Identifying Scope Economies using Demand-Side Data*

Ekaterina Khmelnitskaya† Guillermo Marshall‡ Scott Orr§

March 15, 2023

Abstract

One of the primary reasons a firm may choose to expand its horizontal boundaries is to take advantage of economies of scope, where a firm's per unit costs fall as they produce more products. What is the magnitude of scope economies? What are their impacts on market outcomes? In this paper, we propose a simple, empirically tractable method to identify the importance of economies of scope for market competition. We start with the standard "demand-side" approach to estimating marginal costs, which makes use of a firm's pricing first-order conditions and the estimated demand system. We augment this approach by introducing a micro-founded model of production by a multi-product firm. Our method allows us to generate a set of estimating equations for the cost function parameters that govern the extent of scale and scope economies, together with the distribution of within-firm productivity. We apply this approach to the U.S. beer industry and quantify the importance of scope economies for productive efficiency, prices, and output levels.

Keywords: Economies of scope, boundaries of the firm, productivity, multiproduct firms

*We thank Nate Miller and Matt Weinberg for generously sharing data and code with us. All errors are our own.

†Sauder School of Business, University of British Columbia, ekaterina.khmelnitskaya@sauder.ubc.ca.

‡Sauder School of Business, University of British Columbia, guillermo.marshall@sauder.ubc.ca.

§Sauder School of Business, University of British Columbia, scott.orr@sauder.ubc.ca.

1 Introduction

Multiproduct firms have come to dominate industrial production (Bernard et al., 2010; Goldberg et al., 2010). Economies of scope—cost savings that arise due to the scope of production—have been proposed as one explanation for the existence of multiproduct firms (Panzar and Willig, 1975; Teece, 1980; Panzar and Willig, 1981). Are economies of scope empirically relevant for a firm's decision to expand its horizontal boundaries? We shed light on this question by proposing a new method to estimate economies of scope and scale and applying it to US beer industry data. Using our estimates, we quantify the impact of scope economies on productive efficiency and market outcomes.

We propose a multiproduct model of production that allows for (but does not impose) within-firm productivity dispersion (i.e., productivity variation at the product level) as well as scope and scale economies. From the production model, we derive the cost function of a multiproduct firm, which crucially depends on parameters governing economies of scope and scale. We derive two estimating equations that depend on technology parameters and the output levels, productivities, and marginal costs of the products manufactured by a firm.

We then use our methodology to investigate the existence of scope economies in the US beer industry. This industry is ideal for two reasons. First, the main players are multiproduct firms (e.g., Anheuser-Busch, Molson Coors, SABMiller, Grupo Modelo, among other active firms in our sample period). Second, firms in the industry produce using a small number of plants despite high transportation costs. For example, Molson Coors had two plants (Colorado and Virginia) serving the entire United States up until 2008. This contrasts with other industries where local production is preferred to save on transportation costs (e.g., the US carbonated beverage industry). These facts combined are consistent with scale and scope economies at the brewery level.

Our main data source is the IRI Marketing Dataset (Bronnenberg et al., 2008), which provides price and sales data at the store–week–product level, where a product is defined as a brand–size combination. We focus on the years 2005 to 2008. Our method does not require observing input data either at the firm or firm–product level. In fact, we show that demand-side data (quantities sold, prices, product characteristics, price instruments, etc.) is sufficient to recover all the production technology parameters that enter the marginal cost function, including those governing scale and scope economies. This is a strength of our method, as it can be used in other industries where scanner data (rather than production data) is available.

In estimating our model, we face two econometric challenges. First, we do not observe marginal costs in our data. To deal with this, we use a demand-side approach to recover estimates of each product's marginal cost. Specifically, we estimate demand, set up a product market pricing game, and recover marginal costs from the equilibrium conditions of this game (e.g., see Berry et al., 1995). Second, the firm chooses the output level for each product (in part) based on the productivity of the corresponding product line, which is unobserved. We use demand shifters as instruments to tackle this endogeneity concern.

Our estimates for the US beer industry suggest the existence of both scale economies and scope economies. We use these estimates to measure the impact of scope economies on productive efficiency and market outcomes. In our counterfactual exercise, we shut down scope economies (keeping all other aspects of the production technology fixed) and compute each product's marginal cost at the output levels observed in our data without letting firms reoptimize production. We find that shutting down scope economies increases marginal costs by 25 percent, suggesting that scope economies significantly affect production efficiency.

How does this increase in marginal costs impact pricing and production decisions? On the one hand, the increase in the marginal cost of a product caused by the shutdown of scope economies decreases the marginal incentive to sell an extra unit of that good, incentivizing a price increase. On the other hand, our estimates suggest the existence of economies of scale, making it costly to cut down production, as this would further inflate marginal costs. Economies of scale make price increases costly, creating a tradeoff.

We compute the counterfactual market equilibrium without scope economies to study how these two forces play out in equilibrium. We find that marginal costs increase by 26.3 percent relative to the equilibrium with scope economies. The effect on marginal costs is magnified by a decrease in output (market shares decrease on average by 1.7 percent). That is, scope economies are stronger than scale economies for production decisions, in the sense that firms choose to cut production despite scale economies. We also find that prices increase by 13.7 percent on average in the equilibrium without scope economies. These findings combined suggest that scope economies have a first-order effect on productive efficiency—providing an (at least partial) explanation for why multiproduct production is favored in this industry—and market outcomes.

We contribute to several strands in the literature. First, we contribute to the literature on testing for the existence of non-joint production and scope economies. Previous approaches have either relied on cost function estimation using firm-level cost data (Hall 1973, Kohli 1981, Baumol et al. 1982, Johnes 1997, Zhang and Malikov 2022) or estimation of multi-output technologies using transformation functions (Dhyne et al. 2022, Maican and Orth 2020). These approaches require high-quality data on inputs and costs, which in practice is difficult to find for many industries, and may be prone to measurement error.[1] Our paper,

[1]See, for example, the recent discussion in De Loecker and Syverson (2021).

on the other hand, provides a way to test for and quantify non-joint production by relying only on *demand-side* data, i.e. prices, quantities, and market shares. Importantly, we do not require that a researcher have access to any input or cost data. Instead, our approach builds on the demand-side approach to cost estimation, pioneered in Rosse (1970), and further developed by Berry et al. (1995), Nevo (2000), and Berry and Haile (2014), where a firm's pricing first order conditions are used to back out point estimates of marginal cost. We consider a simple parameterization of a firm's cost function that allows for joint and non-joint production, and show how to generate simple estimating equations for parameters governing scale and scope economies.[2]

In this sense, our paper is closely related to Ding (2020) and Argente et al. (2020), who also provide evidence of scale and scope economies. While we share an interest in many of the same questions, we differ from these papers in a number of important ways. To estimate and quantify scale and scope economies, Ding (2020) proposes a model of joint production driven by public inputs that generate ideas that can be applied to various industries within a multi-industry conglomerate. Argente et al. (2020) considers an alternative model where a firm can invest in firm-wide or product-specific knowledge. Our model largely differs from these papers by relying on a microfoundation for joint production based on public or non-rival production inputs, as in Baumol et al. (1982), rather than scope economies generated by knowledge or idea generation. We also provide a complementary "micro" study — focused on a single industry, beer — to complement the more aggregate "macro", across-industry, approach employed in these studies.

Second, we contribute to the body of work investigating various productivity and competition issues in the US beer industry (Ashenfelter et al., 2015; De Loecker and Scott, 2016; Miller and Weinberg, 2017; Grieco et al., 2018; Miller et al., 2021). Our key contribution is to investigate the existence of scope economies and study their implication for efficiency and market outcomes, which sets us apart from prior work.

[2] The particular form of the cost function we rely on has been used in previous literature (Baumol et al. 1982, Johnes 1997). However, our estimation approach is more flexible than previous approaches, since we explicitly allow for the existence of firm-product-market specific shocks to productivity, which we are able to recover naturally using our marginal cost inversion, relying on insights from Orr (2022) and Cairncross et al. (2023).

2 The US Beer Industry

2.1 Industry

Beer is produced by combining malts, barley, hops, yeast, water, and sometimes other ingredients for added flavor, and letting the resulting liquid ferment in barrels. After that, the drink is packaged and shipped to the market. In the U.S., traditional brewers are associated with the first phase of a so-called "three-tier" system of brewer/distributor or wholesaler/retailer, with regulation severely limiting vertical ownership between the tiers. Large brewers generally are prohibited from selling directly to retailers and are required to sell their products through independent state-licensed distributors.[3] Nevertheless, Ascher (2012) suggests that within this relationship, wholesalers tend to depend on brewers to set final prices for the products. Moreover, while distributors manage shipping and selling products to retailers, brewers incur the bulk of the shipping cost associated with delivering the products to distribution points that are close to the market.

