Identifying Scale and Scope Economies using Product Market Data*

Ekaterina Khmelnitskaya† Guillermo Marshall‡ Scott Orr§

June 18, 2025

Abstract

We propose an empirically tractable method to estimate economies of scale and scope. We start from a micro-founded model of production by a multi-product firm and generate an estimating equation for the parameters governing scale and scope economies, together with the distribution of within-firm productivity. A strength of the method is that the parameters can be estimated using product market data (i.e., quantities, prices, demand shifters) along with production facility information, rather than cost-side accounting data. We apply this approach to the U.S. beer industry and evaluate the impact of scale and scope economies on merger analysis.

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

*We thank Ying Fan (Editor) and three anonymous referees for comments and suggestions that have greatly improved the paper. We thank Nate Miller and Matt Weinberg for generously sharing data and code with us. We also wish to thank Mons Chan, Alon Eizenberg, Paul Grieco, Joseph Paat, Julia Yan, Mert Demirer, and Jonathan Williams for their helpful discussions as well as conference participants at the 2023 CEA, EARIE, and SEA conferences; the 2024 Triangle Micro and Cowles Models and Measurement conferences; and the 2025 ASSA Annual Meeting, as well as seminar participants at UBC (2023), the Competition Bureau of Canada (2024) and Western University (2025), for helpful comments. Financial support from the Social Sciences and Humanities Research Council (SSHRC) is gratefully acknowledged. All estimates and analyses in this paper based on Information Resources Inc. data are by the authors and not by Information Resources Inc. Any 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

A small number of large firms increasingly dominate industrial production (Grullon et al. 2019, Bighelli et al. 2023). While this has generated some concern among economists, increased concentration may be beneficial to final consumers in the presence of scale and scope economies. More precisely, scale economies—where average variable costs fall with output—and scope economies—where cost savings are generated when a firm produces more than one good—are standard economic mechanisms that can make consolidating production in a smaller number of entities efficient (Panzar & Willig 1975, Teece 1980, Panzar & Willig 1981).

Given these potential mechanisms favoring consolidation of production, understanding whether economies of scale and scope matter quantitatively is of first-order importance for antitrust practitioners. Investigating the competitive impact of a merger when these mechanisms matter requires assessing a set of (in)efficiencies as in the classic Williamson tradeoff (Williamson 1968). For example, the price effects of a merger may lead to a decrease in the scale of production of the merging firms, potentially increasing marginal costs of production due to a loss in scale economies. On the other hand, merged firms may choose to consolidate production of many varieties in a smaller number of plants, potentially creating scope economies that lower the costs of production.

In this paper, we propose a new method to estimate scale and scope economies suitable for applied work. We start by setting up a multiproduct cost function at the plant level. The technology allows for (but does not impose) economies of scale and scope at the plant level. Following Baumol et al. (1982), we show that this cost function can be derived from a production technology relying on non-rival inputs, e.g., managerial tasks or maintenance of machinery that can be used to produce many products simultaneously, which give rise to scope economies. The model allows for multiple plants per firm, multiple cities, and transportation costs that make it costly for a firm to ship goods from a plant to a city. In an application of our method, we ask whether accounting for scale and scope effects would substantially change the conclusions of standard merger evaluations.

Our methodology has two key strengths. First, we show that we can identify and estimate all the parameters of the multiproduct cost function using product market data (i.e., quantities, prices, demand shifters) combined with information about production facility locations. No data on inputs, costs, input allocations across products, or input prices are needed. This makes our method relatively easy to implement for industry studies where researchers have access to the data needed to estimate a demand system—with the caveat that the researcher needs to also observe production facility information. Second, multiproduct cost functions can suffer from a dimensionality problem, as the function must specify how an increase in the quantity of product j impacts the marginal cost of product k. This may create a practical problem for the econometrician—in particular, when the set of products is large—since estimation requires one instrument per parameter in the model. We tackle this issue using a parsimonious micro-founded model where two parameters govern scale and scope economies.

Using the marginal cost function for each product, we derive an equation that is linear in the key technology parameters governing scale and scope economies. Using this equation, we can identify the magnitude of scale and scope economies using two sources of exogenous variation. Scale economies are identified from marginal cost changes caused by exogenous variation in the output of a particular production line. Scope economies are then identified from marginal cost changes caused by exogenous variation in the output of *other* product lines, fixing the output of a particular production line.[1]

Estimating our model, however, requires us to resolve three econometric challenges. First, econometricians rarely observe marginal costs in their data. We propose overcoming this issue by estimating the marginal cost for each product using the demand-side approach, pioneered in Rosse (1970), and further developed by Berry et al. (1995), Nevo (2000), and Berry & Haile (2014). The two key elements of this approach are demand system estimates and a product market game specification. Using the system of first-order conditions for profit maximization evaluated at the observed prices, we can recover point estimates of the marginal costs at the city–product level. We use these marginal costs and the observed output levels together with our marginal cost model to estimate the production parameters of interest.

Second, if there are variables that affect costs that are unobservable to the econometrician but observable to the firm—e.g., city–product-specific cost shocks that capture the difficulty of producing different individual products and shipping them to various markets that are not directly observed by the researcher—then realized marginal costs will depend on a *vector* of unobserved cost shocks when there are scope economies. This makes standard estimation approaches—which usually rely on the existence of a single observation-level unobservable to rationalize each observed marginal cost value—not directly applicable. To estimate our model using standard tools, we show that our cost function generates a mapping from observables to a unique city–product-specific unobservable that rationalizes each realized marginal cost value.[2] Importantly, this mapping generates an estimating equation that is linear in the

[1]Relying on this particular source of variation leverages the insight from Baumol et al. (1982) that cost functions incorporating scope economies will generate product-level marginal cost functions that are declining in the output of other products.

[2]In practice, this allows us to *invert* the marginal cost function to recover estimates of the full set of cost shocks, similar to the demand inversion approach developed in Berry (1994) for recovering product-specific

parameters governing scale and scope, leading to straightforward estimation using standard techniques (e.g., generalized method of moments, GMM).

The third challenge we face is that this linear estimating equation leverages variation in the output of various product lines to identify scale and scope parameters. Since these outputs are functions of unobservable cost shocks in equilibrium, estimates of these parameters obtained through ordinary least squares (OLS) will be biased. To tackle this concern, we require instruments that exogenously shift output levels that are uncorrelated with product–city-specific cost shocks. We consider three classes of instruments for this purpose: the number of cities towards which each product is shipped, taste shocks recovered from our demand model (Hausman & Taylor 1983, Wooldridge 2010, MacKay & Miller 2025), as well as instruments constructed based on observable product line characteristics. As noted above, the parsimony of the model also restricts the number of instruments needed, which is a practical strength of our method.

We then apply our methodology to investigate the existence of scale and 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 incurring significant transportation costs. For example, CoorsMolson 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 (e.g., Budwesier, 6-pack). We focus on the years 2005 to 2008, and we model the demand system using a random coefficients logit model (Berry et al. 1995, Nevo 2000), which we estimate using pyBLP (Conlon & Gortmaker 2020).

Our estimates for the US beer industry suggest the existence of both scale and scope economies. We use these estimates to explore the role of these economies in merger analysis. We simulate the impact of the merger between CoorsMolson and SABMiller, which took place in 2008. One efficiency that was cited is that CoorsMolson and SABMiller would be able to leverage the breweries of the other firm, which would, on average, decrease the distance that the products of both firms would need to travel to reach consumers. Our model captures this efficiency gain but also captures an inefficiency that is a direct consequence of this: as production becomes more fragmented (i.e., less output per brewery), firms miss out

demand shocks in discrete choice demand systems.

on scale and/or scope economies. This inefficiency may lessen or overwhelm the shipping cost savings of producing closer to consumers.

A comparison of the equilibria with and without the merger shows that the marginal costs of MillerCoors decreased by 16.2 percent on average in a scenario in which MillerCoors frictionlessly reallocates products across breweries to optimize cost savings—a figure comparable to some of the findings in Miller & Weinberg (2017).[3] This combined effect is a result of multiple factors: cost decreases due to (i) scope and scale economies as some brewing facilities expand their output and range of products, (ii) transportation cost savings, and (iii) production being reallocated to more efficient breweries. These cost savings are partially offset by the enhanced market power of the merged company, resulting in a predicted average price decrease of 8.3 percent in MillerCoors products.[4] These average effects, however, mask significant heterogeneity across products and cities, since while some breweries grew in scale and scope after the merger, others shrank. We show that cost changes—and the corresponding price adjustments—depend on where the production of a given product for a particular city was reallocated, with costs rising in some cases.

We also find that the existence of scale and scope economies does not guarantee post-merger average cost savings. To this end, we provide analysis showing that the conclusions of a merger evaluation of the MillerCoors merger can change significantly depending on the assumptions made regarding the post-merger reallocation of MillerCoors products across breweries. Across scenarios, we find average marginal cost changes ranging between -16.2 percent (frictionless reoptimization) to +1.9 percent (MillerCoors products being reallocated to the brewery closest to the destination city). An implication for competition policy is that evidence in favor of scale and scope economies cannot be the basis of an efficiency defense by the merging parties unless the merging parties can prove that products can be reallocated across breweries in a way that will actually generate cost savings. More broadly, these findings highlight that scale and scope economies are relevant for analyzing mergers, especially where production occurs across multiple plants.

The rest of the paper is organized as follows. The rest of this section presents the literature review. The model is discussed in Section 2, and we present our identification and estimation strategy in Section 3. Section 4 describes our empirical application, which is the U.S. beer industry. Section 5 concludes.

[3]While Miller & Weinberg (2017) do not model scale and scope economies explicitly, they allow marginal costs change after the merger in a before-and-after comparison of marginal costs. We predict similar post-merger marginal cost changes using our cost model, which does not make use of post-merger data.

[4]Our estimates of *price* changes differ from those in Miller & Weinberg (2017) since i) we do not model post-merger price coordination and ii) Miller & Weinberg (2017) do not model scale and scope economies, which make marginal costs change across counterfactual environments.

1.1 Literature Review

We contribute to several strands of literature. First, we contribute to the literature on testing for the existence of non-joint production and scope economies. A part of the literature has relied on cost function estimation using firm-level cost data (Hall 1973, Kohli 1981, Baumol et al. 1982, Johnes 1997, Zhang & Malikov 2022) or estimation of multi-output technologies using transformation functions (Dhyne et al. 2022, Maican & 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.[5]

Other work has instead estimated scale or scope economies relying only on *product market* data, i.e., prices, quantities, and market shares. Along these lines, Berry et al. (1995) estimate scale economies in the automobile industry, and Fan (2013) estimates scale and scope economies in the US daily newspaper industry without explicitly distinguishing between the two mechanisms. More broadly, Berry & Haile (2014) provide conditions under which a non-joint marginal cost function with scale economies at the level of a product is identified when the econometrician has access to product market data only. We contribute to this strand of the literature by proposing a simple parameterization of a firm's cost function that allows for joint production of many products across multiple plants. Within a plant, this parametrization captures both within-product scale effects and across-product spillovers. We further show how to generate a simple estimating equation for the parameters governing scale and scope economies.[6]

Our paper is closely related to Argente et al. (2020), Ding (2022), and Raval (2024) who provide evidence for scale and scope economies for a wide variety of industries. 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 (2022) 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. The public input approach to conceptualizing a firm's scope is also pursued by Raval (2024), who uses data on "non-traceable" inputs from the FTC's line of business data to quantify the importance of non-

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

[6] While the cost function we rely on has been used in previous literature (Baumol et al. 1982, Johnes 1997), our estimation approach is more flexible than previous work since we explicitly allow for the existence of firm–product–city-specific shocks to productivity.

rival inputs in production. While non-rival input shares are unobserved in our data, our cost function estimates allow us to make comparisons with the shares observed in Raval (2024). Beyond these important similarities, we also provide a complementary “micro” study—focused on a single industry, U.S. brewing—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, Asker 2016, Miller & Weinberg 2017, Grieco et al. 2018, Miller et al. 2021, De Loecker & Scott 2024). 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. Fan & Yang (2025) also investigate the existence of scale and scope economies in the US beer industry. Their work complements ours as they study scale and scope economies based on fixed costs of entering a market (i.e., entry may have a firm and product-specific cost), and they estimate entry costs based on observed entry decisions. Our papers differ in the source of scale and scope economies (returns to variable and non-rival inputs versus fixed costs) and the variation used to estimate scope economies (prices versus entry decisions).

Finally, we contribute to the literature on synergies in mergers and acquisitions. While merger-related synergies represent one of the central concerns in antitrust, empirical studies in this area are surprisingly limited (Asker & Nocke 2021). We contribute to a handful of studies investigating merger-related efficiencies empirically (for example, Jeziorski 2014, Ashenfelter et al. 2015, Grieco et al. 2018, Elliott et al. 2025, Chen 2024; Demirer & Karaduman 2024, Eizenberg & Zvuluni 2024, Braguinsky et al. 2015). While many of these studies are relatively context-specific, our approach has the potential to be applied more broadly. Within this literature, our paper is closest to Grieco et al. (2018). Both studies enable the forecasting of post-merger changes in marginal costs using solely pre-merger data, which is particularly valuable from an antitrust perspective. An important distinction is that while Grieco et al. (2018) focuses on scale economies, our analysis also incorporates economies of scope. Moreover, while Grieco et al. (2018) relies on input data, our method solely requires product market data, which is already commonly used in merger evaluations.

2 A Model of Supply with Scale and Scope

We consider a setting where multiple differentiated varieties j are sold in a series of markets (c, t), where c denotes a city and t denotes time (a month).[7] Each variety j is produced

[7] To help the reader keep track of indexes, we always report product-related indexes as superscripts, while location/time/firm-specific indexes are reported as subscripts.

by a single firm f, which owns a series of different production locations—breweries, in our particular empirical application—$b \in \mathbb{B}_f$. We assume throughout that each product–market tuple (j, c, t) is produced in a single production location b, so we can write $b(j, c, t)$, i.e., for any market (c, t) firms do not source the same variety j from multiple separate production locations.[8] We denote the *set* of cities that product j is shipped to from production location b at time t as $c \in \mathbb{C}^j_{bt} \subset \mathbb{C}$ (where $\mathbb{C}$ is the set of all cities), the set of products j produced in production location b at time t as $j \in \mathbb{J}_{bt}$, the set of all products available in market (c, t) as $j \in \mathbb{J}_{ct}$, and the set of varieties sold by firm f in market (c, t) as $j \in \mathbb{J}_{fct}$.

Each firm uses the following CES cost function (Baumol et al. 1982, Johnes 1997), which we define over the set of products j produced at a particular production location b at time t ($\mathbb{J}_{bt}$):

$$C_{bt}(\mathbf{Y}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt}) = g(\mathbf{W}_{bt}) \left(\sum_{j \in \mathbb{J}_{bt}} \left(\frac{Y^j_{bt}}{A^j_{bt}} \right)^{\frac{1}{\alpha}} \right)^{\frac{\alpha}{\phi}}, \tag{1}$$

where $\mathbf{W}_{bt}$ is a vector of input prices faced in location b at time t, $\mathbf{Y}_{bt} = \{Y^j_{bt}\}_{j \in \mathbb{J}_{bt}}$, is the vector of outputs j produced, $\mathbf{A}_{bt} = \{A^j_{bt}\}_{j \in \mathbb{J}_{bt}}$ is a vector of product-specific productivity shifters, and $g(.)$ is a homogenous of degree one function.

This functional form specification offers several advantages. First, it allows for scale and scope economies in a tractable manner (see Lemma 1 below)—only requiring the estimation of two parameters that naturally lend to statistical tests for whether scale or scope economies operate in the data.[9] Particular values of these parameters nest the standard cases of independent and constant marginal costs ($\phi = \alpha = 1$), or models of non-joint production with scale economies at the product-level ($\phi = \alpha$).

Second, it is micro-founded, in that it can be derived from a simple production model where each input X (e.g., labor) can be allocated across two tasks: a private (or rival) task r (e.g., assembly line work), and a public task p (e.g., supervising). More formally, suppose the production function for product-line j in plant b at time t is given by:

$$Y^j_{bt} = \frac{A^j_{bt}}{C} \left(\prod_X \left(X^{rj}_{bt} \right)^{\beta^r_X} \left(X^p_{bt} \right)^{\beta^p_X} \right), \tag{2}$$

where X^{rj}_{bt} denotes the units of X allocated to product line j for the rival task, and X^p_{bt} the units of X allocated to the public task.[10] Note that X^p_{bt} affects all product lines at once, so

[8] Since $b(j, c, t)$ is completely determined by (j, c, t), we will often drop $b(.)$, unless necessary.

[9] Alternatively, a flexible polynomial function for outputs might be used, but this approach suffers from the curse of dimensionality as the number of product lines increases. Another alternative—aggregating outputs of all the other product lines into a single variable—is discussed in Section 3.

[10] $C \equiv (\prod_X (\beta^r_X)^{\beta^r_X} (\beta^P_X)^{\beta^p_X}) / (\prod_X (\beta^r_X + \beta^p_X)^{\beta^r_X + \beta^p_X})$ is a constant.

it does not have a j superscript. We show in Appendix A that this production technology gives rise to the cost function in equation (1), with $\phi \equiv \sum_X \beta_X^r + \beta_X^p$ (returns to scale in all tasks) and $\alpha \equiv \sum_X \beta_X^r$ (returns to scale in rival tasks).

While the Cobb-Douglas formulation of the production function is not strictly necessary to generate this cost function—see Cairncross et al. (2024) for more general nonparametric specifications of a firm's technology that lead to the same class of cost function—an important restriction embodied in this production model is that the production function is the same across all product-lines within a plant *except* through the plant-product productivity shifters A_{bt}^j. This assumption rules out differences in factor intensities across products, which may be important to understand some industries. For example, this model would not appropriately capture firms that produce both very capital-intensive electric cars and more labor-intensive gasoline-powered cars. This is important to bear in mind when choosing whether the proposed model is appropriate for a given research context. When we turn to our empirical application, we consider a setting where identical factor intensities are likely an appropriate assumption: the brewing of beer.

Scope economies are generated by this production model through the public task, p. Note that this model incorporates public inputs only if $\phi > \alpha$ (i.e., the share of inputs allocated to the public task is positive or $\sum_X \beta_X^p > 0$). As a result, whether $\phi > \alpha$ is closely related to whether the cost function in equation (1) captures economies of scope or not, as we show through Lemma 1.

Lemma 1 *For a given vector of outputs $\mathbf{Y}_{bt} > 0$, let $\mathbf{Y}_{bt}^j$ denote a corresponding vector of outputs where all elements are zero except for the j'th element which equals $Y_{bt}^j > 0$ from $\mathbf{Y}_{bt}$. Then,*

- $C(\mathbf{Y}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt}) < \sum_{j \in \mathbb{J}_{bt}} C(\mathbf{Y}_{bt}^j, \mathbf{A}_{bt}, \mathbf{W}_{bt})$ *if* $\phi > \alpha$.
- $C(\mathbf{Y}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt}) = \sum_{j \in \mathbb{J}_{bt}} C(\mathbf{Y}_{bt}^j, \mathbf{A}_{bt}, \mathbf{W}_{bt})$ *if* $\phi = \alpha$;
- $C(\mathbf{Y}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt}) > \sum_{j \in \mathbb{J}_{bt}} C(\mathbf{Y}_{bt}^j, \mathbf{A}_{bt}, \mathbf{W}_{bt})$ *if* $\phi < \alpha$.

