Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

State-Level Beer Production Patterns in the U.S.: The Import of Legacy and Agglomeration Economies

Mitch Kunce[1]

Abstract
This paper considers U.S. beer industry production patterns in light of two leading Economic Geography principles – historical advantages (legacy) and increasing returns. Both theoretically and empirically we establish that historical first moves and subsequent agglomeration forces are dominate influences. Empirical results, using unique panel data for U.S. states, support the importance of legacy in establishing the spatial pattern of beer production and that increasing returns help determine the size of location dominance. Predictably, the unprecedented small-beer expansion, over the last two decades, has produced little effect on production variation across states.

Keywords: Legacy, agglomeration, path dependency, beer industry
JEL Classifications: N60, R12, R13

Introduction
The gold rush of the mid-1800s brought brewing to the front range of the western Kansas Territory Rocky Mountains. In 1859, Rocky Mountain Brewery opened in the Auraria neighborhood of Denver (in today's Denver metro, around Tenth and Larimer streets). Rocky Mountain Brewery is considered by area historians as the first commercial brewery in the eventual Colorado Territory.[2] German-born entrepreneur, Philip Zang bought the growing brewery in 1871 and renamed it, Philip Zang & Company. In 1864, also in Auraria, the beginnings of the Tivoli Brewery were developed by Denver councilman Moritz Sigi. At the time, Zang's and the Tivoli were the largest breweries in the front range. In 1873, Jacob Schueler and Adolph Coors opened the Schueler & Coors Golden Brewery on the banks of Cannonball Creek (now Clear Creek) in Golden City, Colorado Territory. Coors acquired sole interest in the company in 1880 (Baum, 2000). Nearer to the turn of the century, P.H. Schnieder Brewing in Trinidad and Walter Brewing in Pueblo were established. Interestingly, only Coors, Schnieder, Tivoli and Walter survived the Colorado 1916 prohibition along with the national Volstead Act of 1920.[3] Today, the Molson Coors Golden, Colorado legacy brewery is the largest single brewing site in the world.

Of course, this brewing legacy story is not unique to the front range of Colorado. Milwaukee, Wisconsin – with breweries including Owens, Best, Pabst and Schlitz all established in the mid 1840s. Notable brands like Blatz and Miller followed in the 1850s.

---

[1] *Douglas**Mitchell*** Econometric Consulting, Laramie, WY USA. Published in May 2026.
[2] The Colorado Territory was formed in 1861 with Golden City as the capitol. The chief source of this brief history lesson: https://www.historycolorado.org.
[3] Accounts of how many Colorado breweries were in operation before prohibition vary. Some say 20 plus, others cite 40 or so.

Figure 1. Schueler & Coors advertising

Figure 2. The legacy of Coors 2025 advertisement

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

Saint Louis, Missouri – where Lemp and Griesedieck were prominent breweries in the 1840s with Anheuser-Busch opening in 1852. These regional 'first moves' or what New Economic Geography [4] (NEG) coin as "historical advantages" shape the motivation for this 'legacy' take on why variation in U.S. beer production today can be explained by fixed attributes of a handful of states. Our use of the term legacy refers to a level of production agglomeration inherited from the past. Once created, these legacy economies impact the paths of large capital investment, production innovation and productivity growth (David, 1994; Martin and Sunley, 2006). A parallel literature exists in political science showing that legacy political economy continues to influence contemporary outcomes (Charnysh et al., 2023).

This paper aims to explore, both theoretically and empirically, legacy impacts on state-level (jurisdictional) beer production patterns in the U.S. In the next section we build theoretical underpinnings, in a general equilibrium setting, showing how existing agglomeration forces and subsequent integration strength are a dominate influence on jurisdictional production. The third section describes the unique state-level panel data used to show how variation in beer production, of late, is largely explained by fixed or latent attributes of a handful of states. The fourth section presents, with discussion and caveat, empirical results. The last section, of course, provides concluding remarks.

Theoretical underpinnings

We build on the seminal jurisdictional general equilibrium modeling of Zodrow and Mieszkowski (1986), Wilson (1986) and Oates and Schwab (1991) in the presence of agglomeration economies as adapted by Garcia-Milà and McGuire (2001), Fernández (2005) and Krogstrup (2008). The economy modeled is divided into J ($j = 1, . . . , J$) symmetric jurisdictions [5] where historical integration forces are present. A single industry exists within each jurisdiction comprised of a number of identical competitive firms, N_j. Jurisdictions are 'small' enough that their policy dealings are deemed to be 'competitive', yet 'large' enough that the population lives and works within the jurisdiction. Industry output, the numeraire, is produced using the local population (labor), L_j, perfectly mobile capital, K_j, and a public input of the unpaid type, X_j.[6] The public input, X_j, represents an aggregate level of productive factors provided by jurisdictional governments to firms. The publicly provided input is subject to the same degree of rivalry in use as other factors. Following Oates and Schwab (1991), governments ration X_j to firms in direct proportion to their capital stock. Competitive firms view the public input as a given.

In the broader economy, the overall capital stock $\overline{K}$ is fixed with the model concentrating on mobility rather than formation and entry, $\overline{K} = \Sigma_j K_j$.[7] Jurisdictional production is represented as,

[4] See Krugman (1991a; 1991b) and Baldwin and Krugman (2004).
[5] Symmetry avoids Tiebout-type inefficiencies in the model.
[6] See Feehan (1989) for a classification and examples of public inputs to production.
[7] Capital formation and entry modeling adds a layer of unnecessary complexity that ultimately leads to comparable results (Wilson, 1999).

$$F(L_j, K_j, X_j)A(K_j), \tag{1}$$

with each firm in the jurisdiction producing $[1/N_j \cdot F(L_j,K_j,X_j)]A(K_j)$.[8] The last term, $A(K_j)$, denotes a multiplicative augmentation to productivity associated with accumulated levels of capital.[9] This is a simple yet effectual way to model agglomeration. Particularly, it allows for summary measures of the strength of integration forces given the stock of capital invested in the jurisdiction. Initially defining $A(K_j) > 1$ denotes a type of existing external economies that will increase in K_j, $\partial A/\partial K_j > 0$. Profit maximizing firms treat the local aggregate amount of capital invested as a given.

