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

these effects vary across markets? In this paper, we address these questions and study the effects of merger on prices, product portfolios and 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 choice. 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 are endowed with a set of each potential product, and choose the set of products to sell in a market. 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 the second stage of our model are consumer preferences and marginal costs. We estimate beer demand with data on market shares and panel data of individual choices. The first-order conditions for the optimal prices allow us to estimate marginal costs.

The key primitives in the first stage of the model are the fixed costs of product entry into markets. Both the merger outcomes and welfare predictions depend critically on the estimates of 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 and smaller than the probability that $a = 1$ is not a dominated strategy. In the paper, we explain the assumption on firm equilibrium behavior, explain these bounds, and provide details on estimation and inference.

et al. (2020)) or whether to provide non-stop service (Li et al. (2019)). Therefore, we discuss product variety together with firm entry,[4] and we use a new method to address the estimation challenges that arise from allowing for a larger 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. (2020) 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 allow for the latter selection only and address the selection on demand and marginal costs by including a large number of fixed effects in our demand and marginal cost functions. 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. (2020) 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.

In a recent paper, Wang (2020) proposes a hybrid approach that combines the framework in Ciliberto et al. (2020) with probability bounds based on the concept of dominant strategies. The computational burden of such an approach is between our methods and Ciliberto et al. (2020).

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.[5]

identification and estimation details.

In this illustrative model, firms 1 and 2 decide whether to enter market m according to

$$
\begin{aligned}
Y_{1m} &= 1\,[\ \ _{1m}(Y_{2m}) - C(W_{1m};\ \ ) - \ \ _{1m} \geq 0]; \\
Y_{2m} &= 1\,[\ \ _{2m}(Y_{1m}) - C(W_{2m};\ \ ) - \ \ _{2m} \geq 0];
\end{aligned} \tag{1}
$$

where $Y_{jm} = 1$ indicates entry. In (1), $\ _{jm}(Y_{-jm})$, $j \ \ \{1;2\}$ is a variable profit function that is known (or has been estimated), $C(W_{jm};\ \ )$, $j \ \ \{1;2\}$ is the fixed cost of entry, where W_{jm} is a vector of covariates and is a vector of parameters to be estimated, and finally, $\ _{jm}$ is a fixed cost highlight the advantages of using our constructed bounds for estimation. First, these bounds contain useful information on the fixed cost function. Consider the extreme case where $_{jm}(1) = {}_{jm}(0)$. The inequalities collapse into an equality, and estimation becomes GMM estimation of binary choice models (McFadden (1989)). Therefore, intuitively, the usefulness of the inequalities for estimating fixed cost parameters depends on the gap between $_{jm}(1)$ and $_{jm}(0)$, which we explore in the later estimation section and through Monte Carlo exercises. Second, these bounds are one-dimensional CDFs and easy to compute. Computing them 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 there are multiple equilibria, the equilibrium selection mechanism differs across markets, or there is no pure strategy equilibrium.

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. Moreover, California has accommodating distribution laws. 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)). California beer statues do not require breweries to satisfy a burdensome "good cause" clause to terminate a contract with a distributor. Taken together, we find it reasonable to assume that the distribution market is sufficiently competitive,[6] and we do not consider the strategic behaviors of distributors in our model later.[7] 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)). These institutional features motivate our modeling assumption that beer firms make the entry and product variety decisions based on the retail profitability of their products. The simplifications keep our model tractable and allow us to focus on the entry and product variety decision of breweries.

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

[6] In 2020, California has a population of 39 million and 3178 licensed distributors. In comparison, Michigan, a state with beer restrictions similar to the majority of US states, has a population of 10 million and 160 distributors.

[7] 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 reduced 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.

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 equivalent) and adding quantities across weeks within a month and using the quantity-weighted average price across weeks within a month as the product's price in that month.

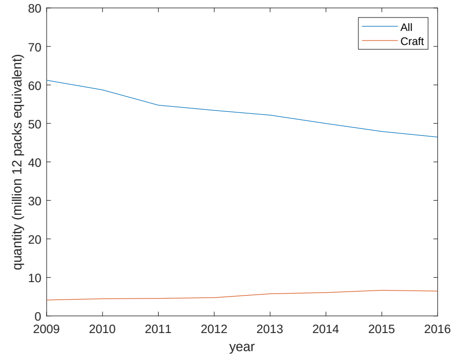
We define a market as a retailer-county pair. Our data suggest cross-retailer shopping appears rare: more than 80% of the households purchased all of their beers from one retailer-county combination in 2016. Similarly, Huang et al. (2020) and Illanes and Moshary (2020) find little evidence of retailer competition in the spirit category.

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 production data. We thus focus on breweries established in the 1990s or earlier. Many of these craft breweries had 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. 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 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)).[8]

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 1. 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 1 (a)). The average price is stable around 11 dollars per unit. The average price of craft products increased from 16 to 17 dollars (Figure 1 (b)). There are an average of 52 firms in

