International Journal of Health Sciences
March 2021, Vol. 9, No. 1, pp. 1-8
ISSN: 2372-5060 (Print), 2372-5079 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/ijhs.v9n1a1
URL: https://doi.org/10.15640/ijhs.v9n1a1

A Factor Variance Decomposition Analysis of Cancer Mortality due to Temporal Shocks in Alcohol and Tobacco Consumption

Zahra N. Masih[1]

ABSTRACT

Identifying the empirical response of cancer mortality to changes in alcohol and tobacco consumption has long been acknowledged as a critical area of preventative cancer strategy and public health policy. This is the first study of its kind that utilizes times series methods to decompose variations in forecasted shocks to cancer mortality into its most significant attributable factors. The analysis uses tobacco expenditure, alcohol expenditure, controlling for health expenditures and aggregate cancer data observed annually over an 80-year period for the US population. Results indicate alcohol has a more dominant effect on explaining cancer mortality regardless of time dimension. As a result, policies that have been previously emphasized toward mitigating tobacco consumption may prove prudent in addressing alcohol as a public health concern with respect to cancer mortality.

Key Words: cancer mortality, tobacco and alcohol consumption, attribuation analysis, long-run relationship, variance decomposition.

1. Introduction

The causative factors contributing to cancer is an area of research that is yet to be fully discovered. Carcinogens, or cancer-causing agents, tend to have both biological and epidemiological roles in determining cancer risk. Alcohol and tobacco, in particular, have been classified as top behavioral/dietary risks responsible for one-third of all cancer-associated deaths (WHO, 2018). Characterizing how such carcinogens truly impact societal health in regard to cancer risk is essential to determining future public policies that may potentially deter cancer's health impact.

The ingestion of alcohol is causally related to a variety of cancers, including breast, colon, liver and pharyngeal cancers (IARC, 2010). The risk associated to ingestion can be most directly tied to the frequency and amount of ethanol consumed. One study, in particular, demonstrates that a low daily dose of ethanol may enhance carcinogenesis around the body (Kirpalani, 2017). Frequency may be an emphasized threat to certain cancers such as breast cancer, where persistent alcohol intake has a positive linear relationship to breast cancer risk. (de Menezes, et. al., 2013). Along with the risks in ingestion frequency, more health consequences reside in the dose-response relationship alcohol maintains with cancer risk. Numerous studies have causally linked higher amounts of alcohol ingestion to increased cancers risk, as seen with an enhancement of alcohol's immunosuppressive and metabolism-altering behaviors. The dose-response relationship reveals that moderate increases in alcohol consumption (1 drink/day) can result in a 20% increase in in pharynx and oral cavity cancer risk (Bagnardi, 2001). The multitude of studies indicate that alcohol's carcinogenic implications are various in its targets. Despite its pronounced effect on cancer risk, alcohol is an avoidable risk factor, where its detrimental impact may be moderated with the application of preventative strategies and analysis.

Aligning to similar concerns is tobacco. Though not considered a carcinogen by itself, the inhalation of the carcinogenic chemicals present in tobacco is leading causes of cancer and cancer deaths (NIH, 2017). Such chemicals are capable of genetically mutating the DNA of bodily cells, increasing the risk of uncontrolled proliferation of damaged cells (Desrichard et. al., 2018). In terms of long term risks, tobacco has been associated to lung, pharynx, larynx, liver, stomach, pancreas cancers and so on (American Cancer Society, 2018). Tobacco is also causally associated to increased lung cancer risk, andcan significantly impact organ functionality (Gandini et. al., 2008). Much like alcohol, tobacco is an avoidable risk factor that may be mediated with preventative strategies.

[1] College of Arts & Sciences, Cornell University, Ithaca NY 14850, USA. And US Metric Association, Windsor CO 80550, USA. Email: znm4@cornell.edu

Given the evident risk it poses in regards to cancer and the severity of its health consequences, tobacco is of significant intertest for specified analysis on population data and preventative strategies.

Previous research suggests tobacco and alcohol consumption has been heavily studied in evaluating cancer risk, however few studies have considered utilizing statistical methods to forecast the aggregate impact. As of recent, one Australia-based study utilized lag structures to suggest a decrease in population-level drinking and smoking reduces cancer mortality over time. This suggests that applying such statistical methods to cancer mortality may provide a detailed account into how alcohol and tobacco affect health on an aggregate scale. In this paper, we attempt to clarify this issue by presenting predicted cancer risk in divided, quantified explanatory variables. This analysis will not only utilize statistical methods to forecast the behavior of cancer risk, but also characterize the degree toward which both alcohol and tobacco explain cancer risk in multiple time dimensions.

