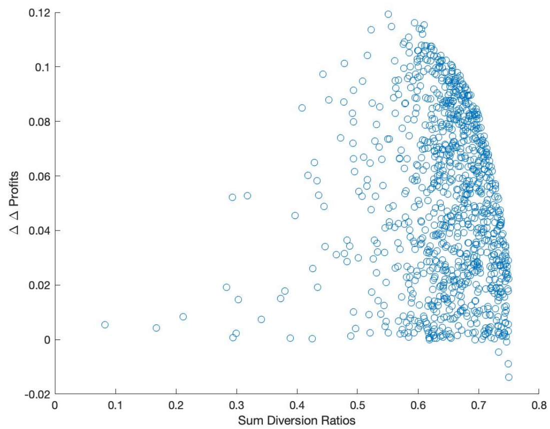


Figure 2.17: Theoretical Simulation: prediction bias in consumer surplus merger effect from the merger between licensees.

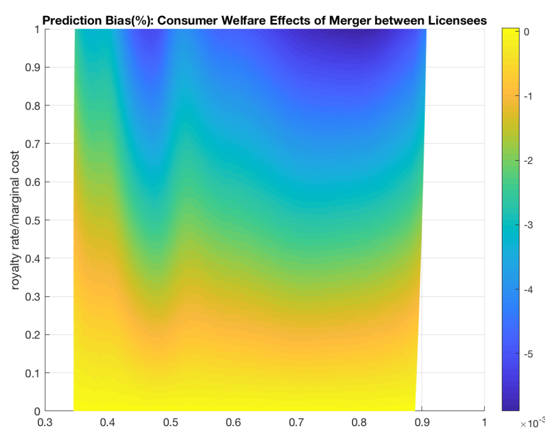


Figure 2.18: Theoretical Simulation: prediction bias in sales-weighted average price merger effect from the merger between licensees.

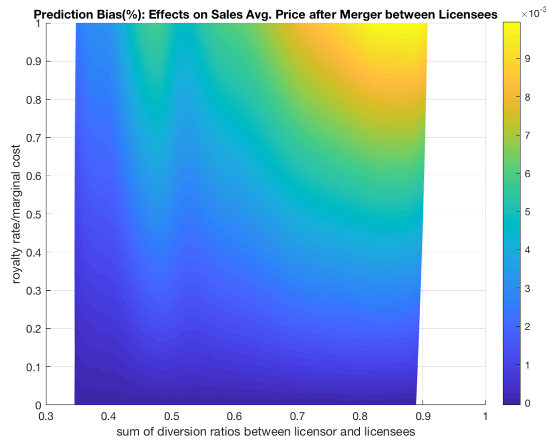


Figure 2.19: Theoretical Simulation: prediction bias in total quantities merger effect from the merger between licensees.

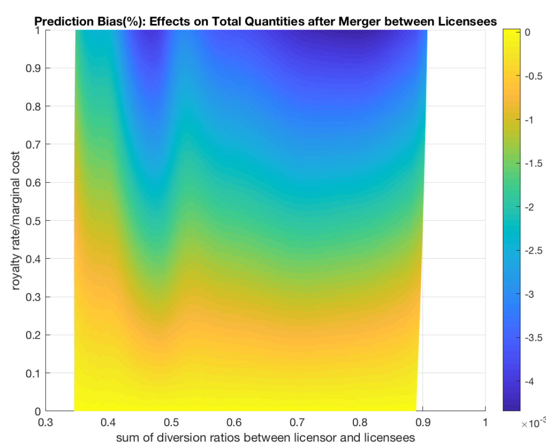


Figure 2.20: Theoretical Simulation: prediction bias in producer surplus merger effect from the merger between licensees.

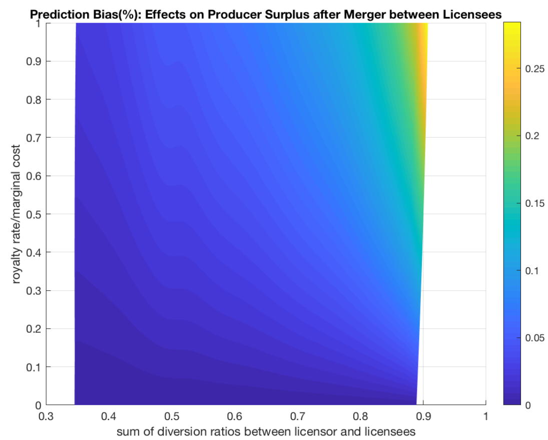


Figure 2.21: Guidance Simulations: prediction bias in the share-weighted average price merger effect from the merger between licensees by licensing royalty rate.

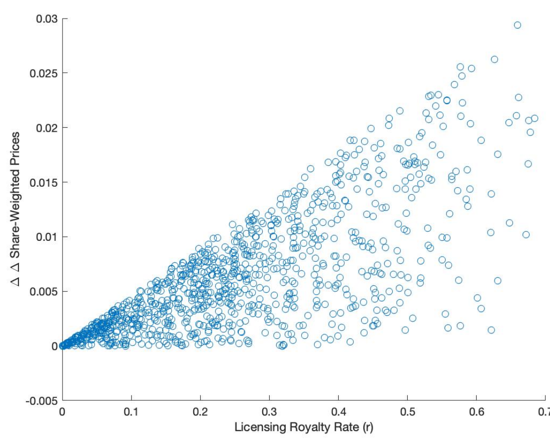


Figure 2.22: Guidance Simulations: prediction bias in the share-weighted price merger effect from the merger between licensees by licensing royalty rate.

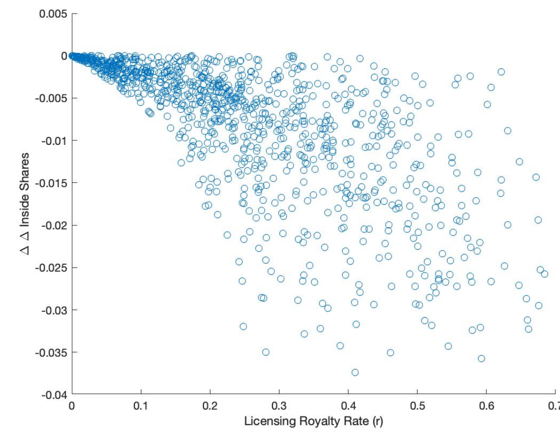


Figure 2.23: Guidance Simulations: prediction bias in the consumer surplus merger effect from the merger between licensees by licensing royalty rate.

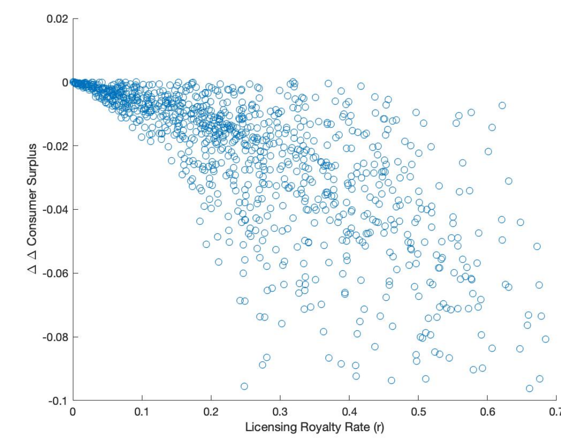


Figure 2.24: Guidance Simulations: prediction bias in the producer surplus merger effect from the merger between licensees by licensing royalty rate.

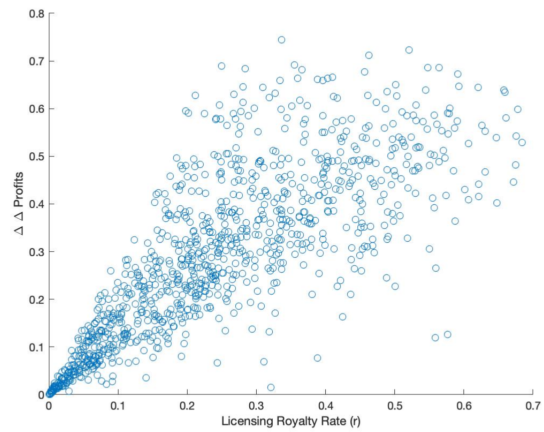


Figure 2.25: Guidance Simulations: prediction bias in the share-weighted average price merger effect from the merger between licensees by sum of diversion ratios.

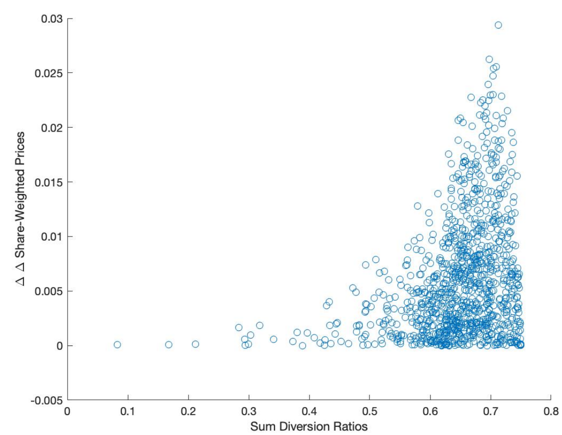


Figure 2.26: Guidance Simulations: prediction bias in the total quantities merger effect from the merger between licensees by sum of diversion ratios.

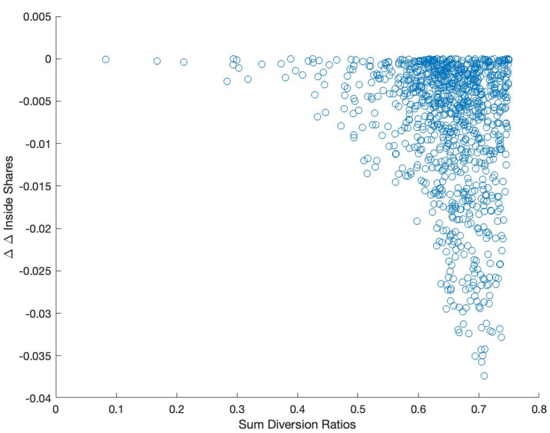


Figure 2.27: Guidance Simulations: prediction bias in the consumer surplus merger effect from the merger between licensees by sum of diversion ratios.

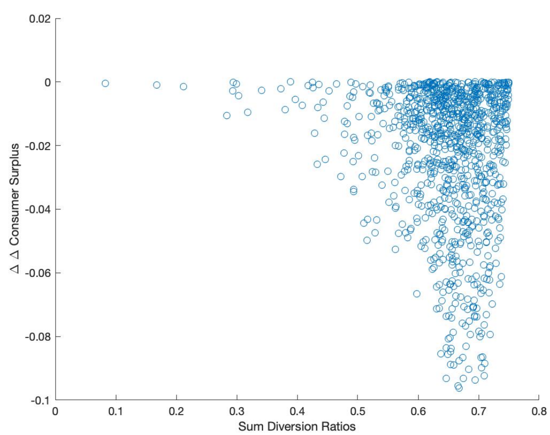
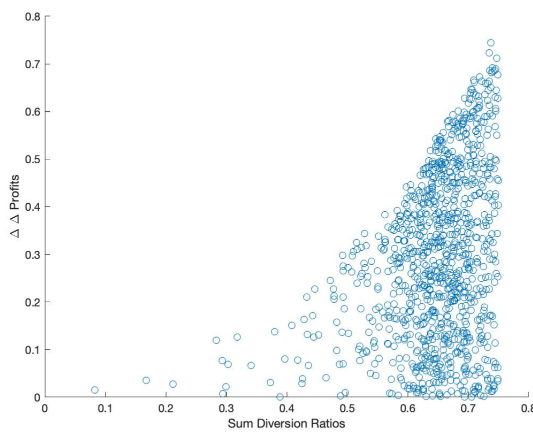


Figure 2.28: Guidance Simulations: prediction bias in the producer surplus merger effect from the merger between licensees by sum of diversion ratios.



Chapter 3

Product Responses to Income-Based Subsidies in the U.S. Infant Formula Industry

3.1 Introduction

Consumption subsidies are a common tool introduced in markets to encourage uptake of a good or service. A complete understanding of the effect these consumption subsidies have on market outcomes is important in informing welfare-enhancing policy.