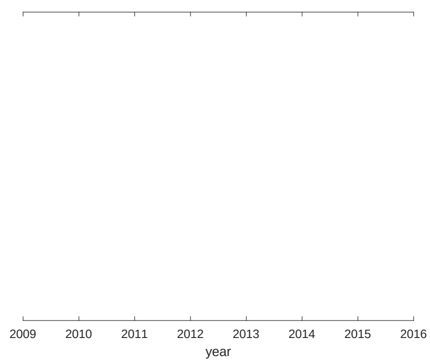
Figure 1: Summary Statistics

(a) Quantity

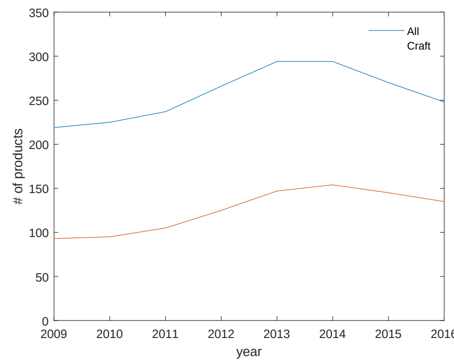


(b) Price

(c) Number of Firms

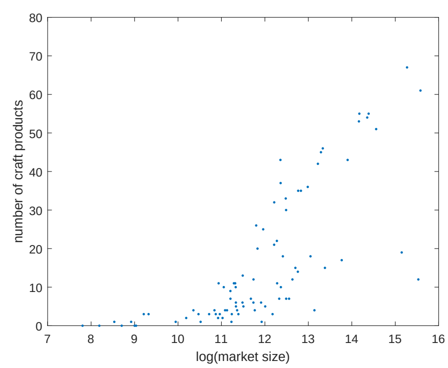


(d) Number of Products



(e) Distance from the Brewery to the Market

(f) Market Size and Number of Craft Products



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. The 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 where y_i is the household income and $^{(\cdot)}_i$ is the household-specific unobserved taste shock to each product attribute, which follows a standard normal distribution and is independent across households. Therefore, the $^{(\cdot)}$ parameters capture the dispersion in unobserved household tastes, and the $^{(\cdot)}$ parameters measure the e□ect of household income on tastes. We also include month fixed e□ects ($_t$) and product-market fixed e□ects ($_{jm}$) to capture unobserved factors that vary across months and product-market pairs. The error term $_{jmt}$ captures the month-to-month variations of demand shocks specific to a product, market and month combination. We do not include mean coe□cients for x_j because they are absorbed in fixed e□ects $_{jm}$. Finally, the last term in (4), "$_{ijmt}$;is the household idiosyncratic taste, which is assumed to be i.i.d. and follows type-1 extreme value distribution.

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 e□ects 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. In each market, firms simultaneously choose which beers, if any, to sell. This product choice is made at the beginning of each year and is fixed through the year. We use $\mathcal{J}_{nm}$ to denote firm n's products in market m in year . Then, in each month t, after observing that month's demand and marginal cost shocks, firms simultaneously choose retail prices.[10] We start from the second stage.

Stage 2. Pricing In month t, firm n chooses prices p_{jmt} for all $j \quad \mathcal{J}_{nm}$ to maximize its total variable profits:

$$\max_{p_{jmt};j \quad \mathcal{J}_{nm}} \sum_{j \quad \mathcal{J}_{nm}} (p_{jmt} - mc_{jmt}) D_{jmt}(p_{jmt}; p_{-jmt}): \qquad (5)$$

The marginal cost mc_{jmt} is decomposed into a product-market e□ect $!_{jm}$ and a product-market-month specific shock $!_{jmt}$:

$$mc_{jmt} = !_t + !_{jm} + !_{jmt}: \qquad (6)$$

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

$$\max_{J_{nm} \quad J_n} \quad _{nm}(J_{nm}\ ; J_{-nm}) - C_{nm}(J_{nm}): \tag{7}$$

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 x_{jm}, time fixed e□ects ($_t$; ! $_t$), product-market characteristics ($_{jm}$; ! $_{jm}$), and fixed costs for any product j $\quad J_n$ and any firm n. After firms make product decisions, the month-to-month transient demand and marginal cost shocks ($_{jmt}$; ! $_{jmt}$)

or dis–economies of scope. We extend the model to allow for this possibility in Appendix H.

5 Estimation

5.1 Estimation of Demand Parameters and Marginal Costs

We combine the aggregate data of product-market-level 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 ( ) and the fixed effect parameters ($_t$; $_{jm}$). The panel data and the correlations between household income and beer purchases help to identify the standard deviations of the unobservable consumer heterogeneity ($^{(\cdot)}$ parameters) and the effect of household income on consumer taste ($^{(\cdot)}$ 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.[11] The identifying assumption is

$$E(\ _{jmt} \mid x_j; V_{jmt}; t = 1, \ldots, 12) = 0.$$

We use these instruments to capture the large month-to-month variations in the ingredient costs (barley, malt, ...) and the heterogeneity in the proportions of these ingredients in different beers. Figure 2 shows the historical prices of barley and malt (in the units of per metric ton prices of barley and an index that linearly reflects the average prices of Malt and related products).

We next specify the micro moments. We construct the first set of micro moments to identify the standard deviation parameters $^{(\cdot)}$. Take ^{craft} as an example. Intuitively, if the parameter ^{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 ^{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., $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$; 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

[11] To be precise, our instruments consist of variables $(\bar{p}_{jmt}; Z_{jmt})$, where if j is available in the given ring, $\bar{p}_{jmt}$ is the Hausman instrument and $Z_{jmt} = 0$. Otherwise $\bar{p}_{jmt} = 0$ and $Z_{jmt} = 1$. The results are robust to further

Figure 2: Prices of Barley and Malt

(a) Barley (b) Malt

^{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 in month t. Matching this moment helps to identify $_0$:

We present the analytic expressions of these moments 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 I, i.e., $E\left[p_{j(i)mt} \mid y_i \quad I\right]$; where $p_{j(i)mt}$ is the price of the product purchased by household i in market m and month t, the income bins I are (0; \$50K], (\$50; \$100K] or (\$100K; \$150K).

$E \sum_{t}^{12}$

market. In Section 2, we have illustrated our bounds using a simple two-firm single-binary-decision model. In what follows, we present the bounds in our more general setting, explain the estimator, and describe implementation details. 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 H, we extend our method for estimating fixed cost functions that allow for (dis–)economies of scope.

5.2.1 Bounds for the Conditional Choice Probability

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

$$_{nm}(J_{nm}; J_{-nm}) - \sum_{j \quad J_{nm}} (_0 + _1 \cdot (\text{in state})_j + _2 \cdot \text{craft}_j + _3 (\text{in state})_j \cdot \text{craft}_j \cdot \text{dist}_{jm} + _{jm})$$

as

$$_n(a_{nm}; a_{-nm}; X_m) - \sum_{j \quad J_n} a_{jm} [c(W_{jm};) + _{jm}];$$

where the vector X_m includes all relevant demand and marginal cost covariates (including the fixed e□ects) 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 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 J_n). The element of a_{nm} that corresponds to product $j \quad J_n$ (denoted by a_{jm}) is 1 if $j \quad J_{nm}$ 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

$$_j(a_{-jm}; X_m) = \quad _n(a_{jm} = 1; a_{-jm}; X_m) - \quad _n(a_{jm} = 0; a_{-jm}; X_m) \tag{10}$$

to be the incremental change in firm n's expected variable profit when product $j \quad J_n$ is included in its product portfolio, given a_{-jm} and X_m. Given the discrete nature of a_{-jm}, .9091 Tf0 n uct j in market m, the bounds in (3) for our illustrative model becomes[12]

$$
\begin{aligned}
& F\left(\underline{\ }_j(X_m) - c(W_{jm};\ );\ \right) \\
\leq\ & \Pr(A_{jm} = 1 \mid X_{jm}; X_{-jm}; W_{jm}; W_{-jm}) \\
\leq\ & F\left(\overline{\ }_j(X_m) - c(W_{jm};\ );\ \right) :
\end{aligned}
\tag{11}
$$

