

Identifying the Temporal Causal Relationship Between Alcohol and Tobacco Consumption with Aggregate Cancer Mortality over Alternative Time Horizons

Zahra N. Masih^{1,2}

¹ College of Arts & Sciences, Cornell University, Ithaca NY 14850, USA

² US Metric Association, Windsor CO 80550, USA

Correspondence: Zahra N. Masih, College of Arts & Sciences, Cornell University, Ithaca, NY 14850, USA.

Received: February 16, 2021; Accepted: March 29, 2021; Published: March 31, 2021

Abstract

Identifying the causal role of tobacco and alcohol has long been acknowledged as a critical area for developing preventative strategies in particular and public policy in general. This study utilizes a unique time series method in an effort to determine the strength of causal relationships between tobacco consumption, alcohol consumption and cancer mortality. By analysing tobacco expenditure, alcohol expenditure, while controlling for health expenditures and aggregate cancer data observed annually over an 80-year period in the US population, we consider the relevant factors explaining and potentially guiding public health concerns going forward. Our results found that while tobacco and alcohol consumption (individually) causally impact cancer mortality, alcohol consumption maintains a stronger, bidirectional impact in comparison to tobacco consumption. From this, we consider explanations from an economic, biological and epidemiological front, gauging the strength of alcohol consumption on societal wellbeing. We find alcohol consumption to be a notable causal factor in cancer mortality that has been neglected from a public policy perspective in comparison to its more mediated tobacco counterpart.

Keywords: cancer mortality, tobacco and alcohol consumption, temporal dynamics, long-run relationship, vector error-correction model

1. Introduction

Alcohol and tobacco significantly contribute to increased cancer risk. As of 2018, the World Health Organization identified tobacco and alcohol as a part of the top behavioural contributors to cancer risk, a cohort responsible for approximately one-third of all cancer deaths (WHO, 2018).

The National Cancer Institute classifies alcohol as a carcinogen capable of detrimental, if not fatal, health consequences including immune susceptibility or death. Its consumption has been causally linked to breast, pharynx, liver and colon cancer (IARC, 2010). Studies indicate alcohol maintains a dose-response relationship with cancer, where higher volumes of alcohol consumed potentially increase cancer risk (Kirpalani, 2017; de Menezes et al., 2013). Meta-analyses on alcohol intake and cancer data indicate that alcohol promotes immunosuppressive processes in the body thus contributing to a higher probability in increased risk (Bagnardi, et al., 2001). Alcohol consumption is therefore a potential threat to cancer risk, where the impact of its consumption on an aggregate scale is yet to be discovered.

The National Cancer Institute also classifies tobacco as a leading cause of cancer and cancer deaths (NIH, 2017). The carcinogenic chemicals released from smoking tobacco may genetically mutate the DNA and alter receptor expression of bodily cells, promoting proliferative or suboptimal performance of cellular function (Desrichard et al, 2018; Peterson et al., 2020). In regard to long-term risks, tobacco is associated to lung, pharynx, larynx, liver, stomach, and pancreatic cancers (American Cancer Society, 2018). While tobacco maintains a causal relationship with lung cancer, it also increases risk for various other cancers as seen with frequent tobacco users (Gandini et al, 2008). Tobacco consumption is of interest due to its versatile and detrimental health effects.

It is evident that in both cases of alcohol and tobacco, mortality and health are correlatively linked to consumption. One study identifies that altering the consumption of alcohol and tobacco per capita may impact population mortality in alcohol-related diseases, but such methods have currently limited application with cancer. The research reveals that a correlated relationship exists between decreasing both population-level alcohol/tobacco

consumption with cancer mortality reduction in Australia. They further consider such relationships can evidently be altered or regulated utilizing public health policy measures (Jiang et al., 2018).