Proof. See Appendix B □

Note that the cost function $C(\mathbf{Y}_{bt}^j, \mathbf{A}_{bt}, \mathbf{W}_{bt})$ corresponds to the cost function for a single-product firm producing Y_{bt}^j. Lemma 1 shows that when $\phi > \alpha$, $C(\mathbf{Y}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt})$ will be strictly smaller than the cost of producing the vector $\mathbf{Y}_{bt}$ through $|\mathbb{J}_{bt}|$ separate production processes. As a result, economies of scope mean that production costs are lower when firms produce multiple products together, potentially providing a rationale for firms to consolidate the manufacturing of many products in a single location. On the other hand, when $\phi = \alpha$, the firm essentially operates $|\mathbb{J}_{bt}|$ 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.

Since we consider a setting with particular production locations b shipping goods to individual cities c, it is also useful to augment this cost function to allow for transportation costs of shipping goods from plant b to city c. In particular, transportation costs are one of the key margins of adjustments that can be relevant for counterfactual analysis if firms reallocate their products across locations. In order to take the transportation costs into account, we distinguish between total quantities produced at a location b, Y^j_{bt}, and the total quantities sold at a particular city c, which we will denote by Q^j_{ct}. To allow for distribution costs associated with shipping goods from the production location b to final consumers in city c, we assume that shipping across locations is constrained by product–city-specific iceberg trade costs, where $\tau^j_{ct} \geq 1$ units of good j must be shipped to market (c, t) for each unit of the good to arrive.[11] More formally, if a firm wishes to sell outputs $\{Q^j_{ct}\}_c$, then total quantity produced at location b, Y^j_{bt} must satisfy:

$$Y^j_{bt} = \sum_{c \in \mathbb{C}^j_{bt}} Q^j_{ct} \tau^j_{ct}, \tag{3}$$

where we have made implicit use of the assumption that every product–city combination is sourced from one brewery at a given time (i.e., $b = b(j, c, t)$).

Combining equation (3) with equation (1) yields the following cost function defined over the outputs chosen by firm f for sale in each city c, $\mathbf{Q}_{bt}$:

$$C_{bt}(\mathbf{Q}_{bt}, \mathbf{A}_{bt}, \mathbf{W}_{bt}, \boldsymbol{\tau}_{bt}) = g(\mathbf{W}_{bt}) \left(\sum_{j \in \mathbb{J}_{bt}} \left(\sum_{c \in \mathbb{C}^j_{bt}} \frac{Q^j_{ct}}{\Omega^j_{ct}} \right)^{\frac{1}{\alpha}} \right)^{\frac{\alpha}{\phi}} \tag{4}$$

where $\Omega^j_{ct} \equiv A^j_{b(j,c,t)t} / \tau^j_{ct}$ is transportation-cost adjusted productivity associated with product j sold in city c at time t.

Note that the potential existence of scale economies in equation (4) generates cross-market interactions in costs. In particular, if there are scale economies at the product level and $\phi > 1$, then equation (4) implies that the cost of selling product j in city c will fall as the firm scales up the output of product j sold in the other cities supplied by the same plant. This also generates potential efficiency gains by consolidating production in a single

[11] This modeling assumption generates cost functions that are multiplicative in transportation costs—as is assumed in Miller & Weinberg (2017), for example—and is standard in quantitative spatial models; see Costinot & Rodríguez-Clare (2014) or Redding (2022).

location.

Given this cost function, the pricing game we assume for each firm is standard, i.e., static Nash-Bertrand pricing. Specifically, each firm f chooses the price of each product j in each city c to maximize its profits, given the prices of its rivals:

$$\max_{\{P_{ct}^j\}_{(j,c,t)\in\left(\bigcup_c \mathbb{J}_{fct}\right)}} \sum_{c\in\mathbb{C}} \sum_{j\in\mathbb{J}_{fct}} P_{ct}^j Q_{ct}^j(\mathbf{P}_{ct}) - \sum_{b\in\mathbb{B}_f} C_{bt}\left(\mathbf{Q}_{bt}(\mathbf{P}_t), \mathbf{A}_{bt}, \mathbf{W}_{bt}, \boldsymbol{\tau}_{bt}\right), \tag{5}$$

where $\mathbf{Q}_{bt}(\mathbf{P}_t)$ denotes the vector of demands for all products and cities served by production location b, which depends on the equilibrium price vector $\mathbf{P}_t$.

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

$$Q_{ct}^j + \sum_{j'\in\mathbb{J}_{fct}} \left(P_{ct}^{j'} - \underbrace{\frac{\partial C_{b(j',c,t)t}\left(\mathbf{Q}_{b(j',c,t)t}(\mathbf{P}_t), \mathbf{A}_{b(j',c,t)t}, \mathbf{W}_{b(j',c,t)t}, \boldsymbol{\tau}_{b(j',c,t)t}\right)}{\partial Q_{ct}^{j'}}}_{\equiv MC_{ct}^{j'}\left(\mathbf{Q}_{b(j',c,t)t}(\mathbf{P}_t), \mathbf{A}_{b(j',c,t)t}, \mathbf{W}_{b(j',c,t)t}, \boldsymbol{\tau}_{b(j',c,t)t}\right)} \right) \frac{\partial Q_{ct}^{j'}(\mathbf{P}_{ct})}{\partial P_{ct}^j} = 0, \tag{6}$$

$\forall j \in \mathbb{J}_{fct}, \forall (f, c)$, which completes the specification of our model.

3 Estimation of Scale and Scope Economies

In this section, we first briefly outline how researchers have previously tackled the problem of estimating the supply side for merger analysis using product market data. We assume throughout that the researcher has access to standard product market data, i.e., prices, quantities, and market shares at (c, j, t) level, and that that the mapping from city–product–time combinations to production locations, $b(j, c, t)$, is known by the researcher. The latter assumption is important, as this allows the researcher to aggregate quantities sold at the city–product–time level, Q_{ct}^j, to aggregate output at the level of a production location b, Y_{bt}^j, where scale and scope economies arise.[12] We then discuss the particular econometric challenges generated by our model with scale and scope, before turning to a complete description of our estimation algorithm.

Standard Approaches Standard product market based approaches to merger analysis (e.g., Hausman et al. 1994, Nevo 2000, among others) often assume that firms face constant

[12] Y_{bt}^j is not observed by the econometrician due to unobservable product–city specific transportation costs (e.g. rivers, shipping agreements for specific brands). We can recover a variable proportional to total output using (3) and a city–product–time-specific marginal cost inversion that accounts for these unobservables.

marginal costs of production following an ownership change. This requires abstracting away from both economies of scale and scope in variable costs, both of which can have important effects on post-merger outcomes.

The key benefit to assuming constant returns to scale is the empirical tractability, as supply-side parameters necessary to simulate merger counterfactuals are exactly identified when demand and each firm's pricing rule are known. For example, for the standard case of Bertrand-Nash price competition, the set of pricing first-order conditions that must be satisfied can be written in matrix form as follows:

$$\mathbf{Q}_{ct} + \Delta_{ct}\left(\mathbf{P}_{ct} - \mathbf{MC}_{ct}\right) = 0, \tag{7}$$

where $(\mathbf{Q}_{ct}, \mathbf{P}_{ct})$ are vectors of quantities and prices charged in each city c at time t, $\mathbf{MC}_{ct}$ is a vector of marginal costs, and Δ_{ct} is the element-by-element product of two matrices: a matrix of cross-price derivatives for the demand system (∂_{ct}) with typical element (j, j') equal to $\frac{\partial Q^j_{ct}}{\partial P^{j'}_{ct}}$, and an ownership matrix $(\mathbb{O}_{ct})$ where element (j, j') equals 1 if product j and j' are produced by the same firm, zero otherwise.

Typically in this setting, the econometrician observes $(\mathbf{Q}_{ct}, \mathbf{P}_{ct})$, with which they can obtain Δ_{ct} using demand estimation techniques. Marginal costs $\mathbf{MC}_{ct}$ are known to the firms but not the econometrician. An estimate of the vector of marginal costs can be obtained by inverting the system of first-order conditions, yielding:

$$\mathbf{MC}_{ct} = \mathbf{P}_{ct} + \Delta^{-1}_{ct}\mathbf{Q}_{ct} \tag{8}$$

Equation (8) provides exactly one marginal cost value for each specific combination of city, time period, and product that rationalizes the firm's pricing decisions under a particular conduct assumption. Under the further restriction of constant marginal cost, knowing this number is sufficient for merger counterfactuals, as marginal costs are invariant to the scale of production, as well as the set of products produced.

However, with scale and scope economies, point identifying marginal costs through equation (8) is no longer sufficient for counterfactual analysis. Rather, marginal costs will be functions of product-line specific output (due to scale economies), as well as the scale of *other* production lines (due to scope economies). As a result, marginal cost *functions* need to be identified to conduct counterfactual merger analysis. While this problem has been studied in non-joint production settings with scale economies at the product-level (e.g., Berry et al. 1995, Berry & Haile 2014), a novel problem generated by the introduction of scope economies is that marginal costs now become functions of the entire set of products produced within a production location b. We outline the specific econometric challenges generated by this type of specification and how we solve them in the following two subsections.

Econometric Challenges and Identification When firms face non-constant marginal costs, equation (8) allows the researcher to instead recover an estimate of the *equilibrium value* of each product–city marginal cost $MC_{ct}^j \equiv \partial C_{b(j,c,t)t}(\mathbf{Q}_{b(j,c,t)t};.)/\partial Q_{ct}^j$. From here, researchers can use these values to estimate a marginal cost function. This can be done simultaneously with demand estimation (e.g., Berry et al. 1995 and Fan 2013) or after demand estimation (Miller & Weinberg 2017). We rely on the latter formulation, taking as given a series of estimated $\{MC_{ct}^j\}_{j,c,t}$ values obtained by inverting an estimated demand system via equation (8).[13]

Before turning to our exact estimation algorithm that will be used for this purpose, it is useful to first outline the type of variation that would help a researcher identify scale and scope economies. To operationalize economies of scope, we rely on the following implication of scope economies, adapted to our notation from Proposition 4B1 of Baumol et al. (1982).[14]

Lemma 2 *Suppose* $\frac{\partial^2 C_{bt}(\mathbf{Y}_{bt};.)}{\partial Y_{bt}^k \partial Y_{bt}^j} \leq 0$*, with this inequality strict for at least one* $k \neq j$*. Then cost function* $C_{bt}(.)$ *involves economies of scope, i.e.,* $C_{bt}(\mathbf{Y}_{bt};.) < \sum_{j \in \mathbb{J}_{bt}} C_{bt}(Y_{bt}^j;.)$

Lemma 2 tells us that one implication of scope economies is that marginal costs of product j will be declining in *other* products' output. Combining this with the idea that scale economies will generate marginal cost functions that are declining in *own* output, one might then look for scale and scope economies in the data by estimating the following cost function at the level of a production unit:

$$\ln MC_{bt}^j = \beta_1 \ln Y_{bt}^j + \beta_2 \ln Y_{bt}^{-j} + \beta_3 X_{bt}^j + \omega_{ct}^j \tag{9}$$

where $Y_{bt}^{-j} \equiv \sum_{k \neq j \in \mathbb{J}_{bt}} Y_{bt}^k$ is total output of *other* product lines in production location b at time t, and X_{bt}^j are various observable characteristics of product j that may affect its costs.

If we assume $\beta_2 = 0$, the cost function in equation (9) is equivalent to the specification of marginal cost functions in Berry et al. (1995). Scale economies are then identified by testing for whether $\beta_1 < 0$. The key threat to identification, however, is that output is a function of cost shocks that are known to the firm but not known to the econometrician, i.e., ω_{ct}^j in equation (9). This means that the identification of scale economies requires instruments that exogenously shift the scale of production of product-line j for reasons unrelated to unobserved

[13]It is straightforward to adapt our approach to simultaneous estimation of demand and supply parameters by simply augmenting the GMM demand system in demand estimation with the supply side moments discussed below.

[14]See Chapter 4 and its Appendix in Baumol et al. (1982) for proofs.

cost differences. Since Y_{bt}^{j} is determined by the interaction of both supply-side and *demand-side* forces, one possibility is relying on *demand shifters* as instruments. Notably, ideal instruments would shift the demand for product j, holding the firm's cost structure fixed.[15]

Similarly, if we want to identify scope economies, i.e., $\beta_2 < 0$, we are looking for whether exogenous increases in the scale of other product lines $k \neq j$ translate into marginal cost decreases for j. Equation (9) suggests these instruments should shift the scale of other product lines while being uncorrelated with the unobserved cost shocks in ω_{ct}^{j}. Recalling again that outputs are determined by the interaction of demand and supply, this suggests that demand shifters for *other* products $k \neq j$ can be used as instruments to identify whether scope economies exist.

By thinking through how one might identify equation (9), we have learned that the identification of scale and scope economies requires demand-side instruments that shift the scale of production of each product independently of unobserved cost shifters. Given this, why not simply estimate equation (9) as a "reduced-form" specification of each firm's marginal cost function? While equation (9) is appealing due to the simple mapping between how (β_1, β_2) would be identified and our intuitive understanding of what scale and scope economies mean, it is an undesirable specification of the cost function for conducting counterfactuals. In particular, this specification of the cost function suffers from internal consistency issues once we start to think more carefully about the interpretation of the structural error term, ω_{ct}^{j}.

Note that equation (9) asserts that scope economies operate entirely through Y_{bt}^{-j}, while also admitting the existence of product-line unobservable cost shocks ω_{bt}^{j}. If such cost shocks are present, the marginal cost of product j should *also* be a function of the marginal cost shocks of other product lines.[16] This interdependence is not easy to capture in equation (9) without re-interpreting what ω_{ct}^{j} actually means: is it a product-line specific shock, or a composite shock affecting all product lines? Resolving this internal consistency issue requires a model that explicitly accounts for unobservable cost shifters at the product level.

[15]The fact that demand shifters are natural instruments for identifying scale economies in cost function estimation is analogous to *cost shifters* being natural candidates in demand estimation.

[16]This is because negative product-line specific cost shocks to product line j' implicitly lead to adjustments in public inputs, which affects the cost of production of all product lines, even holding the scale of j fixed.

Estimation Method Given our specification of the cost function (4), the marginal cost function for every product–city–time combination is given by:

$$
\begin{aligned}
&MC_{ct}^{j}\left(\mathbf{Q}_{b(j,c,t)t}(\mathbf{P}_t), \mathbf{A}_{b(j,c,t)t}, \mathbf{W}_{b(j,c,t)t}, \boldsymbol{\tau}_{b(j,c,t)t}\right) \\
&= \left(\sum_{j' \in \mathbb{J}_{b(j,c,t)t}} \left(\sum_{c' \in \mathbb{C}^{j'}_{b(j,c,t)t}} \frac{Q^{j'}_{c't}}{\Omega^{j'}_{c't}} \right)^{\frac{1}{\alpha}} \right)^{\frac{\alpha}{\phi}-1} \left(\sum_{c' \in \mathbb{C}^{j}_{b(j,c,t)t}} \frac{Q^{j}_{c't}}{\Omega^{j}_{c't}} \right)^{\frac{1}{\alpha}-1} \frac{\frac{1}{\phi} g(\mathbf{W}_{b(j,c,t)t})}{\Omega^{j}_{ct}}.
\end{aligned} \tag{10}
$$

Equations (6) and (10) nest the standard case of Bertrand-Nash pricing with constant marginal costs when $\alpha = \phi = 1$ in which case $MC^j_{ct} = g(\mathbf{W}_{b(j,c,t)t})/\Omega^j_{ct}$. These equations also nest the case of Bertrand-Nash pricing with a cost function with scale economies when $\alpha = \phi$ (Berry et al. 1995). Note, however, that even if there are nonconstant marginal costs as in equation (10), equations (7) and (8) continue to hold in equilibrium, allowing us to continue to rely on (8) to recover a series of realized (c, j, t)-level marginal cost values.[17]

Our microfounded cost function (10) shows that if we start with a model of production incorporating scope economies through non-rival inputs, we should indeed end up with a marginal cost function where all unobservable cost shocks, $\{\{\Omega^k_{ct}\}_{c \in \mathbb{C}^k_{bt}}\}_{k \in \mathbb{J}_{bt}}$, affect the marginal costs of j. Although it is not possible to use this expression directly for estimation—due to the presence of too many unobservables—we can make this particular marginal cost function suitable for GMM estimation by *inverting* out the (j, c, t)-specific marginal cost shocks in equation (10) to form a modified estimating equation.

To generate this inversion, we need to first define the following two cost shares:[18]

$$
S^j_{bt} \equiv \frac{\sum_{c' \in \mathbb{C}^j_{bt}} MC^j_{c't} Q^j_{c't}}{\sum_{j' \in \mathbb{J}_{bt}} \sum_{c' \in \mathbb{C}^{j'}_{bt}} MC^k_{c't} Q^k_{c't}} = \frac{\left(\sum_{c' \in \mathbb{C}^j_{bt}} \frac{Q^j_{c't}}{\Omega^j_{c't}} \right)^{\frac{1}{\alpha}}}{\sum_{j' \in \mathbb{J}_{bt}} \left(\sum_{c' \in \mathbb{C}^{j'}_{bt}} \frac{Q^{j'}_{c't}}{\Omega^{j'}_{c't}} \right)^{\frac{1}{\alpha}}} \tag{11}
$$

and

$$
S^{c|j}_{bt} \equiv \frac{MC^j_{ct} Q^j_{ct}}{\sum_{c' \in \mathbb{C}^j_{bt}} MC^j_{c't} Q^j_{c't}} = \frac{\frac{Q^j_{ct}}{\Omega^j_{ct}}}{\sum_{c' \in \mathbb{C}^j_{bt}} \frac{Q^j_{c't}}{\Omega^j_{c't}}}. \tag{12}
$$

The first cost share is the share of input costs of production line j in plant b relative to the

[17]While this model generates across market interactions through the marginal cost function, these do not show up in Δ_{ct} since $\frac{\partial Q^{j'}_{ct}(\mathbf{P}_{ct})}{\partial P^j_{c't}} = 0$ for $c' \neq c$.