The identification of   through these inequalities is similar to the idea of special regressors in entry games (Ciliberto and Tamer (2009); Lewbel (2019)): how A_{jm} varies with $(X_{jm}; X_{-jm})$ informs us about . For example, if the parameter   is large, the upper and lower bounds would not co-vary strongly with the $(X_m; W_m)$ and the inequalities in (11) are likely to be violated. For instance, if $_{jm}$ follows a symmetric distribution, both bounds in (11) approaches the constant 0:5 as   approaches ∞. On the other hand, as   is close to 0, both bounds simultaneously become close to 0 if $\overline{\ }_j(X_m) - c(W_{jm};\ ) < 0$, or close to 1 if $\underline{\ }_j(X_m) - c(W_{jm};\ ) > 0$ by the Chebyshev's inequality, potentially leading to violations of the inequalities in (11).

5.2.2 Identi cation and Inference

The identification of   through inequalities in (11) is similar to the idea of special regressors in entry games (Ciliberto and Tamer (2009); Lewbel (2019)): how A_{jm} varies with $(X_{jm}; X_{-jm})$ informs us about . For example, if the parameter   is large, the upper and lower bounds would not co-vary strongly with the $(X_m; W_m)$ and the inequalities in (11) are likely to be violated. For instance, if $_{jm}$ follows a symmetric distribution, both bounds in (11) approaches the constant 0:5 as   approaches ∞. On the other hand, as   is close to 0, both bounds simultaneously become close to 0 if $\overline{\ }_j(X_m) - c(W_{jm};\ ) < 0$, or close to 1 if $\underline{\ }_j(X_m) - c(W_{jm};\ ) > 0$

We follow Chernozhukov et al. (2007) and Andrews and Shi (2013) to construct the confidence set for the true values of the parameters and report results from both methods. More details on inference are provided in Appendix C.

5.2.3 Empirical Implementation

To compute the sample analog of the moments, we need to compute $\underline{\quad}_j(X_m)$ and $\overline{\quad}_j(X_m)$ and estimate the conditional choice probability $\Pr(A_{jm} = 1 \mid X_{jm}; X_{-jm})$. We lay out the steps here.

As a reminder, $\underline{\quad}_j(X_m) = \min_{a_{-jm}} \quad _j(a_{-jm}; X_m)$ and $\overline{\quad}_j(X_m) = \max_{a_{-jm}} \quad _j(a_{-jm}; X_m)$. 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 $\quad _j(a_{-jm}; X_m)$ 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

$$\underline{\quad}_j(X_m) \approx \quad _j((1; :::; 1); X_m)$$

and the maximum by

$$\overline{\quad}_j(X_m) \approx \quad _j((0; :::; 0); X_m):$$

These approximate extrema are exact for the model in Ciliberto and Tamer (2009), where the variable profit function is a linear function of and is decreasing in the entry decisions of other products. 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 F).

As mentioned in Section 2, the tightness of our bounds depends on the gap between $\underline{\quad}_j(X_m)$ and $\overline{\quad}_j(X_m)$. We plot the histogram of the ratio $\underline{\quad}_j(X_m) = \overline{\quad}_j$

Figure 3: Tightness of the Bounds: Histogram of

Table 2: Elasticities: Top-5 Main Brands and Top-5 Craft Brands

	Main Brands					Craft Brands					Markup ($)
	-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
Main Brands	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

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

	CHT	AS
Constant ($_0$)	[2075.12, 3139.05]	[2250.49, 4336.24]
Craft ($_1$)	[1102.09, 2076.87]	[616.18, 2932.42]
In State Craft ($_2$)	[-1173.51, 437.03]	[-2118.48, 627.42]
Distance[a] In State Craft ($_3$)	[1216.18, 3451.17]	[770.51, 5529.83]
Std. Dev. ( )	[1313.70, 2303.08]	[1302.35, 4059.56]

[a]: Distance in 1000KM

in-state and out-of-state craft beer fixed costs, which have a relatively wide projected confidence interval. The result on the distance parameter $_3$ shows that the in-state craft breweries' costs increase in the distance from the brewery to the market. The estimation results 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 average fixed cost, which is around 4500 to 5000 dollars.

7 Counterfactual Results

7.1 Counterfactual Designs

Using the estimated model, we consider a counterfactual merger where the largest firm in our sample (a "macro" brewery) 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 about half of the markets. The merger may raise prices, providing incentives for incumbent product adjustment and new firm entry. We allow firms to change their craft products and 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 ($_{nm}(J_{nm}; J_{-nm})$ in (7)). As for fixed cost shocks, we draw from the estimated distribution while taking into account selection. In other words, our fixed cost draws are consistent with the observed pre-merger equilibrium, which is important to ensure that the per- and post-merger outcomes are comparable (Details on how we draw fixed cost shocks can be found in Appendix D).

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

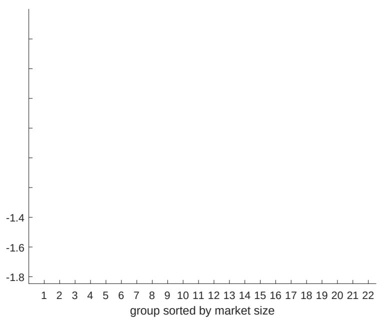
Figure 4: Product Variety, Entry and Prices

(a) Number of Entrants

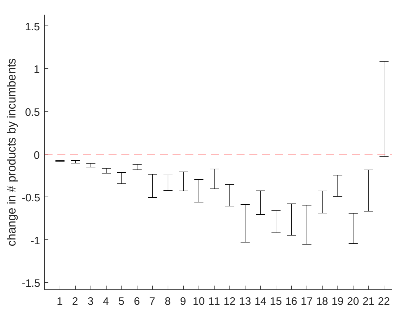
(b) Number of Products Added by Entrants

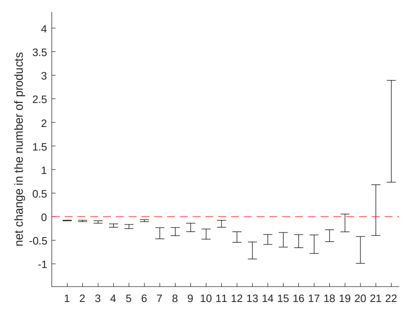
[redacted]