2. Methodology and Modelling

We utilize time series modelling techniques appropriate for using with long time spans. These techniques, while well established in the statistical literature need to be combined with policy inference specific to this area of study. As there is a wealth of literature on the technical side, we provide here a summary of the major features of what these techniques provide.

2.1 Vector Error-Correction Modelling (VECM) and Causality:

Engle and Granger 1987, demonstrated that once a number of variables (say, x_t and y_t) are found to be cointegrated or share a long-run stable relationship, there always exists a corresponding error-correction representation which implies that changes in the dependent variable are a function of the level of disequilibrium in the cointegrating relationship (captured by the error-correction term) as well as changes in other explanatory variable(s). This is the bedrock model that also allows us to test for causal interactions amongst the variables we include in the model. To illustrate we simplify and use a two factor system.

If we exploit the idea that there may exist co-movements between factors related to cancer mortality, such as alcohol and tobacco consumption, and possibilities that they will trend together in finding a long-run stable equilibrium, by the Granger representation theorem we may posit the following testing relationships which constitutes our vector error-correction model:

$$\begin{pmatrix}\Delta x_t\\ \Delta y_t\end{pmatrix}=\begin{pmatrix}d_{11}(L) & d_{12}(L)\\ d_{21}(L)) & d_{22}(L)\end{pmatrix}\begin{pmatrix}\Delta x_t\\ \Delta y_t\end{pmatrix}+\begin{pmatrix}\delta ECT_{t-1}\\ dECT_{t-1}\end{pmatrix}+\begin{pmatrix}c_1\\ c_2\end{pmatrix}+\begin{pmatrix}v_{1t}\\ v_{2t}\end{pmatrix}\cdots(1)$$

where Δ is a difference operator, ECT refers to the error-correction term derived from long-run cointegrating relationship via the Johansen method, c_1 and c_2 are constants, and v_1 and v_2 are serially-uncorrelated random error terms with mean zero.

A consequence of relationships described by (1) is that either Δx_t or Δy_t or both must be caused by ECT_{t-1} which is itself a function of x_{t-1}, y_{t-1}. Intuitively, if y_t and x_t have a common trend, then the current change in x_t (say, the dependent variable) is partly the result of x_t moving into alignment with the trend value of y_t (say, the independent variable). Through the error-correction term, the vector error-correction model opens up an additional channel for Granger-causality (ignored by the standard Granger and Sims tests) to emerge. The Granger-causality (or endogeneity of the dependent variable) can be exposed either through the statistical significance of: (i) the lagged ECTs (δ and d) by a t-test; (ii) a joint test applied to the significance of the sum of the lags of each explanatory variable [$d_{12}(L)$ and $d_{21}(L)$] by a joint F- or Wald χ^2 test; or (iii) a joint test of all the set of terms described in (i) and (ii) by a F- or Wald χ^2 test, i.e. the $d_{12}(L)$ and δ in (1) and $d_{21}(L)$and d in (2). The non-significance of both the t- and F- or Wald tests in the VECM indicates econometric exogeneity of the dependent variable.[2]

2.2 Generalized Variance Decompositions (GVDCs)

[2]Defining the source of causation is still a controversial area particularly when it comes to discerning whether (i) and (ii) constitute long and short-run sources. While lagged changes may be intuitively classified as short-run influences, the ECT involves a constant reconciliation between short-run departures from the long run information channel. This question has been taken up theoretically by Granger (1988).

The VECM can only indicate the presence of Granger-causality of the dependent variable within the sample period but cannot provide the relative contributions of the explanatory variables in explaining the variation in the dependent variable beyond the sample period. The generalized variance decompositions or GVDCs, by partitioning the variance of the forecast error of a certain variable into proportions attributable to innovations (or shocks) in each variable in the system including its own, can provide an indication of relative causal strength. We adopt the partitioning of the variance attributable to a shock in each variable according to the methodology outlined in (Lanne and Nyberg, 2016), who propose a new generalized forecast error variance decomposition with the property that the proportions of the impact accounted for by innovations in each variable sum to unity. Their decomposition methodology is based on the well-established concept of the generalized impulse response function.