[18]The second equality of each expression follows by substituting (10) into the defined cost share.

input costs of all production lines of plant b.[19] The second cost share is simply the share of product j being produced for city c.[20]

Manipulating the definitions of both cost shares, we note that $\sum_{j' \in \mathbb{J}_{bt}} \left(\sum_{c' \in \mathbb{C}^k_{bt}} \frac{Q^{j'}_{c't}}{\Omega^{j'}_{c't}} \right)^{\frac{1}{\alpha}} = \left(\sum_{c' \in \mathbb{C}^j_{bt}} \frac{Q^{j}_{c't}}{\Omega^{j}_{c't}} \right)^{\frac{1}{\alpha}} / S^j_{bt}$ and $\sum_{c' \in \mathbb{C}^j_{bt}} \frac{Q^{j}_{c't}}{\Omega^{j}_{c't}} = \frac{(Q^j_{ct}/\Omega^j_{ct})}{S^{c|j}_{bt}}$. Substituting these expressions into equation (10) and then taking logs, yields, after some minor manipulations:

$$\ln MC^j_{ct} = \frac{\phi - \alpha}{\phi} \ln S^j_{b(j,c,t)t} + \frac{1-\phi}{\phi} \ln \left(\frac{Q^j_{ct}}{S^{c|j}_{b(j,c,t)t}} \right) + \underbrace{\ln \left(\frac{1}{\phi} g(\mathbf{W}_{b(j,c,t)t}) \right) - \frac{1}{\phi} \ln \Omega^j_{ct}}_{\equiv \omega^j_{ct}}, \quad (13)$$

where we make use of the mapping $b = b(j,c,t)$ that identifies the brewery that supplies product j to city c at time t. After conditioning on output and the two cost shares in equations (11) and (12), equation (13) only has a single composite unobservable, $\omega^j_{ct} \equiv \ln \left(\frac{1}{\phi} g(\mathbf{W}_{b(j,c,t)t}) \right) - \frac{1}{\phi} \ln \Omega^j_{ct}$, that rationalizes each marginal cost at the (j,c,t)-level (i.e., product, city, time). This particular formulation of the marginal cost function naturally lends itself to GMM-based estimation using instruments, with ω^j_{ct} functioning as the structural residual.

Equation (13) is linear in two composite parameters; $\frac{1-\phi}{\phi}$, which governs scale economies, and $\frac{\phi-\alpha}{\phi}$, which governs the magnitude of scope economies. We now discuss the key sources of variation that we can use to identify these parameters, which turns out to be very similar to the sources of variation one would use to identify equation (9).

First, consider the identification of scale economies in the absence of scope economies ($\frac{\phi-\alpha}{\phi} = 0$). If each product was shipped to a single city, $\ln S^{c|j}_{b(j,c,t)t} = 0$, we would be in the standard Berry et al. (1995) world where variation in total output identifies the magnitude of scale economies. More precisely, if we see marginal costs fall as output (exogenously) rises, then we will infer the existence of scale economies, since a negative value of $\frac{1-\phi}{\phi}$ implies $\phi > 1$. However, since prices (and therefore outputs) are chosen with knowledge of ω^j_{ct}, one requires instruments that exogenously shift output for reasons unrelated to productivity. As we discussed earlier, natural candidates for this are product-specific demand shifters for product j.

Once we allow firms to ship a particular product to many cities, this opens up further sources of variation to identify scale economies. In particular, now the overall scale of

[19] We show in Appendix H in the Online Appendix that this share, S^j_{bt}, is the share of rival inputs allocated to product j; a novel extension to the identification result in Orr (2022) to a setting with joint production.

[20] To see this, substitute $\Omega^j_{ct} \equiv \frac{A^j_{bt}}{\tau^j_{ct}}$ into this expression, which yields: $\frac{Q^j_{ct}\tau^j_{ct}}{\sum_{c' \in \mathbb{C}^j_{bt}} Q^j_{c't}\tau^j_{c't}}$.

production *across all cities* will govern the degree to which a firm is able to take advantage of scale economies. The addition of the $\ln S^{c|j}_{b(j,c,t)t}$ term in equation (13) implicitly accounts for this, since $\ln Q^j_{ct} - \ln S^{c|j}_{b(j,c,t)t} = \ln\left(\sum_{c' \in \mathbb{C}^j_{b(j,c,t)t}} \frac{Q^j_{c't}}{\Omega^j_{c't}}\right) - \ln(\Omega^j_{ct})$. From this, we can see that instruments that shift the values $\{Q^j_{c't}\}_{c' \in \mathbb{C}^j_{b(j,c,t)t}}$—while holding fixed all $\{\Omega^j_{c't}\}_{c' \in \mathbb{C}^j_{b(j,c,t)t}}$—will allow us to identify whether there are scale economies. This suggests an alternative source of variation for identifying scale economies: the number of cities towards which product j is shipped, $|\mathbb{C}^j_{b(j,c,t)t}|$.[21] In particular, as long as product entry is primarily driven by demand-side factors, rather than product-specific cost differences, then $|\mathbb{C}^j_{b(j,c,t)t}|$ will act as an exogenous shifter of scale.[22] In our empirical application below, we consider product-specific demand shifters, as well as the number of cities a particular product is shipped to, as potential sources of exogenous variation for scale.

Next, consider the identification of $\frac{\phi-\alpha}{\phi}$. According to equation (13), $\frac{\phi-\alpha}{\phi}$ is revealed by exogenous variation in $S^j_{b(j,c,t)t}$ *conditional* on Q^j_{ct} and $S^{c|j}_{b(j,c,t)t}$, which we already used to identify returns to scale. If we wish to shift $S^j_{b(j,c,t)t}$ conditional on Q^j_{ct} and $S^{c|j}_{b(j,c,t)t}$, we need to rely on variation in the scale of *other* product lines. More precisely, if conditional on Q^j_{ct} and $S^{c|j}_{b(j,c,t)t}$, the output of some product $k \neq j$ produced in the same brewery rises, this will tend to decrease $S^j_{b(j,c,t)t}$.[23] As a result, if we observe that exogenous increases in the output of other product lines tend to decrease the marginal cost of product line j, this maps to decreases in $S^j_{b(j,c,t)t}$ decreasing marginal cost, i.e., $\frac{\phi-\alpha}{\phi} > 0$. In general, instruments that would generate this type of exogenous variation are demand shifters for other products $k \neq j$, similar to the instruments we would use to identify β_2 in our reduced form cost function (9). By shifting the scale of other products, conditional on Q^j_{ct} and $S^{c|j}_{b(j,c,t)t}$, these instruments essentially test for the "cost complementarities" from Baumol et al. (1982) discussed earlier, i.e., $\frac{\partial^2 C}{\partial Y^k \partial Y^j} < 0$.

It is worth noting that $\frac{\phi-\alpha}{\phi}$—the coefficient on $S^j_{b(j,c,t)t}$ which reveals whether there are scope economies—can be interpreted as the *cost share of public tasks* according to the production microfoundation underlying our cost function. More precisely, we show in Appendix A that $\frac{\phi-\alpha}{\phi} = \sum_X S^p_{bX} C_{bX}$, where S^p_{bX} is the share of input X allocated to the public task, and C_{bX} is the share of total cost allocated to input X. This is a useful property as this parameter is directly comparable to the common input shares recently observed by Raval

[21]For any set $\mathbb{X}$, we use the operator $|\mathbb{X}|$ to denote the number of elements in $\mathbb{X}$.

[22]Alternatively, cost shocks can be realized after product entry (Eizenberg 2014, Wollmann 2018, Fan & Yang 2020), as is commonly assumed in a demand estimation context to allow "BLP" style instruments to be valid.

[23]Recall, S^j_{bt} is the share of input costs of production line j in plant b relative to the input costs of all production lines of plant b.

(2024) in the FTC's line of business data.[24] This will allow us to compare our estimates of the magnitude of scope economies to some observable benchmarks as a sanity check.

When estimating equation (13), researchers may wish to augment this specification with product or market-level observables to proxy for unobserved marginal costs differences captured by $\omega_{ct}^j \equiv \ln\left(\frac{1}{\phi}g(\mathbf{W}_{b(j,c,t)t})\right) - \frac{1}{\phi}\left(\ln(A_{b(j,c,t)t}^j) - \ln\tau_{ct}^j\right)$. For example, one might assume that transportation costs depend on observables such as distance from production location b to market c or the cost of fuel, i.e., $\ln\tau_{ct}^j = \lambda_1 \mathrm{Z}_{b(j,c,t)ct} + \ln\widetilde{\tau}_{ct}^j$, or that unobservable productivity (cost) differences are related to observable product and brewery characteristics, i.e., $\ln\left(A_{b(j,c,t)t}^j\right) = \lambda_2 Z_j + \lambda_3 Z_{b(j,c,t)} + \ln\widetilde{A}_{b(j,c,t)t}^j$.[25] In practice, this leads to the following final estimating equation:

$$\begin{aligned}
\ln MC_{ct}^j \quad = \quad & \frac{\phi-\alpha}{\phi}\ln S_{b(j,c,t)t}^j + \frac{1-\phi}{\phi}\left(\ln Q_{ct}^j - \ln S_{b(j,c,t)t}^{c|j}\right) + \frac{\lambda_1}{\phi}\mathrm{Z}_{b(j,c,t)ct} - \frac{\lambda_2}{\phi}Z_j - \frac{\lambda_3}{\phi}Z_{b(j,c,t)} \\
& + \ln\left(\frac{1}{\phi}g(\mathbf{W}_{b(j,c,t)t})\right) + \frac{1}{\phi}\ln\widetilde{\tau}_{ct}^j - \frac{1}{\phi}\ln\widetilde{A}_{b(j,c,t)t}^j \qquad (14)
\end{aligned}$$

In our empirical application below, we parametrize the brewery and product observables with fixed effects, and rely on a parametrization of transportation costs from Miller & Weinberg (2017) where observable transportation costs are proportional to the distance from city c to brewery $b(j,c,t)$ times the price of fuel. This allows us to identify transportation cost effects by examining how much marginal costs rise—net of scale and scope effects—as products are shipped further or fuel prices rise.

4 Empirical Application: the MillerCoors Merger

Our empirical application focuses on the well-known joint venture—or merger, as it was effectively treated by the U.S. Department of Justice (DOJ)—between SABMiller and Molson Coors within the U.S. brewing industry, approved in June of 2008. This setting is well-suited for our analysis, as large brewing companies are typically multiproduct firms, producing a diverse range of beer brands. Moreover, despite these firms operating multiple brewing sites, the number of such locations is relatively limited. Ascher (2012) reports that the U.S. brewing industry, initially characterized by thousands of local breweries, saw a dramatic decline to approximately 22 traditional breweries by 2002. This trend suggests a significant

[24] The FTC's line of business data—discontinued in the 1980s—required firms to record the allocation of key inputs by line of business. Firms were also required to record the share of inputs that they could not allocate to particular product lines, which Raval (2024) takes to be the share of public inputs in production.

[25] Brewery characteristics might proxy both for average productivity at the brewery level, or (potentially) unobserved input costs at different breweries, $\ln(g(\mathbf{W}_{bt}))$; for notation sake we load these differences on productivity effects, but in general these are not separately identified unless input prices are observed.

increase in returns to scale and scope at 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).[26] The technological advancements include improvements in the bottling/canning and packaging technology, the automation of brewing processes, enabling large-scale operations with minimal labor, and innovations in the fermentation process (see Keithahn (1978), p. 34-39 for more detail).[27]

The merger between Miller and Coors was approved despite an increase in concentration in an already concentrated industry. Prior to the merger, the top five brands accounted for about 80 percent of sales by 2001, with the merging parties being the second and third largest in the industry (Miller & Weinberg 2017). The merger received approval primarily due to anticipated synergies in shipping costs, as Coors was going to be able to use Miller's more geographically diverse set of brewing facilities, thereby gaining closer access to various geographical markets (Ashenfelter et al. 2015). Our methodology enables an in-depth analysis of the merger's potential impact on prices and other market outcomes, considering not only changes in concentration and shipping costs but also how economies of scale and scope might be affected due to the reallocation of products across breweries.

4.1 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. We restrict attention to the months between January 2005 to May 2008, right before the MillerCoors merger was completed. We focus on this period to abstract away from the price effects of the MillerCoors merger (Miller & Weinberg 2017, Miller et al. 2021).

We define markets as city–month combinations. Products are defined as brand–size combinations (e.g., Bud Light 6-pack), and we focus on three sizes: 6-pack equivalent, 12-pack equivalent, and 24/30-pack equivalent. Following Miller & Weinberg (2017), we express quantities sold (and market shares) in terms of 12-pack equivalent units.[28] That is, one 6-pack or one 24-pack product is equivalent to 0.5 and 2 12-pack equivalent units, respectively.

[26]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 "water-processing equipment, sewage facilities, refrigeration equipment, management, laboratories, and custodial costs".

[27]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 similar labor savings compared to a smaller plant.

[28]We define the market size for each city by computing the maximum number of 12-pack equivalent units sold in that city across all months in our sample period, including all brands and sizes (even those that we exclude from our analysis).

Prices are measured as total revenue divided by 12-pack equivalent units.

We provide a list of the brands included in our sample as well as summary statistics in Table OA.1 in the Online Appendix. As the table shows, there are several sources of price variation. Larger sizes generally have a lower price per volume, but even within a size, the highest average price can be twice as much as the lowest average price, with imported beers generally being more expensive (per volume). Our sample accounts for all major brands, which combined account for 74 percent of all the beer sold (by volume) during our sample period.

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 & Weinberg (2017).

Lastly, we use data on the location of breweries. To estimate transportation costs, we construct a measure of distance based on the interaction of driving miles and diesel fuel prices (see Miller & Weinberg 2017 for details).

4.2 Empirical Model of Demand for Beer

The first step in our approach requires estimates of demand to derive both the own- and cross-price elasticities. These elasticities are then used within equation (8) to estimate marginal costs. To this end, we specify a random coefficients logit demand model (Berry et al. 1995, Nevo 2000).

The conditional indirect utility that consumer m receives from purchasing product $j \in \mathbb{J}_{ct}$ is given by:

$$u^j_{mct} = \delta^j_{ct} - \beta_m P^j_{ct} + \theta^{\text{constant}}_m + \theta^{\text{calories}}_m Cal^{ct}_j + \varepsilon^j_{mct}, \tag{15}$$

where δ^j_{ct} is the mean utility of product j in market (c,t), β_m is consumer m's coefficient on price (P^j_{ct}), $\theta^{\text{constant}}_m$ is a coefficient that is constant across inside goods and specific to consumer m, $\theta^{\text{calories}}_m$ is consumer m's coefficient on calories per 12 ounces (Cal^{ct}_j), and ε^j_{mct} is an idiosyncratic taste shock that is distributed extreme value type 1. The conditional indirect utility that consumer m receives from choosing the outside option in market (c,t) is given by $u^0_{mct} = \varepsilon^0_{mct}$.

The consumer-specific parameters, β_m, $\theta^{\text{calories}}_m$, and $\theta^{\text{constant}}_m$, are parameterized as:

$$\begin{aligned} \beta_m = \exp\{\beta + \pi^{\text{income}}_1 \text{income}_m\},\ \theta^{\text{calories}}_m &= \pi^{\text{income}}_2 \text{income}_m, \\ \text{and}\ \ \theta^{\text{constant}}_m &= \pi^{\text{income}}_3 \text{income}_m, \end{aligned} \tag{16}$$

where income_m is the income of consumer m (standardized to have mean zero and standard deviation 1), drawn from the Public Use Microdata Sample (PUMS) of the American Community Survey. The mean utility, δ^j_{ct}, is parameterized as

$$\delta^j_{ct} = \sigma_t + \sigma_j + \xi^j_{ct},$$

where σ_t and σ_j are time and product fixed effects and ξ^j_{ct} is an unobserved demand shifter.[29] Note that price is not included in δ^j_{ct} because the random coefficient on price is non-linear.

Given these assumptions, the market share of product j in market (c,t),$s^j_{ct}(\mathbf{P}_{ct})$, is given by:

$$s^j_{ct}(\mathbf{P}_{ct}) = \frac{1}{N_{ct}} \sum_{m=1}^{N_{ct}} \frac{\exp\left(\delta^j_{ct} - \beta_m P^j_{ct} + \theta^{\text{constant}}_m + \theta^{\text{calories}}_m Cal^{ct}_j\right)}{1 + \sum_{j' \in \mathbb{J}_{ct}} \exp\left(\delta^{j'}_{ct} - \beta_m P^{j'}_{ct} + \theta^{\text{constant}}_m + \theta^{\text{calories}}_m Cal^{ct}_{j'}\right)} \tag{17}$$

where $\mathbf{P}_{ct}$ is the vector of prices of all products in market (c,t) and N_{ct} is the number of consumers in market (c,t).

Estimation and Demand Estimates We estimate the demand system using pyBLP (Conlon & Gortmaker 2020). For estimation, we combine three sets of instruments that amount to eight instruments in total. The first instrument is the distance to the parent firm's nearest brewery interacted with the price of diesel, which is a marginal cost shifter.

The second set of instruments are differentiation instruments (Berry et al. 1995, Gandhi & Houde 2019) constructed within pyBLP based on two characteristics: size (in ounces) and calories per 12 ounces. For a given product j in market (c,t), and for a given product characteristic X, there are two differentiation instruments: the number of products k in market (c,t) of the same firm such that $|X^j_{c,t} - X^k_{c,t}| < \text{SD}_X$ and the number of products k in market (c,t) of rival firms such that $|X^j_{c,t} - X^k_{c,t}| < \text{SD}_X$, where SD_X is the standard deviation of all pairwise comparisons of characteristic X (across all markets).

Lastly, the third set of instruments is the city-level mean income as well as the city-level mean income interacted with exogenous product characteristics (calories, size). These instruments help identify the parameters governing customer heterogeneity and are valid under the assumption that the structural error is mean independent of average income and the exogenous product characteristics (Miller & Weinberg 2017).

We report the second-stage GMM estimates of the demand parameters $(\beta, \pi^{\text{income}}_1, \pi^{\text{income}}_2)$ in Table 1 after updating the weighting matrix and computing the approximated optimal instruments. Column (1) reports estimates that do not allow for consumer heterogeneity,

[29] Note that product-level fixed effects absorb the components of product-level characteristics (e.g., calories and size), that are treated similarly by all consumers.

Table 1: Demand Estimates

	(1)	(2)
Price (log)		
β	-0.9737	-1.0286
	(0.0424)	(0.1115)
π_1^{income}		-0.4970
		(0.2710)
Calories (linear)		
π_2^{income}		0.8671
		(0.2761)
Constant (linear)		
π_3^{income}		-2.2292
		(0.5954)
Observations	97,006	97,006
Mean own elasticity	-3.7291	-3.8164
Median own elasticity	-3.5142	-3.5886

Notes: The table shows the second-stage GMM estimates of the demand parameters after updating the weighting matrix and computing the approximated optimal instruments. Robust standard errors are in parentheses. All specifications include month and product (brand–size combination) fixed effects. The random coefficient (RC) on price is defined as $\beta_m = \exp\{\beta + \text{income}_m \pi_1^{\text{income}}\}$, whereas the RC on the constant and calories are given by $\theta_m^{\text{constant}} = \pi_3^{\text{income}}$ and $\theta_m^{\text{calories}} = \pi_2^{\text{income}}$, respectively.

whereas column (2) reports the estimates of the full model. The estimates reveal that consumers with a higher income are less price-sensitive, prefer beers with a greater calorie count, and are more willing to choose the outside option.

