Estimating Discrete Games with Many Firms and Many Decisions: An Application to Merger and Product Variety*

Ying Fan†
University of Michigan
CEPR and NBER

Chenyu Yang‡
University of Maryland

May 5, 2022

Abstract

This paper presents a new method for estimating discrete games based on bounds of conditional choice probabilities. The method does not require solving the game and is scalable to models with many firms and many discrete decisions. We apply the method to study merger effects on firm entry and product variety in the retail craft beer market in California. We simulate an acquisition of multiple craft breweries by a large brewery and find that the acquisition would induce firm entry and product entry by non-merging firms. However, these changes are insufficient to offset the negative welfare effects resulting from the higher prices and decreased product offerings by the merging firms.

JEL: D43, L13, L41, L66
Keywords: discrete games, incomplete models, entry, product choice, merger, beer

*We thank Zibin Huang, Sueyoul Kim, and Xinlu Yao for their excellent research assistance and participants at Caltech, Drexel, FTC Microeconomics Conference, ITAM, Johns Hopkins, MIT, NYU IO Day, Penn State, Stanford, Stony Brook, UT Berlin, Yale, and Zhejiang University for their insightful comments. We also thank Ryan Lee, Marc Sorini, and Bart Watson for insights into the craft beer industry. Researchers' own analyses calculated (or derived) based in part on (i) retail measurement/consumer data from Nielsen Consumer LLC ("NielsenIQ"); (ii) media data from The Nielsen Company (US), LLC ("Nielsen"); and (iii) marketing databases provided through the respective NielsenIQ and the Nielsen Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ and 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, Ann Arbor, MI 48109; yingfan@umich.edu.
‡Department of Economics, University of Maryland, College Park, MD 20740; cyang111@umd.edu.

1 Introduction

Discrete games of firm entry or product choice are often used to understand the effects of mergers, divestitures, or industry policies on market structure. In this paper, we consider the estimation of such models when there are many firms or when each firm makes a large set of discrete decisions. For example, a firm may need to choose a set of products to sell in a market from many potential products. In this case, the firm's choice can be represented as a long vector of binary decisions regarding each potential product's entry into the market. However, the estimation of such a model can quickly become challenging as the computational burden of solving the game increases exponentially with the number of firms and firm decisions. In this paper, we propose a computationally tractable estimation method and apply the method to study merger effects on firm entry, product variety, pricing, and welfare in the context of the craft beer market in California.

Our method is based on the bounds for conditional choice probabilities. Consider a binary action $a \in \{0, 1\}$. Assuming no equilibrium action is dominated, we can show that the equilibrium 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. These bounds hold when there is no pure-strategy equilibrium, when there are multiple equilibria, under any equilibrium selection rule, and when the selection rule varies across markets. More importantly, these bounds are easy to compute even with a large number of firms or firm decisions because the bounds can often be reduced to cumulative distribution functions evaluated at certain cutoffs. In this paper, we describe our bounds and provide step-by-step details regarding the estimations and inferences. Using Monte Carlo experiments, we show that our method performs well.

We apply our method to study merger effects on firm entry and product variety. In antitrust litigation, merging parties often argue that the arrival of new entrants mitigates the increased market power resulting from a merger. One assumption behind this argument is that incumbent firms do not change their product offerings. In our paper, we study the effects of mergers by addressing the following questions: Does a merger cause incumbents to add or drop products? Do new firms enter the market after a merger? What is the overall impact of product adjustments and firm entry on welfare? Do any changes in product variety offset the negative price effects on consumer welfare? How do these effects vary across markets?

The US craft beer industry provides an ideal empirical context to study the effects of a merger on the market entry and product variety of multi-product firms. First of all, craft breweries have recently become popular acquisition targets, and these transactions have attracted the attention of antitrust regulators (Codog, 2018). Second, consumer preferences for beers may be highly heterogeneous, implying that product variety is likely an important determinant of consumer welfare in this market. Third, there are rich demographic variations across geographical markets that help to identify consumer tastes and firm costs. In our study, we focus on the state of California, which has the largest number of craft breweries and the highest craft beer production among all US states according to the Brewers Association, a trade group of the craft beer industry.

To address our research questions, we set up a model to describe consumer demand and firm decisions in the retail beer market in California. The demand side is a flexible random coefficient discrete choice model that allows for both observed and unobserved heterogeneity in consumer taste. The supply side is a static two-stage game. In the first stage, each firm is endowed with a set of potential products and chooses the set of products they will sell in a market. If a firm chooses not to enter the market, it chooses the empty set. In the second stage, firms observe demand and marginal cost shocks and choose prices simultaneously.

We use a newly compiled data set to estimate our model. Our main data sources are the Nielsen Retail Scanner Data and Nielsen Consumer Panel from 2010 to 2016. We supplement these data with information on whether a beer is considered a craft beer based on the designation from the Brewers Association. We further augment our data by hand-collecting owner identities and brewery locations.

Our demand estimates reveal substantial unobserved heterogeneity in consumer tastes and little substitution between craft and non-craft products. We obtain these estimates by combining standard macro moments with a new set of micro moments based on the panel structure of the consumer survey data. For example, to identify the dispersion in the unobservable heterogeneity in consumer tastes for craft products, we use the following intuition: if the standard deviation is large, then a household's taste for craft products is highly correlated over time, implying a large expected total purchase of craft beers in a year conditional on the household ever purchasing a craft beer in that year. We back out the marginal costs of beers based on the first-order conditions following the standard approach.

We apply our method based on bounds for conditional choice probabilities to estimate the fixed cost of product entry. This method is well suited to our empirical setting, which features many firms and many potential products. Our empirical setting also features rich market- and product/market-level variations resulting in variations in our bounds. Using the estimated demand and marginal costs to compute the conditions under which selling a product in a market is a dominant or a dominated strategy, we are able to calculate the bounds for the probability that a potential product is observed in a market. Applying our method, we find higher fixed costs of entry for products by independent craft breweries. We also find that both the mean fixed cost of entry and the variance of the fixed cost shock increase with market size.

Using the estimated model, we conduct a counterfactual simulation where the largest macro brewery acquires three large craft breweries. This hypothetical merger case allows us to examine what would happen if the current acquisition trend (i.e., a so-called macro beer firm acquires small craft breweries) continues to the point where the craft beer market becomes as concentrated as the overall beer market. In our simulation, we find that the merger causes new firm entry, which increases product variety. Non-merging incumbents also increase their number of products, at least, in larger markets. However, merging firms drop products. The net changes of variety (due to new entry and incumbents' product adjustments) and associated welfare impacts are both negative. The magnitudes of these effects are heterogeneous across markets, with the largest markets seeing more entries and smaller losses in per-capital consumer surplus. Overall, across markets, new entry occurs, but its positive welfare effect is small relative to the loss of consumer surplus due to reduced product offerings by merging firms. When we consider potential merger efficiencies in the form of a reduction in the fixed costs of the acquired breweries' products, we find that efficiency gains can mitigate but not reverse the overall consumer welfare loss.

Contributions and Literature Review This paper makes two contributions to the literature. First, we develop and evaluate a new method for estimating discrete choice games. Our method can accommodate multiple equilibria, any equilibrium selection rule, the possibility that different equilibrium selection rules are used in different markets, and the possibility of no pure-strategy equilibrium. Moreover, since it requires only the evaluation of one-dimensional cumulative distribution functions and not solving for an equilibrium, it is easily scalable to settings with many firms and firm decisions.

Our method is similar to that of Ciliberto and Tamer (2009), who also estimate a fixed cost shock distribution without specifying equilibrium selection rules. However, our method differs from theirs in the construction of bounds. Ciliberto and Tamer (2009) construct bounds for the probability that an outcome is an equilibrium, where the lower bound is the probability that the outcome is a unique equilibrium and the upper bound is the probability that the outcome is an equilibrium that may not be unique. Computing these bounds requires finding all equilibria to verify an outcome's uniqueness for each draw within a simulated set of fixed cost draws. This can be computationally costly when a model contains many firms and a long vector of decisions for each firm. By contrast, we construct bounds for the entry probability of a single product which are one-dimensional cumulative distribution functions evaluated at certain cutoffs. Computing these cutoffs does not require solving the full game.

Our approach also relates to the methods of estimating discrete games that exploit moment inequalities derived from a necessary equilibrium condition that no firm has an incentive to unilaterally deviate from the observed equilibrium. These approaches typically rely on a mean-zero assumption of non-structural errors (Ho, 2009; Pakes, Porter, Ho and Ishii, 2015; Wollmann, 2018) or support restrictions (Eizenberg, 2014), and do not estimate the distribution of the structural errors associated with the discrete actions. Our approach contributes to this stream of research by estimating the structural error distribution and taking it into account in our counterfactual simulations.

Overall, our approach complements existing methods of estimating discrete games and has advantages when solving for equilibria is costly and when it is important to consider shocks that are known to firms but unobservable to researchers.[1]

Second, we contribute to the literature on mergers, entry response, and product variety. One strand of the literature studies entry defense both theoretically (e.g., Spector, 2003; Anderson, Erkal and Piccinin, 2020; Caradonna, Miller and Sheu, 2021) as well as through simulations and empirical studies (e.g., Werden and Froeb, 1998; Cabral, 2003; Gandhi, Froeb, Tschantz and Werden, 2008; Ciliberto, Murry and Tamer, 2021). We contribute to this strand of the literature by expanding the examination to multi-product firms with endogenous product choice. In our model, firms choose their set of products to sell. Because incumbents can reduce product offerings, it is possible for a merger to decrease product variety while inducing new entry. Another strand within this literature studies how a merger affects product variety and welfare when firm entry is not allowed (e.g., Fan, 2013; Wollmann, 2018; Fan and Yang, 2020; Garrido, 2020; Li, Mazur, Park, Roberts, Sweeting and Zhang, 2022).[2] We contribute to this strand of the literature by jointly studying firm entry responses and incumbent product adjustments after a merger to quantify the net changes in product

[1] There are four other alternative estimation approaches. First, one can obtain a unique equilibrium with additional assumptions and estimate the model via maximum likelihood (Reiss and Spiller, 1989; Garrido, 2020) or a simulated method of moments (Berry, 1992; Li, Mazur, Park, Roberts, Sweeting and Zhang, 2022). Second, Illanes (2017) estimates a dynamic discrete choice problem using a semi-parametric latent variable integration method (Schennach, 2014). This approach also avoids solving a game or an optimization problem, but depends on the availability of certain instruments and can result in relatively wide (and sometimes unbounded) confidence sets of parameters. Third, in their merger simulations, Fan and Yang (2020) make direct assumptions about the distribution of an unobserved fixed cost shock conditional on the observed equilibrium. In comparison, the approach in this paper estimates the distribution. Finally, in a recent paper, Wang (2020) proposes a hybrid approach that combines the Ciliberto and Tamer (2009) bounds with probability bounds based on the concept of dominant strategies. The computational burden of such an approach lies between our method and that of Ciliberto and Tamer (2009).

[2] In the radio industry, a number of papers (e.g., Berry and Waldfogel, 2001; Sweeting, 2010; Jeziorski, 2015) have studied merger effects on firm entry and product variety but do not quantify the impact of a merger on consumer welfare as radio stations do not set prices to listeners. Seim (2006) and Draganska, Mazzeo and Seim (2009) also study entry with endogenous product choice but within the context of an incomplete information framework.

variety. We also investigate the heterogeneity in these responses.

There have also been growing interests in the market structure of the craft beer industry. For example, Tremblay, Iwasaki and Tremblay (2005) document the entry of the microbreweries in the US. Elzinga and McGlothlin (2021) analyze a macro brewery's acquisition of a craft brewery. Bronnenberg, Dubé and Joo (2021) study the formation of preferences for craft beer and the implication for the future market structure of the industry.

The rest of the paper is organized as follows. Section 2 explains our estimation method and presents our Monte Carlo simulation results. Section 3 describes the craft beer market in California and our data. Section 4 presents the empirical model while Section 5 explains the estimation procedure. Section 6 presents the estimation results. Section 7 discusses the counterfactual designs and results. Finally, Section 8 concludes.

2 Discrete Games and Our Estimation Strategy

The estimation of discrete games carries several challenges. First, since there might be multiple equilibria, the maximum likelihood approach may not apply without explicit equilibrium selection rules.[3] Second, a selection issue may complicate a moment inequality approach because the distributions of unobservables conditional on observed actions differ across these actions. We have discussed the existing methods dealing with these issues in the literature review part of the Introduction. In this section, we present our method by starting with a simple model to illustrate how we define our bounds. We then explain our estimation strategy for more general models. We conclude this section with a set of Monte Carlo experiments to evaluate the performance of our method.

2.1 An Illustrative Model and Our Bounds

To illustrate our bounds, we start with a 2×2 model with two firms where each firm makes a single binary decision. We later extend the model to a setting with more firms where each firm makes a vector of binary decisions. In this bivariate model, firms 1 and 2 decide whether to enter market m according to the following:

$$\begin{aligned} Y_{1m} &= \mathbb{1}\left[\pi_{1m}\left(Y_{2m}\right) - C_{1m} - \zeta_{1m} \geq 0\right], \\ Y_{2m} &= \mathbb{1}\left[\pi_{2m}\left(Y_{1m}\right) - C_{2m} - \zeta_{2m} \geq 0\right], \end{aligned} \tag{1}$$

[3] One exception is that Tamer (2003) considers a maximum likelihood estimator in the presence of multiple equilibria for bivariate games without specifying an equilibrium selection rule.

where $Y_{nm} = 1$ indicates entry by firm n in market m, $\pi_{nm}(Y_{-nm})$ is a variable profit function that depends on the rival action Y_{-nm}, C_{nm} is the fixed cost of entry, and ζ_{nm} is a fixed cost shock that follows a distribution, F_ζ.

Our firm behavior assumption is as follows:

Assumption 1. *Y_{nm} is not a dominated strategy for n= 1 or 2.*

In other words, we assume that any observed Y_{nm} is not dominated, which is a weaker assumption than the Nash equilibrium. This level-1 rationality assumption implies the following bounds for $\Pr(Y_{nm} = 1)$:

$$\begin{aligned} & \Pr(Y_{nm} = 1 \text{ is a dominant strategy}) \\ \leq & \Pr(Y_{nm} = 1) \\ \leq & \Pr(Y_{nm} = 1 \text{ is not a dominated strategy}). \end{aligned} \tag{2}$$

Given that $Y_{nm} = 1$ is a dominant strategy if and only if $\zeta_{nm} < \min\{\pi_{1m}(0), \pi_{1m}(1)\} - C_{nm}$, and that $Y_{nm} = 1$ is not a dominated strategy if and only if $\zeta_{nm} < \max\{\pi_{1m}(0), \pi_{1m}(1)\} - C_{nm}$, it follows from (2) that

$$F_\zeta(\min\{\pi_{nm}(0), \pi_{nm}(1)\} - C_{nm}) \leq \Pr(Y_{nm} = 1) \leq F_\zeta(\max\{\pi_{nm}(0), \pi_{nm}(1)\} - C_{nm}).$$

Under the assumption that the entry of the rival reduces a firm's profit, the inequality can be further reduced to

$$F_\zeta(\pi_{nm}(1) - C_{nm}) \leq \Pr(Y_{nm} = 1) \leq F_\zeta(\pi_{nm}(0) - C_{nm}).$$

Before we discuss the general model and estimation, we first compare our bounds with those in the literature and highlight the advantages of our bounds.

Comparison to Bounds in the Literature

Ciliberto and Tamer (2009) Ciliberto and Tamer (2009) (henceforth, CT) assume the outcomes observed in the data are pure-strategy Nash equilibria and construct bounds for the probability of observing an outcome (Y_{1m}, Y_{2m}), denoted by $\Pr(Y_{1m}, Y_{2m})$, as follows:

$$\begin{aligned} & \Pr((Y_{1m}, Y_{2m}) \text{ is a unique pure-strategy Nash equilibrium}) \\ \leq & \Pr(Y_{1m}, Y_{2m}) \\ \leq & \Pr((Y_{1m}, Y_{2m}) \text{ is a pure-strategy Nash equilibrium}). \end{aligned} \tag{3}$$

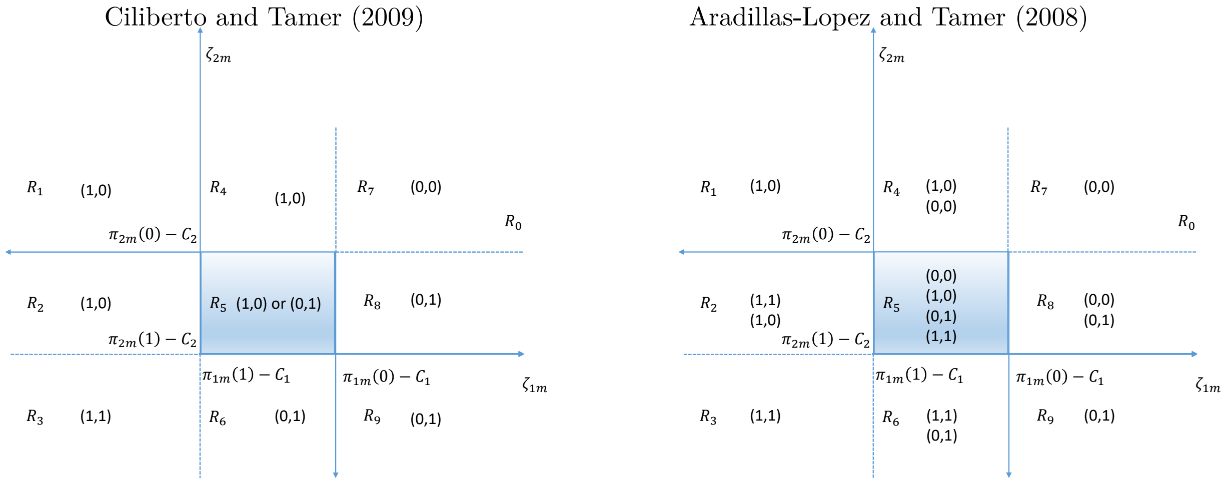

Figure 1: Equilibria Under Different Assumptions

The left graph of Figure 1 lists all possible pure-strategy Nash equilibria in each region of (ζ_{1m}, ζ_{2m}). We use $\Pr(R)$ to represent the probability that (ζ_{1m}, ζ_{2m}) is in region R. We then have the following bounds for the outcome of $(Y_{1m} = 1, Y_{2m} = 0)$:

$$\sum_{\ell=1,2,4} \Pr(R_\ell) \leq \Pr(Y_{1m} = 1, Y_{2m} = 0) \leq \sum_{\ell=1,2,4,5} \Pr(R_\ell), \tag{4}$$

and the following equation for the outcome of $(Y_{1m} = 1, Y_{2m} = 1)$:

$$\Pr(Y_{1m} = 1, Y_{2m} = 1) = \Pr(R_3). \tag{5}$$

Therefore, by adding (4) and (5), we obtain the implied bounds for $\Pr(Y_{1m} = 1)$:

$$\sum_{\ell=1}^{4} \Pr(R_\ell) \leq \Pr(Y_{1m} = 1) \leq \sum_{\ell=1}^{5} \Pr(R_\ell). \tag{6}$$

By contrast, our bounds are

$$\sum_{\ell=1}^{3} \Pr(R_\ell) \leq \Pr(Y_{1m} = 1) \leq \sum_{\ell=1}^{6} \Pr(R_\ell). \tag{7}$$

These bounds are wider than those in (6). This is not surprising given the pure-strategy Nash equilibrium assumed in (6) compared to our weaker assumption of level-1 rationality. However, for a setting with more than two firms or more than a single binary decision, it quickly becomes intractable to partition the space of shocks and find all equilibria in each region. Therefore, one typically simulates the CT bounds by drawing fixed cost shocks (ζ_{1m}, ζ_{2m}) and for each of such draws, enumerating all possible outcomes, and considering whether each outcome is an equilibrium. This procedure could be computationally costly.