Subsidizing the purchase of a good or service will have a direct impact on demand. Demand subsidies may induce consumers to purchase a good or service that exhibits positive externalities who were not previously, such as in education. Alternatively, these subsidies may aim to induce consumers to substitute to an alternative good or service, such as to electronic vehicles, away from one that exhibits negative externalities, such as from combustion engine vehicles.

Due to consumption subsidies having a direct effect on demand, firms, as demand-takers, are likely to respond. Firms can respond to demand-side subsidies in a number of ways, including prices and quality (i.e., product characteristics). For example, firms may respond to a consumption subsidy by increasing price(s) or lowering the quality of their existing product(s). We would expect each of these responses to attenuate consumer welfare relative to when price(s) and quality are fixed, respectively. Thus, to accurately assess the implications of demand-side subsidy policies and resulting welfare, economic models must accurately incorporate relevant supply-side responses.

An additional key margin to consider in evaluating the effects of demand-side subsidies is producer product offering adjustment. That is, firms may adjust the number of products they offer in response to a subsidy policy change, which will impact resulting welfare. Consider a newly-implemented demand-side subsidy. As this subsidy alters the price consumers face, we would expect marginal consumers to adjust purchasing behavior, whether switching to/from another product or to/from purchasing a product at all, depending on their price sensitivity. Thus, given these substitution patterns, the subsidy will affect the aggregate market shares amongst products differently. If consumers shift sufficiently away from a given product due to the implemented demand-side subsidy, it may no longer be profitable for the supplying firm to offer this product. This reduction in the number of product offerings will impact the resulting consumer welfare through available variety.

This chapter will look to understand this product adjustment margin, while incorporating price responses, and the associated effects on welfare. Concisely, what are the (short-run) welfare effects of income-based consumption subsidies when accounting for firm pricing and product offerings responses?

This chapter is related to three strains of the literature. First, the subsidy literature has investigated how firms in imperfect competition respond to demand-side subsidies in a variety of contexts. Decarolis (2015) and Jaffe and Shepard (2018) find pricing distortions due to subsidies in health insurance markets. A number of studies incorporate product quality responses in health insurance markets (see e.g., Starc and Town (2015)). Springel (2017) and Li (2017) evaluate subsidies in the electric car market and the effect demand subsides have on firm investment. This chapter contributes to the subsidy literature by incorporating the role of the product offering adjustment margin, in addition to price responses, in accessing welfare effects of a demand-side subsidy.

Second, this chapter adds to the literature evaluating the U.S. infant formula industry and associated Special Supplemental Nutrition Program for Women, Infants, and Children ("WIC") subsidies. For example, Oliveira, Frazao, and Smallwood (2010) show that between 2004–2006, adjusting for inflation, manufacturer prices of infant formula purchased by vouchers increased significantly. In addition, Oliveira et al. (2004) find, using cross-sectional evidence, that geographic markets with a higher proportion of the population eligible for WIC subsidies face higher prices for infant formula. This chapter contributes to the literature on the U.S. infant formula industry by building a structural model, featuring endogenous product offerings, to evaluate relevant policy counterfactuals.

Third, this chapter uses the framework of the endogenous product choice literature (as reviewed in Crawford (2012)) to evaluate welfare implications of counterfactual policy. Incorporating price responses, this literature has evaluated the policy implications of endogenous product characteristics (e.g., Fan (2013) and Sweeting (2010)) and endogenous product offerings (e.g., Wollman (2017), Fan and Yang (2020), and Draganska, Mazzeo and Seim (2009)). Most closely related to this work is Guglielmo (2016), which builds a structural model that features endogenous coverage quality to back out provider costs for individuals in two substitutable Medicare programs and evaluate a counterfactual subsidy policy. Two main differences exist between Guglielmo (2016) and this chapter. First, as detailed below, in my context, consumers do not choose which subsidy program to enroll in, instead, consumers' subsidization status is (plausibly) exogeneously determined. Second, in my model, I consider the product offerings margin, instead of firms choosing product characteristics. This chapter contributes to the literature by assessing the role of endogenous product choice play in evaluating the welfare effects of a demand-side subsidy.

3.2 Empirical Setting

I will investigate this research question in the context of the U.S. infant formula industry. This industry is an ideal setting to study my research question because it exhibits a number of appealing qualities: products are differentiated by observable physical characteristics, supply is relatively concentrated, and only a portion of consumers receive subsidies.

Infant formula is used as a substitute for breast milk for infants between birth and twelve months of age. While breast milk is recommended as the optimal source of nutrition for infants by the members of the scientific community, such as the American Academy of Pediatrics and regulatory agencies, such as the U.S. Department of Agriculture, due to choice or necessity, for example, mothers with HIV avoiding transmission to their infants, manufactured infant formula can supplement or replace breast milk in an infant's diet.

Infant formula is produced in three mutually exhaustive forms: i) powder, which requires the addition of boiled water, in an approximate ratio of seven parts water to one part powder formula, to reconstitute to a consumable form; ii) liquid concentrate, which also requires the addition of water, at a ratio of one-to-one; and iii) ready-to-feed, which requires no additional preparations prior to consumption.[1] Infant formula, in any of the three forms, is primarily made up of a protein, often derived from cow's milk, a carbohydrate, vegetable fats, and a micronutrient mix, to mimic naturally occurring human breast milk.[2]

Infant formula in the U.S. is largely supplied by three primary national manufacturers. These suppliers, Abbott Laboratories, Mead Johnson, and Nestlé-Gerber, account for over 90 percent of the national market share by volume from 2006–2015 (see Figure **??** below).

Infant formula manufacturers develop new infant formula over a number of years. In order to sell a new infant formula product in the U.S., suppliers must notify the U.S. Food and Drug Administration (FDA) with nutritional content and production process details 90 days prior to marketing and distributing the proposed infant formula product. This development process gives rise to a firm's portfolio of infant formula products. Products in each firm's portfolio are often marketed under the same brand. For example, Abbott markets its products under the Similac brand, such as Similac Pro-Advance or Similac Sensitive.

In the U.S., as a part of the WIC program, households below 185% of the federal poverty level are eligible to receive infant formula vouchers. These vouchers are provided by state WIC agencies and are redeemable for up to 800 reconstituted ounces per month per household of infant formula[3] at no cost to the voucher recipient. However, vouchers are only able to

[1] In my sample, I observe negligible sales of ready-to-feed formula and thus omit this form from the analysis.

[2] See Mead Johnson 2012 10-K filing

[3] Approximately equal to the average infant monthly nutritional intake.

be applied to a single brand of infant formula, which non-recipients may also purchase. The price consumers pay for this single WIC-designated brand without WIC vouchers is federally mandated to be equal to that paid by the state agency via WIC vouchers. This lack of ability to price discriminate between WIC and non-WIC consumers is captured in the model below.

The exclusive right to supply formula for purchase by WIC vouchers varies by state and is allocated by first price auction. Manufacturers submit sealed bids for the amount per unit sold they are willing to rebate the respective state agency for each unit of infant formula purchased with WIC vouchers. The manufacturer that submits the lowest net price bid, defined by the difference between the manufacturer's lowest national wholesale price per unit and a proposed rebate rate, receives the exclusive contract, effective for typically three to five years.

In this chapter, I will focus on the infant formula in Michigan between January 2012 and December 2015. During this time period, Mead Johnson's Enfamil Premium infant formula, in each of the three possible product forms, was eligible for purchase by WIC vouchers in Michigan. For each reconstituted ounce sold, Mead Johnson rebated a publicly known, contractually-fixed amount to the state WIC agency.

3.3 Data

To analyze my research question, I rely upon infant formula sales data from the Nielsen Retail Scanner Dataset, which tracks consumer purchases at point-of-sale for more than 90 retail chains across the U.S. Specifically, the sales data indicates the quantity and total expenditure for a given product at a given store, in a given week. Attributes, such as the form of the infant formula product, are also included in the data. To decrease the likelihood of the model falsely interpreting zero observed sales of a product as one not being offered, instead of in reality, consumers choosing not to purchase the produced product, I follow the literature and aggregate to the month level.[4] Additionally, since geographic information for stores are only available at the Designated Market Area ("DMA"), as defined by Nielsen, I

[4] For example, see Miravete, Seim, and Thurk (2017)

aggregate product-specific quantity and expenditures by DMA. From this, I determine the average price and total quantity purchased for each infant formula product by month-DMA.[5]

Table 3.1 below details summary statistics for the product characteristics observed in the sample. We see that during my sample period, the average price per reconstituted ounce of formula is $0.18. A large majority of product-market-month observations are in powder form in the sample (0.95). To improve the specification of the demand model (detailed below), additional physical product attributes, such as nutritional content, could be collected and implemented in the utility specification.

To measure market size, following Nevo (2001), I define market size as 750 ounces multiplied by the number of infant households in a given DMA-month. Total infant households at the state-level is provided by the Michigan Department of Health and Human Services. Number of infant households amongst the subpopulation eligible for WIC vouchers are provided by USDA Food and Nutritional Service at the state level. I weight the number of infant households in a given subpopulation at the state level by the overall population proportion belonging to each DMA in 2016, as indicated by Nielsen, to calculate the approximate number of infant households in each subpopulation in a given DMA-month.

For the rebate rate paid from Mead Johnson to the Michigan WIC agency for each voucher purchase of Enfamil Premium, I use the contractually agreed upon rate between these two entities. As stated on the contract between the State of Michigan and Mead Johnson & Company LLC, effective November 1, 2011 through October 31, 2016,[6] for each reconstituted ounce of Enfamil Premium purchased by WIC vouchers, Mead Johnson rebated $0.128 for powdered form, $0.151 for liquid concentrate form, or $0.074 for ready-to-feed form.[7] To avoid any potential intertemporal effects from the contract start and end dates, I restrict

[5]Nielsen subdivides Michigan into seven DMAs, one of which, "Alpena", I remove due to having a population less than 100 times smaller than the largest DMA population, "Detroit".

[6]Contract No. 071B1300297

[7]I do not observe any purchases of Enfamil Premium in ready-to-feed form. Federal regulation requires state WIC agencies to provide vouchers for liquid concentrate and powder forms only. Vouchers for ready-to-feed formulation may be granted by the state WIC agency for households with restricted access to sanitary water or refrigeration.

my sample to January 2012 through December 2015. From Table 3.1, we see that, among the 432 (7.3%) product-market-month observations eligible for WIC (i.e., Enfamil Premium) during this time period, the average rebate paid by Mead Johnson to the state was \$0.14 relative to an average price of \$0.18.

3.4 Model

In order to understand how income-based demand subsidies affect product offerings, and thus equilibrium outcomes and welfare measures, I develop a structural model of demand and supply of the infant formula industry that will allow me to conduct counterfactual analyses.

The model is static with firms optimally setting product offerings and prices and households optimally choosing an infant formula or the outside good.

In each period t, in every geographic market m, each firm f:

1. Simultaneously chooses a set of products J_{fmt} to produce from an existing stock of potential products $\mathscr{J}_{fmt}$;
2. Observes demand and marginal cost shocks;
3. Simultaneously sets prices of products chosen in Step 1.

Then, households choose the available product or outside option that yields the highest utility. When making this decision, households account for their WIC status, set of produced products, and their associated product characteristics, including price. The outside option to purchasing and consuming an infant formula product is breast feeding.

The model can be solved by backwards induction, starting with the household's discrete choice problem, then considering the pricing and subsequently product offering firm decisions.

3.4.1 Demand: Household's Problem

On the demand side, infant households in a given geographic market decide between infant formula products or the outside option. Adapting the models of Berry, Carnall, and Spiller (2006) and Berry and Jia (2010) that allow for different underlying preference parameters for different subpopulations, underlying household preferences are allowed to depend on their WIC voucher status. Specifically, household i belonging to subgroup $g \in \{\text{WIC}, \text{non-WIC}\}$ derives indirect utility from purchasing and consuming inside product j during period t in geographic market m, which is given by,

$$U_{igjmt} = x_j \tilde{\beta}_{ig} - \tilde{\alpha}_{ig} p^*_{gjmt} + \lambda_{fmt} + \xi_{jmt} + \epsilon_{igjmt} \tag{3.1}$$

where x_j is a $1 \times K$ vector representing the observable physical product characteristics (i.e., powder or liquid concentrate form).