Table 1 also reports the mean and median own price elasticities implied by each specification. The mean and median own price elasticities in the full model are -3.82 and -3.59, respectively, which are within the range of price elasticity estimates for the US beer industry in the literature.[30] Figure OA.1 in the Online Appendix reports the full distribution of own-price elasticities, which roughly lie between -7 and -2. Table OA.2 in the Online Appendix reports a table of the average own and cross-price elasticities for the set of 12-pack products available in all city–month combinations. The table shows that Bud Light consumers disproportionately substitute towards similar beers (e.g., Coors Light, Miller Lite), and that consumers of imported beers (e.g., Corona Extra, Heineken) disproportionately substitute towards other imported beers.

[30] For example, Asker (2016) and Hidalgo (2023) both report an average own price elasticity of -3.4, Döpper et al. 2022 report an average own price elasticity of -4.2, whereas Miller & Weinberg (2017) report a median own price elasticity of -4.74 using their preferred specification.

4.3 Estimation of the Cost Function Parameters

Given our demand estimates, and assuming that firms engage in Bertrand-Nash pricing, we can recover marginal cost estimates at the firm–product–city–month level using the standard marginal cost inversion in equation (8). The model also requires information about the brewery b that ships product j to city c at time t, i.e., the mapping $b = b(j, c, t)$: we do not observe this information in our data. Our baseline assumption is to allocate products to the firm's nearest brewery, as in Miller & Weinberg (2017). This assumption is reasonable for at least two reasons. The first is that the pre-merger allocation we examine here reflects a long-term equilibrium in which brewery location and size decisions were endogenous, allowing companies to position breweries near key markets. Second, the DOJ stated they were willing to sign off on the merger between Miller and Coors precisely *because* allowing Coors to produce their beer brands closer to key markets was likely to decrease transportation costs, suggesting that transportation costs are economically relevant for the industry.[31] However, to verify that this assumption is not driving our results, we also re-estimate our model using a series of alternative allocations in Online Appendix E and find our results remain fairly similar.

Having allocated product (j, c, t) to breweries, we then estimate the following model:

$$\ln MC_{ct}^{j} = \frac{\phi - \alpha}{\phi} \ln S_{b(j,c,t)t}^{j} + \frac{1-\phi}{\phi} \ln \left(\frac{Q_{ct}^{j}}{S_{b(j,c,t)t}^{c|j}} \right) + \frac{\lambda}{\phi} \text{dist}_{b(j,c,t)c} \times \text{fuel}_t + \\ + \gamma_j + \gamma_{b(j,c,t)} + \gamma_c + \gamma_{y(t)} + \epsilon_{b(j,c,t)ct}^{j}, \tag{18}$$

where $S_{b(j,c,t)t}^{c|j}$ and $S_{b(j,c,t)t}^{j}$ are marginal cost times output shares—see equations (11) and (12)—$\text{dist}_{b(j,c,t)c} \times \text{fuel}_t$ parameterizes transportation costs following Miller & Weinberg (2017) by interacting distance from plant $b(j, c, t)$ to city c with the cost of diesel at time t (fuel_t), and γ_j, $\gamma_{b(j,c,t)}$, γ_c, and $\gamma_{y(t)}$ denote product, brewery, city, and year fixed effects, respectively. We make two adjustments to our sample when estimating this model. First, we exclude Heineken products since they are brewed in Europe in breweries that primarily serve the European market, which we cannot appropriately capture with our cost function and U.S. IRI data.[32] Second, we drop observations with $MC_{ct}^{j} < 0$, which accounts for less than 1% of our observations.[33]

[31] Heyer et al. (2009) note that "Much of the efficiencies involved freight cost savings that were based on the ability of the merged firm to redistribute production of the parties' two brand portfolios across the venture's multiple plants, which were geographically dispersed."

[32] In our counterfactuals, we treat these products as having constant marginal costs.

[33] Note, however, that to avoid dropping these products when measuring scale/scope, we replace any $MC_{ct}^{j} < 0$ with the lowest positive value for marginal cost recovered in the data (0.005) when constructing

Estimating the above model by OLS will likely fail to recover the true parameters of the cost function due to transmission bias: unobserved productivity shocks, transportation costs or input prices—which, according to equation (14) will show up in $\epsilon^{j}_{b(j,c,t)ct}$—will also affect $S^{j}_{b(j,c,t)t}$, Q^{j}_{ct}, and $S^{c|j}_{b(j,c,t)t}$, leading to biased estimates of scale and scope economies. Importantly, we largely expect OLS to overestimate both scale and scope economies. Since positive productivity shocks embodied in Ω^{j}_{ct} tend to increase each product's scale of production while also decreasing their marginal costs, estimation by OLS will mechanically attribute some marginal cost decreases from positive productivity shocks to scale effects, generating overly large estimates of returns to scale. Similarly, negative productivity shocks also will also tend to *increase* $S^{j}_{b(j,c,t)t}$, all else equal.[34] Since negative productivity shocks increase marginal costs and $S^{j}_{b(j,c,t)t}$ at the same time, this will tend to mechanically increase the coefficient on $S^{j}_{b(j,c,t)t}$, leading us to infer larger scope economies than actually exist.

We tackle these endogeneity problems using instrumental variables. We consider three classes of instrumental variables for this purpose, building on the variation we emphasized in Section 3 above. First, we consider simple (log) counts of the number of cities to which each product is shipped to identify scale, as well as counts of the number of *other* product–city combinations brewed and shipped from the same brewery to identify scope. One appealing feature of the above strategy is its simplicity and transparent link between the scale (cities) and scope (product–city combinations). However, the validity of the strategy requires disciplining product entry decisions in ways that may not be appropriate; notably, both the scale and scope measures should not be driven by unobserved cost shocks at either the brewery, product, or market level. To deal with this concern, we include product, city, and brewery fixed effects in many of our regressions to absorb time-invariant variation in the productivity of particular products and breweries, as well as the cost of accessing particular cities.[35] Nevertheless, across-time variation in productivity driving entry of particular products could still serve as a potential threat to identification when relying on this class of instruments.

To make sure this does not drive our final conclusions, we also consider instrumental variable strategies based on instruments that independently shift demand for particular products in each city–time combination. We consider two classes of demand shocks for this purpose. The first is based on unobserved demand shocks constructed from our demand residuals as discussed in Hausman & Taylor (1983) or MacKay & Miller (2025). To construct

the cost shares $S^{j}_{b(j,c,t)t}$ and $S^{c|j}_{b(j,c,t)t}$.

[34] More precisely, if we condition on Q^{j}_{ct} as well as the productivity shifters and scale of all other product lines, then decreasing Ω^{j}_{ct} increases $S^{j}_{b(j,c,t)t}$ according to (11).

[35] Brewery fixed effects also help deal with missing markets (e.g. exports, small cities). If the number of missing markets is the same for all products and constant over time, missing sales affects the denominator of $S^{c|j}_{b(j,c,t)t}$ for all products within a brewery identically; brewery fixed effects then difference out this bias.

this instrument, we first recover δ^j_{ct} for each (j, c, t), and regress this variable on a set of product-fixed effects to then recover the product–city specific unobservable demand shock ξ^j_{ct}. In practice, this instrument takes advantage of across-time variation in the degree to which consumers in different markets like different beers.[36] The second demand shock is the calorie-based differentiation IV we used to identify our demand system. For each of these product–city–time-specific demand shifters, Z^j_{ct}, we construct an instrument for scale using the average value of this demand shifter across all cities $\frac{1}{|\mathbb{C}^j_{bt}|}\sum_{c\in\mathbb{C}^j_{bt}} Z^j_{ct}$. To construct an instrument for scope, we calculate the average value of the demand shifter for all other products $\frac{1}{\sum_{j'\in\mathbb{J}_{bt}/j}|\mathbb{C}^{j'}_{bt}|}\sum_{j'\in(\mathbb{J}_{bt}/j)}\sum_{c\in\mathbb{C}^{j'}_{bt}} Z^{j'}_{ct}$.

Table 2 presents our baseline OLS and IV estimates of equation (14). Panel A presents the regression coefficients, whereas Panel B presents transformations of the regression coefficients to deliver an estimate of each parameter of the cost function. In the brackets below each point estimate, we report 95-percent confidence intervals generated by a bootstrapping procedure where we take into account demand uncertainty by first i) drawing from the estimated asymptotic distribution of our demand parameters, then ii) using these parameters to construct the marginal costs and demand shifters implied by this draw, and finally, iii) using a nonparametric block bootstrap procedure on this new dataset that resamples $(\widehat{MC}^j_{ct}, \widehat{\xi}^j_{ct})$ values with replacement at the brewery-time level, upon which we re-estimate equation (18).[37] The first column reports OLS estimates, while the final three columns report IV estimates using our three sets of instruments. All regressions include product, city, year, and brewery fixed effects. We further report the first stage regressions in Online Appendix D; importantly, we find that our instruments shift the endogenous variables in exactly the direction we would expect.

The cost function IV estimates provide evidence in favor of scale economies, with estimates of ϕ ranging between 1.293 and 1.555, and confidence intervals bounded away from 1. These estimates are comparable to those in Grieco et al. (2018), who estimate a translog production function for the beer sector allowing for heterogeneous returns to scale. They find mean returns to scale of 1.2, with returns to scale at the 75th percentile of 1.36. Note that we find that OLS delivers implausibly large estimates of returns to scale in column 1, consistent with the expected upward bias we discussed earlier.

We also find evidence of scope economies in all specifications, i.e., $\frac{\phi-\alpha}{\phi} > 0$ and we can reject the null hypothesis that $\frac{\phi-\alpha}{\phi} = 0$ at the 95% level. Recall that $\frac{\phi-\alpha}{\phi}$ can be interpreted

[36] Since one source of beer demand variation is seasonality—beer demand is highest in summer months, for reasons generally uncorrelated with costs—we do not include time fixed effects in these specifications. Instead, we only control for year fixed effects to absorb broad trends in overall beer demand over time.

[37] 95% confidence intervals report the 2.5th and 97.5th percentile values of the distribution of estimated parameters over the 500 bootstrap samples generated in this way.

Table 2: Cost Function Estimates: Baseline Specification

	(1) OLS	(2) IV	(3) IV	(4) IV
Panel A: Regression coefficients				
Scale coeff $\frac{1-\phi}{\phi}$	–0.575	–0.227	–0.325	–0.357
	[–0.691 , –0.507]	[–0.326 , –0.049]	[–0.355 , –0.239]	[–0.466 , –0.284]
Scope coeff $\frac{\phi-\alpha}{\phi}$	0.574	0.229	0.327	0.361
	[0.503 , 0.691]	[0.048 , 0.328]	[0.237 , 0.352]	[0.271 , 0.454]
Distance coeff $\frac{\lambda}{\phi}$	0.020	0.011	0.014	0.015
	[0.017 , 0.030]	[0.009 , 0.018]	[0.012 , 0.020]	[0.013 , 0.023]
Panel B: Implied cost function parameters				
ϕ	2.353	1.293	1.481	1.555
	[2.027 , 3.232]	[1.052 , 1.483]	[1.314 , 1.551]	[1.396 , 1.872]
α	1.002	0.997	0.997	0.993
	[0.995 , 1.007]	[0.992 , 1.005]	[0.999 , 1.009]	[0.999 , 1.039]
λ	0.047	0.015	0.020	0.023
	[0.037 , 0.085]	[0.011 , 0.024]	[0.017 , 0.030]	[0.020 , 0.042]
First Stage F-stat Scale	-	30.78	1626.35	101.76
First Stage F-stat Scope	-	40.54	2809.61	75.88
Observations	91,112	91,112	91,112	91,112

Notes: An observation is a product–city–month combination. Panel A presents the regression coefficients from equation (14). Panel B presents the implied cost function parameters (i.e., transformations of the coefficients in Panel A). Bootstrapped 95-percent confidence intervals, taking into account demand system uncertainty, in brackets, based on 500 bootstrap replications. Column (1) reports estimates using OLS, whereas Columns (2) and (3) report IV estimates using the following instruments:

- Column (2): $\ln|\mathbb{C}^j_{bt}|$, and $\ln\left(\sum_{j'\in(\mathbb{J}_{bt}/j)}|\mathbb{C}^{j'}_{bt}|\right)$
- Column (3): $\frac{1}{|\mathbb{C}^j_{bt}|}\sum_{c\in\mathbb{C}^j_{bt}}\xi^j_{ct}$, and $\frac{1}{\sum_{j'\in\mathbb{J}_{bt}/j}|\mathbb{C}^{j'}_{bt}|}\sum_{j'\in(\mathbb{J}_{bt}/j)}\sum_{c\in\mathbb{C}^{j'}_{bt}}\xi^{j'}_{ct}$
- Column (4): $\frac{1}{|\mathbb{C}^j_{bt}|}\sum_{c\in\mathbb{C}^j_{bt}}\text{DiffCal}^j_{ct}$, and $\frac{1}{\sum_{j'\in\mathbb{J}_{bt}/j}|\mathbb{C}^{j'}_{bt}|}\sum_{j'\in(\mathbb{J}_{bt}/j)}\sum_{c\in\mathbb{C}^{j'}_{bt}}\text{DiffCal}^{j'}_{ct}$

where $\text{DiffCal}^j_{ct}\equiv\sum_{k\in\mathbb{J}_{ct}\setminus\mathbb{J}_{f(j)ct}}\mathbb{I}\left(|\text{Calories}^j_{c,t}-\text{Calories}^k_{c,t}|<\text{SD}_{\text{Calories}}\right)$, with $f(j)$ denoting the firm producing product j, $\mathbb{I}(.)$ denoting the indicator function.

First-Stage F-stat Scale is the Sanderson-Windemeijer F-statistic for weak instruments for $\left(\ln Q^j_{ct}-\ln S^{c|j}_{b(j,c,t)t}\right)$, and First-Stage F-stat Scope is the corresponding F-statistic for $\ln S^j_{b(j,c,t)t}$. Heineken products (Heineken and Heineken Light, imported products in our sample) and observations where $MC^j_{it}<0$ (less than 1% of full sample) are excluded from these regressions. All regressions include product, city, brewery, and year fixed effects.

as the share of public, nonrival inputs in production, as we indicated at the end of Section 3. Our point estimates in the IV specification deliver public input shares between 20 to 36 percent. These estimates are slightly larger than the average common inputs shares reported in Raval (2023), who finds a mean expenses shares of 14 percent, although note that 10 percent of his sample report common expense shares of more than 36 percent, similar to the upper end of our estimates. On the other hand, OLS estimates appear to generate implausibly large public input shares—more than 50 percent—again consistent with the upward bias discussed earlier.

Similar to previous work (Miller & Weinberg 2017, Ashenfelter et al. 2015), our analysis shows a positive and significant relationship between distance and marginal costs: increasing the distance measure by one standard deviation (the standard deviation is 1.51) increases marginal costs by between 2 to 3.5 percent on average.

Overall, all three sets of instruments generate evidence for scale and scope economies for the beer sector. While there are differences in the magnitude across our instrument sets—notably, column (2) provides smaller scale and scope estimates, while column (4) provides the largest—these differences are not particularly precise, as we cannot reject the null of equal returns to scale (ϕ) or scope economies ($\frac{\phi-\alpha}{\phi}$) across columns (2) and (3), as well as (3) and (4), at the 95 percent level using bootstrapped confidence intervals.[38] In what follows, we rely on the estimates in column (2), as they are more conservative in terms of the magnitude of scale and scope economies.

Other Cost Function Objects. With additional assumptions on the structure of productivity terms introduced in equation (14), our regression specification allows us to estimate all the other brewery- and product-level components of the cost function necessary for counterfactuals. In particular, any counterfactual involving the reallocation of production must take a stance on how unobserved cost residuals at the product–city level, $\omega_{ct}^{j} \equiv \ln\left(\frac{1}{\phi}g(\mathbf{W}_{b(j,c,t)t})\right) - \frac{1}{\phi}\left(\ln(A_{b(j,c,t)t}^{j}) - \ln\tau_{ct}^{j}\right)$, vary as we reallocate products across production locations. To separate intrinsic product-specific costs from brewery-specific input prices and technology levels, we regress these residuals on product, year, and time-fixed effects and assume product-fixed effects are tied to the products, while brewery-fixed effects are tied to breweries for any given realization of the time fixed effects. We discuss further details in Online Appendix F.

[38] However, we can reject the null of equality for columns (2) and (4), given that these contain both our largest and smallest estimates.

4.4 Scope and Scale Economies in Merger Analysis

A potential efficiency gain of the MillerCoors merger was that Coors and Miller products would have to travel shorter distances to reach consumers, as each merging firm could leverage the network of breweries of the other merging firm (Ashenfelter et al. 2015). However, when scale and scope economies are present, there is a tradeoff between consolidating production and saving on shipping costs. That is, fragmenting production over a larger number of breweries may lessen a different type of efficiency caused by consolidating production: scope and scale economies.[39] The MillerCoors merger thus creates tension between these two forces: scope and scale economies versus shipping cost savings.

Merger Analysis How did the merger impact market outcomes? What is the role of scope and scale economies in explaining these results? To address these questions, we compare market outcomes in the equilibria in which the MillerCoors merger has and has not taken place.

We implement this comparison varying our choices across two dimensions. The first dimension is the specification of the cost function. We either use our estimates of the marginal cost function that allows for scale and scope economies (see equation 10 and the estimates in Table 2) or a simplified cost function that abstracts away from scale and scope economies, where marginal cost is given by

$$\ln MC^j_{ct} = \psi \text{dist}_{b(j,c,t)c} \times \text{fuel}_t + X^j_{b(j,c,t)ct}\gamma + \varepsilon^j_{ct}, \tag{19}$$

where $X^j_{b(j,c,t)ct}$ includes a set of fixed effects (product, brewery, city, and time) and the distance variable ($\text{dist}_{b(j,c,t)c} \times \text{fuel}_t$) is defined in a similar way as in equation (14). We present the estimates of this model in Table OA.3 in the Online Appendix.[40]

The second dimension by which we vary our merger simulations is how MillerCoors products are reallocated across breweries after the merger takes place. We make one of two assumptions: a) MillerCoors products sold in a city are reallocated to the closest brewery to that city (henceforth, *distance-based product reallocation*) or b) MillerCoors solves a cost minimization problem, taking into account scale and scope economies, in which they frictionlessly decide how to reallocate products across the full network of Coors and Miller breweries (henceforth, *optimal product reallocation*). This latter cost minimization problem is given

[39] Earlier, we discussed transportation cost savings as a rationale for assuming products were allocated to the closest breweries in the pre-merger environment. In our counterfactual, we consider short-term outcomes where the feasible capacity expansion of breweries is limited (keeping other aspects of the environment fixed).