3. Data and Estimation Results

Our data set describing cancer mortality (*CAN*) is deaths per 100,000;*ALC* is average per capita alcohol consumption in litres per person per year, *TOB* is total sales of tobacco per adult, per year; and *HLX* is per capita health expenditure (a control factor). All data are observed annuallyfor the US and covers the period 1938 to 2018.For all transformations, annual percentage growth rate of GDP atmarket prices based on constant local currency; aggregates are based on constant 2015 U.S. dollars.All data apart, from health expenditure, are sourced from the "Our World in Data" database (https://ourworldindata.org/). Health expenditure is sourced from OECD Data (https://data.oecd.org/).

Prior to testing for cointegration, we investigated the integrational properties of each of the variables by applying a battery of unit-root testing procedures. These tests verify to what degree the time series under analysis are integrated. We use two tests based on different null hypotheses. Based on the ADF(max) tests and KPSS test for stationarity we could not find any conclusive evidence against the property that all variables were integrated at most of order 1 or $I(1)$.[3]

Given the common integrational properties of these variables, we next proceeded to test for the presence of cointegration in the vector $[CAN_t, ALC_t, TOB_t, HLT_t]$. Results of Johansen's *LR* and trace tests (see Johansen, 1988 and Johansen and Juselius, 1992) are presented in Appendix Table: A2, and indicate that there exists at most one cointegrating realtionship since $r = 0$ is clearly rejected in favour of $r = 1$; but $r \leq 1$ cannot be rejected by the 95% critical values. Given that there exists $(n\text{-}r)$ common trends within the system, we can conclude that there existseveral common trends within the vector of variables.In addition, we tested the long-run restriction that cancer mortality (*CAN*) was statistically zero or insignificant. In all variations of these tests we could not find any conclusive evidence to suggest that *CAN* does not belong in the system. The tests appear in the final column as a chi-square statistic being rejected across all version of the model.

Temporal test results of Granger causality based on the vector error correction model which is described by equations (1) are summarised in Table 2. The table displays test results where each equation in the system is presented in turn pertaining to the tests for causation. We also have results based on each separate set of restrictions that pertain to a different source of causation: the ECT is a proxy for the long-run source of causation as these terms incorporates the long run or level form of the factors; the changes or □ *x*□ or□ □*x*□ □ indicates the short-run source as this captures restriction on purely the changes of the factors; and the joint ECT and □ *x*□ or□ □*x*□ □ indicates the joint source emanating from both short and long-run movements in then factors. It is fairly self-explanatory to note that all three of the channels of Granger causality (short-run, long-run and joint) appear to be significant in the case of alcohol and tobacco consumption having significant impact on aggregate cancer mortality over time. In fact, short and long-run changes in alcohol consumption seem to Granger cause cancer mortality more significantly than when compared to tobacco consumption.Alcohol consumption appears to be comparatively the dominant and consistent driver of cancer deaths.

In contrast, none of the channels of causation are active when *CAN* tries to explain changes in either alcohol or tobacco consumption ie. an absence of reverse causation with the exception of weak joint significance of cancer mortality in the equation where □ ALC_t is dependent. This is evidence indicating that there exists Granger causation mainly in one direction from alcohol consumption to cancer mortality; and from tobacco consumption to cancer mortality. However, a weak relationship exists between cancer mortality explaining changes in alcohol consumption, which will be further analysed in the discussion.

[3]To provide additional evidence, and as a check on the robustness of these tests we performed supplementary tests. Based on augmented Dickey-Fuller and Phillips-Perron tests which are presented in Appendix: Table A1 [see Dickey and Fuller 1981, Perron 1988, Phillips and Perron, 1988], we could not find any significant evidence that the variables in the vector $[CAN_t, ALC_t, TOB_t, HLT_t]$, were not integrated of order one or $I(1)$.

In addition, the interpretation of the *d* parameter being significant implies that when there is a deviation from the equilibrium relationship as measured by the error correction term or ECT ($\square_t = CAN_t - c - ALC_t - TOB_t - HLT_t$) it is y_t that adjusts to restore equilibrium, implying that alcohol and tobacco consumption leads cancer mortality. Furthermore, the novelty of these results is the finding that changes in alcohol and tobacco consumption positively and significantly affects cancer mortality as indicated by the positive coefficients on $\square ALC_t$ and $\square TOB_t$.