Assuming $F(L_j,K_j,X_j)$ is linearly homogeneous in all inputs, using Euler's theorem, production can be written as,

$$\begin{gathered} F(L,K,X)A(K) = [F_L(L,K,X)L + F_K(L,K,X)K + F_X(L,K,X)X]A(K) \\ FA = F_L LA + F_K KA + F_X XA \end{gathered}, \tag{2}$$

where subscripts now represent partial derivatives and jurisdictional indexing going forward will be suppressed, though most expressions are understood to be jurisdiction specific. A single firm operates under constant returns to scale, though its capital employment has productivity effects on all firms in the jurisdiction. Marginal products of $F(L,K,X)$ are all positive and diminish F_{LL}, F_{KK}, $F_{XX} < 0$. Strict concavity of $F(L,K,X)$ requires,

$$F_{LL}F_{KK} - F_{LK}^2 > 0, \tag{3}$$

and,

$$2F_{LK}F_{KX}F_{LX} - F_{LL}F_{KX}^2 - F_{KK}F_{LX}^2 + F_{XX}\left(F_{LL}F_{KK} - F_{LK}^2\right) < 0. \tag{4}$$

The first term in Equation (4) is ambiguous in sign. Of the three cross-partials in this term, the technical relationship between labor and the pubic input is the least obvious.[10] Differentiating $F(L,K,X)$ when set equal to the square-bracketed part of Equation (2) with respect to $\Phi = \{L, K, X\}$ yields,

$$LF_{L\Phi} + KF_{K\Phi} + XF_{X\Phi} = 0, \tag{5}$$

reinforcing that all inputs can be technical complements.

As shown in Equation (2), public inputs generate rents of $F_X XA$. Following Oates and Schwab (1991), rent rationing is defined as,

[8] See Appendix A to this paper for an analogous production per-capita model presentation. The empirical examination below relies on population adjusted state-level variables.
[9] See Fernández (2005) and Krogstrup (2008) for a similar augmentation.
[10] Young's Theorem proves invariance to the order of differentiation, for example, $F_{LX} = F_{XL}$

$$R = \frac{X}{K} F_X A. \tag{6}$$

This rationing mechanism signifies that in addition to the marginal product of capital, public input rents are added to the net return to capital. The condition for profit maximizing capital investment becomes,

$$r = F_K A + R - t, \tag{7}$$

where t is a source-based unit tax. Profit maximizing forces distribute capital until r, the net return to capital, is equalized across all jurisdictions. Competitive firms and 'small' jurisdictions view r as exogenous.

Firms pay each resident (laborer) the marginal product of labor less a free-to-vary local public good tax, τ,

$$w = F_L A - \tau. \tag{8}$$

The jurisdiction's governmental budget constraint for public goods and inputs is defined,

$$G + X = \tau L + tK, \tag{9}$$

where prices of public goods and inputs are normalized to one. Using Equations (2) and (6) through (9) the wage equation can be rewritten as,

$$\begin{aligned} &FA - (w+\tau)L - (r+t)K = 0, \\ &FA - wL - rK - G - X = 0, \\ &w = \frac{1}{L}(FA - rK - G - X). \end{aligned} \tag{10}$$

Jurisdictional income for consumption then follows,

$$C = wL + Y = FA - rK - G - X + Y, \tag{11}$$

where the far-right term, Y, denotes an aggregate amount of exogenous income endowed to residents. Identical residents of a jurisdiction receive utility from consumption and local public goods. Jurisdictional utility takes the form, $U(C,G)$, where U is quasi-concave with marginal utilities U_C , $U_G > 0$. In keeping with the Arrow-Debreu separation assumption for general equilibrium constructs, residents have two distinct roles in the model (Wilson 1999). First, as consumers, they seek to maximize utility over a bundle of goods and public services. Second, they inelastically supply labor input to production and in return receive earned income for consumption. Following Oates and Schwab (1991), the public inputs to firms do not enter resident's utility directly.

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

Jurisdictional authorities, recognizing the presence of external economies, optimize over a policy variable vector $\theta = \{\tau, t, X\}$ that maximizes jurisdictional utility subject to Equations (9) and (11). First order conditions become,

$$\frac{U_G}{U_C}\frac{\partial G}{\partial \theta} + \frac{\partial C}{\partial \theta} = 0, \tag{12}$$

where using Equation (9),

$$\frac{\partial G}{\partial \theta} = L\frac{\partial \tau}{\partial \theta} + t\frac{\partial K}{\partial \theta} + K\frac{\partial t}{\partial \theta} - \frac{\partial X}{\partial \theta}, \tag{13}$$

noting that L is jurisdictionally fixed. Accordingly, using Equation (11),

$$\frac{\partial C}{\partial \theta} = FA_K\frac{\partial K}{\partial \theta} + AF_K\frac{\partial K}{\partial \theta} - r\frac{\partial K}{\partial \theta} + AF_X\frac{\partial X}{\partial \theta} - \frac{\partial G}{\partial \theta} - \frac{\partial X}{\partial \theta}, \tag{14}$$

or by substituting in Equation (7),

$$\frac{\partial C}{\partial \theta} = FA_K\frac{\partial K}{\partial \theta} + t\frac{\partial K}{\partial \theta} - R\frac{\partial K}{\partial \theta} + AF_X\frac{\partial X}{\partial \theta} - \frac{\partial G}{\partial \theta} - \frac{\partial X}{\partial \theta}. \tag{15}$$

Substituting Equations (13) and (15) into Equation (12) yields a more complete set of first order conditions when evaluated for each policy element in θ,

$$\theta = \tau : \quad \frac{U_G}{U_C}L - L = 0. \tag{16}$$

$$\theta = t : \quad \frac{U_G}{U_C}\left(t\frac{\partial K}{\partial t} + K\right) + FA_K\frac{\partial K}{\partial t} - R\frac{\partial K}{\partial t} - K = 0, \tag{17}$$

$$\theta = X : \frac{U_G}{U_C}\left(t\frac{\partial K}{\partial X} - 1\right) + FA_K\frac{\partial K}{\partial X} - R\frac{\partial K}{\partial X} + F_X A = 0. \tag{18}$$

Solving Equations (16) through (18) simultaneously yields the optimal conditions,

$$\frac{U_G}{U_C} = 1, \tag{19}$$

$$t = R - FA_K, \tag{20}$$

$$F_X A = 1. \tag{21}$$

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

Equation (19) represents the familiar 'Samuelson condition' for the efficient provision of local public goods (Wilson 1986; 1999). This optimality condition suggests that the marginal rate of substitution between the public good and consumption equals the marginal cost of providing an incremental increase in the public good. Given how G enters Equation (11), the marginal rate of transformation in this context is one for one. Equation (20) shows the optimal capital tax rule. The capital tax is equal to the marginal benefit to firms from public inputs less the marginal agglomeration benefit of adding capital to a jurisdiction. External economies directly affect t in this optimal condition. Lastly, Equation (21) simply equates the marginal product of public inputs to the normalized price of one.