[40] The estimates of the distance effect are slightly smaller than those in Table 2, with a one standard deviation increase in the distance measure leads to an increase in the marginal cost of 0.8 percent, on average.

by

$$\min_{\mathbf{x}} \sum_{b \in \mathbb{B}_{f_{\text{merge}}}} C_b \left(\mathbf{Q}_b(\mathbf{x}), \mathbf{A}_b(\mathbf{x}), \mathbf{W}, \boldsymbol{\tau}_b(\mathbf{x}) \right) \tag{20}$$

where $\mathbf{Q}_b(\mathbf{x})$ denotes the vector of product–city quantities assigned to brewery b for production, which is a subset of $\mathbf{Q}_{f_{\text{merge}}} = \{Q_c^j(x_c^j) : j \in \mathbb{J}_{f_{\text{merge}}}, c \in \mathbb{C}\}$, representing the quantities sold by Miller and Coors (the combined firm f_{merge}) across all cities c and products j in the *last pre-merger period* t (we omit index t as only one period is considered). The vector $\mathbf{x} = \{x_{bc}^j : j \in \mathbb{J}_{f_{\text{merge}}}, b \in \mathbb{B}_{f_{\text{merge}}}, c \in \mathbb{C}\}$ contains indicator variables where $x_{bc}^j = 1$ if the manufacturing of product j sold in city c is allocated to brewery b, and $x_{bc}^j = 0$ otherwise. $\mathbf{A}_b(\mathbf{x})$ includes a vector of productivity terms for products allocated to brewery b, selected according to $\mathbf{x}$ from $\{A^j : j \in \mathbb{J}_{f_{\text{merge}}}\}$, which contains productivity terms for all products manufactured by Miller or Coors. It also includes a vector of brewery-level productivity terms for all MillerCoors breweries.[41] Finally, $\boldsymbol{\tau}_b(\mathbf{x})$ is a vector of transportation costs associated with shipping product–city quantities allocated to brewery b for production to their respective cities c.[42] $\boldsymbol{\tau}_b(\mathbf{x})$ is a subset (selected according to $\mathbf{x}$) of $\{\tau_{bc} : b \in \mathbb{B}_{f_{\text{merge}}}, c \in \mathbb{C}\}$, which represents the vector of transportation costs from any Miller or Coors brewery to any city c.

The goal of this minimization problem is to identify the cheapest way for Miller and Coors to produce the quantities sold in the last period before the merger if they had access to each other's breweries. It is an approximation to the optimal allocation in the *post-merger* period since it does not take into account price and quantity adjustments associated with the new equilibrium, as we are keeping quantities fixed at the levels of the last month before the merger. However, since such adjustments are not expected to be drastic, the solution to (20) should closely approximate the optimal post-merger allocation.

Despite this simplification, the problem in (20) remains a substantial combinatorial challenge. The merging firms have eight breweries collectively producing hundreds of product–city pairs. To facilitate tractability, we introduce several additional simplifications: (i) instead of reallocating products, we reallocate cities, setting $x_{bc}^j = x_{bc}$ so that all products shipped to a particular city are sourced from a single brewery; (ii) we set brewery capacity constraints based on the maximum observed production in each brewery in the pre-merger period, with an additional buffer of approximately 20% that accommodates the distance-based reallocation discussed above; (iii) we impose relatively relaxed constraints on the distance products may travel from breweries to cities, capping it at roughly the distance

[41] Refer to Online Appendix F for details on the estimation of these components.

[42] For all time-varying components of the cost function, including transportation costs (which change over time due to fluctuations in fuel prices), we use the values from the last pre-merger period to ensure consistency.

from Irwindale, California, to Chicago.[43] We then solve the resulting optimization problem using Gurobi, a state-of-the-art solver for mixed-integer linear programming (Gurobi Optimization, LLC 2024).

Despite the constraints, the number of possible city-to-brewery allocations in our final optimization problem remains large. However, the solver employs a branch-and-bound method that allows it to avoid enumerating all the possible combinations. Ultimately, starting from a distance-based allocation, for most specifications, it finds the optimal solution within a few hours, and for some specifications, it takes only a few minutes.[44] Additional details on how we set up the optimization problem and on the solution algorithm can be found in Online Appendix G.

Compared to the distance-based allocation, the optimal allocation derived from our algorithm results in more concentrated production. The average production in an *active* brewery under the optimal allocation exceeds that in the distance-based allocation by approximately 49.4 thousand units, which corresponds to about a 42% increase.[45] The optimal allocation predicts the closure of three plants—in Colorado, Ohio, and North Carolina.[46] The consolidation of production in the optimal allocation come at the expense of greater travel costs: a unit of beer travels an average of 78.6 miles farther to reach its destination as compared to the distance-based allocation.

Results We compute the equilibrium market outcomes in the equilibria with and without the merger, for every combination of a cost function specification (i.e., with or without scale/scope economies) and a type of post-merger production reallocation (i.e., distance-based or optimal). We compute the impact of the merger on three outcomes (prices, marginal costs, and market shares) separately for the merged (Miller-Coors) and non-merged firms (all other firms) for every choice of cost function and product reallocation. Table 3 presents the results of these comparisons, where the reported values represent the average volume-weighted impact of the merger on each outcome variable. In these comparisons, an observation is a product–city pair for a particular market condition: merger or no merger. We focus on the month prior to the merger taking place: May 2008.

[43] Additionally, if a brewery is located within the city or its suburbs (specifically, within 50 miles of the city center), we assume that products sold in that city are produced at that brewery.

[44] An important limitation of the algorithm is that it can only operate with linear functions. Since our cost function is inherently non-linear, it is linearized during the optimization procedure using a piecewise-linear approximation, with the relative error of the approximation bound to be at most 0.001.

[45] Production is measured in 144-ounce equivalent units (the size of a 12-pack). Note that in the optimal allocation, three breweries are shut down and are not part of the calculation, and that the brewery-level output is within brewery capacity constraints in both allocations.

[46] Notably, in September 2015, MillerCoors actually announced the closure of one of these breweries—the facility in Eden, North Carolina.

Table 3: The Impact of the MillerCoors Merger on Market Outcomes

(1)	(2)	(3)	(4)	(5)	(6)
ΔPrice (in %)		ΔMarginal cost (in %)		ΔMarket share (in %)	
Merged	Non-merged	Merged	Non-merged	Merged	Non-merged
Panel A: Cost model with scale/scope, optimal product reallocation					
-8.30%	0.21%	-16.20%	0.80%	27.60%	-4.39%
Panel B: Cost model with scale/scope, distance-based product reallocation					
3.39%	0.00%	1.92%	-0.21%	-1.41%	0.60%
Panel C: Cost model without scale/scope, optimal product reallocation					
2.19%	0.10%	-0.19%	0.00%	-5.55%	0.93%
Panel D: Cost model without scale/scope, distance-based product reallocation					
2.34%	0.07%	0.05%	0.00%	-5.73%	0.91%

Note: Estimates based on Table 2, column 2. An observation is a product–city pair for a particular market condition: merger or no merger. We compute changes in the volume-weighted mean prices, marginal costs, and market shares of product–city pairs for the merged and non-merged firms (merger versus no merger). Specifically, the volume weights are based on the quantity sold (in 12-pack equivalent units) of each city–product. Merged firms correspond to SABMiller and Coors Molson. Non-merged firms correspond to all other firms. We restrict attention to one time period: May 2008. For the purposes of this table, we drop product–city combinations with a market share of less than 0.000001 in at least one of the counterfactual environments.

Our preferred specification (Table 3, Panel A) features the cost model with scale and scope economies and the optimal post-merger product reallocation. Using this specification, we find that the merger causes a 16.2 percent decrease on average in the marginal costs of MillerCoors products, which combines a series of effects: scope economies (each brewery produces more varieties), scale economies (production is reallocated to increase the scale of some breweries), transportation costs (production moves closer to the destination markets), and the reallocation of products to more efficient plants. The prices of MillerCoors products fall 8.3 percent on average—i.e., the lessening of competition is outweighed by the decrease in marginal costs.

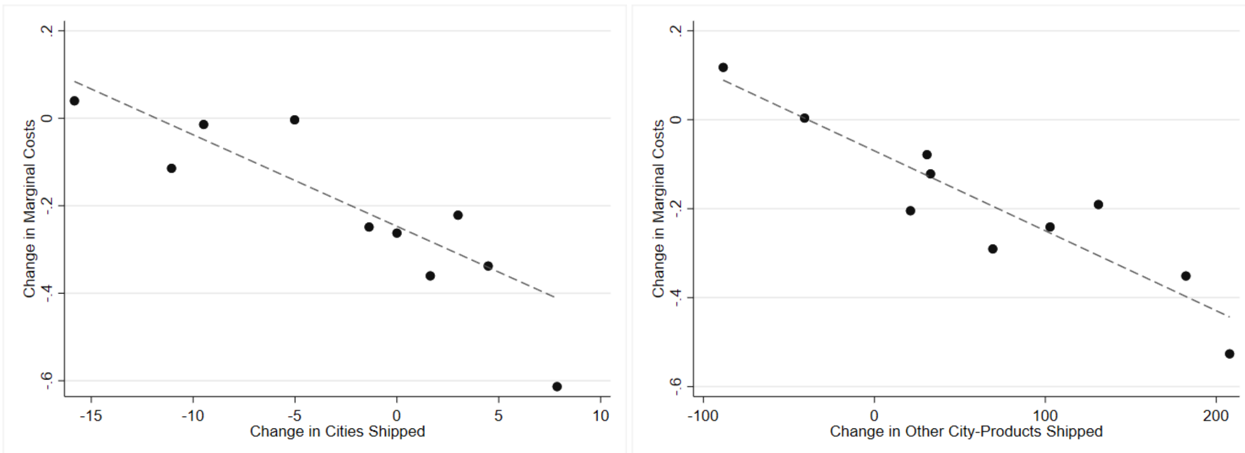
These average marginal cost changes are comparable in magnitude to those in Miller & Weinberg (2017). They find that the marginal cost of MillerCoors products decreased by 52 to 73 cents on average after the merger for reasons *other* than distance-related efficiencies.[47] Using our estimates of the marginal cost of beer, this translates into a decrease in the marginal cost of MillerCoors products of up to 12 percent on average. When also including

[47] These changes in marginal costs are estimated using a linear regression of marginal cost on covariates, which include a post-merger indicator that is specific to MillerCoors. The above-mentioned estimates are the coefficients multiplying this indicator.

Figure 1: Marginal Cost Changes Versus Scale and Scope

(a) Δ MC vs Δ Cities Served

(b) Δ MC vs Δ Other City-Products



Notes: Binned scatter plots of the change in the log of marginal costs (post versus pre-counterfactual for the Panel A optimal product reallocation results of Table 3) versus:

- Panel (a): Change in the number of cities product j is shipped to from brewery $b(j,c,t)$, i.e. $\Delta|\mathbb{C}^{j}_{b(j,c,t)t}|$
- Panel (b): Change in the number of *other* city-products j' within brewery $b(j,c,t)$ i.e. $\Delta\sum_{j'\in(\mathbb{J}_{bt}/j)}|\mathbb{C}^{j'}_{b(j,ct)t}|$

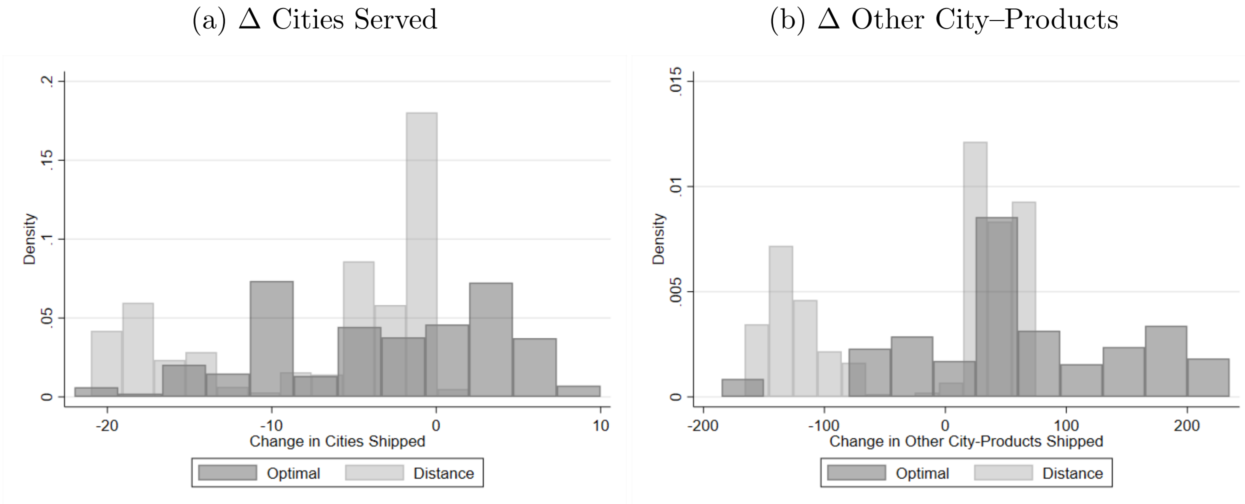
A dot represents the average of the values taken by city-products within 10 equally sized bins based on deciles of the x-axis variable.

transportation cost savings, Miller & Weinberg (2017) predict a 14 percent decrease in the marginal cost of Coors Light. This last figure is similar in magnitude to our estimate of the average change in the marginal cost of MillerCoors products (16.2 percent). It is important to note that these estimates in Miller & Weinberg (2017) rely on a comparison of marginal costs before and after the merger while remaining agnostic about scale and scope economies. That our model is able to reproduce these results without using post-merger data provides support for our model and cost function estimates. On the other hand, our estimates of price changes are somewhat different as i) we do not model post-merger price coordination and ii) Miller & Weinberg (2017) do not model scale and scope economies, which make marginal costs change across counterfactual environments.[48]

Figure 1 illustrates that much of the marginal cost declines we observe are driven by city–products that experience increases in scale and scope due to consolidation of production. Here, we present binned scatter plots at the city-product level of the change in log marginal costs implied by our counterfactual, versus the change in the number of cities a particular product is shipped to (Panel A)—a proxy for scale changes driven by the reallocation of

[48]Note that our model allows us to account for changes in the marginal cost of non-merging firms as a direct consequence of the merger, driven by changes in market shares and, consequently, the scale and scope of their production.

Figure 2: Histogram: Optimal Reallocation Versus Distance Reallocation



Notes: Histograms of:

- Panel (a): Change in the number of cities product j is shipped to from brewery $b(j,c,t)$, i.e. $\Delta|\mathbb{C}^{j}_{b(j,c,t)t}|$
- Panel (b): Change in the number of *other* city-products j' within brewery $b(j,c,t)$ i.e. $\Delta\sum_{j'\in(\mathbb{J}_{bt}/j)}|\mathbb{C}^{j'}_{b(j,ct)t}|$

Dark grey presents these changes for the Panel A optimal product reallocation results of Table 3. Light grey presents changes for the distance-based reallocation results in Panel B of Table 3.

production—as well as the change in the number of *other* city–products shipped from the same brewery (Panel B)—a proxy for scope changes due to the reallocation of production. We see that these changes in scale and scope drive much of the marginal cost changes: specifically, city–products above the 90th percentile of the scale and scope proxies experience declines in marginal cost that are greater than 50%, while products below the 10th percentile experience marginal costs increases between 10% and 20%.

That some city–products experience marginal cost increases is driven by the fact that consolidation of production in one location will tend to involve less consolidated production in other locations. This further implies that reallocating production improperly—i.e., not fully internalizing scale and scope effects—can actually be harmful to costs overall. Table 3, Panel B repeats the merger analysis in Panel A but uses a distance-based product reallocation (as opposed to the optimal product reallocation). Here, we find that marginal costs actually *rise* for the merging firms on average, generating an average increase in marginal costs and prices of 1.9 and 3.4 percent, respectively. This is because distance-based reallocation of products minimizes the distance to the destination city, while ignoring the benefits of producing with sufficient scale and scope.

We further illustrate this in Figure 2, which compares our proxies for reallocation-induced scale and scope changes for the optimal reallocation and the distance-based counterfactual.

Here, we see that the optimal reallocation of production increases scale and scope for a much larger mass of city–products, while the distance-based reallocation *decreases* the scale and scope for many more city–products. This is relevant from a competition policy standpoint, as an efficiency defense cannot be based on evidence in favor of scale and scope economies unless it can be established that it is feasible to reallocate products in a way that leads to cost savings and that changes in production scale will not erase these savings.

How does ignoring scale and scope economies impact our results? We repeat our merger analysis utilizing the simplified cost function (see equation 19) using the optimal and distance-based product reallocations in Panels C and D of Table 3, respectively.[49] Compared to Panel A, we find smaller marginal cost savings for MillerCoors products (i.e., less negative) when using the simplified cost function in Panels C and D. This tends to overpredict the merger's effects on the prices of MillerCoors products.[50] Relative to Panel B, Panels C and D overpredict cost savings and, thus, underpredict the price effects of the merger. That is, ignoring scale and scope economies will bias the merger effects but the direction of the bias will depend on how products are reallocated after the merger. Ultimately, accurately predicting merger effects requires not only understanding the presence of scale and scope economies in production but also how products are reallocated after the merger.

In summary, the existence of scale and scope economies creates the possibility of merger-specific cost savings when the merging firms combine production processes. Whether these cost savings are realized will crucially depend on how products are reallocated across the plants of the merging parties and how production scale changes in these plants. As mentioned, evidence in favor of scale and scope economies cannot be the basis of an efficiency defense by the merging parties unless it can be established that the product reallocation scheme will generate cost savings.[51] As our results show, a merger may increase marginal costs even in the presence of scale and scope economies.

5 Concluding Remarks

We propose a new method to estimate economies of scale and scope suitable for applied work. Our method requires data commonly used for demand estimation (crucially, quantities

[49]Even when not modeling scale and scope economies, products need to be reallocated across breweries. Here, we use the same reallocation schemes as in Table 3, panels A and B.

[50]Panel D shows cost increases because the closest brewery may be less efficient than a farther one, given that our model incorporates brewery-specific productivity. This trade-off is accounted for in the optimal reallocation, which is why cost savings appear only under constant marginal costs in Panel C.

[51]Similarly, evidence of the importance of transportation costs cannot serve as the sole basis for an efficiency defense by arguing for distance-based product reallocation (as was done in the MillerCoors merger case) without also discussing strategies to mitigate potential negative consequences of production disaggregation.

produced and prices for each product–market combination) together with the information on production facility locations, but does not require input data, making it easy to implement in other settings.

We apply our method to the US beer industry, which is an ideal setting to investigate the existence of scale and scope economies, as it features multiproduct firms and production that is consolidated in a small number of plants. Our estimates suggest the existence of both scale and scope economies.

We then use our estimates to explore the implications of scale and scope economies for merger analysis using the MillerCoors merger. The antitrust investigation of the merger considered the tradeoff between two forces: the enhanced market power of MillerCoors and the transportation cost savings that would arise from the merged firm using a more geographically diversified network of production plants. Our framework considers additional effects that are potentially relevant: i) the enhanced market power of the merged firm leads to a decrease in the scale of production, which creates cost increases that at least partially offset the transportation cost savings; ii) conditional on a given post-merger production reallocation, some plants may gain scale (e.g., by supplying products to more cities) and scope (e.g., by producing a larger variety of products), while others may lose one or both of these advantages. This will lead to heterogeneous changes in marginal costs across different plants and, depending on the distribution of these changes, may result in varying average cost outcomes.

