Journal of Wine Economics, Volume 13, Number 1, 2018, Pages 92–98
doi:10.1017/jwe.2018.2

Is There Convergence in National Alcohol Consumption Patterns? Evidence from a Compositional Time Series Approach*

Terence C. Mills [a]

Abstract

Holmes and Anderson (2017a) introduce two extensive data sets on world alcohol consumption and expenditure and with them investigate, among other things, the possible convergence of national alcohol consumption patterns using wine, beer, and spirit shares. Such share data define a composition, on which conventional statistical analysis using covariances and correlations is invalid. This note reanalyses the data using techniques appropriate for a composition and introduces a statistic that can validly track the variation in national shares around the global mean through time. This variability statistic shows that such convergence of national alcohol patterns has clearly taken place over the period 1961 to 2014 and thus confirms Holmes and Anderson's findings using a valid statistical approach. (JEL Classifications: C18, D12, L66)

Keywords: alcohol consumption shares, compositional data analysis.

I. Introduction

Holmes and Anderson (2017a) (henceforth HA) introduce two extensive data sets on world alcohol consumption and expenditure and with them investigate, among other things, the possible convergence of national alcohol consumption patterns.[1] The general argument here is that, with increasing globalization and interactions between cultures, country consumption patterns are converging. Such convergence might raise concerns if consumers in developing countries copy trends in high-income countries that might be considered undesirable from a public health

* The author is grateful for the anonymous comments of a referee.
[a] School of Business and Economics, Loughborough University, Ashby Road, Loughborogh, Leics LE11 3TU, U.K; e-mail: t.c.mills@lboro.ac.uk.
[1] References to these data sets are Holmes and Anderson (2017b) and Anderson and Pinilla (2017). The data series used here were extracted from the megafile_of_global_wine _data_1835_to_2016_031117.xlxs, which is available at www.adelaide.edu.au/wine-econ/databases.

perspective. Tobacco and sugar consumption (the latter also in terms of sweetened non-alcoholic drinks) are obvious concerns, but so too is the consumption of alcohol.

Producers also monitor consumer trends and focus not just on overall levels of consumption but also on changes in composition or mix that might indicate changes in consumer preferences or behaviours. HA report that alcoholic drink producers are clearly aware that wine's share of global recorded alcohol consumption volume has more than halved between 1961 and 2015, falling from 35% to 15%, while beer's share has increased from 29% to 42%, and spirits' share rose from 36% to 43%, over the same period.

The extent to which alcohol consumption patterns are converging across countries has thus been the subject of numerous studies but, prior to HA, such studies have necessarily been limited, because of the paucity of available data, to analysing specific regions or groups of high-income countries or just one or two types of alcohol. HA use their extensive set of annual data from 1961 to 2014 on more than 50 countries and regions to analyse convergence of recorded alcohol consumption and of the beverage mix using a range of convergence indices.

The aim of this article is to add to HA's analysis by utilising the compositional characteristics of wine, beer, and spirits shares in total alcohol consumption to measure convergence over time in national alcohol consumption mixes toward the (changing) world average mix.

II. Compositional Data Analysis of Alcohol Shares

The initial focus here is on the wine, beer, and spirits share of total alcohol consumption for country i in year t. If these shares are denoted $w_{i,t}$, $b_{i,t}$, and $s_{i,t}$ then, as well as the non-negativity constraints $w_{i,t} \geq 0$, $b_{i,t} \geq 0$, and $s_{i,t} \geq 0$, there is also the adding-up constraint $w_{i,t} + b_{i,t} + s_{i,t} = 1$. If there are N countries and T years of observations, then this data may be represented by the set of $N \times 3$ matrices

$$\boldsymbol{X}_t = \begin{bmatrix} w_{1,t} & b_{1,t} & s_{1,t} \\ w_{2,t} & b_{2,t} & s_{2,t} \\ \vdots & \vdots & \vdots \\ w_{N,t} & b_{N,t} & s_{N,t} \end{bmatrix} = [\,\boldsymbol{w}_t \quad \boldsymbol{b}_t \quad \boldsymbol{s}_t\,] \qquad t = 1, 2, \cdots, T,$$

where $\boldsymbol{w}_t + \boldsymbol{b}_t + \boldsymbol{s}_t = \boldsymbol{\iota} = [\,1 \quad 1 \quad \ldots \quad 1\,]'$, the $N \times 1$ unit vector. The sub-matrix

$$\boldsymbol{X}_t^{(2)} = [\,\boldsymbol{w}_t \quad \boldsymbol{b}_t\,],$$

then lies in the two-dimensional *simplex* $\mathcal{S}^2$ embedded in three-dimensional real space $\mathcal{R}^3$ with $\boldsymbol{s}_t = \boldsymbol{\iota} - \boldsymbol{w}_t - \boldsymbol{b}_t$ being the vector of "fill-up" values and $\boldsymbol{X}_t = \left[\boldsymbol{X}_t^{(2)} \quad \boldsymbol{s}_t\right]$.