We can obtain a socially optimal level of capital rule if we use Equation (14) instead of (15) in the set of first order conditions where,

$$r = AF_K + FA_K \,. \qquad (22)$$

Moreover, using Equations (6) and (21), Equation (20) becomes,

$$t = \frac{X}{K} - FA_K \,. \qquad (23)$$

From Equation (22), capital investment is greatest when A is present and recognized by local authorities. Accordingly, the optimal capital tax in Equation (23) is lowest. For comparison, Oates and Schwab (1991) ignore agglomeration effects and find $r = F_K$, $t = R$.

Substituting Equation (23) into (9) yields,

$$\tau L = G + \left(FA_K\right)K \,, \qquad (24)$$

which discloses that the local population pays FA_K per capital unit to subsidize jurisdictional capital enhancement. It is optimal to do so because their welfare will be higher. When A and its effects are not taken into account, the resulting general equilibrium is a restricted solution to the more general problem that internalizes the externality. The existence of A and its many impacts, in this stylized setting, increases jurisdictional production thereby making locals better off. In the next section we describe the unique state-level data used to empirically examine the influences of A on brewing industry production.

Data

Estimation of implicit partial derivatives of the indirect form of Equation (1) establishes the starting point for the empirical model specification. Jurisdictional, brewing industry production is modeled as a function of observed and latent inputs, local fiscal policy, general economic conditions and agglomeration impacts. Data comprise a 13 year (2008-2020), all 50 state and District of Columbia panel. Table 1 below describes all variables, provides sources and key descriptive statistics. State-level industry production, the dependent variable, is modeled in per capita form. In March of 2021, the Department of the Treasury, Alcohol and Tobacco Tax and Trade Bureau (TTB) compiled a rare public report showing taxed removals (in 31 gallon barrels) by reporting breweries in each state for the years 2008-2020. Providing this level of detail is not a common practice of the TTB due to the possibility of taxpayer disclosure. Figure 3 below shows how taxed production relates to overall beer production, in aggregate, for the U.S. The wedge between the two series is comprised of, chiefly, excise tax allowances made for foreign exporting and product spoilage. The Pearson correlation between the two series is 0.974.

Table 1. Data descriptions, sources and descriptive statistics

Taxed production. Dependent variable. Aggregate taxable removals, in gallons per capita, from brewery inventories by state and D.C. for the years 2008-2020. Circumstances that allow breweries to remove product from inventory without being taxed include spoilage and exporting. Sources: Department of the Treasury, Alcohol and Tobacco Tax and Trade Bureau, March 31, 2021 State-level Taxed Removals Report. State population series are sourced from the U.S. Census Bureau. More current, state-level, removals data, post 2020, are available but have been censored for certain states by the TTB because of potential taxpayer disclosure. Mean 12.957, STD 24.125. Maximum: Colorado, 122.160. Minimum: West Virginia, 0.122.
Breweries. Brewery counts, per million of state population, as tallied by the Brewers Association for the years 2008-2020. Mean 20.469, STD 19.433. Maximum: Vermont, 69.573. Minimum: Mississippi, 2.618.
Electric rates. Average commercial electric rates by state in 2018 cents per kilowatt hour, 2008-2020. Source: U.S. Energy Information Administration, State Electric Profiles. Mean 10.722, STD 3.703. Maximum: Hawaii, 29.345. Minimum: Idaho, 7.298.
Tax climate. Beer-specific excise tax collections by state in 2018 dollars per capita. Generally, this compilation includes state and local brewery production and specific sales taxes and fees aimed at malt based fermented products. Interestingly, in some states (e.g. Pennsylvania), separate on-brewery-premise sales taxes are imposed. Source: The Brewer's Almanac, published periodically by The Beer Institute; Bureau of Economic Analysis for GDP deflator data. Mean 6.270, STD 5.436. Maximum: Alabama, 23.079. Minimum: Wyoming, 0.447.
Unemployment. Average annual rates for all states in percent of labor force. U.S. Bureau of Labor Statistics, 2008-2020. Mean 6.088, STD 2.290. Maximum: Nevada, 8.838. Minimum: North Dakota, 3.154.
Wages. Mean annual disposable wage in a state in 1,000s of 2018 dollars. Source: U.S. Bureau of Labor Statistics, State Wage Estimates, 2008-2020. Mean 48.189, STD 7.844. Maximum: DC, 83.916. Minimum: Mississippi, 37.777.
Water/Sewer rates. Water and wastewater rates represented for each state in 2018 dollars per 1,000 gallons. Surveys of select U.S. utilities for commercial/industrial water and wastewater charges, 2008-2020. Sources: American Water Works Association; Pacific Northwest National Laboratory working with the U.S. Department of Energy's Federal Energy Management Program. Mean 6.862, STD 2.851. Maximum: Maryland, 16.537. Minimum: Arkansas, 2.808.