We show that the balance of these effects—and whether the merger causes cost savings—depends crucially on how MillerCoors products are reallocated across breweries after the merger. If the reallocation of products fails to take into account the impacts on production scale and scope, the merger may increase marginal costs, even though a more geographically diversified network of plants helps reduce shipping costs. An implication for competition policy is that evidence in favor of scale and scope economies cannot be the basis of an efficiency defense by the merging parties unless it can be established that the proposed product reallocation scheme will actually generate cost savings.

References

Argente, D., Moreira, S., Oberfield, E. & Venkateswaran, V. (2020), ‘Scalable expertise’, *Working Paper* .

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

Ashenfelter, O. C., Hosken, D. S. & Weinberg, M. C. (2015), 'Efficiencies brewed: pricing and consolidation in the us beer industry', *The RAND Journal of Economics* **46**(2), 328–361.

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

Asker, J. & Nocke, V. (2021), Collusion, mergers, and related antitrust issues, *in* 'Handbook of industrial organization', Vol. 5, Elsevier, pp. 177–279.

Baumol, J. W., Panzar, J. C. & Willig, R. D. (1982), *Contestable Markets and the Theory of Industry Structure*, Harcourt College Pub, New York.

Berry, S., Levinsohn, J. & Pakes, A. (1995), 'Automobile prices in market equilibrium', *Econometrica: Journal of the Econometric Society* pp. 841–890.

Berry, S. T. (1994), 'Estimating discrete-choice models of product differentiation', *The RAND Journal of Economics* **25**, 242–262.

Berry, S. T. & Haile, P. A. (2014), 'Identification in differentiated products markets using market level data', *Econometrica* **82**(5), 1749–1797.

Bighelli, T., Di Mauro, F., Melitz, M. J. & Mertens, M. (2023), 'European firm concentration and aggregate productivity', *Journal of the European Economic Association* **21**(2), 455–483.

Braguinsky, S., Ohyama, A., Okazaki, T. & Syverson, C. (2015), 'Acquisitions, productivity, and profitability: evidence from the japanese cotton spinning industry', *American Economic Review* **105**(7), 2086–2119.

Bronnenberg, B. J., Kruger, M. W. & Mela, C. F. (2008), 'Database paperâthe iri marketing data set', *Marketing science* **27**(4), 745–748.

Cairncross, J., Morrow, P., Orr, S. & Rachapalli, S. (2024), 'Identifying firm vs. product markups using production data: Micro estimates and aggregate implications', *Working Paper* .

Caves, D. W., Christensen, L. R. & Diewert, W. E. (1982), 'The economic theory of index numbers and the measurement of input, output, and productivity', *Econometrica* pp. 1393–1414.

Chen, Y. (2024), 'Network structure and the efficiency gains from mergers: Evidence from us freight railroads', *Available at SSRN 4742990* .

Conlon, C. & Gortmaker, J. (2020), 'Best practices for differentiated products demand estimation with pyblp', *The RAND Journal of Economics* **51**(4), 1108–1161.

Costinot, A. & Rodríguez-Clare, A. (2014), Trade theory with numbers: Quantifying the consequences of globalization, *in* 'Handbook of international economics', Vol. 4, Elsevier, pp. 197–261.

De Loecker, J. & Scott, P. T. (2024), Estimating market power evidence from the us brewing industry, Technical report, Working Paper.

De Loecker, J. & Syverson, C. (2021), An industrial organization perspective on productivity, *in* 'Handbook of Industrial Organization', Vol. 4, Elsevier, pp. 141–223.

Demirer, M. & Karaduman, Ö. (2024), Do mergers and acquisitions improve efficiency? evidence from power plants, Technical report, National Bureau of Economic Research.

Dhyne, E., Petrin, A., Smeets, V. & Warzynski, F. (2022), 'Theory for extending single-product production function estimation to multi-product settings', *Working Paper* .

Ding, X. (2022), 'Industry linkages from joint production', *Work. Pap., Georgetown Univ., Washington, DC* .

Döpper, H., MacKay, A., Miller, N. & Stiebale, J. (2022), 'Rising markups and the role of consumer preferences', *Harvard Business School Strategy Unit Working Paper* (22-025).

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

Eizenberg, A. & Zvuluni, O. (2024), 'An empirical analysis of merger efficiencies', *Working Paper* .

Elliott, J., Houngbonon, G. V., Ivaldi, M. & Scott, P. (2025), 'Market structure, investment and technical efficiencies in mobile telecommunications', *Journal of Political Economy* **133**(5), 1401–1459.

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

Fan, Y. & Yang, C. (2020), 'Competition, product proliferation, and welfare: A study of the us smartphone market', *American Economic Journal: Microeconomics* **12**(2), 99–134.

Fan, Y. & Yang, C. (2025), 'Estimating discrete games with many firms and many decisions: An application to merger and product variety', *Journal of Political Economy* **133**(6), 1886–1931.

Gandhi, A. & Houde, J.-F. (2019), 'Measuring substitution patterns in differentiated-products industries', *NBER Working paper* (w26375).

Grieco, P., Pinkse, J. & Slade, M. (2018), 'Brewed in north america: Mergers, marginal costs, and efficiency', *International Journal of Industrial Organization* **59**, 24–65.

Grullon, G., Larkin, Y. & Michaely, R. (2019), 'Are us industries becoming more concentrated?', *Review of Finance* **23**(4), 697–743.

Gurobi Optimization, LLC (2024), 'Gurobi Optimizer Reference Manual'.
URL: *https://www.gurobi.com*

Hall, R. E. (1973), 'The specification of technology with several kinds of output', *Journal of Political Economy* **81**(4), 878–892.

Hausman, J. A. & Taylor, W. E. (1983), 'Identification in linear simultaneous equations models with covariance restrictions: An instrumental variables interpretation', *Econometrica* pp. 1527–1549.

Hausman, J., Leonard, G. & Zona, J. D. (1994), 'Competitive analysis with differenciated products', *Annales d'Economie et de Statistique* pp. 159–180.

Heyer, K., Shapiro, C. & Wilder, J. (2009), 'The year in review: Economics at the antitrust division, 2008–2009', *Review of Industrial Organization* **35**, 349–367.

Hidalgo, J. (2023), 'This craft's for you! entry and market (ing) competition in the us beer industry'.

Jeziorski, P. (2014), 'Effects of mergers in two-sided markets: The us radio industry', *American Economic Journal: Microeconomics* **6**(4), 35–73.

Johnes, G. (1997), 'Costs and industrial structure in contemporary british higher education', *The Economic Journal* **107**(442), 727–737.

Keithahn, C. F. (1978), *The brewing industry*, Federal Trade Commission, Bureau of Economics.

Kerkvliet, J. R., Nebesky, W., Tremblay, C. H. & Tremblay, V. J. (1998), 'Efficiency and technological change in the us brewing industry', *Journal of Productivity Analysis* **10**(3), 271–288.

Kohli, U. R. (1981), 'Nonjointness and factor intensity in us production', *International Economic Review* pp. 3–18.

MacKay, A. & Miller, N. H. (2025), 'Estimating models of supply and demand: Instruments and covariance restrictions', *American Economic Journal: Microeconomics* **17**(1), 238–281.

Maican, F. & Orth, M. (2020), 'Determinants of economies of scope in retail', *Working Paper* .

Miller, N. H., Sheu, G. & Weinberg, M. C. (2021), 'Oligopolistic price leadership and mergers: The united states beer industry', *American Economic Review* **111**(10), 3123–3159.

Miller, N. H. & Weinberg, M. C. (2017), 'Understanding the price effects of the millercoors joint venture', *Econometrica* **85**(6), 1763–1791.

Nevo, A. (2000), 'Mergers with differentiated products: The case of the ready-to-eat cereal industry', *RAND journal of economics* pp. 395–421.

Orr, S. (2022), 'Within-firm productivity dispersion: Estimates and implications', *Journal of Political Economy* **130**(11), 2771–2828.

Panzar, J. C. & Willig, R. D. (1975), 'Economies of scale and economies of scope in multi-output production', *Bell Laboratories economic discussion paper* **33**.

Panzar, J. C. & Willig, R. D. (1981), 'Economies of scope', *The American Economic Review* **71**(2), 268–272.

Raval, D. (2023), 'Testing the production approach to markup estimation', *Review of Economic Studies* **90**(5), 2592–2611.

Raval, D. (2024), 'Economies of scope and common inputs in multi-output production', *Working Paper* .

Redding, S. J. (2022), Trade and geography, *in* 'Handbook of International Economics', Vol. 5, Elsevier, pp. 147–217.

Rosse, J. N. (1970), 'Estimating cost function parameters without using cost data: Illustrated methodology', *Econometrica* pp. 256–275.

Shephard, R. W. (1970), *Theory of cost and production functions*, Princeton University Press.

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

Williamson, O. E. (1968), 'Economies as an antitrust defense: The welfare tradeoffs', *The American Economic Review* **58**(1), 18–36.

Wollmann, T. G. (2018), 'Trucks without bailouts: Equilibrium product characteristics for commercial vehicles', *American Economic Review* **108**(6), 1364–1406.

Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, second edn, MIT press.

Zhang, J. & Malikov, E. (2022), 'Off-balance sheet activities and scope economies in us banking', *Journal of Banking & Finance* **141**, 106534.

A A Model of Public and Private Tasks

In this Appendix, we derive the cost function (1) for the special case of Cobb-Douglas production.[52] To simplify notation, we suppress time subscripts and consider a single firm/brewery.

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

$$Y_b^j = \frac{A_b^j}{C} \left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p} \right), \tag{21}$$

where $C \equiv \frac{\Pi_X \left(\beta_X^r\right)^{\beta_X^r} \left(\beta_X^p\right)^{\beta_X^p}}{\Pi_X \left(\beta_X^r + \beta_X^p\right)^{\beta_X^r + \beta_X^p}}$ is a constant.

We assume that firms choose the allocation of inputs across tasks, given $\mathbf{X}_b$, to produce the maximal quantities of output feasible; i.e., the firm always operates on their production possibilities frontier. One way to characterize the solution to this problem is by solving for a firm's *output distance function* (Shephard 1970, Caves et al. 1982), which tells us the minimum amount a firm must scale down a given output vector $\mathbf{Y}_b$ to make sure that $(\frac{\mathbf{Y}_b}{\delta}, \mathbf{X}_b) \in \mathbb{P}_b$, where $\mathbb{P}_b$ is the firm's production possibility set, and δ is the minimized scaling factor. For this particular production problem, the firm's output distance function is given by:

$$\begin{aligned} D\left(\mathbf{Y}_b, \mathbf{X}_b, \mathbf{A}_b\right) \equiv & \min_{\delta, \mathbf{X}_b^p, \{\mathbf{X}_b^{rj}\}_j} \quad \delta \\ & \text{s.t.:} \quad \frac{Y_b^j}{\delta} \leq \frac{A_b^j}{C} \left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p} \right) \ \forall j \\ & \qquad X_b^p + \sum_j X_b^{rj} \leq X_b, \ \forall X. \end{aligned} \tag{22}$$

This optimization problem has the following Lagrangian

$$L = \delta + \sum_j \lambda_b^j \left(\frac{Y_b^j}{\delta} - \frac{A_b^j}{C} \left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p} \right) \right) + \sum_X \mu_X \left(\sum_j X_b^{rj} + X_b^p - X_b \right). \tag{23}$$

Since the production functions are strictly increasing in all inputs and generate zero output whenever $X = 0$, all constraints will bind with equality, and therefore $\lambda_b^j > 0 \ \forall j$ and

[52] For more general specifications of the technology, see Cairncross et al. (2024).

$\mu_X > 0 \; \forall X$, with:

$$\frac{Y_b^j}{\delta} = \frac{A_b^j}{C}\left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p}\right) \quad \forall j \tag{24}$$

and

$$X_b^p + \sum_j X_b^{rj} = X_b, \quad \forall X. \tag{25}$$

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

$$\lambda_b^j \beta_X^r \frac{\frac{A_b^j}{C}\left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p}\right)}{X_b^{rj}} = \lambda_b^j \beta_X^r \frac{\frac{Y_b^j}{\delta}}{X_b^{rj}} = \mu_X, \tag{26}$$

whereas the first order condition for X_b^p satisfies

$$\sum_j \lambda_b^j \beta_X^p \frac{\frac{A_b^j}{C}\left(\prod_X \left(X_b^{rj}\right)^{\beta_X^r} \left(X_b^p\right)^{\beta_X^p}\right)}{X_b^p} = \frac{\beta_X^p}{\delta X_b^p}\sum_j \lambda_b^j Y_b^j = \mu_X. \tag{27}$$

To solve this problem, it is useful to first solve for total quantity of inputs going into various rival tasks, $X_b^r \equiv \sum_j X_b^{rj}$, relative to the total quantity of inputs going into non-rival task, X_b^p, for each input X. To do this, we can take the ratio of (26) and (27), which will cancel out the unknown input-specific Lagrangian multiplier μ_X. Rearranging and summing equation (26) for all j, yields:

$$\mu_X X_b^r = \frac{\beta_X^r}{\delta}\sum_j \lambda_b^j Y_b^j. \tag{28}$$

We can then rearrange equation (27) and divide by equation (28) to obtain

$$\frac{X_b^p}{X_b^r} = \frac{\beta_X^p}{\beta_X^r}. \tag{29}$$

Since $X_b = X_b^r + X_b^p$, substitute (29) into this expression, yielding

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

or

$$X_b^r = \frac{\beta_X^r}{\beta_X^r + \beta_X^p} X_b \tag{30}$$

and

$$X_b^p = \frac{\beta_X^p}{\beta_X^r + \beta_X^p} X_b. \tag{31}$$

These two expressions tell us that the share of input X being allocated to the public task is $\frac{\beta_X^p}{\beta_X^r+\beta_X^p}$, while $\frac{\beta_X^r}{\beta_X^r+\beta_X^p}$ is the share of input X allocated to rival tasks.

We can use this latter fact to obtain the quantity of input X being allocated to the rival task in product line j by rearranging equation (26) and dividing by equation (28), yielding:

$$X_b^{rj} = \frac{\lambda_b^j Y_b^j}{\sum_k \lambda_b^k Y_b^k} X_b^r. \tag{32}$$

At this point, we have managed to characterize the optimal quantities of each input X being allocated to the public task and the various rival tasks by product-line. While the product-line specific rival input shares depend on unknown Lagrangian multipliers λ_b^j, we do not need to solve for the Lagrangian multipliers to characterize δ, the value of the output distance function given $(\mathbf{Y}_b, \mathbf{X}_b)$. To see this, substitute equations (30), (31) and (32) into equation (24), which yields:

$$\frac{Y_b^j}{\delta} = \frac{A_b^j}{C}\left(\prod_X \left(\frac{\lambda_b^j Y_b^j}{\sum_k \lambda_b^k Y_b^k}\frac{\beta_X^r}{\beta_X^r+\beta_X^p}X_b\right)^{\beta_X^r}\left(\frac{\beta_X^p}{\beta_X^r+\beta_X^p}X_b\right)^{\beta_X^p}\right).$$

By defining $\alpha \equiv \sum_X \beta_X^r$, $\beta_X = \beta_X^r + \beta_X^p$, and $\phi \equiv \sum_X \beta_X$, and rearranging and canceling out terms in the equation above, we obtain

$$\left(\frac{Y_b^j}{\delta A_b^j}\right)^{\frac{1}{\alpha}} = \frac{\lambda_b^j Y_i^j}{\sum_k \lambda_b^k Y_b^k}\left(\prod_X (X_b)^{\beta_X}\right)^{\frac{1}{\alpha}}.$$

By summing over all j, we obtain

$$\frac{1}{\delta^{\frac{1}{\alpha}}}\sum_j \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}} = \left(\prod_X (X_b)^{\beta_X}\right)^{\frac{1}{\alpha}}$$

or

$$\delta = \frac{\left(\sum_j \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{\prod_X (X_b)^{\beta_X}},$$

which establishes that the firm's distance function is $D(\mathbf{Y}_b, \mathbf{X}_b, \mathbf{A}_b) = \frac{\left(\sum_j \left(\frac{Y_b^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{\prod_X (X_b)^{\beta_X}}$.[53]

Assuming that a firm operates on its production possibility frontier (i.e., does not waste

[53]While we did not use the first order condition for δ, which implies $\delta^2 = \sum_j \lambda_b^j Y_b^j$, we would use this expression to solve for μ_X and λ_b^j

any inputs when producing some desired output vector $\mathbf{Y}_b$) means that the firm will only produce $(\mathbf{Y}_b, \mathbf{X}_b)$ satisfying $D(\mathbf{Y}_b, \mathbf{X}_b, \mathbf{A}_b) = 1$. This implies that

$$\frac{\left(\sum_j \left(\frac{Y_i^j}{A_i^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{\prod_X (X_i)^{\beta_X}} = 1 \tag{33}$$

To generate the relevant cost function, it is useful to rearrange equation (33) as follows:

$$\mathcal{Y}_b \equiv \left(\sum_{j\in\mathbb{J}_b} \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha} = \prod_X (X_b)^{\beta_X}. \tag{34}$$

Equation (34) provides a "psuedo" Cobb-Douglas production function for the output aggregator $\mathcal{Y}_b \equiv \left(\sum_j \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}$. Under the further assumption that all inputs $\mathbf{X}_b$ are obtained from perfectly competitive markets at prices $\mathbf{W}_b$, then it is well known that the cost function for production function (34) is given by

$$C(\mathcal{Y}_b, \mathbf{W}_b) = \underbrace{K\left(\prod_X (W_{bX})^{\frac{\beta_X}{\sum \beta_X}}\right)}_{\equiv g(\mathbf{W}_b)} (\mathcal{Y}_b)^{\frac{1}{\phi}} \tag{35}$$

where K is a constant that depends on the various β_X terms.

Substituting the definition of the output aggregator—$\mathcal{Y}_b \equiv \left(\sum_j \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}$—into equation (36) then yields:

$$C(\mathbf{Y}_b, \mathbf{A}_b, \mathbf{W}_b) = g(\mathbf{W}_b)\left(\sum_j \left(\frac{Y_b^j}{A_b^j}\right)^{\frac{1}{\alpha}}\right)^{\frac{\alpha}{\phi}}. \tag{36}$$

Note that from this derivation of the firm's cost function, we can see that α, which we defined as $\alpha \equiv \sum_X \beta_X^r$, should be interpreted as the returns to scale in *rival* or *private* tasks, while overall returns to scale, $\phi \equiv \sum_X \beta_X$ depends on both private and public tasks, so $\alpha \leq \phi$. However, our key estimating equation determines whether there are economies of scope by examining whether $\frac{\phi-\alpha}{\phi} > 0$. We now show that this parameter can be interpreted as a cost-share weighted average of the share of each input X allocated to the public task.

