Published in final edited form as:

Alcohol Clin Exp Res. 2011 January ; 35(1): 108–115. doi:10.1111/j.1530-0277.2010.01327.x.

Alcohol and liver cirrhosis mortality in the US: Comparison of methods for the analyses of time-series panel data models

Yu Ye, M.A. and

Alcohol Research Group, 6475 Christie Avenue, Suite 400, Emeryville, CA 94608 USA

William C. Kerr, Ph.D.

Alcohol Research Group, 6475 Christie Avenue, Suite 400, Emeryville, CA 94608 USA

Abstract

Background—To explore various model specifications in estimating relationships between liver cirrhosis mortality rates and per capita alcohol consumption in aggregate level cross-section time-series data.

Methods—Using series of liver cirrhosis mortality rates from 1950 to 2002 for 47 US states, the effects of alcohol consumption were estimated from pooled ARIMA models and four types of panel data models: GEE, GLS, fixed effect and multi-level models. Various specifications of error term structure under each type of model were also examined. Different approaches controlling for time trends and for using concurrent or accumulated consumption as predictors were also evaluated.

Results—When cirrhosis mortality was predicted by total alcohol, highly consistent estimates were found between ARIMA and panel data analyses, with an average overall effect of 0.07–0.09. Less consistent estimates were derived using spirits, beer and wine consumption as predictors.

Conclusions—When multiple geographic time series are combined as panel data, none of existent models could accommodate all sources of heterogeneity such that any type of panel model must employ some form of generalization. Different types of panel data models should thus be estimated to examine the robustness of findings. We also suggest cautious interpretation when beverage-specific volumes are used as predictors.

Keywords

cirrhosis; mortality; alcohol consumption; time series; panel data

INTRODUCTION

There is a long tradition of studying the relationship between alcohol consumption and mortality rates using aggregate time-series data at the country or state (province) level (Norström and Ramstedt, 2005; Rehm, 2001). Aggregate data allows investigation of indirect relationships, which normally cannot be modeled using individual level data, such as traffic accidents or homicides caused by others' drinking, and facilitates comparative analyses of structural and societal effects (Norström and Skog, 2001). While earlier research mainly focused on data from a single country, data from multiple countries/states have

become more available in recent years, not only increasing the sample size for analysis but making cross-regional comparisons possible.

Modeling time-series data across multiple geographic units has typically taken one of two general approaches in alcohol epidemiology. In the first approach, time-series models of each aggregate unit are estimated separately and the individual estimates are then pooled. Autoregressive integrated moving average (ARIMA) methods are most commonly used. This approach was applied in European Comparative Alcohol Study of 14 EU countries on the relationship between per capita alcohol consumption and mortality rates for liver cirrhosis (Ramstedt, 2001b), heart disease (Hemström, 2001), accidents (Skog, 2001), suicide (Ramstedt, 2001a), homicide (Rossow, 2001) and all-cause mortality (Norström, 2001), and similarly to study mortality rates of Canadian provinces (Ramstedt, 2003). The ARIMA model is used to reduce the risk of spurious trend relationships (Norström and Skog, 2001; Skog, 1988), and to fully specify the structure of the error term to accommodate cross-country/state differences in variance (i.e. heteroskedasticity) and correlation of error terms.