The random coefficients on product attributes, $\tilde{\beta}$ and $\tilde{\alpha}$, allow within-subgroup heterogeneity in preferences over product attributes to yield reasonable substitution patterns. Accurately capturing substitution patterns is critical in understanding demand response to changes in product offerings. The random coefficient on physical product characteristic $k \in \{1, ..., K\}$ is given by,

$$\tilde{\beta}_{igk} = \beta_{gk} + \sigma_{\beta_{gk}} \nu_{\beta_{gk}}, \text{ where } \nu_{\beta_{gk}} \sim N(0, 1) \tag{3.2}$$

β_{gk} captures the mean preference for product characteristic k amongst the subgroup g and $\sigma_{\beta_{gk}}$ the variation in this preference within g.

Price per reconstituted ounce in a given market-time period depends on the WIC status of the consumer and the product. WIC voucher holders purchasing a WIC-designated infant formula product face no price, while non-WIC products and non-WIC consumers face manufacturer-set prices. That is,

$$p^*_{gjmt} = \begin{cases} 0, & \text{if } g = WIC, j = WIC \\ p_{gjmt}, & \text{otherwise} \end{cases} \tag{3.3}$$

I make the reasonable assumption that WIC consumers only consider WIC products or the outside good (i.e., they do not purchase non-WIC products). Therefore, given WIC households purchasing WIC products do not face prices, I only consider a random coefficient on price for the non-WIC subpopulation, which is given by,

$$\tilde{\alpha}_{igk} = \alpha_g + \sigma_{\alpha_g}\nu_{\alpha_g}, \text{ where } \nu_{\alpha_g} \sim N(0,1) \text{ for } g = \text{WIC} \tag{3.4}$$

Similar to the random coefficient on product characteristics, within the subgroup $g =$ WIC, mean utility for price is given by α_g and variance σ_{α_g}.

To capture a firm-specific component to utility, λ_{fmt} contains a firm fixed effect. Similarly, market and time fixed effects are included to capture differences in preferences across geographic markets and over time. Thus, the demand shock ξ_{mtj} represents the market-time-specific unobservable deviation from mean valuation.[8] Following the discrete choice literature, ϵ_{igjmt} follows the Type I extreme value distribution and is independently and identically distribution (i.i.d.) across households. The utility of the outside good, breast feeding, is normalized to zero, i.e.,

$$U_{ig0mt} = \epsilon_{ig0mt} \tag{3.5}$$

where, as with the inside good utilities, ϵ_{ig0mt} follows the Type I extreme value distribution and i.i.d.

For notational ease, let $\theta_{WIC} \equiv (\alpha_{WIC}, \beta_{WIC}, \sigma_{\alpha_{WIC}}, \sigma_{\beta_{WIC}})$,
$\theta_{non-WIC} \equiv (\beta_{non-WIC}, \sigma_{\beta_{non-WIC}})$, and $\theta \equiv (\theta_{WIC}, \theta_{\text{non-}WIC}, \lambda)$.

Given the distributional assumption on the error term, ϵ_{igjmt}, the model implies average purchase probabilities for each subgroup g by integrating over the joint distribution of random coefficients, F. Subgroup-specific average purchase probabilities are given by,

$$s^g_{jmt}(x, p, \xi; \theta_g, \lambda_f) = \int \frac{exp(x_j\tilde{\beta}_{ig} - \tilde{\alpha}_{ig}p_{gjmt} + \lambda_f + \xi_{jmt})}{1 + \sum_k exp(x_k\tilde{\beta}_{ig} - \tilde{\alpha}_{ig}p_{gkmt} + \lambda_f + \xi_{kmt})} dF_g \tag{3.6}$$

[8] See discussion in Nevo (2001)

Ideally, estimation of θ_g would come from taking this expression to data; however, subgroup-specific market shares are not observed in the data. Therefore, to bring the demand side of the model to observed overall market shares, subgroup purchase probabilities in equation (6) must be weighted by the observed proportion of households belonging to each subgroup, ζ_{gmt},

$$s_{jmt}(x, p, \xi; \theta) = \sum_g \zeta_{gmt}\ s^g_{jmt}(x, p, \xi; \theta_g, \lambda_f) \tag{3.7}$$

where $\sum_g \zeta_{gmt} = 1$ for all m and t.

In a given market mt, the change in consumer surplus for a sub-group g for a counterfactual scenario, with associated values denoted by x^1, relative to the observed scenario, denoted x^0, can be extended from McFadden (1973) as,

$$\Delta CS_g(p, p') =$$

$$M_{mt}\zeta_{gmt} \int \frac{1}{|\alpha_{ig}|} \left[\log\left(\sum_{j'=1}^{J^1_{mt}} \exp(U^1_{igj'mt}) \right) - \log\left(\sum_{j=1}^{J^0_{mt}} \exp(U^0_{igjmt}) \right) \right] dF_g(\alpha_{ig}, \beta_{ig}) \tag{3.8}$$

where U_{igjmt} is given by Equation 3.1.

3.4.2 Supply: Firm's Problem

The supply model can be solved by backwards induction, taking market-level demand as given from the household's problem above in Equation (3.7).

Step 3: Pricing Decision

At each time period t, in each geographic market m, for each product in the set of products $J_{fmt} \subseteq \mathscr{J}_{fmt}$ that each firm f chooses to produce in Step 1, firms simultaneously set prices that maximize profits in Nash-Bertrand equilibrium. Define variable profits for firm f in period t and market m as,

$$\tilde{\pi}_{fmt} = \sum_{j \in J_{fmt}} \sum_{g} \left\{ M_{mt} \zeta_{gmt} s^{g}_{jmt}(x, p, \xi; \theta_g, \lambda_f)(p_{jmt} - mc_{jmt} - \mathbb{I}_{\{j,g \in WIC\}} reb_j) \right\} \tag{3.9}$$

where M_{mt} is the total number of infant households and ζ_{gmt} is defined as in the demand model above. Per reconsitituted ounce variable profits are given by the price charged, p_{jmt}, less the marginal cost of producing an additional ounce of the product, mc_{jmt}, and, if applicable, the contractually-set rebate paid to the state WIC agency, reb_j.

Total profits in a given market-time period, which incorporates product fixed costs are then

$$\pi_{fmt} = \tilde{\pi}_{fmt} - \sum_{j \in J_{fmt}} F_{jmt} \tag{3.10}$$

where F_{jmt} is the per-market, per-period fixed cost of supplying product j.

To back out product-specific marginal costs, we can take the first order condition of the profit function in equation (8) (or similarly to equation (9)) with respect to price and collapse into matrix notation, which yields,

$$mc_{jmt} = p^{*}_{jmt} + \underbrace{\left\{ \Big(\sum_{g} \zeta_g (\nabla_p s^g) \Big)^{-1} \big(s - \zeta_{WIC} (\nabla_p s^{WIC}) reb \big) \right\}_j}_{\equiv \mu_{jmt}(x,p,\xi;\theta)} \tag{3.11}$$

where p^{*}_{jmt} represents the profit-maximizing price, $\nabla_p s^g$ represents the gradient matrix of subgroup g market shares with respect to prices times an ownership matrix for firm f that contains indicators for all products produced by f (as in Miravete, Seim, Thurk (2017)). By setting $\zeta_{WIC} = 0$, we can see that equation (10) nests the standard expression for marginal cost in terms of price and the optimal markup.

Then marginal cost can be parameterized as,

$$\ln\big(p_{jmt} + \mu_{jmt}(x, p, \xi, \theta)\big) = W_j \gamma + \omega_{jmt} \tag{3.12}$$

where W_j are observable cost shifters, such as the price of cow's milk and observed physical product attributes, γ is a parameter (vector) to be estimated, and ω_j is an unobservable (to the econometrician) marginal cost shock.

Step 2: Demand and Cost Shock Realizations

Prior to setting prices, firms observe marginal cost shocks ω and demand shifters ξ. As in Berry, Levinsohn, and Pakes (1995), the timing of this step aids identification. As detailed further below, we require exogeneity of demand shocks ξ and product characteristics x, i.e., $\mathbb{E}[\xi|x] = 0$.

Step 1: Product Choice

Prior to observing demand and marginal cost shocks, firms must choose which products to supply in each market. We assume firms offer the set of products that bring the highest profits in expectation.

At each time period t, in every market m, firm produces J^*_{fmt} such that

$$J^*_{fmt} \in \underset{J_{fmt} \subseteq \mathscr{J}_{ft}}{\operatorname{argmax}} \ \mathbb{E}[\pi_{fmt}(J_{fmt})] \tag{3.13}$$

where the expectation of profits, defined in equation (9) above, is with respect to demand and cost shocks. The portfolio of available products, $\mathscr{J}_{ft}$, is assumed to be any product produced by firm f prior to and including time period t.

3.5 Estimation

In order to estimate demand and supply parameters of the model detailed above, we can follow the discrete choice demand estimation literature (Berry, Levinsohn, and Pakes (1995), Nevo (2001)). To eventually simulate counterfactual policies, as detailed below, of interest are the demand parameters governing the household's problem, namely, the mean and variance of the random coefficients, β and α, and the fixed effects, λ. Additionally, we are interested in recovering product-specific marginal costs, mc, and cost shifter parameters γ, and product fixed costs, F.

3.5.1 Demand

To estimate the true underlying demand parameters, θ_0, we want to determine the value of θ that is consistent with the observed infant formula sales data. We can do so by taking generalized method of moments (GMM) approach that takes the model to observed data to infer demand parameter values.

To develop moment conditions, we rely on the identification assumption that demand shock ξ is exogenous to a chosen instrument Z. Following Berry, Levinsohn, and Pakes (1995), the excluded element of Z is constructed, for a given product-market-time produced by firm f, by summing the characteristics of products produced by competing firms, x_{-f}. The included elements of the instrument are the characteristics for of firm f's product, x_j, and a constant. As indicated in Step 2 of the supply model, the timing assumption that demand shocks ξ are observed only after product choices are made (Step 1) implies the excluded and included elements of this instrument is independent of ξ. Further, the instrument is relevant, as firm f sets prices with respect to product space competition, that is, it takes x_f and x_{-f} into account.

We can then use this exogeneity of the instrument and demand shocks to form the required moment condition,

$$\mathbb{E}[\xi_{jmt}(x, p, s; \theta)|Z_{jmt}] = 0 \Rightarrow \mathbb{E}[Z'_{jmt}\ \xi_{jmt}(x, p, s; \theta)] = 0 \tag{3.14}$$

Then, to determine the estimate $\hat{\theta}$ of the true parameter θ_0, we search over the parameter space and evaluate the GMM objective function. Specifically, for a given guess of θ,

1. Determine ξ_{jmt} that is consistent with observed market shares s_{jmt}, prices p^*_{jmt}, and product characteristics x_j, and the guess of demand parameters, θ, by employing a

contraction mapping, in the spirit of Berry (1994),

$$\xi_{jmt}^{M} = \xi_{jmt}^{M-1} + log(s_{jmt}) - log(s_{jmt}(x, p, \xi; \theta)) \tag{3.15}$$

where s_{jmt} is observed market share of product j in market m at time t and $s_{jmt}(x, p, \xi; \theta)$ is its model-implied analogue. Convergence is achieved when $||\xi_{jmt}^{M} - \xi_{jmt}^{M-1}|| \leq \varepsilon_{tol}$, $\forall j, m, t$ for a desired level of tolerance, ε_{tol}, which occurs when the model market shares (closely) match observed market shares. Let $\xi(x, p, s; \theta)$ represent the vector of product-market-time demand shocks that result from the contraction mapping in Equation 3.15.

2. Use the value of $\xi(x, p, s; \theta)$ from Step 1 to generate a vector of moment conditions as a sample analogue to the second expression in equation (13),

$$m(x, p, s; \theta) = \frac{1}{N} \sum_{jmt=1}^{N} Z'_{jmt} \xi_{jmt}(x, p, s; \theta) \tag{3.16}$$

where N represents the total number of product-market-time periods. The number of (included and excluded) instruments equals the dimension of moment condition vector, $m(\cdot)$. To identify θ, we require the number of moments, and thus the number of instruments, to be at least as large as the dimension of θ. As evident from (13), we can generate additional excluded instruments by taking arbitrary functions of the excluded elements of Z. In my estimation routine, excluded instruments are calculated as other firms' product characteristics, other firms' product characteristics squared, and interaction terms of other firms' product characteristics such that rank(Z) > dim(θ).