While the beer industry was fragmented at its inception, with thousands of local breweries, the concentration has been increasing. Ascher (2012) reports that by 2002, the number of traditional breweries dropped to around 22. Market concentration in the industry has also been increasing, with the top five overarching brands accounting for about 80% of sales by 2001 (Miller and Weinberg, 2017).[4] While there are many non-traditional or craft breweries, they tend to be small and local, in total accounting for less than 6% of total sales by volume during the time period in our data (Ascher, 2012). These trends are described in more detail in Grieco et al. (2018).

The trend toward the concentration of traditional brewing in a small number of locations might suggest a significant increase in returns to scale and scope on the brewery level. Multiple sources agree that the technological shift in the 1960s and 1970s induced these changes (Kerkvliet et al., 1998; Ascher, 2012; Keithahn, 1978).[5] The technological advancements include improvements in the bottling/canning and packaging technology; automation of the brewing technology, which allowed large brewhouses to use the same amount of labor as much smaller ones; and innovations in the fermentation process (see Keithahn (1978), p.

[3] Many states make exemptions for microbreweries.

[4] Miller and Weinberg (2017) use the following 5 overarching brand categories: ABI, Miller, Coors, Modelo, and Heineken. These brand categories each aggregate multiple smaller brands. For example, Coors category includes brands such as Coors and Coors Light, among others.

[5] Before that, scale and scope economies were present but moderate, originating mostly from general brewery overhead and utilities. Keithahn (1978, p. 33) suggests that those included "the cost of wells, water-processing equipment, sewage facilities, refrigeration equipment, management, laboratories, and custodial costs".

34-39 for more detail).[6]. On the national level, Ascher (2012) points towards the importance of economies of scale in advertising.

2.2 Data

Our main data source is the IRI Marketing Dataset (see Bronnenberg et al., 2008 for a detailed description), which provides price and sales data at the store–week–product level for the years 2001 to 2012. A product is defined as a brand–size combination. We follow the replication package of Miller and Weinberg (2017) to prepare the data for estimation and restrict attention to January 2005 to May 2008, right before the Miller–Coors joint venture was completed. We focus on this period to abstract away from the price effects of the Miller–Coors joint venture (Miller and Weinberg, 2017; Miller et al., 2021).

Following Miller and Weinberg (2017), we complement these data with the Public Use Microdata Sample (PUMS) of the American Community Survey. We use these data to incorporate demographic variables into the demand system. For every geographic area in the IRI data, we use 500 draws of the distribution of income per person. We use the same draws used by Miller and Weinberg (2017). Lastly, we use data on the distance from every geographic market in the IRI dataset to the nearest brewery making each product in the data, as described in Miller and Weinberg (2017).

Table 1 presents summary statistics of the products in our sample. A product is defined as a brand–size combination, and we focus on three sizes: 6-pack equivalent, 12-pack equivalent, and 24/30-pack equivalent. The table shows that firms in the industry produce multiple beer brands and sizes for each brand, making all firms in our sample multiproduct firms.

3 Theory

3.1 Demand

Our empirical model uses the random coefficients nested logit model estimated in Miller and Weinberg (2017), which we briefly summarize here. The conditional indirect utility that consumer c receives from purchasing product $(i, j) \in \Omega_t$, where i indexes a firm and j indexes a particular product in a market t (a city-time period combination) is given by:

$$u^j_{cit} = \delta^j_{it} + \mu^j_{cit} + \zeta_{cg(i,j)t}(\varrho) + (1 - \varrho)\epsilon^j_{cit}, \tag{1}$$

[6] Notice that some of these technologies might allow for both economies of scale and scope. For example, brewing large quantities of the same type of beer or brewing multiple different types of beers in a large brewery might result in the same labor savings compared to a smaller plant.

Table 1: Summary Statistics of the US Beer Industry in 2007

Brand	(1) Parent	(2) Average price	(3) Share of revenue	(4) Sizes	(5) Observations
Blue Moon Belgian White Ale	Molson Coors	13.721	0.006	3	898
Bud Light	Anheuser-Busch	9.114	0.083	3	1,535
Budweiser	Anheuser-Busch	9.116	0.040	3	1,513
Budweiser Select	Anheuser-Busch	9.186	0.007	3	1,417
Busch	Anheuser-Busch	6.774	0.010	3	1,300
Busch Light	Anheuser-Busch	6.733	0.014	3	1,306
Coors	Molson Coors	9.110	0.004	3	1,338
Coors Light	Molson Coors	9.106	0.042	3	1,521
Corona Extra	Grupo Modelo	14.325	0.050	3	1,212
Corona Light	Grupo Modelo	14.540	0.015	3	1,049
Heineken	Heineken USA	14.293	0.026	3	1,065
Heineken Premium Light Lager	Heineken USA	14.621	0.007	3	982
Keystone Light	Molson Coors	5.977	0.007	3	1,102
Michelob Light	Anheuser-Busch	10.553	0.004	3	899
Michelob Ultra	Anheuser-Busch	10.587	0.014	3	1,012
Miller Genuine Draft	SABMiller	9.042	0.012	3	1,506
Miller High Life	SABMiller	6.916	0.012	3	1,436
Miller Lite	SABMiller	9.067	0.052	3	1,521
Modelo Especial	Grupo Modelo	13.894	0.006	3	878
Natural Ice	Anheuser-Busch	6.149	0.006	3	1,207
Natural Light	Anheuser-Busch	6.192	0.019	3	1,368
Tecate	FEMSA	11.323	0.008	3	895

Notes: An observation is a brand–size–city–month combination. The table makes use of data from year 2007. For every brand, average price is the average price across all observations. Share of revenue is a brand's share of the overall revenue of the beer industry in 2007. Observations is the number of observations of each brand.

where δ_{it}^{j} is the mean utility of product (i, j) in market t, μ_{cit}^{j} is the consumer specific deviation in the valuation of product (i, j) from its market-specific mean which depends on product characteristics and consumer demographics, and $\zeta_{cg(i,j)t}(\varrho) + (1 - \varrho)\epsilon_{cit}^{j}$ is the remaining consumer taste heterogeneity that is distributed extreme value. As in Miller and Weinberg (2017), this structure of the unobserved consumer heterogeneity follows the assumptions of a two-level nested logit model and allows substitution patterns within a group (or a nest) to differ from substitution patterns across groups. The size of this difference is determined by the nesting parameter $0 \leq \varrho < 1$. Since that $\zeta_{cg(i,j)t}(\varrho)$ is common to all products in group g, and ϵ_{cit}^{j} is i.i.d. extreme value, larger values of ϱ correspond to stronger correlation in preferences for products within the same group.[7] As in Miller and Weinberg (2017), here we use two groups $g = 0, 1$, where group 0 includes only the outside option $(i, j) = (0, 0) \in \Omega_{0t}$ (e.g. buy no beer) and group 1 includes all the other products $(i, j) \in \Omega_t$, which we denote by Ω_{1t}.

The mean and consumer specific utilities δ_{it}^{j} and μ_{cit}^{j} are parameterized as follows:

$$\delta_{it}^{j} = X_i^j \beta + \gamma P_{it}^{j} + \sigma_i^j + \sigma_\tau + \xi_{it}^{j} \tag{2}$$

$$\mu_{cit}^{j} = [P_{it}^{j}, X_i^j] \Pi D_c \tag{3}$$

where X_i^j denotes observable product characteristics that in our analysis include calories and a constant term, P_{jt} is the price of product j in market t, σ_i^j is a product specific intercept, σ_τ is a time period τ specific intercept, and ξ_{jt} is the unobserved product quality. (γ, β) then denote the average valuation of price and various product characteristics, respectively. D_c denotes (demeaned) consumer income, and Π is a vector of parameters governing how (γ, β) vary across consumers according to their demeaned income. The outside option payoff is normalized to zero so that $\delta_{0t}^{0} + \mu_{c0t}^{0} = 0$.