Since the VECM is estimated by ordinary least squares regression and our test statistics may be prone to inconsistencies due to violations of its underlying assumptions, a battery of diagnostic tests was applied for each equation and appears in summarised form in Appendix Table: A3. These checks are for serially uncorrelated errors of first and second orders, constance of variance of errors, misspecification of function form and non-normality of errors respectively.In summary, across all these tests, given the power for which they are designed over the sample, we could not find any significant evidence of departures from standard assumptions.Such tests for robustness provide statistical confidence when it comes to conducting appropriate inference.

The results from the generalized variance decompositions go a step further and identify the causal relativities in explaining a shock to cancer mortality. Further, as the partitioning is of the forecast variance of the focus variable (*CAN*), we can ascertain using a quantities approach; to what extent the expanators explain the forecasted variance over time. Results clearly indicate the dominance of alcohol consumption over the other variables: higher than 20% of the forecast variance of cancer mortality is explained by the variance attributable to alcohol consumption using a 10-year post-shock horizon. This figure compares with just over 8% in the case of tobacco consumption.

Finally, the charts illustrate the response path taken by cancer mortality when we independently subject alcohol and tobacco consumption to a once off positive shock. These are displayed in Figures 1A and 1B, for response of cancer mortality from shocks to alcohol and tobacco consumption; and Figures 2A and 2B for the factor shocks related to the joint behaviour of alcohol and tobacco consumption. We find that cancer mortality displays persistence and its path dependency varies considerably between the shock factors of alcohol and tobacco consumption. Note that while shocks are independent, the technique itself allows us to treat the system as simultaneous ie. While we shock a certain factor, say alcohol consumption, the system assumes that other shocks are not in some way "switched off" allowing the dynamics of the system to react as a function of the individual shock.

4. Discussion and Policy Implications

Our variance decomposition of the varitions in cancer mortality due to tobacco consumption, alcohol consumption and health expenditure reveal: (i) over time alcohol dominates the explanation of variations in cancer mortality above all other variables, and (ii) tobacco does explain some variation in cancer mortality however its impact becomes stagnant over time in comparison to alcohol. These results are best explained through an analysis of the biological and political implications regarding cancer mortality.

Firstly, in addressing alcohol (see Table 2) results indicate that even after a 10 year period, up to 20% of cancer mortality is explained through alcohol consumption. This may be determined through the long-term health consequences associated with alcohol ingestion. Numerous studies have pointed to acetaldehyde as the primary contributor to detrimental health consequences, due to its role in the metabolic process. Acetaldehyde can cause various forms of DNA damage and mutations which may ultimately deter cellular performance. This may ultimately lead to carcinogenesis and weakened immunity (Mizumoto, et. al., 2017, Kwo et al., 1998). Acetaldehyde exists temporarily in the body when alcohol is consumed, in which alcohol dehydrogenase enzymes are responsible for its expulsion (Paton, 2005). Studies indicate that with larger amounts of alcohol consumed, higher concentrations of acetaldehyde may exist in the body, potentially increasing the presence of acetaldehyde in the body and promoting carcinogenesis (Seitz et. al., 2007). This may explain the data by indicating that alcohol consumption over time may increase the probability of cancer risk based on implications of acetaldehyde presence in the body.

Our results also indicate that tobacco impacts cancer mortality variability up to 8%, comparatively less than alcohol over the 10 year span. The explanation of impact has been highlighted in other epidemiological studies, in which decreases in tobacco consumption was correlatively linked to a 16% reduction in overall cancer mortality over a 20-year period (Jiang et. al., 2018). However, our data indicates alcohol to have a much more pronounced effect on cancer mortality and this may be explained by the varying levels of acetaldehyde in each substance. Along with the direct metabolism of alcohol, acetaldehyde can also enter the body through inhaled cigarette smoke. Though both substances are related to acetaldehyde, its concentrations vary depending on the degree of substance used.

