Merger, Product Repositioning and Firm Entry: the Retail Craft Beer Market in California*

Ying Fan† University of Michigan, CEPR and NBER
Chenyu Yang‡ University of Maryland

October 5, 2020

Abstract

We study the effects of merger on firm entry, product repositioning and prices in the retail craft beer market in California. To deal with selection on unobserved fixed cost shocks, we develop a new method to estimate multiple-discrete choice models. The method is based on bounds of conditional choice probabilities and does not require solving the game. Using the estimated model, we simulate a counterfactual merger where a large brewery acquires multiple craft breweries. In most markets, we find that new firms enter, non-merging incumbents add products, and merging firms drop products. However, the net effects of product variety from firm entry and product repositioning differ considerably across markets. Larger markets are more likely to see an increase in product variety, which moderates the loss of consumer surplus from the merger's price effects. In a majority of smaller markets, product variety decreases, exacerbating the welfare loss from the price effects.

1 Introduction

In antitrust litigation, merging parties could defend a merger proposal by using the potential entry argument: if a merger increases prices, the resulting greater profit opportunity will attract entry, which then compensates for the lost competition and thus curbs the price increase, mitigating the negative effect of the merger. One assumption behind this argument is that the incumbent firms do not change their product offerings. Does a merger cause the incumbents to add or drop products? How likely does entry occur? What is the overall impact of product repositioning and entry on welfare? Does the changes of product variety offset the negative price effects? How do all these effects vary across markets? In this paper, we address these questions and study the effects of

*We want to thank the excellent research assistance of Zibin Huang, Sueyoul Kim and Xinlu Yao. The results in the paper are calculated based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.
†Department of Economics, University of Michigan, 611 Tappan Street, Ann Arbor, MI 48109; yingfan@umich.edu.
‡Department of Economics, University of Maryland, College Park, MD 20740; cyang111@umd.edu.

merger on prices, product repositioning and new firm entry in the context of the retail craft beer market in California.

The craft beer industry provides an ideal empirical context to study the effects of merger on the entry and product variety of multi-product firms. The craft industry is a growing segment of the beer industry. Between 2006 and 2016, while major breweries such as ABI and MillerCoors saw the sales of their main (non-craft) products plateau, the craft beer market experienced substantial growth in variety and sales (Hart and Alston (2019)). Craft breweries have thus become popular targets of acquisitions, attracting the concerns of the antitrust regulators (Codog (2018)). In addition, there are rich demographic variations across geographical markets that help to identify consumer tastes and incentives of entry and product positioning. We focus on the state of California, which has the largest number of craft breweries and the highest craft beer production among the US states, with 462 craft breweries (12% of all US craft breweries) and 43 million barrels of production (18% of all US craft beer production) in 2015, according to the Brewers Association, a trade group in the beer industry.

To address our research questions, we set up a model to describe demand and firm decisions in the retail beer market in California. The demand side is a discrete choice model where we allow for both observed and unobserved heterogeneity in consumer tastes. The supply side is a static two-stage model. In the first stage, firms observe the characteristics and fixed cost of each potential product, and choose the set of products to sell in a market (defined as a retailer-county pair). A firm can choose an empty set, indicating no entry. In the second stage, firms observe shocks to demand and marginal costs and choose prices simultaneously. The structure of the game is similar to the prior empirical work on product variety (e.g., Eizenberg (2014); Wollmann (2018); Fan and Yang (2020)).

Our main data sources are Nielsen Retail Scanner Data and Nielsen Consumer Panel from 2009 to 2016. We supplement the data with information on whether a beer is considered craft using data from the Brewers Association. We further augment the data by hand-collecting the owner and brewery identities and the location of the brewery for each beer.

The key primitives in our model are consumer preferences, marginal costs and fixed costs of entry into a market. We estimate beer demand with data on market shares and individual choices. The first-order conditions for the optimal prices allow us to estimate marginal costs. Firms' product choices inform the fixed costs. We estimate the distribution of fixed costs with a new method that takes into account selection on fixed cost unobservables. Our method relies on a construction of two-sided bounds for the probability that a product is in a market. The construction is based on the following intuition: for a binary action a, the equilibrium choice probability of $a = 1$ is larger than the probability that $a = 1$ is a dominant strategy (so that $a = 1$ *must* be a best response) and smaller than the probability that $a = 0$ is not a dominant strategy (so that $a = 1$ *can* be a best response).

Using the estimated model, we simulate the effects of merger for a number of markets in California in 2016. In our baseline simulations, the largest macro brewery acquires 8 large craft breweries.

We find that the merger causes both new firm entry and product repositioning. In particular, merging firms tend to drop products, while the non-merging incumbents tend to add products. However, the net changes of variety (from new entry and product repositioning) and the associated welfare impacts are quite heterogeneous across markets. Specifically, the number of products increases in 25 markets and decreases in 114 markets out of the 149 markets we simulate. Moreover, larger markets are more likely to see an increase in the number of products and a positive welfare effect attributed to variety changes. In this case, we find that the positive product variety effects offsets the negative price effects, but only partially. In smaller markets, the net change in product variety tends to be negative, which exacerbates the negative price effects. Across most markets, new entry does occur and improves consumer welfare, but it often does not fully compensate for the decline in product variety from incumbent product repositioning or offsets the price effects.

Our contributions are two-fold. First, we contribute to the literature of merger and product variety. In this paper, we consider the entry of firms and products in the context of multi-product firms. Entry defense has long been recognized in policy guidelines and investigated in both theoretical (for example, Spector (2003); Anderson et al. (2020); Caradonna et al. (2020)) and simulation or empirical studies (for example, Werden and Froeb (1998); Cabral (2003); Gandhi et al. (2008); Ciliberto et al. (2018)). Different from these papers, which focus on entry of single-product firms, we consider multi-product firms. Therefore, it is possible in our model for a merger to decrease product variety even when the merger causes entry, because the incumbents can reduce product offerings. The US Horizontal Merger Guidelines have started to recognize the roles of product variety in a merger.[1] Recent academic work has sought to empirically quantify how merger affects product repositioning and the associated welfare impact (Fan (2013); Wollmann (2018); Li et al. (2019); Fan and Yang (2020); Garrido (2020)) while disallowing firm entry.[2] For the retail craft beer market in California, we show significant heterogeneity in entry, repositioning and the net changes of variety across markets. We show that the size of the market, the market power of the merging firms and the fixed costs of the merging firms and their rivals are important determinants of the merger outcomes.

Our second contribution is a new method for estimating multiple-discrete choice games. The simulations to quantify the extent of entry and repositioning critically depend on the estimates of fixed costs. An important step in our analysis is to account for the following selection issue: the fixed costs (unobserved by the researcher but observed by the firms) associated with products currently not in the market are likely higher than those in the market. Specifically, we estimate the distribution of unobserved fixed cost shocks by constructing bounds for the conditional choice

[1] The 2010 Guidelines state, "(t)he Agencies also consider whether a merger is likely to give the merged firm an incentive to cease offering one of the relevant products sold by the merging parties. Reductions in variety following a merger may or may not be anticompetitive. Mergers can lead to the efficient consolidation of products when variety offers little in value to customers. In other cases, a merger may increase variety by encouraging the merged firm to reposition its products to be more differentiated from one another". In comparison, the 1997 US Merger Guidelines did not explicitly mention the roles of product variety.

[2] In the radio industry, a number of papers (e.g., Berry and Waldfogel (2001); Sweeting (2010); Jeziorski (2015)) have studied merger, entry and variety but cannot quantify the impact on consumer welfare because radio stations do not set prices to listeners.

probability (CCP) that a product is in a market conditional on the primitives of the market. The lower bound is the probability of a sufficient condition for choosing the product: the fixed cost shock is so low that choosing the product is a dominant strategy. The upper bound is the probability of a necessary condition: the fixed cost shock is not too high, and not choosing the product is not a dominant strategy. These bounds hold for any equilibrium selection rules and even when the selection rules vary across markets. Computing these bounds does not require solving for an equilibrium of product choices.

Our method is similar to Ciliberto and Tamer (2009) in two dimensions: both approaches (1) specify the distribution of the fixed cost shocks and (2) rely on bounds of the probability that an observed outcome satisfies certain equilibrium conditions. The key difference between the two approaches is the construction of the bounds. Ciliberto and Tamer (2009) uses the probability that an outcome is the unique equilibrium of the game and the probability that the outcome is one of multiple equilibria as the lower and upper bounds. Therefore, to compute these bounds in estimation, one has to simulate multiple fixed cost draws, and for each simulation draw, solve for all equilibria. This can be computationally costly if not prohibitive for a multi-product firm in a game with many firms. In our approach, we construct bounds for the CCP of a single action: whether a given product is in a market. These bounds are easier to compute: the bounds are probabilities of whether a one-dimensional shock is above or below certain cutoffs. Computing these cutoffs does not require solving the game. Our approach therefore is especially applicable to estimating games with many players or actions. By using empirical estimates of CCP in moment inequalities, our method is related to the CCP approach in the estimation of dynamic models (Bajari et al. (2007)).

Our paper is closely related to two recent papers on entry (Ciliberto et al. (2018)) and product repositioning (Li et al. (2019)) in the airline industry. In both papers, a firm chooses a binary action in a first stage and then chooses prices in a second stage. In Ciliberto et al. (2018), the binary action is whether to enter. In Li et al. (2019), the action is whether to provide non-stop service. There are three modeling differences. First, we consider multi-product firms where the first stage action for each firm consists of a vector of binary actions. Therefore, we discuss product variety together with firm entry, and we use a new method to address the estimation challenges that arise from allowing for a greater action space. In our model, even when a merger causes entry, the overall product variety can decrease when the (multi-product) incumbents reduce product offerings. Second, Ciliberto et al. (2018) and Li et al. (2019) assume that firms observe demand and marginal cost shocks as well as fixed cost shocks when the firms make decisions on entry, thus accounting for selection on unobserved demand and marginal costs, in addition to the selection on unobserved fixed cost shocks. We address the selection on demand and marginal costs by including a large number of fixed effects. The remaining unobservables are transient shocks, and we find it reasonable to assume firms do not observe them when making product choices. We also show that these transient shocks are small. Third, related to the second point, Ciliberto et al. (2018) and Li et al. (2019) allow for correlations among unobserved demand, marginal cost and fixed cost shocks. We include common observable covariates in demand, marginal cost and fixed cost shocks to allow for correlation through observables.

Another approach in the literature of estimating discrete games exploits moment inequalities derived from a necessary equilibrium condition that no firm has an incentive to unilaterally deviate from the observed equilibrium (e.g., Ho (2009); Eizenberg (2014); Pakes et al. (2015); Wollmann (2018)). Such an approach typically relies on a mean zero or conditional mean zero assumption, and does not require specifying a distribution for the fixed cost unobservables. This approach also does not estimate this distribution. In our empirical application, we find that the variance of the fixed-cost shocks is substantial. Ignoring this variance and simply using the mean fixed cost for counterfactual simulations may give an incomplete answer to the questions of interest when, for example, it would be useful to know the range of the effects of the merger.[3]

Overall, we consider our approach complementary to existing papers on estimating discrete games. Our approach is suitable for a setting where solving for equilibria is costly and the unobserved shock to the fixed-cost of entry is potentially important to address the research questions of interest.

The rest of the paper is organized as follows. We first use an illustrative model to explain our bounds in Section 2. Section 3 discusses the craft beer market in California and the data. Section 4 presents the model. Section 5 explains the estimation strategy and presents the estimation results. Section 6 describes the counterfactual designs and results. Finally, we conclude in Section 7.

2 An Illustrative Model

In this section, we use a two-firm entry model to illustrate how we construct our bounds used for estimation. We consider two firms 1 and 2 that decide whether to enter in markets $m = 1, \dots, M$. We assume that the firm decisions form a complete information Nash equilibrium. In market m, the binary entry decisions Y_{1m} and Y_{2m} satisfy the following simultaneous equations

$$Y_{1m} = 1\left[\pi_{1m}\left(Y_{2m}\right) - C\left(W_{1m}, \theta\right) - \zeta_{1m} \geq 0\right]$$

$$Y_{2m} = 1\left[\pi_{2m}\left(Y_{1m}\right) - C\left(W_{2m}, \theta\right) - \zeta_{2m} \geq 0\right],$$

[3] There are three other alternative estimation approaches. First, one can invoke additional assumptions so that the equilibrium of the game is unique (or unique in some aspects) and the action space of players is small, which makes it possible to estimate the model with maximum likelihood (Reiss and Spiller (1989); Garrido (2020)) or simulated method of moments (Berry (1992); Li et al. (2019)). These assumptions may not be desirable in other empirical contexts, and for applications as large as ours, the likelihood function can be intractable and the simulations costly. Alternatively, Illanes (2017) estimates a dynamic discrete choice problem using a semi-parametric latent variable integration method (Schennach (2014)) to deal with selection in unobservables. Like us, this approach also avoids solving a game or an optimization problem, but can result in relatively wide (and sometimes unbounded) confidence sets of parameters. Finally, Fan and Yang (2020) directly make assumptions about the distribution of the unobserved fixed cost shock conditional on the observed equilibrium for their merger simulations. In comparison, the approach in this paper estimates this distribution.

where $\pi_{im}(Y_{-im})$, $i \in \{1,2\}$ is a variable profit function known to the firms and researchers, $C(W_{im},\theta)$ is the fixed cost of entry, W_{im} is a vector of covariates, θ is a vector of parameters to be estimated, ζ_{im} is a fixed cost shock known to the firms but unobserved by the researchers. The shock ζ_{im} is distributed $F_{\zeta}(\zeta,\sigma)$ where σ is a vector of distributional parameters to be estimated.

This model poses well–known estimation challenges due to the presence of multiple equilibria and unspecified equilibrium selection rule (Tamer (2003)). When the number of firms is small and it is feasible to enumerate all possible equilibria, there exist several approaches to estimate the fixed cost function without fully specifying equilibrium selection rules (as reviewed in the Introduction). Our approach below is more scalable for games with more players.

Specifically, we construct bounds for $\Pr(Y_{im}=1\,|W_m)$, where we use W_m to denote (W_{1m},W_{2m}) to save notation. We consider a behavioral assumption that is weaker than Nash equilibrium:

Assumption 1. *Y_{im} for $i \in \{1,2\}$ is not a dominated strategy.*

The assumption implies the following bounds

$$\begin{aligned}
&\Pr(Y_{im}=1 \text{ is a dominant strategy})\\
\le&\Pr(Y_{im}=1\,|W_{im},W_{-im})\\
\le&\Pr(Y_{im}=1 \text{ is not a dominated strategy}).
\end{aligned}$$

We further assume that rival entry reduces the own profit:

Assumption 2. $\pi_{im}(1) < \pi_{im}(0)$.

Under this assumption, the entry of firm 1 ($Y_{1m}=1$) is a dominant strategy when firm 1 enters even if firm 2 enters, and entry is a dominated strategy when firm 1 does not enter even if firm 2 does not enter. Our bounds thus become

$$\begin{aligned}
&\Pr(\zeta_{im} < \pi_{im}(1) - C(W_{im},\theta))\\
\le&\Pr(Y_{im}=1\,|W_{im},W_{-im})\\
\le&1-\Pr(\zeta_{im} > \pi_{im}(0) - C(W_{im},\theta)),
\end{aligned}$$

where the middle term is the observed conditional probability of entry, which is estimable from data. These bounds have several advantages. First, these bounds contain useful information on the fixed cost function. In the extreme case where $\pi_{im}(1)=\pi_{im}(0)$, the inequalities become equalities, and estimation based on them becomes GMM estimation of binary choice models (McFadden (1989)). The usefulness of the inequalities depends on the difference of $\pi_{im}(1)$ and $\pi_{im}(0)$, which we explore in the later estimation section and Monte Carlo exercises. Second, these bounds are one-dimensional CDFs and easy to compute. The numerical integration does not suffer from a dimensionality problem in a game with more firms. Third, the bounds do not rely on equilibrium selection assumptions. Specifically, these bounds hold when (1) there are multiple equilibria, (2) the equilibrium selection mechanism differs across markets, or (3) firms use mixed or pure strategies.