Alternatively, multiple series of the same dependent variable across geographic units can be analyzed simultaneously using cross-sectional time-series approaches, also called panel data models. Panel data models greatly improve efficiency through increased sample size when the data are pooled and analyzed together. Applications of panel data models in the alcohol field have involved a variety of specifications. For example, using fixed effects models, Gruenewald and Ponicki (1995) analyzed a panel of 50 US states from 1976 to 1985 including cirrhosis mortality rates and the consumption of spirits, wine and beer, and recently, re-examined an updated US state panel, also including alcohol taxes, comparing fixed effects and random effects models (Ponicki and Gruenewald, 2006). Another method, the generalized least-square procedure (GLS), was used by Kerr, Fillmore and Marvy (2000) to examine alcohol and cirrhosis mortality relationships of five countries, allowing for heteroskedastic error structure, cross-sectional correlation and panel-specific first-order autocorrelation. A similar approach was applied by Her and Rehm (1998) to study all-cause mortality. To our knowledge, there has been only one methodological study conducting comprehensive comparisons of different models of alcohol-related mortality rates. Using data of 15 European countries, Gmel, Rehm and Frick (2001) compared estimates from Generalized Estimating Equation (GEE), GLS and hierarchical linear models, with a variety of error term specifications explored for specific models. In addition to examining model specifications, comparisons were made to examine whether a linear time trend could be used for removing common time trends.

In the current analysis, a similar methodological study comparison is undertaken for liver cirrhosis mortality rates in 47 US states from 1950 to 2002. In addition to the panel data specifications previously examined (Gmel et al., 2001), we will further explore how estimates from panel models compare with results pooled from ARIMA models of each state. We will also examine the effect of consumption from spirits, wine and beer, in addition to total alcohol, as differential relationships with liver cirrhosis by beverage type have been found in earlier research, where significant effects with spirits, but not beer or wine have been found in US data (Gruenewald and Ponicki, 1995; Roizen et al., 1999; Terris, 1967). While the control of historic time will be examined with linear and quadratic terms, first-differenced dependent variables and predictors, a method used in ARIMA models, will also be used. Finally, to address the potential time lag problem sometimes observed in modeling chronic diseases and particularly liver cirrhosis (Norström, 1987), we will compare models utilizing concurrent alcohol consumption with those utilizing a distributed lag.

MATERIALS AND METHODS

Data and Measures

Liver cirrhosis mortality rates were taken from the National Center for Health Statistics Compressed Mortality File series (National Center for Health Statistics, 2000; National Center for Health Statistics, 2003; National Center for Health Statistics, 2004) for the years 1968 to 2002 and from the Vital Statistics of the United States for each year from 1950 to 1967 (National Center for Health Statistics, 1961–1969; National Office of Vital Statistics, 1954–1960). The number of deaths in each specific age group along with population estimates for that group were used to directly age standardize rates to the 2000 US population. Population estimates for each age group by state are from the same sources of mortality data from 1968 to 2002. The 1968 population estimates, together with those from the 1950 and 1960 Census, were then used to calculate population estimates for years in between, assuming a constant rate of increase (or decrease). Liver Cirrhosis was defined across several revisions of the ICD using code 581 for ICD 6 and 7, code 571 for ICD 8 and 9 and codes K70, K73, and K74 for ICD 10. Data for per capita consumption of ethanol from beer, wine, spirits (liters of ethanol) and combined for total alcohol are described in detail in Kerr et al (Kerr et al., 2006). Four states (Alaska, Hawaii, Mississippi and Oklahoma) that do not have complete 1950–2002 series data of either mortality or alcohol consumption were excluded from the current analysis to retain balanced panel as required for fitting some specific models. The resulting panel includes 47 states (including DC) over 53 years.

ARIMA modeling

The ARIMA modeling technique was used to evaluate relationships between cirrhosis mortality and per capita alcohol consumption for each state separately. The first difference of log mortality rates and consumption series was used to remove unit roots. Semi-log models were fit, regressing log transformed age-standardized mortality rates on total alcohol consumption volume, with separate models where spirits, wine and beer volume were entered together. A distributed lag structure developed by Norström (1987) and employed by others (Ramstedt, 2001b; Ramstedt, 2003) was used to transform the volume series to model the cumulative alcohol effect, truncated after 5 years. The error structure was modeled using autoregressive (AR) and moving average (MA) terms as appropriate in each model. The final models were tested for any remaining autoregressive error structure using the Ljung-Box Q statistic and for normality of residuals in the final model using the Jarque-Bera test and were estimated using EViews version 5 (EViews, 2004). Two approaches were used to pool the effects across states for a combined estimate. The first was the simple average as employed in similar ARIMA analysis in European countries (Norström and Skog, 2001; Ramstedt, 2001b). The second approach applies a method from meta-analysis by deriving the pooled fixed and random effect estimates as well as performing a test of homogeneity to determine which effect is more appropriate (Sutton et al., 2000). The meta-analytic pooled effect size is a weighted average taking account of the standard error of individual state estimates.