Aradillas-Lopez and Tamer (2008) Similar to CT, Aradillas-Lopez and Tamer (2008) also construct bounds for the probabilities of equilibrium outcomes but consider assumptions weaker than Nash, such as the level-1 rationality assumption. The right panel of Figure 1 shows all equilibria consistent with the level-1 rationality assumption. Aradillas-Lopez and Tamer (2008) consider the following bounds for the outcomes $(Y_{1m} = 1, Y_{2m} = 1)$ and $(Y_{1m} = 1, Y_{2m} = 0)$, respectively:

$$\Pr(R_3) \leq \Pr(Y_{1m} = 1, Y_{2m} = 1) \leq \sum_{\ell=2,3,5,6} \Pr(R_\ell), \tag{8}$$

$$\Pr(R_1) \leq \Pr(Y_{1m} = 1, Y_{2m} = 0) \leq \sum_{\ell=1,2,4,5} \Pr(R_\ell),$$

Although their bounds are derived from the same assumption as ours, adding the two inequalities in (8) yields a wider bound for the probability of a single firm's action $\Pr(Y_{1m} = 1)$ compared to our bounds in (7). This is because while their inequalities in (8) focus on the uniqueness of an equilibrium, they do not exploit the uniqueness of a firm's action. For example, although the equilibrium is not unique in R_2, firm 1 always chooses the dominant strategy $Y_{1m} = 1$ in R_2. Furthermore, given that the number of regions increases exponentially with the number of players, one again has to simulate these bounds, which can be computationally costly in scenarios with a large numbers of players or decisions.

Advantages of Our Bounds

There are two advantages of using our bounds for estimating discrete games. First, since our bounds are one-dimensional CDFs, they are easy to compute even in settings with many firms and when each firm makes multiple binary decisions simultaneously (such as product portfolio decisions). Second, our bounds do not rely on equilibrium selection assumptions. Specifically, they hold when there are multiple equilibria, when the equilibrium selection mechanisms differ across markets, and when there is no pure strategy equilibrium for some values of fixed cost shocks.

Our bounds are also intuitive. In a single-agent binary choice model, the inequalities collapse into an equality used in the standard GMM estimator (McFadden, 1989). Thus, our approach can be considered an extension of the GMM estimation of binary choice models to a game setting.

2.2 General Models and Estimation Using Our Bounds

In this section, we describe a general model and explain how to estimate the model using our bounds. We consider M markets and N firms. Each firm n makes a vector of binary decisions, Y_{nm}, in each market m. For example, firms decide whether to enter a market and, if so, which subset of products from a potential set of products to sell. In this setting, $Y_{nm} = (Y_{jm}, j \in \mathcal{J}_n)$, where $\mathcal{J}_n$ is the set of potential products for firm n and $Y_{jm} \in \{0, 1\}$ indicates whether product j is in market m. Let $\mathcal{J} = \cup_n \mathcal{J}_n$. We use $Y_m = (Y_{jm}, j \in \mathcal{J})$ to denote all firm decisions in the market and $\pi_n(Y_m, X_{nm})$ to denote the variable profit function of firm n, which depends on both its own decisions and its rivals' decisions in the market as well as a set of observable covariates, X_{nm}. This function is either known or has been estimated. For example, it can be the variable profit function at a Nash-Bertrand price equilibrium given separately estimated demand and marginal costs. We further assume that there is a cost associated with choosing $Y_{jm} = 1$. This cost is $c(W_{jm}, \theta) + \zeta_{jm}$, where W_{jm} is a set of exogenous covariates and θ is a vector of parameters to be estimated. The unobserved cost shock, ζ_{jm}, is assumed to be i.i.d. and follow the distribution $F_\zeta(\cdot, \sigma_\zeta)$, where σ_ζ represents the distributional parameters to be estimated.

We define the change in a firm's variable profit when Y_{jm} turns from 0 to 1:

$$\varDelta_j(Y_{-jm}, X_{nm}) = \pi_n(Y_{jm} = 1, Y_{-jm}, X_{nm}) - \pi_n(Y_{jm} = 0, Y_{-jm}, X_{nm}), \tag{9}$$

where $Y_{-jm} = (Y_{j'm}, j' \in \mathcal{J}, j' \neq j)$ represents the product entry outcome in market m regarding all potential products except product j. Given the discrete nature of Y_{-jm}, the following minimum and maximum changes in variable profits exist: $\underline{\varDelta}_j(X_{nm}) = \min_{Y_{-jm}} \varDelta_j(Y_{-jm}, X_{nm})$ and $\overline{\varDelta}_j(X_{nm}) = \max_{Y_{-jm}} \varDelta_j(Y_{-jm}, X_{nm})$.

Following the discussion in the previous section, we can see that the bounds of the conditional probability $Y_{jm} = 1$ given X_{nm} and W_{jm} are:

$$\begin{aligned} & F_\zeta\left(\underline{\varDelta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta\right) \\ & \leq \Pr(Y_{jm} = 1 \,|X_{nm}, W_{jm}) \\ & \leq F_\zeta\left(\overline{\varDelta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta\right). \end{aligned} \tag{10}$$

We define the following moment functions:

$$\begin{aligned} L(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) &= F_\zeta\left(\underline{\varDelta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta\right) - \mathbb{1}(Y_{jm} = 1), \\ H(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta) &= \mathbb{1}(Y_{jm} = 1) - F_\zeta\left(\overline{\varDelta}_j(X_{nm}) - c(W_{jm}, \theta), \sigma_\zeta\right). \end{aligned} \tag{11}$$

The inequalities in (10) imply the following conditional moment inequalities:

$$E\left(L\left(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta\right)|X_{nm}, W_{jm}\right) \le 0, \quad (12)$$
$$E\left(H\left(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta\right)|X_{nm}, W_{jm}\right) \le 0.$$

Following the literature on inference based on conditional moment inequalities, we transform these conditional moment inequalities into unconditional ones:

$$E\left(\frac{1}{\#\mathcal{J}}\sum_{j\in\mathcal{J}} L\left(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta\right)\cdot g^{(k)}\left(X_{nm}, W_{jm}\right)\right) \le 0, \quad (13)$$
$$E\left(\frac{1}{\#\mathcal{J}}\sum_{j\in\mathcal{J}} H\left(Y_{jm}, X_{nm}, W_{jm}, \theta, \sigma_\zeta\right)\cdot g^{(k)}\left(X_{nm}, W_{jm}\right)\right) \le 0,$$

where $g^{(k)}, k = 1, ..., K$ are non-negative functions of (X_{nm}, W_{jm}) that capture the information contained within the conditioning variables. In (13), we average over potential products and exploit variations across markets. We do so because, even under the independence assumption related to the fixed cost shock ζ_{jm} across j, Y_{jm} across j within the same market m may be correlated due to strategic interdependence among firms. However, the entry decisions Y_{jm} are independent across markets.

We provide step-by-step details on how we carry out inferences in both Section 2.3 for our Monte Carlo experiments and Section 5.2 for our empirical application.

Identification The identification of (θ, σ_ζ) based on the inequalities in (10) is similar to the idea of special regressors in entry games (Ciliberto and Tamer, 2009; Lewbel, 2019). To identify our parameters, we exploit exogenous variations in X_{nm} and W_{jm}. For example, to identify the coefficient in θ that corresponds to an indicator variable, we first define a non-negative function $g^{(k)}$ equal to the indicator and then compare the entry probability conditional on this indicator being 1 versus 0, holding other covariates fixed. Similarly, to identify the coefficient of a continuous variable, we first define a set of non-negative indicator functions that indicate whether the value of the continuous variable falls within a certain range and then examine how entry probability varies across these ranges. Exogenous variations in X_{nm} and W_{jm} also allow us to identify the distribution parameters σ_ζ. For example, consider a distribution of ζ_{jm} that is fully specified by its variance. If the variance is large, the upper and lower bounds will show little to no co-variance with the covariates. In the special case where ζ_{jm} follows a symmetric distribution, both bounds in (10) approach 0.5 as the variance increases. On the other hand, when the variance is close to 0, both bounds are close to 0 if $\overline{\varDelta}_j\left(X_{nm}\right) - c\left(W_{jm}, \theta\right) < 0$ or close to 1 if $\underline{\varDelta}_j\left(X_{nm}\right) - c\left(W_{jm}, \theta\right) > 0$, by Chebyshev's Inequality. Therefore, with a very small variance, the model will predict large jumps in entry probabilities even with small changes in the covariates. Both the lack of sensitivity of entry probability to covariates (in the case of a large variance of ζ_{jm}) and the high sensitivity (in the case of a small variance) can be tested by data.

Extensions Here, we assume that the total cost associated with a vector of binary decisions is the sum of the cost associated with each decision. In Supplemental Appendix D, we extend our method to estimate a model allowing for economies or dis-economies of scope. We also assume here that the unobservable cost shock is i.i.d. In Supplemental Appendix E, we extend our method to estimate correlations in unobservable cost shocks.

2.3 Monte Carlo Experiments

We use Monte Carlo experiments to evaluate the performance of our estimation method. In this section, we first explain the setup of our Monte Carlo experiments. We then provide step-by-step details about our inference procedure, examine the probability that our 95% confidence set covers parameters different from the true parameter values, and compare both the performance and the computational burden of our method with those of CT's.

2.3.1 Monte Carlo Experiment Setup

In our Monte Carlo experiments, we consider an entry game with N potential entrants. Each firm n makes a binary decision $Y_{nm} \in \{0, 1\}$, where $Y_{nm} = 1$ represents entering market m. Firm n's post-entry variable profit is:

$$\pi_n\left(Y_{-nm}, X_{nm}\right) = O_m \cdot \prod\nolimits_{r \neq n} x_{rnm}^{Y_{rm}},$$

where $Y_{-nm} = (Y_{rm}, r \neq n)$ represents rival firm entry decisions, $O_m > 0$ represents a market-level profit shifter such as market size, and $x_{rnm} \in (0, 1)$ represents the competitive impact of firm r's entry on firm n's profit. We collect covariates in $X_{nm} = \left(\left(x_{rnm}, r \neq n\right), O_m\right)$.

The fixed cost of entry is $C + \sigma\zeta_{nm}$, where the unobservable cost shock ζ_{nm} is assumed to be a standard normal random variable and i.i.d across both firms and markets. The mean fixed cost parameter C and the standard deviation σ are the parameters to be estimated.

In our Monte Carlo experiments, we draw O_m from a uniform distribution between 0 and 2, and x_{rnm} from a to 1. As we explain later, varying $a \in (0, 1)$ changes the tightness of our bounds. We set $C = \sigma = 1$. For each draw of $(X_{nm}, n = 1, ..., N)$, we compute the Nash equilibrium. In the case of multiple equilibria, the equilibrium with the highest total profit is selected. For each Monte Carlo experiment, we simulate 500 data sets.

2.3.2 Inference

Since the profit function π_n decreases in the entry decision of a firm's rivals, we have

$$\min_{Y_{-nm}} \pi_n \left(Y_{-nm}, X_{nm}\right) = \pi_{nm} \left(\left(1, \ldots, 1\right), X_{nm}\right) = O_m \cdot \prod\nolimits_{r \neq n} x_{rnm},$$
$$\max_{Y_{-nm}} \pi_n \left(Y_{-nm}, X_{nm}\right) = \pi_{nm} \left(\left(0, \ldots, 0\right), X_{nm}\right) = O_m.$$

Based on (10), we use the following bounds in our estimation:

$$\Phi\left(\left[O_m \cdot \prod\nolimits_{r \neq n} x_{rnm} - C\right] / \sigma\right) \leq \Pr\left(Y_{nm} = 1 \left| X_{nm}\right.\right) \leq \Phi\left(\left[O_m - C\right] / \sigma\right),$$

where $\Phi(\cdot)$ is the standard normal distribution function. Note that the gap between the lower and upper bounds depends on the difference between $O_m \cdot \prod_{r \neq n} x_{rnm}$ and O_m. Therefore, as we vary a and thus the range from which we draw x_{rnm}, we tighten or widen the bounds.

Moments

In our experiments, we specify the following moment functions:

$$L\left(Y_{nm}, X_{nm}, C, \sigma\right) = \Phi\left(\left[O_m \cdot \prod\nolimits_{r \neq n} x_{rnm} - C\right] / \sigma\right) - \mathbb{1}\left(Y_{nm} = 1\right),$$
$$H\left(Y_{nm}, X_{nm}, C, \sigma\right) = \mathbb{1}\left(Y_{nm} = 1\right) - \Phi\left(\left[O_m - C\right] / \sigma\right).$$

The conditional moment inequalities are:

$$E\left[L\left(Y_{nm}, X_{nm}, C, \sigma\right) \left| X_{nm}\right.\right] \leq 0,$$
$$E\left[H\left(Y_{nm}, X_{nm}, C, \sigma\right) \left| X_{nm}\right.\right] \leq 0.$$

We transform the above moment inequalities into unconditional ones as follows:

$$E\left[\frac{1}{N} \sum_n L\left(Y_{nm}, X_{nm}, C, \sigma\right) \cdot g^{(k)}\left(X_{nm}\right)\right] \leq 0,$$
$$E\left[\frac{1}{N} \sum_n H\left(Y_{nm}, X_{nm}, C, \sigma\right) \cdot g^{(k)}\left(X_{nm}\right)\right] \leq 0,$$

where $g^{(k)}\left(X_{nm}\right)$ for $k = 1, \ldots, K$ is a non-negative function of X_{nm}, defined as:

$$g^{(k)}\left(X_{nm}\right) = \mathbb{1}\left(O_m \geq b_1^{(k)},\ O_m \cdot \prod\nolimits_{r \neq n} x_{rnm} \geq b_2^{(k)}\right),$$

where $\left(b_1^{(k)}, b_2^{(k)}\right) = (0.5, 0.5), (0.5, 1), (1, 0.5)$, or $(1, 1)$ for $k = 1, ..., 4$.

Confidence Set

We construct our confidence set by inverting the test in Chernozhukov, Chetverikov and Kato (2019) (CCK), which does not require a tuning parameter for its one-step critical value.[4] The CCK test statistic is based on the maximum of t-type statistics that correspond with each moment. Specifically, for each market $m = 1, \ldots, M$, we define the vector

$$Z_m(C,\sigma) = \left(\frac{1}{N} \sum_n L\left(Y_{nm}, X_{nm}, C, \sigma\right) \cdot g^{(k)}\left(X_{nm}\right), \ \frac{1}{N} \sum_n H\left(Y_{nm}, X_{nm}, C, \sigma\right) \cdot g^{(k)}\left(X_{nm}\right) \right)_{k=1}^{K}.$$

We use $Z_{\tilde{k}m}(C,\sigma)$ to denote a component of the vector, where $\tilde{k} = 1, \ldots, 2K$. We denote the sample mean and sample standard deviation of the sample moment $\tilde{k}$ as:

$$\hat{\mu}_{\tilde{k}}(C,\sigma) = \frac{1}{M} \sum_{m=1}^{M} Z_{\tilde{k}m}(C,\sigma) \text{ and } \hat{\sigma}_{\tilde{k}}(C,\sigma) = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \left(Z_{\tilde{k}m}(C,\sigma) - \hat{\mu}_{\tilde{k}}(C,\sigma)\right)^2}.$$

The test statistic is given by:

$$\max_{1 \le \tilde{k} \le 2K} \frac{\sqrt{M} \hat{\mu}_{\tilde{k}}(C,\sigma)}{\hat{\sigma}_{\tilde{k}}(C,\sigma)}. \tag{14}$$

The critical value at the significance level of α is:

$$\frac{\Phi^{-1}\left(1 - \alpha/2K\right)}{\sqrt{1 - \Phi^{-1}\left(1 - \alpha/2K\right)^2 / M}}.$$

2.3.3 Monte Carlo Experiment Results

To evaluate the performance of our method, we first present the coverage probability that our 95% confidence set contains parameter values of C and σ in the neighborhood of the true parameter value (i.e., $(C,\sigma) = (1,1)$). Specifically, we consider candidate parameter values ranging from 0 to 2. We then compare the statistical power and computational times of the CCK test statistic based on our bounds versus the CT bounds.