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

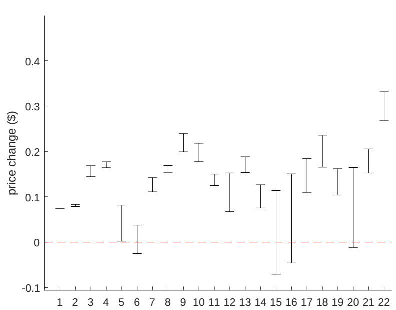


(e) Change in the Number of Products by Incumbent





(g) Change in Craft Prices



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. For the pre-merger craft consumers, however, the average consumer surplus loss is the range of -10 and -35 dollars (Panel (b)).[15] In panel (c), we break out the average consumer welfare effect attributed to the variety changes and find that the product variety effect of the merger recovers some consumer surplus loss in the larger markets, but is either ambiguous or exacerbates the negative price effects in smaller markets. A further decomposition of the product variety effect on consumer welfare indicates that the product variety effect due to new entry is positive (panel (d)) but, consistent with the product changes, often offset by the product variety changes of the incumbents (panel (e)).

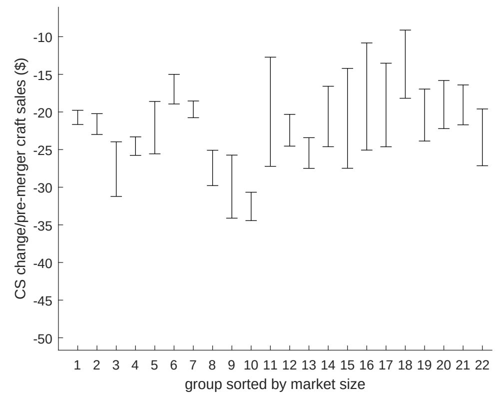
7.2.2 Correlation between the Merger Effects and Market Characteristics

The heterogeneity of the merger effects on product variety and welfare is associated with heterogeneity in market characteristics. For example, 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

Figure 5: Average Consumer Surplus

(a) Change in the Average Consumer Surplus

(b) Change in the Average Craft Consumer Surplus



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

Table 4: 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.

the 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.

7.2.3 Aggregate Merger Effects

Having established the heterogeneity in the merger effects across markets and documenting the correlation between the merger effects and market characteristics, we now turn to the aggregate merger effects across the simulated 149 markets. Similar to Section 7.2.1, we again report the range (across parameter values) of the average merger effects (averaged across simulation draws of the shocks for each parameter value).

In the left panel for Rows (1)–(9), we report the weighted average across markets where the weights are the market sizes. In the right panel for Rows (10)–(20), we report the sum. The 95% confidence interval of the (weighted) average number of expected new entrants is $[0.2, 0.6]$. On average, the merged firms drop between 0.5 to 0.7 products, while the rival incumbents add 0.1 to 0.4 products and new entrants adds 0.2 to 0.7 products. The average beer prices are barely affected, but the merger increases the craft prices by about 15 cents. In comparison, if the variety is held fixed, the average craft price increases by 22 cents (not reported in Table 5). The average price of the merging firms increases by about 40 cents. The decrease in quantity and increase in prices lead to a total welfare loss ranging from 1.4 to 1.8 million dollars, aggregated across the simulated markets. With entry and incumbent product adjustment, the profit of the merging firms changes

Table 5: Aggregate Post-Merger Outcomes

Average Change Per Market			Aggregate Change Across Markets		
(1)	# of rms	[-2.7, -2.3]	(10)	quantity (1000)	[-325.8, -255.8]
(2)	# new entrants	[0.2, 0.6]	(11)	craft	[-302.7, -224.8]
(3)	# of products	[-0.3, 0.4]	(12)	craft, merging rms	[-429.4, -413.1]
(4)	merging rms	[-0.7, -0.5]	(13)	consumer surplus ($1000)	[-1826.0, -1370.3]
(5)	non-merging incumbents	[0.1, 0.4]	(14)	craft beer pro ts ($1000)	[192.4, 307.7]
(6)	new entrants	[0.2, 0.7]	(15)	merging rms	[-3.1, 53.8]
(7)	average price ($)	[0.0, 0.0]	(16)	total surplus ($1000)	[-1518.2, -1175.5]
(8)	craft products ($)	[0.1, 0.2]		CS decomposition ($1000)	
(9)	craft, merging rms ($)	[0.4, 0.4]	(18)	due to variety change	[-162.4, 293.3]
			(19)	due to entry	[124.2, 385.0]
			(20)	due to incumbent product adjustments	[-286.6, -69.4]

across markets, the effects of entry and product variety are positive but account for a minor part 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, 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, Luís 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, 2020.

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.

Draganska, Michaela, Michael Mazzeo, and Katja Seim , "Beyond plain vanilla: Modeling joint product assortment and pricing decisions," QME, 2009, 7 (2), 105–146.

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, Przemysªaw , "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

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.

Wang, Shuang, "Price Competition with Endogenous Entry: The Effects of Marriott & Starwood Merger in Texas," 2020.

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

In this section, we explain how the micro-moment $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$ explained in Section 5.1 is computed. The calculation of the model implication of other micro-moments is similar. The moment $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0\right)$ is the average annual purchase of a household of a certain type of beer conditional on ever purchasing this type of beer in the year, where q_{it}^f is household i's 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.

Let $s_{jmt}(\nu; y)$ denote the Logit choice probability of product j when the vector of unobserved tastes and income are $(\nu; y)$ and $G_m(\nu; y)$ denote the distribution of the unobserved preferences and income, which can vary across markets and is thus indexed by m. We first compute the conditional mean in each market m: $E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0; m\right)$. We assume that each consumer has $R = 8$ opportunities per month to buy beer (where on each trip the consumer buys 1 or 0 products), which is the average number of trips to the stores per household in the Nielsen Consumer Panel data.[16] Then, the probability of purchasing at least one beer of type f in market m conditional on $(\nu; y)$ is

$$\tilde{\pi}_m^f(\nu; y) = 1 - \prod_t \left(1 - \sum_{j \text{ s.t. } x_j^f = 1} s_{jmt}(\nu; y)\right)^R;$$

where the summation $\sum_{j \text{ s.t. } x_j^f = 1}$ is the sum over all products in the choice set in market m, month t where the product is of type f. Therefore, the expected total quantity of purchase conditional buying f at least once in market m is

$$E\left(\sum_{t=1}^{12} q_{it}^f \mid \sum_{t=1}^{12} q_{it}^f > 0; m\right) = \int_{\nu; y} \frac{R \sum_{j \text{ s.t. } x_j^f = 1} s_{jmt}(\nu; y)}{\tilde{\pi}_m^f(\nu; y)} dG_m(\nu; y):$$