We can search over the parameter space to find the value of θ that is most consistent with the assumed moment condition. That is, the estimate of θ is that which minimizes the GMM objective function,

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \ (m(x, p, s; \theta))' \ W \ m(x, p, s; \theta) \tag{3.17}$$

where W represents a weight matrix. In estimation, I employ a two-step routine to first determine $\hat{\theta}_1$ that satisfies expression (16) with W as the identity matrix. From this estimate $\hat{\theta}_1$, I calculate the sample analogue of the optimal weight matrix,

$$W(\hat{\theta}_1) = \left(\frac{1}{N} \sum_{jmt=1}^{N} Z'_{jmt} \xi_{jmt}(x, p, s; \hat{\theta}_1) \xi_{jmt}(x, p, s; \hat{\theta}_1)' Z_{jmt} \right)^{-1} \tag{3.18}$$

Then, I re-run the GMM procedure with $W = W(\hat{\theta}_1)$ to determine $\hat{\theta}_2$, a consistent estimate of θ_0.

3.5.2 Supply

Marginal Costs. Estimation of the marginal cost parameter γ can come from a similar exogeneity assumption as demand. For a given guess of γ, the marginal cost shock ω, as a function of data and demand parameter θ, can be backed out from Equation 3.11.

Fixed Costs. Once demand and marginal cost parameters are estimated, fixed costs, F_j, can be set identified exploiting necessary expected profit-maximizing conditions on product offerings from Step 1 of the supply model.[9]

The necessary profit-maximizing condition states that, for firm f, fixing competing firms' strategies, any alternative set of products $J'_{fmt} \subseteq \mathscr{J}_{ft}$ than the observed choice J_{fmt}, would, in expectation, yield lower profits to the firm (with re-optimized prices for all $j \in J'_{fmt}$). This follows from the optimal product set choice detailed in expression [X] above.

We see that, in a given market and time, if firm f was forced to produce and sell a product j' that was not produced in actuality, generating variable profits but incurring a fixed cost from j', the firm's expected profits, with respect to demand and marginal cost shocks, must be less than the observed profits.

$$\mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt}; \hat{\theta}, \hat{\gamma})] \geq \mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt} \cup j'; \hat{\theta}, \hat{\gamma})] - F_{j'mt}, \forall j' \in \mathscr{J}_{ft} \setminus J_{fmt} \tag{3.19}$$

[9]For example, see Eizenberg (2014) and Fan and Yang (2020)

where $\tilde{\pi}_{fmt}(J_{fmt}; \hat{\theta}, \hat{\gamma})$ represents variable profits, which does not include fixed costs (as in Equation 3.9 above). Rearranging expression 3.19 can form a lower bound on the per-period fixed cost of producing j' in market m,

$$F_{j'mt} \geq \mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt} \cup j'; \hat{\theta}) - \tilde{\pi}_{fmt}(J_{fmt}; \hat{\theta})] \text{ for } j' \in \mathscr{J}_{ft} \setminus J_{fmt} \tag{3.20}$$

Similarly, we can construct upper bounds on a product j' observed to be supplied in a given market at a given time by simulating counterfactual variable profits without j'.

$$\mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt}; \hat{\theta})] - F_{j'mt} \geq \mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt} \setminus j'; \hat{\theta})] \text{ for } j' \in J_{fmt} \tag{3.21}$$

Rearranging gives an upper bound on per-period fixed costs for j',

$$F_{j'mt} \leq \mathbb{E}[\tilde{\pi}_{fmt}(J_{fmt}; \hat{\theta}) - \tilde{\pi}_{fmt}(J_{fmt} \setminus j'; \hat{\theta})] \text{ for } j' \in J_{fmt} \tag{3.22}$$

For a given product j', one side of the set identified fixed cost interval can be estimated using simulated profits for alternative, unobserved product offering sets to generate empirical counterparts to the bounds given by the RHS of inequalities 3.20 and 3.22. For a product in the observed product set, i.e., $j' \in J_{fmt}$, an upper bound on per-period fixed costs are given by the magnitude of change in variable profits for firm f if j' were to be dropped, and all else equal. Similarly, for a product not in firm f's product set, a lower bound on fixed cost is given by the magnitude of change in firm f's variables profits were that product to be supplied by f, all else equal.

In practice, to "fill in" the remaining fixed cost bounds, to be conservative, I use the extremums of fixed costs derived via inequality approach described above. For example, for a product j' in the observed product set, its lower bound is achieved by taking the minimum estimated lower bound for all other products in the market in the given time period.[10]

[10]For the results included in this chapter, to ensure product fixed cost lower bounds are less than upper bounds, I conservatively use one-half the minimum of calculated lower bounds and twice the maximum of

I estimate costs only for the three primary manufacturers in the sample (Abbott, Mead Johnson, and Nestle-Gerber), while assumming all other manufacturers are passive in the product choice stage, and thus do not require fixed cost bounds for counterfactuals that endogenize firm product choice.

3.6 Results

3.6.1 Demand Parameter Estimates

The demand parameter estimates indicate the "best" guess, given observed sales data, of underlying parameters that govern consumer tastes. Table 3.2 below can be read as, for each product attribute, for each subpopulation, what are the estimates of the mean and variance parameters that govern the distribution of tastes, in terms of added (or diminished) utility for a marginal ounce of infant formula that exhibits a given product attribute.

As the utility of the outside good, breast feeding, is normalized to zero, the interpretation of the constant term allows for comparison between inside goods – available infant formula – and breast feeding. For the subpopulation eligible for WIC vouchers, mean taste for inside infant formula goods is less than zero, with some heterogeneity, suggesting this subpopulation largelprefers breast feeding relative to infant formula. The distribution of taste for inside goods within the non-WIC subpopulation is also mean negative, though to a lesser magnitude ($\beta_{0,\text{non-WIC}} = -0.48$), and the estimate for $\sigma_{\beta_{0,\text{non-WIC}}}$ also suggests heterogeneity in this taste within this subpopulation.

Given the assumption that WIC consumers only consider inside goods eligible for WIC vouchers in their choice set, and thus do not face product prices, we cannot identify a price coefficient (mean nor random coefficient) for this subpopulation. We are, however, able to identify a price coefficient for the non-WIC subpopulation, given this group of consumers do face variation in prices. We see that the estimated mean for the non-WIC subpopulation price coefficient is negative ($\alpha_{\text{non-WIC}} = -28.21$), implying disutility from higher prices; however,

calculated upper bounds. In future work, I hope to employ the finite sample adjustment for these bounds described in Eizenberg (2014).

the random coefficient estimate implies no heterogeneity within this subpopulation.

The parameter estimates for the product attribute "powder" can be interpreted as the relative taste for powder infant formula products relative to liquid concentrate (i.e., the omitted group). The mean utility parameter estimates suggest that, relative to the non-WIC subpopulation, the WIC-eligible subpopulation more strongly prefers powder form to liquid concentrate. Further, the mean utility parameter point estimate (and standard error) for powder form for the non-WIC subpopulation suggest the mean taste is not statistically significantly different from zero at the 95% confidence level. Though, the random coefficient estimate ($\sigma_{\beta_{\text{powder,non-WIC}}} = 0.02$) does suggest slight heterogeneity in this preference for powder formula within the non-WIC subpopulation

Overall, the model's demand parameter estimates suggested by the data seems to exhibit reasonable taste patterns with some heterogeneity in preferences within subpopulations.

3.7 Counterfactuals

After we have recovered demand preference parameter estimates and marginal and fixed cost supply parameter estimates by taking the structural model to data, we can run counterfactual simulations to evaluate the welfare implications of infant formula subsidies. Specifically, using the structural model, I assess consumer and producer responses to a policy in a single geographical market in which a larger proportion of the population is eligible for WIC infant formula vouchers and evaluate the resulting effects on welfare measures.

In the first set of counterfactuals, I allow firms to re-optimize only prices and compare counterfactual welfare measures to those of the factual environment. For a given increase in the proportion of WIC-eligible households, I determine new prices, which follow firms' first order pricing condition, as given in Equation 3.11, as an optimal markup over marginal costs. I then calculate changes in welfare measures, including consumer surplus and producer profits.

I then allow for the three primary firms, Mead Johnson, the supplier of the WIC eligible brand in the sample, Abbott, and Nestle-Gerber to optimally adjust their product portfolio and associated prices in response to a change in WIC eligibility. I follow the iterative best-response algorithm developed in Fan and Yang (2020) to determine equilibrium product choices for each draw of a vector of product fixed costs. For simplicity in this chapter, I set demand and cost shocks to zero.[11]

Table 3.3 shows the surplus measures associated with increasing the eligible WIC population by either 20% or 40%, relative to the observed factual scenario, under both price-only supply-side adjustments (top panel) and price and product adjustments (bottom panel) in the Detroit DMA during the last month of the sample (i.e., December 2015).[12] From these results, we see that, only considering price adjustments, an increase in the proportion of the population that is eligible for WIC voucher would suggest consumer surplus losses and producer surplus losses, each increasing in magnitude as the WIC proportion increases.[13] However, we see that, if both product and price changes were accounted for, the associated counterfactuals with increasing WIC populations suggests the firms' optimal response is to reduce the number of product offerings (from 28 products to, on average, 27.3 and 26.2, respectively), which drives an additional reduction in consumer surplus. In summary, a price-only response generates total surplus losses that increase in the counterfactual proportion of WIC-eligible; however, these losses are less in magnitude than if, perhaps more realistically, firms adjust in both product and price margins, resulting in larger losses in

[11] As is done in Fan and Yang (2020), demand and cost shocks can be taken over the recovered empirical distribution of these objects or a fitted distribution (e.g., normal) derived from the respective empirical distributions. I look to incoporate these steps in future versions of this work.

[12] The baseline proportion of WIC-eligible households is 17.04%. Increasing the fraction of WIC-eligible household by 20% results in an increase by 3.41 percentage points to 20.45% (middle column counterfactuals). Increasing the baseline proportion 40% results in an increase by 6.82 percentage points to 23.86% (last column counterfactuals).

[13] In the calculation of consumer surplus, since WIC consumers face a price of zero for their assumed inside good choice set (i.e., WIC products), and, in the product and price adjustment counterfactuals, WIC products are not able to be removed, the change in consumer surplus for consumers that are WIC-eligible in both the factual and counterfactual is zero. Further, I do not include changes in surplus for the population that is newly eligible for WIC vouchers since I am not able to recover consumer surplus for this subpopulation since a price parametner (i.e., $\tilde{\alpha}_{\text{WIC}}$) is not identified for this subpopulation. Even if we assumed a price coefficient for this subpopulation (e.g., $\tilde{\alpha}_{\text{WIC}} = \tilde{\alpha}_{\text{non-WIC}}$), the model could suggest negative surplus given the reduction in choices between the factual and counterfactuals for these consumers stemming from the assumption WIC consumers only choose between WIC products. As a result, the reported consumer surplus measures in this chapter can be thought of as welfare for the subpopulation of consumers that are not WIC eligible in both the factual and relevant counterfactual scenario.

total surplus, though still increasing in magnitude with the counterfactual proportion of WIC-eligible consumers.

3.8 Conclusion

In this chapter, I have developed a model of demand and supply in the infant formula industry to understand how demand-side subsidies effect short-run welfare when accounting for supply-side pricing and product responses. The demand model allows for heterogeneity of consumer tastes by allowing for subpopulation-specific random coefficient parameters that govern the distributions on taste for observable product attributes. The supply model allows for price and product responses by firms to counterfactual policies that increase the eligible population for WIC infant formula vouchers.

In these counterfactual exercises, I find that consumer, producer, and total surplus diminish when increasing the WIC-eligible population in Detroit during December 2015. Further, when incorporating product adjustments, along with price, the magnitude in each welfare measure increases. This result provides support for the need to incorporate both pricing and product responses when accurately evaluating policies in differentiated product industries.