One study indicates alcohol may promote higher concentrations of acetaldehyde in the body through an analysis of frequent drinking and frequent smoking. Their results demonstrate that ethanol concentrations over a range of 53 to 33 mM/acetaldehyde can lead to an average of 42.7 μM acetaldehyde in the blood. Whereas, cigarette smoke [typically maintaining 1μg of acetaldehyde] was almost "undetectable" in the blood at moderate consumption (Korsten et. al., 1975). Therefore, it may be argued that with higher concentrations of alcohol in the body, acetaldehyde has a higher probability at promoting carcinogenesis and therefore may lead to increased cancer risk over time, compared to tobacco which has particularly lower concentrations and bodily exposure levels of acetaldehyde.

The dominance in the explanation of alcohol argues the need for better public health policies surrounding its consumption. Along with analyzing the significance of each variable on cancer mortality, we analyzed the sales for tobacco and alcohol over an 88-year period (see Figure 1) in considering the impact of public policy on these variables respectively. As our data suggests, the mediation of tobacco and its rather stagnant impact over the 10 year period could be due to how tobacco consumption has changed as a result of policy implications. American policy has continuously acknowledged the pressing health concerns of tobacco with the Surgeon General's overall review of tobacco risk factors, adding smoke-free environments to public facilities, reduced advertising, adding warning labels to merchandise, etc. (Stratton et. al., 2001). The direct impact of such policies and awareness are seen in the steady decline of tobacco sales post 1980, where the decline could be explained with the increased awareness for tobacco's negative health consequences and banning of cigarettes in multiple public areas. Alcohol, when compared to tobacco, demonstrates a steady incline in sales past 1980 and has less intervention strategies than tobacco. Alcohol consumption is therefore relatively unregulated, where regulation is said to be "self-imposed" and "voluntary at the federal level" despite its known carcinogenic implications (American Addiction Centers, 2020). Recognizing alcohol, as our results find, as explaining nearly a fifth of cancer mortality is necessary to determining effective health policy strategies going forward.

Limitations and Conclusions

Our results provide a seminal attempt at predicting both the behaviors and explanatory variables associated with aggregate cancer mortality in the long run. The dominance of alcohol, contributing to approximately a fifth in explaining cancer mortality, is a public health concern that has yet to be mediated. The dominance of alcohol could be dissected through a variety of pathways. The variability in our results could be due to the type of sample size being utilized, as the data set was on overall mortality as opposed to considering specific cancers. With this comes the acknowledgement that tobacco and alcohol may impact particular cancers differently, or at varying degrees. For example, tobacco may have a larger degree of impact with lung cancer as opposed to alcohol's impact. Similar limitations are seen with the age groups and gender differences unaccounted for in the study. Future studies should consider further validating the long-run relationship between alcohol and cancer mortality with a particular emphasis on specified populations to determine if the risk is truly persistent or versatile in its impact.

References

American Cancer Society. Health Risks of Smoking Tobacco. *Cancer.org.*

Bagnardi, V., Blangiardo, M., La Vecchia, C., & Corrao, G. Alcohol consumption and the risk of cancer: A meta-analysis. *Alcohol Research and Health*. 2001;**25**(4), 263–270.

de Menezes, R. F., Bergmann, A., & Thuler, L. C. S. Alcohol consumption and risk of cancer: A systematic literature review. *Asian Pacific Journal of Cancer Prevention*. 2013; **14**(9), 4965–4972.

Desrichard, A., Kuo, F., Chowell, D., Lee, K.-W., Riaz, N., Wong, R. J., Chan, T. A., & Morris, L. G. T. Tobacco Smoking-Associated Alterations in the Immune Microenvironment of Squamous Cell Carcinomas. *Journal of the National Cancer Institute*. 2018; *110*(12), 1386–1392.

Dickey, D.A. and Fuller W.A. Distribution of the estimators for autoregressive time-series with a unit root, *Journal of the American Statistical Association*.1979; **74**, 427-431.

Dickey, D.A., and Fuller W.A., Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*. 1981;**49**, 1057-1072.

Editorial Staff. Rules & Regulations About Marketing Alcohol. *American Addiction Centers.* 2020

Engle, R.F. and C.W.J. Granger. Cointegration and error correction: Representation, estimation, and testing, *Econometrica*.1987;**55**, 251-276.

Gandini, S., Edoardo, B., Simona, I., Mathieu, B., Lowenfels, A. B., Patrick, M., Peter., B. Tobacco Smoking and Cancer: A Meta-Analysis. *International Journal of Cancer.* 2008; **122**(1): 155–64.