To compute the coverage probability, we simulate 500 data sets with $N = 2$ and $M = 2000$. For each candidate parameter value, the coverage probability is the fraction of the 500 data sets where the test statistic evaluated at this candidate parameter value is below the critical value at the 5% significance level. We repeat this exercise for three different values of a: 1/3, 1/2, and 3/4, where $(a, 1)$ is the range for the covariate x_{rnm}. As explained above, this range determines the size of the gap between our lower and upper bounds.

For visibility, we plot the coverage probabilities for C and σ separately in Figure 2. For

[4] The confidence set based on CCK's two-step critical value is very similar.

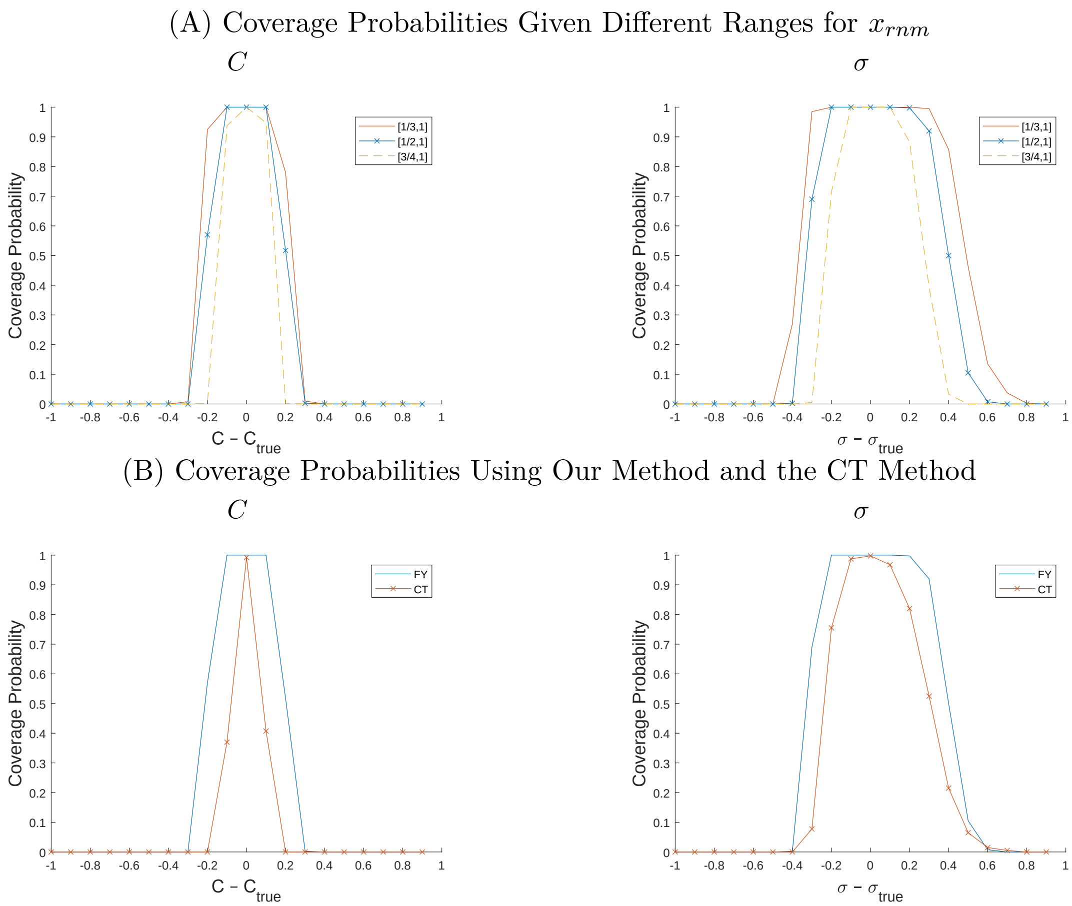

Figure 2: Coverage Probabilities

example, for the coverage probability of $C = 0.8$, we fix the value of C at 0.8, minimize the test statistic with respect to σ, and report the fraction of the data sets where there exists a value of σ such that the test statistic falls below the critical value. We do so for candidate values of C from 0 to 2 and analogously for candidate values of σ from 0 to 2.

Figure 2(A) shows that the coverage probability for a parameter value in the neighborhood of the true parameter value is smaller when the covariate x_{rnm} is drawn from a narrower range and that this probability decreases quickly for parameter values further away from the true values, even when x_{rnm} is drawn from the widest range.

Figure 2(B) compares the coverage probabilities using our method versus the CT method again for $C \in [0, 2]$ and $\sigma \in [0, 2]$.[5] We use the simulated data sets based on $a = 1/2$ for this comparison. The figure shows that the coverage probabilities using the CT method are moderately smaller than those based on our method (labeled FY).

[5]To construct the CCK estimator using the CT bounds, we use the same $g^{(k)}$ functions given in Section 2.3.2 to make the estimates based on the two methods comparable. For each possible outcome of (Y_{1m}, Y_{2m}), we interact the moment function corresponding to the lower bound of $\Pr(Y_{1m}, Y_{2m})$ with $g^{(k)}(X_{1m})$ as well as $g^{(k)}(X_{2m})$. We do the same for the upper bound.

Table 1: Comparison to CT Bounds: Computational Time

N (#Firms)	M (#Markets)	FY (s)	CT (s)
2	2000	0.0003	0.0139
5	800	0.0003	0.0469
10	400	0.0002	1.3606
15	267	0.0004	43.0739

Notes: the table reports the computational time to evaluate the CCK test statistic corresponding with our bounds (FY) and Ciliberto and Tamer (2009) (CT) bounds.

However, our method comes with a computational advantage which grows exponentially with the number of firms in a game. To compare the computational burden of the two methods, we report the time needed to evaluate the test statistic once using our bounds versus the CT bounds in Table 1.[6] From the table, we can see that when $N = 2$ and $M = 2000$, the computation time for evaluating the test statistic once based on our FY bounds is about 1/40th of the time based on the CT bounds, although both methods seem relatively fast when there are only two firms in the model. To see how the computational advantage of our method varies with the number of firms, we increase the number of firms from 2 to 15. We simultaneously decrease the number of markets to keep the number of firm-market combinations roughly constant in order to rule out mechanical increases in computation time. From Table 1, we can see that the computation time using the FY bounds remains stable from the first row with $(N = 2, M = 2000)$ to the last row with $(N = 15, M = 267)$. This is not surprising because evaluating our moment functions only involves evaluating one-dimensional CDFs. By comparison, the computation time using the CT method increases from 0.0139 seconds to 43 seconds because it requires checking all possible market outcomes to find all equilibria and the number of all possible market outcomes increases exponentially as the number of firms increases.

Overall, the results from our Monte Carlo experiments indicate that our method performs well and is a good alternative to existing methods when the number of players is large.

3 Empirical Background and Data

We apply our method to study merger effects on firm entry and product variety in the retail craft beer market in California. According to the 2015 Brewers Association estimates, California accounted for 18% of craft beer volume and 12% of craft breweries in the nation,

[6]The results are computed on an Intel Core i7-8700K Processor (3.70GHz). We use 100 simulation draws of (ζ_{1m}, ζ_{2m}) to simulate the CT bounds.

the highest among all US states. California has its own tied-house laws that expand on federal statues prohibiting "tied-houses".[7] In addition, California passed competition laws that further prohibit payments for stocking products (Croxall, 2019). These institutional features motivate our assumption that breweries make their own entry and product variety decisions. This simplification keeps our model tractable.

Our analysis is based on a new data set that we compiled from various sources. The primary data sets are the market-level data in the Nielsen Retail Scanner Data and the micro-level panel data in the Nielsen Consumer Panel between 2010 and 2016. We define a product to be a brand in the Nielsen data (e.g., Samuel Adams Boston Lager). We aggregate the Nielsen scanner data from its original UPC/week level into a product/month-level data set by homogenizing the size of a product (a unit represents a 12-ounce-12-pack equivalent), adding quantities across weeks within a month, and using the quantity-weighted average price across weeks within a month as a given product's price in a given month. We then supplement the data set with information on whether a beer is considered a craft beer based on the designation by the Brewers Association. We also add hand-collected data on the identities of the corporate owner and the brewery as well as the location of the production facility for each product in our data set. For example, Samuel Adams Boston Lager is owned by Boston Beer Company and produced at Samuel Adams Boston Brewery in Boston. We define a firm as a corporate owner (e.g., Boston Beer Company). A firm can own multiple breweries and products. Finally, we merge the data with county demographics obtained from the US Census.

In our analyses, we define a market as a retailer-county pair. The Nielsen consumer panel data suggest that cross-retailer shopping is rare. For example, we find that more than 80% of the households in our study purchase all of their beer from one retailer-county combination in 2016. This finding is consistent with those in Huang, Ellickson and Lovett (2021) and Illanes and Moshary (2020) who find little evidence of retailer competition in the spirits category. In our estimation, we define 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 per month in the panel data.[8]

We consider a product to be available in a market in a calendar year if the product's monthly sales are more than 20 units for more than 6 months in the market in the year. Moreover, for craft products, we keep those produced by the top 60 craft breweries according to their national volumes in the 2015 Brewers Association production data. We thus focus on breweries established in the 1990s or earlier. In the end, our sample covers 83% of California's

[7]"Tied-houses" refer to vertical relationships between manufacturers and retailers that exclude small manufacturers such as craft breweries from placing their products with retailers.

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

Table 2: Annual Total Quantity, Prices, and Numbers of Firms and Products

	Total Quantity (12 pk equiv)	Avg. Price (2016 $)	# Firms Per Year	# Products Per Year
Craft	4,914,209	17	36	135
All	53,465,658	11	54	269

Notes: for each year from 2010 to 2016, we first calculate a year's total and craft beer quantities, quantity-weighted average prices, number of firms and number of products, and then take the average across years.

Table 3: Shares of Total Quantity and Number of Products by Beer Types

		Ale	Lager	Light
Quantity				
	Craft	71.53%	27.13%	0.44%
	All	12.50%	46.42%	40.02%
Number of products				
	Craft	66.33%	26.87%	0.41%
	All	44.19%	39.76%	7.23%

Notes: the shares reflect the respective proportions of the total quantity of beer from 2010 to 2016, or of the total number of unique products in these years.

craft beer quantity in the Nielsen Scanner Data across our sample periods.[9]

We define a firm's set of potential products in a year as all products owned by the firm available in any market in the year. Note that we do not consider new brewery or brand creation but rather focus on a firm's decision to sell an existing product in a market, a decision far less costly than a *de novo* entry. Therefore, our setting can be considered favorable for firm or product entry. As we see later, even in this favorable setting where new entrants bring in new products, the net merger effect on product variety is negative.

Table 2 reports the summary statistics based on 110 markets present in the data every year from 2010 to 2016. These markets account for 82% of the total quantity from all markets and years. From Table 2, we see that the annual craft beer sales in the sample are, on average, about 5 million units (12-ounce-12-pack), accounting for about 10% of the total beer sales for a given year. We can also see that the average craft beer price is around 17 dollars per unit (in 2016 dollars), which is higher than the average beer price of 11 dollars per unit. Although craft beers account for around 10% of the total sales, the number of craft firms and craft products account for over half of the market.

[9]Although our retail data precludes a direct comparison of the retail beer market with the "on-premises" market (such as taprooms, bars, and restaurants), the Brewers Association suggests that the retail channel accounts for 65% of craft beer volume (Watson, 2016). Likely due to similar data limitations, prior studies on the beer industry have also focused on the retail segment (Ashenfelter, Hosken and Weinberg, 2015; Asker, 2016; Miller and Weinberg, 2017; Miller, Sheu and Weinberg, 2019).

Table 3 provides a breakdown of the sales and number of products by beer types. Among craft products, ales account for 66% of the product counts and 72% of sales. Lagers account for 27% (46%) of craft (overall beer) market share. Finally, while light beers account for 40% of the overall beer sales, their market share within craft products is only 0.44%.

A key primitive in the product variety decisions is the fixed cost of product entry. According to our interviews with industry experts,[10] the main cost of product entry is a flow cost of the marketing support that a firm needs to provide to a retailer in a local market. By contrast, the sunk cost of convincing a retailer to carry a brewery's products or contracting with a distributor seems negligible compared to the fixed cost of marketing support. For the craft products studied in this paper, it is illegal at both the federal and state level and extremely rare for grocers or distributors to charge slotting fees.

4 Model

4.1 Demand

We use a random coefficient discrete choice model to describe consumer demand for beer. A product's characteristics include its flavor type (ale, lager, light, and others),[11] whether it is imported from outside North America, and whether it is designated as a craft product. Note that these characteristics can overlap. For example, Bud Light is a light, North American, non-craft beer, while Samuel Adams Lager is a lager, North American, craft beer. These characteristics of product j are captured by a vector of indicator variables $\boldsymbol{x}_j = \left(x^{\text{ale}}, x_j^{\text{lager}}, x_j^{\text{light}}, x_j^{\text{import}}, x_j^{\text{craft}}\right)$. We allow both household income and unobservable heterogeneity to affect preferences. We specify the utility function of household i in market m from product j in month t as

$$\begin{aligned} u_{ijmt} = {} & (\sigma_0 \nu_i + \kappa_0 y_i) + (\alpha + \kappa_\alpha y_i)\, p_{jmt} \\ & + \sigma^{\text{ale}} \nu_i^{\text{ale}} x_j^{\text{ale}} + \sigma^{\text{lager}} \nu_i^{\text{lager}} x_j^{\text{lager}} + \sigma^{\text{light}} \nu_i^{\text{light}} x_j^{\text{light}} \\ & + \sigma^{\text{import}} \nu_i^{\text{import}} x_j^{\text{import}} + (\sigma^{\text{craft}} \nu_i^{\text{craft}} + \kappa^{\text{craft}} y_i) x_j^{\text{craft}} \\ & + \beta X_{jm} + FE_j^{\text{demand}} + FE_m^{\text{demand}} + FE_t^{\text{demand}} + \xi_{jmt} + \varepsilon_{ijmt}, \end{aligned} \tag{15}$$

where y_i is the logarithm of household i's annual income and $\nu_i^{(\cdot)}$ is the household-specific unobserved taste shock, which follows a normal distribution and is independent across households. Therefore, the σ parameters capture the dispersion in unobserved household tastes

[10] They are the Chief Economist Bart Watson and General Counsel Marc Sorini at the Brewers Association.

[11] Some examples in the category of "others" include stout, porter, and near beers, which collectively account for 0.9% of the craft quantities.

while the κ parameters measure the effect of household income on tastes and the price coefficient. Note that we do not include mean coefficients for $\boldsymbol{x}_j$ because they are absorbed in product fixed effects. The covariates X_{jm} represent a set of indicator functions for whether the distance from the brewery's nearest production facility to the market falls within a certain distance range. Distance potentially plays an important role in demand as a local beer may lack name recognition outside of its local market (see, for example, Tamayo, 2009). We also include in our model product fixed effects FE_j^{demand}, market fixed effects FE_m^{demand}, and month fixed effects FE_t^{demand} to capture unobserved factors that may vary at these levels. The error term ξ_{jmt}, therefore, captures the transient, month-to-month variations of demand shocks specific to a product, market, and month combination. Finally, the last term in (15), ε_{ijmt}, is a household's idiosyncratic taste, which is assumed to be i.i.d. and follows a type-1 extreme value distribution.

Overall, our demand specification gives us the market share $s_{jmt}\left(p_{jmt}, p_{-jmt}\right)$ 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 then gives us the demand for product j, $D_{jmt}\left(p_{jmt}, p_{-jmt}\right)$.

4.2 Supply

We model the supply side as a two-stage static game. In each market, firms simultaneously choose which beer products, if any, to sell. This product choice is made at the beginning of each year τ and is fixed throughout the year. We use $\mathcal{J}_{nm\tau}$ to denote 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 the retail prices for their products.

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

$$\max_{p_{jmt}, j \in \mathcal{J}_{nm\tau}} \sum_{j \in \mathcal{J}_{nm\tau}} \left(p_{jmt} - mc_{jmt}\right) D_{jmt}\left(p_{jmt}, p_{-jmt}\right). \tag{16}$$

The marginal cost mc_{jmt} is decomposed into a product fixed effect FE_j^{mc}, a market fixed effect FE_m^{mc}, a month fixed effect FE_t^{mc}, the effect of facility-market distance γX_{jm} to account for any transportation cost, and a product-market-month specific shock ω_{jmt}:

$$mc_{jmt} = FE_j^{\text{mc}} + FE_m^{\text{mc}} + FE_t^{\text{mc}} + \gamma X_{jm} + \omega_{jmt}. \tag{17}$$

For computational simplicity, we assume that firms directly set the retail prices for their products. In reality, breweries publish a price menu every few months, and distributors purchase the products and sell them to the retailers. We include product-, market-, and month-specific fixed effects in our specification of the brewery marginal cost to capture the distributor and retailer markups in a reduced-form way. Therefore, the underlying assumption for this simplification is that the markups charged by the distributors and retailers can vary at the product, market, and month levels but do not change in our counterfactual simulations. Miller and Weinberg (2017) show that a double marginalization model where a brewery first sells to retailers does not significantly change their merger simulation results.

Stage 1. Entry and Product Decisions At the beginning of year τ, firm n is endowed with a set of potential products $\mathcal{J}_{n\tau}$ and decides on the set of products $\mathcal{J}_{nm\tau}$ to offer in market m. Firms observe the product characteristics, product-, market-, and time-specific fixed effects, and fixed costs for all potential products when making their product decisions. However, the month-to-month transient demand and marginal cost shocks $(\xi_{jmt}, \omega_{jmt})$ are realized only in the second stage, after firms have chosen which products to sell.

The profit-maximization problem at this stage is:

$$\max_{\mathcal{J}_{nm\tau} \subseteq \mathcal{J}_{n\tau}} \pi_{nm}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right) - C_{nm}\left(\mathcal{J}_{nm\tau}\right), \tag{18}$$

where $\pi_{nm}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)$ is the expected variable profit and $C_{nm}\left(\mathcal{J}_{nm\tau}\right)$ is the fixed cost. We now derive the former and specify the latter.