As discussed above, the reported consumer surplus measures in this chapter do not incoporate surplus from newly-WIC eligible consumers. Therefore, from the social planner's perspective, the magnitude of the decreases in total surplus when increasing the WIC population in counterfactual simulations in this paper can be thought of as a lower bound of required gains in surplus from newly-WIC eligible consumers. For example, when increasing the number of WIC-eligible consumers by 20% and allowing for firm pricing and product responses, total surplus decreases by $9,750 (Table 3.3, bottom panel, middle column). Thus, the social planner would enact this policy only if the resulting increase in consumer surplus for those newly WIC-eligible was greater than $9,750.

The framework developed in this chapter can further be used to evaluate additional proposed policies in the infant formula industry, as well as more broadly be used to understand demand-side subsidies in differentiated product markets. For example, in the factual setting, only one brand of formula is eligible for purchase by WIC vouchers. A natural extension is to consider the welfare effects of relaxing this policy to allow for multiple eligible brands, while incorporating firms' responses in both the price and product dimensions. I leave this, and related questions, to future research.

Table 3.1: Summary Statistics

Variable	Mean	Standard Deviation	Min	Max
Price ($/reconstituted oz.)	0.18	0.04	0.00	0.60
Powder indicator	0.95	0.22	0.00	1.00
Liquid concentrate indicator	0.05	0.22	0.00	1.00
WIC eligible indicator	0.07	0.26	0.00	1.00
Total observations		5,888		
Price (for WIC eligible)	0.18	0.03	0.10	0.29
Rebate (for WIC eligible)	0.14	0.01	0.13	0.15
WIC eligible observations		432		

Note: Observations at the product-market-month level.

Table 3.2: Demand Parameter Estimates

	Subpopulation (g)	Mean Utility (β_g, α_g)	Random Coeff. ($\sigma_{\beta_g}, \sigma_{\alpha_g}$)
Constant	WIC	-7.6431	0.4470
		(0.0644)	(0.0076)
	non-WIC	-0.4766	0.3048
		(0.7676)	(0.0454)
Price	WIC	-	-
		-	-
	non-WIC	-28.2088	0.0000
		(1.1380)	(0.0000)
Powder	WIC	5.2943	0.0000
		(0.0601)	(0.0000)
	non-WIC	1.3202	0.0169
		(0.7355)	(0.0002)

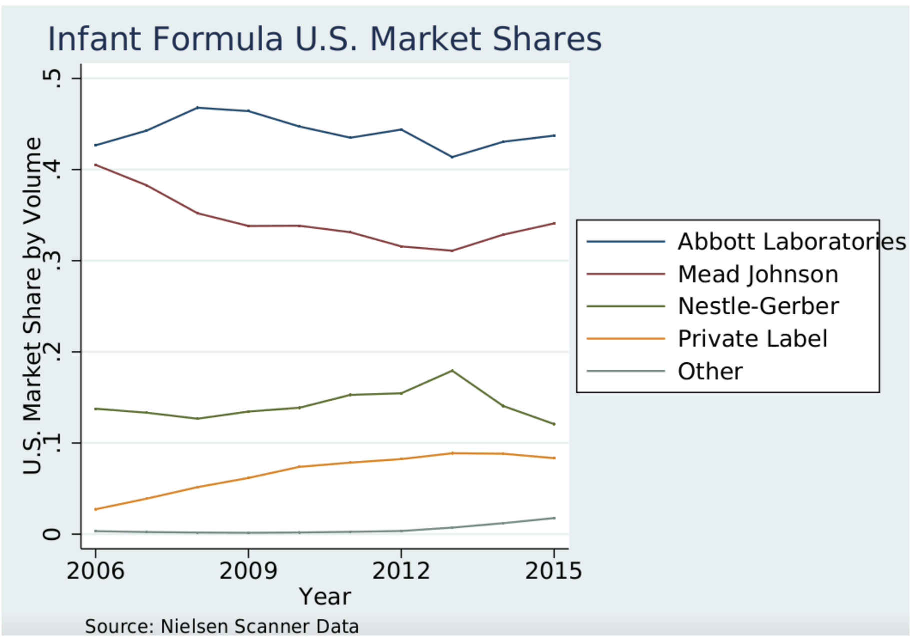
Note: Standard errors reported in parenthesis. Estimates based on GMM estimation on 5,888 product-market-month observations in Michigan from January 2012 – December 2015. Concentrate characteristic ommitted given all observations in the sample are either concentrate or powder form. Firm, year, and DMA fixed effects included. The non-WIC population's mean utility on price is constrained in estimation procedure to be less than zero. Excluded instruments are functions of other firms' products' characteristics in the same market and same firm's other products' charateristics.

Table 3.3: Counterfactual Results

	Obs. WIC Pop. ($\zeta' = \zeta_0$)	Increase 1 ($\zeta' = 1.2\zeta_0$)	Increase 2 ($\zeta' = 1.4\zeta_0$)
Price-Only Adjustment			
Number Products	28.0	28.0	28.0
Δ Consumer Surplus ($)	-	-5,573	-11,073
Δ Profits	-	-3,048	-6,070
Δ Rebate Payments	-	581	1,161
Δ Voucher Payments	-	-380	-781
Δ Total Surplus	-	-8,420	-16,764
Product and Price Adjustment			
Number Products	28.0	27.3	26.2
Δ Consumer Surplus ($)	-	-6,941	-24,192
Δ Variable Profits	-	-4,848	-10,952
Δ Fixed Costs	-	1,835	5,070
Δ Rebate Payments	-	581	1,161
Δ Voucher Payments	-	-377	-768
Δ Total Surplus	-	-9,750	-29,681

Note: Counterfactuals evaluated for the Detroit DMA during the last month of the sample (December 2015) with changes relative to the observed factual scenario (i.e., under ζ_0 in December 2015) in US dollars ($). For product and price adjustment counterfactuals (bottom panel), for each evaluated increase in WIC eligibility, I draw a product fixed cost vector from the recovered empirical bounds and evaluate each respective counterfactual. I repeat this process for 25 fixed cost draws and average over welfare measures. Changes in consumer surplus represent changes in welfare for consumers not eligible for WIC vouchers in both factual and respective counterfactual scenarios. Changes in fixed costs (bottom panel) expressed in terms of producer surplus gains due to fixed costs saved. Voucher payments represent total revenues for WIC eligible products and rebate payments represent payments from the WIC provider to the state. Both of these transfers are accounted for in variable profit measures, which allows for (changes in) total surplus to be taken as the sum of each respective sub-column.

Figure 3.1: Infant Formula Manufacturer Market Shares over Time



Appendices

Appendix A

Appendix to Chapter 1

A.1 Appendix

A.1.1 Additional Comparative Statics

This section provides additional detail to the comparative static exercises using the theoretical model developed in Section 1.2. As demonstrated in Figure A.1, when adding a vertically integrated store brand product, retailers and consumers both unambiguously benefit, while the sign in the change in manufacturer profits varies on model parameters (Figure A.2). Particularly, vectors of model parameters that generate equilibria where retailers gain more correspond to larger decreases in manufacturer profits.

When a vertically integrated store brand replaces a national brand product supplied at arm's-length, the sign of the change in retailer profits corresponds to the sign of the change in consumer surplus. For example, as demonstrated in Figure A.3, when the replacement by a store brand would yield a decrease in retailer profits, consumers would likewise be worse off. Alternatively, a profitable replacement yields an increase in surplus. On the other hand, upon the replacement of a national brand product by a retailer's store brands, manufacturer profits nearly always decrease (Figure A.4).

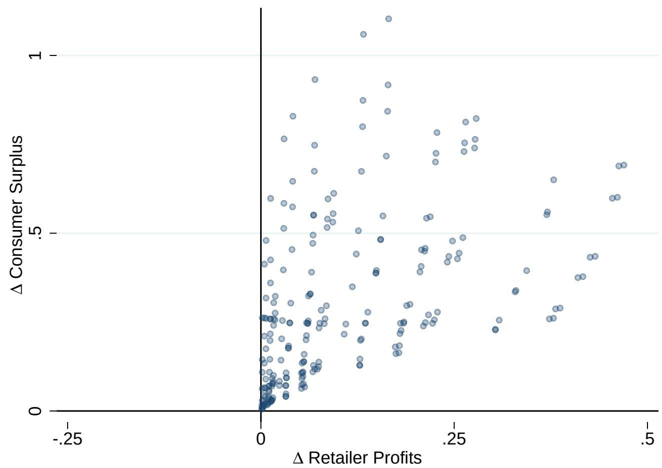
A.1.2 Additional Event Studies

I conduct additional event studies that assess retailer responses in the product and price margins around store brand product line introductions.

Wholesale Prices

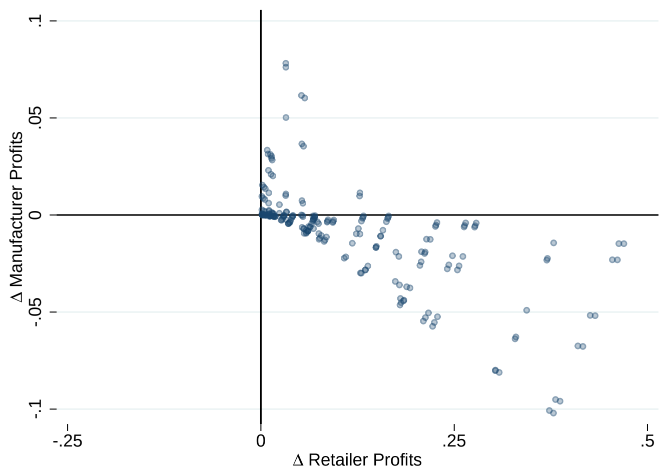
I also conduct an additional event study on the wholesale price of national brand product lines around the time of store brand introductions. To do so, I replace the lefthand side of the regression specification given in Equation (1.7) with wholesale prices from Promo Data PriceTrak. Coefficients from this regression are plotted in Figure A.5. Given the coefficient estimates, there does not seem to be a detectable change in wholesale prices at time of entry. This negative result could be due to two possibilties. First, theory would predict an ambiguous effect of store brand introduction on national brand wholesale prices, depending on the substitutability of introduced store brand with existing national brands. Second, and perhaps more likely, the data I have on wholesale prices is relatively coarse – from one distributor in only a select number of geographic markets. So the event study is assessing the impact of a retailer's store brand entry on a representative wholesale transaction price, not necessarily by the "introducing" retailer.

Figure A.1: A1 to B: Consumer Surplus



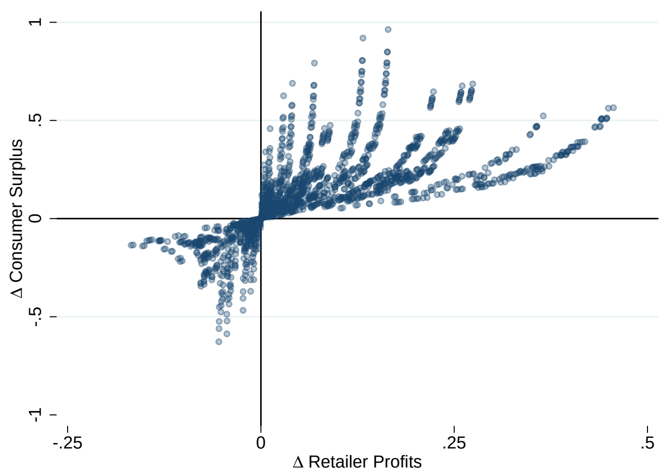
Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A1 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Change in consumer surplus is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1,2,3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1,2,3$. Equilibria associated with duplicative vectors with variation only in parameters pertaining to product $j = 2$, which do not impact the comparative static exercise between Scenarios A1 and B, are removed for visual clarity.