Panel data modeling

Since a number of model characteristics were to be tested, we arranged our panel data analysis in two steps. First, we compared the estimates for total alcohol consumption from different models examining model specification, time trend control methods and the inclusion of a distributed lag of consumption. Next, we investigated estimates of spirits, wine and beer consumption, entered either one at a time or simultaneously, across different panel model specifications.

Panel model specifications—Four different panel models were evaluated separately: GEE, GLS, fixed effects and multi-level models. GEE is an extension of the Generalized Linear model (GLM) which constructs a general likelihood-based estimating equation for outcome variables belonging to exponential family distribution. GEE relaxes the independence assumption of GLM and allows modeling of various forms of within-panel correlation. In this study, we examined three types of error correlation structures: fixed within-panel correlation across time (exchangeable correlation), first-order autoregression (AR1) and second-order autoregression (AR2).

GLS is a generalized method specifying the variance-covariance matrix of the error structure. In addition to differences in estimation methods, GLS varies from GEE in specification of error structure, allowing the modeling of differences in variances across panels (heteroskedasticity) as well as cross-panel correlation. Six types of error specification were examined in this exercise; (1) exchangeable correlation, (2) exchangeable correlation plus AR1, (3) heteroskedastic, (4) heteroskedastic and common AR1, (5) heteroskedastic and panel specific AR1, and (6) heteroskedastic and panel specific AR1 plus panel correlation. Note that the first two models in GLS (exchangeable correlation without or with AR1) only specify within-panel error structure and are called random effects models in the Econometric literature. Here we put them together with other GLS models as their estimation methods all belong to the same family.

Fixed effect models explicitly control for cross-state differences in mortality using a state fixed effect, so that the coefficient estimates are within-panel effects, leaving out cross-sectional relationships. The Hausman specification test between fixed effect and random effects estimates (the first 2 models of GLS) is typically used to examine whether estimates from random effects are inconsistent. GEE, GLS and fixed effect models were fit using STATA version 10 (Stata Corp, 2007).

The three models described above can all be called “population averaged” models in which a fixed coefficient for the effect of alcohol consumption on mortality rates is specified and estimated. The multi-level model allows both mortality rates and the mortality/alcohol relationship to vary across panels (i.e. states) by specifying a random intercept and random slope. In addition to generating an overall average effect, the predicted effect of alcohol on mortality for each state can thus be estimated, which allows us to test the concordance of state-specific effects derived from the multilevel model with those from ARIMA models. We examined three specifications under multilevel models: (1) random alcohol effect, (2) random alcohol effect plus AR1 and (3) random alcohol effect and time trend, with all three models also specifying random intercept. Under various names (mixed effect, random coefficient or hierarchical linear model), multilevel models can be estimated by different software packages. In current study, we used SuperMix (Hedeker et al., 2008) for its powerful capabilities for modeling auto-correlated error structures.

Controlling for time trend—Two approaches were employed to adjust for time trends, which could confound the relationship between mortality and alcohol. The first approach used the linear and quadratic terms of time trend as predictors. This assumes a common time trend across states. The second approach, using the first difference of both mortality rates (log transformed) and the consumption volume, as in ARIMA modeling, can be expected to remove state-specific time trends.

Lag structure of alcohol volume—While the accumulative alcohol effect on cirrhosis mortality at aggregate levels was established and tested using data from Nordic countries (Norström, 1987; Skog, 1985), earlier panel data studies on US cirrhosis mortality tended to use either current year consumption (Gruenewald and Ponicki, 1995) or a 1-year lag

together with current volume (Ponicki and Gruenewald, 2006). This study will thus compare the estimates using current volume and a distributed lag structure of consumption as described above.