The purpose of this paper is to characterize the impact of aggregate alcohol and tobacco consumption on overall cancer mortality in the United States. Stemming from the correlated associations identified in Jiang et. al., this study aims to distinguish a causative relationship, if any, between alcohol/tobacco and cancer mortality, and if so identify which of the substances have a stronger effect on general cancer risk within the population. This study will not only utilize lag structures to characterize population dynamics, but will also test a formal hypothesis of whether alcohol consumption and/or tobacco consumption lead to causal effects (both short and long-run) in cancer mortality. This is of particular importance to suggesting directions for American public health policies and contributing to the growing understanding of alcohol and tobacco’s societal impact.

The contributions of this paper are as follows: (i) To the author’s knowledge, this paper provides a seminal attempt at understanding the dynamic causal relationships that co-exist between cancer mortality and its determinants over time; (ii) The methods we apply unveils more than just simple correlations; (iii) Finally we examine the relationships over alternative time periods that is imperative for policy making both from an epidemiological and public health standpoint.

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.

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, a representation in the form of an error-correction representation can be formulated. This formulation results in expressing the dependent variable as a function of the degree of departure in the long-run relationship, in addition to fluctuations in the other factors included in the representation. 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.

Using the notion that factors related to cancer mortality, such as alcohol and tobacco consumption, can co-exist 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:

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

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. (Note 1)

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 annually for the US and covers the period 1938 to 2018. For all transformations, annual percentage growth rate of GDP at market 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/>).

The first step in testing for the existence of a long-run relationship is to identify the properties of the data at a univariate level. These tests verify to what degree the time series under analysis are integrated. We use two tests based on different null hypotheses. The KPSS test due to Kwiatkowski, Phillips, Schmidt and Shin, 1992 is used for testing that the null hypothesis of the series is stationary around a deterministic linear trend. The Augmented Dickey-Fuller or ADF(max) tests due to Dickey and Fuller, 1979 and 1981 tests the null hypothesis that the series in question are non-stationary. 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 $I(1)$. 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 $I(1)$. (Note 2)

As we derived evidence supporting common integrational properties for the factors, 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 relationship since $r = 0$ is clearly rejected in favour of $r = 1$; but $r \leq 1$ cannot be rejected by the 95% critical values. Hence, we find evidence of a number of common trends within the system. 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 Δx or Δz indicates the short-run source as this captures restriction on purely the changes of the factors; and the joint ECT and Δx or Δz indicates the joint source emanating from both short and long-run movements in then factors. 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 ΔALC_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 analyzed in the discussion. The error correction term or ECT ($\epsilon_t = CAN_t - c - ALC_t - TOB_t - HLT_t$) being significant means that the long-run adjustment parameter given by d allows the dependent variable to adjust to restore equilibrium, implying that alcohol and tobacco consumption leads cancer

mortality. Furthermore, the unique finding that changes in alcohol and tobacco consumption positively and significantly affects cancer mortality as indicated by the positive coefficients on ΔALC_t and ΔTOB_t , is noteworthy.

As the vector error correction model is estimated using least squares regression, we tested its assumptions using a number of standard tests already established in the literature. The results of these tests are in summarized form in Appendix Table: A3. These checks are for serially uncorrelated errors of first and second orders, constancy of variance of errors, misspecification of function form and non-normality of errors respectively. In conclusion, none of the tests could detect significant evidence of a departure from standard assumptions. Such tests for robustness provide statistical confidence when it comes to conducting appropriate inference.

4. Discussion and Policy Implications

Our data analyzes the causal relationships between tobacco consumption and alcohol consumption on cancer mortality, controlling for health expenditure. The results in Table 2 indicate the strength of short and long run Granger-causing relationships between the variables tested. Our findings suggest that (i) there exists significant short and long run causation from alcohol consumption to cancer mortality, (ii) a unidirectional relationship exists from tobacco consumption to cancer mortality, with its impact being comparatively less than that of alcohol, and (iii) cancer, in the long run, significantly causes changes in health expenditure.

The strength of alcohol consumption Granger causing cancer mortality may be due to classified carcinogen, acetaldehyde. When alcohol is metabolized, acetaldehyde exists temporarily before being expelled by the body. With increased alcohol consumption, a higher concentration of acetaldehyde may damage organ functionality and weaken immunity (Seitz et al., 2007; Paton, 2005). One study indicates acetaldehyde can increase cancer-causing cellular behaviors by actively interfering in replication and repairment of DNA, contributing to the proliferation of damaged cells (Kwo et al., 1998). This, in general indicates that the consumption of alcohol increases the concentration of such carcinogenic substances in the blood, contributing to decreased cellular functions and therefore increased cancer risk.