3 Industry and Data

Our empirical analysis focuses on the retail craft beer market in the state of California. According to the 2015 Brewers Association estimates, California accounted for 18% of craft beer volume and 12% of craft breweries in the nation, the highest among all US states. Like the rest of the US, California has a "three-tier" system of beer distribution, consisting of breweries, distributors and retailers. To simplify our analysis and focus on the entry and product variety decisions of breweries, we abstract from the strategic behaviors of distributors and retailers. Several institutional features of the California craft beer market help to justify the simplification. First, California places no cap on the volume a brewery can distribute its products without a third-party distributor, essentially empowering breweries to become distributors if the costs of distributing through third parties are too high (Anhalt (2016)). Second, compared with other states, California beer statues do not require breweries to satisfy a burdensome "good cause" clause to terminate a contract with a distributor. The staff of DOJ and FTC on two occasions objected to legislation that sought to add such "good cause" provisions, viewing them as anti-competitive, increasing the distributors' market power and harmful to the growth of the craft beer industry (Ohlhausen et al. (2005); Sayyed et al. (2020)). Taken together, we find it reasonable to assume that the distribution market is sufficiently competitive, and we do not consider the strategic behaviors of distributors.[4] Furthermore, in addition to federal statues that prohibit "tied-houses", vertical relationships between manufacturers and retailers that exclude small alcoholic beverage makers such as craft breweries from retailers, California additionally passed its own "tied-house" laws and unfair competition laws to prevent exclusion (Croxall (2019)). Theses institutional features motivate our modeling assumption that craft breweries play the main role in making the entry and product variety decisions based on the retail profitability of their products.[5]

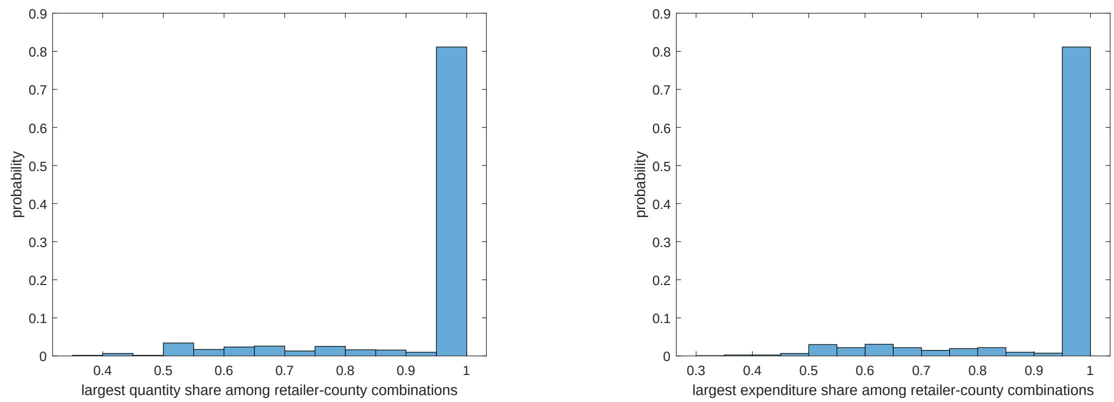
Our analysis is based on the product, sales and price information of beers sold in major retailer chains in the Nielsen data. We use both the aggregate data in the Nielsen Retail Scanner Data and the micro-level panel data in the Nielsen Consumer Panel between 2009 and 2016. We supplement the data with information on whether a beer is considered craft based on the designation by the Brewers Association. We further add hand-collected data on the identities of the owner and brewery and the location of the brewery of each product in our data. Finally, we merge the data with county demographics from the Census. We define a firm to be a corporate owner (e.g., Boston Beer Company) and a product to be a brand (e.g., Samuel Adams Boston Lager). A firm can own multiple breweries and products. We aggregate the Nielsen data from its original UPC/week level to the product/month level by homogenizing the size of a product (so that a unit is a 12-ounce-12-pack

[4] We note that the loosening of craft beer distribution laws has become a recent national trends. For example, in 2019, North Carolina enacted the Craft Beer Distribution and Modernization Act (HB 363) to increase the quota that a craft brewery can self distribute without a distributor, and Maryland both increased the quota and lessened the burden of a "good cause" termination with Brewery Modernization Act and Beer Franchise Law (HB1010, HB1080). Similar changes have occurred or are proposed in Illinois, Massachusetts, Tennessee, Texas and other states.

[5] Another recent trend is the acquisition of independent distributors by large breweries (Sayyed et al. (2020)). Our analysis focuses on the case where the cost of distributing products to a market does not significantly change in our counterfactual analysis (we simulate the scenario where a large brewery acquires multiple craft breweries).

Figure 1: Cross–Market Beer Purchases in 2016

(a) Quantity (b) Expenditure



equivalent).[6]

We define a market as a retailer-county pair. This definition allows us to interpret the fixed cost of product entry in our model as the cost of entry into a retailer chain in a county (e.g. overhead costs of transportation and storage, which the brewery pays directly if it self distributes, or it pays through a distributor in a competitive distributor market). We do not consider retailer competition. Our data suggest cross-retailer shopping appears rare. In Figure 1, we show the distribution of a household's largest shares of beer quantities and expenditures at a retailer-county combination among all retailer-county combinations in 2016. About 80% of the households purchased all of their beers from one retailer-county combination.[7]

We consider a product was "in" a market in a calendar year if the product sold more than 20 units in a month for more than 6 months in the market in the year. Moreover, for craft products, we keep those by the top 60 craft breweries (by national volume in 2015) in the Brewers Association data. We thus focus on breweries established in the 1990s or earlier. Many of these craft breweries have sold beers on their own premise and through other avenues before entering the retailers in the Nielsen data. We do not consider the potentially dynamic problem of new brewery or brand creation.[8] In the end, our sample covers 83% of California craft beer quantity in the Nielsen Scanner Data. Although it is not possible to directly compare the importance of the retail craft beer market

[6]We add quantities across weeks within a month and use the quantity weighted price as the product's price in that month.

[7]Huang et al. (2020) and Illanes and Moshary (2020) find little evidence of retailer competition in the spirit category.

[8]In the radio industry, Berry and Waldfogel (2001) shows that a preemption motive may lead to an increase in variety post-merger.

with the “on-premise” market (such as taprooms, bars and restaurants) using our data, which cover just the retail segment, the Brewers Association suggested that on-premise channels account for 35% of the craft volume (Watson (2016)).[9]

The number of markets, firms and products vary across the years. In 2016, there are 178 markets, 51 firms, 37 craft firms, 255 products and 111,219 product-market-month observations. We provide summary statistics year by year in Figure 2. To make the time-series comparable, we condition on the 109 markets present in every year from 2009 to 2016. All dollar values are in 2016 dollars. The total annual beer sales from these markets decreased from 61 to 46 million units (12-ounce-12-pack), while the craft sales increased from 4.1 million in 2009 to 6.4 million units in 2016 (Figure 2 (a)). The average price is stable around 11 dollars per unit. The average price of craft products increased from 16 to 17 dollars (Figure 2 (b)). There are an average of 52 firms in total and 35 craft firms (Figure 2 (c)). Both the numbers of firms and products first increased and then decreased over time, driven by changes in the numbers of craft firms and products. In the 109 markets present through all years in our data, there were 31 craft firms and 93 craft products in 2009, compared with 36 craft firms and 135 craft products in 2016. (Figure 2 (c) and Figure 2 (d)). The general trend is similar if we include all the markets in each year.

We make two additional observations that help to motivate our modeling and identification strategies. First, in Figure 2 (e), we plot the histogram of the distance from a product's brewery to a market in 2016. The y–axis shows the count of the unique product-market pairs in the data. The majority of the in-state craft breweries distribute close to their breweries. Distance potentially plays important roles in the demand, marginal cost and fixed cost of a craft beer: a popular local beer may struggle to gain traction in markets further away, because it lacks the name recognition of the well-known national brands (Tamayo (2009)); the transportation cost factors into the marginal cost of a beer (Ashenfelter et al. (2015)); it may be hard to secure a reliable long-distance distributor. We explicitly account for these product-market effects in our demand, marginal cost and fixed cost estimation. This effect of distance on entry is less obvious for the out-of-state craft beers produced by larger breweries.

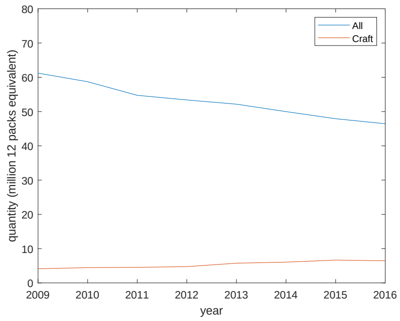
In Figure 2 (f), we show that there is a strong positive correlation between the size of a market and the number of craft products in the market. We define the market size as the average monthly alcohol sales in a market (in the unit of a 12-once-12-pack equivalent) times 8, which is the median number of household trips to a retail store in the panel data.[10] While Figure 2 (f) only plots market size and the number of products for our 178 markets in 2016, the plots for other years are similar. The local population and retailer chains drive the differences in market sizes, which could affect profits and thus product choice and entry. We leverage this variation (in addition to the variations of other covariates such as distances from a brewery to different markets) to identify the fixed costs of entry.

[9] Probably due to similar data limitations, prior work on the beer industry (Ashenfelter et al. (2015); Asker (2016); Miller and Weinberg (2017); Miller et al. (2019)) have also focused on the retail segment.

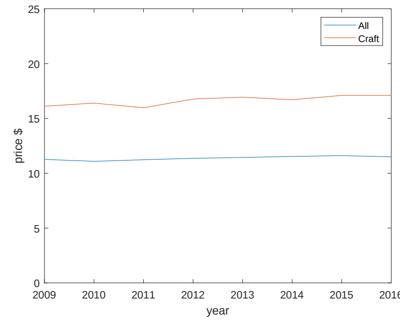
[10] Our results are robust to alternative scaling factors.

Figure 2: Summary Statistics

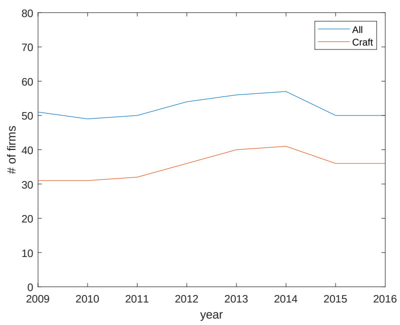
(a) Quantity



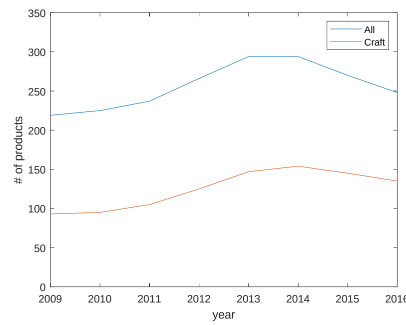
(b) Price



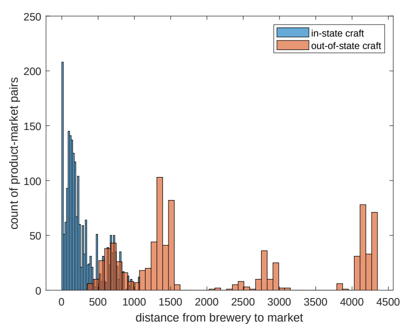
(c) Number of Firms



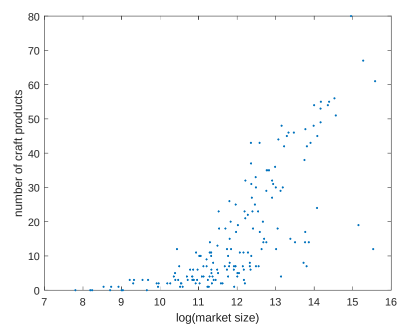
(d) Number of Products



(e) Distance from the Brewery to the Market



(f) Market Size and Number of Craft Products



4 Model

4.1 Demand

We describe the demand for beer with a random-coefficient discrete-choice model. A product's characteristics include its flavor type (lager, light and others), whether they are designated as craft products, and whether they are imported from outside North America. For example, Bud Light is a light, non-craft, North American beer, while Samuel Adams Lager is a lager, craft, North American beer. These characteristics of product j are captured by a vector of indicator variables $\boldsymbol{x}_j = \left(x_j^{\text{lager}}, x_j^{\text{light}}, x_j^{\text{craft}}, x_j^{\text{import}}\right)$. We allow both household income and unobserved heterogeneity to affect preferences. We specify the utility function of household i in market m from product j in month t as

$$\begin{aligned} u_{ijmt} =& (\sigma_0\nu_i + \kappa_0 y_i) + (\alpha + \kappa_1 y_i)\, p_{jmt} \\ & + \sigma^{\text{lager}}\nu_i^{\text{lager}} x_j^{\text{lager}} + \sigma^{\text{light}}\nu_i^{\text{light}} x_j^{\text{light}} + \sigma^{\text{import}}\nu_i^{\text{import}} x_j^{\text{import}} \\ & + (\sigma^{\text{craft}}\nu_i^{\text{craft}} + \kappa^{\text{craft}} y_i) x_j^{\text{craft}} \\ & + \xi_t + \xi_{jm} + \xi_{jmt}, \end{aligned} \tag{1}$$

where y_i is the household income and $\nu_i^{(\cdot)}$ is the household-specific unobserved taste shock to each product attribute, which follows a standard normal distribution and is independent across households. Therefore, the $\sigma^{(\cdot)}$ parameters capture the dispersion in unobserved household tastes, and the $\kappa^{(\cdot)}$ parameters measure the effect of household income on tastes. We also include month fixed effects (ξ_t) and product-market fixed effects (ξ_{jm}) to capture unobserved factors that vary across months and product-market pairs. Finally, the error term ξ_{jmt} captures the month-to-month variations of demand shocks specific to a product, market and month combination. We do not include mean coefficients for $\boldsymbol{x}_j$ because they are absorbed in fixed effects ξ_{jm}.

This specification gives us the market share $s_{jmt}(p_{jmt}, p_{-jmt})$, corresponding with the familiar mixed logit choice probability formula (Berry et al. (1995); Nevo (2001)), of product j in month t and market m, where p_{-jmt} is a vector of the prices of all other products in market m and month t. Other determinants of demand (product characteristics, fixed effects and demand shocks of all products in the market) are absorbed by the subscript jmt of the function $s_{jmt}(\cdot,\cdot)$. Multiplying the market share by the corresponding market size gives us the demand for product j, $D_{jmt}(p_{jmt}, p_{-jmt})$.

4.2 Supply

The supply side is a two-stage static model. At the beginning of each year τ, firms simultaneously choose which beers, if any, to sell in each market. This product choice is fixed through the year and is denoted by $\mathcal{J}_{nm\tau}$ for firm n's products in market m in year τ. Then, in each month t, after observing that month's demand and marginal cost shocks, firms simultaneously choose retail prices for t.[11] We start from the second stage.

Stage 2. Pricing In month t, firm n observes the month effect ξ_t, product characteristics $(\boldsymbol{x}_{jm}, \xi_{jm}, \xi_{jmt})$ and the marginal cost for each product j in the market m. The firm then chooses prices p_{jmt} for all $j \in \mathcal{J}_{nm\tau}$ to maximize its total variable profits:

$$\max_{p_{jmt}, j \in \mathcal{J}_{nm\tau}} \sum_{j \in \mathcal{J}_{nm\tau}} (p_{jmt} - mc_{jmt}) \, D_{jmt}(p_{jmt}, p_{-jmt}) \, . \tag{2}$$

The marginal cost mc_{jmt} is decomposed into a product-market effect ω_{jm} and a product-market-time specific shock ω_{jmt}:

$$mc_{jmt} = \omega_t + \omega_{jm} + \omega_{jmt}. \tag{3}$$

Stage 1. Entry and Product Decisions At the beginning of each year τ, each firm n is endowed with a set of potential products $\mathcal{J}_{n\tau}$. In the first stage, each firm decides on its set of products $\mathcal{J}_{nm\tau}$ in market m for the year τ to maximize the expected profit, which is the difference between the expected variable profit π_{nm} and the fixed cost C_{nm}:

$$\max_{\mathcal{J}_{nm\tau} \subseteq \mathcal{J}_{n\tau}} \pi_{nm}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}) - C_{nm}(\mathcal{J}_{nm\tau}) \, . \tag{4}$$

We consider a simultaneous-move, complete information Nash equilibrium: firms observe π_{nm}, C_{nm} and $\mathcal{J}_{n\tau}$ for any n when choosing products. Firm can use pure or mixed strategies.

We now specify the expected variable profit and the fixed cost. We first make a timing assumption: when making product decisions, firms observe the product characteristics $\boldsymbol{x}_{jm}$, time fixed effects (ξ_t, ω_t), product-market characteristics (ξ_{jm}, ω_{jm}), and fixed costs for any product $j \in \mathcal{J}_{n\tau}$ and any firm n. After firms make product decisions, the month-to-month transient demand and marginal cost shocks $(\xi_{jmt}, \omega_{jmt})$ realize in the second stage. Given this timing assumption, the expected variable profit is the sum of expected values in (2) across the months in the year, and the expectation is taken over all $(\xi_{jmt}, \omega_{jmt})$.