RESULTS

Figure 1 shows the age-standardized liver cirrhosis mortality rates (per 100,000) for the five largest states (based on average population) from 1950–2002 (Figure 1A) and five smallest states (Figure 1B) with both including overall rates for the US. First, a quadratic time trend was observed universally for the US as a whole, for all five large states, and for the five small states. Second, stable time series were seen for large states, while small states' rates moved up and down over time. These differences in variances across states suggest a heteroskedastic error structure and that the autocorrelation of error terms for small states might be better modeled by moving average (MA) terms, rather than autoregressive (AR) terms which are likely to be appropriate for large states.

We first fit a semi-log ARIMA model on log-transformed cirrhosis mortality rates for each state, predicted by alcohol consumption volume using a distributed lag to account for the accumulative effects of alcohol. Table 1 shows the pooled alcohol effect sizes from both the simple average and meta-analytic methods. Since the stability of mortality time series varied greatly across states as, shown in Figure 1, so are expected cross-state estimates. For this reason, a weighted average like pooled effect from meta-analysis would be more appropriate as it takes account of the standard error estimates of individual state effects. The average pooled effect of total alcohol from meta-analysis was 0.083 (fixed effect, given non-significant homogeneity test), which was significantly different from zero ($p < 0.01$). In a separate model, spirits, wine and beer volume were entered together to predict mortality rates. The average meta-analytic random effect was 0.238 ($p < 0.01$) from spirits and an insignificant -0.023 from beer. For wine the average meta-analytic fixed effect was -0.095 ($p < 0.05$).

We then estimated panel models of cirrhosis mortality rates predicted by total alcohol consumption, examining various model specifications, time trend adjustments and lag structures. The coefficient estimates are arranged to be presented from five models as shown Table 2. Model 1 is a raw model with log-transformed mortality rates predicted by current-year alcohol volume. Model 2 and 3 are adjusted for time trend, with M2 controlling for linear and quadratic time trends and M3 modeling the first-difference of both mortality rates and volume. Model 4 and 5 further examine the lag structure of volume together with time trend adjustments (M4 quadratic time trend and M5 difference). Each model was fitted with GEE, GLS, fixed effect and multilevel methods, with various error specifications under each.

As shown in Table 2, different panel model specifications for the raw model (M1) generate quite consistent estimates of total alcohol consumption's effect on cirrhosis mortality rates, ranging from 0.06 to 0.11, all significantly larger than zero. Similar and quite consistent results were observed when quadratic time trends were controlled (M2), with effect sizes ranging from 0.06 to 0.09, and there was little difference using accumulated volume when the distributed lag structure was introduced (M4, estimates 0.07–0.09). The effects, however, were considerably reduced when mortality rates and volume were transformed using first differences (M3), with coefficient estimates ranging from 0.01 to 0.06, with some estimates not significant. When the differenced series were modeled together with accumulative volume (M5), the lower estimates observed in M3 were not found, with estimates now ranging from 0.04 to 0.09. Multilevel models using differenced series (M3

and M5) failed to converge, which might have been caused by reduced random variance when variables were transformed by first differences.

Comparing ARIMA and panel model results showed good consistency between the two approaches. The overall weighted average estimate (using the meta-analytic pooling method) from ARIMA modeling was 0.083, in comparison the overall estimates of 0.07–0.09 when the distributed lag structure is similarly used and time trend is controlled (panel data M4) and also similar to estimates from models using concurrent consumption (M1 and M2). In addition to average estimates, we evaluated the comparability of individual state estimates between ARIMA models and predicted effects from multi-level models. Here we used the multi-level model of random alcohol slope without AR, controlling for quadratic time trend (M2). As shown in Figure 2, high consistency was observed between the two approaches at individual state level, with correlation coefficient 0.80.