Acetaldehyde also exists in tobacco, in which it enters the body through the burning and inhalation of tobacco smoke. Acetaldehyde's critical role in tobacco consumption may explain the unidirectional relationship it maintains with cancer mortality in the long run. Despite such findings, our results suggest alcohol consumption has a stronger causal relationship to cancer mortality in comparison to tobacco consumption. This, too, can be explained by the varying levels of acetaldehyde in both substances. Though both tobacco and alcohol contain acetaldehyde, alcohol may contain more acetaldehyde over a greater frequency compared to tobacco. One study demonstrates this by speaking toward the impact of alcohol consumption on bodily concentration. They identified that ethanol concentrations over a range of 53 to 33 mM/acetaldehyde can lead to an average of 42.7 μ M acetaldehyde in the blood. Whereas, cigarettes maintain approximately 1 μ g of acetaldehyde, and its levels were found almost undetectable in the blood at moderate consumption (Korsten et al., 1975). Tobacco may also maintain a weaker relationship with cancer mortality due to the aggregate sample of data being used in the model. Our results consider causation for overall cancer mortality instead of targeting specific cancers. This can be of particular importance given the location of acetaldehyde in the consumptive process. With mainstream smoke found in commercialized cigarettes, acetaldehyde tends to settle in the upper respiratory tract and mouth before being expelled by the body (Seeman et al., 2002). Whereas, alcohol (and therefore acetaldehyde) may metabolize in the liver, pancreas or GI tract (NIH, 2007). In looking at overall cancers, we generalize the findings to a broader scale which may not highlight the individual impacts tobacco/alcohol may have on specific cancers. Furthermore, epidemiological studies have both identified the causal relationship between tobacco and lung cancer, but also introduce increased risk in cancer for frequent tobacco users, where stronger probabilities of cancer exist for the lung, laryngeal, pharyngeal, oral and upper digestive regions (Gandini et al., 2008). Tobacco may be less versatile in the types of range of cancers it impacts compared to that of alcohol, contributing to a weaker relationship.

The degree of alcohol explaining cancer mortality (in the long-run) and its maintenance of a bidirectional relationship indicates that both cancer and alcohol consumption impact each other. The factors influencing alcohol consumption on cancer mortality have been discussed among public health policy and biological aspects, however little has been focused on how cancer mortality impacts alcohol consumption. Though weak in its relation, unnoted factors may be influencing this result, such as stress or psychological factors. It may be possible that the stress underlying the diagnoses of cancer may coerce patients/affected loved ones to drink more. Alcohol has been ingested with the intention to reduce stress as proposed by the tension reduction theory. One study highlights this by showing that college students with high levels of anxiety drink more than those with low levels of anxiety. This

application of stress-inducement and drinking supports the relation between cancer, a severe stress-causing factor, and alcohol consumption (Kalodner et al., 1989). This has been further highlighted in suggested changes to medical protocols to include mental health services, as one study considered the impact of a medical patient's stress (Simonton, 1981). Regardless of the bidirectionality of cancer mortality to alcohol, results indicate that this has an overall greater effect than the relationships tested between cancer mortality and tobacco.

Finally, our results suggest that cancer Granger causes health care expenditure in the long run (an increase in cancer mortality increases expenditure). Since increases in alcohol consumption Granger causes cancer mortality, it may also be an indirect factor contributing to higher health expenditure. Adjusting public health policies to address alcohol consumption will therefore benefit society on both a health and financial scale by not only controlling for cancer risk, but also initiating efficient healthcare spending. As noted in Figure 1, post 1980 tobacco spending has maintained a steady decline in aggregate spending whereas alcohol remains increasing. Currently there is less legislation and intervention strategies for alcohol in comparison to tobacco. Stemming from tobacco being nationally emphasized as a health risk factor, regulations such as discontinuing cigarette rationing, smoke-free locations and restricted promotional advertising were employed (Stratton et al., 2001). Whereas, alcohol has experienced limited regulation aside from a mandated drinking age. Despite its carcinogenicity, regulations surrounding alcohol are said to be "self-imposed" and "voluntary at the federal level" (American Addiction Centers, 2020). Given this, it is an even more pronounced reason explaining the intensity of alcohol on societal expenditure and consequently cancer risk.