To derive firm n's expected variable profit $\pi_{nm}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)$, we plug the second-stage equilibrium prices into its profit function, take the expectation over the transitory demand and marginal cost shocks, and sum over all months in a year. Formally, we use $\mathcal{J}_{-nm\tau}$ to denote the set of products that firm n's competitors sell in market m. Let $p_{jmt}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)$ and $Q_{jmt}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)$ denote the second-stage equilibrium price and quantity, respectively, which depends on the observable covariates $(\boldsymbol{x}_j,\ X_{jm})$, fixed effects $(FE_j^{\text{demand}},\ FE_m^{\text{demand}}, FE_t^{\text{demand}}, FE_j^{\text{mc}}, FE_m^{\text{mc}},\ FE_t^{\text{mc}})$ as well as the shocks $(\xi_{jmt}, \omega_{jmt})$ for all products in market m. Let $\xi_{mt} = (\xi_{jmt}, j \in \mathcal{J}_{nm\tau} \cup \mathcal{J}_{-nm\tau})$ be the collection of demand shocks for all products in market m and define ω_{mt} for the marginal cost shocks analogously. Let $\mathcal{T}_\tau$ represent all months of year τ. Firm n's expected variable profit, $\pi_{nm}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)$ in (18) is:

$$\begin{aligned} &\pi_{nm}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right) \\ = &\sum_{t \in \mathcal{T}_\tau} E_{\xi_{mt}, \omega_{mt}}\left(\sum_{j \in \mathcal{J}_{nm\tau}} \left[p_{jmt}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right) - mc_{jmt}\right] \cdot Q_{jmt}\left(\mathcal{J}_{nm\tau}, \mathcal{J}_{-nm\tau}\right)\right). \end{aligned} \tag{19}$$

The fixed cost function in (18) is specified as

$$C_{nm}\left(\mathcal{J}_{nm\tau}\right) = \sum_{j\in\mathcal{J}_{nm\tau}} \left(W_{jm}\theta + \sigma_m \zeta_{jm\tau}\right), \tag{20}$$

where W_{jm} is a vector of covariates including, for example, whether product j is a craft beer product. While we assume the fixed cost shock $\zeta_{jm\tau}$ to be i.i.d. and follows a standard normal distribution, we allow the parameter σ_m to vary with market size.

In this baseline specification, we assume additive separability of fixed costs across products, a common assumption in the literature of estimating discrete games. This assumption, however, rules out economies or dis-economies of scope. In Supplemental Appendix D, we extend our model and estimation method to allow for this possibility. Our baseline specification also assumes unobservable shocks are uncorrelated. In Supplemental Appendix E, we consider an extension allowing for such correlations. Our results are robust to both changes.

5 Estimation

5.1 Estimation of Demand Parameters and Marginal Costs

We combine the aggregate product/market/month-level data on prices, product characteristics, and market shares with the individual/month-level panel data on household purchases to estimate demand parameters. Specifically, we rely on the market share data to identify the mean taste coefficients (α, β) and fixed effects $(FE_j^{\text{demand}}, FE_m^{\text{demand}}, FE_t^{\text{demand}})$. We exploit the panel data and the correlations between household income and beer purchases to identify the standard deviations of the unobservable consumer heterogeneity ($\sigma^{(\cdot)}$ parameters) as well as the effect of household income on consumer tastes ($\kappa^{(\cdot)}$ parameters). We estimate these parameters using the Generalized Method of Moments approach where we combine a set of macro moments with two sets of micro moments.

We construct macro moments based on instrumental variables consisting of the interactions of global barley prices with beer types to address potential price endogeneity. We choose the price of barley because it is an ingredient in almost all beers. We interact the price of barley with beer types because its ingredient proportion differs by beer type. The historical prices of barley in dollars per metric ton in Figure 3 shows that the price series displays fairly large monthly variations.[12]

We construct a new set of micro moments based on the persistence of a household's purchasing decisions to identify the standard deviation parameters $\sigma^{(\cdot)}$. For example, a

[12] Data source: https://fred.stlouisfed.org/series/PBARLUSDM

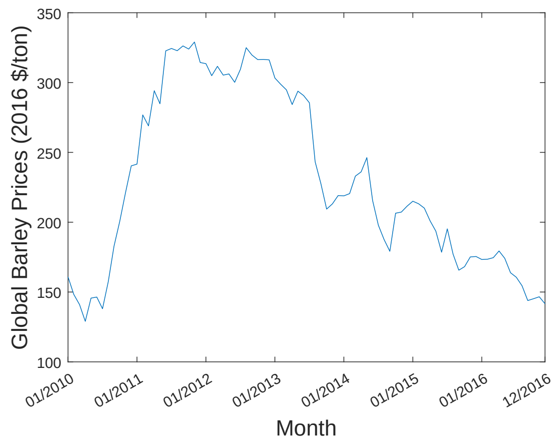

Figure 3: Price of Barley

large σ^{craft} suggests a high correlation in a household's preference for craft products across months. In other words, conditional on a household ever purchasing a craft product in a year, that household is likely to have purchased many craft products throughout the year. More generally, if we use $q_{i\tau}^{f} = \sum_{t \in \mathcal{T}_\tau} q_{it}^{f}$ to denote a household's purchase of a beer type ($f \in \{\text{ale}, \text{lager}, \ldots\}$) in year τ, then matching the conditional mean $E\left(q_{i\tau}^{f} \,|\, q_{i\tau}^{f} \geq 1\right)$ helps to identify the parameter σ^{f}.

Similar moments are also useful for identifying the correlation between taste shocks. For example, if a household that prefers type-f products tends to dislike type-f' products, then conditional on a household ever purchasing a type-f product, the household should buy few if any type-f' beers throughout the year.

Specifically, in constructing these micro moments, we match the model predictions of the following moments to their empirical counterparts (see Supplemental Appendix A for details on calculation):

- A household's expected annual purchase of a certain type of beer conditional on purchasing at least one unit of this type of beer in the year, i.e., $E\left(q_{i\tau}^{f} \,|\, q_{i\tau}^{f} \geq 1\right)$. Matching these moments helps to identify σ^{f}.

- A household's expected annual purchase of a certain type of beer conditional on purchasing at least one unit of craft beer in the year, i.e., $E\left(q_{i\tau}^{f} \,|\, q_{i\tau}^{\text{craft}} \geq 1\right)$. Matching these moments help to identify the taste correlation between craft and type-f beer.

- A household's expected annual purchase of beer conditional on purchasing at least one unit of beer in the year, i.e., $E\left(q_{i\tau} \,|\, q_{i\tau} \geq 1\right)$, where $q_{i\tau}$ is a household's total beer purchase amount over a year. Matching this moment helps to identify σ_0.

We construct a second set of micro moments similar to those in Petrin (2002) to identify the effect of household income on consumer tastes:

- The ratio of average expenditure over average purchase quantity in a year among households whose income falls into a bin $\mathcal{I}$, i.e., $E\left(\text{expenditure}_{i\tau}\,|y_i \in \mathcal{I}\right)/E\left(q_{i\tau}\,|y_i \in \mathcal{I}\right)$, where the log-income bins $\mathcal{I}$ are $log\,(0, \$50K]$, $log\,(\$50K, \$100K]$, or $log\,(\$100K, +\infty)$.[13] Matching these moments helps to identify the income effect on price sensitivity, κ_1.
- $E\left(q_{i\tau}^{\text{craft}}\,|\,q_{i\tau}^{\text{craft}} \geq 1, y_i \in \mathcal{I}\right)$, which helps to identify κ^{craft}.
- $E\left(q_{i\tau}\,|\,q_{i\tau} \geq 1, y_i \in \mathcal{I}\right)$, which helps to identify κ_0.

Our estimation of marginal costs is standard and follows Berry, Levinsohn and Pakes (1995): we back out marginal costs based on the first-order conditions of the profit maximization problem in (16).

5.2 Estimation of Fixed Cost Parameters

We follow the estimation procedure described in Section 2.2 to estimate the fixed cost parameters, which include the parameters of the mean fixed cost (θ) and the standard deviations of the fixed cost shock (σ_m). In this section, we explain how we reformulate our empirical model to be consistent with the model outlined in Section 2.2, provide estimation details, and importantly, present data patterns that help with identification. We estimate the fixed cost parameters for each year separately and thus suppress the year subscript τ in this section for exposition simplicity.

Reformulation of the Model

To be consistent with the model outlined in Section 2.2, we rewrite the profit function in (18), i.e., $\pi_{nm}\left(\mathcal{J}_{nm}, \mathcal{J}_{-nm}\right) - \sum_{j \in \mathcal{J}_{nm}} \left(W_{jm}\theta + \sigma_m \zeta_{jm}\right)$ as

$$\pi_n\left(Y_{nm}, Y_{-nm}, X_{nm}\right) - \sum_{j \in \mathcal{J}_n} Y_{jm}\left(W_{jm}\theta + \sigma_m \zeta_{jm}\right). \tag{21}$$

Specifically, we now use a vector of indicators $Y_{nm} \in \{0,1\}^{\#\mathcal{J}_n}$ to denote a firm's product portfolio $\mathcal{J}_{nm} \subseteq \mathcal{J}_n$, where $\mathcal{J}_n$ represents the potential products that firm n is endowed with. We use Y_{jm} to denote the element of Y_{nm} that corresponds to product $j \in \mathcal{J}_n$, where $Y_{jm} = 1$

[13] An alternative moment is the average price $E\left(\frac{\text{expenditure}_{i\tau}}{q_{i\tau}}\,|\,y_i \in \mathcal{I}\right)$. However, this moment is computationally cumbersome as it requires drawing both v_i^f and ε_{ij} to simulate it but only v_i^f to simulate the moment in the text.

if $j \in \mathcal{J}_{nm}$ and 0 otherwise. Therefore, the expected variable profit $\pi_{nm}(\mathcal{J}_{nm}, \mathcal{J}_{-nm})$ can be written as $\pi_n(Y_{nm}, Y_{-nm}, X_{nm})$, where the vector X_{nm} includes all demand and marginal cost covariates (including fixed effects). Similarly, the total fixed cost $\sum_{j \in \mathcal{J}_{nm}} (\theta W_{jm} + \zeta_{jm})$ can be written as the summation over all products with $Y_{jm} = 1$, i.e., $\sum_{j \in \mathcal{J}_n} Y_{jm} (\theta W_{jm} + \zeta_{jm})$.

Estimation Details

To compute the moment functions, we need to compute $\underline{\Delta}_j(X_{nm}) = \min_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$ and $\overline{\Delta}_j(X_{nm}) = \max_{Y_{-jm}} \Delta_j(Y_{-jm}, X_{nm})$, where $\Delta_j(Y_{-jm}, X_{nm})$ is the change in firm n's expected variable profit when product j joins the market. Directly solving for the minimum and maximum of the expected profit across all possible values of Y_{-jm} may be computationally prohibitive given $2^{(\text{length of } Y_{-jm})}$ possible values of Y_{-jm} and the need to solve a pricing game for multiple simulated draws of demand and marginal cost shocks in order to compute $\Delta_j(Y_{-jm}, X_{nm})$ for each value of Y_{-jm}. Economic intuition suggests that, because products are substitutes, we can approximate the minimum and maximum by, respectively, the following:

$$\underline{\Delta}_j(X_{nm}) \approx \Delta_j((1, ..., 1), X_{nm}) \text{ and } \overline{\Delta}_j(X_{nm}) \approx \Delta_j((0, ..., 0), X_{nm}).$$

These approximate extrema are exact for entry games such as those in Berry (1992), Seim (2006), Ciliberto and Tamer (2009), Sweeting (2013), and Berry, Eizenberg and Waldfogel (2016). For more general demand and pricing models such as ours, we find that the approximate extrema coincide with the true values in all of our computational experiments.[14]

We construct the 95% confidence set following Chernozhukov, Chetverikov and Kato (2019). The inference procedure is outlined in Section 2. We consider 5000 grid points of parameters and use each grid point as a starting point to minimize the test statistic in (14) until it just falls below the critical value. The resulting set of parameter values comprises the 95% confidence set. More details, such as our construction of the non-negative functions $g^{(k)}(X_{nm}, W_{jm})$, are given in Supplemental Appendix B.

Data Patterns That Help with Identification

In Section 2, we discuss identification in abstract for a general model. Here, we present data patterns in our empirical setting that help with identification. As mentioned, we estimate

[14] We randomly sample 100 markets and K potential products. For each selected potential product j in a selected market, we hold fixed the entry outcomes of the products not included in the K selected products, and enumerate all possible outcomes for the other $K-1$ products (i.e., all possible Y_{-jm}) to find the actual extrema. For all sampled markets and K products when K takes the value of 6, 8, or 10, we find that the approximations coincide with the actual extrema.

(A) Histogram of $\underline{\Delta}_j (X_{nm}) / \overline{\Delta}_j (X_{nm})$

(B) Variations in $\underline{\Delta}_j (X_{nm}), \overline{\Delta}_j (X_{nm})$ and Entry Probability

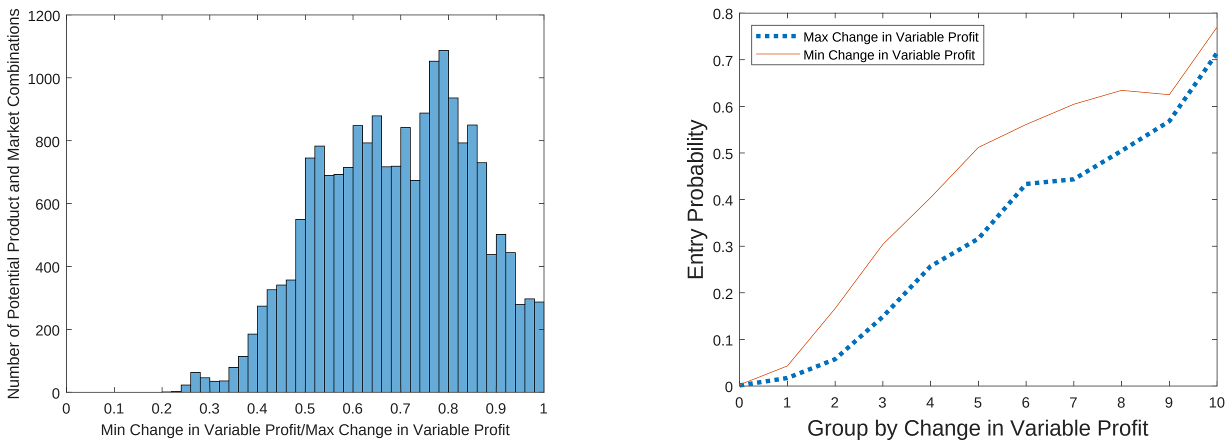

Figure 4: Data Patterns Aiding Identification

the fixed cost parameters year by year. We present the patterns based on the 2016 data.

First, for a large proportion of the observations, the minimum and maximum changes in variable profit, i.e., $\underline{\Delta}_j (X_{nm})$ and $\overline{\Delta}_j (X_{nm})$, are relatively close, resulting in tight conditional choice probability bounds for these products. We plot the histogram of the ratio $\underline{\Delta}_j (X_{nm}) / \overline{\Delta}_j (X_{nm})$ across all combinations of potential products and markets in Panel (A) of Figure 4. A larger ratio reflects a smaller difference between the minimum and maximum. The median of the ratio is around 0.7.

Second, there are rich variations in $\underline{\Delta}_j (X_{nm})$ and $\overline{\Delta}_j (X_{nm})$, and moreover, these variations are informative about variations in entry probabilities. To see the variations in $\underline{\Delta}_j (X_{nm})$ and $\overline{\Delta}_j (X_{nm})$, we note that their 25%, 50%, and 75% percentiles are, respectively, (\$78.10, \$239.32, \$799.75) and (\$116.60, \$360.46, \$1187.09), where the 75% percentiles are more than 10 times of the 25% percentiles. To see the association between these variations and variations in entry probabilities, we discretize $\underline{\Delta}_j (X_{nm})$ into 10 groups, i.e., [0,240], [240,480],..., [2160,2400] and [2400,∞). For each group g, we compute the average entry probability for observations jm such that $\underline{\Delta}_j (X_{nm})$ is in this group as $\frac{\sum_{j,m} \mathbb{1}\left(\underline{\Delta}_j (X_{nm}) \in g\right) \cdot Y_{jm}}{\sum_{j,m} \mathbb{1}\left(\underline{\Delta}_j (X_{nm}) \in g\right)}$, where $Y_{jm} \in \{0, 1\}$ is the observed entry outcome. We repeat this exercise for the association between entry probabilities and $\overline{\Delta}_j (X_{nm})$ analogously. In Panel (B) of Figure 4, the red solid line represents the entry probabilities associated with $\underline{\Delta}_j (X_{nm})$ while the blue dotted line represents those associated with $\overline{\Delta}_j (X_{nm})$. From the figure, we can see that the average entry probabilities increase in both $\underline{\Delta}_j (X_{nm})$ and $\overline{\Delta}_j (X_{nm})$.

(A) Market Size and # Craft Products

(B) Distribution of Distance

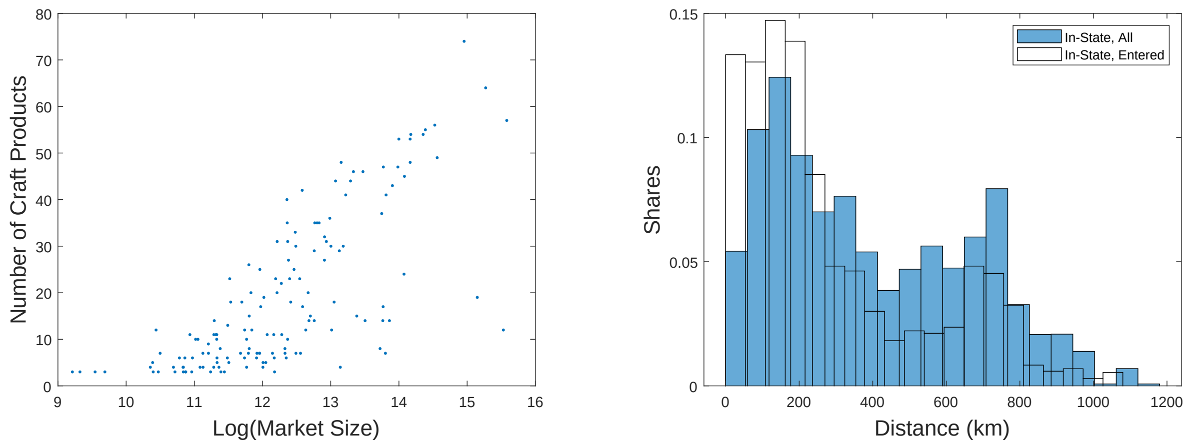

Figure 5: Exogenous Variations Aiding Identification

What exogenous variations in X_{nm} generates the variations in $\underline{\Delta}_j\left(X_{nm}\right)$ and $\overline{\Delta}_j\left(X_{nm}\right)$? In addition to variations in product characteristics as well as the fixed effects in the demand and marginal cost functions, variations in market sizes also play an important role because everything else being equal, the returns to entry increase in the size of a market. This can be seen in Panel (A) of Figure 5, which depicts a strong positive correlation between the log of market sizes and the number of craft products in a market.