Analyses with beverage-specific series

In these analyses, the focus was on how the effects differed when the beverages were evaluated separately as compared to when they are entered together as predictors. To simplify analysis and presentation, we will present results for concurrent alcohol beverage volume measures controlling for quadratic time trend, as in M2 above with total consumption. Earlier analysis on total consumption showed this simple model generated estimates similar to those from more complicated specifications.

Estimates of beverage-specific effects on cirrhosis mortality rates from panel models are presented in Table 3. When spirits consumption was used alone, positive and significant estimates were derived from all models, with estimates ranging from 0.09 to 0.13, except for multi-level model without an AR term, where a high estimate of 0.20 was found. When spirits were included together with beer and wine consumption, estimates were slightly reduced across models, ranging from 0.07 to 0.12 and 0.18 for the multi-level model. While these average estimates from panel data models are quite consistent, they are much lower than those from ARIMA models, where the weighted average estimate was 0.24 (Table 1). Comparing the individual state estimates from ARIMA models and the multi-level models generating the closest average estimate (0.18), the correlation coefficient is just a moderate 0.50, much lower than 0.80 observed earlier for total consumption.

Panel data estimates where beer was used as the lone predictor ranged from 0.06 to 0.13, in comparison to estimates of 0.03 to 0.09 when beer was entered together with spirits and wine consumption. In both cases, significant positive effects were observed across all panel models, whereas a non-significant average beer effect was seen in ARIMA models (weighted average from random effect -0.02 , Table 1). Highly inconsistent estimates were observed for wine. When only wine was examined, significant positive estimates were seen for nearly all models and ranged from 0.10 to 0.40. The estimates when spirits and beer consumption were also entered were highly unstable, with significant negative effects, non-significant effects and significant positive effects seen across models ranging from -0.11 to 0.20. A significant negative effect was derived from ARIMA models (-0.10).

DISCUSSION

It has been a common practice for researchers to choose the model of their preference when investigating relationships between aggregate mortality and alcohol consumption using panel data and comparison studies between various model specifications are rarely conducted. As illustrated in this example, combining multiple geographic time series, which could vary greatly in trend, error term distribution and form of autocorrelation, into panel data demands sophisticated data modeling techniques. As complex as panel data can be,

none of the existent models could accommodate all sources of heterogeneity and therefore any type of model must employ some form of generalization. A question facing all analysts is: which model is most appropriate when estimating relationships in this type of data?

Issues and considerations in modeling mortality and alcohol using panel data

Statistical models are mathematical constructs designed to represent the behavior of data, and different model specifications reflect varied assumptions of underlying data-generating process. In panel data models, for example, fixed effects models assume that the stochastic outcome is generated following a linear functional form for each fixed cross-section (panel) unit, while models like GEE and GLS assume that the outcome variable is conditional on exogenous independent predictors at both cross-section and temporal dimensions. Further complications arise regarding assumptions of the error term structure: whether the residual is generated following some complicated within-panel autoregressive pattern (i.e 2nd order AR term handled by GEE), whether the variance and auto-correlation of residuals vary across panels as specified in GLS, or whether the error structure is generated by the random intercept and slope process implied in multi-level models. Although ideally each model should be appropriately specified following the relevant underlying data-generating process, in reality the assumptions are difficult or impossible to test and empirical researchers generally end up choosing *the* model they believe.

Another key model specification issue is the choice of whether to adjust for a common trend as a function of time or to difference each series. Neither approach is perfect. By eliminating both cross-sectional variation and long-term series-specific trends, the differencing method is more conservative and consequently subject to a higher likelihood of type-II error (Skog, 2001). Also, differencing trend-stationary series can induce spurious auto-correlation (Rehm, 2001), and therefore stationarity tests should be used to indicate whether differencing is appropriate. On the other hand, adjusting for a function of time in a panel model assumes a common trend exists across all series in the panel, which might not be true. In essence, the two approaches model different relationships, one between levels and one between changes, thus the resulting estimates might not necessarily converge.