There are several difficulties with analysing $\boldsymbol{X}_t$ within the simplex sample space, these being a consequence of the summation constraint rendering standard covariance and correlation analysis invalid. These were foreseen by Pearson (1897) but were brought to the attention of applied statisticians by Aitchison (1982), who subsequently followed this article with his book (Aitchison, 2003). For example, consider the *crude covariance matrix* of the shares for a particular year (the subscript t is now omitted for notational convenience)

$$\boldsymbol{K} = \begin{bmatrix} \sigma_w^2 & \sigma_{wb} & \sigma_{ws} \\ \sigma_{wb} & \sigma_b^2 & \sigma_{bs} \\ \sigma_{ws} & \sigma_{bs} & \sigma_s^2 \end{bmatrix},$$

with accompanying *crude correlations*

$$\rho_{wb} = \sigma_{wb}/\sigma_w\sigma_b \qquad \rho_{ws} = \sigma_{ws}/\sigma_w\sigma_s \qquad \rho_{bs} = \sigma_{bs}/\sigma_b\sigma_s.$$

Now consider the covariance between $\boldsymbol{w}$ and $\boldsymbol{w}+\boldsymbol{b}+\boldsymbol{s}$, $\sigma_{w,w+b+s}$. Because of the unit sum constraint this is the covariance between $\boldsymbol{w}$ and $\boldsymbol{\iota}$, which is clearly zero. However, this then implies that, since $\sigma_{w,w+b+s} = \sigma_w^2 + \sigma_{wb} + \sigma_{ws}$, it must be the case that

$$\sigma_{wb} + \sigma_{ws} = -\sigma_w^2. \tag{1}$$

The right-hand side of equation (1) must be negative except for the trivial situation where the wine share is constant across all countries. Thus at least one of the covariances on the left must be negative or, equivalently, that there must be at least one negative element in the first row of the crude covariance matrix $\boldsymbol{K}$. The same negative bias will similarly occur in the other rows of $\boldsymbol{K}$ so that all three covariances must be negative.

In terms of the correlations, it would normally be the case that the three correlations ρ_{wb}, ρ_{ws}, and ρ_{bs} could take any values in the interval –1 to + 1 subject to the non-negative definiteness condition

$$1 + 2\rho_{wb}\rho_{ws}\rho_{ws} - \rho_{wb}^2 - \rho_{ws}^2 - \rho_{bs}^2 \geq 0.$$

For example, since $1 + 2\rho^2 - 3\rho^3 \geq 0$ for $0 \leq \rho \leq 1$, it should be perfectly feasible for all three correlations to be equal and positive. Such a set of correlations for a three-part composition is, however, prohibited by the negative bias property.

These difficulties also affect the coefficient of variation employed by HA as one of their indicators of convergence. The coefficients of variation for the three shares are defined as

$$cv(w) = \sigma_w/\mu_w \qquad cv(b) = \sigma_b/\mu_b \qquad cv(s) = \sigma_s/\mu_s,$$

where μ_w, μ_b, and μ_s are the mean shares. Given the unit sum constraint it is straightforward to show that since, for example, $\mu_s = 1 - \mu_b - \mu_w$ and $\sigma_s^2 = -(\sigma_{bs} + \sigma_{ws})$, then

$$cv(s) = \sqrt{-(\sigma_{bs} + \sigma_{ws})}/\left(1 - \mu_b - \mu_w\right) = \sqrt{-(\sigma_{bs} + \sigma_{ws})}/(1 - \sigma_b cv(b) - \sigma_w cv(w)),$$

and dependencies therefore exist between the three coefficients of variation, thus rendering problematic any conclusions about convergence based on these statistics.

Because of the difficulties in interpreting covariances and correlations calculated from $\boldsymbol{X}_t$ within the simplex sample space, Aitchison (1982, 2003) thus proposed mapping $\boldsymbol{X}_t^{(2)}$ from $\mathcal{S}^2$ to the two-dimensional real space $\mathcal{R}^2$ and then examining the statistical properties of the transformed data within $\mathcal{R}^2$.[2] Several transformations have been proposed for doing this, the most popular being the *additive log-ratio*, defined, for the general three-shares composition

$$\boldsymbol{X}_t = [\,\boldsymbol{x}_{1,t} \quad \boldsymbol{x}_{2,t} \quad \boldsymbol{x}_{3,t}\,] = \left(\boldsymbol{X}_t^{(2)} \quad \boldsymbol{x}_{3,t}\right),$$