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

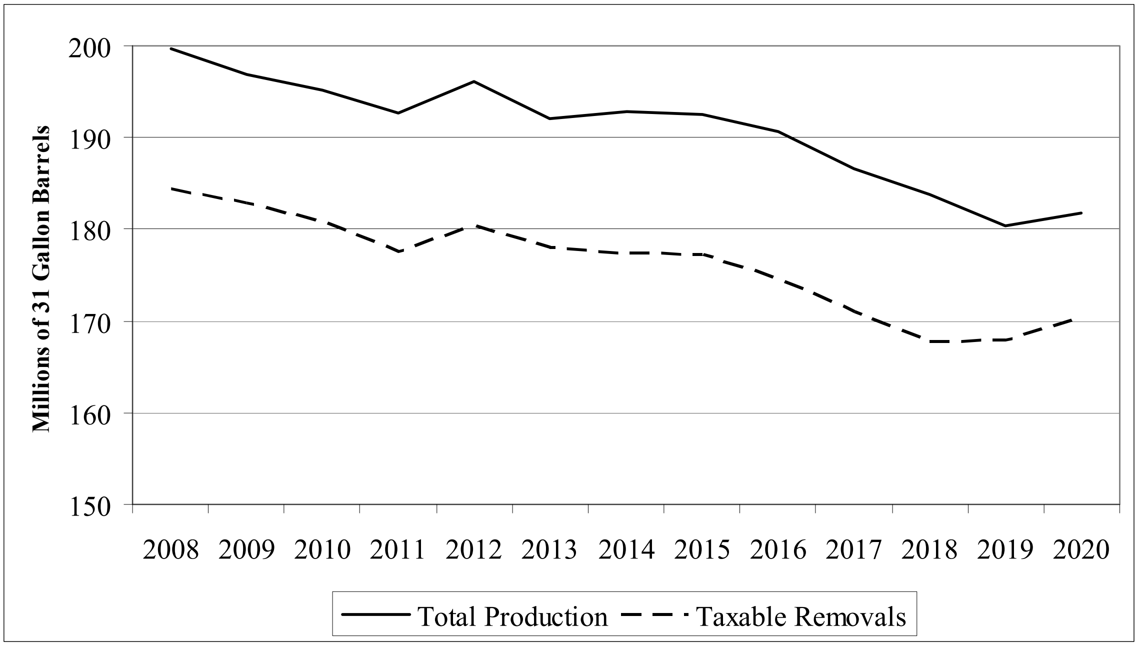


Figure 3. Aggregate U.S. beer production and taxable removals, 2008-2020.

Regarding the right-hand-side (RHS), Table 1 describes state-specific controls for recent industry concentration, key utility costs, industry targeted tax burden, business cycles, labor costs and vital material costs. Within-state variation in these regressors was a key characteristic influencing inclusion. State brewery counts, per one million in population, account for recent industry expansion. As shown in Figure 4, permitted reporting brewery counts, in the U.S., increased by over 4-fold from 2008 to 2020.[11] State beer excise tax collections, per capita, proxy the local tax climate that effects both firms in the

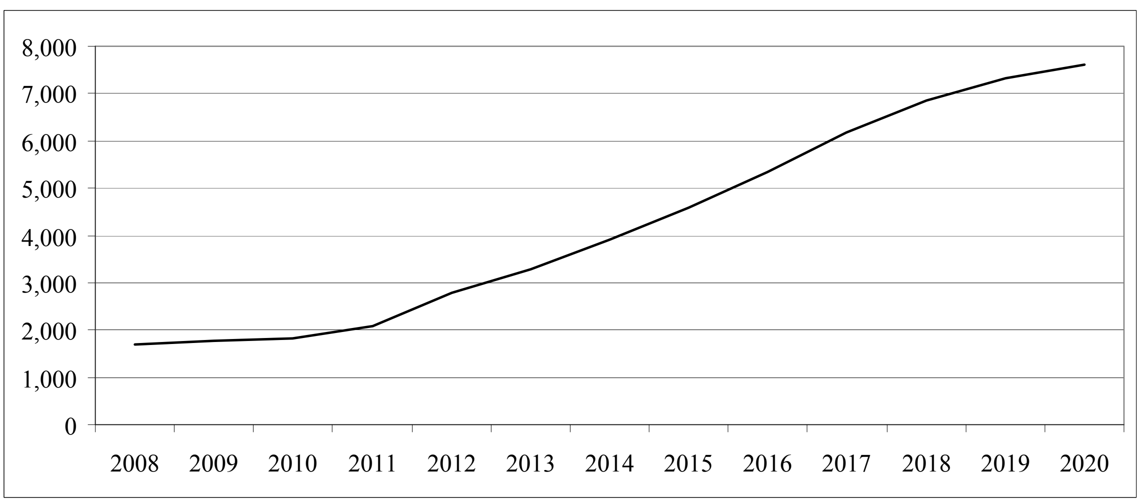


Figure 4. TTB reporting brewery counts for the U.S., 2008-2020.

[11] The late 2000s was the start of the unprecedented growth of small brewing facilities in the U.S. (Elzinga et al., 2015; Kunce, 2023).

industry and local populations. The concept of a state-level 'business cycle' is overly broad and researchers generally focus on key indicators of economic fluctuations. Unemployment is one of (if not) the most important business cycle indicators. For example, Hatzius and Stehn (2012) referred to the unemployment rate as their "desert island economic indicator", the one they would choose if they had to choose only one indicator to provide information about a local economy. Lastly, as one might expect, electric power costs along with commercial/industrial water and wastewater charges impact beer production profoundly. Particularly, water use and discharge is crucial in all stages of the brewing process. In the U.S., it is estimated that, on average, it takes roughly seven gallons of water to produce one gallon of finished beer (Brewers Association, 2013). Moreover, U.S. breweries discharge around 70% of incoming water as effluent. Accordingly, the data show that wastewater marginal costs are much higher than water supply costs for most states.

In addition to the variable particulars denoted in Table 1 above, Table 2 depicts the variable correlation matrix and variance inflation factors for all RHS regressors. While multicollinearity problems are certainly a matter of degree, the risk of detrimental issues within this sample appears small. Lastly, non-stationary variables should not be used in conventional regression models. Regrettably, the time dimension in our panel is too short for proper scrutiny. Unit root and cointegration diagnostics suffer when the variable time series is short (here $T = 13$).

Table 2. RHS correlation matrix and centered variance inflation factors

	Breweries	Electric	Taxes	Wage	Water	VIF
Breweries	1.000					1.20
Electric	0.121	1.000				1.47
Taxes	-0.083	0.295	1.000			1.36
Wage	0.172	0.401	-0.228	1.000		1.88
Water	0.181	0.371	0.083	0.451	1.000	1.58
Unemployment	-0.348	-0.025	-0.028	0.001	0.032	1.15

Empirical model

Consider the conventional two-way time-series and cross-section model,

$$Y_{it} = X_{it}\beta + \varepsilon_{it} \qquad (i = 1, \ldots, N;\ t = 1, \ldots, T;\ n = NT), \tag{25}$$

where Y_{it} is the dependent variable, X_{it} are observable variables that vary across states i and more importantly over years t, β is a k vector of estimated coefficients and ε_{it} denotes the overall error term.[12] The error term is comprised of three components,

$$\varepsilon_{it} = \mu_i + \lambda_t + \nu_{it}, \tag{26}$$

[12] Many presentations of this familiar model may include a scalar constant term, α. All repeated variable representations are meant to be section specific and not conflated with variables laid out in the theoretical underpinnings section above.

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

where μ_i denotes the unobservable state-specific effects, λ_t is the latent year-specific effects and v_{it} is the remainder stochastic disturbance. The component μ_i is time-invariant and will account for state-specific effects not included in the RHS. These effects are expected to pick up any legacy economies. The year-specific effects, λ_t, are state-invariant and will account for any time-specific effects not included in the regression.[13] The remainder disturbance v_{it} varies with groups and time and is assumed orthogonal to X, μ and λ with a mean of zero and a constant variance σ_v^2.