In particular, equations (30) and (31) imply for each input X, the quantity share of that input allocated to the public task is $S_{bX}^p \equiv \frac{X_b^p}{X_b} = \frac{\beta_X^p}{\beta_X^r+\beta_X^p}$, while the share allocated to the private task is $S_{bX}^r \equiv \frac{X_b^r}{X_b} = \frac{\beta_X^r}{\beta_X^r+\beta_X^p}$. From the definition of α and ϕ, we have:

$$\frac{\phi-\alpha}{\phi}=\frac{\sum_X \beta_X^p}{\sum_X \beta_p^r+\beta_X^p}. \tag{37}$$

This immediately implies that if public task shares do not depend on X, i.e., $\beta_X^p=\beta^p$ and $\beta_X^r=\beta^r\ \forall X$, then $\frac{\phi-\alpha}{\phi}$ is simply the public task share $S^p=\frac{\beta^p}{\beta^r+\beta^p}$.

When public and private task shares vary across inputs, $\frac{\phi-\alpha}{\phi}$ measures a cost share weighted average of the input specific public task shares. To see this, note that since the output aggregator $\mathcal{Y}_b$ is produced using a Cobb-Douglas production function, equilibrium cost shares for each input satisfy $C_{bX}\equiv\frac{W_{bX}X}{\sum_{X'}W_{X'b}X'}=\frac{\beta_X}{\sum_{X'}\beta_{X'}}$. Since $\beta_X=\beta_X^r+\beta_X^p$ and $\sum_X \beta_X=\phi$, we also have $C_{bX}=\frac{\beta_X^r+\beta_X^p}{\phi}$, or, $\phi=\frac{\beta_X^r+\beta_X^p}{C_{bX}}$. Substituting this into (37) then yields:

$$\frac{\phi-\alpha}{\phi}=\frac{\sum_X \beta_X^p}{\phi}=\sum_X \frac{\beta_X^p}{\beta_X^r+\beta_X^p}C_{bX}=\sum_X S_{bX}^p C_{bX}. \tag{38}$$

B Properties of the Cost Function

Proof of Lemma 1

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

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

$$\left(\sum_j\left(\frac{Y_{bt}^j}{A_{bt}^j}\right)^{\frac{1}{\alpha}}\right)^{\alpha}<\left(\sum_j\left(\frac{Y_{bt}^j}{A_{bt}^j}\right)^{\frac{1}{\phi}}\right)^{\phi},$$

which holds true given that

$$\left(\sum_j\left(\frac{Y_{bt}^j}{A_{bt}^j}\right)^{\frac{1}{x}}\right)^{x}$$

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

□

Online Appendix: Not For Publication

Identifying Scale and Scope Economies using Product Market Data

Ekaterina Khmelnitskaya, Guillermo Marshall, and Scott Orr

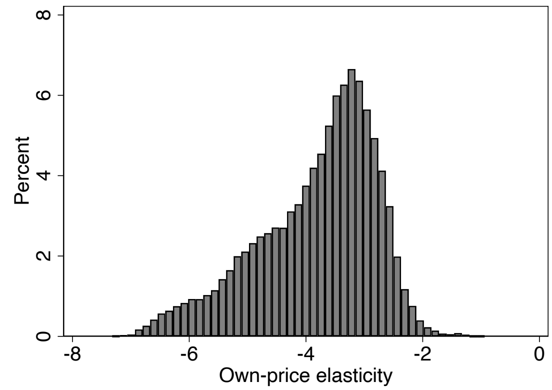
C Additional Tables and Figures

Table OA.1: Summary Statistics

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		6-pack			12-pack			24/30-pack		
Brand	Parent	Average price	Share	Observations	Average price	Share	Observations	Average price	Share	Observations
Blue Moon Belgian White Ale	Molson Coors	14.807	0.002	1,552	13.796	0.002	1,304	7.299	0.000	5
Bud Light	Anheuser-Busch	11.329	0.007	1,558	9.688	0.031	1,558	8.237	0.040	2,111
Budweiser	Anheuser-Busch	11.309	0.005	1,558	9.670	0.016	1,558	8.234	0.021	2,076
Budweiser Select	Anheuser-Busch	11.330	0.001	1,521	9.698	0.005	1,522	8.151	0.004	1,696
Busch	Anheuser-Busch	8.316	0.001	1,136	7.334	0.005	1,557	6.308	0.009	1,815
Busch Light	Anheuser-Busch	8.327	0.001	1,080	7.391	0.007	1,414	6.327	0.016	1,953
Coors	Molson Coors	11.374	0.000	1,366	9.671	0.002	1,557	8.123	0.002	1,801
Coors Light	Molson Coors	11.325	0.003	1,558	9.694	0.014	1,558	8.168	0.019	2,153
Corona Extra	Grupo Modelo	16.505	0.005	1,558	14.256	0.021	1,558	13.459	0.004	893
Corona Light	Grupo Modelo	16.510	0.002	1,558	14.279	0.007	1,558	13.240	0.000	393
Heineken	Heineken	16.479	0.004	1,558	14.171	0.013	1,558	12.056	0.002	456
Heineken Premium Light Lager	Heineken	16.450	0.001	1,035	14.160	0.003	1,029	12.604	0.001	151
Keystone Light	Molson Coors	7.958	0.000	211	6.721	0.003	1,456	5.854	0.007	1,984
Labatt Blue	InBev	11.242	0.001	123	10.320	0.004	642	8.097	0.011	429
Michelob Light	Anheuser-Busch	12.042	0.001	1,538	10.606	0.004	1,450	7.997	0.001	73
Michelob Ultra	Anheuser-Busch	12.136	0.003	1,558	10.785	0.010	1,558	9.941	0.001	339
Miller Genuine Draft	SABMiller	11.380	0.002	1,523	9.684	0.005	1,558	8.051	0.005	2,020
Miller High Life	SABMiller	8.777	0.001	1,373	7.347	0.008	1,558	6.288	0.009	1,975
Miller Lite	SABMiller	11.326	0.004	1,558	9.649	0.018	1,558	8.152	0.029	2,103
Modelo Especial	Grupo Modelo	15.309	0.001	1,464	13.601	0.002	1,422	12.015	0.001	39
Natural Ice	Anheuser-Busch	7.487	0.001	886	6.617	0.005	1,523	5.812	0.004	1,827
Natural Light	Anheuser-Busch	7.645	0.003	1,018	6.644	0.011	1,553	5.915	0.014	2,208
Pabst Blue Ribbon	S&P	8.222	0.000	898	6.887	0.004	1,543	6.056	0.004	1,514
Tecate	FEMSA	13.315	0.001	1,484	11.178	0.003	1,201	8.286	0.013	307
Yuengling Traditional Lager	DG Yuengling	10.962	0.003	459	9.569	0.010	459	8.427	0.004	342

Note: An observation is a brand–size–city–month combination. We measure market shares based on 144-ounce equivalent units (the size of a 12-pack). Prices are also expressed in terms of the size of a 12-pack.

Figure OA.1: Distribution of own-price elasticities



Notes: An observation is a product–city–month combination. The elasticities are constructed based on the estimates in Table 1 (Column 2).

Table OA.2: Mean Elasticities for 12-Pack Products

	Brand	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1)	Bud Light	-4.193	0.059	0.081	0.053	0.050	0.030	0.064	0.016	0.026	0.103
(2)	Budweiser	0.110	-3.282	0.051	0.090	0.038	0.052	0.039	0.018	0.023	0.061
(3)	Coors Light	0.171	0.057	-4.457	0.046	0.051	0.026	0.068	0.016	0.026	0.111
(4)	Corona Extra	0.053	0.050	0.021	-3.278	0.021	0.071	0.016	0.016	0.016	0.023
(5)	Corona Light	0.150	0.062	0.072	0.063	-5.934	0.037	0.056	0.018	0.026	0.088
(6)	Heineken	0.052	0.049	0.020	0.122	0.020	-3.274	0.016	0.015	0.016	0.022
(7)	Michelob Ultra	0.172	0.057	0.088	0.045	0.052	0.025	-5.012	0.016	0.026	0.112
(8)	Miller Genuine Draft	0.100	0.058	0.045	0.096	0.035	0.056	0.035	-3.147	0.022	0.054
(9)	Miller High Life	0.127	0.059	0.061	0.078	0.042	0.045	0.047	0.017	-2.760	0.073
(10)	Miller Lite	0.177	0.056	0.091	0.041	0.052	0.023	0.071	0.015	0.026	-4.532

Notes: The table shows the elasticity of product j (row) with respect to the price of product k (column). The labels of columns are the same as those for rows. Results are based on the estimates in Table 1 (Column 2). We include products available in all city-month combinations.

Table OA.3: Estimates of the Cost Function Without Scale and Scope Economies

	(1)
	$\ln MC_{ct}^{j}$
ψ (distance measure)	0.0050
	(0.0009)
N	96,899

Notes: Robust standard errors in parentheses. The table presents estimates for equation (19). Observations where $MC_{it}^{j} < 0$ (less than 1% of full sample) are excluded from these regressions. All regressions include product, city, brewery, and year fixed effects.

D First Stage Regressions

Table OA.4: Cost Function Estimates: First Stages

	(1) Scale	(2) Scale	(3) Scale	(4) Scope	(5) Scope	(6) Scope
Scale IV	2.327	0.852	–0.151	2.276	0.835	–0.161
	[2.252 , 2.405]	[0.841 , 0.869]	[–0.170 , –0.135]	[2.195 , 2.345]	[0.817 , 0.845]	[–0.176 , –0.140]
Scope IV	0.069	–0.486	0.170	–0.775	–0.853	0.109
	[–0.281 , 0.378]	[–0.515 , –0.462]	[0.144 , 0.191]	[–0.943 , –0.627]	[–0.876 , –0.837]	[0.085 , 0.128]
Transportation	0.005	0.010	0.026	–0.012	0.005	0.002
	[–0.010 , 0.013]	[–0.006 , 0.019]	[0.002 , 0.040]	[–0.016 , –0.007]	[–0.003 , 0.012]	[–0.006 , 0.008]
Observations	91,112	91,112	91,112	91,112	91,112	91,112

Notes: Reports first-stage regressions for the IV regression in columns (2), (3), and (4) of Table 2. The outcomes of interest are $\left(\ln Q^j_{ct} - \ln S^{c|j}_{b(j,c,t)t}\right)$ in columns (1) - (3) (Scale), $\ln S^j_{b(j,c,t)t}$ in columns (4) - (6) (Scope). Independent variables are $\text{Transportation}_{b(j,c,t),t} = \text{dist}_{b(j,c,t)c} \times \text{fuel}_t$ and:

- (1) and (4): $\text{SCALE\_IV}^j_{ct} = \ln|\mathbb{C}^j_{bt}|$; $\text{SCOPE\_IV}^j_{ct} = \ln\left(\sum_{j'\in(\mathbb{J}_{bt}/j)} |\mathbb{C}^{j'}_{bt}|\right)$
- (2) and (5): $\text{SCALE\_IV}^j_{ct} = \frac{1}{|\mathbb{C}^j_{bt}|}\sum_{c\in\mathbb{C}^j_{bt}} \xi^j_{ct}$; $\text{SCOPE\_IV}^j_{ct} = \frac{1}{\sum_{j'\in\mathbb{J}_{bt}/j}|\mathbb{C}^{j'}_{bt}|}\sum_{j'\in(\mathbb{J}_{bt}/j)}\sum_{c\in\mathbb{C}^{j'}_{bt}} \xi^{j'}_{ct}$
- (3) and (6): $\text{SCALE\_IV}^j_{ct} = \frac{1}{|\mathbb{C}^j_{bt}|}\sum_{c\in\mathbb{C}^j_{bt}} \text{DiffCal}^j_{ct}$, and $\text{SCOPE\_IV}^j_{ct}\frac{1}{\sum_{j'\in\mathbb{J}_{bt}/j}|\mathbb{C}^{j'}_{bt}|}\sum_{j'\in(\mathbb{J}_{bt}/j)}\sum_{c\in\mathbb{C}^{j'}_{bt}} \text{DiffCal}^{j'}_{ct}$

where $\text{DiffCal}^j_{ct} \equiv \sum_{k\in\mathbb{J}_{ct}\setminus\mathbb{J}_{f(j)ct}} \mathbb{I}\left(|\text{Calories}^j_{c,t} - \text{Calories}^k_{c,t}| < \text{SD}_{\text{Calories}}\right)$, with $f(j)$ denoting the firm producing product j, $\mathbb{I}(.)$ denoting the indicator function. Bootstrapped 95-percent confidence intervals, taking into account demand system uncertainty, in brackets, based on 500 bootstrap replications. Heineken products (Heineken and Heineken Light) and observations where $MC^j_{it} < 0$ (less than 1% of full sample) are excluded from these regressions. All regressions include product, city, brewery, and year fixed effects.

Table OA.4 reports the corresponding first-stage regressions. The first row reports coefficients for what we call our "Scale IV"; either the log count of the number of cities each product is shipped to from brewery $b(j,c,t)$ or the average of a demand shock across all those cities. The first three columns report the first stage regressions on $\ln Q^j_{ct} - \ln S^{c|j}_{b(j,c,t)t}$, which we refer to as the "Scale" variable since it identifies returns to scale. Since $\ln Q^j_{ct} - \ln S^{c|j}_{b(j,c,t)t} = \ln\left(\sum_{c'\in\mathbb{C}^j_{b(j,c,t)t}} \frac{Q^j_{c't}}{\Omega^j_{c't}}\right) - \ln(\Omega^j_{ct})$, we expect our positive scale shifters—i.e., the number of cities and the residual demand shock ξ^j_{ct}—to be positively correlated with the scale, while our negative demand shifter—the differentiation IV based on calories, which acts as a negative demand shifter by increasing the degree of competition product j faces in market (c,t)—should be negatively correlated with scale. Reassuringly, we find both instruments shift scale in the way we would expect. In the second row, we report coefficients for our

"Scope IV", which is either counts of the number of city–products $k \neq j$ brewed in brewery $b(j, c, t)$, or averages of demand shifters for all other city–products $k \neq j$ produced in the same brewery. Columns (4) through (6) report first-stage regressions for $\ln S^{j}_{b(j,c,t)t}$, which we refer to as the "Scope" variable as it identifies the degree of scope economies through $\frac{\phi-\alpha}{\phi}$, and will tend to decrease as the scale of other product lines within the same brewery rise. Again, we find the expected signs for our instruments, with the positive demand shifters—city–product counts as well as mean ξ^{k}_{ct}— decreasing the scope variable, while the negative demand shifter generated by our differentiation IV increases the scope variable. These results reassure us that the IV specifications are identifying the scale and scope parameters by relying on the type of exogeneous variation we expected.[54]

[54]While one might also interpret the Scope IV for the Scale first stages or the Scale IV for the Scope regressions, correctly predicting these signs will generally depend on how the scale of other product lines k interact with the scale of product line j in equilibrium. In general, this will depend on the magnitude of scale and scope economies, as well as the degree of substitutability across product lines, and is therefore hard to predict *ex-ante*. Note, however, that we emphasized in Section 4.3 that the scope instruments should be understood as variables that shift scope *conditional* of the scale of j, which is consistent *any* pattern of correlation between the Scope IV and Scale, or the Scale IVs and Scope.

E Alternative allocations of products to breweries

In the main text, we assumed that each product is produced at the closest brewery as in Miller & Weinberg (2017). In practice, however, the existence of scale and scope economies can lead firms to want to consolidate production in a way that does not actually involve sourcing production from the closest brewery. While we consider exactly these type of reallocations in our counterfactual experiment, it is worth noting that if our assumed product and shipping sets at the brewery level are incorrectly specified, we may obtain biased estimates of scale and scope economies, due to mismeasurement of the production sets.[55]

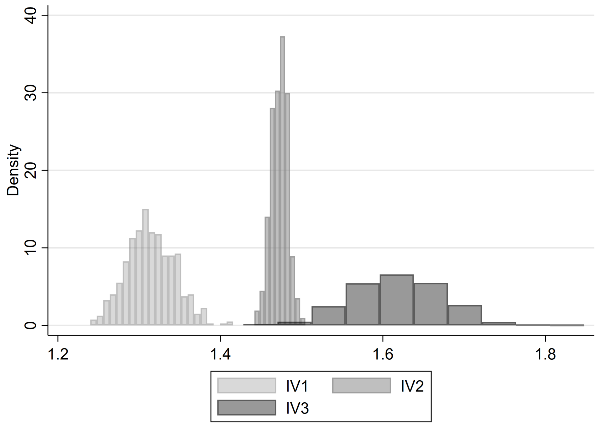
To get a sense of how much this might matter, we instead re-estimate our model by randomly allocating a subset of the products to the second closest brewery that is owned by the same firm with 1000 miles.[56] The key idea behind this exercise is to determine how sensitive our estimates are to potentially mismeasuring our product sets. Note, however, that since transportation costs are well known to be an important part of this industry, we assume that mismeasurement on this front is likely to be towards a not too distant brewery. Figure OA.2 reports histograms of the estimated cost function parameters for 500 random reallocations of products to breweries for each set of instruments used to identify the cost function. All permutations generate point estimates consistent with increasing returns to scale and economies of scope, and are generally centered around our baseline estimates. This leads us to conclude that allocating products to the closest brewery is unlikely to be driving our final conclusions.

[55] We expect smaller trade-offs between scale, scope, and shipping costs in the original pre-merger allocation, as that allocation likely reflects a long-term equilibrium where brewery location and size decisions were endogenous, allowing companies to position larger breweries near key markets.

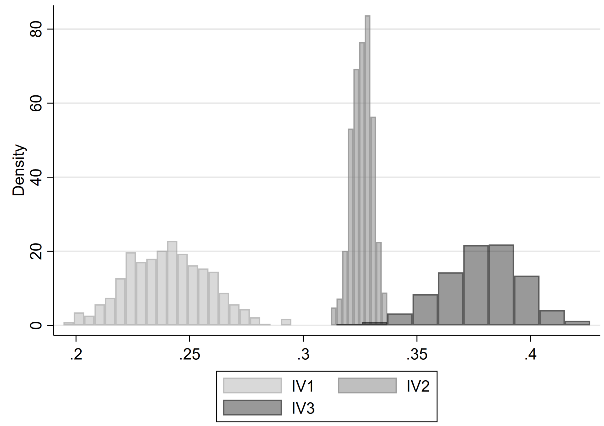
[56] Reallocations are done for all products sold by the same firm in each city, and the likelihood of producing at the closest, or second closest brewery, is assumed to be equal.

