Online Appendix: Not For Publication

Identifying Scale and Scope Economies using Product Market Data

Ekaterina Khmelnitskaya, Guillermo Marshall, and Scott Orr

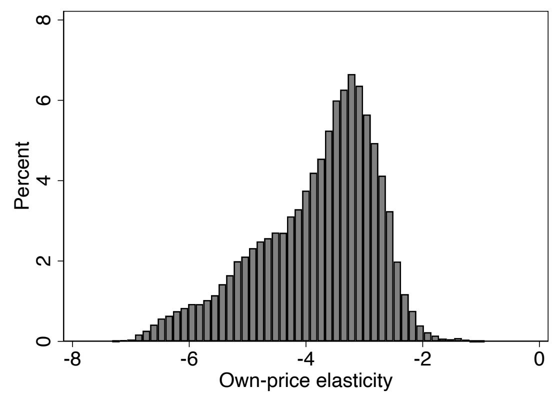
C Additional Tables and Figures

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]

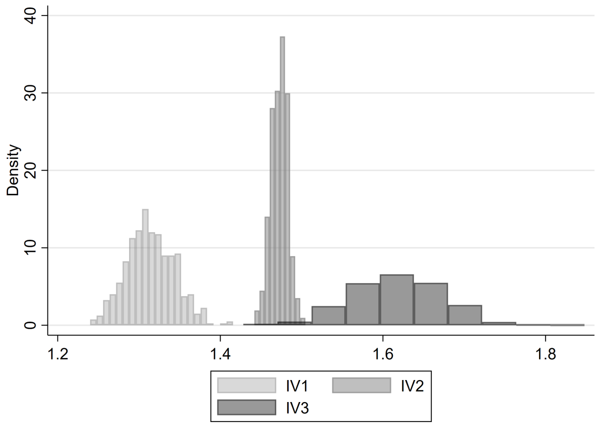
To get a sense of how much this might matter, we instead re-estimate our model by randomly allocating a subset of the products to the second closest brewery that is owned by the same firm with 1000 miles.[56] The key idea behind this exercise is to determine how sensitive our estimates are to potentially 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.

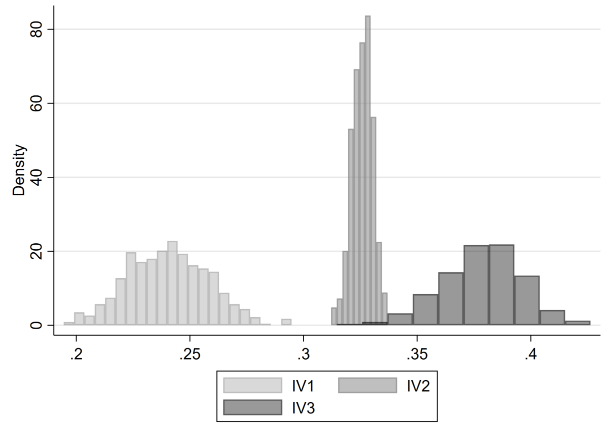
[56] Reallocations are done for all products sold by the same firm in each city, and the likelihood of producing at the closest, or second closest brewery, is assumed to be equal.

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

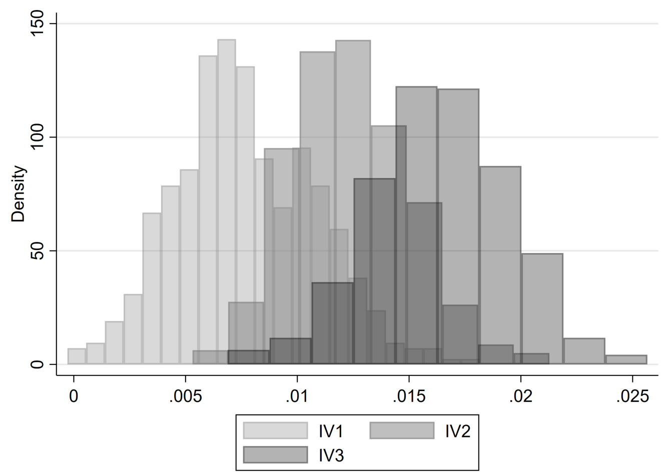
(a) Returns to Scale (ϕ)



(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.

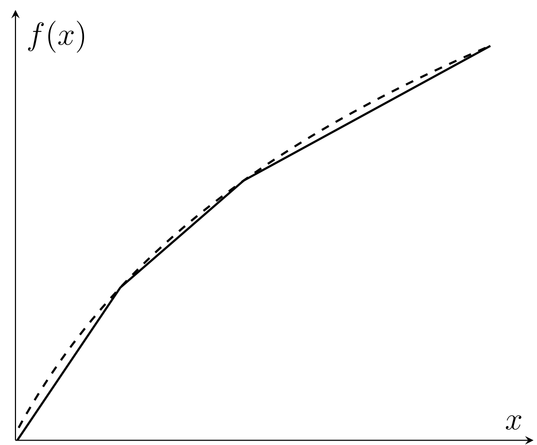


Figure OA.3: An example of a piecewise-linear approximation (solid line) of an arbitrary concave function $f(x)$ (dashed line).

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 $S_{bt}^{jr} = \frac{\left(\frac{Y_{bt}^j}{A_{bt}^j}\right)^{\frac{1}{\alpha}}}{\sum_k \left(\frac{Y_{bt}^k}{A_{bt}^k}\right)^{\frac{1}{\alpha}}}$