Specifically, for product j in the set $\mathcal{J}_{nm\tau}$ and month t in year τ, we define $p_{jmt}(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$ to be the second-stage equilibrium prices that depend on product characteristics, product-market fixed effects and transient shocks. We use $\mathcal{J}_{-nm\tau}$ to denote the set of products that firm n's

[11] As mentioned in Section 2, we do not model retailer problems. On the technical side, this simplification allows us to avoid excessive computational burdens, especially in merger simulations. The underlying assumption is that efficient contracting resolves the double marginalization problem conditional on product entry. We show in Table 2 of Section 5.1 that our markup estimates are reasonable.

competitors sell in market m. Firm n's expected annual profit π_{nm} in (4) is, therefore,

$$\begin{aligned}&\pi_{nm}\left(\mathcal{J}_{nm\tau},\mathcal{J}_{-nm\tau}\right)\\=\sum_{t=1}^{12}E_{\xi_{jmt},\omega_{jmt}}\Bigg\{&\sum_{j\in\mathcal{J}_{nm\tau}}\left(p_{jmt}(\mathcal{J}_{nm\tau},\mathcal{J}_{-nm\tau})-mc_{jmt}\right)\\&\cdot D_{jmt}\left(p_{jmt}(\mathcal{J}_{nm\tau},\mathcal{J}_{-nm\tau}),p_{-jmt}(\mathcal{J}_{nm\tau},\mathcal{J}_{-nm\tau})\right)\Bigg\}.\end{aligned}\tag{5}$$

The fixed cost function in (4) is specified as

$$C_{nm}\left(\mathcal{J}_{nm\tau}\right)=\sum_{j\in\mathcal{J}_{nm\tau}}\Big(\theta_0+\theta_1\text{craft}_j+\theta_2(\text{in state})_j\cdot\text{craft}_j+\theta_3(\text{in state})_j\cdot\text{craft}_j\cdot\text{dist}_{jm}+\zeta_{jm\tau}\Big),\tag{6}$$

where craft_j indicates whether the product is a craft beer, $(\text{in state})_j$ indicates whether product j is produced in California, and dist_{jm} is the distance between product j's brewery and market m. The specification is motivated by the evidence in Section 3 that distance matters for the in-state craft breweries, but less for the out-of-state ones. Finally, the fixed cost shock $\zeta_{jm\tau}$ is assumed to be i.i.d. across j, m, τ and follows a normal distribution with mean 0 and standard deviation σ_ζ. In this specification, θ_1 measures the difference between craft and non-craft breweries in the fixed cost, θ_2 the difference between in-state and out-of-state craft breweries, and θ_3 the effect of distance on the fixed cost for the in-state craft breweries. This baseline specification of the fixed cost function rules out economies or diseconomies of scope. We extend the model to allow for this possibility in Appendix E.

5 Estimation

5.1 Estimation of Demand Parameters and Marginal Costs

We combine the aggregate data of product-market shares and individual-level panel data of consumer purchases to estimate demand parameters. Specifically, we rely on the market share data to identify the mean price coefficient (α) and the fixed effect parameters (ξ_t, ξ_{jm}). The panel data and the correlations between household income and beer purchases help to identify the standard deviations of the unobservable consumer heterogeneity ($\sigma^{(\cdot)}$ parameters) and the effect of household income on consumer taste ($\kappa^{(\cdot)}$ parameters). We estimate these parameters using the Generalized Method of Moments approach where we combine a set of macro moments and two sets of micro moments.

To deal with the price endogeneity, the macro moments are based on the Hausman instrument V_{jmt}, which is the average price of product j in a "ring" consisting of markets more than 500KM from market m but less than 1000KM away.[12] The identifying assumption is

$$E\left(\xi_{jmt} \mid \boldsymbol{x}_j, V_{jmt}, t=1,\ldots,12\right) = 0.$$

To construct these macro moments, we first invert out the mean utility component (Berry et al. (1995)) and construct the residuals $\xi_{jmt} - \bar{\xi}_{jm}$ after a within transformation, where $\bar{\xi}_{jm}$ is the average ξ_{jmt} across the months j is in market m. We then interact the transformed residuals with $\boldsymbol{x}_j$ as well as the Hausman instrument.[13]

We next specify the micro moments. We construct the first set of micro moments to identify the standard deviation parameters $\sigma^{(\cdot)}$. We first provide some intuition behind our construction: if the parameter σ^{craft} is large, a consumer's preference for craft products should be highly correlated across months, and therefore we should expect strong correlations of a consumer's purchase decisions across months. The implication is that conditional on a consumer ever purchasing a craft product, the consumer should purchase many craft products throughout the year if σ^{craft} is large.

We thus match the model predictions and the empirical counterparts of the following moments:

- A household i's expected annual purchase of a certain type of beer conditional on ever purchasing this type of beer in the year, i.e., $E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$, where q_{it}^f is household i's total quantity of beer with a certain flavor (f = lager or f = light) or of a certain characteristic (f = import or f = craft) in month t. Matching these moments helps to identify σ^f.

- A household i's expected annual purchase of beer conditional on purchasing beer in the year, i.e., $E\left(\sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} > 0\right)$, where q_{it} is household i's total beer purchase. Matching this moment helps to identify σ_0.

We present the analytic expressions of the moments above in Appendix A.

We also use a second set of micro moments that are helpful to identify the income effect on consumer tastes:

- The average price of the purchased beer among households whose income falls into a bin $\mathcal{I}$, i.e., $E\left(p_{j(i)mt} \mid y_i \in \mathcal{I}\right)$, where $p_{j(i)mt}$ is the price of the product purchased by household i in market m and month t, the income bins $\mathcal{I}$ are $(0, \$50K]$, $(\$50, \$100K]$ or $(\$100K, \$150K)$. Matching these moments helps to identify the income effect on price sensitivity κ_1.

- $E\left(\sum_{t=1}^{12} q_{it}^{\text{craft}} \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} > 0, y_i \in \mathcal{I}\right)$, which helps to identify κ^{craft}.

- $E\left(\sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} > 0, y_i \in \mathcal{I}\right)$, which help to identify κ_0.

[12] To be precise, our instruments consist of variables $\left(\overline{p}_{jmt}, Z_{jmt}\right)$, where if j is available in the given "ring", $\overline{p}_{jmt}$ is the Hausman instrument and $Z_{jmt} = 0$. Otherwise $\overline{p}_{jmt} = 0$ and $Z_{jmt} = 1$. The results are robust to further increasing the distances or using the simple average or quantity-weighted average of prices.

[13] Miller and Weinberg (2017) propose using the brewery–market distance as a price instrument. From our discussion in Section 3, distance may not be an excluded variable for craft beer demand. We thus include a product-market component in our demand model and difference out this component after the within transformation. Distance would no longer be useful as an instrument because it is at the product-market level.

Table 1: Demand Estimates

		2009	2010	2011	2012	2013	2014	2015	2016
unobs. heterogeneity	σ_0	0.71	0.02	0.01	0.08	0.91	0.04	0.56	0.32
		(0.02)	(0.22)	(0.11)	(0.05)	(0.00)	(0.01)	(0.01)	(0.05)
	σ^{lager}	0.72	0.08	0.01	0.01	3.28	3.83	1.60	0.25
		(0.03)	(0.07)	(0.09)	(0.17)	(0.00)	(0.00)	(0.01)	(0.09)
	σ^{light}	4.46	4.95	5.00	3.69	5.99	4.86	4.94	6.67
		(0.01)	(0.03)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)	(0.01)
	σ^{craft}	9.31	16.44	16.44	8.03	9.78	5.29	4.90	4.68
		(0.02)	(0.03)	(0.05)	(0.02)	(0.01)	(0.01)	(0.00)	(0.01)
	σ^{import}	1.09	0.24	2.58	2.81	4.26	5.22	0.03	0.88
		(0.01)	(0.09)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)
income effect	κ_0	-5.82	-2.30	-3.09	-3.20	-5.46	-4.31	-6.05	-6.89
		(0.02)	(0.02)	(0.03)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)
	κ^{craft}	3.26	1.88	4.58	8.84	1.25	0.98	0.53	0.54
		(0.07)	(0.12)	(0.12)	(0.15)	(0.1)	(0.02)	(0.02)	(0.01)
	κ_1	0.58	0.28	0.31	0.33	0.32	0.55	0.35	0.37
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
mean price coef.	α	-0.97	-0.60	-0.64	-0.73	-0.61	-0.86	-0.58	-0.60
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

By using these micro-moments, the identification of demand relies on the variations of demographic characteristics, with the mean price coefficient identified by the Hausman instrument (Berry and Haile (2020)). We therefore estimate demand for each year separately, which also helps to flexibly allow for potential changes in consumer tastes over time. We report the estimates for 2009-2016 in Table 1.

The estimates indicate significant heterogeneity in preferences. In particular, although high-income households are less likely to purchase beer ($\hat{\kappa}_0 < 0$), conditional on a purchase, they are less price sensitive ($\hat{\kappa}_1 < 0$) and prefer craft products $\left(\hat{\kappa}^{\text{craft}} > 0\right)$. The estimates also show changes over time. In particular, the dispersion of tastes for craft ($\hat{\sigma}^{\text{craft}}$) as well as the positive effect of income on craft ($\hat{\kappa}^{\text{craft}}$) are larger in the early years of the sample. To see the substitution between craft and non-craft beers, we calculate the own and cross elasticities among the top-5 non-craft brands and top-5 craft brands in 2009 and 2016 (see Table 2).[14] We find that the substitution within craft brands is much larger than those across the two groups in both years, and the cross elasticities between the non-craft brands and craft brands are slightly higher in 2016 than in 2009 (and remain small).

The estimation of the marginal costs is standard and follows Berry et al. (1995): we back out marginal costs based on the first-order condition of the profit maximization problem in (2). We report the markup estimates of the top products (quantity-weighted average across markets where the product is available) in Table 2. In 2016, the median markup is \$2.99, and the median craft markup is \$3.89. All terms are 2016 dollars.

[14] Per data contract with Nielsen, we refrain from discussing the specific identities of beers or breweries in the data.

Table 2: Elasticities: Top-5 Main Brands and Top-5 Craft Brands in 2009 and 2016

	Main Brands					Craft Brands					Markup ($)
	2009										
Main Brands	-3.77	0.80	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	2.81
	0.94	-3.91	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	2.75
	0.03	0.03	-5.05	0.06	0.06	0.00	0.00	0.00	0.00	0.00	1.93
	0.02	0.01	0.05	-4.94	0.02	0.00	0.00	0.00	0.00	0.00	3.09
	0.03	0.03	0.11	0.06	-5.1	0.00	0.00	0.00	0.00	0.00	1.88
Craft	0.00	0.00	0.01	0.01	0.00	-4.76	0.20	0.25	0.19	0.14	4.25
	0.00	0.00	0.01	0.01	0.00	0.53	-4.71	0.22	0.19	0.13	3.92
	0.00	0.00	0.01	0.01	0.00	0.60	0.19	-5.13	0.19	0.14	3.90
	0.00	0.00	0.01	0.01	0.00	0.53	0.21	0.24	-4.66	0.13	4.14
	0.00	0.00	0.01	0.01	0.00	0.57	0.20	0.24	0.19	-5.01	3.67
	2016										
Main Brands	-2.18	0.77	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	4.49
	0.80	-2.21	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	4.48
	0.01	0.01	-5.18	0.06	0.06	0.00	0.01	0.00	0.00	0.00	2.67
	0.01	0.01	0.07	-5.26	0.06	0.00	0.01	0.00	0.00	0.00	2.72
	0.02	0.02	0.11	0.09	-3.89	0.01	0.01	0.01	0.00	0.00	2.22
Craft	0.00	0.00	0.03	0.02	0.02	-4.52	0.26	0.26	0.11	0.07	3.94
	0.00	0.00	0.02	0.02	0.02	0.26	-4.53	0.26	0.12	0.08	3.80
	0.00	0.00	0.02	0.02	0.01	0.23	0.26	-4.54	0.11	0.08	3.73
	0.00	0.00	0.02	0.02	0.01	0.23	0.28	0.25	-4.68	0.08	3.92
	0.00	0.00	0.02	0.02	0.01	0.23	0.27	0.25	0.11	-4.66	3.63

5.2 Estimation of Fixed Cost Parameters

The main challenge in our analysis is estimating the fixed cost parameters, which include the parameters in the fixed cost function (i.e., $\theta_0, \theta_1, \theta_2, \theta_3$) and the standard deviation of the fixed cost shock (i.e., σ_ζ). The challenge stems from (1) potential multiple equilibria in product choice and (2) the selection issue that the fixed cost shocks of the observed products and those of the un-chosen products are not from the same distribution, because firms observe the fixed cost shocks before making product choices. To deal with these challenges, we develop a new method that relies on bounds of the probability a product is in the market. In what follows, we present the bounds, explain the estimator, describe the issues in implementation and present the estimation results. We provide additional details in Appendix C.

The construction of the bounds partly relies on the additive-separability of the fixed costs across products, which is a common assumption in the literature of estimating discrete games. In Appendix E, we extend our method for estimating fixed cost functions that allow for (dis-)economies of scope.

5.2.1 Bounds for the Choice Probabilities

To simplify the exposition, and also because we estimate the fixed cost parameters for each year separately, we suppress the subscript τ in the remainder of this section whenever it is clear, and we rewrite the profit function

$$\pi_{nm}\left(\mathcal{J}_{nm}, \mathcal{J}_{-nm}\right) - \sum_{j \in \mathcal{J}_{nm}} \left(\theta_0 + \theta_1 \cdot (\text{in state})_j + \theta_2 \cdot \text{craft}_j + \theta_3 (\text{in state})_j \cdot \text{craft}_j \cdot \text{dist}_{jm} + \zeta_{jm}\right)$$

as

$$\Pi_n(a_{nm}, a_{-nm}, X_m) - \sum_{j \in \mathcal{J}_n} a_{jm}\left[c_{jm}(W_{jm}, \theta) + \zeta_{jm}\right],$$

where the vector X_m includes all relevant demand and marginal cost covariates (including the fixed effects) in market m, while the vector W_{jm} includes all fixed cost covariates. Moreover, we now use a vector of indicators a_{nm} to denote a firm's product portfolio $\mathcal{J}_{nm}$. Specifically, the length of a_{nm} equals the number of potential products that firm n is endowed with (i.e, the size of $\mathcal{J}_n$). The element of a_{nm} that corresponds to product $j \in \mathcal{J}_n$ (denoted by a_{jm}) is 1 if $j \in \mathcal{J}_{nm}$ and 0 otherwise. Furthermore, we use a_{-jm} to denote firm n's decision on its products other than j and the product choices of firm n's rival. We define

$$\Delta_{jn}(a_{-jm}, X_m) = \Pi_n(a_{jm} = 1, a_{-jm}, X_m) - \Pi_n(a_{jm} = 0, a_{-jm}, X_m) \tag{7}$$

to be the incremental change in firm n's expected variable profit when product j is included in its product portfolio, given a_{-jm}. Given the discrete nature of a_{-jm}, the following minimum and maximum exist: $\underline{\Delta}_{jn}(X_m) = \min_{a_{-jm}} \Delta_{jn}(a_{-jm}, X_m)$ and $\overline{\Delta}_{jn}(X_m) = \max_{a_{-jm}} \Delta_{jn}(a_{-jm}, X_m)$. These extrema imply that when $\zeta_{jm} > \overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)$, firm n will always choose $a_{jm} = 0$ regardless of the cost shocks to n's other products, or the actions of its rivals. Similarly, when

$\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)$, firm n will always choose $a_{jm} = 1$.

Using these conditions, we construct the bounds for the probability that $a_{jm} = 1$ is an equilibrium action in Theorem 1 below. The lower bound is the probability that firm n chooses $a_{jm} = 1$ given X_m, W_{jm} and any a_{-jm}. The upper bound is 1 minus the probability that firm n chooses $a_{jm} = 0$ given X_m, W_{jm} and any a_{-jm}. To formally construct the bounds, we first define ζ_m as the vector of fixed cost shocks of all potential products in market m. We define W_m analogously. We use $A_{jm}(X_m, W_m, \zeta_m, \theta) \in \{0, 1\}$ to denote the decision on product j in a complete information, simultaneous move Nash equilibrium.

Theorem 1. *For any market m, firm n and product j of firm n,*

$$\Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right) \leq \Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \mid X_m, W_m\right)$$
$$\leq 1 - \Pr\left(\zeta_{jm} > \overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right).$$

Proof. We first note that

$$\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \,|\, \zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta), X_m, W_m\right) = 1.$$

Therefore, we have

$$\begin{aligned}
&\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \mid X_m, W_m\right) \\
&= \Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right) \cdot \\
&\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \,|\, \zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta), X_m, W_m\right) \\
&+ \Pr\left(\zeta_{jm} > \underline{\Delta}_{j}(X_m) - c_{jm}(W_{jm}, \theta)\right) \cdot \\
&\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \,|\, \zeta_{jm} > \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta), X_m, W_m\right) \\
&\geq \Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right).
\end{aligned}$$

Similarly,

$$\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 0 \mid X_m, W_m\right) \geq \Pr\left(\zeta_{jm} > \overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right),$$

which implies

$$\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \mid X_m, W_m\right) \leq 1 - \Pr\left(\zeta_{jm} > \overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right).$$

□

We note that the inequalities in Theorem 1 hold when there are multiple equilibria and even when the equilibrium selection mechanism is not the same across markets. They also hold when there does not exist a pure strategy equilibrium for some values of ζ_m.

The identification of σ_ζ through these inequalities is similar to the idea of special regressors in entry games (Ciliberto and Tamer (2009); Lewbel (2019)). Relative to the variations of Π_n and c_{jm}, a low σ_ζ in the true data generating process would imply a high correlation between A_{jm} and the observed covariates (X_m, W_m). If the estimated σ_ζ is too large, the upper and lower bounds would not co-vary strongly with the (X_m, W_m) and are likely to be violated. For example, if ζ_{jm} is a symmetric distribution, both bounds in Theorem 1 would approach the constant 0.5 if the estimated σ_ζ is too large. On the other hand, if the estimated σ_ζ is too small, both bounds would simultaneously be close to 0 if $\overline{\Delta}_{jn}(X_m)-c_{jm}(W_{jm},\theta) < 0$, or close to 1 if $\underline{\Delta}_{jn}(X_m)-c_{jm}(W_{jm},\theta) > 0$ by the Chebyshev's inequality, potentially leading to violations of the inequalities in Theorem 1.