4. Limitations and Conclusions

Our findings identify alcohol to be a dominant causative factor on cancer mortality compared to tobacco consumption. Not only does this hold relevance to the biological factors associated with aggregate health, but it also addresses the impact potential policies could have on societal finances. This further characterizes the positive temporal associations between tobacco and alcohol discussed in Jiang et. al. by applying causative analysis on a unique population. However, this analysis is not without its limitations. While this study has addressed two variables on cancer mortality, more emphasis needs to be directed toward the impact of gender differences, ethnicity and ages of the sample. This is necessary given how mortality risk associated to frequent tobacco/alcohol consumption varies with age (Thun et al., 1997). Our study also looks at aggregate cancer data which may impact how tobacco and alcohol consumption affect cancer on a biological scale. For example, certain cancers may be more prone to impact on tobacco than alcohol unnoted in the study (lung cancer impacted by tobacco smoke). Future directions should not only employ the utilization of diverse samples/cancers, but also aim toward suggesting policy measures ample for the regulation of alcohol consumption.

Acknowledgements: The author would like to acknowledge the support of Prof. Mansur Masih for his encouragement and very valuable comments on an earlier draft of this paper.

References

- American Cancer Society. Health Risks of Smoking Tobacco. *Cancer.org*.
- Bagnardi, V., Blangiardo, M., La Vecchia, C., & Corrao, G. (2001). Alcohol consumption and the risk of cancer: A meta-analysis. *Alcohol Research and Health*, 25(4), 263-270. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/11910703/>
- de Menezes, R. F., Bergmann, A., & Thuler, L. C. S. (n.d.). Alcohol consumption and risk of cancer: A systematic literature review. *Asian Pacific Journal of Cancer Prevention*, 14(9), 4965–4972. <https://doi.org/10.7314/apjcp.2013.14.9.4965>
- Desrichard, A., Kuo, F., Chowell, D., Lee, K.-W., Riaz, N., Wong, R. J., Chan, T. A., & Morris, L. G. T. (2018). Tobacco Smoking-Associated Alterations in the Immune Microenvironment of Squamous Cell Carcinomas. *Journal of the National Cancer Institute*, 110(12), 1386-1392. <https://doi.org/10.1093/jnci/djy060>
- Dickey, D. A., & Fuller W. A. (1979), Distribution of the estimators for autoregressive time-series with a unit root, *Journal of the American Statistical Association*, 74, 427-431. <https://doi.org/10.2307/2286348>
- Dickey, D. A., & Fuller W. A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 1057-1072. <https://doi.org/10.2307/1912517>
- Editorial Staff. (2020). Rules & Regulations About Marketing Alcohol. *American Addiction Centers*. Retrieved from <https://www.alcohol.org/laws/marketing-to-the-public/>
- Engle, R. F., & Granger, C. W. J. (1987). Cointegration and error correction: Representation, estimation, and testing. *Econometrica*, 55, 251-276. <https://doi.org/10.2307/1913236>