Granger, C.W.J., Some recent developments in a concept of causality, *Journal of Econometrics*. 1988;**39**, 199-211.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans., & International Agency for Research on Cancer. (2010). *Alcohol consumption and ethyl carbamate / IARC Working Group on the Evaluation of Carcinogenic Risks to Humans*. ix.

Jiang, H., Livingston, M., Room, R., Chenhall, R., & English, D. R. (2018). Temporal Associations of Alcohol and Tobacco ConsumptionWith Cancer Mortality. *JAMA Network Open*, **1**(3), e180713.

Johansen, S. Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*. 1988; **12**, 231-254.

Johansen, S., and K. Juselius. Maximum likelihood estimation and inference on cointegration with applications to money demand, *Oxford Bulletin of Economics and Statistics*. 1990;**52**, 169-210.

Lanne, M., and H. Nyberg.Generalised Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models,*Oxford Bulletin of Economics and Statistics*. 2016;**78**, 595–603.

Kirpalani, P. Alcohol & Cancer. Mimeo, University of North Carolina. 2017; http://www.researchgate.net/publication/321225698

Korsten, M. A., Matsuzaki, S., Feinman, L., & Lieber, C. S. High Blood Acetaldehyde Levels after Ethanol Administration - Difference between Alcoholic and Nonalcoholic Subjects. *New England Journal of Medicine*. 1975; **292**(8), 386–389.

Kwo, P.Y.; Ramchandani, V.A.; O'Connor, S.; et al. Gender differences in alcohol metabolism: Relationship to liver volume and effect of adjusting for body mass. *Gastroenterology*. 1998; 115, 1552–1557.

Paton, A. Alcohol in the Body. *The BMJ*.2005;330(7482): 85–87

Perron, P. Trends and random walks in macroeconomic time series, *Journal of Economic Dynamics and Control*. 1988;**12**, 297-332.

Peterson, L. A., Balbo, S., Fujioka, N., Hatsukami, D. K., Hecht, S. S., Murphy, S. E., Stepanov, I., Tretyakova, N., Turesky, R. J., & Villalta, P. W. Applying tobacco, environmental and dietary-related biomarkers to understand cancer etiology and evaluate prevention strategies. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*.2020.

Phillips, P.C.B., and P. Perron Testing for a unit root in time series regression, *Biometrika*. 1988;**75**, 335-346.

Seitz, H. K., & Becker, P. Alcohol Metabolism and Cancer Risk. *Alcohol Research & Health*. 2007; **30**(1), 38–47.

Stratton, K., Shetty, P., Wallace, R., Bondurant S. Clearing the smoke: Assessing the science base for tobacco harm reduction. *ADDICTION*, 2001; *97*(11), 1480–1481.

World Health Organization. Cancer. *World Health Organization.* 2018.

Table 1. Summary of Stationarity Properties

Variable	ADF(max)	KPSS	Conclusion
Cancer Mort (*CAN*)	-1.325	0.943**	*I*(1)
Alcohol Cons (*ALC*)	-2.475	0.658**	*I*(1)
Tobacco Cons (*TOB*)	-1.098	1.142**	*I*(1)
Health Exp (*HLX*)	-2.001	0.981*	*I*(1)