5.2.2 Estimation

In the data, for each market m and firm n, we observe the variable profit and fixed cost covariates X_m, W_m and equilibrium product choices $(A_{jm})_{j\in\mathcal{J}_n}$. Theorem 1 implies that, at the true value of the parameters $\left(\theta^0, \sigma_\zeta^0\right)$, for any product $j \in \mathcal{J}_n$

$$\begin{aligned}\left[\Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c(W_{jm}, \theta^0)\right) - \Pr\left(A_{jm} = 1 \mid X_m, W_m\right)\right] F_{jm}(X_m, W_m) \leq 0,\\ \left[1 - \Pr\left(\zeta_{jm} > \overline{\Delta}_{jn}(X_m) - c(W_{jm}, \theta^0)\right) - \Pr\left(A_{jm} = 1 \mid X_m, W_m\right)\right] F_{jm}(X_m, W_m) \geq 0,\end{aligned} \tag{8}$$

for any non-negative function $F_{jm}(\cdot)$. We define

$$L_{jm}\left(A_m, X_m, W_m, \theta, \sigma_\zeta\right) = \Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right) - \Pr\left(A_{jm} = 1 \mid X_m, W_m\right)$$

and

$$H_{jm}\left(A_m, X_m, W_m, \theta, \sigma_\zeta\right) = 1 - \Pr\left(\zeta_{jm} > \bar{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta)\right) - \Pr\left(A_{jnm} = 1 \mid X_m, W_m\right),$$

where A_m is the vector of the equilibrium product choices. The inequalities in (8) become

$$L_{jm}\left(A_m, X_m, W_m, \theta^0, \sigma_\zeta^0\right) F_{jm}\left(X_m, W_m\right) \leq 0,$$

$$H_{jm}\left(A_m, X_m, W_m, \theta^0, \sigma_\zeta^0\right) F_{jm}\left(X_m, W_m\right) \geq 0.$$

The objective function we use for estimation is

$$\begin{aligned}Q\left(\theta, \sigma_\zeta\right) = \frac{1}{\#jm}\sum_{k=1}^{K}\Bigg[\left\|\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j\in\mathcal{J}_n} L_{jm}\left(A_m, X_m, W_m, \theta, \sigma_\zeta\right) F_{jm}^{(k)}\left(X_m, W_m\right)\right\|_+^2 \\ + \left\|\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{j\in\mathcal{J}_n} H_{jm}\left(A_m, X_m, W_m, \theta, \sigma_\zeta\right) F_{jm}^{(k)}\left(X_m, W_m\right)\right\|_-^2\Bigg].\end{aligned} \tag{9}$$

where $\#jm$ indicates the number of observations in a year, where an observation is a combination of a potential product of a firm and a market and k indexes one of the K exogenous functions. In (9), $\left\|L_{jm}F_{jm}^{(k)}\right\|_{+} = L_{jm}F_{jm}^{(k)} \cdot 1\left(L_{jm}F_{jm}^{(k)} > 0\right)$, $\left\|H_{jm}F_{jm}^{(k)}\right\|_{-} = H_{jm}F_{jm}^{(k)} \cdot 1\left(H_{jm}F_{jm}^{(k)} < 0\right)$.

5.2.3 Empirical Implementation

To use the GMM objective function in (9) to carry out inference, we need to

1. define the set of potential products $\mathcal{J}_n$ for each firm n,
2. compute $\underline{\Delta}_{jn}(X_{nm})$ and $\overline{\Delta}_{jn}(X_{nm})$,
3. estimate the conditional choice probability $\Pr\left(A_{jm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) = 1 \mid X_{nm}, W_{nm}\right)$, and
4. construct the non-negative functions $F_{jm}\left(X_{nm}, W_{nm}\right)$.

We address these four things in this subsection. For inference, we use the subsampling procedure in Chernozhukov et al. (2007) (CHT) and the bootstrap procedure in Andrews and Shi (2013) (AS) and report results from both methods. The details on estimation and inference are provided in Appendix C.

First, we define the set of firm n's potential products in year τ, $\mathcal{J}_{n\tau}$, as such that $j \in \mathcal{J}_{n\tau}$ if j is owned by firm n and j is in any market in year τ in our sample. Given our focus on craft products and the very small cross-elasticity between the craft and other products (Table 2), we construct our objective function with craft products plus the non-craft products by Blue Moon (owned by MillerCoors) and Shock Top (owned by ABI).[15] Despite not being designated as craft by the Brewers Association, the packaging and marketing of Blue Moon brands are allegedly similar to craft products (Field (2019)), and ABI introduced Shock Top brands to compete with Blue Moon (Shears (2014)). Including these products allows us to estimate the parameters involving the indicator of whether a product is craft. Finally, we focus on markets where at least one craft product was observed. These restrictions still leave us with a large number of inequalities for estimation. In 2016, there are 95 potential products, 149 markets and 14,155 product-market combinations. To ease the computational burden, we randomly sample 8000 observations (product-market combinations) for each year and estimate the fixed costs separately.

Next, we discuss how to compute $\underline{\Delta}_{jn}(X_m)$ and $\overline{\Delta}_{jn}(X_m)$. Directly solving for the minimum and the maximum of the expected profits over all possible values of a_{-jm} is computationally costly, because there are $2^{\text{length of } a_{-jm}}$ possible values of a_{-jm}. and computing $\Delta_{jn}(a_{-jm}, X_m)$ for each a_{-jm} involves solving stage-2 price games for multiple simulated draws of demand and marginal cost shocks. However, economic intuition suggests that because products are substitutes, we can approximate the minimum by

$$\underline{\Delta}_{jn}\left(X_{nm}\right) \approx \Delta_{jn}\left((1, \ldots, 1), X_m\right)$$

[15] In our sample period, Goose Island and a few other craft products were acquired by large breweries such as ABI, and the Brewers Association removed their "craft" designations. We continue to designate these products as craft but assume their acquirers set their prices. The underlying assumption is that consumers base purchase decisions on tastes: the craft beers are still produced from the same facilities and we assume the chemical compositions remain the same as the pre-merger products.

and the maximum by

$$\overline{\Delta}_{jn}\left(X_{nm}\right) \approx \Delta_{jn}\left(\left(0, ..., 0\right), X_m\right).$$

These approximate extrema are exact for the model in Ciliberto and Tamer (2009), where the variable profit function is a linear function of the entry decisions of rivals and when the competitive effects are negative. For more general demand and pricing models such as ours, we conduct Monte Carlo simulations and find the approximate extrema also coincide with the true ones across a variety of parameter specifications (see Appendix H).

To calculate $\Delta_{jn}\left(\left(\cdot\right), X_m\right)$ for each potential product in each market, we solve for the corresponding stage-2 pricing equilibrium based on the estimated demand and marginal costs. For the demand and marginal cost estimation, we try to be general and include the product-market fixed effects ξ_{jm} and ω_{jm} to capture systematic variations at the product-market levels. However, this specification does not directly give us ξ_{jm} or ω_{jm} when j is not observed in market m. Therefore, at this stage of estimation, we impose more structures and decomposes ξ_{jm} into a product fixed effect, a market fixed effect and a linear combination of market-product specific variables, which are indicators based on the distance from j's brewery to the market m following the motivation in Section 3. Appendix B provides the detailed estimation procedure and evidence that our approach has reasonably good in-sample and out-of-sample fit.

We use a nearest-neighbor approach to estimate the empirical distribution $\Pr\left(A_{jnm} = 1 \mid X_m, W_m\right)$ in (8). Appendix C provides details on this procedure. By using empirical estimates of conditional choice probabilities and conditional moment inequalities, our method bears some semblance to the CCP approach in the estimation of dynamic models (Bajari et al. (2007)).

We construct the exogenous function $F_{jm}(X_m, W_m)$ as follows.[16] For each observation jm, we consider the following variables: $\underline{\Delta}_{jn}(X_m)$, $\overline{\Delta}_{jn}(X_m)$ and W_{jm}. Some of these variables are continuous and the rest are binary. We use Z^c_{jm} and Z^b_{jm} to denote the continuous and binary sub–vectors. Following the construction of instruments in Andrews and Shi (2013), we standardize Z^c_{jm} and transform the vector so that each element lies in $(0, 1)$. Specifically, the transformed vector is $\tilde{Z}^c_{jm} = \Phi\left(\mathit{\Gamma} \cdot Z^c_{jm}\right)$, where $\mathit{\Gamma}$ is the Cholesky decomposition of $Var^{-1}\left(Z^c_{jm}\right)$ and $\Phi\left(\cdot\right)$ is the standard normal distribution applied to each element of the vector. We then divide the set $(0, 1)^{\left|\tilde{Z}^c_{jm}\right|} \times \{0, 1\}^{\left|Z^b_{jm}\right|}$ (where $\left|\tilde{Z}^c_{jm}\right|$ and $\left|Z^b_{jm}\right|$ denote the dimensions of $\tilde{Z}^c_{jm}$ and Z^b_{jm}) into $(2r)^{\left|\tilde{Z}^c_{jm}\right|} \cdot 2^{\left|Z^b_{jm}\right|}$ hypercubes, where each interval $(0, 1)$ is divided into $\left(0, \frac{1}{2r}\right), \left[\frac{1}{2r}, \frac{2}{2r}\right), \ldots \left[\frac{2r-1}{2r}, 1\right)$.[17] We denote the set of hypercubes by $\boldsymbol{\mathcal{C}}$. We construct $F_{jm}(X_m, W_m)$ as indicator functions $1\left(\left(\tilde{Z}^c_m, Z^b_m\right) \in \mathcal{C}\right)$ for each element $\mathcal{C} \in \boldsymbol{\mathcal{C}}$.

[16] AS objective function (a Cramér–von Mises statistic) is slightly different and uses a weighted sum of the objective functions in (9) where L and H functions are interacted with different sets of F functions.

[17] We report the results with $r = 3$. The results are also robust for $r = 2$ or 4.

Table 3: Estimates of Fixed Costs: Projected 95% Confidence Intervals, 2016

	CHT	AS
Constant (θ_0)	[2075.12, 3139.05]	[2250.49, 4336.24]
Craft (θ_1)	[1102.09, 2076.87]	[616.18, 2932.42]
In State× Craft (θ_2)	[-1173.51, 437.03]	[-2118.48, 627.42]
Distance^a×In State× Craft (θ_3)	[1216.18, 3451.17]	[770.51, 5529.83]
Std. Dev. (σ_ς)	[1313.70, 2303.08]	[1302.35, 4059.56]

^{a}: Distance in 1000KM

5.2.4 Estimation Results

We report the estimates and the 95% confidence set projected to each parameter from 2016 data in Table 3. All estimates are in 2016 US dollars. The CIs of CHT estimates are slightly shifted and narrower compared with the AS estimates. The 95% CI of the fixed cost intercept θ_0 based on CHT ranges from \$2075 to \$3139, and the CI based on AS ranges from \$2230 to \$4336. To put the figures in context, we note that the change in the expected variable profit $\Delta_{jn}(a_{-jm}, X_m)$ in the estimation sample has a mean of 3,755 dollars and a standard deviation of 12,925 dollars in 2016. The intercept θ_0 is the mean fixed cost of a Blue Moon or Shock Top brand. The mean fixed cost of an out-of-state craft brewery, which in our sample is typically one of the largest craft brewery in the nation, is $\theta_0 + \theta_1$. The estimates show that even these breweries face additional costs to distribute a product than Blue Moon or Shock Top. We interpret the estimates of θ_2 as the difference between in-state and out-of-state craft beer fixed costs, which have a relatively wide CI. The distance parameter θ_3 shows that the in-state craft breweries face additional costs that increase in distance from the brewery to the market. The estimates also indicate sizable variance of the unobserved fixed cost shock. The estimate of the standard deviation of the unobserved shocks is of comparable magnitude to the fixed cost intercept θ_0.

6 Counterfactual Results

6.1 Counterfactual Simulations

Using the estimated model, we consider a counterfactual merger where the largest firm (a "macro" brewery) in our sample acquires 3 largest craft firms (excluding Boston Beer Company and Sierra Nevada Brewing, which are unlikely merger targets given their sizes) in 2016. We simulate the merger in 149 markets where at least one craft product is observed in data. Through the acquisition, the large brewery would have acquired more than 50% of the craft beer shares in many markets. The merger may raise prices, providing incentives for repositioning and new firm entry. We allow firms to change their craft products, and we allow craft breweries to enter or exit. We hold the non-craft product choices fixed as observed in data to ease the computation burden, but we allow their prices to adjust. The simplification is justified by the estimated small substitution between the craft and non-craft products.

There are three types of shocks in the models: demand, marginal cost and fixed cost shocks. We draw demand and marginal cost shocks directly from their estimated distributions (Appendix B) to calculate the expected variable profits given product choices (π_{nm} in (4)). To make the pre- and post-merger outcomes comparable, we use fixed cost draws consistent with the observed pre-merger equilibrium by drawing from the estimated distribution while taking into account selection (details can be found in Appendix F). We compute the post-merger product equilibrium using the algorithm in Fan and Yang (2020). In the simulation, a decision maker (or a potential entrant) is a firm that is observed in any market in our sample. Each firm is endowed with a set of potential products, consisting of the firm's craft products observed in the sample. In each market, a firm chooses a subset from its potential products. Choosing a non-empty subset implies entry into the market. The merging firms choose products from the union of the sets of their potential products. Additional computational and simulation details are in Appendix G.

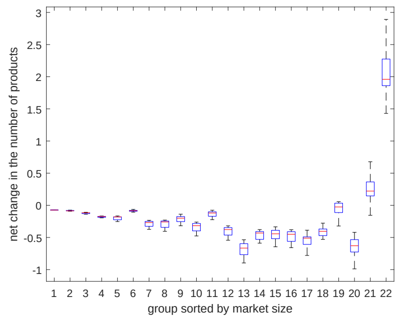
6.2 Counterfactual Results for All Simulated Markets

Overall, the effects of the merger are highly heterogeneous across markets. Figure 3 presents the effects of the merger on the number of products from 149 markets. We first sort markets by market sizes, and group them from small to large in 22 approximately equal sized groups. We plot various merger outcomes against for each group. Specifically, we sample 15 vectors from the estimated confidence set (based on the AS approach) and simulate the merger outcomes using 20 simulated fixed cost draws for each market. Then we calculate the expected outcomes across fixed cost draws for each parameter and market, and within each group, we average again across markets, weighted by market sizes. The end points of the broken lines indicate the range of the outcomes across the sampled parameters. The box indicates the 25–75% quantiles. In Figure 3 (a), we show that the number of products tends to increase in large markets and decrease in smaller markets, with the most significant decline in the medium-sized markets. We decompose this change in product variety attributed to entry, repositioning by non-merging incumbents and repositioning by merging firms in Panels (b)–(e). The number of added products by entrants conditional on entry for most markets is fewer than 0.5. New entrants add more products in the largest markets, and the pattern of firm entry is similar. The non-merging incumbents add fewer products. The merging firms drop products, and the largest decrease occurs in medium sized markets. On average, craft beer prices increase by 15 cents.