Figure A.2: A1 to B: Manufacturer Profits



Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A1 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Change in manufacturer profits is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1, 2, 3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1, 2, 3.$

Figure A.3: A2 to B: Consumer Surplus

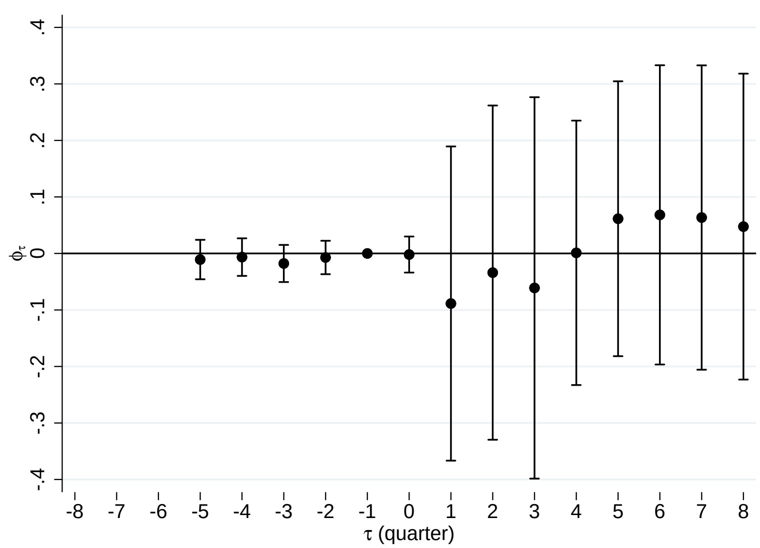


Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A2 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Change in consumer surplus is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1, 2, 3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1, 2, 3$. Equilibria associated with duplicative vectors with variation only in parameters pertaining to product $j = 2$, which do not impact the comparative static exercise between Scenarios A1 and B, are removed for visual clarity.

Figure A.4: A2 to B: Manufacturer Profits

Note: Each dot represents changes in equilibrium welfare measures for a given vector of model parameters. The x-axis represents the change in retailer profits when moving from the baseline scenario A2 to the comparative scenario B. Positive x-values represent an increase in profits for the retailer in moving from the baseline to comparative scenario. Change in manufacturer profits is plotted on the y-axis. Positive y-values represent increasing total surplus when moving from the baseline to comparative scenario. Parameter values: $\alpha = -5, c_j = 0, \gamma = 0.1, \omega_j = 0$ and $(\beta, x_j, \xi_j) \in \{1, 2, 3\} \times \{1, 1.5, 2\} \times \{-2, 0, 2\},\ \forall j = 1, 2, 3$.

Figure A.5: Average Percentage Change in Wholesale Price of National Brand Product Lines around Store Brand Product Line Introductions

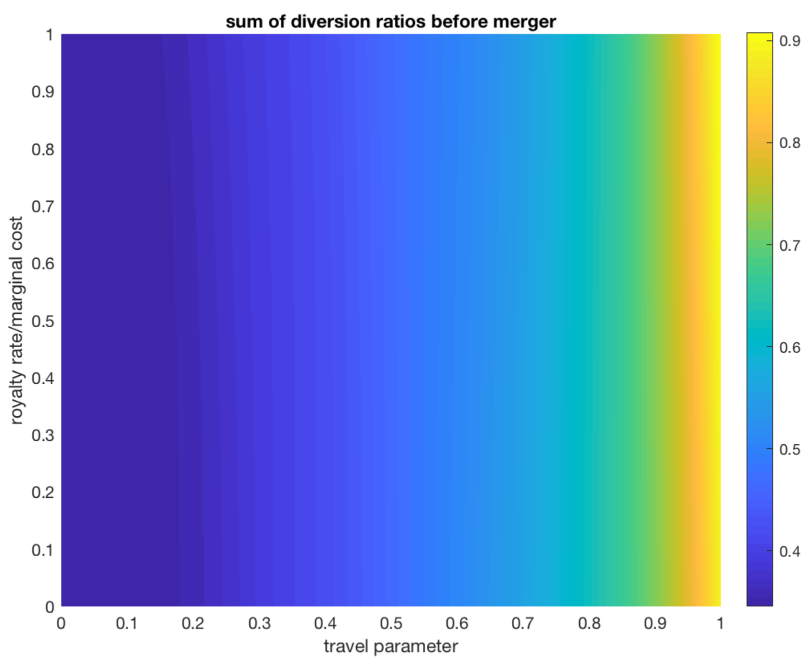


Appendix B

Appendix to Chapter 2

B.1 Additional Results

Figure B.1: Theoretical Simulation: The Sum of Diversion Ratios between the Licensor and the Two Licensees before Merger. This figure shows that the equilibrium sum of diversion ratios is determined by and increases with respect to the travel parameter ρ.



Here, we can observe the relationship between the bias in predicted changes in equilibrium objects and model primitives. In Figure B.9, we see that, for a given sum of diversion ratios between the licensor and each respective licensee, the overstated decrease in consumer surplus from the licensor-licensee merger is increasing in the royalty rate, r. In terms of producer surplus, for a given sum of diversion ratios, the overstated increase in producer surplus from the merger is increasing in the royalty rate, r.

Figure B.2: Theoretical Simulation: Prediction Bias on Consumer Welfare Effect of the Merger between the Licensor A and the Licensee B when royalty rate is 80% of the licensee's marginal costs. This figure illustrates the non-monotonicity of the prediction bias with respect to the sum of diversion ratios.

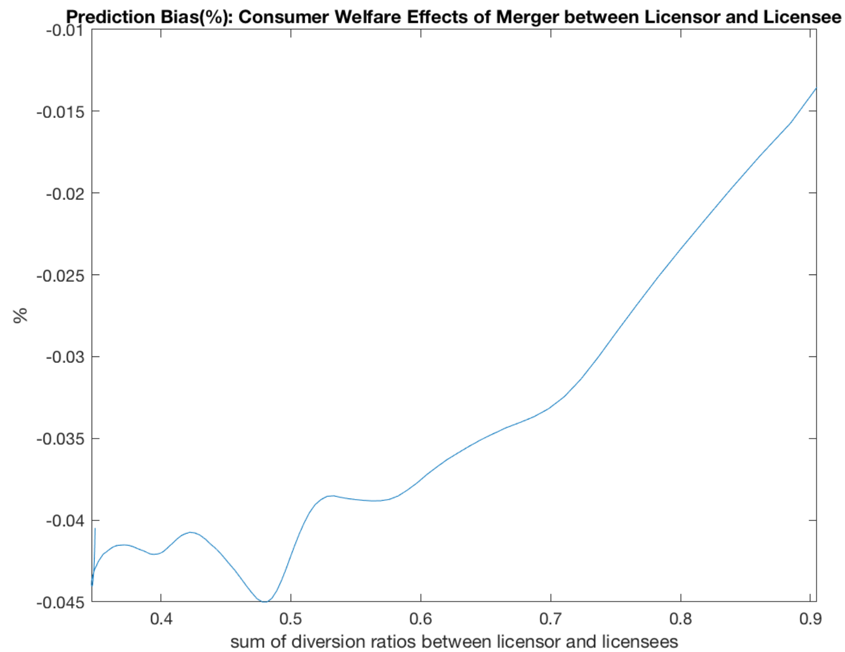


Figure B.3: Theoretical Simulation: Magnitude of the alignment effect divided by the estimation bias in the licensor's marginal costs.

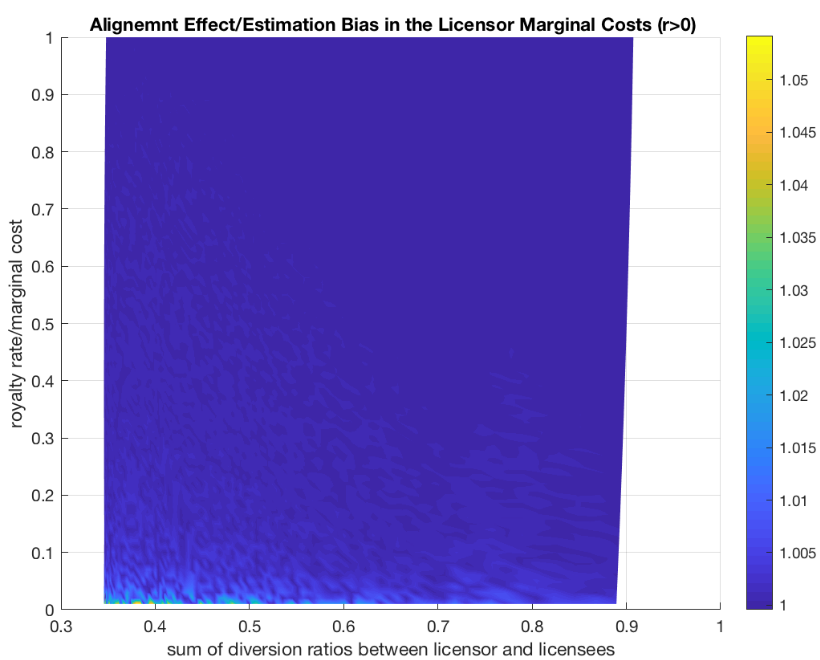


Figure B.4: Theoretical Simulation: Estimation Bias in the Marginal Costs of Licensee B.

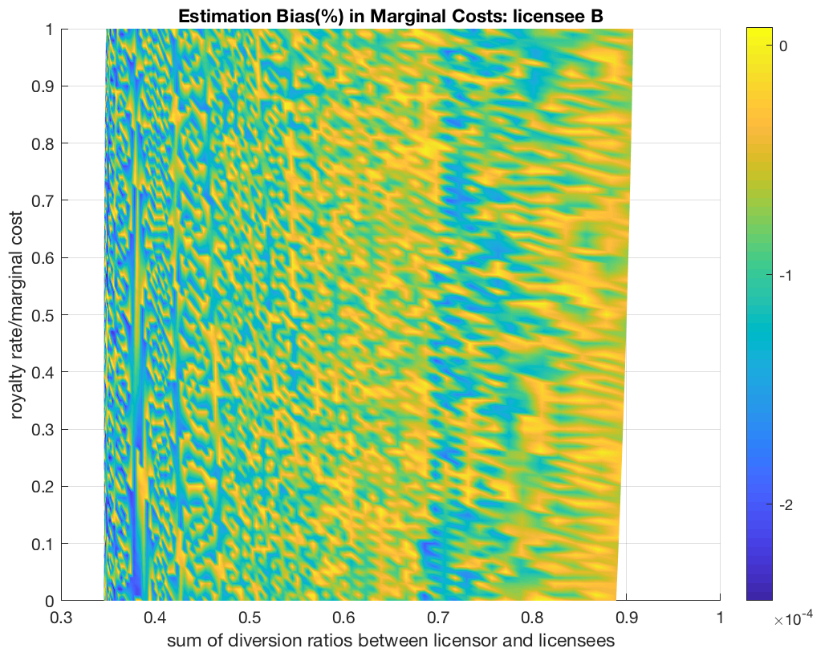


Figure B.5: Theoretical Simulation: Estimation Bias in the Marginal Costs of Licensee C.