where

$$\boldsymbol{x}_{i,t} = (\,x_{i1,t} \quad \cdots \quad x_{iN,t}\,)' \qquad i = 1,\ 2,\ 3,$$

as

$$\boldsymbol{Y}_t = \left[\boldsymbol{y}_{1,t} \quad \boldsymbol{y}_{2,t}\right] = a_2\left(\boldsymbol{X}_t^{(2)}\right) = \left[\,log\left(\boldsymbol{x}_{1,t}/\boldsymbol{x}_{3,t}\right) \quad log\left(\boldsymbol{x}_{2,t}/\boldsymbol{x}_{3,t}\right)\right].$$

The inverse transformation, known as the *additive-logistic*, is

$$\boldsymbol{X}_t^{(2)} = a_2^{-1}(\boldsymbol{Y_t}) = \left[\,exp\left(\boldsymbol{y}_{1,t}\right)/\boldsymbol{y}_t \quad exp\left(\boldsymbol{y}_{2,t}\right)/\boldsymbol{y}_t \quad exp\left(\boldsymbol{y}_{3,t}\right)/\boldsymbol{y}_t\,\right]$$

$$\boldsymbol{x}_{3,t} = 1/\boldsymbol{y}_t,$$

where

$$\boldsymbol{y}_t = 1 + exp\left(\boldsymbol{y}_{1,t}\right) + exp\left(\boldsymbol{y}_{2,t}\right).$$

Although it is important to note that $a_d(\cdot)$ is invariant to the choice of fill-up share, so that nothing in the statistical analysis hinges on which of the shares is chosen for this role, it is possible to avoid choosing a fill-up by using the *centred log-ratio* transformation, defined as

$$\boldsymbol{Z}_t = c_d\left(\boldsymbol{X}_t^{(2)}\right) = \left[\,log\left(\boldsymbol{x}_{1,t}/g(\boldsymbol{X}_t)\right) \quad log\left(\boldsymbol{x}_{2,t}/g(\boldsymbol{X}_t)\right) \quad log\left(\boldsymbol{x}_{3,t}/g(\boldsymbol{X}_t)\right)\right],$$

[2] The following results were presented by Aitchison (2003) for the general d-dimensional simplex; see also Mills (2007, 2009, 2010).

https://doi.org/10.1017/jwe.2018.2 Published online by Cambridge University Press https://doi.org/10.1017/jwe.2018.2 Published online by Cambridge University Press

where

$$g(\boldsymbol{X}_t) = \begin{bmatrix} (x_{11,t}x_{21,t}x_{31,t})^{1/3} \\ (x_{12,t}x_{22,t}x_{32,t})^{1/3} \\ \vdots \\ (x_{1N,t}x_{2N,t}x_{3N,t})^{1/3} \end{bmatrix},$$

is the vector of geometric means.

Although the centred log-ratio transformation introduces a singularity since $\boldsymbol{Z}_t\boldsymbol{\iota} = \boldsymbol{0}$, which has limited its application in favour of the additive log-ratio transformation, it does have a symmetric, albeit singular, covariance matrix

$$\boldsymbol{\Gamma}_t = \left[\gamma_{ij,t}\right] = \left[cov\left\{log\left(\boldsymbol{x}_{i,t}/g(\boldsymbol{X}_t)\right), log\left(\boldsymbol{x}_{j,t}/g(\boldsymbol{X}_t)\right)\right\} : i,j = 1,2,3\right].$$

In the application considered here this is a major advantage over the additive log-ratio transform, for a measure of the total variability of $\boldsymbol{X}_t$ is given by the trace of $\boldsymbol{\Gamma}_t$

$$tr(\boldsymbol{\Gamma}_t) = \lambda_1 + \lambda_2,$$

where $\lambda_1 > \lambda_2$ are the eigenvalues of $\boldsymbol{\Gamma}_t$ (Aitchison, 2003, pp. 190–191, Definition 8.1 and Property 8.1).