- Gandini, S., Edoardo, B., Simona, I., Mathieu, B., Lowenfels, A. B., Patrick, M., & Peter, B. (2008). Tobacco Smoking and Cancer: A Meta-Analysis. *International Journal of Cancer*, 122(1), 155-164. <https://doi.org/10.1002/ijc.23033>
- Granger, C. W. J. (1988), Some recent developments in a concept of causality, *Journal of Econometrics*, 39, 199-211. [https://doi.org/10.1016/0304-4076\(88\)90045-0](https://doi.org/10.1016/0304-4076(88)90045-0)
- 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. Retrieved from <https://www.ncbi.nlm.nih.gov/books/NBK326568/>
- Jiang, H., Livingston, M., Room, R., Chenhall, R., & English, D. R. (2018). Temporal Associations of Alcohol and Tobacco Consumption With Cancer Mortality. *JAMA Network Open*, 1(3). <https://doi.org/10.1001/jamanetworkopen.2018.0713>
- Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12, 231-254. [https://doi.org/10.1016/0165-1889\(88\)90041-3](https://doi.org/10.1016/0165-1889(88)90041-3)
- Johansen, S., & Juselius, K. (1990). Maximum likelihood estimation and inference on cointegration with applications to money demand. *Oxford Bulletin of Economics and Statistics*, 52, 169-210. <https://doi.org/10.1111/j.1468-0084.1990.mp52002003.x>
- Kalodner, C. R., Delucia, J. L., & Ursprung, A. W. (1989), An examination of the tension reduction hypothesis: The relationship between anxiety and alcohol in college students, *Addictive Behaviors*, 14(6), 649-654. [https://doi.org/10.1016/0306-4603\(89\)90007-5](https://doi.org/10.1016/0306-4603(89)90007-5)
- Kirpalani, P. (2017). Alcohol & Cancer. Mimeo, University of North Carolina. Retrieved from <http://www.researchgate.net/publication/321225698>
- Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54(1-3), 159-178. [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y)
- Korsten, M. A., Matsuzaki, S., Feinman, L., & Lieber, C. S. (1975). High Blood Acetaldehyde Levels after Ethanol Administration - Difference between Alcoholic and Nonalcoholic Subjects. *New England Journal of Medicine*, 292(8), 386-389. <https://doi.org/10.1056/NEJM197502202920802>
- Kwo, P. Y., Ramchandani, V. A., O'Connor, S., et al. (1998). Gender differences in alcohol metabolism: Relationship to liver volume and effect of adjusting for body mass. *Gastroenterology*, 115, 1552-1557. [https://doi.org/10.1016/s0016-5085\(98\)70035-6](https://doi.org/10.1016/s0016-5085(98)70035-6)
- National Cancer Institute. (2017). Tobacco. *National Institute of Health*. Retrieved from <https://www.cancer.gov/about-cancer/causes-prevention/risk/tobacco>
- National Institute on Alcohol Abuse and Alcoholism. (2007). National Alcohol Metabolism: An Update, *U.S. Department of Health and Human Services*, 72. <https://pubs.niaaa.nih.gov/publications/aa72/aa72.htm>
- Nelson, C. R., & Plosser, C. (1982). Trends and random walks in macroeconomic time series, *Journal of Monetary Economics*, 10, 129-62. <https://doi.org/10.1016/j.jmacro.2011.10.001>
- Paton, A. (2005). Alcohol in the Body. *The BMJ*. 330(7482), 85-87. <https://doi.org/10.1136/bmj.330.7482.85>
- Perron, P. (1988). Trends and random walks in macroeconomic time series, *Journal of Economic Dynamics and Control*, 12, 297-332. [https://doi.org/10.1016/0165-1889\(88\)90043-7](https://doi.org/10.1016/0165-1889(88)90043-7)
- Peterson, L. A., Balbo, S., Fujioka, N., Hatsukami, D. K., ... Villalta, P. W. (2020). 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*. <https://doi.org/10.1158/1055-9965.EPI-19-1356>
- Phillips, P. C. B., & Perron P. (1988). Testing for a unit root in time series regression, *Biometrika*, 75, 335-346. <https://doi.org/10.2307/2336182>
- Seeman, J., Dixon, M., & Haussman, H. (2002). Acetaldehyde in mainstream tobacco smoke: formation and occurrence in smoke and bioavailability in the smoker. *Chemical Research in Toxicology*, 15(11), 1331-1350. <https://doi.org/10.1021/tx020069f>
- Seitz, H. K., & Becker, P. (2007). Alcohol Metabolism and Cancer Risk. *Alcohol Research & Health*, 30(1), 38-

47. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/17718399/>

Simonton, O. C., & Matthews-Simonton S. (1981). Cancer and Stress Counselling the Cancer Patient. *The Medical Journal of Australia*, 1(13), 679-683. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/7278751/>