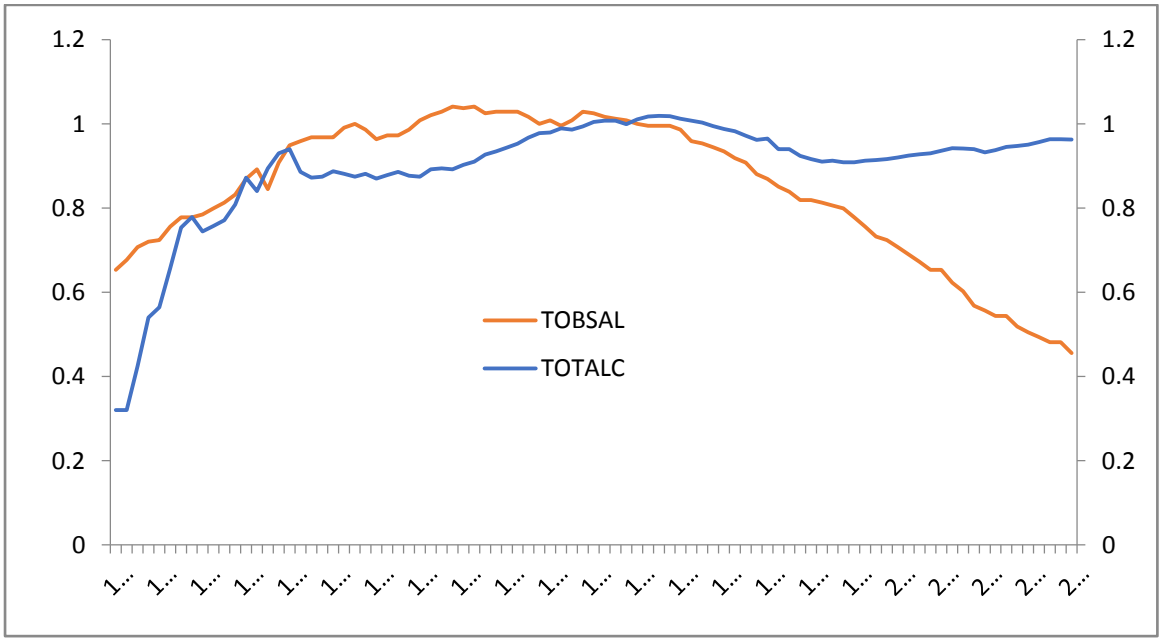
*Notes:*Using the appropriate notation, a series x_t is said to be integrated of order *d*, if it has an invertible ARMA representation after being differenced *d* times. For example, a stationary series is indicated by *I*(0), whereas a non-stationary series in levels, but stationary in first differences is indicated by *I*(1). Data observed annually from 1938 to 2018 inclusive.

Table 2.Generalized Variance Decompositions from Multivariate Models of Cancer Mortality, Alcohol and Tobacco Consumption, and Health Expenditures

Relative Variance of □ *CAN*				
Years	Percentage of Forecast Variance Explained by Innovations in:			
	□ *CAN*	□ *ALC*	□ *TOB*	□ *HLX*
1	93.45 (98.86)	1.18 (1.14)	3.90 (0.00)	1.47 (100.00)
2	84.48 (96.12)	11.82 (3.88)	8.61 (0.13)	1.82 (99.86)
5	75.27 (85.88)	13.19 (14.12)	8.22 (32.40)	3.63 (67.35)
10	60.85 (87.26)	20.35 (12.74)	8.17 (36.34)	4.12 (63.66)

Notes: Decompositions based on generalized variance decompositions, robust to alternative ordering for the vector [□ *CAN,* □ *ALC,* □ *TOB,* □ *HLT*]. All figures are estimates rounded to two decimal places — rounding errors may prevent a perfect percentage decomposition in some cases.

Figure 1. Total Alcohol Consumption and Tobacco Sales in the United States: 1930-2018



Notes: TOBSAL is total tobacco sales; TOTALC is total aggregate alcohol sales. Time plot appears in logs of actual figures observed annually.

Appendix Table A1. Tests of the Unit Root Hypothesis

	AugDickey-Fuller				*Phillips-Perron*				
	□□	□□	$Z(\square)$	$Z(t\square)$	$Z(\square_1)$	$Z(\square^*)$	$Z(t\square^*)$	$Z(\square_2)$	$Z(\square_3)$
Levels									
CAN	-1.63	-2.33	-4.14	-2.17	2.07	-5.07	-1.85	4.15*	5.02
ALC	-1.04	-2.79	-2.35	-1.42	2.15	-4.35	-2.29	2.38	3.19
TOB	-1.02	-2.15	-3.77	-0.99	2.79	-7.05**	-1.88	2.13	2.38
HLX	-0.74	-2.06	-0.90	-2.52	2.19	-4.96	-2.63	2.58	4.65
First Differences (□)									
□ *CAN*	-4.18	-6.54	-8.65	-4.81	6.07	-5.15	-5.09	4.32	5.57
□ *ALC*	-3.62	-5.77	-4.47	-3.68	4.54	-6.58	-4.39	3.98	4.02
□ *TOB*	-4.53	-5.06	-4.62	-5.69	4.36	-7.25**	-4.20	3.34	4.55
□ *HLX*	-3.45	-4.32	-5.95	-4.33	4.34	-8.91	-4.00	3.56	4.12