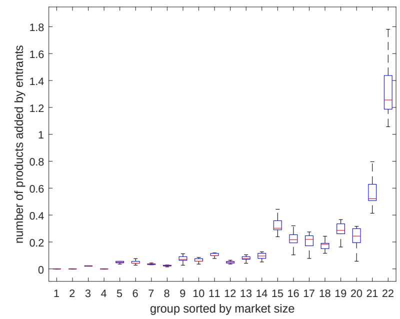
The heterogeneity of the merger's effect on product variety is associated with heterogeneity in market characteristics. Intuitively, if the products of the merging firm are close substitutes, the merged firm is more likely to drop some products and increase prices significantly. At the same time, pronounced price increases attract firm entry and product entry by non-merging firms. Therefore, a measure of the merging firms' market power may be positively correlated with both the number of added products and the number of dropped products, although the sign is less clear with the net change of the number of products. Furthermore, high fixed costs should be associated with a low number of products. In addition, everything else equal, we should expect more entry

Figure 3: Product Variety, Entry and Prices

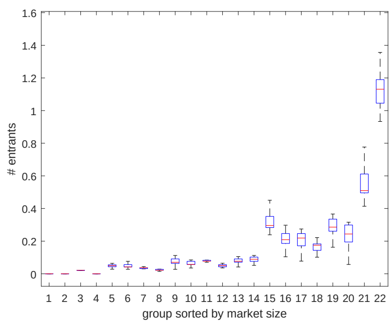
(a) Net Change in the Number of Products



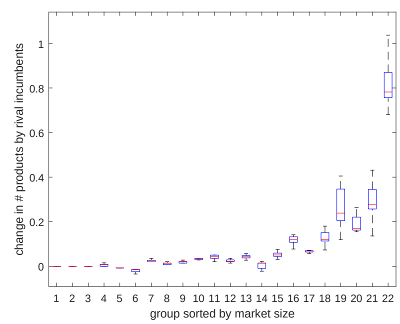
(b) Number of Products Added by Entrants



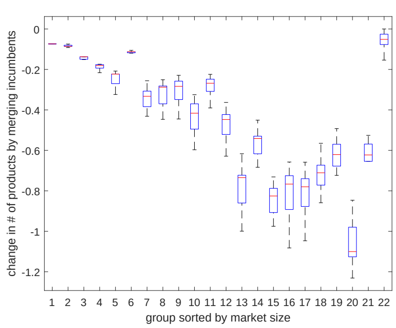
(e) Number of Entrants



(c) Change in the Number of Products by Incumbent Non-merging Firms



(d) Change in the Number of Products by Merging Firms



(f) Craft Prices

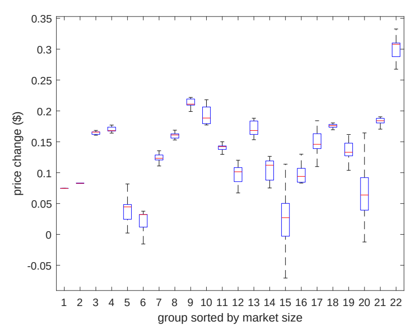
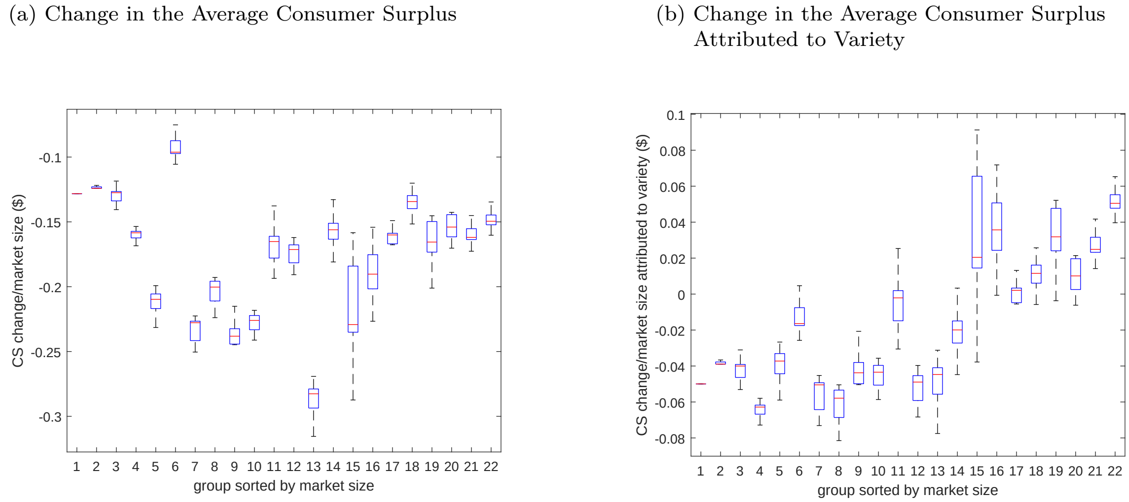


Figure 4: Average Consumer Surplus



in larger markets. In Appendix D, we use regression analysis to document these correlations. We find that market power, measured by the average price increase of the merging firms' products in the price-adjustment-only counterfactual simulation, is positively correlated with both product exit by the merging firms and product entry by new firms or non-merging incumbents, and negatively correlated with the net change in the number of products. We also find that fixed costs are correlated with a reduced number of products and that market sizes are positively correlated with product entry.

We find that the welfare changes are consistent with the product changes. In Panel (a) of Figure 4, we show the changes of average consumer surplus, which is defined as the change of consumer surplus in a market divided by its total annual alcohol sales. The average consumer surplus ranges from close to 0 to -0.3 dollars. Large declines in the average welfare correspond with markets where the net change of products is negative. In Panel (b), we break out the average consumer welfare changes attributed to the variety changes (due to entry and repositioning), where we first calculate the consumer surplus change when we allow firms to change prices but not product variety, and then subtract this quantity from the consumer surplus change when we allow both prices and variety to change (Panel (a)). The variety changes recover some consumer surplus loss in the top 8 groups of markets for all sampled parameters. In the rest of the markets, there exist some parameters in the confidence set such that the variety changes further increase the loss of consumer surplus from higher prices. In 10 of these markets, the variety changes decrease consumer welfare for all sampled parameters.

In Appendix I, we conduct two additional sets of simulations. In Appendix I, we allow the mean and the standard deviation of the fixed costs to differ across market sizes. In Appendix E, we estimate a fixed cost function that allows for (dis–) economies of scale. Our results are robust.

7 Conclusion

We develop a new method to estimate multiple-discrete choice games where firms decide on entry and product choices. We apply the method to study the effects of merger on firm entry, product choices and prices in the retail craft beer market in California. We make two contributions. On the methodological front, our method combines the ideas of conditional choice probabilities and conditional moment inequalities. By estimating bounds for a single action's probability instead of the probability of a game's outcome, our estimator is easy to compute and scalable to large games (with many firms or a large action set for each firm). On the substantive front, the paper adds to the literature of merger and endogenous product choices. For the retail craft beer market in California, we find significant heterogeneity in merger outcomes across markets. Large markets are likely to see sufficient post-merger product entry and firm entry that curb the consumer surplus loss from increased prices, but in a majority of markets, the net number of products decreases, which tends to further reduce consumer welfare. Aggregated across markets, the effects of entry and repositioning are positive but small relative to the total loss of consumer surplus.

References

Anderson, Simon P, Nisvan Erkal, and Daniel Piccinin, "Aggregative games and oligopoly theory: short-run and long-run analysis," *The RAND Journal of Economics*, 2020, *51* (2), 470–495.

Andrews, Donald WK and Xiaoxia Shi, "Inference based on conditional moment inequalities," *Econometrica*, 2013, *81* (2), 609–666.

Anhalt, Brian D, "Crafting a Model State Law for Today's Beer Industry," *Roger Williams UL Rev.*, 2016, *21*, 162.

Ashenfelter, Orley C, Daniel S Hosken, and Matthew C Weinberg, "Efficiencies brewed: pricing and consolidation in the US beer industry," *The RAND Journal of Economics*, 2015, *46* (2), 328–361.

Asker, John, "Diagnosing foreclosure due to exclusive dealing," *The Journal of Industrial Economics*, 2016, *64* (3), 375–410.

Bajari, Patrick, C Lanier Benkard, and Jonathan Levin, "Estimating dynamic models of imperfect competition," *Econometrica*, 2007, *75* (5), 1331–1370.

Berry, Steven and Phil Haile, "Nonparametric Identification of Differentiated Products Demand Using Micro Data," 2020.

_ , James Levinsohn, and Ariel Pakes, "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, 1995, pp. 841–890.

Berry, Steven T, "Estimation of a Model of Entry in the Airline Industry," *Econometrica: Journal of the Econometric Society*, 1992, pp. 889–917.

_ **and Joel Waldfogel**, "Do mergers increase product variety? Evidence from radio broadcasting," *The Quarterly Journal of Economics*, 2001, *116* (3), 1009–1025.

Cabral, Luıs MB, "Horizontal mergers with free-entry: why cost efficiencies may be a weak defense and asset sales a poor remedy," *International Journal of Industrial Organization*, 2003, *21* (5), 607–623.

Caradonna, Peter, Nathan Miller, and Gloria Sheu, "Mergers, Entry, and Consumer Welfare," *Available at SSRN 3537135*, 2020.

Chernozhukov, Victor, Han Hong, and Elie Tamer, "Estimation and confidence regions for parameter sets in econometric models," *Econometrica*, 2007, *75* (5), 1243–1284.

Ciliberto, Federico and Elie Tamer, "Market structure and multiple equilibria in airline markets," *Econometrica*, 2009, *77* (6), 1791–1828.

_ **, Charles Murry, and Elie T Tamer**, "Market structure and competition in airline markets," *Available at SSRN 2777820*, 2018.

Codog, Aric, "The Antitrust Roadblock: Preventing Consolidation of the Craft Beer Market," *McGeorge L. Rev.*, 2018, *50*, 403.

Croxall, Daniel J, "Helping Craft Beer Maintain and Grow Market Shares with Private Enforcement of Tied-House and False Advertising Laws," *Gonz. L. Rev.*, 2019, *55*, 167.

Eizenberg, Alon, "Upstream innovation and product variety in the us home pc market," *Review of Economic Studies*, 2014, *81* (3), 1003–1045.

Fan, Ying, "Ownership consolidation and product characteristics: A study of the US daily newspaper market," *American Economic Review*, 2013, *103* (5), 1598–1628.

_ **and Chenyu Yang**, "Competition, product proliferation and welfare: A study of the us smartphone market," *American Economic Journal: Microeconomics*, 2020.

Field, Emily, "Blue Moon Looks Like A Craft Beer, But Is Not, Court Told," *https://www.law360.com/articles/747836*, 2019.

Gandhi, Amit, Luke Froeb, Steven Tschantz, and Gregory J Werden, "Post-Merger Product Repositioning," *The Journal of Industrial Economics*, 2008, *56* (1), 49–67.

Garrido, Francisco Andres, "Mergers between Multi-Product Firms with Endogenous Variety: Theory and an Application to the Ready-To-Eat Cereal Industry." PhD dissertation, Georgetown University 2020.

Hart, Jarrett and Julian M Alston, "Persistent patterns in the US alcohol market: Looking at the link between demographics and drinking," *Journal of Wine Economics*, 2019, *14* (4), 356–364.

Ho, Katherine, "Insurer-provider networks in the medical care market," *American Economic Review*, 2009, *99* (1), 393–430.

Hotz, V Joseph and Robert A Miller, "Conditional choice probabilities and the estimation of dynamic models," *The Review of Economic Studies*, 1993, *60* (3), 497–529.

Huang, Yufeng, Paul B Ellickson, and Mitchell J Lovett, "Learning to Set Prices," 2020.

Illanes, Gastón, "Switching costs in pension plan choice," *Unpublished manuscript*, 2017.

_ **and Sarah Moshary**, "Market Structure and Product Assortment: Evidence from a Natural Experiment in Liquor Licensure," Technical Report, National Bureau of Economic Research 2020.

Jeziorski, Przemysław, "Empirical Model of Dynamic Merger Enforcement–Choosing Ownership Caps in US Radio," 2015.

Lewbel, Arthur, "The identification zoo: Meanings of identification in econometrics," *Journal of Economic Literature*, 2019, *57* (4), 835–903.

Li, Sophia, Joe Mazur, Yongjoon Park, James Roberts, Andrew Sweeting, and Jun Zhang, "Repositioning and market power after airline mergers," 2019.

McFadden, Daniel, "A method of simulated moments for estimation of discrete response models without numerical integration," *Econometrica: Journal of the Econometric Society*, 1989, pp. 995–1026.

Miller, Nathan, Gloria Sheu, and Matthew Weinberg, "Oligopolistic price leadership and mergers: The united states beer industry," *Available at SSRN 3239248*, 2019.

Miller, Nathan H and Matthew C Weinberg, "Understanding the price effects of the Miller-Coors joint venture," *Econometrica*, 2017, *85* (6), 1763–1791.

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

Ohlhausen, Maureen K, Michael A Salinger, and Susan A Creighton, "Re: Comment on Proposed Beer Francise Act," 2005, *https://www.ftc.gov/sites/default/files/documents/advocacy_documents/ftc-staff-comment-honorable-wesley-chesbro-concerning-proposed-california-franchise-act-govern/050826beerfranchiseact.pdf.*

Pakes, Ariel, Jack Porter, Kate Ho, and Joy Ishii, "Moment inequalities and their application," *Econometrica*, 2015, *83* (1), 315–334.

Reiss, Peter C and Pablo T Spiller, "Competition and entry in small airline markets," *The Journal of Law and Economics*, 1989, *32* (2, Part 2), S179–S202.

Ryan, Stephen P, "The costs of environmental regulation in a concentrated industry," *Econometrica*, 2012, *80* (3), 1019–1061.

Sayyed, Bilal K., Andrew Sweeting, Ian Conner, David Lawrence, Eric D Welsh, and Elizabeth Armington, *https://www.ftc.gov/system/files/documents/advocacy_documents/joint-comment-ftc-staff-doj-antitrust-division-staff-california-state-assembly-concerning-california/v200008_california_beer_distribution_advocacy_2020.pdf*, 2020.

Schennach, Susanne M, "Entropic latent variable integration via simulation," *Econometrica*, 2014, *82* (1), 345–385.

Shears, Andrew, "Local to national and back again: beer, wisconsin & scale," in "The geography of beer," Springer, 2014, pp. 45–56.

Spector, David, "Horizontal mergers, entry, and efficiency defences," *International Journal of Industrial Organization*, 2003, *21* (10), 1591–1600.

Sweeting, Andrew, "The effects of mergers on product positioning: evidence from the music radio industry," *The RAND Journal of Economics*, 2010, *41* (2), 372–397.

Tamayo, Andrew, "What's Brewing in the Old North State: An Analysis of the Beer Distribution Laws Regulating North Carolina's Craft Breweries," *NCL Rev.*, 2009, *88*, 2198.

Tamer, Elie, "Incomplete simultaneous discrete response model with multiple equilibria," *The Review of Economic Studies*, 2003, *70* (1), 147–165.

Watson, Bart, "On-Premise Beer Data and Craft," *https://www.brewersassociation.org/insights/importance-on-premise-craft-brewers/*, 2016.

Werden, Gregory J and Luke M Froeb, "The entry-inducing effects of horizontal mergers: An exploratory analysis," *The Journal of Industrial Economics*, 1998, *46* (4), 525–543.

Wollmann, Thomas G, "Trucks without bailouts: Equilibrium product characteristics for commercial vehicles," *American Economic Review*, 2018, *108* (6), 1364–1406.

A Details on Micro-Moments

Our first moments are the average annual purchase of a household of a certain type of beer conditional on ever purchasing this type of beer in the year, i.e., $E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$, where i is a household that has purchased beer at least once, q_{it}^f is household i's total quantity of beer with a certain flavor (f =lager or f = light) or of a certain characteristic (f =Import or f =craft) in month t. To save notation, we denote the market share by

$$s_{jmt} = \int \tilde{s}_{jmt}\left(\boldsymbol{\nu}, y\right) dG_m\left(\boldsymbol{\nu}, y\right),$$

where $\tilde{s}_{jmt}\left(\boldsymbol{\nu}, y\right)$ is the logit choice probability of j when the vector of unobserved tastes and income are $(\boldsymbol{\nu}, y)$. The distribution G_m of the unobserved preferences and income can vary across markets and is indexed by m. To calculate the first moments $E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$, we first compute the average household purchase of type f conditional on purchasing f in market m in year. We assume that each consumer has $N = 8$ opportunities per month to buy beer (where on each trip the consumer buys 1 or 0 products), the average number of trips to the stores in Nielsen data per household.[18] We first define $\tilde{\rho}_{mt}^f\left(\boldsymbol{\nu}, y\right)$ as the probability of purchasing at least one beer of type f conditional on $(\boldsymbol{\nu}, y)$:

$$\tilde{\rho}_m^f\left(\nu, y\right) = 1 - \prod_t \left(1 - \sum_{j \text{ s.t. } x_j^f=1} \tilde{s}_{jmt}\left(\boldsymbol{\nu}, y\right)\right)^N .$$

The expected total quantity of purchase conditional buying f at least once is

$$E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0, m\right) = \int_{\boldsymbol{\nu}, y} \frac{N \sum_{\text{ s.t. } x_j^f=1} \tilde{s}_{jmt}\left(\boldsymbol{\nu}, y\right)}{\tilde{\rho}_m^f\left(\nu, y\right)} dG_m\left(\boldsymbol{\nu}, y\right).$$

where the summation $\sum_{\text{ s.t. } x_j^f=1}$ is the sum over all products in the choice set in market m, month t where the product is of type f, and all trips. To obtain the average across markets, we weigh these conditional means by the expected number of households who purchases products of type f at least once. We define the unconditional probability of purchasing at least once beer of type f as

$$\rho_m^f = \int \tilde{\rho}_m^f\left(\nu, y\right) dG_m\left(\boldsymbol{\nu}, y\right),$$

which implies the total number of households that purchase at least one product of type f in market m is $O_m \cdot \rho_m^f$, where O_m is the market size of m. Therefore the expected purchase of type f conditional on having at least one purchase of type f is

$$E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right) = \frac{1}{\sum_m O_m \cdot \rho_m^f} \cdot \left(\sum_m E_i\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0, m\right) \cdot O_m \cdot \rho_m^f\right),$$

For the second moments, the average annual purchase by a household of beer conditional on purchasing beer in the year, we can replace $\sum_{j \text{ s.t. } x_j^f=1}$ in the above with $\sum_j$, the summation of

[18] The demand estimates are robust to $N = 5, 6, 7$.

all products in the choice set.