Generally, two specifications of Equation (25) are considered and differ based on their treatment of the specific effects. First, fixed effects (FE) treats μ_i and λ_t as fixed but unknown constants differing across states and time. This specification is easily estimated by including dummy variables in the RHS (Least Squares Dummy Variable (LSDV) estimator). However, as in our case (N = 51, T = 13), LSDV suffers from the loss of precious degrees of freedom. Alternatively, estimates can be obtained by transforming the data into deviations from respective state-means (within estimator) and by including T - 1 year dummies. The two fixed effects estimation methods described reveal two crucial defects: (i) time-invariant variables are picked up by the fixed effect, and (ii) the estimator is not fully efficient because, in certain cases, it ignores variation across states and/or time.

Second, random effects (RE) assumes that the μ_i are random variables, distributed independently across groups with variance σ_μ^2. When T is short, including year dummies for λ_t facilitates a two-way model. Estimates of this specification are based on transformations of the data into deviations from weighted respective state-means where the weights are based on, generally, the estimated variances of the components in Equation (26) and T (Feasible General Least Squares (FGLS) estimator). Specifically, the weight on the state-means takes the form,

$$\hat{\theta} = 1 - \frac{\hat{\sigma}_v}{\sqrt{T\hat{\sigma}_\mu^2 + \hat{\sigma}_v^2}}. \tag{27}$$

Note, if $\hat{\theta} = 1$, random effects is the within fixed effects estimator. Unbiased robust estimates of the variance components are best obtained from pooled OLS and LSDV estimators. The potential correlation of the specific effects with the variables in X_{it} is a defect of the random effects construct. If these correlations are present, random effects estimation yields biased and inconsistent estimates of β. Conversely, by transforming the data into deviations from the simple state-means the fixed effects estimator is not impacted by this lack of orthogonality.

Hausman (1978) and Kang (1985) outlined specification tests of the null hypotheses of orthogonality between the effects and the regressors. By failing to reject the null, both

---

[13] The model is presented in the most general form. As careful trial will show below, year-effects test jointly insignificant in all specifications therefore the one-way model is estimated and presented.

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

fixed effects and random effects are unbiased and consistent, but fixed effects is less efficient. When the null is rejected, fixed effects is unbiased and consistent but random effects is not. The random effects specification requires exogeneity of all regressors and the components in Equation (26). Conversely, the fixed effects model allows for endogeneity of all the regressors and the specific effects, but ignores time-invariance on the RHS.

A common problem in the estimation of panels with a relatively large number of cross-sectional group units, is cointegrating relationships among the units (Baltagi et al., 2007). Presuming that errors are cross-sectionally independent can lead to incorrect inference and inconsistent estimates. Pesaran (2004; 2015; 2021) proposed the following test statistic,

$$CD = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right), \tag{28}$$

where $\hat{\rho}_{ij}$ denotes the pair-wise correlation coefficients from regression residuals. This CD statistic is distributed, generally, asymptotically standard normal and the null hypothesis depends on the relative expansion rates of N and T. Under general conditions, even when T is small, the appropriate null is weak dependence (Pesaran, 2015). In many cases, an easy correction for cross-section dependence is the fixed effects or within estimator. However, Juodis and Reese (2022) revealed that the Pesaran CD test, when applied to residuals from a model correcting with fixed effects (FE), tends to diverge in favor of the null hypothesis.[14] The authors forward a weighted form of pair-wise correlation,

$$\hat{\rho}_{ij} = \sum_{t=1}^{T} w_i \hat{v}_{it} w_j \hat{v}_{jt}, \tag{29}$$

where $\hat{v}_{it}$ denotes the regression residuals and the weights, w_i, are iid Rademacher. Under the null, the statistic is again asymptotically normal. We will apply the weighted form in formal residual testing below.

Empirical results

Results from pooled OLS and one-way fixed effects specifications are presented in Table 3 below. The OLS results, presented for comparison, reflect a very small $\overline{R}^2$ (0.05) and are easily challenged by the rejection of state homogeneity at the less than one percent level. Controlling for year-specific effects, after accounting for state effects, tests jointly insignificant at all conventional levels ($F(12, 594) = 0.64$). The $\overline{R}^2$ of the one-way specification is quite large, 0.98. A Hausman test rejects orthogonality between the effects and the regressors, at the less than one percent level, favoring the one-way fixed

[14] Juodis and Reese (2022) also look at other correction forms such as common-correlated effect (CCE) models.

effects specification. Cross-section dependence, post fixed effects correction with residual weighting, tests out as weak.

An initial observation concerns the explanatory power of the one-way specification. This estimator explains roughly 98 percent of the variation in beer production per capita, yet included regressors contribute very little to the goodness of fit. Only two time-varying observables are significant at conventional levels. Higher state excise taxes and, to a lesser degree of significance, higher water and sewer costs negatively impact state beer production. An increase in beer excise tax burden, by one standard deviation, reduces per capita production by roughly 4 percent, for the average state. Interestingly, increasing brewery counts, during this sampled period of unprecedented expansion, is not statistically significant. To put this seemingly odd result into context, of the over 7,600 breweries permitted and reporting to the TTB in 2020, 80 percent reported production for the year of 0 to 1,000 barrels. The average production for these small breweries (over 6,000 reporting) in 2020 was roughly 200 barrels with over 900 breweries reporting no production at all. In contrast, the Molson Coors brewery in Golden produces an estimated 9.7 million barrels annually (Molson Coors, 2026). Consequently, the small-beer boom depicted in Figure 4 has had little impact (statistically) on the variation of overall state-level beer production in the U.S.

Table 3. Estimation results

	OLS	**One-Way FE** [a]
Constant	-0.351	[b]
Breweries	0.230***	-0.069
Electric	0.256	0.202
Taxes	0.222	-0.099**
Wage	0.348*	-0.076
Water	-0.986**	-0.266*
Unemployment	0.446	0.101
$\overline{R}^2$	0.050	0.979
CD Test	18.821***	1.934* [c]
F-Test, One-Way FE vs OLS, df(50,606)	-	536.158***
F-Test, adding year effects, df(12,594)	-	0.640
Hausman Test	-	16.587***

Level of significance: (***) < 1%, (**) < 5%, (*) < 10% .
[a]Robust White (cluster) variance-covariance matrix used. This estimator is designed to accommodate arbitrary heteroskedasticity and within cross-section serial correlation.
[b]See Table 4 for State fixed effects estimates.
[c]Juodis and Reese (2022) weighted Pesaran statistic.