Given these distributional assumptions, the market share of product (i, j) in market t is given by $S_{it}^{j}(\mathbf{P}_t)$, where $\mathbf{P}_t$ is the vector of prices of all products in market t, can be written as:[8]

$$S_{it}^{j}(\mathbf{P}_t) = \frac{1}{N_t} \sum_{c=1}^{N_t} \frac{\exp\left(\frac{\delta_{it}^{j}(P_{it}^{j}) + \mu_{cit}^{j}(P_{it}^{j})}{1-\rho}\right)}{\exp\left(\frac{I_{cgt}}{1-\rho}\right)} \frac{\exp\left(I_{cgt}\right)}{\exp\left(I_{ct}\right)} \tag{4}$$

[7] $\zeta_{cig(j)t}(\varrho)$ follows a distribution, which depends on ϱ, that makes $\zeta_{cg(i,j)t}(\varrho) + (1 - \varrho)\epsilon_{cit}^{j}$ extreme value

[8] Here, we assume the number of consumers is large enough so that we can "integrate out" $\zeta_{cg(i,j)t}(\varrho) + (1 - \varrho)\epsilon_{cit}^{j}$.

where N_t is the number of consumers in market t, $I_{cgt} = (1-\rho)\sum_{j\in\Omega_{gt}} \exp\left(\frac{\delta^j_{it}(P^j_{it})+\mu^j_{cit}(P^j_{it})}{1-\rho}\right)$ is the inclusive value for groups $g = 0, 1$ and $I_{ct} = \log(1 + \exp(I_{c1t}))$ is the inclusive value for the entire market. Here, we write $\delta^j_{it}(P^j_{it})$ and $\mu^j_{cit}(P^j_{it})$ to emphasize that the consumer-specific payoffs to each good depend are functions of each good's price (through equations 2 and 3).

3.2 Pricing Game

Consider market t (a city–time period combination in our analysis). Firm i has a portfolio $\mathbb{J}_{it}$ of products in market t. We assume that the product market game takes the form of a simultaneous-move Bertrand pricing game.

Firm i maximizes its profit given beliefs about the behavior of rivals, i.e.,

$$\max_{\{P^j_{it}\}_{j\in\mathbb{J}_{ft}}} N_t \sum_{j\in\mathbb{J}_{it}} P^J_{it} S^j_{it}(\mathbf{P}_t) - C(\mathbf{Y}_{it}), \tag{5}$$

where N_t is the market size of market t, $Y^j_{it} = N_t S^j_{it}(\mathbf{P}_t)$, and $C(\mathbf{Y}_{it})$ is the cost function of producing vector $\mathbf{Y}_{it}$ (of cardinality $|\mathbb{J}_{it}|$).

The equilibrium vector of prices, $\mathbf{P}_t$, solves the system of first-order conditions,

$$S^j_{it} + \sum_{k\in J_{it}} \frac{\partial S^k_{it}(\mathbf{P}_t)}{\partial P^j_{it}} \left(P_{kt} - \frac{\partial C(\mathbf{Y}_{it})}{\partial Y^j_{it}}\right) = 0, \quad \forall j \in \mathbb{J}_{it}, \forall i \tag{6}$$

3.3 Technologies and Cost Functions

Each product (i, j, t) is assumed to be produced at a single brewery $b(i, j, t)$, although multiple products can be produced in the same brewery b.[9] We now discuss a specification of technologies that allow for joint production and economies of scope, an important feature of multi-product production emphasized by Baumol et al. (1982), at the level of a brewery, b.

For this purpose, we start by characterizing a firm's production possibility frontier from its production possibility set $(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt}) \in \mathbb{P}_{ibt}$, where $\mathbf{Y}_{ibt}$ is a vector of outputs produced by the firm i at brewery b for market t, and $\mathbf{X}_{ibt}$ is a vector of inputs. A firm's production possibility frontier is the *maximal* vector of outputs that a firm can produce, $\mathbf{Y}_{ibt}$, given a vector of inputs $\mathbf{X}_{ibt}$. One useful way to characterize a firm's production possibility frontier is by characterizing their *output distance function* (Caves et al. 1982), which tells us the

[9]We allocate products to the same brewery b if the distance to the brewery variable in Miller and Weinberg (2017) is the same.

minimum amount a firm must scale down a given output vector $\mathbf{Y}_{ibt}$ to make sure that $(\frac{\mathbf{Y}_{ibt}}{\delta}, \mathbf{X}_{ibt}) \in \mathbb{P}_{ibt}$, where δ is the minimized scaling factor. More formally, the output distance function is:

$$D_{ibt}(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt}) \equiv \min_{\delta} \delta \quad s.t. \left(\frac{\mathbf{Y}_{ibt}}{\delta}, \mathbf{X}_{ibt}\right) \in \mathbb{P}_{ibt}. \tag{7}$$

The output distance function is a convenient way to represent a firm's technology—in particular, to characterize their production possibility frontier, one needs to look at the set of $(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt})$ such that $D_{ibt}(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt}) = 1$.

Cairncross et al. (2023) show that the following class of output distance functions can be derived from a class of firm technologies based on the allocation of inputs across private, rival production tasks, and public, non-rival tasks:[10]

$$D(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt}; \mathbf{A}_{ibt}) \equiv \frac{\left(\sum_{j\in\mathbb{B}_{ibt}} \left(\frac{Y^j_{ibt}}{A^j_{ibt}}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{F(\mathbf{X}_{ibt})} \tag{8}$$

where $F(\mathbf{X}_{ibt})$ is homogeneous of degree ϕ, and $\mathbb{B}_{ibt}$ denotes the set of products produced in brewery b owned by firm i for market t. Note that technologies are assumed to differ across firms according to the vector of productivity shocks $\mathbf{A}_{ibt} = \{A^j_{ibt}\}_{j\in\mathbb{B}_{ibt}}$.[11]

As we show in Appendix B for the case where $F(\mathbf{X}_{ibt})$ is Cobb-Douglas, α and ϕ together determine whether a technology involves the use of public, non-rival inputs, which can generate scope economies. In particular, we show that $\beta_p \equiv \phi - \alpha$ captures the intensity of public, non-rival tasks in production. As a result, this technology generates economies of scope whenever $\beta_p > 0$, i.e. non-rival tasks contribute to overall production, and nests the standard case of non-joint production whenever $\beta_p = 0$.

To go from this characterization of a firm's production possibility set to their cost function, we assume that all inputs $\mathbf{X}_{ibt}$ are obtained from perfectly competitive markets. Since cost-minimizing firms will always operate on their production possibility frontier, and therefore $D_i(\mathbf{Y}_{ibt}, \mathbf{X}_{ibt}) = 1$, we can write:

$$\mathbb{Y}_{ibt} \equiv \left(\sum_{j\in\mathbb{B}_{ibt}} \left(\frac{Y^j_{ibt}}{A^j_{ibt}}\right)^{\frac{1}{\alpha}}\right)^{\alpha} = F(\mathbf{X}_{ibt}) \tag{9}$$

Equation (9) provides a "psuedo" production function for the output aggregator $\mathbb{Y}_{ibt} \equiv$

[10]See also Section 15G of Baumol et al. (1982) for a representation of this class of technology purely in terms of cost functions.

[11]As shown in Cairncross et al. (2023), these product specific productivity shifters govern the opportunity cost of each good within each firm.

$\left(\sum_j \left(\frac{Y^j_{ibt}}{A^j_{ibt}}\right)^{\frac{1}{\alpha}}\right)^{\alpha}$. With this in mind, since $F(\mathbf{X}_{ibt})$ is homogeneous of degree ϕ, we know that the cost function for $\mathbb{Y}_{ibt}$ will be given by:

$$C(\mathbb{Y}_{ibt}, \mathbf{W}_{ibt}) = g(\mathbf{W}_{ibt})\left(\mathbb{Y}_{ibt}\right)^{\frac{1}{\phi}} \tag{10}$$

$g(\mathbf{W}_{ibt})$ is a function that is homogeneous of degree 1 in input prices.

Substituting in the definition of $\mathbb{Y}_{ibt}$,we have:

$$C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt}) = g(\mathbf{W}_{ibt})\left(\sum_{j\in\mathbb{B}_{ibt}} \left(\frac{Y^j_{ibt}}{A^j_{ibt}}\right)^{\frac{1}{\alpha}}\right)^{\frac{\alpha}{\phi}} \tag{11}$$

Since this cost function represents a technology that nests the case of joint production through public inputs whether $\phi - \alpha > 0$, this cost function generates economies of scope, in the sense of Baumol et al. (1982) whenever $\phi > \alpha$. To show this, let $\mathbf{Y}^j_{ibt}$ denote a vector of outputs where all elements are zero except for the j'th element.