B Details on Computing Δ_{jn}

This section discusses how to compute the incremental variable profit $\Delta_{jn}(a_{-jm}, X_{nm})$ in (7), which is the change in variable profits of firm n when product j is added to market m. Given a_{-jm} and X_{nm}, one can compute

$$\Delta_{jn}(a_{-jm}, X_{nm}) = \pi_{nm}(\mathcal{J}_{nm\tau} \cup j, \mathcal{J}_{-nm\tau}) - \pi_{nm}(\mathcal{J}_{nm\tau} \cup j, \mathcal{J}_{-nm\tau})$$

using (5), where $(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau})$ is the set of products in the market corresponding with $(a_{jm} = 0, a_{-jm})$. We therefore focus on how to estimate X_{nm}, the vector of all relevant demand and marginal cost covariates. The vector consists of the following components:

$$X_{nm} = \left\{ \left(\xi_{\tilde{j}m}, \omega_{\tilde{j}m}, \boldsymbol{x}_{\tilde{j}} \right)_{\tilde{j}: a_{\tilde{j}m}=1}, G_m(\boldsymbol{\nu}, y) \right\},$$

where $G_m(\boldsymbol{\nu}, y)$ is the market specific distribution of unobserved tastes and income. The observables $\boldsymbol{x}_{\tilde{j}}$ are given in data, and $G_m(\boldsymbol{\nu}, y)$ given by the demand estimates. The rest of the section concerns the estimation of the demand and marginal cost unobservables ξ and ω.

In the demand estimation, we estimate ξ_t as parameters, and the inversion following Berry et al. (1995) allows us to estimate $\hat{\xi}_{jm}$ for observed product-market combination jm. To estimate $\Delta_{jn}(a_{-jm}, X_{nm})$ such that j is not observed in m, we further parameterize $\xi_{jm} = \xi_j + \xi_m + \gamma Z_{jm}$, where γ is a vector of parameters and Z_{jm} is a vector of product-market specific observables. In practice, we construct Z_{jm} as indicators of whether the distance is in the bins of 0–50KM, 50–150KM, 150–300KM, 300–600KM, 600–1200KM, and greater than 1200KM. We define μ_{jmt} as the inverted mean utility net of the time fixed effects from the Berry inversion

$$\mu_{jmt} = \xi_{jm} + \xi_{jmt}.$$

We then estimate (ξ_j, ξ_m, γ) as parameters from the regression

$$\mu_{jmt} = \xi_j + \xi_m + \gamma Z_{jm} + \xi_{jmt},$$

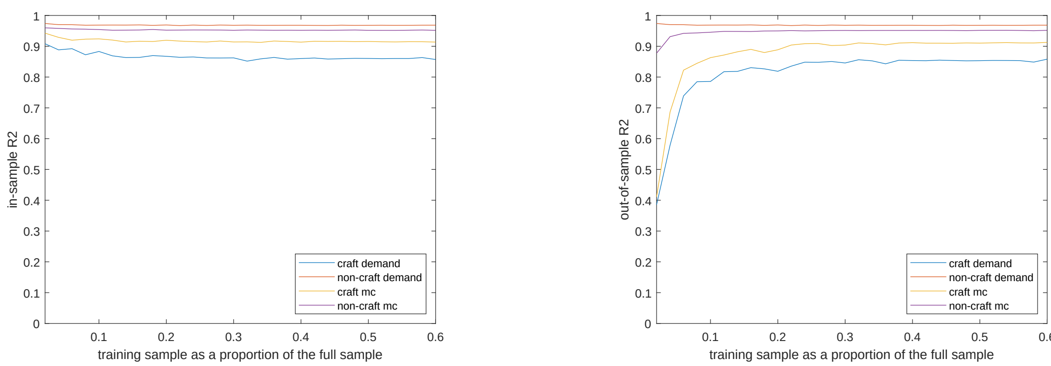
and construct ξ_{jm} from the parameter estimates $\hat{\xi}_j + \hat{\xi}_m + \hat{\gamma} Z_{jm}$. We can similarly construct ω_{jm}.

To simulate the expected profit in (7), we use the empirical distribution of the residuals $\left(\hat{\xi}_{jmt}, \hat{\omega}_{jmt} \right)$ from the estimating equations above. The identification of price coefficients relies on the assumption that product choices are uncorrelated with the transient shocks $(\xi_{jmt}, \omega_{jmt})$. This assumption also is sufficient for the consistency of the procedures above.

We next show that $(\xi_{jmt}, \omega_{jmt})$ accounts for a small proportion of the variations in mean utility and marginal costs. Specifically, we examine the in-sample and out-of-sample fits. We randomly sample a percentage ρ of the entire demand data (at the product-market-month level). This is

Figure B.1: In- and Out-of-Sample Fit for μ_{jmt} and mc_{jmt}

(a) In-sample R^2 (b) Out-of-sample R^2



our training sample. We estimate the regression above, and then calculate the R^2 as a measure of in-sample fit. We then use the estimates to predict μ_{jmt} and mc_{jmt} on the other $1-\rho$ sample. For example, the prediction of μ_{jmt} is

$$\hat{\xi}_j + \hat{\xi}_m + \hat{\gamma} Z_{jm},$$

where we assume the predicted value of the transient shock ξ_{jmt} is 0. If the prediction involves a j or m whose ξ_j, ξ_m values are not estimated, we also set these values to 0. The prediction error is

$$\mu_{jmt} - \left(\hat{\xi}_j + \hat{\xi}_m + \hat{\gamma} Z_{jm}\right).$$

We calculate an "out-of-sample" R^2, which is 1 minus the variance of the prediction errors divided by the variance of the corresponding outcome variables.

We estimate the models for craft and non-craft separately, allowing for different estimates of ξ_m and γ for craft and non-craft products. The craft products account for 31424 of 111219 observations in 2016. Using the 2016 estimates, we plot the in- and out-of- sample fit of μ_{jmt} and mc_{jmt}, where the size of the training data is a proportion of ρ of the full sample, in Figure B.1 (a) and (b) as ρ varies from 0.01 to 0.6. The in-sample R^2 shows that the transient shocks account for no more than 14% of the variations in μ_{jmt} and less than 10% for mc_{jmt} of craft products, and no more than 5% for non-craft products. The out-of-sample performance is similar: with the training sample including as little as 20% of the data, the out-of-sample R^2 becomes comparable to the in-sample ones. The reasonably good fit provides us with further justification to use this specification to construct the mean utility and marginal costs in X_{nm}.

C Details on the Fixed Cost Estimation and Inference

In this section, we give step-by-step guides on the estimation procedures, including how we estimate the conditional choice probability $\Pr\left(A_{jm} = 1 \,|\, X_m, W_m\right)$ and carry out the inference of the fixed cost parameters. We estimate the confidence set in two ways, based on Chernozhukov et al. (2007) (CHT) and Andrews and Shi (2013) (AS).

1 Index product/market combinations by jm, and calculate the corresponding extrema $\underline{\Delta}_{jn}(X_{nm})$ and $\overline{\Delta}_{jn}(X_{nm})$ (Appendix B).

2 For each jm, estimate $\Pr(A_{jm}=1 \mid X_m, W_m)$, where (X_{nm}, W_{nm}) is a potentially high-dimensional vector. Similar to the conditional choice probability approach for estimating dynamic models (Hotz and Miller (1993); Bajari et al. (2007); Ryan (2012)), we assume that it is sufficient to condition on a few summary statistics. Specifically, we consider $\left(\overline{\Delta}_{jn}(X_m), \underline{\Delta}_{jn}(X_m), W_{jm}\right)$. Some of these variables are continuous and the rest are binary. We use Z^c_{jm} to denote the continuous subvector as Z^b_{jm} and the binary vector as Z^b_{jm}. Following the construction of instruments in Andrews and Shi (2013), we standardize Z^c_{jm} and transform the vector so that each element lies in $[0,1]$. Specifically, the transformed vector is $\tilde{Z}^c_{jm} = \Phi\left(\Gamma \cdot Z^c_{jm}\right)$, where Γ is the Cholesky decomposition of $Var^{-1}\left(Z^c_{jm}\right)$ and $\Phi(\cdot)$ is the standard normal distribution applied to each element of the vector. For each jm, we estimate $\Pr(A_{jm}=1 \mid X_{nm}, W_{nm})$ by the average of A_{jm} across the closest 20 observations measured by Euclidean distance in $\left(\tilde{Z}^c_{jm}, Z^b_{jm}\right)$.

3(CHT) Estimate $(\hat{\theta}, \hat{\sigma}_\zeta)$ by minimizing the objective function Q.[19]

4(CHT) For inference, we use subsampling to take into account the noise introduced in estimating the choice probability in Step 2. We use 20% of the data to construct 200 subsamples. For each subsample, we re-estimate the choice probabilities in Step 2 and re-compute the objective function in Step 3 at $\left(\hat{\theta}, \hat{\sigma}_\zeta\right)$. We obtain an initial estimate of critical value c_1 as the 95% quantile of $Q^r\left(\hat{\theta}, \hat{\sigma}_\zeta\right)$, where Q^r is the objective function on the rth sample.

5(CHT) Construct a set of (θ, σ) such that $Q(\theta, \sigma) < c_1$. We first obtain a set Θ of 8000 candidate parameters by adding shocks to $\left(\hat{\theta}, \hat{\sigma}_\zeta\right)$. We then discard parameters such that $Q(\theta, \sigma) > c_1$.

6(CHT) For each vector of remaining parameters (θ, σ), we obtain a critical value $c_1(\theta, \sigma)$ as the 95% quantile of $Q^r(\theta, \sigma)$ across $r = 1, \ldots, 200$. Define $c_2 = \max_{(\theta,\sigma) \text{ s.t. } Q(\theta,\sigma)<c_1} c_1(\theta, \sigma)$. We construct the confidence set as $\{(\theta, \sigma) : Q(\theta, \sigma) < c_2\}$ by discarding parameters in Θ such that $Q(\theta, \sigma) > c_2$.

3(AS) Use $F^{(k),r}_{jm}$ to denote the non-negative functions defined in Section 5.2.3, where r is the integer that controls the fineness of the grid. For a candidate vector of parameters

[19] In practice, to deal with over-identification, we use $Q^r(\theta, \sigma) - \min_{(t_1,t_2)} Q^r(t_1, t_2)$ as the objective function (Chernozhukov et al. (2007)). The same applies to the calculation of the objective functions on the subsamples in step 4.

(θ, σ_ζ), define the test statistic T in Andrews and Shi (2013):

$$\begin{aligned} T(\theta, \sigma_\zeta) = \sum_{r=1}^{3} \frac{(2r)^{-d_x}}{(r^2+100)} \cdot (\#jm) \cdot \\ \sum_{k=1}^{K} \Bigg[\left\| \frac{\varpi_{L,kr}^{-1}}{\#jm} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{j \in \mathcal{J}_n} L_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta) F_{jm}^{(k),r}(X_m, W_m) \right\|_+^2 \\ + \left\| \frac{\varpi_{H,kr}^{-1}}{\#jm} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{j \in \mathcal{J}_n} H_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta) F_{jm}^{(k),r}(X_m, W_m) \right\|_-^2 \Bigg]. \end{aligned}$$

We choose the conditioning weights $\frac{(2r)^{-d_x}}{(r^2+100)}$, where d_x is the dimension of conditioned observables, and other tuning parameters following AS. The weights $\varpi_{L,kr}$ and $\varpi_{H,kr}$ are standard deviations of the terms

$$m_{kr}^L = \frac{1}{\#jm} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{j \in \mathcal{J}_n} L_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta) F_{jm}^{(k),r}(X_m, W_m)$$

and

$$m_{kr}^H = \frac{1}{\#jm} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{j \in \mathcal{J}_n} H_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta) F_{jm}^{(k),r}(X_m, W_m).$$

This CvM test statistic is a weighted sum of the objective functions in Step 3, with each objective function using a finer definition of the box instruments.

4(AS) Generate bootstrap samples and compute the "re-centered" statistic on each sample

$$\begin{aligned} T^\star(\theta, \sigma_\zeta) = \sum_{r=1}^{3} \frac{(2r)^{-d_x}}{(r^2+100)} \cdot (\#jm) \cdot \\ \sum_{k=1}^{K} \Bigg[\left\| \frac{m_{kr}^{\star L} - m_{kr}^L}{\varpi_{L,kr}^\star} - \frac{\varpi_{L,kr}}{\varpi_{L,kr}^\star} \frac{\varphi_{k,r}^L}{\sqrt{\#jm}} \right\|_+^2 \\ + \left\| \frac{m_{kr}^{\star H} - m_{kr}^H}{\varpi_{H,kr}^\star} - \frac{\varpi_{H,kr}}{\varpi_{H,kr}^\star} \frac{\varphi_{k,r}^L}{\sqrt{\#jm}} \right\|_-^2 \Bigg], \end{aligned}$$

where the superscript $\star$ indicates the quantities calculated on the bootstrapped samples. The shifts of moments φ is defined in AS. The idea is to reduce the effects of the moments that are not binding at (θ, σ) but have large variances.

5(AS) Take the 95% critical value $c_{95}(\theta, \sigma)$ to be the 95% quantile of the bootstrapped statistic. Reject (θ, σ) if $T(\theta, \sigma) > c_{95}(\theta, \sigma)$.

Table D.1: Changes in the Number of Products[a]

	(a) Merging Firms	(b) Other Firms	(c) Other Firms	(d) Market	(e) Market
average price increase, merging firms ($), variety fixed	[-2.82, -1.93]*	[0.27, 0.68]*	[0.30, 0.70]*	[-2.15, -1.40]*	[-2.16, -1.38]*
average fixed cost ($1000), merging firms	[-0.68, -0.12]			[-0.63, -0.05]	[-0.64, 0.01]
average fixed cost ($1000), other firms		[-1.32, -0.32]*	[-1.32, -0.32]*	[-1.74, -0.31]	[-1.74, -0.40]
average household income ($10,000)			[-0.01, 0.01]		[-0.03, 0.01]
market size (10^6)	[-0.07, 0.08]	[0.16, 0.48]*	[0.16, 0.48]*	[0.09, 0.57]*	[0.09, 0.57]*
R^2	[0.28, 0.33]	[0.62, 0.76]	[0.62, 0.76]	[0.19, 0.42]	[0.20, 0.42]
N	149	149	149	149	149

[a]Linear regressions where the dependent variables are the changes in the numbers of products by merging firms and other firms and the net change in a market. Each observation is a market. We report the range of estimates from the parameters in the confidence set. * indicates significance above 95% confidence level for all parameters in the sampled confidence set.

D Correlation Between Merger Effects on Product Variety and Market Characteristics

In this section, we use regression analysis to document the correlation between the merger effects on the number of products and a set of market characteristics motivated by the discussion in Section 6.2. Our measure of market power is the quantity-weighted average price increase of the merging firms' products in a market, holding the variety fixed as observed in the data.[20] We calculate the fixed costs by

$$\hat{\theta}_0 + \hat{\theta}_1 \cdot (\text{in state})_j + \hat{\theta}_2 \cdot \text{craft}_j + \hat{\theta}_3 \cdot (\text{in state})_j \cdot \text{craft}_j \cdot \text{dist}_{jm},$$

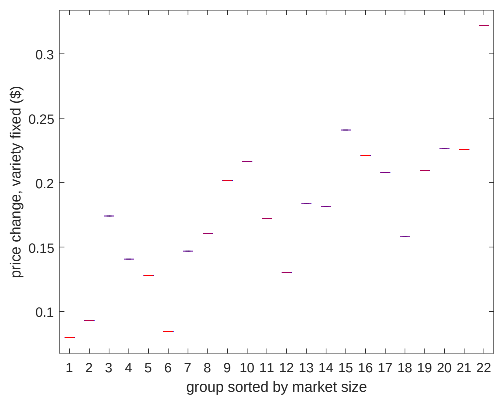
i.e. not including the unobserved fixed cost shocks, so that these fixed costs reflect the underlying distances between the in-state craft breweries to the markets. The average fixed cost of merging firms in a market is defined as the average of the fixed costs of all potential products by the merging firms for the market. The average fixed cost of other firms is defined analogously. We provide the summary statistics of these variables in Figure D.1. The largest average increase in prices within a group is 0.32 dollars. The increases are more pronounced in medium-sized and large markets, where we see both large increases and decreases in the number of products. The mean components of the fixed costs are about 4,500 dollars.

We regress the changes in the numbers of products by merging firms (so dropping products corresponds with a negative change), by other firms and the net change in a market on our measure of market power, mean fixed costs and market sizes. We carry out the regressions for each sampled parameter in the confidence set and report the range of the parameter estimates in Table D.1. The signs in Columns (a) and (b), corresponding with the changes in the number of products by merging firms and other firms, are consistent with the discussions in Section 6.2. The estimates are significant at 5% level for all parameters in the confidence set for variety-fixed price increases in Columns (a) and (b), and significant for the fixed costs in Column (b). In Columns (d), the

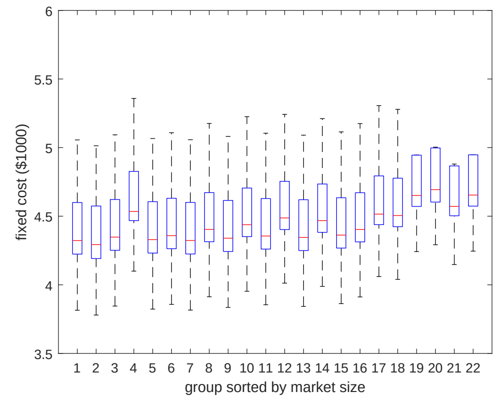
[20]An alternative is the market share of the merging firms' craft products among all craft products. We find similar but weaker correlations.