Figure OA.2: Cost Function Estimates: Randomizing to Second Closest Brewery

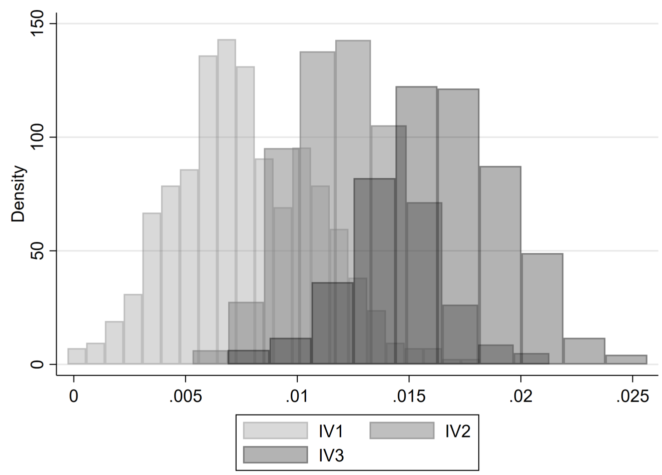
(a) Returns to Scale (ϕ)



(b) Scope $\left(\frac{\phi-\alpha}{\phi}\right)$



(c) Transportation costs (λ)



Notes: The above histograms plot histograms of returns to scale (ϕ) in panel (a), scope $\left(\frac{\phi-\alpha}{\phi}\right)$ in panel (b), and transportation costs (λ), for 500 iterations of our cost estimation routine. Each iteration re-estimates (18) with the same set of instruments used in Table 2, after randomly allocating a subset of city-firm production sets to the *second* closest brewery within 1000 miles. IV1 refers to the estimates based on the instruments used in column 2 of Table 2, IV2 refers to the instruments used in column 3, and IV3 refers to the instruments in column 4. The probability a given firm-product pair is produced at the closest, or second closest, brewery owned by that firm, is 50-50.

F Other Cost Function Objects

With additional assumptions on the structure of productivity terms introduced in equation (14), our regression specification allows us to estimate all the other necessary components of the cost function. Specifically, we assume that the productivity term $\ln A^j_{b(j,c,t)t}$ is a sum of product, location, and time-specific shocks: $\ln A^j_{b(j,c,t)t} = \ln A^j + \ln A_{b(j,c,t)} + \ln A_t + \widetilde{A}^j_{b(j,c,t)t}$. Additionally, we assume that the unit input cost function can be decomposed into brewery and time-specific components: $g(\mathbf{W}_{b(j,c,t)t}) = G_{b(j,c,t)} \times g(\mathbf{W}_t)$. Note that given our cost structure (see Equation 1), all the brewery and time-specific components can be aggregated into the augmented unit cost function $g^{\text{aug}}_b(\mathbf{W}_t) = \left(\frac{1}{A_{b(j,c,t)}}^{\frac{1}{\phi}} G_{b(j,c,t)}\right) \times \left(\frac{1}{A_t}^{\frac{1}{\phi}} g(\mathbf{W}_t)\right)$, which remains linked to a specific brewery in any counterfactual product allocation. The product-specific productivity term A^j remains linked to the corresponding product.

To estimate these remaining structural components of the model, we first project the residuals $\widetilde{\omega}^j_{ct} \equiv \ln MC^j_{ct} - \left(\frac{\phi-\alpha}{\phi} \ln S^j_{b(j,c,t)t} + \frac{1-\phi}{\phi} \ln\left(\frac{Q^j_{ct}}{S^{c|j}_{b(j,c,t)t}}\right) + \frac{\lambda}{\phi}\text{dist}_{b(j,c,t)c} \times \text{fuel}_t\right)$ from equation (18) onto product, brewery, and time (year-month) dummy variables. We denote these estimates as $\bar{\gamma}_j$, $\bar{\gamma}_{b(j,c,t)}$, and $\bar{\gamma}_t$, respectively. We denote the residual from this projection as $\bar{\epsilon}^j_{b(j,c,t)ct}$.

Under the assumptions discussed above, equation (14) can be written as

$$\begin{aligned} \ln MC^j_{ct} = {} & \frac{\phi-\alpha}{\phi} \ln S^j_{b(j,c,t)t} + \frac{1-\phi}{\phi}\left(\ln Q^j_{ct} - \ln S^{c|j}_{b(j,c,t)}\right) + \frac{\lambda}{\phi}\text{dist}_{b(j,c,t)c} \times \text{fuel}_t \\ & \underbrace{-\frac{1}{\phi} lnA^j}_{\equiv\bar{\gamma}_j} + \underbrace{\ln\left(\frac{1}{\phi}\frac{1}{A_{b(j,c,t)}}^{\frac{1}{\phi}} G_{b(j,c,t)}\right)}_{\equiv\bar{\gamma}_{b(j,c,t)}} + \underbrace{\ln\left(\frac{1}{A_t}^{\frac{1}{\phi}} g(\mathbf{W}_t)\right)}_{\equiv\bar{\gamma}_t} + \underbrace{\frac{1}{\phi}\ln\widetilde{\tau}^j_{ct} - \frac{1}{\phi}\ln\widetilde{A}^j_{b(j,c,t)t}}_{\equiv\bar{\epsilon}^j_{b(j,c,t)ct}} \end{aligned} \tag{39}$$

Referring to the equation (39) above, we can infer that the product fixed effect is related to the product-level productivity term as follows: $\ln A^j = -\bar{\gamma}_j \times \phi$. Similarly, the brewery fixed effect corresponds to the brewery-specific component of the augmented unit cost function: $\frac{1}{A_{b(j,c,t)}}^{\frac{1}{\phi}} G_{b(j,c,t)} = \exp(\bar{\gamma}_{b(j,c,t)}) \times \phi$, and the time fixed effect corresponds to the time-specific component: $\frac{1}{A_t}^{\frac{1}{\phi}} g(\mathbf{W}_t) = \exp(\bar{\gamma}_t)$. With estimated values of the fixed effects and ϕ, we can compute the values on the left-hand sides of these equations. Finally, the regression residual estimates yield a combination of the unobserved productivity and shipping cost components: $\ln\widetilde{\tau}^j_{ct} - \ln\widetilde{A}^j_{b(j,c,t)t} = \phi \times \bar{\epsilon}^j_{b(j,c,t)ct}$. This allows us to estimate the value of Ω^j_{ct} (after removing the brewery- and time-specific components of productivity, now aggregated in the augmented unit cost function): $\ln\Omega^j_{ct} = \ln A^j - \lambda \times \text{dist}_{b(j,c,t)c} \times \text{fuel}_t - (\ln\widetilde{\tau}^j_{ct} - \ln\widetilde{A}^j_{b(j,c,t)t})$. With estimates of the product fixed effects, regression residual, ϕ, and λ, we can calculate Ω^j_{ct}.

Given this accounting of all terms in the marginal cost function, we remove the impact of the unobserved term $\ln \widetilde{\tau}^j_{ct} - \ln \widetilde{A}^j_{b(j,c,t)t}$ from our counterfactual simulations. In practice, this is done by setting $\ln \widetilde{\tau}^j_{ct} - \ln \widetilde{A}^j_{b(j,c,t)t} = 0$ in the marginal cost function, and then solving for both the pre- and post-merger value of prices and quantities numerically.

G Solving for the approximately optimal allocation

To obtain an approximately optimal allocation, we formulate our problem as a mixed-integer programming problem with both linear and nonlinear constraints. Specifically, we write down the following:

$$
\begin{aligned}
\min_{\mathbf{x}} \quad & \sum_b g_b^{\text{aug}} \times C_b \\
\text{s.t.} \quad & x_{bc} \in \{0,1\} \quad \text{Sets up the binary allocation variables} \\
& \sum_b x_{bc} = 1 \quad \text{A particular city can only be allocated to one brewery} \\
& Q_{bc}^j = Q_c^j \times x_{bc} \quad \forall\, b, c, j \quad \text{Demand from a product-city pair is allocated according to } x_{b,c} \\
& Y_{bc}^j = Q_{bc}^j \times \tau_{bc} \quad \forall\, b, c, j \quad \text{Translates demand into output, } \tau_{bc} = \lambda \times \text{dist}_{bc} \times \text{fuel} \\
& \sum_{c,j} Y_{bc}^j \leq \bar{Y}_b \quad \forall\, b \quad \text{Brewery capacity constraints} \\
& \sum_c \frac{Y_{bc}^j}{A^j} = \Upsilon_b^j \quad \forall\, b, j \quad \text{Defines auxiliary variables} \\
& \left(\Upsilon_b^j\right)^{\frac{1}{\alpha}} = U_b^j \quad \forall\, b, j \quad \text{First nonlinear constraint} \\
& \sum_j U_b^j = Z_b \quad \forall\, b \quad \text{Defines auxiliary variables} \\
& (Z_b)^{\frac{\alpha}{\phi}} = C_b \quad \forall\, b \quad \text{Second nonlinear constraint,}
\end{aligned}
$$

where g_b^{aug} in the objective function is a brewery-specific constant, see Online Appendix F for details on how it is defined.

Finally, we include a set of distance-based constraints that set certain x_{bc} to 1 or 0. Specifically, we impose constraints on the distance products may travel from breweries to cities, capping it at approximately the distance from Irwindale, California, to Chicago. These constraints are implemented by setting $x_{bc} = 0$ whenever the distance between b and c exceeds this cutoff. Additionally, if a brewery is located within the city or its suburbs (defined as within 50 miles of the city center), we assume that the products sold in that city are produced at that brewery. This is implemented by setting $x_{bc} = 1$ whenever the distance between b and c is less than 50 miles.

We then pass this formulation to the solver. Gurobi, primarily a Mixed Integer Linear Programming (MILP) solver, handles most nonlinear constraints by converting them into a set of piecewise-linear constraints. For example, consider the first nonlinear constraint $(\Upsilon)^{\frac{1}{\alpha}} = U$. The solver approximates this constraint by selecting a set of points

$(\Upsilon_1, U_1), (\Upsilon_2, U_2), ..., (\Upsilon_n, U_n)$ and generating a corresponding set of linear constraints. The solver then ensures that the variables satisfy these constraints within their respective ranges. Figure OA.3 provides an example of a piecewise-linear approximation of a concave function. In our specific application, the solver constructs the piecewise-linear approximation such that the relative error (e.g., the maximum relative difference between the original nonlinear function and its piecewise approximation) does not exceed 0.001.

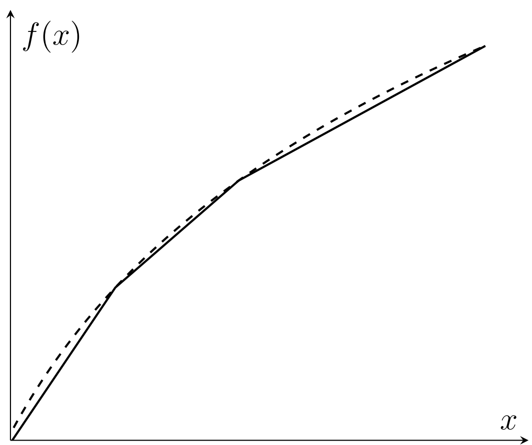


Figure OA.3: An example of a piecewise-linear approximation (solid line) of an arbitrary concave function $f(x)$ (dashed line).

The solver then proceeds to find the optimal solution using the latest tools from the mixed integer programming theory. The backbone algorithm that it uses is called branch-and-bound. The method involves solving computationally easy continuous relaxations of the original problem, where integer constraints are temporarily removed. It then generates subproblems (branches) by adding inequality constraints that exclude the optimal fractional values found in the relaxation.[57] Note that solving continuous (unconstrained) relaxations improves efficiency: if a solution to relaxation in one branch is worse than an integer solution found in another, the entire former branch can be discarded without the need to explore integer solutions further down that branch. More information on this method can be found here: `www.gurobi.com/resources/mixed-integer-programming-mip-a-primer-on-the-basics`.

[57]For instance, if a variable x has a relaxed solution of 5.7, two branches will be created: one with $x \leq 5$ and another with $x \geq 6$.

H S^j_{bt} is the share of rival inputs allocated to j

Define $S^j_{bt} \equiv \frac{\sum_{c\in\mathbb{C}^j_{bt}} MC^j_{ct}Q^j_{ct}}{\sum_{k\in\mathbb{J}_{bt}}\sum_{c\in\mathbb{C}^k_{bt}} MC^k_{ct}Q^k_{ct}}$ as in the main text. In this appendix, we now show that this share is equal to the share of inputs allocated to the rival task for good j, i.e. $S^j_{bt} = \frac{X^{rj}_{bt}}{\sum_{k\in\mathbb{J}_{bt}} X^{rk}_{bt}} = \frac{X^{rj}_{bt}}{X^r_{bt}}\ \forall X$, a generalization of the input share inversion result in Orr (2022) to joint production settings.

The result can be obtained by noting that the relevant rival input shares were already determined when we characterized a firm's output distance function in Appendix A; specifically, equation (32) tells us that $\frac{X^{rj}_{bt}}{X^r_{bt}} = \frac{\lambda^j_{bt}Y^j_{bt}}{\sum_k \lambda^k_i Y^k_{bt}}$, with the difference that we now include a time subindex. We can obtain the desired result by showing that $\lambda^j_{bt}Y^j_{bt}$ is proportional to $\sum_{c\in\mathbb{C}^j_{bt}} MC^j_{ct}Q^j_{ct}$; this can be done by applying the envelope theorem to a firm's cost minimization problem, as well as the output distance function problem.

First, note that a firm's cost function can be recovered from the following cost minimization problem:

$$\begin{aligned} C\left(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt}\right) \equiv & \min_{\mathbf{X}_{bt}} \sum_X W_{btX}X_{bt} \\ & \text{s.t.: } D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt}) \leq 1 \end{aligned} \tag{40}$$

where $D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})$ is the output distance function corresponding to the cost function used in the main text.[58]

This problem has the following Lagrangian:

$$L = \sum_X W_{btX}X_{bt} + \theta_{bt}(D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{btX},\boldsymbol{\tau}_{bt}) - 1)$$

From the envelope theorem, it follows that:

$$MC^j_{ct} \equiv \frac{\partial C(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})}{\partial Q^j_{ct}} = \theta_{bt}\frac{\partial D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})}{\partial Q^j_{ct}} \tag{41}$$

From Appendix B, we know that the cost function used in the main text has an output distance function defined by (40). Applying the envelope theorem to its associated

[58]Note that we have written the output distance function in terms of quantities sold in each market c that are produced by brewery b, $\mathbf{Q}_{bt}$. Since $D\left(\mathbf{Y}_{bt},\mathbf{X}_{bt},\mathbf{A}_{bt}\right) = \frac{\left(\sum_j\left(\frac{Y^j_{bt}}{A^j_{bt}}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{\prod_X (X_{bt})^{\beta_X}}$ from Appendix A, this becomes $D\left(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt}\right) = \frac{\left(\sum_j\left(\frac{\sum_c Q^j_{ct}\tau^j_{ct}}{A^j_{bt}}\right)^{\frac{1}{\alpha}}\right)^{\alpha}}{\prod_X (X_{bt})^{\beta_X}}$ once we replace aggregate factory-level outputs with market-specific sales through the iceberg transportation constraint (3).

Lagrangian (equation 23) yields:

$$\frac{\partial D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})}{\partial Q^j_{ct}} = \frac{\partial D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})}{\partial Y^j_{bt}}\frac{\partial Y^j_{bt}}{\partial Q^j_{ct}} = \frac{\lambda^j_{bt}}{\delta}\tau^j_{ct} = \lambda^j_{bt}\tau^j_{ct} \tag{42}$$

where the third equality uses (3), and the fourth equality uses the fact that $\delta = 1$ when firms cost minimize.

Note that (41) and (42) together imply that:

$$MC^j_{ct}Q^j_{ct} = \theta_{bt}\frac{\partial D(\mathbf{Q}_{bt},\mathbf{A}_{bt},\mathbf{W}_{bt},\boldsymbol{\tau}_{bt})}{\partial Q^j_{ct}}Q^j_{ct} = \theta_{bt}\tau^j_{ct}\lambda^j_{bt}Q^j_{ct} \tag{43}$$

Summing over all $c \in \mathbb{C}^j_{bt}$ then yields:

$$\sum_{c\in\mathbb{C}^j_{bt}} MC^j_{ct}Q^j_{ct} = \theta_{bt}\lambda^j_{bt}\sum_{c\in\mathbb{C}^j_{bt}}\tau^j_{ct}Q^j_{ct} = \theta_{bt}\lambda^j_{bt}Y^j_{bt} \tag{44}$$

where the last equality follows from (3).

Substituting equation (44) into (32) then yields:

$$\frac{X^{rj}_{bt}}{X^r_{bt}} = \frac{\lambda^j_{bt}Y^j_{bt}}{\sum_{k\in\mathbb{J}_{bt}}\lambda^k_{bt}Y^k_{bt}} = \frac{\frac{\sum_{c\in\mathbb{C}^j_{bt}} MC^j_{ct}Q^j_{ct}}{\theta_{bt}}}{\sum_{k\in\mathbb{J}_{bt}}\frac{\sum_{c\in\mathbb{C}^k_{bt}} MC^k_{ct}Q^k_{ct}}{\theta_{bt}}} = \frac{\sum_{c\in\mathbb{C}^j_{bt}} MC^j_{ct}Q^j_{ct}}{\sum_{k\in\mathbb{J}_{bt}}\sum_{c\in\mathbb{C}^k_{bt}} MC^k_{ct}Q^k_{ct}} = S^j_{bt} \tag{45}$$

Intuitively, this result comes from our assumption that production functions are homogeneous of degree $\phi > 0$ with no differences in factor intensities across product lines. Since factor intensities do not differ across product lines, we no longer need to worry about input shares by a factor but rather can simply track a single aggregate share of rival inputs by product line. This makes the overall production technology behave "as if" we had a single composite input.

When production is analogous to a setting with a single composite input, this means input shares must vary proportionally to output over product-line productivity. This implies that we can only rationalize observed output levels if rival input shares satisfy $\frac{\left(\frac{Y^j_{bt}}{A^j_{bt}}\right)^{\frac{1}{\alpha}}}{\sum_k\left(\frac{Y^k_{bt}}{A^k_{bt}}\right)^{\frac{1}{\alpha}}}$.[59] The duality results we have shown in this appendix simply establish that variation in $Y^j_{bt}MC^j_{bt}$

[59]More formally, if we have a rival input that we can use to rationalize different production levels across product lines, our model simplifies to $Y^j_{bt} = (S^{jr}_{bt})^\alpha A^j_{bt}G_{bt}$, where $G_{bt}(.)$ is a brewery-specific scalar capturing the production technology and non-rival inputs to production, and α is total returns to scale in rival inputs.

within a brewery is proportional to $\left(\frac{Y_{bt}^{j}}{A_{bt}^{j}}\right)^{\frac{1}{\alpha}}$. This is because within-brewery marginal costs will change as we scale up and down the quantity of rival inputs used in a particular product line. Since increases in rival inputs change marginal costs proportionally to returns to scale in rival inputs and the productivity level of product-line j, $Y_{bt}^{j} MC_{bt}^{j} \propto \left(\frac{Y_{bt}^{j}}{A_{bt}^{j}}\right)^{\frac{1}{\alpha}}$ within a brewery.

Since rearranging this expression yields $S_{bt}^{jr} = \left(\frac{Y_{bt}^{j}}{A_{bt}^{j} G_{bt}}\right)^{\frac{1}{\alpha}}$ and $\sum_j S_{bt}^{jr} = 1$ it must be that $S_{bt}^{jr} = \frac{\left(\frac{Y_{bt}^{j}}{A_{bt}^{j}}\right)^{\frac{1}{\alpha}}}{\sum_k \left(\frac{Y_{bt}^{k}}{A_{bt}^{k}}\right)^{\frac{1}{\alpha}}}$