Lemma 1 *Consider a vector* $\mathbf{Y}_{ibt} = (Y^1_{ibt}, \ldots, Y^J_{ibt})$ *with* $Y^j_{ibt} > 0$ *for some* j *and* $\sum_j \mathbf{Y}^j_{ibt} = \mathbf{Y}_{ibt}$. *Then,*

- $C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt}) < \sum_j C(\mathbf{Y}^j_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt})$ *if* $\phi > \alpha$.
- $C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt}) = \sum_j C(\mathbf{Y}^j_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt})$ *if* $\phi = \alpha$;
- $C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt}) > \sum_j C(\mathbf{Y}^j_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt})$ *if* $\phi < \alpha$.

Proof. See Appendix A □

Note that the cost function $C(\mathbf{Y}^j_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt})$ corresponds to the cost function for a single product firm producing Y^j_{ibt}.[12] Lemma 1 shows that when $\phi > \alpha$, $C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt})$ will be strictly lower than producing the vector $\mathbf{Y}_{ibt}$ through J separate production processes. As a result, economies of scope and production costs are lower when firms produce multiple products together. On the other hand, when $\phi = \alpha$, the firm essentially operates J separate production processes, and as a result, there are no cost savings to producing multiple goods together. Finally, the case $\phi < \alpha$ involves *diseconomies of scope*, where production costs rise when a firm produces many outputs.[13]

[12] It is always possible for a firm to produce "as if" they were a single product when a firm has output distance function (8), since a firm can always choose to produce only a single product j with some bundle of inputs $\mathbf{X}^j_{ibt}$, in which case $D\left((0,0,...,Y^j_{ibt},...0), \mathbf{X}^j_{ibt}; \mathbf{A}_{ibt}\right) = 1$ implies $Y^j_{ibt} = A^j_{ibt}F(\mathbf{X}^j_{ibt})$.

[13] We do not expect this case to arise empirically since, in these situations, a firm would choose to operate the J independent product lines, which would generate lower costs.

4 Estimation

Having specified demand, the pricing game, and each firm's cost function, we now turn to estimation to conduct counterfactual experiments. Our primary interest is conducting counterfactuals that quantify the importance of scope economies. As a result, core to our analysis will be the estimation of (α, ϕ), which govern the magnitude of scope economies.

Since we rely on Miller and Weinberg (2017)'s demand model, we simply take their estimates of (γ, β, Π) as given. Appendix C contains more information about the estimates of the demand parameters we use. Given those estimates, we can construct δ_{it}^{j} using the information on $S_{it}^{j}(\mathbf{P}_t)$ and the standard Berry (1994) inversion. Since we assume that firms engage in Bertrand-Nash pricing, we can then recover an estimate of marginal cost by firm-product-market using a standard marginal cost inversion (e.g. Berry et al., 1995, Nevo, 2000). In particular, after stacking the $|\Omega_t| - 1$ first order conditions in (6) for each market t as follows:

$$\mathbf{S}_t + \Delta_t \left(\mathbf{P}_t - \mathbf{MC}_t\right) = 0, \tag{12}$$

where $\Delta_t \equiv \mathbb{O}_t \circ \partial_t$, where ∂_t is a$(|\Omega_t| - 1) \times (|\Omega_t| - 1)$ matrix of with typical element (m, n) equal to $\frac{\partial S_{it}^{m}}{\partial P_{i't}^{n}}$, and $\mathbb{O}_t$, and is an $(|\Omega_t| - 1) \times (|\Omega_t| - 1)$ ownership matrix, where element (m, n) equals 1 is product (i, m) and (i', n) are produced by the same firm (so $i' = i$), zero otherwise. We can then solve for the vector of marginal costs for all active products using:

$$\mathbf{MC}_t = \mathbf{P}_t + \Delta_t^{-1}\mathbf{S}_t \tag{13}$$

Note that (13) allows us to recover an estimate of the *equilibrium value* of each product's marginal cost $MC_{it}^{j} \equiv \frac{\partial C(\mathbf{Y}_{it})}{\partial Y_{it}^{j}}$. This value would only be sufficient for counterfactuals where outputs and prices change if there was no joint production and constant returns to scale.[14] If we wish to conduct such counterfactuals when firms face cost functions (11), we will need to estimate the components of their full marginal cost function, given by:[15]

$$MC_{ibt}^{j} = \frac{1}{\phi} g(\mathbf{W}_{ibt}) \left(\sum_{j \in \in \mathbb{B}_{ibt}} \left(\frac{Y_{ibt}^{j}}{A_{ibt}^{j}} \right)^{\frac{1}{\alpha}} \right)^{\frac{\alpha}{\phi} - 1} \left(\frac{Y_{ibt}^{j}}{A_{ibt}^{j}} \right)^{\frac{1}{\alpha}} \frac{1}{Y_{ibt}^{j}} \tag{14}$$

For this purpose, it will be useful be rewrite the marginal cost function as follows:

---

[14] In particular, imposing those conditions on our cost function by setting $\alpha = \phi = 1$ yields $C(\mathbf{Y}_{ibt}, \mathbf{A}_{ibt}, \mathbf{W}_{ibt}) = g(\mathbf{W}_{ibt}) \sum_j \frac{Y_{ibt}^{j}}{A_{ibt}^{j}}$, which then leads to $MC_{it}^{j} = \frac{g(\mathbf{W}_{ibt})}{A_{it}^{j}}$, which is sufficient to recover the full cost function.

[15] We show in Appendix A, that the marginal costs are *lower* for every product when there are scope economies and $\alpha < \phi$.

$$MC_{ibt}^{j} = \underbrace{\frac{g(\mathbf{W}_{ibt})}{\phi A_{ibt}^{\frac{1}{\phi}}}}_{\equiv K_{ibt}} \left(\sum_{j \in \mathbb{B}_{ibt}} \left(\frac{Y_{ibt}^{j}}{\widehat{A}_{ibt}^{j}} \right)^{\frac{1}{\alpha}} \right)^{\frac{\alpha}{\phi}-1} \left(\frac{Y_{ibt}^{j}}{\widehat{A}_{ibt}^{j}} \right)^{\frac{1}{\alpha}} \frac{1}{Y_{ibt}^{j}} \tag{15}$$

where we use the fact that we can substitute $A_{ibt}^{j} = A_{ibt}\widehat{A}_{int}^{j}$ where $\ln A_{ibt} \equiv \frac{1}{|\mathbb{B}_{ibt}|}\sum_{j\in\mathbb{B}_{ibt}} \ln A_{ibt}^{j}$ and $\ln \widehat{A}_{ibt}^{j} \equiv \ln A_{int}^{j} - \ln A_{int}$. Since $K_{ibt} \equiv \frac{g(\mathbf{W}_{ibt})}{\phi A_{ibt}^{\frac{1}{\phi}}}$ shifts the marginal costs of all products equally, we can fully recover the marginal cost function for the purpose of running counterfactuals by estimating $(\alpha, \phi, \{K_{ibt}\}_{(i,b,t)}, \{\widehat{A}_{ibt}^{j}\}_{(j,i,b,t)})$. In the following subsections, we describe our step-by-step procedure for estimating each of the objects. However, before turning to this procedure, we briefly discuss the intuitive sources of variation available to identify (α, ϕ), which together govern the magnitude of scope economies.

4.1 Intuition for identification of scale and scope economies

To generate a straightforward estimating equation for (α, ϕ), first multiply (15) by Y_{ibt}^{j}, and divide this expression by its sum over all $j \in \mathbb{B}_{ibt}$, yielding:

$$\frac{MC_{ibt}^{j}Y_{ibt}^{j}}{\sum_{j\in\mathbb{B}_{ibt}} MC_{ibt}^{j}Y_{ibt}^{j}} = \frac{\left(\frac{Y_{ibt}^{j}}{\widehat{A}_{ibt}^{j}}\right)^{\frac{1}{\alpha}}}{\sum_{j\in\mathbb{B}_{ibt}}\left(\frac{Y_{ibt}^{j}}{\widehat{A}_{ibt}^{j}}\right)^{\frac{1}{\alpha}}} = \left(\frac{Y_{ibt}^{j}}{A_{ibt}^{j}F(\mathbf{X}_{ibt})}\right)^{\frac{1}{\alpha}} \tag{16}$$

where the second equality follows from (9) and $A_{ibt}^{j} = A_{ibt}\widehat{A}_{ibt}^{j}$. We can rearrange this expression to generate the following estimating equation

$$\ln Y_{ibt}^{j} = \alpha \ln \left(\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}} MC_{ibt}^{j}Y_{ibt}^{j}} \right) + \ln F(\mathbf{X}_{ibt}) + \ln(A_{ibt}^{j}) \tag{17}$$

Note that since $F(\mathbf{X}_{ibt})$ is homogenous of degree ϕ, this expression is identical to the estimating equation derived in Orr (2022) when $\alpha = \phi$, in which case we can write:

$$\begin{aligned} \ln Y_{ibt}^{j} &= \phi \ln \left(\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}} MC_{ibt}^{j}Y_{ibt}^{j}} \right) + \ln F(\mathbf{X}_{ibt}) + \ln(A_{ibt}^{j}) \\ &= \ln F \left(\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}} MC_{ibt}^{j}Y_{ibt}^{j}} \mathbf{X}_{ibt} \right) + \ln(A_{ibt}^{j}) \end{aligned} \tag{18}$$

where Orr (2022) shows that $\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}}MC_{ibt}^{j}Y_{ibt}^{j}}\mathbf{X}_{ibt}$ will exactly equal the quantity of inputs allocated to product line j under the assumption of non-joint production.