Figure D.1: Market Power and Fixed Costs

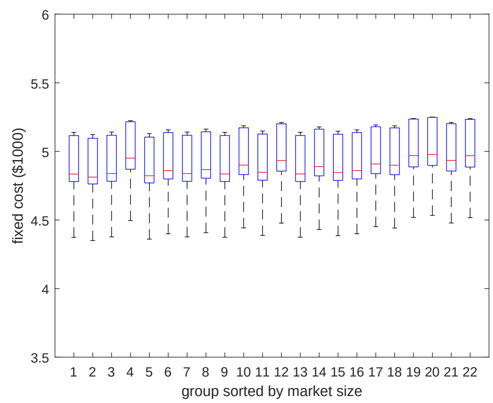
(a) Mean Price Increase ($) of Merging Firms



(b) Merging Firms Average Fixed Costs ($1000)



(c) Other Firms: Average Fixed Costs ($1000)



correlation between our measure of market power and the number of products is weaker than (a) and (b), reflecting the countervailing nature of how market power both induces product entry and causes the merging firms to drop products. The sign is negative, indicating that market power is correlated with a net decrease in the number of products. In the three regressions, fixed costs are correlated with a reduced number of products. We also note that market sizes are positively correlated with product entry. In Columns (c) and (e), we also control for the average income of the county of the market, which does not appear to be correlated with the changes in variety. Finally, the measure of market power and average fixed costs account for a significant portions of variations in the changes of the number of products. By regressing the dependent variables on just market sizes, the R^2s drop to $[0.01, 0.09]$, $[0.59, 0.72]$ and $[0.00, 0.26]$ for columns (a), (b) and (d), respectively.

E (Dis)Economies of Scope

To allow for potential economies or diseconomies of scope in fixed costs, we modify the baseline fixed cost function in Section 4 by adding a parameter θ_4:

$$\begin{aligned} &\theta_4 1\left(\sum_{j\in\mathcal{J}_n} a_{jm} > 0\right) \\ &+ \sum_{j\in\mathcal{J}_n} a_{jm}\left(c_{jm}(W_{jm},\theta) + \zeta_{jm}\right), \end{aligned} \tag{10}$$

where

$$c_{jm}(W_{jm},\theta) = \theta_0 + \theta_1 \text{craft}_j + \theta_2 (\text{in state})_j \cdot \text{craft}_j + \theta_3 (\text{in state})_j \cdot \text{craft}_j \cdot \text{dist}_{jm}.$$

This fixed cost function in (10) is no longer additive in the fixed cost of each product. If $\theta_4 > 0\,(<0)$, the fixed cost exhibits (dis)-economies of scope.

The bounds in Theorem 1 now need to take into account the part $\theta_4 1\left(\sum_{j\in\mathcal{J}_n} a_{jm} > 0\right)$. We consider the following modification:

$$\begin{aligned} &\Pr\left(\zeta_{jm} < \underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm},\theta) - |\theta_4|_+\right) \\ \le &\Pr\left(A_{jm}(X_m, W_m, \zeta_m, \theta) = 1 \mid X_m, W_m\right) \\ \le &1 - \Pr\left(\zeta_{jm} > \bar{\Delta}_{jn}(X_m) - c_{jm}(W_{jm},\theta) + |\theta_4|_-\right), \end{aligned} \tag{11}$$

where $|\theta_4|_+$ is θ_4 if $\theta_4 > 0$ and 0 otherwise; $|\theta_4|_-$ is $-\theta_4$ if $\theta_4 < 0$ and 0 otherwise. The bounds hold for any value of θ_4: when product j is added, the increased fixed cost is either $c_{jm} + \zeta_{jm} + \theta_4$ or $c_{jm} + \zeta_{jm}$; therefore for the left-hand side, we consider the minimum of the two and for the right-hand side, we consider the maximum of the two. However, the bounds in (11) alone are not sufficient to estimate the fixed cost function, because the bounds become weakly looser when $|\theta_4|$

increases.

To estimate θ_4, we also consider the following additional inequalities related to the probability that a firm has at least one product in a market:

$$\Pr\left(\sum_{j\in\mathcal{J}_n} A_{jm}\left(X_m, W_m, \zeta_m, \theta\right) > 0 \,|\, X_{nm}, W_{nm}\right)$$

$$= \Pr\left(\max_{\substack{a_{nm} \\ \text{s.t. } \sum_{j\in\mathcal{J}_n} a_{jm} > 0}} \left[\Pi_{nm}\left(a_{nm}, a_{-nm}, X_m\right) - \sum_{j\in\mathcal{J}_n} a_{jm}\left(c_{jm}(W_{jm},\theta) + \zeta_{jm}\right)\right] - \theta_4 > 0\right). \tag{12}$$

By the definition of $\overline{\Delta}_{jn}(X_m)$, we have

$$\Pi_{nm}\left(a_{nm}, a_{-nm}, X_m\right) \leq \sum_{j\in\mathcal{J}_n} a_{jm}\overline{\Delta}_{jn}(X_m).$$

This gives us the upper bound for the probability in (12)

$$\Pr\left(\max_{\substack{a_{nm} \\ \text{s.t. } \sum_{j\in\mathcal{J}_n} a_{jm} > 0}} \left\{\sum_{j\in\mathcal{J}_n} a_{jm}\left(\overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm},\theta) - \zeta_{jm}\right)\right\} - \theta_4 > 0\right) \tag{13}$$

Similarly, we can construct the following lower bound:

$$\Pr\left(\max_{\substack{a_{nm} \\ \text{s.t. } \sum_{j\in\mathcal{J}_n} a_{jm} > 0}} \left\{\sum_{j\in\mathcal{J}_n} a_{jm}\left(\underline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm},\theta) - \zeta_{jm}\right)\right\} - \theta_4 > 0\right) \tag{14}$$

We can similarly derive bounds for "firm entry with at least two products":

$$\Pr\left(\max_{\substack{a_{nm} \\ \text{s.t. } \sum_{j\in\mathcal{J}_n} a_{jm} > 1}} \left\{\Pi_{nm}\left(a_{nm}, a_{-nm}, X_m\right) - \sum_{j\in\mathcal{J}_n} a_{jm}\left(c_{jm}(W_{jm},\theta) + \zeta_{jm}\right)\right\} - \theta_4 > 0\right).$$

Table E.1: Estimates of Fixed Costs: 2016

Constant (θ_0)	[2163.79, 3598.07]
Craft (θ_1)	[1866.87, 3357.79]
In State× Craft (θ_2)	[-1288.92, 669.20]
Distance^a×In State× Craft (θ_3)	[-154.33, 2904.70]
Scope (θ_4)	[-2237.18, -628.20]
Std. Dev. (σ_ζ)	[862.97, 2329.04]

^{a}: Distance in 1000KM

The bounds above give us additional $L_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta)$ and $H_{jm}(A_m, X_m, W_m, \theta, \sigma_\zeta)$ terms in the GMM objective function (12). The bounds of the probabilities that a firm has at least one product or two products in the market do not have analytical forms, and in estimation, we simulate these bounds. The simulation is still faster than CT because the optimizations (13) and (14) are simple linear optimization problems. Specifically, for (13), we start with a counter $B_{nm} = 0$. We take L draws of ζs, and for $\zeta_{jm}^{(\ell)}$ of the draw ℓ, we check whether

$$\overline{\Delta}_{jn}(X_m) - c_{jm}(W_{jm}, \theta) - \zeta_{jm}^{(\ell)}$$

is positive for each $j \in \mathcal{J}_n$. We then add up all the positive terms and subtract θ_4. If the resulting quantity is positive, we add 1 to B_{nm}. Finally, the probability in (13) is simulated as $\frac{B_{nm}}{L}$. The probability in (14) is simulated analogously.

One additional challenge is to construct the exogenous function to interact with the bounds of firm-level entry probabilities. We use the following firm-level characteristics

$$\left(\overline{\Delta}_{jn}\left(X_m, W_m\right), \underline{\Delta}_{jn}\left(X_m, W_m\right), W_{jm}\right)_{j \in \mathcal{J}_n}$$

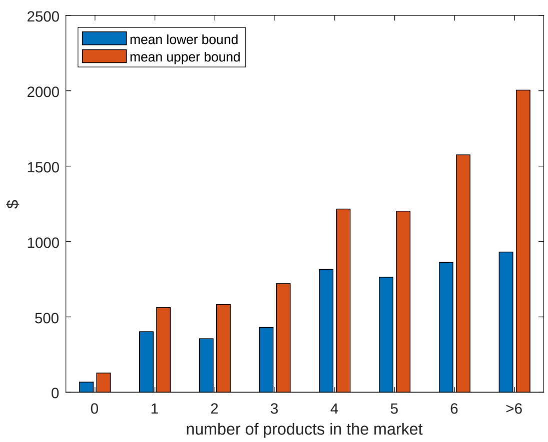
to construct F_{nm} functions, analogous to the F_{jm} functions in (9). We first group firms by the number of potential products. Within each group, we use a kmeans algorithm to classify firms by the vector $\left(\tilde{Z}_{jm}^c\right)_{j \in \mathcal{J}_n}$ into r classes, which are transformed continuous product characteristics (Section 5.2.3). We use $\|\mathcal{J}_n\|$ to denote n's number of potential products, and $\phi(n) \in \{1, \ldots, r\}$ to denote the kmeans class of firm n. We denote an element in the set

$$\tilde{\boldsymbol{\mathcal{C}}} = \left\{1, \ldots, \bar{J}\right\} \times \{1, \ldots, r\} \times \{0, 1\}^{|Z_{jm}^b|}$$

as $\tilde{\mathcal{C}}$, where $\bar{J}$ is the highest number of products per firm. We construct indicator functions F_{nm} by whether $\left(\|\mathcal{J}_n\|, \phi(n), Z_{jm}^b\right) = \tilde{\mathcal{C}}$ for each element in $\tilde{\boldsymbol{\mathcal{C}}}$.

Table E.1 reports the estimation results. We use CHT for inference to reduce the computational burden. We find dis–economies of scope: the projected CI of the θ_4 is $[-2237.18, -628.20]$. To see what data pattern leads to this estimate, we compute the average $\overline{\Delta}_{jn}(X_m)$ and $\underline{\Delta}_{jn}(X_m)$ across all jm pairs such that the number of products that firm n has in market m is h. We plot these averages for $h = 0$,1,...,6 and >6 in Figure E.1. The figure suggests that for firms that place more

Figure E.1: Average $\underline{\Delta}_{jn}(X_m)$ and $\overline{\Delta}_{jn}(X_m)$, Conditional on the Number of Products by a Firm in a Market



products in a market, the average incremental variable profit are higher. This pattern is suggestive evidence of dis–economies of scope: everything else equal, if the average fixed cost increases in the number of products, in equilibrium we should see that the average variable profit is positively correlated with the number of products by a firm.

In the merger simulation, we assume that the common owner of the acquired craft breweries coordinates pricing and product entry decisions, but does not change the underlying distribution of fixed costs. Therefore the fixed cost reduction from θ_4 applies to each acquired brewery that enters a market.

To see the robustness of our results, we report results in Figures E.2 and E.3, corresponding with Figures 3 and 4 and 4. The results are similar.

F Fixed Cost Simulation Draws Conditional on Observed Equilibrium Outcomes

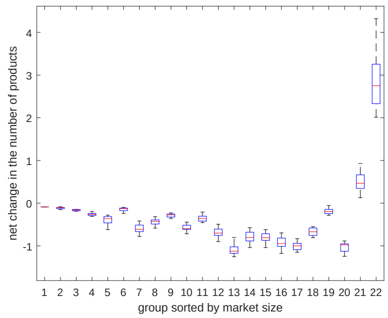
We draw the fixed costs that are consistent with both the estimated underlying distribution of fixed cost and the pre-merger, observed outcome as a pure-strategy equilibrium. As explained in Section 6, it is important to take into account the latter requirement, which is essentially a selection issue. To obtain one such set of draws in market m, we proceed with the following steps:

1. For each potential product j, we calculate an upper and a lower bound of the fixed cost shock ζ_{jm} as follows. If j is in the market before the merger,

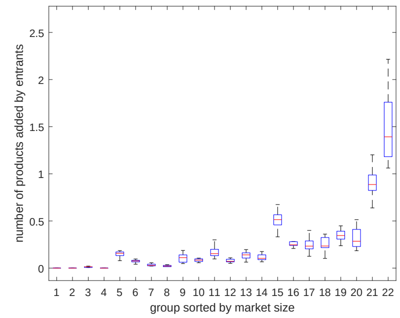
$$\tilde{L}_{jm} = -\infty, \tilde{H}_{jm} = \underline{\Delta}_{jn}(X_m)$$

Figure E.2: Product Variety, Entry and Prices

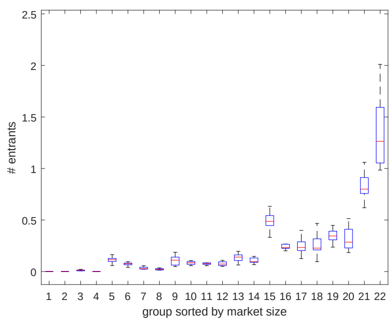
(a) Net Change in the Number of Products



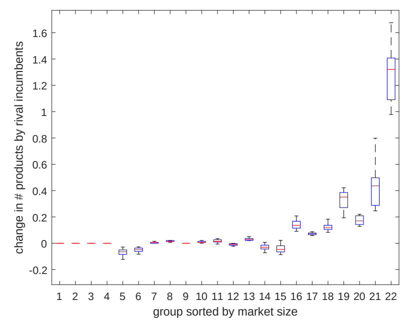
(b) Number of Products Added by Entrants



(c) Number of Entrants



(d) Change in the Number of Products by Incumbent Non-merging Firms



(e) Change in the

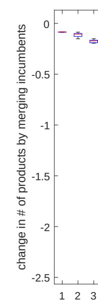
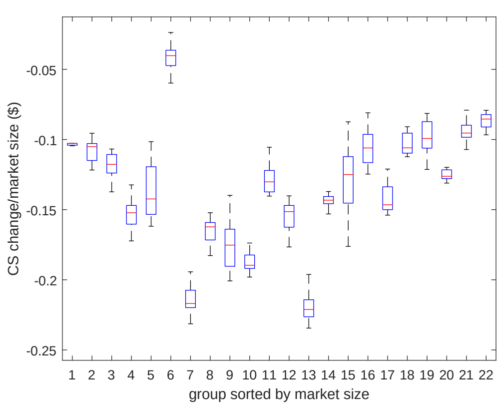
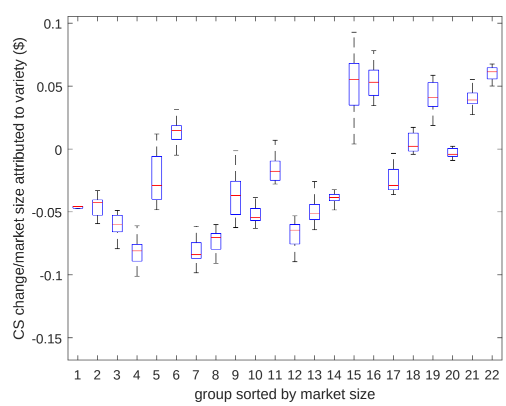


Figure E.3: Average Consumer Surplus

(a) Change in the Average Consumer Surplus



(b) Change in the Average Consumer Surplus Attributed to Variety



and if product j is not in the market,

$$\tilde{L}_{jm} = \underline{\Delta}_{jn}(X_m), \tilde{H}_{jm} = \infty.$$

2. We simulate draws of the fixed cost shocks from a truncated normal distribution with the underlying normal distribution parameterized by mean 0 and variance $\hat{\sigma}_\zeta^2$ and the truncation being $\tilde{L}_{jm} < \zeta_{jm} < \tilde{H}_{jm}$. These draws satisfy the necessary conditions for the observed equilibrium.

3. We next verify these draws indeed support the equilibrium. To do so, we employ the algorithm in Appendix G with the starting points $\mathcal{J}_{nm}^0 = \emptyset$ and $\mathcal{J}_{nm}^0 = \mathcal{J}_n$ and check whether the algorithm converges to $\mathcal{J}_{nm}$, holding $\mathcal{J}_{-nm}$ fixed. If the algorithm converges to $\mathcal{J}_{nm}$ from both starting points, we keep the set of draws for n. If at least one of the starting points does not lead to $\mathcal{J}_{nm}$, we go back to Step 2 and re-draw the fixed costs.