Another source of exogenous variation is the distance between a production facility and a market. In Panel (B) of Figure 5, we plot the unconditional distribution of distances for all in-state craft potential product/market combinations and the conditional distribution for observed in-state craft product/market combinations (i.e., the product is in the market in the data).[15] Panel (B) shows that the conditional distribution has more probability mass at shorter distances than the unconditional one, suggesting a negative correlation between distance and entry. We account for these variations in variable profits by including controls for distances in both our demand and marginal cost functions.

6 Estimation Results

6.1 Demand and Markup

Table 4 reports the demand estimation results. The estimated $\sigma^{(\cdot)}$ parameters indicate significant heterogeneity in preferences for craft products, imported products, and flavor

[15] Out-of-state craft products tend to be widely distributed and thus less affected by the distance between the production facility and market. Some of the more popular and widely distributed craft products in California are brewed on or near the East Coast.

Table 4: Demand Estimates

Unobserved	σ_0	0.00	Income Effect	κ_0	-2.15
Heterogeneity		(0.02)			(0.02)
	σ^{ale}	1.98		κ^{craft}	1.08
		(<0.01)			(0.02)
	σ^{lager}	0.89		κ_α	0.15
		(<0.01)			(<0.01)
	σ^{light}	2.67			
		(<0.01)	Price Coefficient	α	-2.26
	σ^{import}	2.14			(0.03)
		(<0.01)			
	σ^{craft}	2.44	Distance bin FE		Yes
		(<0.01)	Product FE		Yes
	$\rho^{\text{craft-light}}$	-0.28	Market FE		Yes
		(<0.01)	Time FE		Yes

Note: Standard errors are in parentheses.

types. For example, the estimated standard deviation of the unobservable heterogeneity in consumer taste for craft products $\hat{\sigma}^{\text{craft}}$ is 2.44. To understand the magnitude of this estimate, we compare it to the price coefficient of a household with an income of \$50,000, which is $-2.26 + 0.15 \cdot log\,(\$50,000) = -0.64$. Therefore, the estimated $\hat{\sigma}^{\text{craft}}$ is equivalent to a price discount of 2.44/0.64, or 3.81 dollars.

The dispersion parameters $\sigma^{(\cdot)}$ are estimated by matching the micro moments that capture the persistence in a household's purchasing decisions. Table 5 shows the model fit for these micro moments. For example, from Row (6), we see that the average per-household annual craft purchase among households that purchase at least one unit of craft beers is 3.93. Compared with the unconditional average per-household annual craft purchase of 0.38 units, this micro moment implies that craft beers are purchased by a set of dedicated craft consumers, leading to a significant estimate of σ^{craft}.

The estimation results also indicate a negative correlation between consumer taste for craft and light beers ($\hat{\rho}^{\text{craft-light}} = -0.28$). This finding is consistent with the summary statistics in Table 3, which show that light craft beers account for only 7.23% of the craft beer sales while light beers in general account for 40% of all beer sales. We find that allowing for a correlation between ν_i^{light} and ν_i^{craft} is helpful for matching the conditional purchases of light beers given at least one craft purchase. The moment $E\left(q_{i\tau}^{\text{light}} \,|\, q_{i\tau}^{\text{craft}} \geq 1\right)$ is 1.27 in the data and 1.07 according to our estimated model, and it would be 2.14 if such a correlation were not allowed in our model.

Moreover, we find heterogeneity in consumer tastes across income levels. Specifically, high-income households are less likely to purchase beer ($\hat{\kappa}_0 < 0$), have a stronger preference

Table 5: Micro Moments on Persistence in Purchasing Decisions

		Data	Model
(1)	$E\left(\sum_{t=1}^{12} q_{it} \mid \sum_{t=1}^{12} q_{it} \geq 1\right)$	7.50	7.80
(2)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{ale}} \mid \sum_{t=1}^{12} q_{it}^{\text{ale}} \geq 1\right)$	3.10	3.91
(3)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{lager}} \mid \sum_{t=1}^{12} q_{it}^{\text{lager}} \geq 1\right)$	5.56	4.17
(4)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{light}} \mid \sum_{t=1}^{12} q_{it}^{\text{light}} \geq 1\right)$	8.03	8.25
(5)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{import}} \mid \sum_{t=1}^{12} q_{it}^{\text{import}} \geq 1\right)$	2.86	2.96
(6)	$E\left(\sum_{t=1}^{12} q_{it}^{\text{craft}} \mid \sum_{t=1}^{12} q_{it}^{\text{craft}} \geq 1\right)$	3.93	4.17

for craft products $\left(\hat{\kappa}^{\text{craft}} > 0\right)$ if they do purchase beer, and are less price sensitive ($\hat{\kappa}_\alpha < 0$).

Overall, the estimated demand parameters imply that the substitution within craft products is much larger than the substitution between craft and non-craft products. Table 6 reports the own and cross elasticities among the top-3 non-craft and top-3 craft products in 2016.[16] These elasticities suggest little substitution between craft and non-craft products. Similarly, Figure 6 presents the histogram of the diversion ratio for a craft product to non-craft products (Panel (A)) and that for a craft product to other craft products (Panel (B)). Panel (A) shows that for most craft products, almost no sales would be captured by non-craft products if the focal craft product's price is increased. By contrast, the distribution of the diversion ratio to other craft products in Panel (B) has a mode of around 20%.

Table 6: Elasticities: Top-3 Craft Products and Top-3 Main Products (%)

	Craft			Main		
	-10.09	0.14	0.02	0.01	0.01	0.01
Craft	0.22	-9.52	0.02	0.01	0.01	0.01
	0.04	0.03	-9.16	0.01	0.03	0.01
	<0.01	<0.01	<0.01	-5.87	0.04	0.67
Main	<0.01	<0.01	<0.01	0.08	-6.81	0.08
	<0.01	<0.01	<0.01	0.68	0.04	-5.88

We back out the marginal costs using the first-order conditions at the pricing stage of the game and present the distribution of the quantity-weighted markup in Panel (C) of Figure 6. The median markup of craft beers is about \$1.7 in 2016 dollars. Some industry sources (e.g., Satran, 2014) put the brewer's margin at 8% of the retail price, or \$1.4 for an average price of \$17, in line with our estimates.

Finally, we note that our observed explanatory variables account for the majority of the

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

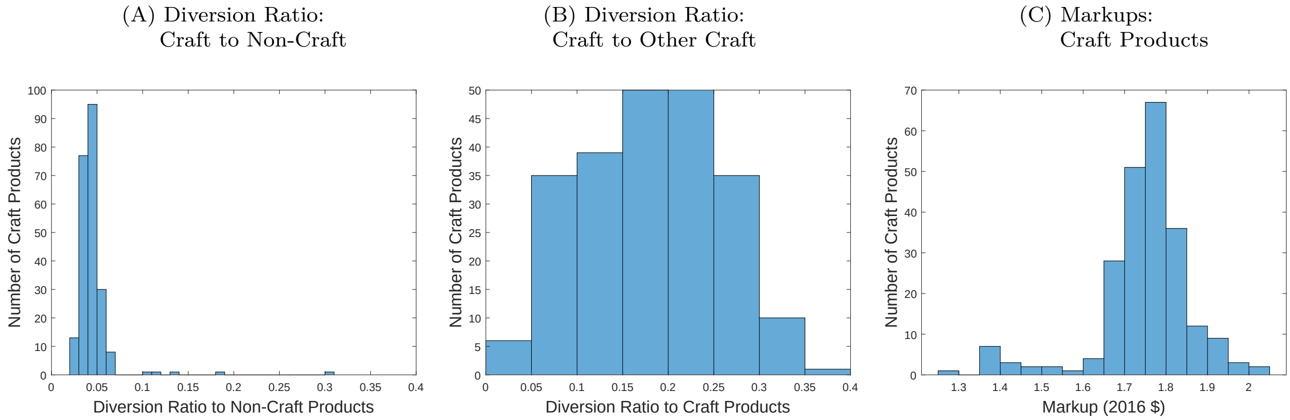

Figure 6: Histograms of Diversion Ratios and Markups

variations in the utility levels and marginal costs compared to the month-to-month transient shocks. The R^2's from regressing the mean utility and marginal cost on observable covariates and fixed effects are both above 0.9, implying that transient month-to-month shocks play, at most, a small role. Two recent papers on entry or product repositioning (Ciliberto, Murry and Tamer, 2021 and Li, Mazur, Park, Roberts, Sweeting and Zhang, 2022) assume that firms observe demand and marginal cost shocks as well as fixed cost shocks when making entry decisions. In other words, they account for selection based on unobserved demand and marginal costs as well as on unobserved fixed cost shocks. By contrast, we allow for selection based only on unobserved fixed cost shocks and address selection based on unobservable demand and marginal cost shocks by including a large number of fixed effects in our demand and marginal cost functions. The remaining unobservables are month-to-month product/market-level transient shocks. We find it reasonable to assume firms do not observe them when making product choices. The finding that these shocks play a small role in explaining demand and marginal cost implies that even when our timing assumption is violated, the resulting bias is likely to be small.

6.2 Fixed Costs

We estimate the fixed cost parameters year by year. In this section, we focus on craft products and present our fixed cost estimation results using the 2016 data for consistency with our later counterfactual analyses. In our estimation, one unit of observation is a potential product j and market m combination. In this part of the estimation, we exclude markets with no craft products, yielding a total of 95 potential products, 149 markets, and 14,155 potential product/market combinations in 2016. We follow the inference procedure described in Section 2 and the estimation details provided in Section 5, and report the 95% confidence

Table 7: Estimates of Fixed Costs: Projected 95% Confidence Intervals, 2016

Craft (θ_1)	[229.14, 1093.24]
In State× Craft (θ_2)	[-387.82, 208.18]
Market-size specific fixed cost (θ_3)	
Small market	[308.95, 938.33]
Medium market	[1027.77, 1468.10]
Large market	[3325.71, 4177.69]
Market-size specific std. dev. (σ_ς)	
Small market	[0.00, 522.79]
Medium market	[679.41, 863.25]
Large market	[2511.65, 3424.06]

Note: Estimates in 2016 US dollars.

set projected to each parameter in Table 7.

We find a higher fixed cost for independent craft breweries and larger markets. The 95% confidence set projected to the coefficient of the craft indicator is [$229, $1093], indicating that craft breweries incur higher fixed costs than non-craft breweries. This parameter is identified by the data pattern that products of craft breweries acquired by macro breweries are more likely to enter a market than those of independent craft breweries.[17] To study whether fixed costs vary with market size, we categorize markets into small, median, or large bins according to whether the market size is below 10^5, between 10^5 and 5×10^5, or above 5×10^5 units and allow fixed costs to differ across bins.[18] We find that fixed costs are higher in larger markets and that the standard deviation of the unobservable fixed cost shock also increases in market size.

7 Counterfactual Results

7.1 Counterfactual Designs

We consider a counterfactual merger where the largest firm in our sample (a so-called macro brewery) acquires the three largest craft firms (excluding Boston Beer Company and Sierra

[17] We make a distinction between products of independent craft breweries and the products of (former) craft breweries currently owned by large breweries. We set the craft dummy to 0 for these latter products in our fixed cost function, but still designate these products as craft in our demand and marginal cost functions. The underlying assumption is that consumers make purchase decisions based on taste preferences for craft products. Given that the craft beers acquired by large firms are still produced by the same facilities with the same ingredients and procedures, they are likely considered craft by consumers. However, these products may benefit from the distribution and marketing networks of the large firms and therefore have a different fixed cost.

[18] The 25% and 75% quantiles of the market sizes are 0.84×10^5 and 4.5×10^5 units.

Nevada Brewing, which are unlikely merger targets given their sizes) in 2016. During our sample period, there are four observed acquisitions where a different large brewery acquired an independent craft brewery in our sample.[19] For these mergers, our model predicts little to no change in prices and, in the absence of merger efficiency, entry as well as product variety.[20] In our counterfactual, we study an acquisition of three large craft firms by the largest macro beer firm, i.e., a scenario where this trend of acquisitions continues to a point that the concentration of the craft market approaches the level in the overall beer market.

In our simulation, we allow firms to adjust their craft products. For computational ease, we hold the non-craft product choices fixed as observed in the data but allow their prices to change. The simplification is justified by the estimated small substitution between craft and non-craft products.

We consider a decision maker to be a firm that is observed in any market in our sample. Each firm is endowed with a set of potential products comprised of the firm's craft products observed in any market in the 2016 data. In each market, a firm chooses a subset from its potential products. An empty subset denotes no entry. The potential product set for the merged firm consists of the combined set of potential products. We assume that firms maximize profits in both the product choice and pricing stages. We compute the post-merger product equilibrium using the algorithm outlined in Fan and Yang (2020).[21]

To quantify the effects of the merger and to decompose the overall effects into those due to price versus product variety adjustments, we conduct three counterfactual simulations. Specifically, in the counterfactual simulation described above (CF1), we allow for three adjustment margins—new entry, product adjustments, and price adjustments. In the second counterfactual (CF2), we allow for only incumbent product adjustment and price adjustment by removing the products added by new entrants in CF1 and recomputing the pricing equilibrium. In the third counterfactual (CF3), we allow for only the price effect of the

[19] There are more mergers involving smaller craft breweries not in our sample.

[20] Of the four observed acquisitions, one brewery was not present in our sample prior to the transaction. Among the other three acquisitions, we observe an increase in entry and product variety post merger, but we do not observe a consistent change in prices. After accounting for merger efficiency as described in Section 7.3, the merger effects on product variety and entry are consistent with what we observe in the data.

[21] Fan and Yang (2020) develops a heuristic algorithm to find a firm's best-response product portfolio given the portfolios of its competitors, and embed this optimization algorithm in a best-response iteration to solve for the post-merger product-choice equilibrium. To find a firm's best response, we start with an initial vector of product decisions for this firm and check whether it is profitable to add a potential product or drop a product. If so, we update the vector of product decisions to be the most profitable deviation and check again whether there is profitable one-product deviation from the updated vector. This process continues until there is no more profitable one-product deviation. To check for multiple equilibria, we use different initial vectors of product decisions as starting point for this algorithm. We find identical results across two starting points we use (one corresponds to the observed product decisions in the data and the other corresponds to a vector of ones for each firm n while its rivals' initial product decisions are consistent with those in the data).

merger by restoring pre-merger market products and recomputing the pricing equilibrium. The difference between the outcomes in CF1 and CF3 gives us the overall product variety effect of the merger, which can be further decomposed into the product variety effect due to new entry (CF1 - CF2) and that due to incumbent product adjustments (CF2 - CF3).

For all simulations, we use the point estimates of the demand and marginal cost parameters and sample 30 vectors of fixed cost parameters from their 95% confidence set. We ensure that the sampled parameter vectors include those on the boundary of the confidence set.

We draw three sets of shocks: demand, marginal cost, and fixed cost shocks. We draw demand and marginal cost shocks directly from their estimated distributions. For each sampled fixed cost parameter vector, we draw fixed cost shocks from the estimated distribution conditional on the observed pre-merger equilibrium to ensure that pre- and post-merger outcomes are comparable (details on how we draw our fixed cost shocks can be found in Supplemental Appendix C).

For each sampled fixed cost parameter vector, we compute the simulated merger effects averaged across the simulation draws of the demand, marginal cost, and fixed cost shocks. We report the range of this average effect across the sampled fixed cost parameter vectors as the 95% confidence interval of the average effects.

7.2 Counterfactual Results

7.2.1 Heterogeneous Merger Outcomes Across Markets

In this section, we show how merger effects vary across markets. For the purpose of presentation, we sort markets into 10 groups according to market size from group 1 (smallest) to group 10 (largest). There are 15 markets in groups 1 to 9 and 14 largest markets in group 10. Within each group, we average counterfactual outcomes across markets weighted by their market size.