Table 4 below shows the fixed effects estimates for each state and the District of Columbia. Fifteen are significant at conventional levels with ten, listed by size in descending order Colorado, New Hampshire, Missouri, Wisconsin, Virginia, Ohio, Georgia, Texas, New Jersey and North Carolina, showing significance at the less than one percent level. This handful of states' time-invariant or latent attributes explain over 90 percent of the variation in per capita beer production sampled. We propose that these significant state effects are capturing impacts of legacy economies. Focusing on the 10

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

states highlighted in Table 4, all boast legacy sites, subsequent acquisitions or logistic expansions that pre-date the time period of our sample.[15] Table 5 below shows the results from a ten-state effects specification. Predictably, estimates are analogous to those in Table 3 with a slightly lower $\overline{R}^2$.

Table 4. State fixed effects

State	Fixed Effect	State	Fixed Effect
Alabama	-5.791	Montana	5.595
Alaska	4.668	Nebraska	0.664
Arizona	0.154	Nevada	-0.619
Arkansas	0.042	**New Hampshire**	**70.592*** **
California	14.663*	**New Jersey**	**18.114*** **
Colorado	**125.209*** **	New Mexico	0.885
Connecticut	0.318	New York	11.645
Delaware	7.500	**North Carolina**	**15.623*** **
District	1.841	North Dakota	1.212
Florida	13.369*	**Ohio**	**50.641*** **
Georgia	**42.005*** **	Oklahoma	-1.163
Hawaii	-7.650	Oregon	12.318**
Idaho	1.878	Pennsylvania	11.845*
Illinois	2.522	Rhode Island	0.068
Indiana	0.553	South Carolina	-3.020
Iowa	1.122	South Dakota	0.424
Kansas	-0.095	Tennessee	1.346
Kentucky	-0.202	**Texas**	**20.976*** **
Louisiana	-1.326	Utah	1.223
Maine	6.827	Vermont	14.967**
Maryland	-1.251	**Virginia**	**53.179*** **
Massachusetts	0.208	Washington	1.993
Michigan	3.040	West Virginia	-0.559
Minnesota	5.514	**Wisconsin**	**55.415*** **
Mississippi	-3.447	Wyoming	4.289
Missouri	**69.618*** **		

Level of significance: (***) < 1%, (**) < 5%, (*) < 10% .

[15] Oregon and Vermont, significant at < 5% level, are considered vibrant, contemporary small-beer states.

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

Table 5. Ten-state effects estimation results

	Ten-State FE [d]
Constant	-3.244**
Breweries	0.053
Electric	0.116
Taxes	-0.034*
Wage	-0.056
Water	-0.212*
Unemployment	0.121
$\overline{R}^2$	0.950

Level of significance: (***) < 1%, (**) < 5%, (*) < 10% .
[d]Robust White (cluster) variance-covariance matrix used. State effects included: Colorado, New Hampshire, Missouri, Wisconsin, Virginia, Ohio, Georgia, Texas, New Jersey and North Carolina.

What is it about this group of ten states that capture brewing legacy? Coors was a single-site company with limited western state distribution until the late 1970s. A one hundred year-old taproot, if you will. As part of a 50 state distribution expansion campaign, Coors built a $70 million dollar eastern packaging facility in Virginia with plans to build a new brewery in Rockingham County (L.A. Times, 1985; Baum, 2000). The Elkton, Virginia brewery was completed in 1987. In 2005, Canada's Molson Brewing and Coors merged. By 2008, Molson, Coors and SABMiller (U.S.) formed a joint venture, MillerCoors. Miller's key locations at the time included Milwaukee, Wisconsin (Miller's legacy site), Georgia, North Carolina, Texas and Ohio. In 2016, Molson Coors acquired Miller (U.S.) outright. Coors and Miller's legacy – seven states. The same taproot story follows for Anheuser-Busch. AB's legacy site in St. Louis, Missouri with branch roots expanding to Colorado, Georgia, Texas, Ohio, New Hampshire, New Jersey and Virginia.[16] Note the co-agglomeration in Colorado, Georgia, Texas, Ohio and Virginia.

Another taproot worth mentioning is Pabst Brewing Company, founded in Milwaukee by Jacob Best in 1844. Over recent decades, Pabst has become a brand acquirer with a diverse portfolio of beer legacy. Blatz, Colt 45, Falstaff, Lone Star, Schlitz, Schmidt, Old English, Old Style and many more oldies. Pabst closed its only brewing facility in 1996 but continued brewing operations through a contract arrangement with Stroh Brewing (Wisconsin) until 2001. Miller acquired the Pabst contract in 2001 and Molson Coors honored the arrangement until 2019. City Brewing Company, who purchased the historic Stroh facilities in Wisconsin, contract brews for Pabst to date (Pabst, 2026).

A methodological note. Because our sample is contemporary and does not encompass legacy origins, an argument could be made that we are introducing a form of post-treatment bias. Current-day fixed effects may induce post-treatment bias if any observable or latent factor that varies across states is linked (directly or indirectly) to the legacy breweries (Homola et al., 2024). Two remedies have been forwarded.[17] The first involves a sequential estimation approach which requires pre- and post-treatment data for

[16] AB breweries are also in California, Florida and New York. In 2026 AB-InBev closed their New Hampshire and New Jersey breweries. Molson Coors closed their North Carolina brewery in 2016.
[17] Combinations of the two methods are commonly used (Homola et al., 2024).

the first stage (Acharya et al., 2016). In most economic settings, a unicorn data set. A second method suggests the use of pre-treatment state effects. This correction presumes that the historic borders were significantly different. For the U.S., the state or territory borders were principally set by the late 1800s (see Appendix B for historical maps). Arguably, reducing jurisdictional breadth may be the best way to capture localized effects and minimize biases, again a very rare data set (Henderson, 1988; Glaeser et al., 1992).

Concluding remarks

The central query in the New Economic Geography literature involves explaining the distribution of production across space. We explore two principal theories, historical location advantages and increasing returns. With a stylized inter-jurisdictional general equilibrium model, we show that both the existence and strength of agglomeration forces can dominate all factors impacting local production. In light of these underpinnings, we examine beer industry specific state-level production data from a period of unprecedented firm count expansion. Historical advantage (legacy) emerges as the force of import. Over 90 percent of the variation in per capita beer production in the U.S. is explained by the fixed or latent attributes of roughly ten states. What these states possess is a long-standing industry taproot or an offshoot. We do find co-agglomeration patterns for the larger breweries but the small brewery 'boom' (shown in Figure 4 above) has had little impact on spatial production variation.