16 5 11.2m3 r7d a t0u3(c)estitio ccR

. To obtain the average across markets, we weigh these conditional means in each market 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

$$ {}^{f}_{m} = \int \tilde{\ }^{f}_{m} (\ ; y) \, dG_m (\ ; y) \, ; $$

which implies the total number of households that purchase at least one product of type f in market m is $O_m \cdot \ ^{f}_{m}$, 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

$$ E \left(\sum_{t=1}^{12} q^{f}_{it} \mid \sum_{t=1}^{12} q^{f}_{it} > 0 \right) = \frac{1}{\sum_m O_m \cdot \ ^{f}_{m}} $$]TJeat0\8 3

We then estimate ($_j$; $_m$;) as parameters from the regression

$$_{jmt} = \ _j + \ _m + \ Z_{jm} + \ _{jmt} ;$$

and construct $_{jm}$ from the parameter estimates $\hat{}_j + \hat{}_m + \hat{}Z_{jm}$. We can similarly construct $!_{jm}$.

To simulate the expected profit in (10), we use the empirical distribution of the residuals $\hat{}_{jmt} ; \hat{!}_{jmt}$ from the estimating equations above. The identification of price coe□cients relies on the assumption that product choices are uncorrelated with the transient shocks ($_{jmt}$; $!_{jmt}$). This assumption also is su□cient for the consistency of the procedures above.

We next show that ($_{jmt}$; $!_{jmt}$) 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 of the entire demand data (at the product-market-month level). This is 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 $_{jmt}$ and mc_{jmt} on the other $1 -$ sample. For example, the prediction of $_{jmt}$ is

$$\hat{}_j + \hat{}_m + \hat{}Z_{jm} ;$$

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

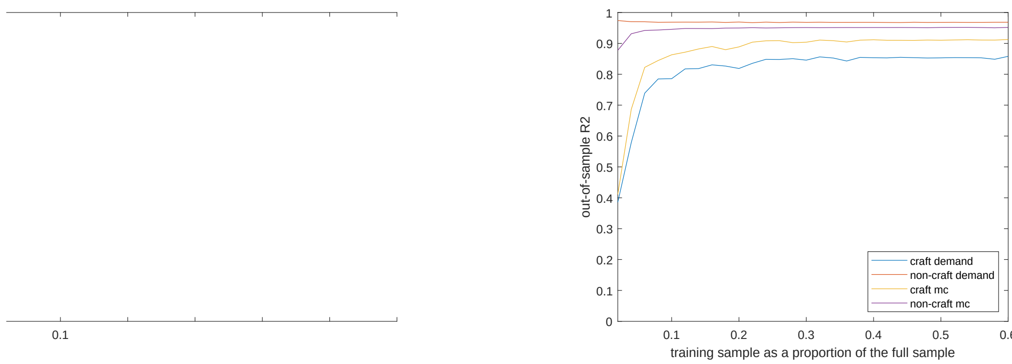
$$_{jmt} - \ \hat{}_j + \hat{}_m + \hat{}Z_{jm} \ :$$

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 di□erent estimates of $_m$ and for craft and non-craft products. The craft products account for 31424of 111219observations in 2016. Using the 2016 estimates, we plot the in- and out-of- sample fit of $_{jmt}$ and mc_{jmt} , where the size of the training data is a proportion of of the full sample, in Figure B.1 (a) and (b) as varies from 0.01 to 0.6. The in-sample R^2 shows that the transient shocks account for no The in-sample

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

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



Common steps:

1 For each potential product/market combination jm, we calculate the corresponding extrema $\underline{\ }_j(X_m)$ and $\overline{\ }_j(X_m)$.

2 We estimate $Pr(A_{jm} = 1 \mid X_{jm}; X_{-jm})$. Recall that we use X_{jm} to represent both the variable profit covariates X_{jm} and the fixed cost covariates W_{jm}: Therefore, the condition variable vector $(X_{jm}; X_{-jm})$ is a potentially high-dimensional vector. Similar to the conditional choice probability approach for estimating dynamic models (e.g., 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 $\overline{\ }_j(X_m); \underline{\ }_j(X_m); W_{jm}$. Some of these variables are continuous and the rest are binary. We use Z^c_{jm} to denote the continuous subvector and Z^b_{jm} the binary subvector. 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 $\tilde{Z}^c_{jm} = \quad \cdot Z^c_{jm}$, where is the Cholesky decomposition of $Var^{-1}\ Z^c_{jm}$ and $(\cdot)$ is the standard normal distribution applied to each element of the vector. We estimate $Pr(A_{jm} = 1 \mid X_{jm}; X_{-jm})$ by the average of A_{jm} across the 20 observations closest to $\tilde{Z}^c_{jm}; Z^b_{jm}$ measured by Euclidean distance.

For inference following CHT, we define a sample objective function $Q(\ )$ as

$$Q(\ ) = \frac{1}{\# jm} \sum_{k=1}^{K} \left[\sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{j \in J_n} h(X_{jm}; X_{-jm};\ ) F^{(k)}(X_{jm}; X_{-jm}) \right]_+^2 ; \qquad (12)$$

where $\# jm$ indicates the number of observations, where an observation is a combination of a potential product of a firm and a market and $F^{(k)}(X_{jm}; X_{-jm}); k = 1; :::; K$ are non-negative exogenous functions that are constructed as follows: we consider the same transformed vectors

Z^c_{jm} and Z^b_{jm} as in Step 2 above. We then divide the set $(0;1)^{|Z^c_{jm}|} \times \{0;1\}^{|Z^b_{jm}|}$ (where Z^c_{jm} and Z^b_{jm} denote the dimensions of Z^c_{jm} and Z^b_{jm}) into $(2r)^{|Z^c_{jm}|} \cdot 2^{|Z^b_{jm}|}$ hypercubes, where each interval $(0;1)$ is divided into $0;\frac{1}{2r}$; $\frac{1}{2r};\frac{2}{2r}$;::: $\frac{2r-1}{2r};1$. We denote the set of hypercubes by $\{C_1;:::;C_K\}$ and construct $F^{(k)}(X_{jm};X_{-jm})$ as indicator functions 1 $Z^c_m;Z^b_m$ C for each hypercube C_k. For inference, we use subsampling to take into account the noise introduced in estimating the conditional choice probability in Step 2.