Figure 7 shows that entry occurs across all market groups, and that the largest markets in group 10 see more entries than do the other groups (Panel (A)). The number of added products by new entrants (Panel (B)) is almost identical to the number of new entrants, implying that new entrants, on average, enter with one product. As for the incumbents, the merging incumbent firms drop products (Panel (C)) while non-merging incumbents add products in larger markets (Panel (D)). The net change in the number of products by the incumbents is always negative (Panel (E)), which includes the changes in Panels (C) and (D). Furthermore, even with the added products by new entrants, the overall number of products decreases (Panel (F)), which includes the changes in Panels (B)–(D). The increase in the quantity-weighted average craft beer price (Panel (G)) is centered around 5 cents, but

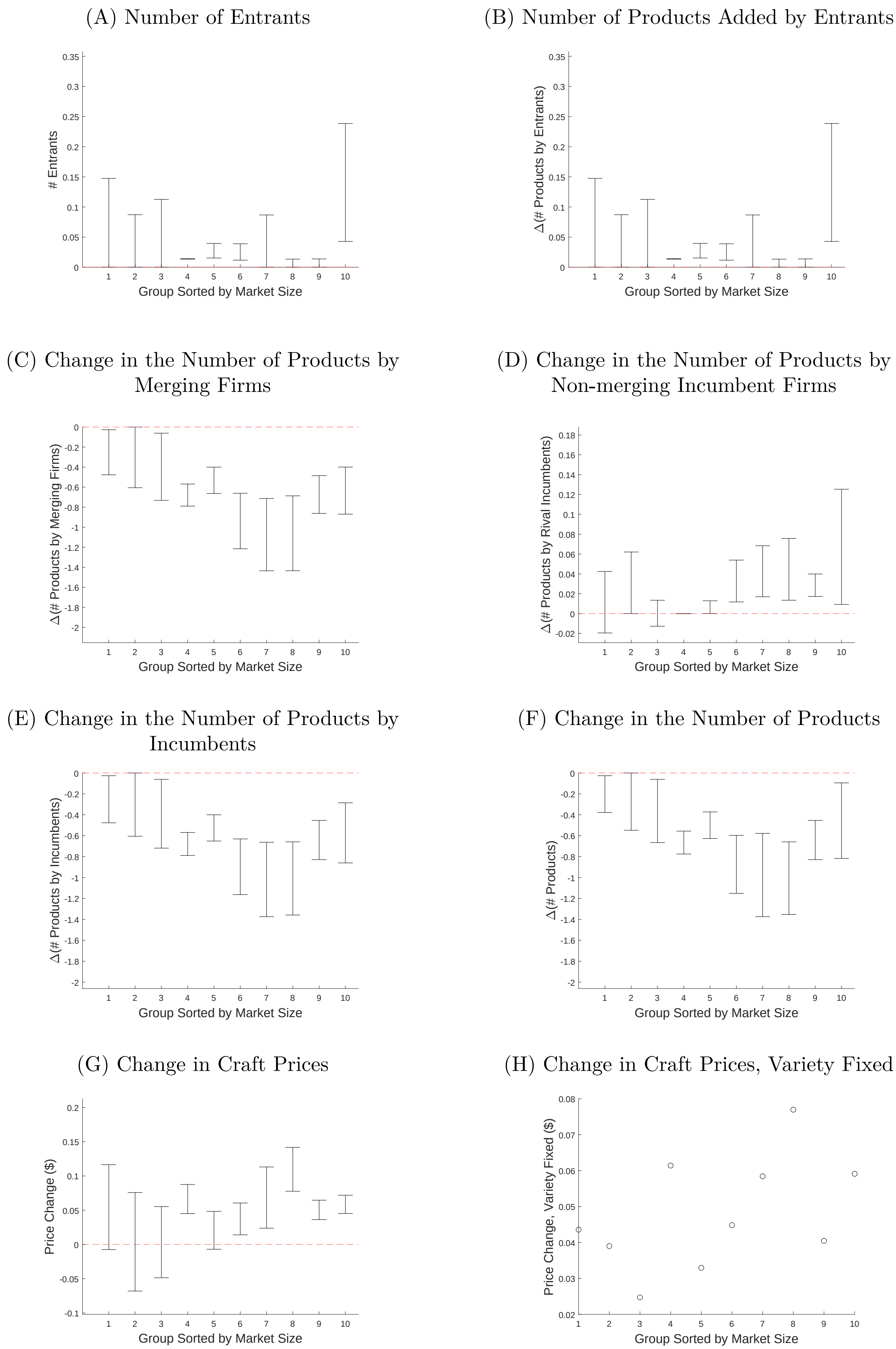

Figure 7: Merger Effects on Entry, Product Variety, and Prices

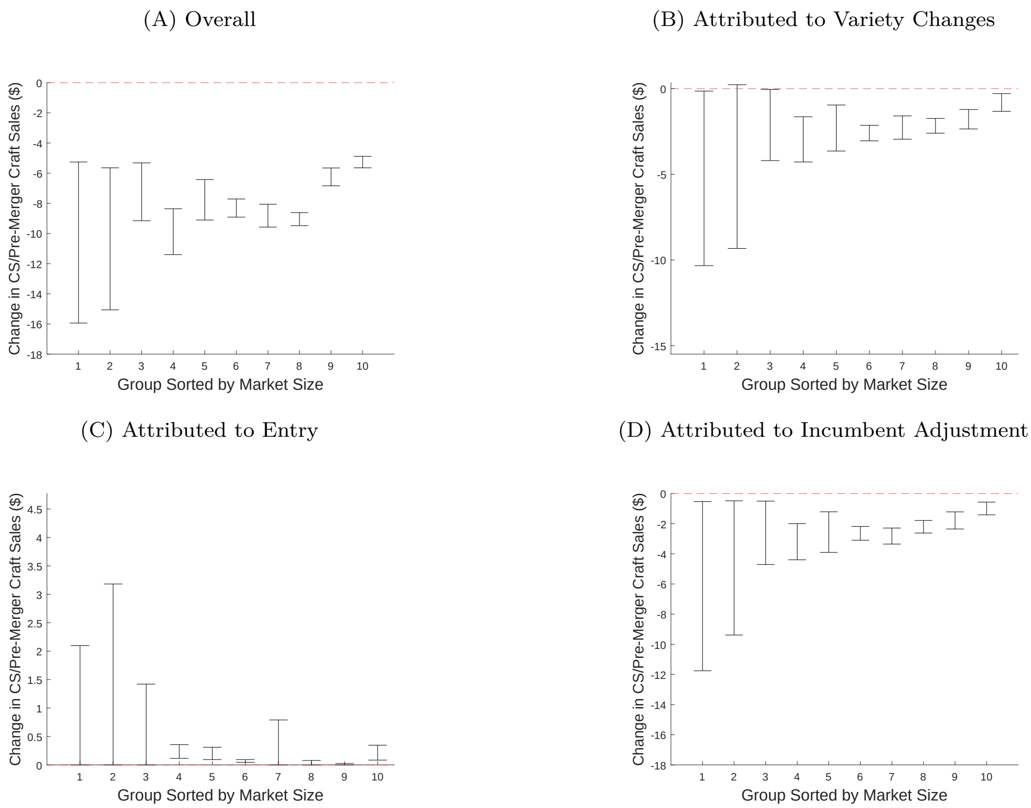

Figure 8: Change in CS/Pre-Merger Craft Sales

could be as large as nearly 15 cents, 9% of an average markup of $1.7. Panel (H) shows that price changes under fixed variety have similar magnitudes.

Figure 8 presents the effects of the merger on consumer welfare, measured as the change in the total consumer surplus in a market divided by the market's pre-merger craft beer sales. We choose this measure due to the low substitution between craft and non-craft beers. As a result, craft beer consumers are the more relevant consumer base in our the welfare analysis. Panel (A) shows that the average annual loss of craft consumer surplus ranges from 5 dollars to 16 dollars with smaller per-consumer losses in larger markets. In panel (B), we break out the average loss attributable to the variety changes, which, according to this panel, is an important part of the welfare loss. A further decomposition of the effects of variety changes indicates that the effect due to new entries is positive (Panel (C)) and only partially offsets the negative effect due to the product adjustments by the incumbents (Panel (D)), resulting in the negative net change documented in Panel (B).

7.2.2 Correlation between Merger Effects and Market Characteristics

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

Table 8: Regressions of Changes in the Number of Products

	(A) Merging Firms	(B) Other Firms	(C) Market
Variety-Fixed Avg Price Increase, Merging Firms ($)	[-14.64, -4.06]*	[0.33, 0.95]*	[-13.73, -2.88]*
Avg Fixed Cost, Merging Firms ($1000)	[-0.02, 0.00]	[-0.00, 0.00]	[-0.05, 0.00]
Avg Fixed Cost, Other Firms ($1000)	[-0.16, 0.05]	[-0.01, 0.01]	[-0.19, 0.01]
Market Size (10^6)	[-0.04, 0.08]	[-0.01, 0.04]	[-0.02, 0.21]
Avg Household Income ($10,000)	[-0.00, 0.03]	[-0.01, 0.00]	[-0.01, 0.04]
R^2	[0.15, 0.58]	[0.07, 0.31]	[0.13, 0.53]
N	149	149	149

Notes: the dependent variables are the change in the number of products by merging firms (Column (A)), the change in the number by other firms (Column (B)), and the net change in a market (Column (C)). Each observation is a market. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set. * indicates significance at the 5% level for all sampled vectors of fixed cost parameters.

merging firms are close substitutes, the merged firm is more likely to drop products and increase prices significantly. At the same time, pronounced price increases attract new entries and/or product entry by non-merging incumbents. Therefore, the merging firms' market power may be positively correlated with both the number of dropped merging firm products and the number of added non-merging firm products, although the sign is less clear with the net change of the product count. Another relevant market characteristic is market size. All else being equal, we expect more entries in larger markets. Finally, higher fixed costs may be associated with a smaller increase or larger decline in the number of products.

To document the correlation between merger effects on product numbers and market characteristics, we employ a series of regressions. In our regressions, one observation is a market. For each market, we measure the merging firms' market power as the increase in the quantity-weighted average price of their products in CF3 (i.e., when the set of products is fixed). One could consider this measure to be the realized upward pricing pressure. We compute the average fixed cost of the merging firms in market m as the average of $W_{jm}\hat{\theta}$ across all merging firms' potential products. Since the average fixed cost does not include fixed cost unobservables, it reflects market features. The average fixed costs of other firms are defined analogously.

Table 8 reports the regression results. We regress the change in the number of products by merging firms in a market on our respective measures of market power, average fixed costs, and market size. We collect the regression results for each sampled fixed cost parameter vector and report the range for our regression estimates in Column (A). In Column (B), we do the same using the change in the number of products by non-merging firms as the dependent variable. Column (C) reports the results for the net change.

Table 8 shows that the estimated coefficient of the variable "variety-fixed price increase" is

Table 9: Aggregate Post-Merger Outcomes

Average Change Per Market			Aggregate Change Across Markets		
(1)	# of firms	[-2.93, -2.82]	(10)	quantity (1000)	[-266.94, -251.57]
(2)	# new entrants	[0.02, 0.14]	(11)	craft	[-249.76, -230.46]
(3)	# of products	[-0.86, -0.33]	(12)	craft, merging firms	[-301.64, -283.55]
(4)	merging firms	[-0.90, -0.49]	(13)	consumer surplus ($1000)	[-639.00, -602.81]
(5)	non-merging incumbents	[0.02, 0.08]	(14)	craft beer profits ($1000)	[97.95, 111.07]
(6)	new entrants	[0.02, 0.14]	(15)	merging firms	[24.78, 27.68]
(7)	average price ($)	[0.00, 0.00]	(16)	total surplus ($1000)	[-533.03, -504.86]
(8)	craft products ($)	[0.04, 0.07]	ΔCS decomposition ($1000)		
(9)	craft, merging firms ($)	[0.13, 0.15]	(17)	due to variety change	[-155.43, -106.54]
			(18)	due to entry	[6.85, 26.09]
			(19)	due to incumbent product adj.	[-164.55, -123.48]

Notes: rows (1)–(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)–(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.

significant at the 5% level for all sampled fixed cost parameter values across all three columns. The signs are consistent with the discussion above. Specifically, this variable is negatively (positively) correlated with the change in the merging firms' (non-merging firms') products. The correlation with the net change is negative. For the other market characteristics, we find that the estimates are less precise for at least some sampled fixed cost parameter vectors. We also note that the measure of the merging firms' market power accounts for a nontrivial share of the variations in the outcome variables. Specifically, the range of R^2 values for the three regressions are $[0.15, 0.58]$, $[0.08, 0.31]$, and $[0.13, 0.53]$, respectively. Regressing on the variable "variety-fixed price increase" alone yields only a moderate decrease in R^2 to $[0.13, 0.56]$, $[0.03, 0.13]$, and $[0.08, 0.52]$.

7.2.3 Aggregate Effects

Having established the heterogeneity of merger effects across markets and documented the correlation between the merger effects and market characteristics, we now turn to the aggregate effects across the simulated 149 markets in our counterfactual analysis.

For outcomes in the left panel of Table 9 (i.e., Rows (1)–(9)), we report the average changes across markets weighted by market size. For outcomes in the right panel (i.e., Rows (10)–(19)), we report the sum of the changes across markets. The 95% confidence interval of the (weighted) average number of new entrants is $[0.02, 0.14]$ per market (Row (2)). The merged firms drop between 0.5 to 0.9 products per market post merger (Row (4)), while rival incumbents add 0.02 to 0.08 products (Row (5)) and new entrants add 0.02 to 0.14 products in a market (Row (6)). Regarding prices, we find that the average overall beer price is barely affected by the merger. However, the merger increases the average craft beer price by about 4 to 7 cents (Row (8)). The average price of the merging firms' craft products increases by about 13 to 15 cents (Row (9)). Together, the drop in the number of products, the increase in prices, and the decrease in quantities (Rows (10)–(12)) lead to a total welfare loss of about 0.5 million dollars, aggregated across markets (Row (16)). In particular, consumer surplus decreases by \$602,810 to \$639,000 (Row (13)). Incumbent product adjustments exacerbate the loss of consumer surplus by \$123,480 to \$164,550 (Row (19)) while new entries recover between \$6,850 to \$26,090 (Row (18)), resulting in a net loss of \$106,540 to \$155,430 dollars of consumer welfare due to product variety changes after the merger (Row (17)). This loss constitutes about one quarter of the total consumer welfare loss in Row (13).

7.3 Merger Efficiency

In this section, we consider potential efficiency gains from a merger and how they affect the merger effects. We first note that mergers involving craft breweries are unlikely to realize efficiency gains in marginal costs as craft breweries often remain operationally independent and their beers continue to be brewed at the same facilities. This arrangement stands in contrast to mergers among macro breweries, where the merged firms relocate production and economize on transportation costs from production facilities to markets (Ashenfelter, Hosken and Weinberg, 2015).

However, craft breweries could benefit from using the marketing networks of the acquirer and enjoy reduced fixed costs (Elzinga and McGlothlin (2021)). To quantify the equilibrium effects of a fixed cost reduction, we consider a scenario where the fixed costs of the acquired craft products decrease by the estimated θ_1, which is the extra fixed cost faced by independent craft breweries. Recall that this parameter is identified by the difference in product decisions between independent craft breweries and craft breweries acquired by the macro breweries, i.e., products by the latter breweries are more likely to enter a market.

We report the new aggregate results in Table 10. We find that, on average, the merging firms now add products (Row (4)), but the number of new entrants decreases (Row (2)). Overall, the number of products now increases (Row (3)). At the same time, the merging firms increase prices by a bigger margin (Row (9)). In the end, the total consumer welfare loss is reduced but not reversed, with a lowered range of \$320,010 to \$512,790 (Row (13)). Unlike the merger efficiency associated with marginal costs, the efficiency associated with fixed costs appears to cause countervailing effects on the number of products: synergies in fixed costs induce new product entry by merging firms but discourage product entry by new

Table 10: Aggregate Post-Merger Outcomes: Merger Efficiencies

Average Change Per Market			Aggregate Change Across Markets		
(1)	# of firms	[-2.95, -2.93]	(10)	quantity (1000)	[-213.04, -127.44]
(2)	# new entrants	[0.00, 0.03]	(11)	craft	[-190.51, -99.72]
(3)	# of products	[0.00, 1.43]	(12)	craft, merging firms	[-229.06, -122.52]
(4)	merging firms	[-0.04, 1.42]	(13)	consumer surplus ($1000)	[-512.79, -320.01]
(5)	non-merging incumbents	[0.00, 0.01]	(14)	craft beer profits ($1000)	[89.96, 145.94]
(6)	new entrants	[0.00, 0.03]	(15)	merging firms	[26.77, 103.40]
(7)	average price ($)	[0.00, 0.01]	(16)	total surplus ($1000)	[-422.83, -174.07]
(8)	craft products ($)	[0.07, 0.10]	ΔCS decomposition ($1000)		
(9)	craft, merging firms ($)	[0.14, 0.19]	(17)	due to variety change	[-16.52, 176.26]
			(18)	due to entry	[0.59, 7.25]
			(19)	due to incumbent product adj.	[-23.77, 175.67]

Notes: this table reports the results when we take into account reductions in fixed costs when a craft brewery is acquired by a macro brewery. Rows (1)–(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)–(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.

entrants and non-merging incumbents, limiting the overall positive effect of efficiency gains on product variety and consumer welfare. This is consistent with the intuition that while prices are often strategic complements, product offerings tend to be strategic substitutes.

8 Conclusion

We propose a new method to estimate discrete games and apply it to study merger effects on firm entry, product choice, and prices in the retail craft beer market in California. The paper makes two contributions. Methodologically, we construct bounds on the probability of a single action instead of a market equilibrium, resulting in an estimator that is easy to compute and scalable to games with many firms or many firm decisions. Empirically, the paper adds to the literature of merger, entry, and product variety with a joint study of merger effects on firm entry, product variety, and prices. We consider a merger that significantly increases the concentration of the craft beer market. We find that new entries occur after such a merger. However, even in our entry-favorable setting, the impact of the new entries is insufficient to offset the overall negative effect of the merger. Potential merger efficiencies can reduce but do not reverse the consumer surplus loss.

References

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

Aradillas-Lopez, Andres and Elie Tamer (2008), "The identification power of equilibrium in simple games." *Journal of Business & Economic Statistics*, 26, 261–283.

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

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

Berry, Steven (1992), "Estimation of a model of entry in the airline industry." *Econometrica*, 60, 889–917.

Berry, Steven, Alon Eizenberg, and Joel Waldfogel (2016), "Optimal product variety in radio markets." *The RAND Journal of Economics*, 47, 463–497.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile prices in market equilibrium." *Econometrica*, 63, 841–890.

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

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Joonhwi Joo (2021), "Millennials and the take-off of craft brands: Preference formation in the US beer industry." working paper 28618, National Bureau of Economic Research.

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

Caradonna, Peter, Nathan Miller, and Gloria Sheu (2021), "Mergers, entry, and consumer welfare." *Georgetown McDonough School of Business Research Paper*.

Chernozhukov, Victor, Denis Chetverikov, and Kengo Kato (2019), "Inference on causal and structural parameters using many moment inequalities." *The Review of Economic Studies*, 86, 1867–1900.

Ciliberto, Federico, Charles Murry, and Elie Tamer (2021), "Market structure and competition in airline markets." *Journal of Political Economy*, 129, 2995–3038.

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

Codog, Aric (2018), "The antitrust roadblock: Preventing consolidation of the craft beer market." *University of the Pacific Law Review*, 50, 403.