Equation (17) provides an estimating equation that generalizes equation (18) by allowing input shares to enter with a freely variable coefficient α. More importantly, if a researcher were to estimate $\alpha < \phi$, this provides evidence that allocating inputs across product lines in a purely non-joint way does not scale up output in the same proportion as allocating inputs through $F(.)$. This implies returns to scale in $F(.)$ must capture something *beyond* the allocation of non-joint inputs across product lines. As the derivation of this technology based on public and private tasks makes clear (see Appendix B), this is because $F(.)$ also captures returns to *public* tasks, which are not captured by $\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}}MC_{ibt}^{j}Y_{ibt}^{j}}\mathbf{X}_{ibt}$. In particular, according to this model, α is equal to the overall intensity of private, rival tasks, β_r, while $\phi - \alpha$ is equal to the overall intensity of public tasks, β_p in production.

With this in mind, equation (17) suggests two separate sources of variation for identifying ϕ and α. Since ϕ captures overall returns to scale in all inputs, regardless of whether they are allocated to rival or non-rival tasks, exogenous variation in $\mathbf{X}_{ibt}$ can be used to identify the production function $F(.)$, which will allow a researcher to recover overall returns to scale. Since $\mathbf{X}_{ibt}$ are chosen by the firm with knowledge of $\{A_{ibt}^{j}\}_{j\in\mathbb{B}_{ibt}}$, which makes estimation by OLS infeasible, a researcher will generally need to find some way to instrument for $\mathbf{X}_{ibt}$. Ideal instruments would include anything that shifts up the overall scale of the firm, $\mathbf{X}_{ibt}$, holding overall productivity $\{A_{ibt}^{j}\}_{j\in\mathbb{B}_{ibt}}$ fixed.

To then identify α, a researcher needs to leverage exogenous variation in the marginal cost times output shares $\frac{MC_{ibt}^{j}Y_{i}^{jbt}}{\sum_{j\in\mathbb{B}_{ibt}}MC_{ibt}^{j}Y_{ibt}^{j}}$. Note, however, that Y_{ibt}^{j} and MC_{ibt}^{j} both directly depend on A_{ibt}^{j}, which again suggests the need for instruments. To find appropriate instruments, note that Y_{ibt}^{j} is determined by the interaction of both supply side and *demand-side* forces through a firm's pricing first-order conditions (6). This suggests relying on product-specific demand shifters as a way to shift $\frac{MC_{ibt}^{j}Y_{ibt}^{j}}{\sum_{j\in\mathbb{B}_{ibt}}MC_{ibt}^{j}Y_{ibt}^{j}}$ within a firm, and thereby identify α.

While (17) suggests a straightforward instrumental variables estimation strategy to identify (ϕ, α), this particular strategy is unavailable when researchers only rely on demand-side data, as $\mathbf{X}_{ibt}$ will be unobserved. In the following subsection, we propose a two-step estimator that leverages similar sources of variation for identifying (ϕ, α), while not requiring that $\mathbf{X}_{ibt}$ be known.

4.2 A two-step procedure

Step 1.

For simplicity of notation, rewrite equation (17) in the following way:

$$\ln Y_{ibt}^{j} = \alpha \ln \left(\rho_{ibt}^{j}\right) + \ln F(\mathbf{X}_{ibt}) + \ln(A_{ibt}^{j}) \tag{19}$$

where:

$$\rho_{ibt}^{j} \equiv \frac{MC_{ibt}^{j} Y_{i}^{jbt}}{\sum_{j \in \mathbb{B}_{ibt}} MC_{ibt}^{j} Y_{ibt}^{j}} \tag{20}$$

Notice that $F(\mathbf{X}_{ibt})$ is constant within a firm-brewery-market combination. Therefore, we can use product-level variation within these production sets in order to estimate α without requiring information about $F(.)$ or $\mathbf{X}_{ibt}$. We can rewrite Equation 19 in the following way:

$$\ln \widehat{Y}_{ibt}^{j} = \alpha \ln \left(\widehat{\rho}_{ibt}^{j}\right) + \ln(\widehat{A}_{ibt}^{j}) \tag{21}$$

where $\ln \widehat{Var}_{ibt}^{j} = \ln Var_{ibt}^{j} - \frac{1}{|\mathbb{B}_{ibt}|} \sum_{j \in \mathbb{B}_{ibt}} \ln Var_{ibt}^{j}$

Having recovered the marginal costs using Equation 13, we can construct $\ln\left(\widehat{\rho}_{ibt}^{j}\right)$. Notice that now we have the data to estimate α from equation (21) using OLS. However, as discussed earlier, the estimate would likely be biased, since $\widehat{\rho}_{ibt}^{j}$ would generally be correlated with $\widehat{A}_{ibt}^{j}$, the (demeaned) product-specific productivity shock. Notice that the bias may be either positive or negative, depending on the correlation between the equilibrium output of product j and all the characteristics of the production process captured in A_{ibt}^{j}. For example, Jaumandreu and Yin (2016), Forlani et al. (2016), Orr (2022) show that (quantity) productivity may be negatively correlated with product quality. This is because quality tends to be costly to produce (Verhoogen 2008, Kugler and Verhoogen 2012), and therefore firms might use smaller amounts of certain private inputs or less sophisticated processes to make the same amount of lower-quality goods, which will then be captured by a larger productivity term. Lower quality, at the same time, would then drive down the demand for the product. In this case, the sign of the correlation between A_{ibt}^{j} and the equilibrium Y_{ibt}^{j} would depend on the market conditions. Overall, *ex-ante*, it is not clear whether higher product-specific productivity would be associated with larger or smaller output.

Given the endogeneity of $\widehat{A}_{ibt}^{j}$, we use instrumental variables to estimate α. Specifically, we use the demand shocks that affect the equilibrium output of product j but are not correlated with the productivity term. As discussed earlier, given the estimated discrete choice demand model, we can recover the mean utility of product (i, j) in market t, $\widehat{\delta}_{it}^{j}$, and

Table 2: Scale and Scope Parameter Estimates

	(1) OLS	(2) IV	(3) OLS	(4) IV
$\widehat{\alpha}$	0.977***	1.020***		
	(0.007)	(0.016)		
$\frac{1}{\widehat{\phi}}$			0.792***	0.900***
			(0.002)	(0.016)
First Stage F-stat	-	548.37	-	275.76
$\widehat{\phi}$	-	-	1.263	1.111
	-	-	(.029)	(.022)
$\widehat{\phi} - \widehat{\alpha}$	-	-	-	.091
	-	-	-	(.022)
N	88468	88468	10920	10920

Notes: Standard errors, clustered by city, in parentheses. P-value $< .01$: $***$, P-value$< .05$: $**$, P-value$< .1$: $*$

obtain an estimate of the unobserved product appeal $\widehat{\xi}^j_{it}$ of product (i,j) on market t (see Equation 2). We will construct our instrument based on $\widehat{\xi}^j_{it}$. The first issue that we need to resolve is that $\widehat{\xi}^j_{it}$ may not be comparable across markets, as, given the way the demand model is estimated, $\widehat{\xi}^j_{it}$ should in fact be interpreted as the difference in the product's appeal for consumers on market t relative to that specific market's outside option. As a result, if the outside option differs by market, $\widehat{\xi}^j_{it}$ may not be properly comparable across markets. To deal with this, we define $\widehat{\xi}^{jR}_{it} \equiv \widehat{\xi}^j_{it} - \widehat{\xi}^r_{it}$ as the difference in the product appeal relative to a reference good r that is offered in all markets, which we take to be 6-packs of bug light.