4(CHT) We construct 200 subsamples and each subsample consists of 20% of the data. For each subsample, we re-estimate the conditional choice probabilities in Step 2 and re-compute the objective function in Step 3 at $\hat{};\hat{}$. We obtain an initial estimate of critical value c_1 as the 95% quantile of Q^r $\hat{};\hat{}$, where Q^r is the objective function on the rth sample.[17]

5(CHT) We construct a set of 10000 candidate parameters by adding shocks to $\hat{}$ and then define $_1 = \{ \quad : Q( ) < c_1\}$.

6(CHT) For each $_1$;we obtain a critical value $c_1( )$ as the 95% quantile of $Q^r( )$ across $r = 1;:::;200$. Define $c_2 = \max_{\ 1} c_1( )$. We construct the confidence set as $\{ \quad : Q( ) < c_2\}$.

For inference following AS, we now use $F^{(k);r}(X_{jm};X_{-jm})$ to denote the non-negative exogenous functions defined above, where ālrol 0 Td [non-ne25r3iu0n Q

This Cramér–von Mises test statistic is a weighted sum of the objective function in (12), with each objective function using a finer definition of the box instruments.

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

$$T^{?}(; \quad) = \sum_{r=1}^{3} \left\{ \frac{(2r)^{-d_x}}{(r^2+100)} \cdot \sum_{k=1}^{K(r)} \left[\frac{\sqrt{\# jm}}{\$^{?}_{kr}} (m^{?}_{kr} - m_{kr}) + \frac{\$_{kr}}{\$^{?}_{kr}} '_{kr} \right]_{+}^{2} \right\};$$

where the superscript ? indicates the quantities calculated on the bootstrapped samples. The shifts of moments $'_{kr}$ is defined in AS. The idea is to reduce the eects of the moments that are not binding at but have large variances.

5(AS) We take the 95% critical value $c_{95}(\ )$ to be the 95% quantile of the bootstrapped statistic. The 95% confidence set is $\{ \quad : T(\ ) < c_{95}(\ ) \}$:

D 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 7, 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 $_{jm}$ as follows. If j is in the market before the merger,

$$\underline{L}_{jm} = -\infty; \overline{H}_{jm} = \_\__j (X_m)$$

   and if product j is not in the market,

$$\underline{L}_{jm} = \_\__j (X_m); \overline{H}_{jm} = \infty:$$

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 $\hat{\ }^2$ and the truncation being $\underline{L}_{jm} < \ _{jm} < \overline{H}_{jm}$. 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 E with the starting points $J^0_{nm} = \quad$ and $J^0_{nm} = J_n$ and check whether the algorithm converges to J_{nm}, holding J_{-nm} fixed. If the algorithm converges to 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 J_{nm}, we go back to Step 2 and re-draw the fixed costs.

4. We repeat this process for every firm n.

E 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 J^0_{nm} J_n, we compute firm n's profit from each of the following deviations from J^0_{nm}: $J^0_{nm} \setminus k$ for k J^0_{nm} or J^0_{nm} k for k $J_n \setminus J^0_{nm}$. Each deviation differs from J^0_{nm} in only one product: either a product in J^0_{nm} is removed or a potential product not in J^0_{nm} is added. There are $\# J_n$ such deviations. Let J^1_{nm} be the highest-profit deviating product portfolio. If firm n's profit corresponding to J^1_{nm} is smaller than that corresponding to J^0_{nm}, this procedure stops and returns J^0_{nm} as the best response. Otherwise, we compute n's profit from any one-product deviation from J^1_{nm} by either adding a potential product to or dropping a product from J^1_{nm}. 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 $2^{\# J_n}$ product portfolios) into one whose action space grows linearly (in each step, evaluating $\# J_n$ portfolios).

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

F 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 $\overline{\ }_j$ and $\underline{\ }_j$ work well; second, we show that our estimation method generally result in conservative but still reasonably tight confidence set 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 J_m in market m. The market share of good j in market m is given by

$$s_{jm} = \frac{\exp(\ _{jm}/(1-\ ))}{\sum_{j \in J_m} \exp(\ _{jm}/(1-\ ))} \frac{\exp\left((1-\ )\sum_{j \in J_m} \exp(\ _{jm}/(1-\ ))\right)}{1+\exp\left((1-\ )\sum_{j \in J_m} \exp(\ _{jm}/(1-\ ))\right)};$$

where is the nesting parameter and the mean utility is $_{jm} = \ p_{jm} + \ x_{jm}$. 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

$$_0 + \ _1 W_{jm} + \ _{jm};$$

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

We set $= -0.5$, $= 0.2$ 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, $W_{jm} \quad N(0, 3)$. 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 $(O; \ ; \ ; \ )$

Table F.1: Monte Carlo Simulations: 2 firms

	O					# products/firm	med $\frac{\overline{\ }_{jm}}{_{jm}}$	Multi Equi %	Cvg Rate
(1)	50000	-5	1.5	1.5	1.5	2	0.6	0.03%	98.00%
(2)	100	0	12	3	1.5	2	0.5	1.25%	100.00%
(3)	50	0.8	15	6	1.5	2	0.3	2.50%	100.00%
(4)	50000	-5	1.5	2.5	0.5	2	0.6	0.03%	98.00%
(5)	50000	-5	1.5	0.5	2.5	2	0.6	0.02%	99.00%
(6)	50	0.8	15	15	1.5	2	0.3	1.40%	100.00%
(7)	50	0.8	15	15	1.5	3	0.2	1.27%	100.00%

set in the last column. The coverage is higher than 95%, which is expected, and indeed decreases in med $\frac{\overline{\ }_{jm}}{_{jm}}$.