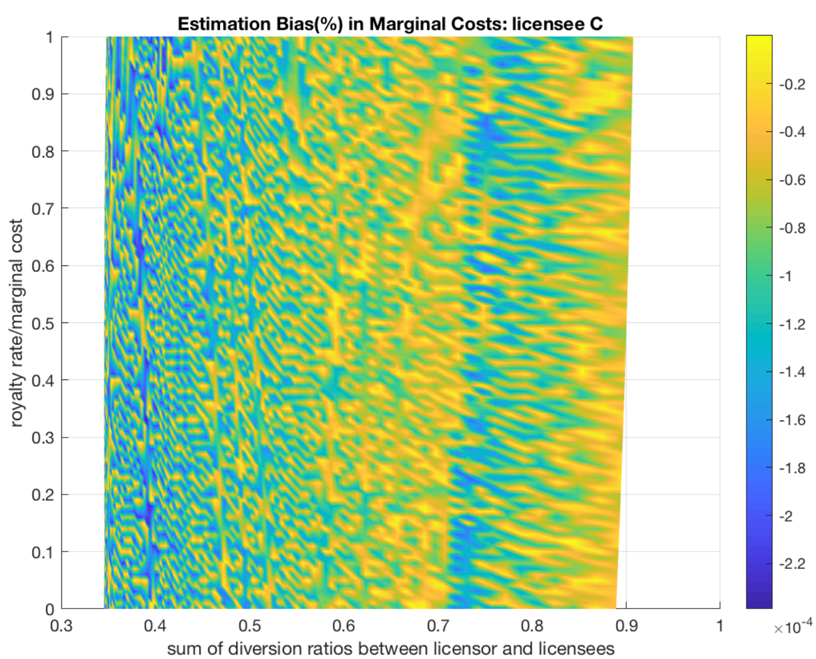


Figure B.6: Guidance Simulations: estimation bias in costs between mis-specified and true models for firm A (licensor) against royalty rate, r

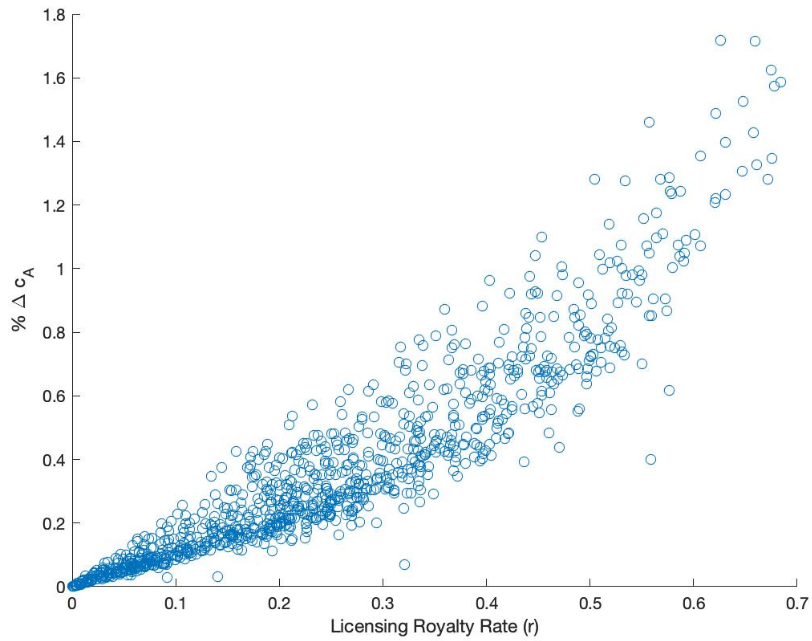


Figure B.7: Guidance Simulations: prediction bias in the share-weighted average price effects of the merger between licensor and licensee.

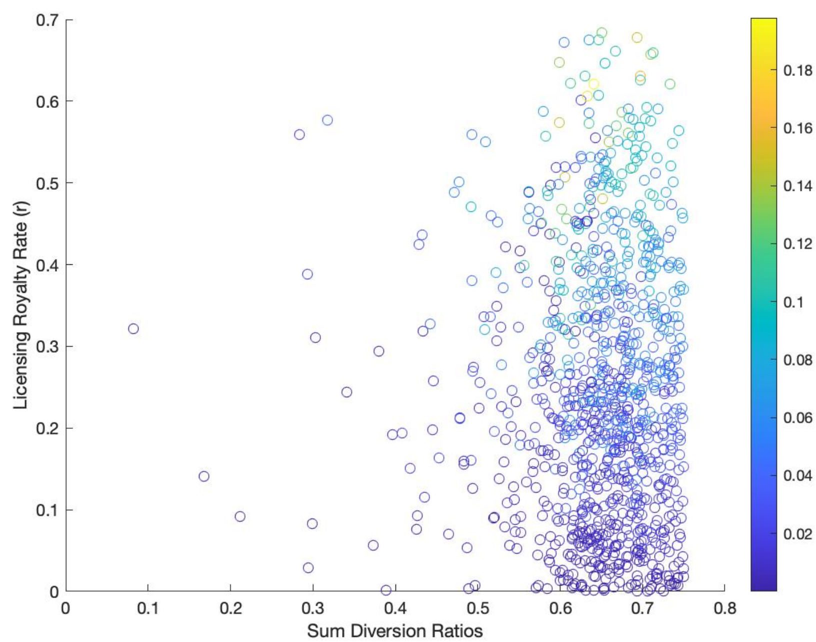


Figure B.8: Guidance Simulations: prediction bias in the total quantity effects of the merger between licensor and licensee.

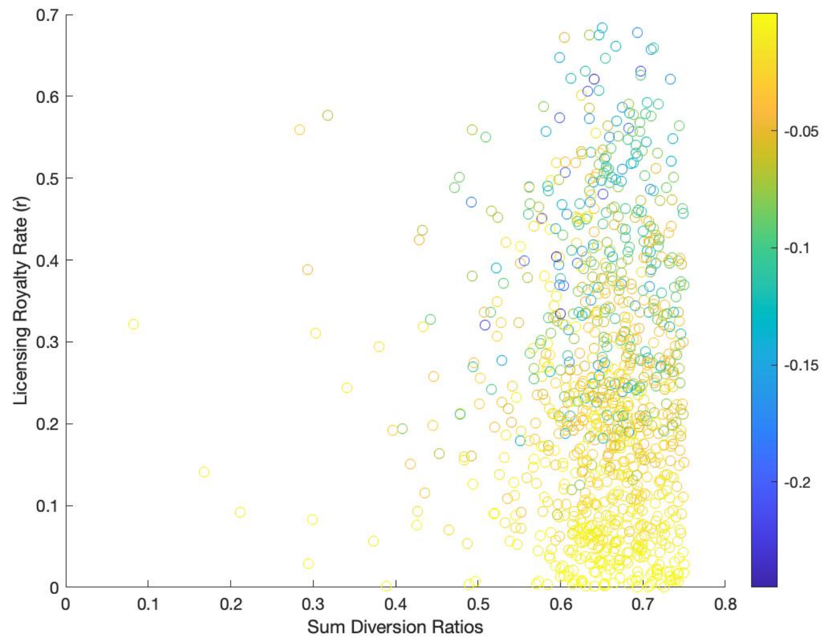


Figure B.9: Guidance Simulations: prediction bias in the consumer surplus effects of the merger between licensor and licensee.

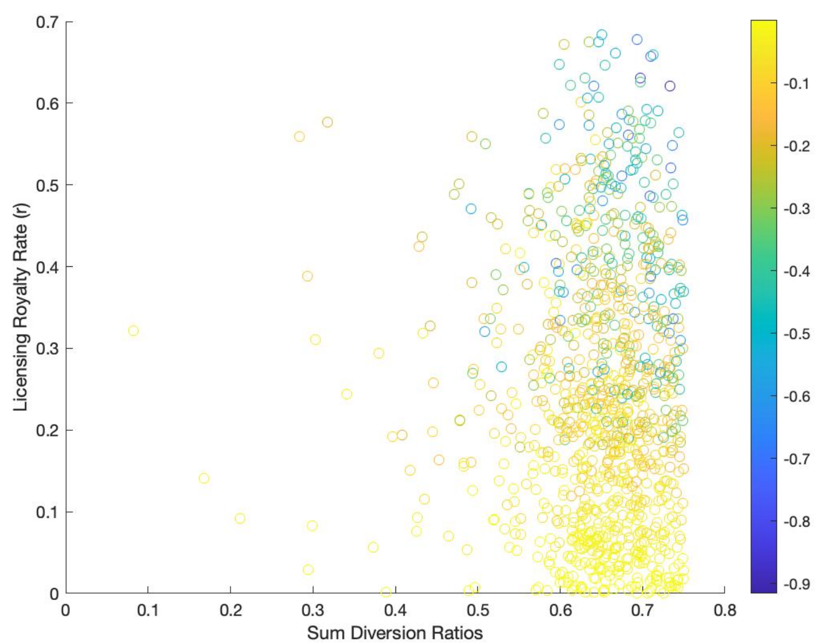


Figure B.10: Guidance Simulations: prediction bias in the producer surplus effects of the merger between licensor and licensee.

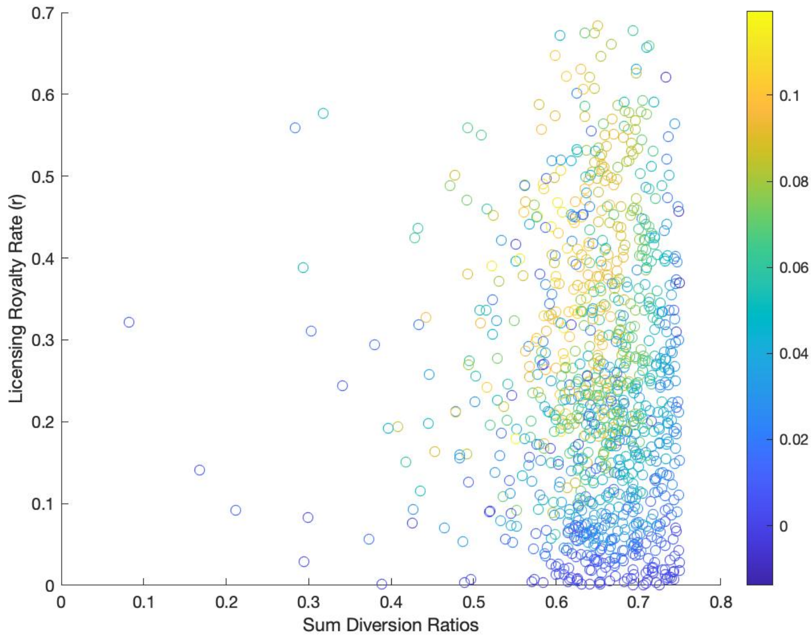


Figure B.11: Guidance Simulations: prediction bias in the share-weighted average price effects of the merger between licensees.

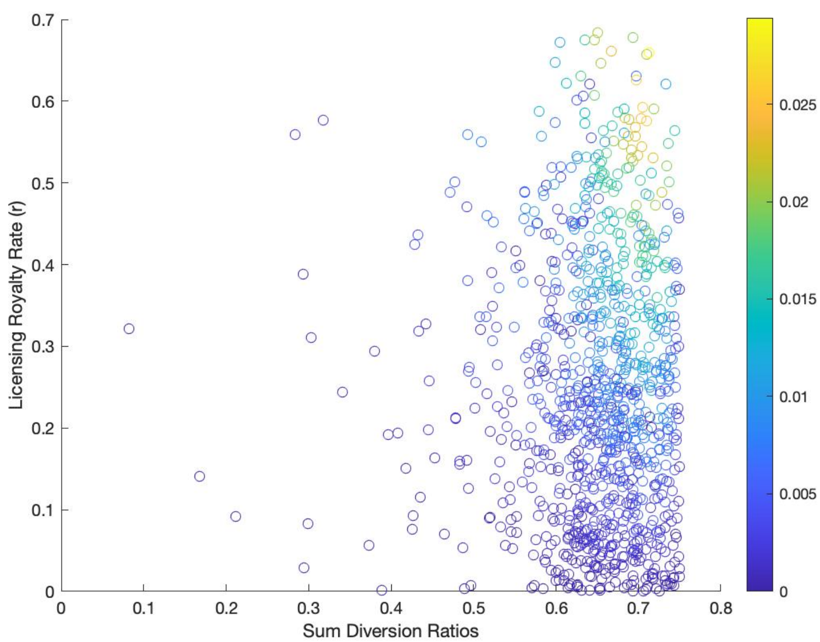


Figure B.12: Guidance Simulations: prediction bias in the total quantity effects of the merger between licensees.

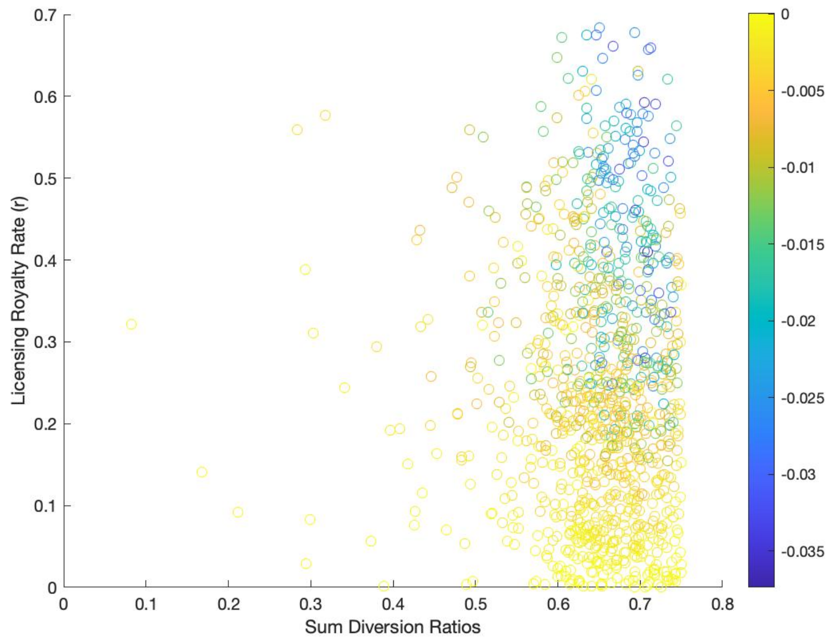


Figure B.13: Guidance Simulations: prediction bias in the consumer surplus effects of the merger between licensees.