Secondly, note that the estimate of $\widehat{\xi}^{jR}_{it}$ would reflect both the unobserved product quality as well as the market (e.g., city$\times$time) specific taste shocks. To construct our instrument, we remove the quality component from $\widehat{\xi}^{jR}_{it}$ since, as discussed above, it may be correlated with productivity. Specifically, we regress $\widehat{\xi}^{jR}_{it}$ on a series of product$\times$time fixed effects and obtain the new residual, which takes out product-specific quality as well as any time-varying dimensions to perceived quality (e.g. changes in packaging, national advertising campaigns, etc.). The remaining variation $\widehat{\xi}^{jRDM}_{it}$ captures market variation in the perceived appeal of different products due to local taste shocks, and therefore it should be uncorrelated with supply conditions. In the end, we use $\widehat{\xi}^{jRDM}_{it}$ as our instrument, leveraging local taste shocks relative to the baseline reference product r by market (e.g., city$\times$time).

OLS and IV results can be found in columns (1) and (2) of Table 2 below. Notice that the instrumental variable approach leads to a larger estimate of α. That implies that the OLS estimate was negatively biased, which is consistent with the productivity term A^j_{ibt} being negatively correlated with product quality and eventually with the output.

After estimating α, equation (21) also allows us to obtain the productivity shocks $\{\widehat{A}^j_{ibt}\}_{(j,i,b,t)}$, demeaned at the firm-brewery-market level. We will use these estimates as inputs into the second step of our procedure to evaluate the remaining parameters of the marginal cost function.

Step 2.

We can use the obtained estimate of α and $\{\widehat{A}^j_{ibt}\}_{(j,i,b,t)}$ and plug them into equation (15). We can then multiply both sides of the equation by Y^j_{ibt} and sum over all $j \in \mathbb{B}_{ibt}$ to get the following expression:

$$\sum_{j\in\mathbb{B}_{ibt}} MC^j_{ibt}Y^j_{ibt} = K_{ibt}\left(\sum_{j\in\mathbb{B}_{ibt}}\left(\frac{Y^j_{ibt}}{\widehat{A}^j_{ibt}}\right)^{\frac{1}{\widehat{\alpha}}}\right)^{\frac{\widehat{\alpha}}{\phi}} \tag{22}$$

where $K_{ibt} \equiv \frac{g(\mathbf{W}_{ibt})}{\phi A_{ibt}^{\frac{1}{\phi}}}$

Taking logs, we can rewrite the equation in the following way:

$$\ln\left(\sum_{j\in\mathbb{B}_{ibt}} MC^j_{ibt}Y^j_{ibt}\right) = \frac{1}{\phi}\ln\widehat{\mathbb{Y}}_{ibt} + \ln K_{ibt} \tag{23}$$

where $\widehat{\mathbb{Y}}_{ibt} \equiv \left(\sum_j\left(\frac{Y^j_{ibt}}{\widehat{A}^j_{ibt}}\right)^{\frac{1}{\widehat{\alpha}}}\right)^{\widehat{\alpha}}$ is the aggregator.

Again, while we can now estimate equation (23) using OLS, the estimate will be biased since the output aggregator is likely to be correlated with input prices $\mathbf{W}_{ibt}$, as well as the average productivity within a production set (e.g., firm-brewery-market) A_{ibt}. For this purpose, we rely on production set-level *averages* of relative local taste shifters $\widehat{\xi}^{RDM}_{it} = \frac{1}{|\mathbb{B}_{ibt}|}\sum_{j\in\mathbb{B}_{ibt}}\widehat{\xi}^{jRDM}_{it}$ as our instrument.

Column (3) of Table 2 reports the OLS estimate of $\frac{1}{\phi}$ based on Equation 23 while column (4) reports the IV estimate. Both columns use the IV estimate of α from column (2) of the same table to construct the aggregator $\widehat{\mathbb{Y}}_{ibt}$. Each specification also includes time, firm, and city fixed effects. Notice that the OLS estimate of $\frac{1}{\phi}$ is negatively biased since the unit cost $g(\mathbf{W}_{ibt})$ tends to be negatively correlated with the output.

At the bottom part of the table, we present point estimates and standard errors for $\widehat{\phi}$ and $\widehat{\phi} - \widehat{\alpha}$. As in Grieco et al. (2018), we also find evidence of scale economies, although ours (1.11) are slightly smaller than their preferred average returns to scale estimates (1.17 -

1.20). To account for uncertainty in the first step of the estimation, these standard errors are constructed using block bootstrap, where we sample cities with replacement and conduct 100 replications of the two-step procedure for each bootstrap sample. Standard errors are based on the sample standard deviation of the relevant statistic. Notice that using our estimates and bootstrapped standard errors, we can construct a t-statistic for whether $\widehat{\phi}$ is strictly greater than $\widehat{\alpha}$ and reject the null of no economies of scope in the beer industry. Finally, equation (23) also allows us to obtain the estimates of K_{ibt}.

5 The Impact of Scope Economies on Market Outcomes

How do scope economies impact market outcomes in the US beer industry? We compare the observed equilibrium with a counterfactual equilibrium in which we shut down economies of scope (i.e., no joint production occurs, but the production technology otherwise stays the same). In this counterfactual scenario, the marginal cost of producing good j by firm i at brewery b for market t becomes

$$MC_{ibt}^{\text{non-joint},j} = \frac{1}{\phi} g(\mathbf{W}_{ibt}) \left(\frac{Y_{ibt}^j}{A_{ibt}^j} \right)^{\frac{1}{\phi}} \frac{1}{Y_{ibt}^j} = K_{ibt} \left(\frac{Y_{ibt}^j}{\widehat{A}_{ibt}^j} \right)^{\frac{1}{\phi}} \frac{1}{Y_{ibt}^j}. \tag{24}$$

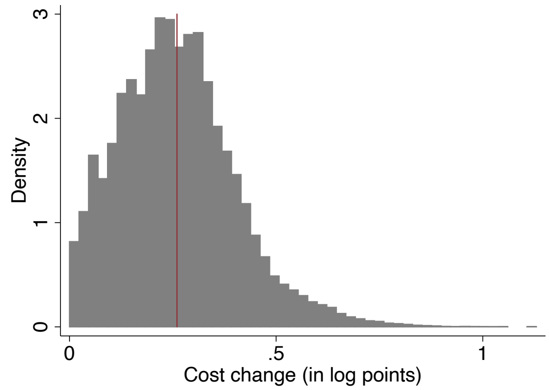
(see equation 14 for the marginal cost of product j with joint production). Lemma 2 in the Appendix shows that the marginal cost of production of a good is lower with joint production when $\phi > \alpha$, which is what our estimates suggest for the US beer industry.

We quantify the impact of joint production on marginal costs in two steps. We first hold quantities produced fixed, and compute the counterfactual marginal costs using equation (24). We present the results in Figure 1, which shows cost changes (in log points) for each product–market combination in our data when shutting down scope economies. The figure shows that the marginal cost of production of a product in our sample increases by about 25 percent on average (at the observed quantities).

How do these cost increases impact pricing incentives? On the one hand, the increase in the marginal cost of a product decreases the firm's marginal incentives to sell it, which creates an incentive to increase the price of the good to lower the quantity sold. On the other hand, the existence of increasing returns to scale (i.e., $\hat{\phi} > 1$ in the US beer industry) suggests the existence of a tradeoff: an increase in price decreases quantity, which further increases the marginal cost of production. This makes a price increase costly for the firm.

To quantify how these pricing incentives play out in equilibrium, Table 3 compares the

Figure 1: Cost Change when Shutting Down Scope Economies at Observed Quantities



Notes: The histogram displays the distribution of $\log(MC^{\text{no scope economies}}) - \log(MC^{\text{observed}})$, where MC^{observed} is the outcome in the observed equilibrium.

counterfactual equilibrium in which scope economies are shut down with the observed equilibrium. In the counterfactual equilibrium, firms fully adjust prices and production. To compute the counterfactual prices, we solve the system of first-order conditions of the pricing game in Equation 6, replacing the marginal cost of production of good j with $MC_{ibt}^{\text{non-joint},j}$ (see Equation 24).

Table 3 shows that shutting down scope economies causes prices to increase by 13.7 percent on average. The effects are heterogeneous across firms—the larger price increases are among the firms with the largest number of products (Anheuser-Busch, Molson Coors, and SABMiller), which have the most to lose when shutting down scope economies. Market shares on average decrease by 1.7 percent, which makes production even less efficient, as firms miss out on scale economies. Shutting down scope economies (including the interaction with scale economies), causes an increase in the marginal cost of production to 26.3 percent on average. As with prices, Anheuser-Busch, Molson Coors, and SABMiller are the firms with the largest effects on their market shares and marginal costs of production.