We next examine the false coverage probabilities of the projected 95% confidence intervals. Typically authors report the false coverage probability defined as the probability that a confidence set includes a point just outside the identified set (for example, Andrews and Shi (2013)). 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 set tends to be, we instead calculate the probabilities that points around the true parameter values are covered by the 95% confidence interval. In Figure F.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 F.1, where the tightness of the bounds is similar to the empirical setting, and Row (7) of Table F.1, where the bounds are wider. The y-axis of Figure F.1 shows how often a point is included in the confidence interval. When the bounds are tight (med $\frac{\overline{\ }_{jm}}{_{jm}}$ $= 0{:}63$), the results show the projected confidence intervals are tight: for example, in Figure F.1 (1), the vertical line indicates that the true value of the fixed cost intercept is 1.5, and the probability of the confidence interval covering 3.5 or 0.5 is close to 0. When the bounds are more relaxed, the confidence intervals are proportionally wider for (;) (relative to the values of the true parameters). The interval is often twice as wide for $_1$. In both cases, the false coverage probabilities of 0 are reasonably small for all parameters across both specifications.

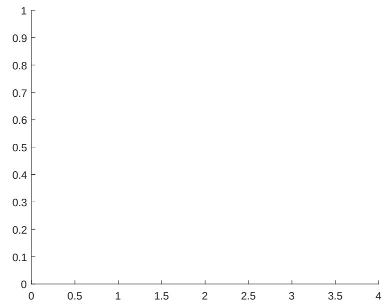
G Market Size-Dependent Fixed Costs

In this section, we present the results where we allow fixed costs to depend on market sizes. Specifi-

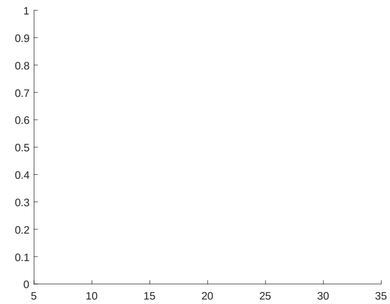
Figure F.1: False Coverage Probability. Left: $\text{med}\left(\frac{|m|}{|m|}\right) = 0.63$; Right: $\text{med}\left(\frac{|m|}{|m|}\right) = 0.18$

FC Intercept

(1) Specification (1)

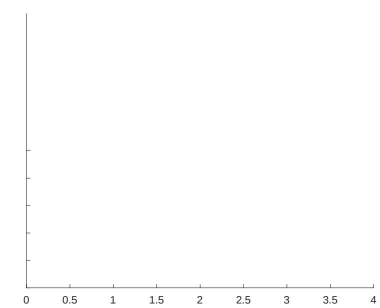


(2) Specification (7)

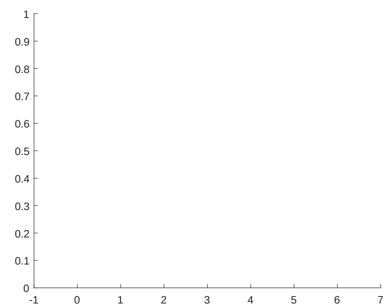


FC Covariate Coefficient

(3) Specification (1)

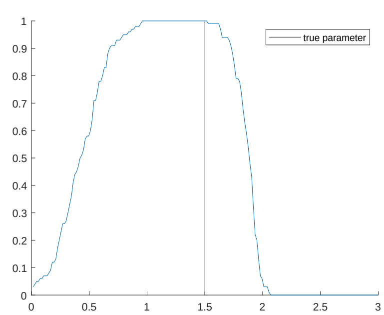


(4) Specification (7)



FC Unobservable Standard Deviation

(5) Specification (1)



(6) Specification (7)

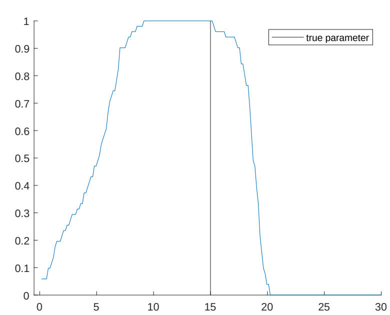


Table G.1: Estimates of Fixed Costs

Craft ($_1$)	[485.88, 2168.68]
In State Craft ($_2$)	[-656.55, 837.75]
Distance[a] In State Craft ($_3$)	[814.51, 2586.09]
Constant ($^k_0$)	
Small Market	[839.50, 2131.35]
Medium Market	[1400.49, 2384.31]
Large Market	[3655.23, 5815.32]
Std. Dev. (^{k})	
Small Market	[90.42, 1218.75]
Medium Market	[471.74, 1305.17]
Large Market	[3465.14, 4917.70]

[a]: Distance in 1000KM

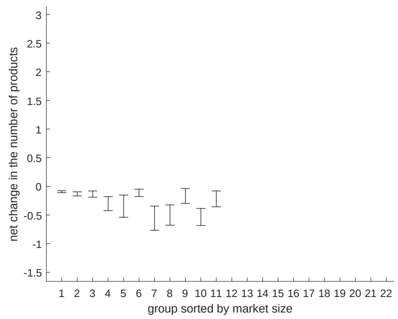
on market sizes:

$$c(W_{jm}\,;\ ) = \ ^k_0 + \ _1 craft_j + \ _2 (in\ state)_j \cdot craft_j + \ _3 (in\ state)_j \cdot craft_j \cdot dist_{jm} + \ ^k_{jm}\,;$$

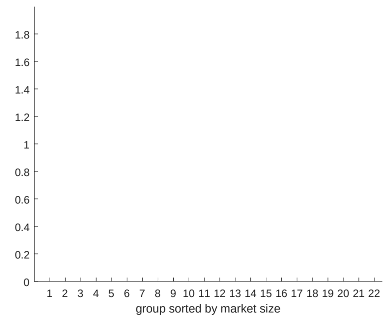
where $k \quad \{small; medium; large\}$ and $^k_{jm} \qquad 0;$

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

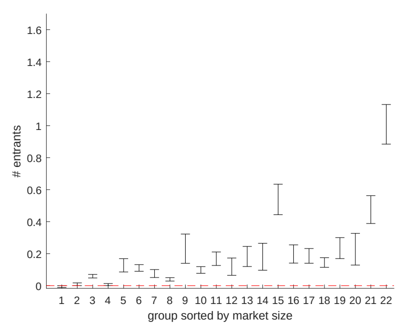
(a) Net Change in the Number of Products



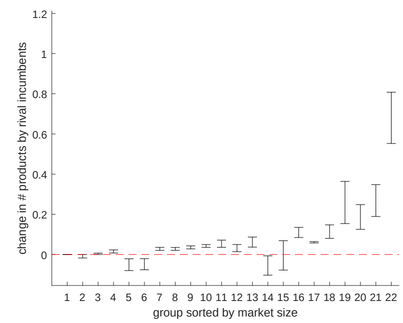
(b) Number of Products Added by Entrants