References

AB-InBev. (2026). https://www.ab-inbev.com

Acharya, A., Blackwell, M. and Maya, S. (2016). The political legacy of American slavery. *Journal of Politics*, *78*(3), 621-641.

Baldwin, R. and Krugman, P. (2004). Agglomeration, integration and tax harmonisation. *European Economic Review*, *48*, 1-23.

Baltagi, B., Song, S., Jung, B. and Kuh, W. (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, *140*, 5-51.

Baum, D. (2000). *Citizen Coors: An American Dynasty*. William Morrow.

Beer Institute. (2024). https://www.beerinstitute.org

Brewers Association. (2013). Water and wastewater: Treatment/Volume reduction manual. www.brewersassociation.org

Charnysh, V., Finkel, E. and Gehlbach, S. (2023). Historical political economy: Past, present, and future. *Annual Review of Political Science*, *26*, 175-191.

David, P.A. (1994). Why are institutions the 'carriers of history'?: Path dependence and the evolution of conventions, organizations and institutions. *Structural Change and Economic Dynamics*, *5*(2), 205–220.

Department of the Treasury, Alcohol and Tobacco Tax and Trade Bureau (TTB). (2026). http://www.ttb.gov

Elzinga, K., Tremblay, C. and Tremblay, V. (2015). Craft beer in the United States: History, numbers, and geography. *Journal of Wine Economics*, *10*(3), 242-274.

Feehan, J. (1989). Pareto-efficiency with three varieties of public input. *Public Finance*, *44*, 237-248.

Fernández, G. (2005). A note on tax competition in the presence of agglomeration economies. *Regional Science and Urban Economics*, *35*, 837-847.

Garcia-Milà, T. and McGuire, T. (2001). Tax incentives and the city. Paper prepared for the Brookings-Wharton conference on urban affairs, Washington, DC.

Glaeser, E., Kallal, H., Scheinkman, J., and Shleifer, A. (1992) Growth in cities. *Journal of Political Economy*, *36*, 117-143.

Hatzius, J. and Stehn, S. (2012). Comment on: The ins and outs of forecasting unemployment: using labor force flows to forecast the labor market, by Barnichon, R. and Nekarda, C. *Brookings Papers on Economic Activity*, *45*(2), 118-123.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, *46*(6), 1251-1272.

Henderson, J. (1988). *Urban Development. Theory, Fact and Illusion*. Oxford University Press.

Homola, J., Pereira, M. and Tavits, M. (2024). Fixed effects and post-treatment bias in legacy studies. *American Political Science Review*, *118*(1), 537-544.

Juodis, A. and Reese, S. (2022). The incidental parameters problem in testing for remaining cross-section correlation. *Journal of Business and Economic Statistics*, *40*(3), 1191-1203.

Kang, S. (1985). A note on the equivalence of specification tests in the two-factor multivariate variance components model. *Journal of Econometrics*, *28*, 193-203.

Krogstrup, S. (2008). Standard tax competition and increasing returns. *Journal of Public Economic Theory*, *10*(4), 547-561.

Krugman, P. (1991a). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), 483-499.

Krugman, P. (1991b). History and industry location: The case of the manufacturing belt. *American Economic Review*, *81*(2), 80-83.

Kunce, M. (2023). State-level impacts on new brewery location choices: The beer-Snob effect. *Journal of Beer Economics*, *1*(1), 1-18.

L.A. Times. (1985). Seeks nationwide distribution by 1990: Coors plans to build brewery in Virginia. https://www.latimes.com

Martin, R. and Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, *6*(4), 395-437.

Molson Coors. (2026). https://www.molsoncoors.com

Oates, W. and Schwab, R. (1991). The allocative and distributive implications of local fiscal competition. Reprinted in: *Environmental Policy and Fiscal Federalism, Selected Essays of Wallace E. Oates.* 2004. Edward Elgar Publishing.

Pabst Brewing Company. (2026). https://pabst.com

Pesaran, M. (2004). General diagnostic tests for cross section dependence in panels. University of Cambridge Working Papers in Economics No. 0435.

Pesaran, M. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, *34*, 1089-1117.

Pesaran, M. (2021). General diagnostic tests for cross section dependence in panels. *Empirical Economics*, *60*, 13-50.

Wilson, J. (1986). A theory of interregional tax competition. *Journal of Urban Economics*, *19*, 296-315.

Wilson, J. (1999). Theories of tax competition. *National Tax Journal*, *52*, 269-304.

Zodrow, G. and Mieszkowski, P. (1986). Pigou, Tiebout, property taxation, and the underprovision of local public goods. *Journal of Urban Economics*, *19*, 356-370.

Appendix A

Using the left side of text Equation (2), jurisdictional production can be re-written as,

$$F\left(\frac{L}{L},\frac{K}{L},\frac{X}{L}\right)A\left(\frac{K}{L}\right)=f(k,x)a(k)=fa\,, \tag{A1}$$

which denotes jurisdictional production to labor or more generally a per-capita presentation. All variable representations are Appendix specific. Using Euler's theorem, the net of tax wage equation becomes,

$$w=fa-f_k ka-f_x xa-\tau\,. \tag{A2}$$

Net of tax returns to capital are,

$$r=f_k a+\frac{f_x xa}{k}-t\,. \tag{A3}$$

The local government's budget constraint is defined,

$$g+x=\tau+tk\,. \tag{A4}$$

Substituting Equation (A3) into (A2) yields,

$$w=fa-(r+t)k-\tau\,, \tag{A5}$$

with consumption defined as,

$$c=fa-rk-tk-\tau+y\,. \tag{A6}$$

The solution to the local government's problem coincides with the solution to the Lagrangean,

$$\text{Max}_{c,g,k,x,t,\tau}\;\; u(c,g)+\lambda(fa-rk-tk-\tau+y-c)+\mu(\tau+tk-g-x)\,, \tag{A7}$$