Notes: The sample consists of logged-annual time-series observations (1938-2019). The optimal lag used for conducting the Augmented Dickey-Fuller test statistic was selected based on an optimal criteria [Akaike's Final Prediction Error (FPE)], using a range of lags. The truncation lag parameter l used for the Phillips-Perron tests was selected using a window choice of $w(s, l) = 1 - [s/(l + 1)]$ where the order is the highest significant lag from either the autocorrelation or partial autocorrelation function of the first differenced series. Relevant test equations and related technical descriptions for all unit root testing procedures appear in Appendix: A1. Presented for *levels* tests only: *** , ** and * indicate significance at the 1%, 5% and 10% levels respectively.

Appendix Table A2. Multivariate Johansen Tests for Cointegrating Relationships Between Cancer Mortality, Alcohol and Tobacco Consumption and Health Expenditures

	Optimal Lag	Test Statistic	χ^2 Test of	
H_0 and H_1	Used in VAR	Max Eigenvalue	TraceRestriction	
No Intercepts; No Trends				
$r = 0$ $r > 0$	2	40.53**	70.69**	9.27† 12.03]
$r \leq 1$ $r = 2$		16.44	30.15	
$r = 2$ $r > 3$	13.42	13.72		
$r \leq 3$ $r = 4$		0.31	0.29	
Restricted Intercepts; No Trends				
$r = 0$ $r > 0$	2	45.44**	67.37**	15.13† [11.22]
$r \leq 1$ $r = 2$		19.14	17.21	
$r = 2$ $r > 3$		15.64	7.13	
$r \leq 3$ $r = 4$		4.42	4.47	
Unrestricted Intercepts; No Trends				
$r = 0$ $r > 0$	1	47.46**	64.55**	11.14† [9.58]
$r \leq 1$ $r = 2$		10.01	9.01	
$r = 2$ $r > 3$		9.42	11.37	
$r \leq 3$ $r = 4$		4.55	4.01	
Unrestricted Intercepts; Restricted Trends				
$r = 0$ $r > 0$	2	45.16**	62.33**	7.73† [15.64]
$r \leq 1$ $r = 2$		19.77	3.01	
$r = 2$ $r > 3$		12.04	14.73	
$r \leq 3$ $r = 4$		5.25	5.15	
Unrestricted Intercepts; Unrestricted Trends				
$r = 0$ $r > 0$	1	48.57**	68.78**	13.11† [11.94]
$r \leq 1$ $r = 2$		7.54	11.24	
$r = 2$ $r > 3$		15.64	20.37	
$r \leq 3$ $r = 4$		5.77	5.25	

Notes: r indicates the number of cointegrating relationships. The optimal lag structure for the VAR was selected by minimising the Akaike's FPE criteria. All estimated coefficients of the cointegrating vectors are available on request. Critical values are sourced from Johansen and Juselius (1990). ** indicates rejection at the 95% critical values. † and ‡ indicate significance at the 1 and 5 per cent levels, associated with a chi-square statistic testing the restriction that *ALC* [*TOB*] in the cointegrating vector [*CAN, ALC, TOB, HLT*] is equivalent to zero statistically.

Appendix Table A3. Summary of Diagnostic Tests for Equations used in Vector Error Correction Model

	Serial Correlation		*Heteroskedasticity*		*Funct Form*	*Normality*
Country/Eq	LM(1)	LM(2)	Het	ARCH	RESET	JB
□ *CAN*	0.55	1.16	0.16	0.33	4.16 (2,15)	1.13
□ *ALC*	0.78	1.33	1.36	0.06	0.56 (3,12)	1.17
□ *TOB*	1.42	1.55	0.77	1.09	3.22 (3,12)	2.14
□ *HLX*	0.74	2.12	1.27	1.27	4.67 (3,12)	2.55

Notes: Distributional properties of diagnostics are respectively: LM(1) and LM(2) as □ ^{2}(1) and □ ^{2}(2) testing for the null of no first and no fourth order serial correlation amongst the residuals; Het: a □ ^{2}(1) test based on regression of squared residuals on a constant and squares of the fitted values; a □ ^{2}(1) test for first-order ARCH effects; Ramsey's REgression Specification Error (*F*) Test with (*n, m*) degrees of freedom; and the Jarque-Bera □ ^{2}(2) LM test for normality of residuals.