III. Convergence of Alcohol Shares through Time

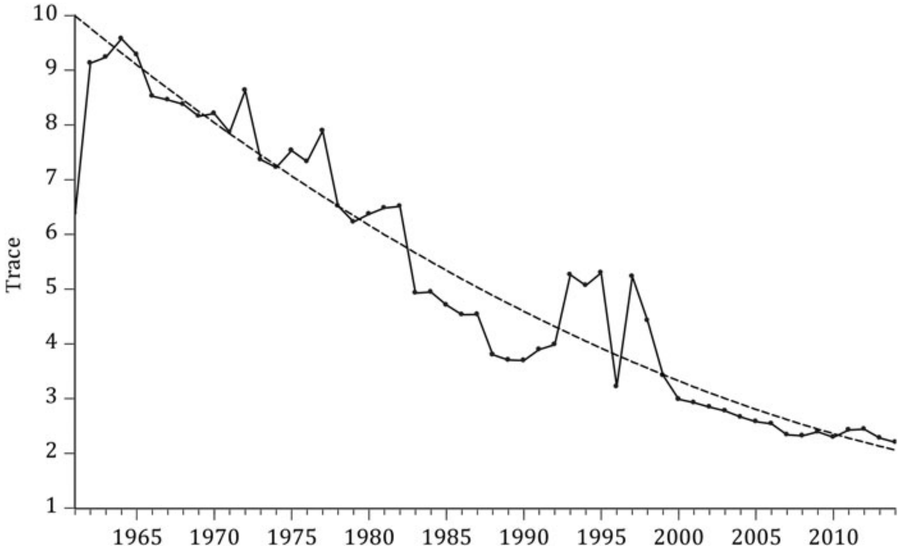
The trace of the centred log-ratio covariance matrix, $tr(\boldsymbol{\Gamma}_t)$, thus provides a measure of the total variation in alcohol shares in year t, and computation of this statistic across the T years for which data on alcohol shares are available will therefore allow an assessment of the extent of convergence to the (changing) mean shares through time.

Annual data from 1961 to 2014 on wine, beer, and spirits shares of total alcohol consumption volume is available for up to 53 countries and regions each year in Tables 53(a)–(c) of Anderson and Pinilla (2017).[3] For each year $t = 1961, \cdots, 2014$ the centred log-ratios were calculated for the set of country shares available and the covariance matrix $\boldsymbol{\Gamma}_t$ constructed. From this the eigenvalues were computed and $tr(\boldsymbol{\Gamma}_t)$ obtained. The sequence of 54 traces are plotted in Figure 1 with the fitted quadratic trend

[3] A small number of countries have occasional zero shares in some years. The presence of a zero share obviously precludes using a log-ratio transformation. Although a literature does exist on methods for dealing with zero shares (see Mills, 2007, for a survey), they all essentially replace the zero share with a small positive value and reduce the other shares by appropriate amounts. Because of the small number of shares affected, zero shares were replaced by 0.001%, with the other two shares for that country being decreased by 0.0005% in each case.

Figure 1

Trace of Log-Ratio Covariance Matrix, 1961–2014, with Fitted Quadratic Trend Superimposed



$$\widehat{tr(\boldsymbol{\Gamma}_t)} = 9.99 - 0.230t + 0.0015t^2,$$

superimposed to emphasise the decline over time in this measure of total variability.

It is thus clear that, using this metric of total variation based on the compositional character of shares data, that the variation in alcohol shares around the global mean shares has declined substantially over the half-century from the early 1960s, thus confirming, in a statistically valid way, the findings of HA.

IV. Conclusions

In this article we have utilised the compositional properties of country alcohol share data to compute a statistically valid measure of variation in alcohol shares for each year from 1961 to 2014. This measure shows a quadratically declining trend in variation over this period and thus confirms HA's finding that the cross-country alcohol mix has converged substantially towards the global alcohol mix over the last half century or so.

References

Aitchison, J. (1982). The statistical analysis of compositional data (with discussion). *Journal of the Royal Statistical Society, Series B* 44(2), 139–177.

Aitchison, J. (2003). *The Statistical Analysis of Compositional Data*. Caldwell, NJ: Blackburn Press.

Anderson, K., and Pinilla, V. (with the assistance of A. J. Holmes) (2017). *Annual Database of Global Wine Markets, 1835 to 2016*. Wine Economics Research Centre, University of Adelaide. Available at www.adelaide.edu.au/wine-econ/databases/global-wine-history.

Holmes, A. J., and Anderson, K. (2017a). Convergence in national alcohol consumption patterns: New global indicators. *Journal of Wine Economics*, 12(2), 117–148.

Holmes, A. J., and Anderson, K. (2017b). *Annual Database of National Beverage Consumption Volumes and Expenditures, 1950 to 2015*. Wine Economics Research Centre, University of Adelaide. Available at www.adelaide.edu.au/wine-econ/databases/alcohol-consumption.

Mills, T. C. (2007). Modelling compositional economic data. *Indian Economic Journal*, 55(3), 99–115.

Mills, T. C. (2009). Forecasting obesity trends in England. *Journal of the Royal Statistical Society, Series A* 172(1), 107–117.

Mills, T. C. (2010). Forecasting compositional time series. *Quality and Quantity*, 44(4), 673–690.

Pearson, K. (1897). Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London, Series A* 60, 489–498.