Croxall, Daniel (2019), "Helping craft beer maintain and grow market shares with private enforcement of tied-house and false advertising laws." *Gonzaga Law Review*, 55, 167.

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

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

Elzinga, Kenneth G and Alexander J McGlothlin (2021), "Has Anheuser-Busch let the steam out of craft beer? The economics of acquiring craft brewers." *Review of Industrial Organization*, 1–27.

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

Fan, Ying and Chenyu Yang (2020), "Competition, product proliferation, and welfare: A study of the US smartphone market." *American Economic Journal: Microeconomics*, 12, 99–134.

Gandhi, Amit, Luke Froeb, Steven Tschantz, and Gregory Werden (2008), "Post-merger product repositioning." *The Journal of Industrial Economics*, 56, 49–67.

Garrido, Francisco Andres (2020), "Mergers between multi-product firms with endogenous variety: Theory and an application to the ready-to-eat cereal industry." working paper, ITAM.

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

Huang, Yufeng, Paul B Ellickson, and Mitchell J Lovett (2021), "Learning to set prices." *Journal of Marketing Research, forthcoming.*

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

Illanes, Gastón and Sarah Moshary (2020), "Market structure and product assortment: Evidence from a natural experiment in liquor licensure." working paper 27016, National Bureau of Economic Research.

Jeziorski, Przemysław (2015), "Empirical model of dynamic merger enforcement–choosing ownership caps in US radio." working paper, University of California, Berkeley.

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

Li, Sophia, Joe Mazur, Yongjoon Park, James Roberts, Andrew Sweeting, and Jun Zhang (2022), "Repositioning and market power after airline mergers." *The RAND Journal of Economics*, 53, 166–199.

McFadden, Daniel (1989), "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica*, 57, 995–1026.

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

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

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

Petrin, Amil (2002), "Quantifying the benefits of new products: The case of the minivan." *Journal of Political Economy*, 110, 705–729.

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

Satran, Joe (2014), "Here's how a six-pack of craft beer ends up costing \$12." *https://www.huffpost.com/entry/craft-beer-expensive-cost_n_5670015*.

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

Seim, Katja (2006), "An empirical model of firm entry with endogenous product-type choices." *The RAND Journal of Economics*, 37, 619–640.

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

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

Sweeting, Andrew (2013), “Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry.” *Econometrica*, 81, 1763–1803.

Tamayo, Andrew (2009), “What’s brewing in the old north state: An analysis of the beer distribution laws regulating North Carolina’s craft breweries.” *NCL Rev.*, 88, 2198.

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

Tremblay, Victor J, Natsuko Iwasaki, and Carol Horton Tremblay (2005), “The dynamics of industry concentration for US micro and macro brewers.” *Review of Industrial Organization*, 26, 307–324.

Wang, Shuang (2020), “Price competition with endogenous entry: The effects of Marriott & Starwood merger in Texas.” working paper.

Watson, Bart (2016), “On-premise beer data and craft.” *https://www.brewersassociation.org/insights/importance-on-premise-craft-brewers/*.

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

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

A Details on Micro Moments

In this section, we explain how we compute the model prediction for the micro moment $E\left(q_{i\tau}^{f'} \mid q_{i\tau}^{f} \geq 1\right)$. The calculation for other micro-moments in Section 5.1 is similar. In this moment, $q_{i\tau}^{f}$ is household i's quantity of beer with a certain flavor (f = ale, lager or light) or of a certain characteristic (f = import or craft) in year τ, and f' could be the same or a different type.

Let $s_{jmt}(\boldsymbol{\nu}, y)$ denote the Logit choice probability of product j in month t when the vector of unobserved tastes and log-income is $(\boldsymbol{\nu}, y)$. Let $G_m(\boldsymbol{\nu}, y)$ denote the distribution of $(\boldsymbol{\nu}, y)$, which can vary across markets and is thus indexed by m. We assume that each consumer has 8 opportunities to buy beer per month, which is the average per-household number of trips to the stores in the Nielsen Consumer Panel data.[22] Then, the probability that a household with values $(\boldsymbol{\nu}, y)$ buys type-f products in market m in year τ is

$$\psi_{m\tau}^{f}(\boldsymbol{\nu}, y) = 1 - \prod_{t \in \mathcal{T}_\tau} \left(1 - \sum_{j \in \mathcal{J}_{m\tau}^{f}} s_{jmt}(\boldsymbol{\nu}, y)\right)^{8},$$

where $\mathcal{J}_{m\tau}^{f}$ is the collection of all type-f products in market m in year τ. The conditional expectation of the annual purchase of type-f' beers for households in market m and year τ is, therefore,

$$E_{m\tau}\left(q_{i\tau}^{f'} \mid q_{i\tau}^{f} \geq 1\right) = \int_{\boldsymbol{\nu}, y} \frac{\sum_{t \in \mathcal{T}_\tau} \sum_{j \in \mathcal{J}_{m\tau}^{\tilde{f}}} 8 \cdot s_{jmt}(\boldsymbol{\nu}, y)}{\psi_{m\tau}^{f}(\nu, y)} dG_m(\boldsymbol{\nu}, y),$$

where $E_{m\tau}$ is the expectation conditional on a market m and a year τ. To obtain the average across market/year combinations, we weigh these conditional means in each market/year combination by the expected number of households who purchase type-f products, which is the product of the market size and the unconditional probability of purchasing type-f products in a market/year, i.e.,

$$weight_{m\tau} = MktSize_{m\tau} \cdot \int \psi_{m\tau}^{f}(\nu, y)\, dG_m(\boldsymbol{\nu}, y).$$

Therefore, the expected purchase of type-f' conditional on having at least one purchase of type f is

$$E\left(q_{i\tau}^{f'} \mid q_{i\tau}^{f} \geq 1\right) = \frac{\sum_{m\tau} E_{m\tau}\left(q_{i\tau}^{f'} \mid q_{i\tau}^{f} \geq 1\right) \cdot weight_{m\tau}}{\sum_{m\tau} weight_{m\tau}}.$$

[22]The demand estimates are robust to $5, 6, 7$ opportunities of purchase per month.

B Details on the Fixed Cost Estimation and Inference

We define our non-negative functions $g^{(k)}(X_{nm}, W_{jm})$ as functions of the extrema of the variable profit change $\left(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm})\right)$ and W_{jm}, where $\left(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm})\right)$ are summary statistics of a long vector of covariates capturing market structure and the characteristics of each potential product. The covariates W_{jm} are indicator variables, while the extrema of the variable profits $\left(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm})\right)$ are continuous variables. To define the functions $g^{(k)}(X_{nm}, W_{jm})$, we specify a series of cutoffs for $\underline{\Delta}_j(X_{nm})$ and $\overline{\Delta}_j(X_{nm})$, which are

$$D^{\Delta} = \{1,200, 4,800, 8,400, 12,000\}.$$

We use $W_{jm\lambda}$ to denote the λ^{th} component of the vector W_{jm}, $\lambda = 1, \ldots, \Lambda$. The indicator functions are

$$\mathbb{1}\left(\underline{\Delta}_j(X_{nm}) \geq D^{\Delta}_{l_1}, \overline{\Delta}_j(X_{nm}) \geq D^{\Delta}_{l_2}\right), \text{ for } l_1 = 1, ..., 4, l_2 = 1, ..., 4;$$

$$\mathbb{1}\left(\underline{\Delta}_j(X_{nm}) \geq D^{\Delta}_{l_1}, W_{jm\lambda} > 0\right), \text{ for } l_1 = 1, ..., 4, \lambda = 1, \ldots, \Lambda;$$

$$\mathbb{1}\left(\overline{\Delta}_j(X_{nm}) \geq D^{\Delta}_{l_2}, W_{jm\lambda} > 0\right), \text{ for } l_2 = 1, ..., 4, \lambda = 1, \ldots, \Lambda.$$

Defining the indicator functions in a pairwise fashion helps to avoid having too few non-zero elements in each moment, which could result in a noisy estimate of the moment's standard deviation and under- or over-coverage of the confidence set.

C 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 observed pre-merger 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 of firm n, we calculate the change of firm n's expected variable profit when product j enters the market, as defined by equation (9):

$$\Delta_j(Y_{-jm}, X_{nm}) = \pi_n(Y_{jm} = 1, Y_{-jm}, X_{nm}) - \pi_n(Y_{jm} = 0, Y_{-jm}, X_{nm}).$$

   Here, $Y_{-jm} = (Y_{j'm}, j' \in \mathcal{J}, j' \neq j)$ represents the observed entry outcomes of all potential products other than j in market m.

If j is in the market before the merger, we define a range $(-\infty, \Delta_j(X_{nm}))$, and otherwise, we define a range $(\Delta_j(X_{nm}), \infty)$.

2. We simulate draws of the fixed costs for firm n from a truncated normal distribution with the underlying normal distribution parameterized by mean $W_{jm}\hat{\theta}$ and variance $\hat{\sigma}_m^2$. The support of the truncated distribution is defined by the ranges in Step 1. These draws satisfy the necessary conditions for the observed equilibrium.

3. For each draw from Step 2, we check whether firm n's best response to Y_{-nm} is indeed Y_{nm}, where Y_{nm} is firm n's product decisions in market m and Y_{-nm} represents its opponents' decisions in the market. We find each firm's best response by employing the algorithm in Fan and Yang (2020) using the starting points $Y_{nm}^0 = (0, ..., 0)$ and $Y_{nm}^0 = (1, ..., 1)$. If the algorithm converges to Y_{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 Y_{nm}, we go back to Step 2 and re-draw the fixed costs.

4. We repeat this process for every firm n.

D (Dis-)Economies of Scope

In this section, we extend our model to allow for economies or dis-economies of scope in fixed costs and derive a new set of inequalities bounding the entry probability of a firm in addition to that of a product in order to estimate the additional parameter. We find economies of scope in the medium and large markets. Our merger simulation results are, however, robust.

Our fixed cost specification in the main text of the paper is additively separable across products and, thus, does not allow for economies or dis-economies of scope. We consider the following extension of the fixed cost function:

$$\theta_{0m}\mathbb{1}\left(\sum_{j\in\mathcal{J}_n} Y_{jm} > 0\right) + \sum_{j\in\mathcal{J}_n} Y_{jm}\left(W_{jm}\theta + \sigma_m\zeta_{jm}\right),$$

which is no longer additive in the fixed cost of each product. If $\theta_{0m} > 0$, the firm faces a firm-level entry cost in addition to the product-level entry costs, and the fixed cost exhibits economies of scope. Conversely, if $\theta_{0m} < 0$, the cost function exhibits dis-economies of scope.

To estimate θ_{0m}, we additionally consider bounds for the conditional probability that a firm has at least one product in a market, denoted by $\Pr(\sum_{j\in\mathcal{J}_n} Y_{jm} > 0 \,|X_{nm}, W_{nm})$, where

$W_{nm} = (W_{jm}, j \in \mathcal{J}_n)$ is the collection of the fixed cost covariates for firm n's potential products in market m. We define $\zeta_{nm} = (\zeta_{jm}, j \in \mathcal{J}_n)$ analogously. Also, let $Y_{nm} = (Y_{jm}, j \in \mathcal{J}_n)$ be firm n's product decision in market m and Y_{-nm} be the opponents' decisions. Finally, we denote firm n's maximum profit from entering the market for given $(Y_{-nm}, X_{nm}, W_{nm})$ by

$$\begin{aligned} &\Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm}, \theta) \\ =& \max_{Y_{nm} \text{ s.t. } \sum_{j\in\mathcal{J}_n} Y_{jm}>0} \pi_n(Y_{nm}, Y_{-nm}, X_{nm}) - \sum_{j\in\mathcal{J}_n} Y_{jm}\left(W_{jm}\theta + \sigma_m \zeta_{jm}\right) - \theta_{0m}. \end{aligned}$$

We define the minimum

$$\underline{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta) = \min_{Y_{-nm}} \Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm}, \theta)$$

and the maximum

$$\overline{\Gamma}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta) = \max_{Y_{-nm}} \Gamma_{nm}(Y_{-nm}, X_{nm}, W_{nm}, \zeta_{nm}, \theta).$$

Under the assumption that the observed brewery entry decisions are not dominated, the bounds for $\Pr(\sum_{j\in\mathcal{J}_n} Y_{jm} > 0 \,|X_{nm}, W_{nm})$ are

$$\begin{aligned} &\Pr\left(\underline{\Gamma}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) > 0\right) \\ &\le \Pr(\sum_{j\in\mathcal{J}_n} Y_{jm} > 0 \,|X_{nm}, W_{nm}) \\ &\le \Pr\left(\overline{\Gamma}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) > 0\right). \end{aligned}$$

Note that $\pi_n((1,1,0), Y_{-nm}, X_{nm})$ can be written as the sum of two differences: $\pi_n((1,1,0), Y_{-nm}, X_{nm}) - \pi_n((1,0,0), Y_{-nm}, X_{nm})$ and $\pi_n((1,0,0), Y_{-nm}, X_{nm}) - \pi_n((0,0,0), Y_{-nm}, X_{nm})$. By this observation and by the definition of $\underline{\Delta}_j(X_{nm})$ and $\overline{\Delta}_j\left(X_{nm}\right)$ in Section 2.2, we have

$$\sum_{j\in\mathcal{J}_n} Y_{jm}\underline{\Delta}_j\left(X_{nm}\right) \le \pi_n\left(Y_{nm}, Y_{-nm}, X_m\right) \le \sum_{j\in\mathcal{J}_n} Y_{jm}\overline{\Delta}_j\left(X_m\right).$$

Define

$$\begin{aligned} \tilde{\underline{\Gamma}}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) &= \max_{Y_{nm} \text{ s.t. } \sum_{j\in\mathcal{J}_n} Y_{jm}>0} \sum_{j\in\mathcal{J}_n} Y_{jm}\left(\underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m\zeta_{jm}\right) - \theta_{0m}, \\ \tilde{\overline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta) &= \max_{Y_{nm} \text{ s.t. } \sum_{j\in\mathcal{J}_n} Y_{jm}>0} \sum_{j\in\mathcal{J}_n} Y_{jm}\left(\overline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m\zeta_{jm}\right) - \theta_{0m}. \end{aligned} \tag{D.1}$$

We have $\underline{\tilde{\Gamma}}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) \leq \underline{\Gamma}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right)$ by the max-min inequality and $\overline{\Gamma}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) \leq \tilde{\overline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta)$ by the definition of $\tilde{\overline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta)$.

The integer programming problem in (D.1) can be solved quickly given the additive structure. Specifically, $\underline{\tilde{\Gamma}}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right)$ can be written as

$$
\begin{aligned}
&\underline{\tilde{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta) \\
=& -\theta_{0m} + \begin{cases} \sum_{j \in \mathcal{J}_n} \left\| \underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} \right\|_+ & \text{if } \exists j \text{ s.t. } \underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} > 0, \\ \max_{j \in \mathcal{J}_n} \left\{ \underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} \right\} & \text{otherwise.} \end{cases}
\end{aligned}
$$

In other words, we simply need to calculate the values of each $\underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm}$, sum up all the positive terms, and subtract θ_{0m}. If $\underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} < 0$ for all $j \in \mathcal{J}_n$, we calculate the bound as the maximum of $\underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} - \theta_{0m}$. Similarly, $\tilde{\overline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta)$ is given by

$$
\begin{aligned}
&\tilde{\overline{\Gamma}}_{nm}(X_{nm}, W_{nm}, \zeta_{nm}, \theta) \\
=& -\theta_{0m} + \begin{cases} \sum_{j \in \mathcal{J}_n} \left\| \overline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} \right\|_+ & \text{if } \exists j \text{ s.t. } \overline{\Delta}_j\left(X_{nm}\right) - \theta W_{jm} - \zeta_{jm} > 0, \\ \max_{j \in \mathcal{J}_n} \left\{ \overline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \sigma_m \zeta_{jm} \right\} & \text{otherwise.} \end{cases}
\end{aligned}
$$

In the end, we use the following as the lower and upper bounds of the entry probability

$$
\begin{aligned}
&\Pr\left(\underline{\tilde{\Gamma}}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) > 0\right) \\
&\leq \Pr\left(\sum_{j \in \mathcal{J}_n} Y_{jm} > 0 \,|X_{nm}, W_{nm}\right) \\
&\leq \Pr\left(\tilde{\overline{\Gamma}}_{nm}\left(X_{nm}, W_{nm}, \zeta_{nm}, \theta\right) > 0\right).
\end{aligned}
$$

These bounds do not have analytic expressions and we simulate them by taking 100 draws of ζ_{nm}.

Estimation Details

We construct unconditional moments similar to those in Section 2. Specifically, we use $\left(\underline{\Delta}_j(X_{nm}), \overline{\Delta}_j(X_{nm})\right)$ of the top three most profitable products of firm n to define the indicator functions $g^{(k)}(\cdot)$ as in Supplemental Appendix B. If a firm has only one or two potential products, we set indicator functions corresponding with the unavailable products to be 0.

We also modify the bounds of the conditional choice probability of an individual product's

Table D.1: Fixed Cost Estimates: Projected 95% Confidence Interval, Allowing for (Dis-)Economies of Scope

	Baseline Model	(Dis-)Economies of Scope
Craft (θ_1)	[229.14, 1093.24]	[949.06, 1590.55]
In State× Craft (θ_2)	[-387.82, 208.18]	[-1733.91, -1209.29]
Market-size specific fixed cost (θ_3)		
Small market	[308.95, 938.33]	[658.96, 1644.17]
Medium market	[1027.77, 1468.10]	[655.01, 952.30]
Large market	[3325.71, 4177.69]	[1966.51, 3307.71]
Market-size specific std. dev. (σ_ζ)		
Small market	[0.00, 522.79]	[0.00, 205.29]
Medium market	[679.41, 863.25]	[0.00, 464.35]
Large market	[2511.65, 3424.06]	[1912.77, 2449.15]
Market-size specific firm entry cost (θ_0)		
Small market		[-826.28, 277.07]
Medium market		[433.22, 952.24]
Large market		[865.57, 2494.82]

Notes: Estimates in 2016 US dollars.

outcome to take into account θ_{0m}:

$$
\begin{aligned}
& F_\zeta\left(\zeta_{jm} < \underline{\Delta}_j\left(X_{nm}\right) - W_{jm}\theta - \left\|\theta_{0m}\right\|_+, \sigma_\zeta\right) \\
\leq & \Pr\left(Y_{jm} = 1 \left| X_{nm}, W_{jm}\right.\right) \\
\leq & F_\zeta\left(\zeta_{jm} < \overline{\Delta}_j\left(X_{nm}\right) - \theta W_{jm} - \left\|\theta_{0m}\right\|_-, \sigma_\zeta\right).
\end{aligned}
$$

The same set of g functions are used to construct the moments associated with individual products.