with first order conditions,

$$c:\, u_c-\lambda=0\,, \tag{A8}$$

$$g:\, u_g-\mu=0\,, \tag{A9}$$

$$k:\ \lambda(fa_k + af_k - r - t) + \mu t = 0, \tag{A10}$$

\ /

$-\dfrac{f_x xa}{k}$ from Equation (A3),

$$x:\ \lambda f_x a - \mu = 0, \tag{A11}$$

$$t:\ -\lambda k + \mu k = 0, \tag{A12}$$

$$\tau:\ -\lambda + \mu = 0. \tag{A13}$$

From Equations (A12) and (A13), $\lambda = \mu$, using this equality of the multipliers, the simultaneous solution becomes,

$$\frac{u_g}{u_c} = 1, \tag{A14}$$

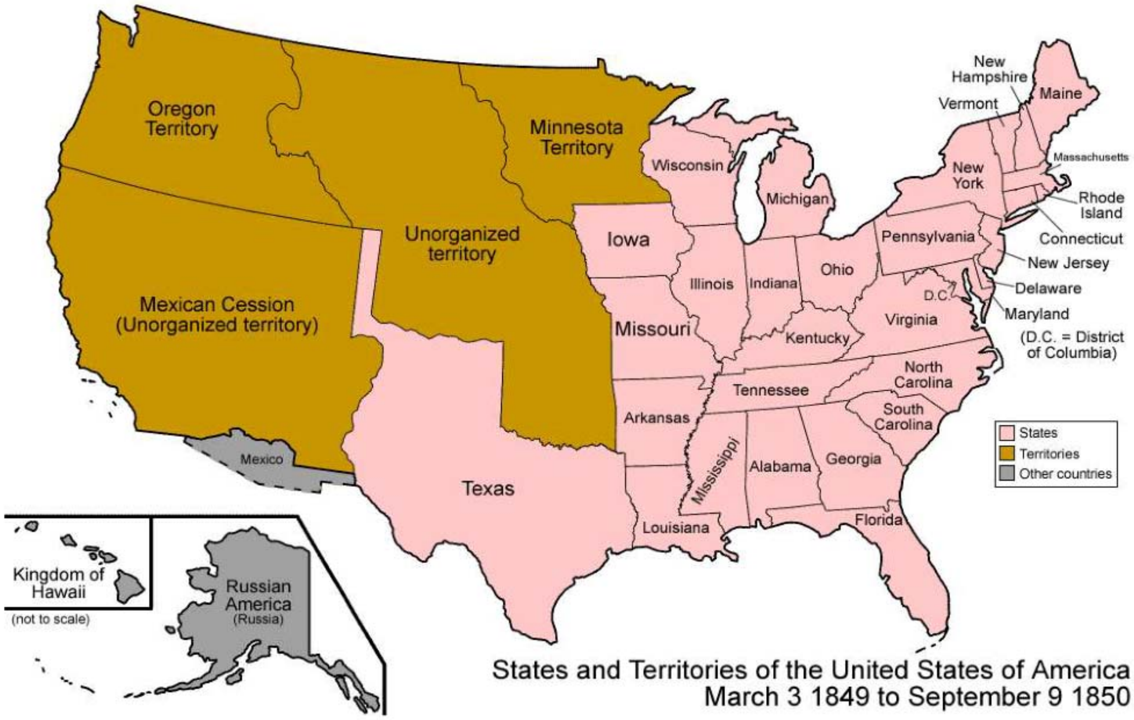
$$t = \frac{f_x xa}{k} - fa_k, \tag{A15}$$

$$f_x a = 1, \tag{A16}$$

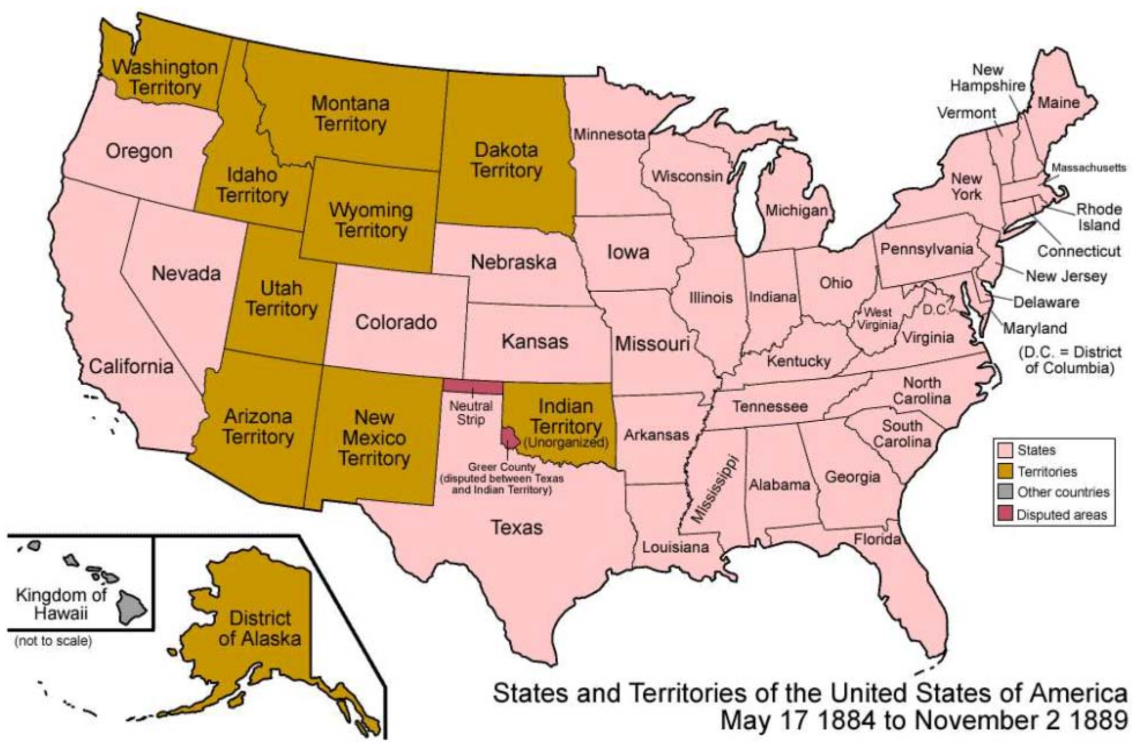
which coincide with text Equations (19) - (21).

Journal of Beer Economics, Vol. 4, No. 1, 2026, 24-44.
ISSN: 2993-2521
https://douglasmitchellconsulting.eu5.org/Journal.html

Appendix B



USA map 1850: Territorial expansion to the Pacific



USA map 1900: The United States at the dawn of the 20th century

Figure B1. Historical state and territory borders