Thun, M. J., Peto, R., Lopez, A. D., Monaco, J. H., Henley, S. J., Heath, C. W., & Doll, R. (1997). Alcohol Consumption and Mortality among Middle-Aged and Elderly U.S. Adults. *The New England Journal of Medicine*, 337(24). <https://doi.org/10.1056/NEJM199712113372401>

World Health Organization. (2018). Cancer. *World Health Organization*. Retrieved from <https://www.who.int/news-room/fact-sheets/detail/cancer>

Notes

Note 1. Causality is a topic of great debate in econometrics and statistical science. In time series, the short-run is captured through changes in the factors, whereas the long-run is captured with the factors in their level form. For a more detailed elaboration of this debate, see Granger (1988).

Note 2. Factor that appear or are used in their level form are referred to as their ‘long-run’ component. See Granger (1988) for further details.

Table 1. Summary of Stationarity Properties

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

Notes: An *I*(1) indicates that the variable is stationary in its first difference form. All variables are observed annually from 1938 to 2018 inclusive.

Table 2. Test Results for Causality Between Alcohol/Tobacco Consumption with Cancer Mortality based on a VECM Source of Causation

Equation	Short-RunJoint:			Short-RunJoint:		
	ECT	Δz	(ECT & Δz)	ECT	x	(ECT & Δx)
	$\Delta z = ALC$			$\Delta x = TOB$		
ΔCAN	-3.77**	2.25**	3.17**	-1.77*	1.63	1.95*
	$\Delta z = CAN$			$\Delta x = TOB$		
ΔALC	-1.28	1.68	1.84*	-2.07*	0.97	2.18*
	$\Delta z = CAN$			$\Delta x = ALC$		
ΔTOB	-0.86	0.87	1.19	-2.15*	1.27	2.32**
	$\Delta z = CAN$			$\Delta x = ALC$		
ΔHLX	-2.92**	0.44	1.87	-0.47	0.98	1.02

Notes: The 'ECT' indicates to *t*-statistics testing the null that the error-correction term is statistically insignificant. The error correction terms are formulated with the use of Johansen tests. 'Short-Run' are F-statistics and indicates significance levels associated testing the null that lags of the variable indicated are jointly equivalent to zero; 'Joint' refers to significance levels associated with a joint *F*-test of the ECT and all lags of the variable indicated, being statistically insignificant. A full summary of diagnostic tests conducted for each of these equations appear in Appendix Table: A3. ***, ** indicates statistical significance at the 1, 5 per cent levels.