We combine moments associated with firm and product entry in the estimation. We report the 95% projected confidence interval in Table D.1. For comparison, we also copy the results from the baseline specification in the first column. Using the specification allowing for (dis-) economies of scope, we estimate a larger cost for craft, but also larger savings for in-state crafts. We find evidence for economies of scope in the medium and large markets.

The counterfactual results on the heterogeneous merger effects in Figures D.1 and D.2 are noisier but comparable to the baseline results. For example, we still find that medium-sized markets suffer more per-capita consumer welfare loss than the largest ones. The aggregate results in Table D.2 are also similar to the baseline results.

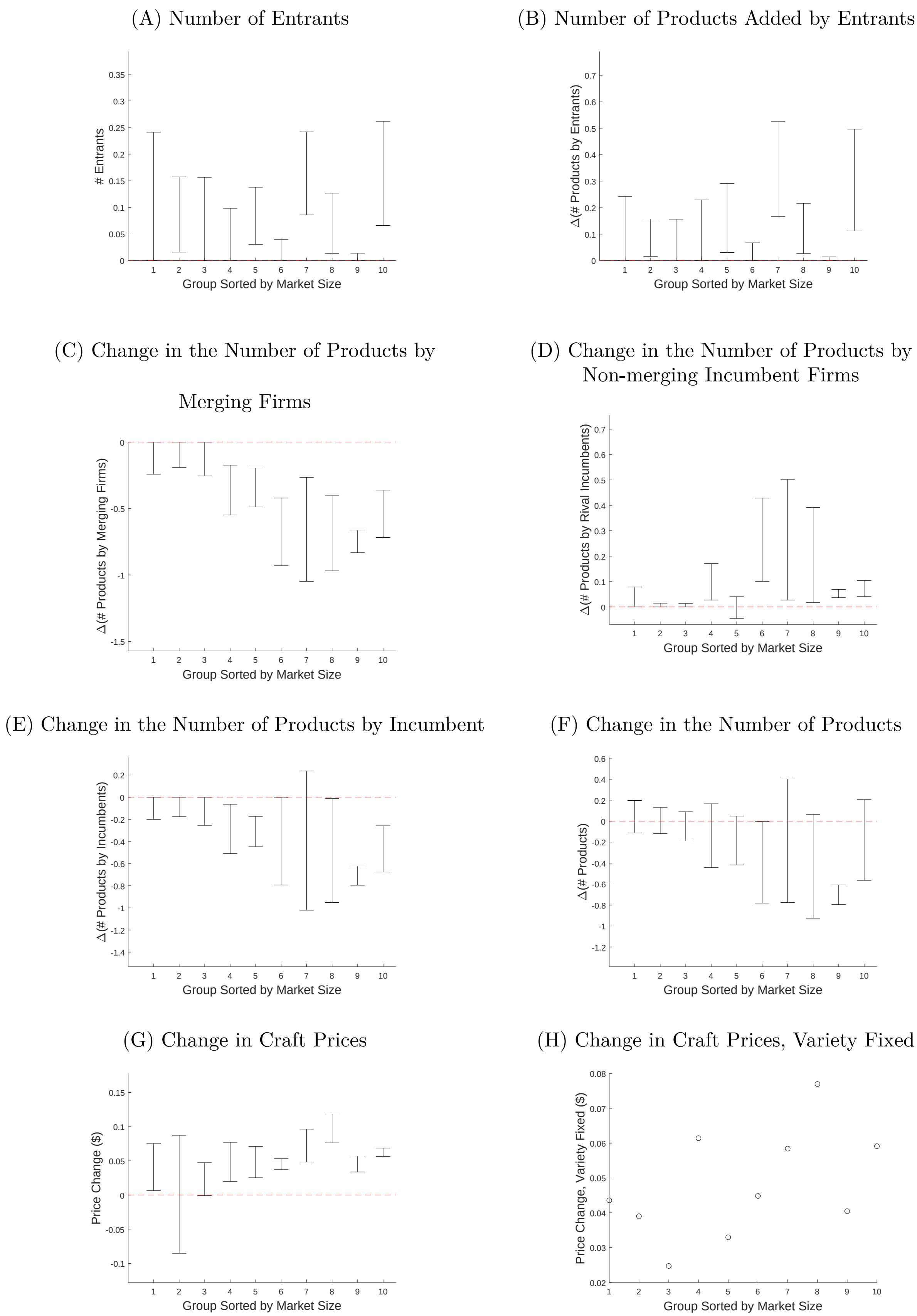

Figure D.1: Product Variety, Entry and Prices

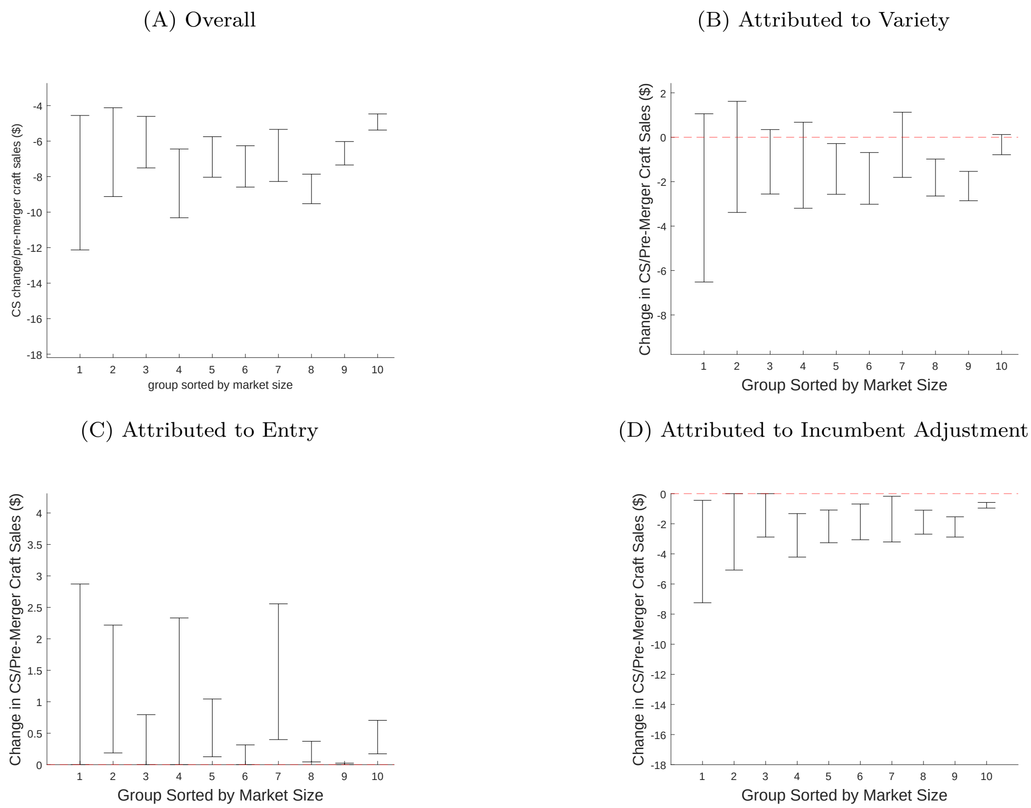

Figure D.2: Change in CS/Pre-Merger Craft Sales

Table D.2: Aggregate Post-Merger Outcomes: Economies of Scope

Average Change Per Market		Aggregate Change Across Markets	
(1) # of firms	[-2.90, -2.79]	(10) quantity (1000)	[-254.19, -225.98]
(2) # new entrants	[0.05, 0.16]	(11) craft	[-232.74, -204.14]
(3) # of products	[-0.54, -0.05]	(12) craft, merging firms	[-292.32, -266.06]
(4) merging firms	[-0.69, -0.48]	(13) consumer surplus ($1000)	[-606.00, -537.74]
(5) non-merging incumbents	[0.06, 0.15]	(14) craft beer profits ($1000)	[88.91, 94.92]
(6) new entrants	[0.09, 0.31]	(15) merging firms	[12.41, 18.63]
(7) average price ($)	[0.00, 0.00]	(16) total surplus ($1000)	[-511.07, -446.51]
(8) craft products ($)	[0.05, 0.07]	ΔCS decomposition ($1000)	
(9) craft, merging firms ($)	[0.14, 0.14]	(17) due to variety change	[-109.73, -41.47]
		(18) due to entry	[19.01, 50.59]
		(19) due to incumbent product adj.	[-139.09, -79.06]

Notes: Rows (1)–(9) on the left report the weighted average changes, where the simulated expected changes in each market are weighted by the market size. Rows (10)–(19) on the right report the total changes, where the simulated expected changes are summed across markets. We report the range of estimates across the vectors of fixed cost parameters sampled from their 95% confidence set.

E Correlated Unobserved Fixed Costs

In the main text of the paper, we assume that the unobservable fixed cost shocks are i.i.d. In this section, we consider an extension to allow for correlated shocks. We start with explaining how an additional set of bounds can be constructed and used to identify the correlation using the illustrative example in Section 2.1. We then extend our empirical model in Section 4 and show that our merger simulation results are robust to allowing for a market-wide shock common to all products.

Extension to the Illustrative Model and Additional Bounds

To estimate the correlation in unobserved cost shocks, we bound the joint entry probabilities of multiple firms or products. We first explain the extension using the illustrative example in Section 2.1. We consider bounding the probability of $\Pr\left(Y_{1m}=1, Y_{2m}=1\right)$. The level-1 rationality assumption implies the following:

$$
\begin{aligned}
&\Pr\left(Y_{nm}=1 \text{ is a dominant strategy for both } n=1,2\right)\\
\leq &\Pr\left(Y_{1m}=1, Y_{2m}=1\right)\\
\leq &\Pr\left(Y_{nm}=1 \text{ is not a dominant strategy for both } n=1,2\right).
\end{aligned}
$$

To save notation, we define $\overline{\pi}_{nm}=\max\left(\pi_{nm}\left(0\right), \pi_{nm}\left(1\right)\right)$ and $\underline{\pi}_{nm}=\min\left(\pi_{nm}\left(0\right), \pi_{nm}\left(1\right)\right)$. We also use $F_{\zeta_1\zeta_2}$ to denote the joint distribution of $\left(\zeta_{1m}, \zeta_{2m}\right)$. The bounds above can be expressed as

$$
\begin{aligned}
&F_{\zeta_1\zeta_2}\left(\underline{\pi}_{1m}-C_{1m}, \underline{\pi}_{2m}-C_{2m}\right)\\
\leq &\Pr\left(Y_{1m}=1, Y_{2m}=1\right)\\
\leq &F_{\zeta_1\zeta_2}\left(\overline{\pi}_{1m}-C_{1m}, \overline{\pi}_{2m}-C_{2m}\right),
\end{aligned} \tag{E.1}
$$

which can be rewritten as

$$
\begin{aligned}
&\Pr\left(\zeta_{1m}<\underline{\pi}_{1m}-C_{1m}\left|\zeta_{2m}<\underline{\pi}_{2m}-C_{2m}\right.\right)\cdot\Pr\left(\zeta_{2m}<\underline{\pi}_{2m}-C_{2m}\right)\\
\leq &\Pr\left(Y_{1m}=1, Y_{2m}=1\right)\\
\leq &\Pr\left(\zeta_{1m}<\overline{\pi}_{1m}-C_{1m}\left|\zeta_{2m}<\overline{\pi}_{2m}-C_{2m}\right.\right)\cdot\Pr\left(\zeta_{2m}<\overline{\pi}_{2m}-C_{2m}\right).
\end{aligned}
$$

As the correlation between ζ_{1m} and ζ_{2m} increases, both $\Pr(\zeta_{1m}<\underline{\pi}_{1m}-C_{1m}\left|\zeta_{2m}<\underline{\pi}_{2m}-C_{2m}\right.)$ and $\Pr\left(\zeta_{1m}<\overline{\pi}_{1m}-C_{1m}\left|\zeta_{2m}<\overline{\pi}_{2m}-C_{2m}\right.\right)$ increase, making the lower bound more likely to be violated. Conversely, as the correlation decreases, both conditional probabilities decrease, making the upper bound more likely to be violated.

Extension to the Empirical Model and Robustness of Our Results

We now extend the fixed cost specification to allow for a market level unobserved cost component ζ_m:

$$W_{jm}\theta + \sigma_m \zeta_{jm} + \sigma \zeta_m,$$

and estimate the standard deviation σ. We assume ζ_m is i.i.d. across markets and has a standard normal distribution.

In this model, for two potential products (j, j'), the analogy of inequality (E.1) is:

$$\begin{aligned}
&\Phi\left(\underline{\Delta}_j(X_{nm}) - c(W_{jm},\theta), \underline{\Delta}_{j'}(X_{nm}) - c(W_{j'm},\theta); \Sigma_m\right) \\
\leq &\Pr\left(Y_{jm}=1, Y_{j'm}=1 \,|X_{nm}, W_{jm}, W_{j'm}\right) \\
\leq &\Phi\left(\overline{\Delta}_j(X_{nm}) - c(W_{jm},\theta), \overline{\Delta}_{j'}(X_{nm}) - c(W_{j'm},\theta); \Sigma_m\right),
\end{aligned}$$

where Φ is the bivariate normal distribution with zero means and the covariance matrix

$$\Sigma_m = \begin{pmatrix} \sigma_m^2 + \sigma^2 & \sigma^2 \\ \sigma^2 & \sigma_m^2 + \sigma^2 \end{pmatrix}.$$

The resulting new moment functions are

$$\begin{aligned}
&L\left(Y_{jm}, Y_{j'm}, X_{nm}, W_{jm}, W_{j'm}, \Sigma_m, \theta\right) \\
&= \Phi\left(\underline{\Delta}_j(X_{nm}) - c(W_{jm},\theta), \underline{\Delta}_{j'}(X_{nm}) - c(W_{j'm},\theta); \Sigma_m\right) \\
&- \mathbb{1}\left(Y_{jm}=1, Y_{j'm}=1\right)
\end{aligned}$$

and

$$\begin{aligned}
&H\left(Y_{jm}, Y_{j'm}, X_{nm}, W_{jm}, W_{j'm}, \Sigma_m, \theta\right) \\
&= \mathbb{1}\left(Y_{jm}=1, Y_{j'm}=1\right) \\
&- \Phi\left(\overline{\Delta}_j(X_{nm}) - c(W_{jm},\theta), \overline{\Delta}_{j'}(X_{nm}) - c(W_{j'm},\theta); \Sigma_m\right).
\end{aligned}$$

We combine the moments based on a single product's entry probability in the main text of

Table E.1: Fixed Cost Estimates: Projected 95% Confidence Interval, Allowing for a Market-Level Shock

	Baseline Model	With Market-level Unobserved Costs
Craft (θ_1)	[229.14, 1093.24]	[288.43, 980.55]
In State× Craft (θ_2)	[-387.82, 208.18]	[-397.04, 206.30]
Market-size specific fixed cost (θ_3)		
Small market	[308.95, 938.33]	[407.24, 718.54]
Medium market	[1027.77, 1468.10]	[1094.22, 1409.10]
Large market	[3325.71, 4177.69]	[2927.78, 4116.15]
Market-size specific std. dev. (σ_ζ)		
Small market	[0.00, 522.79]	[0.00, 9.95]
Medium market	[679.41, 863.25]	[659.03, 777.36]
Large market	[2511.65, 3424.06]	[2173.87, 3277.92]
Market-level unobserved cost std. dev. (σ)		[18.59, 85.94]

Notes: Estimates in 2016 US dollars.

the paper with the following additional moments in estimation:

$$E\left[\frac{1}{\frac{1}{2}|J|\left(|J|-1\right)}\sum_{j=2}^{J}\sum_{j'=1}^{j-1}L\left(Y_{jm},Y_{j'm},X_{nm},W_{jm},W_{j'm},\Sigma_m,\theta\right)\cdot g^{(k)}\left(X_{nm},W_{jm}\right)\right]\leq 0,$$

$$E\left[\frac{1}{\frac{1}{2}|J|\left(|J|-1\right)}\sum_{j=2}^{J}\sum_{j'=1}^{j-1}L\left(Y_{jm},Y_{j'm},X_{nm},W_{jm},W_{j'm},\Sigma_m,\theta\right)\cdot g^{(k)}\left(X_{nm},W_{j'm}\right)\right]\leq 0,$$

$$E\left[\frac{1}{\frac{1}{2}|J|\left(|J|-1\right)}\sum_{j=2}^{J}\sum_{j'=1}^{j-1}H\left(Y_{jm},Y_{j'm},X_{nm},W_{jm},W_{j'm},\Sigma_m,\theta\right)\cdot g^{(k)}\left(X_{nm},W_{jm}\right)\right]\leq 0,$$

$$E\left[\frac{1}{\frac{1}{2}|J|\left(|J|-1\right)}\sum_{j=2}^{J}\sum_{j'=1}^{j-1}H\left(Y_{jm},Y_{j'm},X_{nm},W_{jm},W_{j'm},\Sigma_m,\theta\right)\cdot g^{(k)}\left(X_{nm},W_{j'm}\right)\right]\leq 0,$$

where $g^{(k)}$ is defined in Supplemental Appendix B.

Table E.1 reports our estimation results. Compared with the baseline results in the first column, accounting for market-level cost unobservables shifts the confidence intervals for the standard deviation of the product/market-level shock ζ_{jm}, although the confidence intervals still overlap with the baseline estimates. The standard deviation of the market-level shock ζ_m is estimated to be fairly modest compared to the standard deviation of the product/market-level shock.