Our estimates of economies of scope suggest that these have a first-order effect on market outcomes. The results also show that scale and scope economies reinforce each other—when both are present, they interact, making it cheaper for firms to sell each additional unit. Our findings also show that scope economies and joint production have an economically significant impact on productive efficiency, providing an (at least partial) explanation for the existence of multiproduct firms in the US beer industry.

Table 3: The Impact of Scope Economies on Market Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Price change (in log points)		Cost change (in log points)		Share change (in log points)	
Overall	0.137		0.263		-0.017	
	(0.001)		(0.001)		(0.001)	
Anheuser-Busch		0.134		0.312		-0.030
		(0.000)		(0.001)		(0.000)
FEMSA		0.025		0.078		0.019
		(0.001)		(0.001)		(0.001)
Grupo Modelo		0.111		0.214		0.040
		(0.001)		(0.001)		(0.001)
Coors Molson		0.173		0.258		-0.015
		(0.001)		(0.001)		(0.001)
Pabst		0.120		0.118		-0.015
		(0.001)		(0.002)		(0.001)
SABMiller		0.158		0.234		-0.036
		(0.001)		(0.001)		(0.001)
Observations	86,392	86,392	86,392	86,392	86,392	86,392

Notes: Standard errors in parentheses. Each column displays regression coefficients of $\log(X^{\text{no scope economies}}) - \log(X^{\text{observed}})$ on a constant (columns 1, 3, and 6) or firm-level indicators, for $X \in \{\text{price}, \text{marginal cost}, \text{market share}\}$, and where X^{observed} is the outcome in the observed equilibrium.

6 Concluding Remarks

Why do multiproduct firms exist? Are economies of scope an empirically relevant explanation for the existence of multiproduct firms? We shed light on these questions by proposing a new method to estimate economies of scope and using the method to evaluate the impact of economies of scope on market outcomes in the US beer industry. Our method requires data commonly used for demand estimation (crucially, quantities produced and prices for each product–market combination) but does not require input data, making it easy to implement in other settings. We find that shutting down economies of scope (i.e., no joint production takes place, but the production technology otherwise stays the same) leads to price and marginal cost increases of 13 percent and 26 percent on average, respectively, which leads to an average decrease in a product's market share of 1.7 percent. Our findings explain the prevalence of multiproduct firms in the US beer industry.

References

Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran, "Scalable Expertise," 2020.

Ascher, Bernard, "Global beer: The road to monopoly," *American Antitrust Institute*, 2012, *1* (1), 1–36.

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

Baumol, J. William, John C. Panzar, and Robert D. Willig, *Contestable Markets and the Theory of Industry Structure*, New York: Harcourt College Pub, 1982.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott, "Multiple-product firms and product switching," *American economic review*, 2010, *100* (1), 70–97.

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

Berry, Steven T., "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25*, 242–262.

_ **and Philip A. Haile**, "Identification in differentiated products markets using market level data," *Econometrica*, 2014, *82* (5), 1749–1797.

Bronnenberg, Bart J, Michael W Kruger, and Carl F Mela, "Database paperâThe IRI marketing data set," *Marketing science*, 2008, *27* (4), 745–748.

Cairncross, John, Peter Morrow, Scott Orr, and Swapnika Rachapalli, "Multi-product Markups," *Working Paper*, 2023.

Caves, Douglas W., Laurits R. Christensen, and W. Erwin Diewert, "The economic theory of index numbers and the measurement of input, output, and productivity," *Econometrica*, 1982, pp. 1393–1414.

Dhyne, Emmanuel, Amil Petrin, Valerie Smeets, and Frederic Warzynski, "Theory for Extending Single-Product Production Function Estimation to Multi-Product Settings," *Working Paper*, 2022.

Ding, Xiang, "Industry linkages from joint production," *Work. Pap., Georgetown Univ., Washington, DC*, 2020.

Forlani, Emanuele, Ralf Martin, Giordiano Mion, and Mirabelle Muûls, "Unraveling firms: Demand, productivity and markups heterogeneity," *Working Paper*, 2016.

Goldberg, Pinelopi K, Amit K Khandelwal, Nina Pavcnik, and Petia Topalova, "Multiproduct firms and product turnover in the developing world: Evidence from India," *The Review of Economics and Statistics*, 2010, *92* (4), 1042–1049.

Grieco, Paul, Joris Pinkse, and Margaret Slade, "Brewed in North America: Mergers, marginal costs, and efficiency," *International Journal of Industrial Organization*, 2018, *59*, 24–65.

Hall, Robert E, "The specification of technology with several kinds of output," *Journal of*

Political Economy, 1973, *81* (4), 878–892.

Jaumandreu, Jordi and Heng Yin, "Cost and product advantages: A firm-level model for the chinese exports and industry growth," *Working Paper*, 2016.

Johnes, Geraint, "Costs and industrial structure in contemporary British higher education," *The Economic Journal*, 1997, *107* (442), 727–737.

Keithahn, Charles F, *The brewing industry*, Federal Trade Commission, Bureau of Economics, 1978.

Kerkvliet, Joe R, William Nebesky, Carol Horton Tremblay, and Victor J Tremblay, "Efficiency and technological change in the US brewing industry," *Journal of Productivity Analysis*, 1998, *10* (3), 271–288.

Kohli, Ulrich R, "Nonjointness and factor intensity in US production," *International Economic Review*, 1981, pp. 3–18.

Kugler, Maurice and Eric Verhoogen, "Prices, plant size, and product quality," *Review of Economic Studies*, 2012, *79* (1), 307–339.

Loecker, Jan De and Chad Syverson, "An industrial organization perspective on productivity," in "Handbook of Industrial Organization," Vol. 4, Elsevier, 2021, pp. 141–223.

_ **and Paul T Scott**, "Estimating market power Evidence from the US Brewing Industry," Technical Report, National Bureau of Economic Research 2016.

Maican, Florin and Matilda Orth, "Determinants of Economies of Scope in Retail," *Working Paper*, 2020.

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

_ **, Gloria Sheu, and Matthew C Weinberg**, "Oligopolistic price leadership and mergers: The united states beer industry," *American Economic Review*, 2021, *111* (10), 3123–3159.

Nevo, Aviv, "A Practitionerâs Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 2000, *9*, 513–548.

Orr, Scott, "Within-Firm Productivity Dispersion: Estimates and Implications," *Journal of Political Economy*, 2022, *130* (11), 2771–2828.

Panzar, John C and Robert D Willig, "Economies of scale and economies of scope in multi-output production," *Bell Laboratories economic discussion paper*, 1975, *33.*

_ **and** _ , "Economies of scope," *The American Economic Review*, 1981, *71* (2), 268–272.

Rosse, James N., "Estimating cost function parameters without using cost data: Illustrated methodology," *Econometrica*, 1970, pp. 256–275.

Teece, David J, "Economies of scope and the scope of the enterprise," *Journal of economic behavior & organization*, 1980, *1* (3), 223–247.

Verhoogen, Eric A, "Trade, quality upgrading, and wage inequality in the Mexican manufacturing sector," *Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

Zhang, Jingfang and Emir Malikov, "Off-balance sheet activities and scope economies in US banking," *Journal of Banking & Finance*, 2022, *141*, 106534.

A Properties of the Cost Function

Proof of Lemma 1

Proof. When $\phi > \alpha$, it follows that

$$C(\mathbf{Y}_i, \mathbf{A}_i, \mathbf{W}_i) < \sum_j C(\mathbf{Y}_i^j, \mathbf{A}_i, \mathbf{W}_i) \quad \Leftrightarrow$$

$$\left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha} < \left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\phi}}\right)^{\phi},$$

which holds true given that

$$\left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{x}}\right)^{x}$$

is strictly increasing in x. A similar argument can be used to prove the other claims.