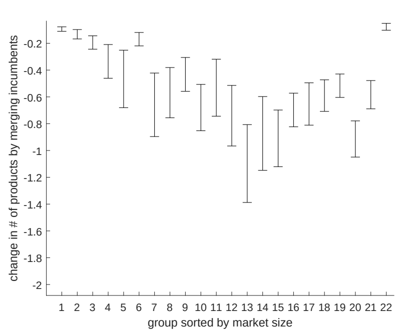
(c) Number of Entrants



(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

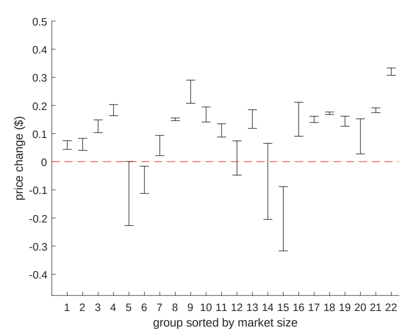
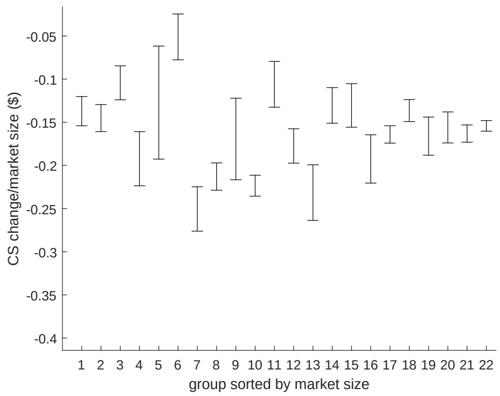


Figure G.2: Average Consumer Surplus

(a) Change in the Average Consumer Surplus



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

The inequality becomes

$$\Pr \quad _{jm} < \underline{\quad}_j (X_m) - c(W_{jm}; \ ) - | \ _4|_+$$
$$\leq \Pr(A_{jm} = 1 \mid X_{jm}; X_{-jm}$$

By the definition of $\bar{\pi}_j(X_m)$, we have

$$\pi_{nm}(a_{nm}; a_{-nm}; X_m) \leq \sum_{j \in J_n} a_{jm} \bar{\pi}_j(X_m).$$

This gives us the upper bound for the probability in (15)

$$\Pr\left(\max_{a_{nm} \text{ s.t. } \#(a_{nm})>0} \left\{ \sum_{j \in J_n} a_{jm} \left(\bar{\pi}_j(X_m) - c(W_{jm};\theta) - \varepsilon_{jm} \right) \right\} - \delta_4 > 0 \,|\, X_{jm}; X_{-jm}; W_{jm}; W_{-jm} \right) \tag{16}$$

Similarly, we can construct the following lower bound:

$$\Pr\left(\max_{a_{nm} \text{ s.t. } \#(a_{nm})>0} \left\{ \sum_{j \in J_n} a_{jm} \left(\underline{\pi}_j(X_m) - c(W_{jm};\theta) - \varepsilon_{jm} \right) \right\} - \delta_4 > 0 \,|\, X_{jm}; X_{-jm}; W_{jm}; W_{-jm} \right) \tag{17}$$

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

$$\Pr\left(\max_{a_{nm} \text{ s.t. } \#(a_{nm})>1} \left\{ \pi_{nm}(a_{nm}; a_{-nm}; X_m) - \sum_{j \in J_n} a_{jm} \left(c(W_{jm};\theta) + \varepsilon_{jm} \right) \right\} - \delta_4 > 0 \,|\, X_{jm}; X_{-jm}; W_{jm}; W_{-jm} \right).$$

The bounds above give us additional moments. These bounds do not have analytical forms. We simulate them as follows. We take NS draws of ε_{jm}'s and index the l^{th} draw for jm combination by $\varepsilon_{jm}^{(l)}$. The simulation for (16) is

$$\frac{1}{NS} \sum_{l=1}^{NS} 1\left(\sum_{j \in J_n} \left(\bar{\pi}_j(X_m) - c(W_{jm};\theta) - \varepsilon_{jm}^{(l)} \right)_+ - \delta_4 > 0 \right).$$

The probability in (17) is simulated analogously. Therefore, the computation is still fast.

Since the bounds are for firm-level entry probabilities, we construct the exogenous non-negative functions $F^{(k)}$ in Appendix C using firm-level variables. Specifically, we first group firms by the number of potential products. Within each group, we use a kmeans algorithm to classify firms by the vector $\left(\tilde{Z}_{jm}^c \right)_{j \in J_n}$ into r classes, which are transformed continuous product characteristics. We use $|J_n|$ to denote firm n's number of potential products, and $\kappa(n) \in \{1, \ldots, r\}$ to denote the kmeans class of firm n. Define

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

where J is the highest number of products per firm. We denote an element of $\mathbf{C}$ as $\tilde{C}_k$ and construct indicator functions $F^{(k)}$ by whether $\left(|J_n|, \kappa(n), Z_{jm}^b\right) = \tilde{C}_k$.

Table H.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 δ_4 is $[-2237.18, -628.20]$. To see what data pattern leads to this estimate, we compute the average $\bar{\pi}_j(X_m)$ and $\underline{\pi}_j(X_m)$ across all jm pairs such that the number of products that firm n has in market m is h. We plot these

Table H.1: Estimates of Fixed Costs: 2016	
Constant ($_0$)	[2163.79, 3598.07]
Craft ($_1$)	[1866.87, 3357.79]
In State Craft ($_2$)	[-1288.92, 669.20]
Distance[a] In State Craft ($_3$)	[-154.33, 2904.70]
Scope($_4$)	[-2237.18, -628.20]
Std. Dev. ( )	[862.97, 2329.04]

[a]: Distance in 1000KM

Figure H.1: Average $\_{j}$ (X_m) and $\bar{\ }_j$ (X_m), Conditional on the Number of Products by a Firm in a Market

averages for $h = 0; 1; :::; 6$ and > 6 in Figure H.1. The figure suggests that for firms that place more 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.

To see the robustness of our results, we repeat the counterfactual simulations in Section (7) using this fixed cost specification and the corresponding confidence set of the parameters. 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 $_4$ applies to each acquired brewery that enters a market. We report results in Figures H.2 and H.3, corresponding with Figures 4 and 5 and 5. The results are similar.

Figure H.3: Average Consumer Surplus

(a) Change in the Average Consumer Surplus

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