4. We repeat this process for every firm n.

G Equilibrium Computation Details

We conduct the counterfactual simulations using the algorithm in Fan and Yang (2020). At a high level, for each market, we solve for the equilibrium using best response iterations where firms take turns to choose their products until no firm has an incentive to deviate. We use two orderings of firms (i.e., ascending and descending order based on the observed annual sales) and find the same equilibrium. Embedded in the best-response iteration procedure is an optimization problem to determine a firm's best response. Similar to Fan and Yang (2020), firms in our model often face too large a choice set. For example, the merging firms in our counterfactual analysis has a total of 51 potential products, leading to a choice set of $2^{51} \approx 10^{15}$. We thus use a heuristic algorithm described below to determine each firm's best response. Additional discussions and Monte Carlo simulations demonstrating the performance of the algorithm can be found in Fan and Yang (2020).

In the following, we use firm n and market m as an example. Starting with a product portfolio $\mathcal{J}_{nm}^0 \subseteq \mathcal{J}_n$, we compute firm n's profit from each of the following deviations from $\mathcal{J}_{nm}^0$: $\mathcal{J}_{nm}^0 \backslash k$ for $k \in \mathcal{J}_{nm}^0$ or $\mathcal{J}_{nm}^0 \cup k$ for $k \in \mathcal{J}_n \backslash \mathcal{J}_{nm}^0$. Each deviation differs from $\mathcal{J}_{nm}^0$ in only one product: either a product in $\mathcal{J}_{nm}^0$ is removed or a potential product not in $\mathcal{J}_{nm}^0$ is added. There are $\#\mathcal{J}_n$ such deviations. Let $\mathcal{J}_{nm}^1$ be the highest-profit deviating product portfolio. If firm n's profit corresponding to $\mathcal{J}_{nm}^1$ is smaller than that corresponding to $\mathcal{J}_{nm}^0$, this procedure stops and returns $\mathcal{J}_{nm}^0$ as the best response. Otherwise, we compute n's profit from any one-product deviation from $\mathcal{J}_{nm}^1$ by either adding a potential product to or dropping a product from $\mathcal{J}_{nm}^1$. We continue this process until firm n's profit no longer increases. This algorithm allows us to translate a problem whose action space grows exponentially in the number of potential products (choosing from $2^{\#\mathcal{J}_n}$ product portfolios) into one whose action space grows linearly (in each step, evaluating $\#\mathcal{J}_n$ portfolios).

Due to the computational burden, we simulate 35 sets of draws of fixed costs and use them to conduct counterfactuals.

H Monte Carlo Simulations

In this appendix, we use Monte Carlo simulations to examine the performance of our estimation procedure for estimating the fixed cost parameters. Our purposes are two-fold: first, we show that our approximations of $\overline{\Delta}_{jn}$ and $\underline{\Delta}_{jn}$ work well; second, we show that our estimation method generally result in conservative but still reasonably tight confidence intervals of the true parameters.

We first describe the data generating process. The demand is given by a nested Logit model with a nest over all inside products $\mathcal{J}_m$ in market m. The market share of good j in market m is given by

$$s_{jm} = \frac{\exp\left(\delta_{jm}/\left(1-\rho\right)\right)}{\sum_{j'\in\mathcal{J}_m}\exp\left(\delta_{j'm}/\left(1-\rho\right)\right)}\frac{\exp\Big(\left(1-\rho\right)\sum_{j'\in\mathcal{J}_m}\exp\left(\delta_{j'm}/\left(1-\rho\right)\right)\Big)}{1+\exp\Big(\left(1-\rho\right)\sum_{j'\in\mathcal{J}_m}\exp\left(\delta_{j'm}/\left(1-\rho\right)\right)\Big)},$$

where ρ is the nesting parameter and the mean utility is $\delta_{jm} = \alpha p_{jm} + \beta x_{jm}$. We define O_m to be the market size, and the demand for j in market m is $O_m \cdot s_{jm}$. The marginal cost is mc_{jm}. The fixed cost is

$$\theta_0 + \theta_1 W_{jm} + \zeta_{jm},$$

where ζ_{jm} follows a normally distribution with mean 0 and variance σ_ζ^2.

We set $\alpha = -0.5$, $\rho = 0.2$ and draw x_{jm} and mc_{jm} from a normal distribution with a mean of 2 and a standard deviation of 0.25, truncated on the support of $[1.5, 2.5]$. In the fixed cost function, $W_{jm} \sim N\left(0, 3\right)$. The market size O_m is uniformly drawn from $(0, O)$. There are two firms per market and they have the same number of potential products. We consider 7 specifications (7 different values for $(O, \beta, \sigma_\zeta, \mu_\zeta, \theta)$ and the number of potential products per firm), and for each specification, we simulate 1000 markets. For each market, we find the equilibrium by enumerating and checking all possible outcomes. When there are multiple equilibria, we assign equal probability to each equilibrium and randomly choose one.

With the simulated data, we conduct two exercises. First, we examine how close our approximations are to the true extrema of the incremental profits. We compute the true minimum and maximum of the change in variable profit $\underline{\Delta}_{jn} = \min_{a_{-jm}} \Delta_{jn}\left(a_{-jm}\right)$ and $\bar{\Delta}_{jn} = \max_{a_{-jm}} \Delta_{jn}\left(a_{-jm}\right)$ as well as our approximations of them, i.e., $\Delta_{jn}\left(1, ..., 1\right)$ and $\Delta_{jn}\left(0, ..., 0\right)$. We find that they match perfectly for all product-market combinations in all specifications.

Second, we estimate the fixed cost parameters $(\theta_0, \theta_1, \sigma_\zeta)$ and obtain the 95% confidence set following the estimation and inference procedure in Appendix C. We report the coverage probability that the confidence set covers the true parameter from simulating the competition between 2 firms in 1000 markets for 7 specifications.[21]

[21] Unlike the empirical model, in the Monte Carlo simulation we use fewer observations and less rich covariates.

Table H.1: Monte Carlo Simulations: 2 firms

	O	β	σ_ζ	μ_ζ	θ	# products/firm	$\text{med}\left(\frac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}\right)$	Multi Equi %	Cvg Rate
(1)	50000	-5	1.5	1.5	1.5	2	1.6	0.03%	98.00%
(2)	100	0	12	3	1.5	2	2	1.25%	100.00%
(3)	50	0.8	15	6	1.5	2	3.6	2.50%	100.00%
(4)	50000	-5	1.5	2.5	0.5	2	1.6	0.03%	98.00%
(5)	50000	-5	1.5	0.5	2.5	2	1.6	0.02%	99.00%
(6)	50	0.8	15	15	1.5	2	3.6	1.40%	100.00%
(7)	50	0.8	15	15	1.5	3	5.5	1.27%	100.00%

We report the results in Table H.1. We vary our parameter values to adjust the tightness of our bounds, which we measure by the median of the ratio $\frac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}$. The tightness of the bounds has a direct impact on estimation. Wide bounds may not be informative of the underlying parameters, while tight bounds may be more likely to result in point identification: intuitively, a single agent discrete choice problem is identified under weaker conditions than discrete games. To put the number in perspective, the median ratio is 1.53 in the estimation sample in Section 5.2. We also report the percentage of markets with multiple equilibria in the Column Multi Equi %. The frequency of multiple equilibria overall amounts to a small fraction of the simulated markets (2.50% in the most extreme cases). Finally, because the identified set does not have an analytic form, and simulating it involves solving both the price and product placement equilibria and is computationally costly, we therefore report the coverage probability for the true parameter instead of the identified set in the last column. The coverage is higher than 95%, which is expected, and increases in $\text{med}\left(\frac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}\right)$.

We next examine the false coverage probabilities of the projected 95% confidence intervals. Typically (for example Andrews and Shi (2013)) authors report the false coverage probability defined as the probability that a confidence set includes a point just outside the identified set. As explained above, the identified set of the model in our Monte Carlo exercise does not have an analytic form, and so it is not possible to precisely pin down the boundary of the identified set. To still give a sense of how "wide" confidence intervals tend to be, we instead calculate the probabilities that points around the true parameter values are covered by the 95% confidence interval. In Figure H.1, we plot the coverage probabilities by the projected confidence intervals from the confidence sets for points in intervals around the true parameters. We do so for the specifications in Row (1) of Table H.1, where the tightness of the bounds is similar to the empirical setting, and Row (7) of Table H.1, where the bounds are wider. By construction, the y-axis shows how often a point is included in the confidence interval. When the bounds are tight ($\text{med}\left(\frac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}\right) = 1.6$), the results show the projected confidence intervals are tight: for example, in Figure H.1 (1), the vertical line indicates that the true value of the fixed cost intercept μ is 1.5, and the probability of the confidence

As a result, we sometimes do not have point identification and find multiple parameters that reduce the objective function to 0.

Table I.1: Estimates of Fixed Costs: 2016

Craft (θ_1)	[485.88, 2168.68]
In State$\times$ Craft (θ_2)	[-656.55, 837.75]
Distance^a$\times$In State$\times$ Craft (θ_3)	[814.51, 2586.09]
Constant (θ_0^k)	
Small Market	[839.50, 2131.35]
Medium Market	[1400.49, 2384.31]
Large Market	[3655.23, 5815.32]
Std. Dev. (σ_ζ^k)	
Small Market	[90.42, 1218.75]
Medium Market	[471.74, 1305.17]
Large Market	[3465.14, 4917.70]

^{a}: Distance in 1000KM

interval covering 3.5 or 0.5 is close to 0. When the bounds are more relaxed, the confidence intervals are proportionally wider for $(\sigma_\zeta, \mu_\zeta)$ (relative to the values of the true parameters). The interval is often twice as wide for θ_1. In both cases, the false coverage probabilities of 0 are reasonably small for all parameters across both specifications.

I Market Size-Dependent Fixed Costs

In this section, we present the results where we allow fixed costs to depend on market sizes. Specifically, we define market sizes below the 25% quantile as "small", between 25%–75% as "medium" and above 75% as"large" markets. We modify the baseline fixed cost function of (6) to allow depend on market sizes:

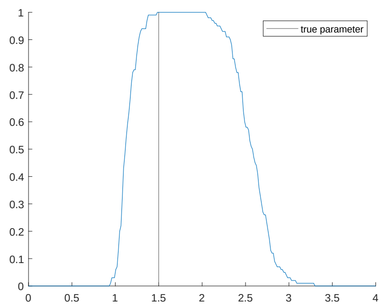
$$C_{nm}^k\left(\mathcal{J}_{nm\tau}\right) = \sum_{j\in\mathcal{J}_{nm\tau}} \left(\theta_0^k + \theta_1 \text{craft}_j + \theta_2 (\text{in state})_j \cdot \text{craft}_j + \theta_3 (\text{in state})_j \cdot \text{craft}_j \cdot \text{dist}_{jm} + \zeta_{jm\tau}^k\right),$$

where $k \in \{\text{small, medium, large}\}$and $\zeta_{jm\tau}^k \sim \Phi\left(0, \sigma_\zeta^k\right)$. We use CHT for inference to lessen the computational burden. Table I.1 reports the estimates of the projected 95% confidence sets. Figures I.1 and I.2 report the counterfactual results, which are similar to our baseline.

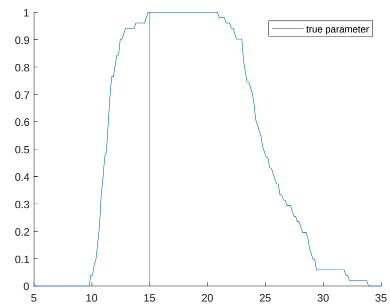
Figure H.1: False Coverage Probability. Left: $\text{med}\left(\dfrac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}\right) = 1.6$; Right: $\text{med}\left(\dfrac{\overline{\Delta}_{jm}}{\underline{\Delta}_{jm}}\right) = 3.6$

FC Intercept μ

(1) Specification (1)

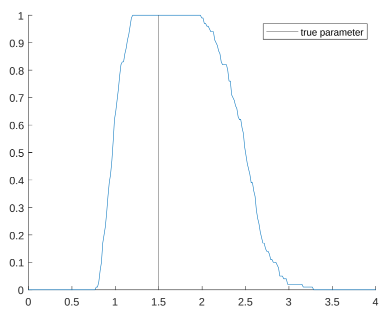


(2) Specification (7)

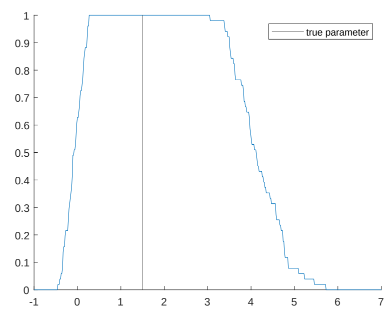


FC Covariate Coefficient θ

(3) Specification (1)

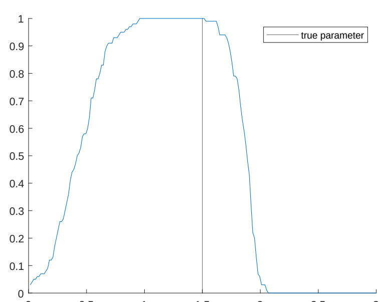


(4) Specification (7)



FC Unobservable Standard Deviation σ_ζ

(5) Specification (1)



(6) Specification (7)

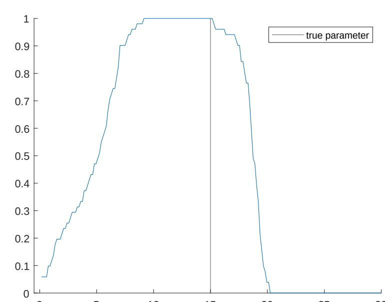
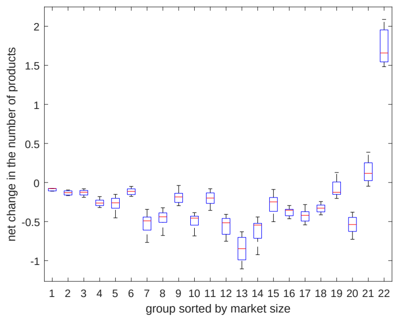
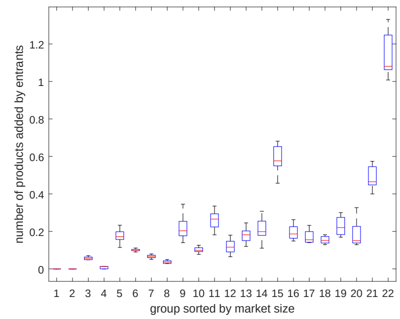


Figure I.1: Effects of Merger on the Number of Products

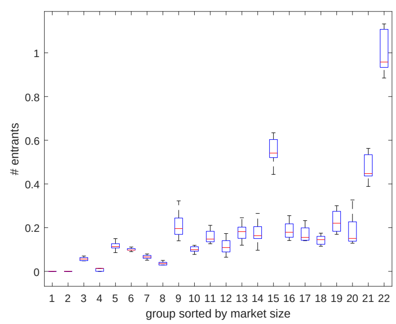
(a) Net Change in the Number of Products



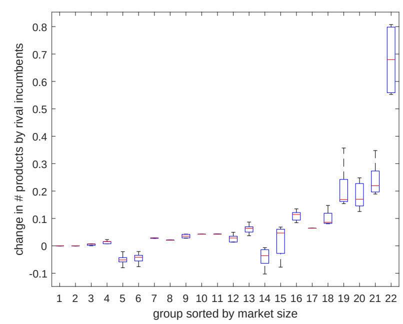
(b) Number of Products Added by Entrants



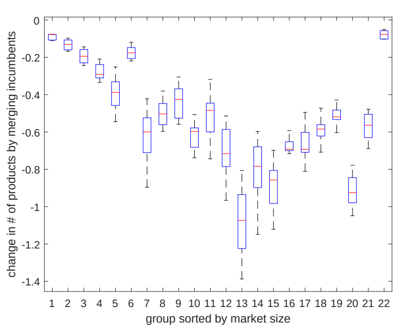
(c) Number of Entrants



(d) Change in the Number of Products by Incumbent Non-merging Firms



(e) Change in the Number of Products by Merging Firms



(f) Craft Prices

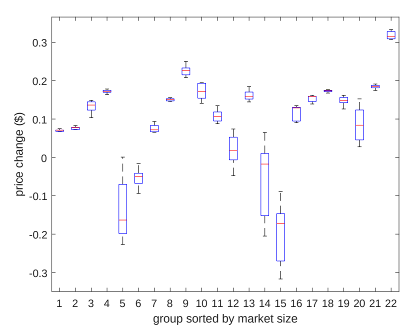
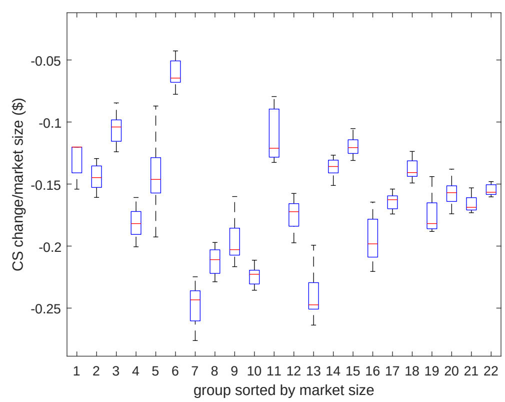


Figure I.2: Average Consumer Surplus

(a) Change in the Average Consumer Surplus



(b) Change in the Average Consumer Surplus Attributed to Variety