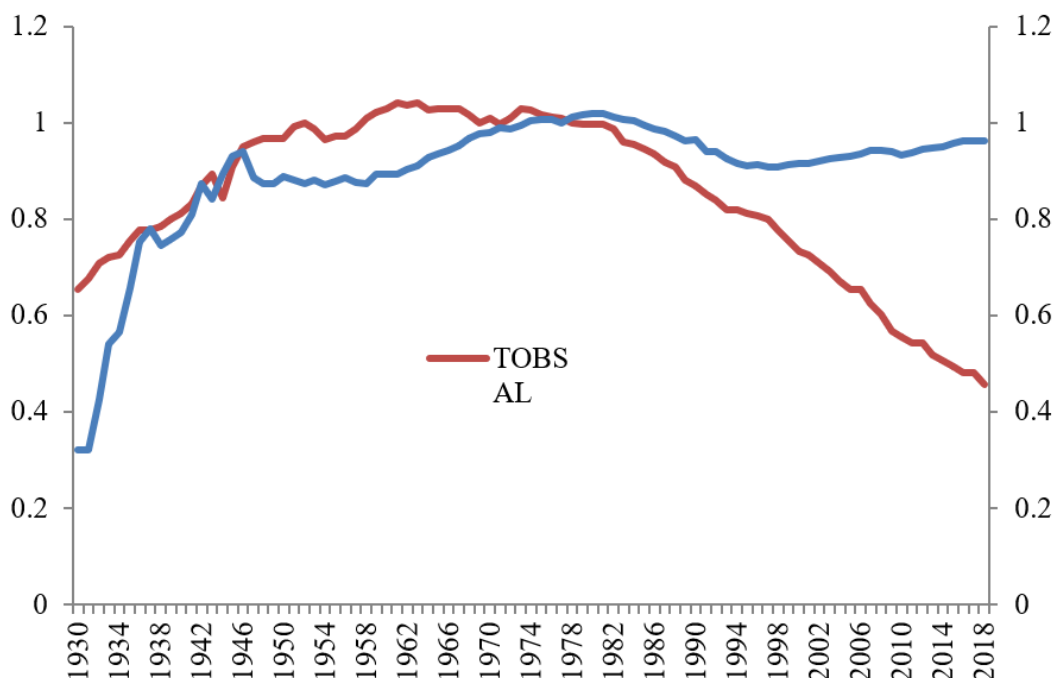


Figure 1. Total Alcohol Consumption and Tobacco Sales in the United States: 1930-2018

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

	<i>Aug Dickey-Fuller</i>				<i>Phillips-Perron</i>				
			$Z(\)$	$Z(t\)$	$Z(\ 1)$	$Z(\ *)$	$Z(t\ *)$	$Z(\ 2)$	$Z(\ 3)$
<i>Levels</i>									
<i>CAN</i>	-1.63	-2.33	-4.14	-2.17	2.07	-5.07	-1.85	4.15*	5.02
<i>ALC</i>	-1.04	-2.79	-2.35	-1.42	2.15	-4.35	-2.29	2.38	3.19
<i>TOB</i>	-1.02	-2.15	-3.77	-0.99	2.79	-7.05**	-1.88	2.13	2.38
<i>HLX</i>	-0.74	-2.06	-0.9	-2.52	2.19	-4.96	-2.63	2.58	4.65
<i>First Differences (D)</i>									
<i>DCAN</i>	-4.18	-6.54	-8.65	-4.81	6.07	-5.15	-5.09	4.32	5.57
<i>DALC</i>	-3.62	-5.77	-4.47	-3.68	4.54	-6.58	-4.39	3.98	4.02
<i>DTOB</i>	-4.53	-5.06	-4.62	-5.69	4.36	-7.25**	-4.2	3.34	4.55
<i>DHLX</i>	-3.45	-4.32	-5.95	-4.33	4.34	-8.91	-4	3.56	4.12

Notes: Akaike’s Final Prediction Error was used to select the optimal lag length for the ADF test; and the truncation lag parameter for the PP tests is based on window choice of $w(s, l) = 1 - [s/(l + 1)]$. The order refers to the lag with the highest significance from either the autocorrelation or partial autocorrelation function. ***, ** and * indicate significance at the 1%, 5% and 10% levels respectively. See Appendix A1 for technical details.

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

H_0 and H_1	Optimal Lag		Test Statistic		χ^2 Test of
	Used in VAR		Max Eigenvalue	Trace	Restriction
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. Akaike’s FPE was used to select the optimal lag structure for the VAR, with critical values 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 long-run vector [CAN , ALC , TOB , HLT] is statistically insignificant.

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

Country/Eq	Serial Correlation		Heteroskedasticity		Funct Form	Normality
	LM(1)	LM(2)	Het	ARCH	RESET	JB
ΔCAN	0.55	1.16	0.16	0.33	4.16 (2,15)	1.13
ΔALC	0.78	1.33	1.36	0.06	0.56 (3,12)	1.17
ΔTOB	1.42	1.55	0.77	1.09	3.22 (3,12)	2.14
ΔHLX	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 $\chi^2(1)$ and $\chi^2(2)$ testing for the null of no first and no fourth order serial correlation amongst the residuals; Het: a $\chi^2(1)$ test based on regression of squared residuals on a constant and squares of the fitted values; a $\chi^2(1)$ test for first-order ARCH effects; Ramsey's REgression Specification Error (F) Test with (n , m) degrees of freedom; and the Jarque-Bera $\chi^2(2)$ LM test for normality of residuals.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).