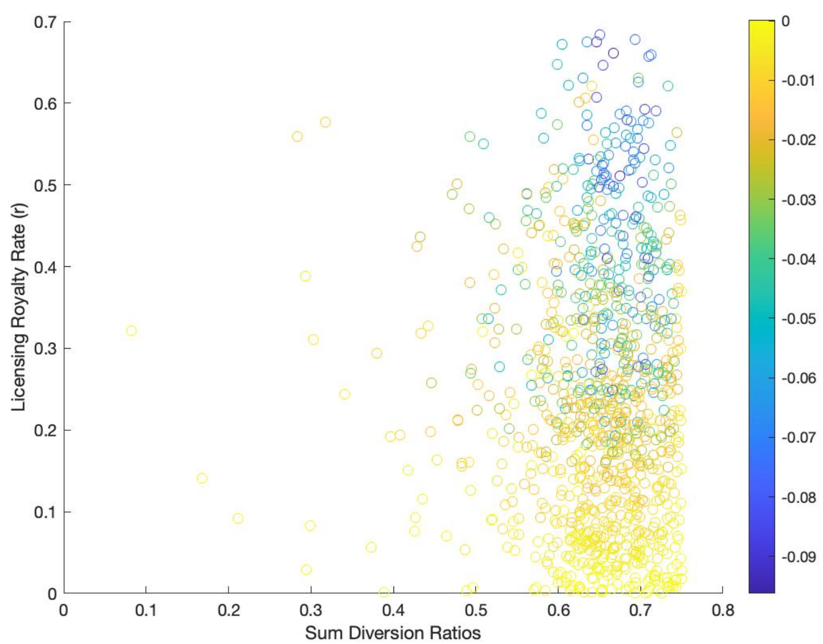
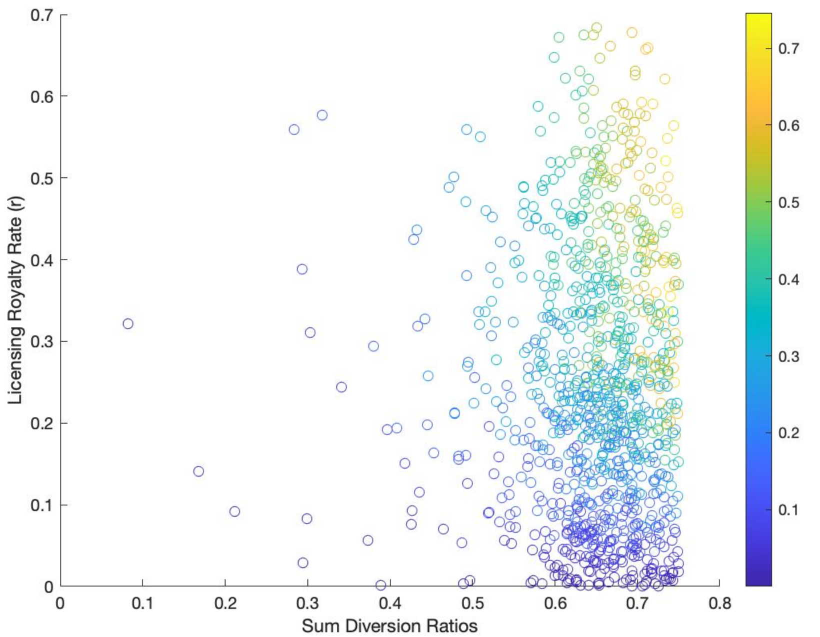


Figure B.14: Guidance Simulations: prediction bias in the producer surplus effects of the merger between licensees.



References

Aksoy-Pierson, M., Allon, G., and Federgruen, A. (2013) "Price competition under mixed multinomial logit demand functions" *Management Science*, 59(8):18171835.

Asker, J. (2016) "Diagnosing Foreclosure Due to Exclusive Dealing" *The Journal of Industrial Economics*, 64(3), 375-410.

Balan, D. and Brand, K. (2018) "Simulating hospital merger simulations" Available at SSRN 3153109.

Berry, S. (1994) "Estimating Discrete-Choice Models of Product Differentiation" *RAND Journal of Economics*, 25(2), 242-262.

Berry, S. and Jia, P. (2010) "Tracing the Woes: An Empirical Analysis of the Airline Industry" *American Economic Journal: Microeconomics* 2(3), 1-43.

Berry, S., Carnall, M., and Spiller, P. (2006) "Airline Hubs: Costs, Markups and the Implications of Customer Heterogeneity." In Advances in Airline Economics: Competition Policy and Antitrust, Vol. 1, ed. Darin Lee, 183-214. Amsterdam: Elsevier.

Berry, S., Levinsohn, J., and Pakes, A. (1995). "Automobile Prices in Market Equilibrium" *Econometrica* 63(4), 841890.

Bonfrer, A., and Chintagunta, P. (2004) "Store Brands: Who Buys Them and What Happens to Retail Prices When They Are Introduced?" *Review of Industrial Organization*, 24, 195-218.

Caplin, A. and Nalebuff, B. (1991) "Aggregation and imperfect competition: On the existence of equilibrium." *Econometrica: Journal of the Econometric Society*, pages 2559.

Crooke, P., Froeb, L., Tschantz, S., and Werden, G. (1999). "Effects of assumed demand form on simulated postmerger equilibria" *Review of Industrial Organization*, 15(3):205217.

Das Varma, G., and De Stefano, M. (2020) "Equilibrium analysis of vertical mergers" *The Antitrust Bulletin*, 65(3): 445458.

Decarolis (2015). "Medicare Part D: Are Insurers Gaming the Low Income Subsidy Design?" *American Economic Review* 105(4), 1547-1580.

Domnenko G., and Sibley D., (2020) "Simulating vertical mergers and the vertical GUPPI approach." Available at SSRN 3606641.

Draganska, M., Mazzeo, M., and Seim, K. (2009) "Beyond Plain Vanilla: Modelling Joint

Product Assortment and Pricing Decisions" *Quantitative Marketing and Economics* 7, 105-146.

Dubé, J., Fox, J., and Su, C. (2012) "Improving the Numerical Performance of Static and Dynamic Aggregate Discrete Choice Random Coefficients Demand Estimation" *Econometrica* 80(5), 2231-2267.

Eizenberg, A. (2014) "Upstream innovation and product variety in the U.S. home PC market" *Review of Economic Studies*, 81, 1003-1045.

Ellickson, P., Kong, P., and Lovett, M. (2018) "Private Labels and Retailer Profitability: Bilateral Bargaining in the Grocery Channel" Working Paper.

Fan, Y. (2013). "Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market", *American Economic Review*, 103(5), 1598-1628.

Fan, Y., and Yang, C. (2020)"Competition, product proliferation, and welfare: A study of the us smartphone market". *American Economic Journal: Microeconomics*, 12(2):99134.

Ghandi, A. and Houde, J. (2019) "Measuring Substitution Patterns in Differentiated Products Industries" NBER Working Paper No. 26375.

Gross, A. (2019) "Private Labels and Bargaining in the Supply Chain: The Case of Wine" Working Paper.

Guglielmo, A. "Competition and Costs in Medicare Advantage" *Working Paper*.

Hausman, J. (1994) "Valuation of New Goods Under Perfect and Imperfect Competition" NBER Working Paper No. 4970.

Hristakeva, S. (2019) "Vertical Contracts with Endogenous Product Selection: An Empirical Analysis of Vendor-Allowance Contracts" Working Paper.

Jaffe, S. and Shepard, M. (2018) "Price-Linked Subsidies and Imperfect Competition in Health Insurance", NBER Working Paper No. 23104

Layne-Farrar, A., and Lerner, J. (2011) "To join or not to join: Examining patent pool participation and rent sharing rules." *International Journal of Industrial Organization*, 29(2):294303.

Lerner, J. and Tirole, J. (2004) "Efficient patent pools" *American Economic Review*, 94(3): 691711.

Lerner, J., Strojwas, M., and Tirole, J. (2007) "The design of patent pools: The determinants of licensing rules." *The RAND Journal of Economics*, 38(3):610-625.

Li, J. (2017). "Compatibility and Investment in the U.S. Electric Vehicle Market" *Working Paper*.

Luco, F. and Marshall, G. (2020) "The Competitive Impact of Vertical Integration by Multiproduct Firms" *American Economics Review*, 110(7), 2041-64.

Mazzeo, M., Seim, K., and Varela, M. (2018)"The welfare consequences of mergers with endogenous product choice" *The Journal of Industrial Economics*, 66(4):980-1016, 2018.

McFadden, D. (1973) "Conditional Logit Analysis of Qualitative Choice Behavior" in Zarembka, P., ed., *Frontiers in Econometrics* (New York: Academic Press, 105-142).

Meza, S. and Sudhir, K. (2010) "Do private labels increase retailer bargaining power?" Quantitative Marketing and Economics

Miller, N. H. Remer, M., Ryan, C., and Sheu, G. (2016). "Pass-through and the prediction of merger price effects" *The Journal of Industrial Economics*, 64(4):683-709.

Miller, N. H. Remer, M., Ryan, C., and Sheu, G. (2017) "Upward pricing pressure as a predictor of merger price effects" *International Journal of Industrial Organization*, 52:216-247.

Miravete, E., Seim, K., and Thurk, J. (2017) "One Markup to Rule Them All: Taxation by Liquor Pricing Regulation" NBER Working Paper No. 24124.

Nevo (2001) "Measuring market power in the ready-to-eat cereal industry". *Econometrica*, 69(2):307-342.

Nevo, A. (2000) "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand". *Journal of Economics and Management Strategy*, 9(4), 513548.

Oliveira, V., Frazao, E., and Smallwood, D. (2010) "Rising Infant Formula Costs to the WIC Program: Recent Trends in Rebates and Wholesale Prices" USDA Economic Research Report No. 93.

Oliveira, V., Prell, M., Frazao, E., and Smallwood, D. (2004) "WIC and the Retail Price of Infant Formula" USDA Food Assistance and Nutrition Research Report No. 39.

Petrin, A. (2002) "Quantifying the Benefits of New Products: The Case of the Minivan" *Journal of Political Economy*, 110(4), 705-725.

Reisinger, M., and Tarantino, E. (2019)"Patent pools, vertical integration, and downstream competition." *The RAND Journal of Economics*, 50(1):168-200.

Sayman, S., Hoch, S., and Raju, J. (2002) "Positioning of Store Brands" *Marketing Science*, 21(4), 378-397.

Scott-Morton, F., and Zettelmeyer, F. (2004) "The Strategic Positioning of Store Brands in RetailerManufacturer Negotiations" *Review of Industrial Organization*, 24, 161-194.

Sheu, G. and Taragin, C. (2020) "Simulating mergers in a vertical supply chain with bargaining." *The RAND Journal of Economics*.

Springel, K. (2017). "Network Externality and Subsidy Structure in Two-Sided Markets: Evidence from Electric Vehicle Incentives" *Working Paper*.

Starc, A. and Town, R. (2015) "Internalizing Behavioral Externalities: Benefit Integration in Health Insurance" NBER Working Paper.

Sweeting, A. (2010) "The Effects of Mergers on Product Positioning: Evidence from the

Music Radio Industry." *RAND Journal of Economics* 41(2), 372-97.

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

Vives, X. (2001) "Oligopoly pricing: Old ideas and new tools. MIT Press Books, 1.

Wollman, T. (2017) "Trucks without Bailouts: Equilibrium Product Characteristics for Commercial Vehicles" Working Paper, The University of Chicago Booth School of Business.