□

Statement and Proof of Lemma 2

Lemma 2 *Consider a vector $(Y_i^1, \ldots, Y_i^J)$ with $Y_i^j > 0$. Then,*

- *the marginal cost of product j is lower under joint production (with a strict inequality if $Y_i^k > 0$ for some $k \neq j$) when $\phi > \alpha$;*
- *the marginal cost of product j is grester under joint production (with a strict inequality if $Y_i^k > 0$ for some $k \neq j$) when $\alpha > \phi$.*

Proof. The marginal cost of production with joint and non-joint production is given by

$$\begin{aligned} MC_i^{\text{joint},j} &= \frac{1}{\phi} g(\mathbf{W}_i) \left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\frac{\alpha}{\phi}-1} \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}} \frac{1}{Y_i^j}, \\ MC_i^{\text{non-joint},j} &= \frac{1}{\phi} g(\mathbf{W}_i) \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\phi}} \frac{1}{Y_i^j}, \end{aligned}$$

respectively, for $(Y_i^1, \ldots, Y_i^J) \geq 0$. When $\phi > \alpha$, it follows that

$$\begin{aligned} MC_i^{\text{non-joint},j} &\geq MC_i^{\text{joint},j} \quad &\Leftrightarrow \\ \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\phi}} &\geq \left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\frac{\alpha}{\phi}-1} \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}} \quad &\Leftrightarrow \\ 1 &\geq \left(\left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\frac{\phi-\alpha}{\phi}} \Big/ \left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\frac{\phi-\alpha}{\phi}}, \end{aligned}$$

where the inequality is strict if $Y_i^k > 0$ for some $k \neq j$. When $\phi < \alpha$, it is straightforward to establish that the reverse inequality holds. □

B A Model of Public and Private Tasks

In this Appendix, we derive the output distance function (8) for the special case of Cobb-Douglas production. For more general functional forms, see Cairncross et al. (2023).

For each input X, there are two tasks- a private task r, and a public task p. A firm allocates X_i^{rj} units of X to product line j doing the private task (e.g. construction), and X_i^p units of X to the public task (supervising), which affects all product lines at once. Output of product line j is determined by the following production function

$$Y_i^j = \frac{A_{it}^j}{C}\left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p}\right) \tag{25}$$

Where $C \equiv \frac{\prod_X \left(\beta_X^r\right)^{\beta_X^r}\left(\beta_x^P\right)^{\beta_X^p}}{\prod_X \left(\beta_X^r+\beta_X^p\right)^{\beta_X^r+\beta_X^p}}$. Given $\mathbf{X}_i$, the vector of aggregate inputs for the firm, we can characterize the firm's production possibility set by solving for their output distance function, given by:

$$\begin{aligned} D\left(\mathbf{Y}_i, \mathbf{X}_i, \mathbf{A}_i\right) \equiv & \min_{\delta, \mathbf{X}_i, \{\mathbf{X}_i^{rj}\}_j} \quad \delta \\ & \text{s.t.:} \quad \frac{Y^j}{\delta} \leq \frac{A_{it}^j}{C}\left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p}\right) \ \forall j \\ & \qquad X_i^p + \sum_j X_i^{rj} \leq X_i, \ \forall X \end{aligned} \tag{26}$$

This optimization problem has the following Lagrangian

$$L = \delta + \sum_j \lambda_i^j \left(\frac{A_i^j}{C} \left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p} \right) - \frac{Y_{it}^j}{\delta} \right) + \sum_X \mu_X \left(X_i - X_i^p + \sum_j X_i^{rj} \right)$$

Since the production functions are increasing in all inputs, all constraints will bind with equality, and therefore $\lambda_i^j > 0 \; \forall j$ and $\mu_X > 0 \; \forall X$, with:

$$\frac{Y^j}{\delta} = \frac{A_{it}^j}{C} \left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p} \right) \quad \forall j \tag{27}$$

and

$$X_i^p + \sum_j X_i^{rj} = X_i \quad \forall X \tag{28}$$

Taking the first order condition for X_i^{rj} yields

$$\lambda_i^j \beta_X^r \frac{\frac{A_i^j}{C} \left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p} \right)}{X_i^{rj}} = \lambda_i^j \beta_X^r \frac{\frac{Y_i^j}{\delta}}{X_i^{rj}} = \mu_X \tag{29}$$

The first order condition for X_{it}^p satisfies

$$\sum_j \lambda_i^j \beta_X^p \frac{\frac{A_i^j}{C} \left(\prod_X \left(X_i^{rj}\right)^{\beta_X^r} \left(X_i^p\right)^{\beta_X^p} \right)}{X_i^p} = \frac{\beta_X^p}{\delta X_i^p} \sum_j \lambda_i^j Y_i^j = \mu_X \tag{30}$$

Let $X_i^r \equiv \sum_j X_i^{rj}$. Rearrange and sum (29) for all j, yielding:

$$\mu_X X_i^r = \frac{\beta_X^r}{\delta} \sum_j \lambda_i^j Y_i^j \tag{31}$$

Rearrange (30) and divide by (31)

$$\frac{X_i^p}{X_i^r} = \frac{\beta_X^p}{\beta_X^r} \tag{32}$$

Since $X_i = X_i^r + X_i^p$, substitute (32) into this expression, yielding:

$$X_i^r + \frac{\beta_X^p}{\beta_X^r} X_i^r = X_i$$

or:

$$X_i^r = \frac{\beta_X^r}{\beta_X^r + \beta_X^p} X_i \tag{33}$$

and:

$$X_i^p = \frac{\beta_X^p}{\beta_X^r + \beta_X^p} X_i \tag{34}$$

Next, rearrange (29) and divide by (31), yielding:

$$X_i^{rj} = \frac{\lambda_i^j Y_{it}^j}{\sum_k \lambda_i^k Y_i^k} X_i^r \tag{35}$$

Substitute into (33), (34) and (35) into (27), which yields :

$$\frac{Y_i^j}{\delta} = \frac{A_i^j}{C} \left(\prod_X \left(\frac{\lambda_i^j Y_i^j}{\sum_k \lambda_i^k Y_i^k} \frac{\beta_X^r}{\beta_X^r + \beta_X^p} X_i \right)^{\beta_X^r} \left(\frac{\beta_X^p}{\beta_X^r + \beta_X^p} X_i \right)^{\beta_X^p} \right)$$

Define $\alpha \equiv \sum_X \beta_X^r$, $\beta_X = \beta_X^r + \beta_X^p$, and $\phi \equiv \sum_X \beta_X$. Rearranging and cancelling out terms in the above yields::

$$\left(\frac{Y_i^j}{\delta A_i^j} \right)^{\frac{1}{\alpha}} = \frac{\lambda_i^j Y_i^j}{\sum_k \lambda_i^k Y_i^k} \left(\prod_X (X_i)^{\beta_X} \right)^{\frac{1}{\alpha}}$$

Sum over all j :

$$\frac{1}{\delta^{\frac{1}{\alpha}}} \sum_j \left(\frac{Y_i^j}{A_i^j} \right)^{\frac{1}{\alpha}} = \left(\prod_X (X_i)^{\beta_X} \right)^{\frac{1}{\alpha}}$$

Or:

$$\delta = \frac{\left(\sum_j \left(\frac{Y_i^j}{A_i^j} \right)^{\frac{1}{\alpha}} \right)^{\alpha}}{\prod_X (X_i)^{\beta_X}}$$

This establishes that the firm's distance function is $D(\mathbf{Y}_i, \mathbf{X}_i, \mathbf{A}_i) = \frac{\left(\sum_j \left(\frac{Y_i^j}{A_i^j} \right)^{\frac{1}{\alpha}} \right)^{\alpha}}{\prod_X (X_i)^{\beta_X}}$.[16] Therefore a firm's production possibility frontier can be characterized by $(\mathbf{Y}_i, \mathbf{X}_i)$ satisfying:

$$\frac{\left(\sum_j \left(\frac{Y_i^j}{A_i^j} \right)^{\frac{1}{\alpha}} \right)^{\alpha}}{\prod_X (X_i)^{\beta_X}} = 1 \tag{36}$$

Note that from this derivation of the firm's distance function, we can see that α, which

[16]While we did not use the first order condition for δ, which implies $\delta^2 = \sum_j \lambda_i^j Y_i^j$, we would use this expression to solve for μ_X and λ_i^j

we defined as $\alpha \equiv \sum_X \beta_X^r$, should be interpreted as the share of *rival* or *private* inputs in production, while overall returns to scale, $\phi \equiv \sum_X \beta_X$ depends both private and public tasks, so $\alpha \leq \phi$.

C Estimates of Demand Parameters

We use the estimates of the demand parameters directly from Miller and Weinberg (2017). Specifically, we use the estimates from their baseline RCNL-1 model (see Table IV in Miller and Weinberg (2017)).

Variables	**Parameter**	**Estimate**
Price	γ	-0.0887
Nesting Parameter	ϱ	0.8299
Demographic interactions		
Income × Price	Π_1	0.0007
Income × Constant	Π_2	0.0143
Income × Calories	Π_3	0.0043