In addition to model selection and trend adjustment, a separate issue is whether a distributed lag transformation of consumption should be used to model the accumulative effect of alcohol when modeling chronic disease mortality, like cirrhosis. Given all the uncertainties described above, we argue different types of panel data models should be estimated to check the robustness of findings, and to avoid conclusions drawn from fitting a single model.

Results from US cirrhosis mortality and their interpretations

In this study, across 47 US state series of liver cirrhosis mortality rates from 1950 to 2002, the effects of total alcohol consumption on cirrhosis mortality rates from the four types of panel data models with different error specifications were highly consistent. Similar results were also reached when controlling for quadratic time trend and when accumulated volume was used as predictor through a distributed lag structure. The main discrepancy arose when the effect was modeled using the first difference of the dependent and predictor series with concurrent volume. One interpretation of this is that when the level, rather than the first difference of volume is used, concurrent volume itself already incorporates information on accumulated consumption because the current consumption level is reached by building up or down from earlier consumption, while the difference only captures the *change* to the current year and so a distributed lag transformation is needed to model the accumulative effect. While this survey of panel data modeling methods is not exhaustive, the consistency reached across various model specifications and estimation methods confirms the robustness of the findings, which is further strengthened when compared to similar results from

ARIMA models. In addition to the average overall effect, highly consistent individual state effects were derived from the ARIMA models and predicted estimates from the multi-level model, the only panel model treating state effects as random.

Quite different findings were observed when mortality rates were predicted by beverages-specific volumes. While moderate consistency was seen across different types of panel models for spirits and beer effects, their panel estimates were different from those from ARIMA models. Of particular interest is that significant positive coefficients were seen for beer, controlling for spirits and wine, across panel models, though effect sizes were quite variable (0.031–0.094), while a non-significant negative effect was observed in pooled ARIMA estimates (−0.023). When panel models were re-estimated using differenced series, results were quite similar to those from ARIMA models (not shown), suggesting that the use of level or differenced series might be a major underlying cause of discrepancies. Substantial discrepancies were found for wine effects, with variation in estimates across different panel models and in comparison to ARIMA models. The unreliability of wine effect estimates was not unexpected. Wine consumption has been low compared to spirits and beer in the US and this smaller variation would result in bigger standard error estimates.

Some guidance and recommendations for model selection and estimation

This study of alcohol consumption and cirrhosis mortality across states of the US found highly consistent estimates across different model specifications when total consumption was used as predictor, but discrepancies for beverage specific effects, particularly for beer and wine. Two natural questions arise: 1) when large differences found across various model estimations, is there any approach to investigate and reconcile the discrepancies, and 2) are some model specifications preferred in modeling this type of data?

When different estimates are generated, extensive efforts should be made to investigate the causes of the discrepancies. For example, when fixed effect estimates or pooled estimates from ARIMA models differ from GEE or GLS estimates, a potential cause might be varied relationships across and within panels and one might think of group-centering the predictor variables, or alternatively using panel dummy variables, in the pooled models. When different estimates from GEE and GLS are derived, a further comparison with ARIMA estimates might reveal whether heteroskedasticity or panel-specific auto-correlation might be the underlying cause. Here, one issue is how to appropriately specify the error structure in a panel model and what is the impact of misspecification. Theoretically, fitting panel data with even Ordinary Least Square (OLS) and assuming homoskedasticity and independent errors generates coefficient estimates that are asymptotically consistent and unbiased (but not efficient) in a finite sample (Greene, 2000). In reality, however, it is not unusual for differences to be observed in coefficient estimates with varied error structure specifications. Further, special care should be given to controlling time trends using differenced or untransformed level series, as both approaches have some potential problems and might have different meanings. This is best exemplified by the varied beer estimates observed using these two approaches. One might argue both estimates are correct: it is the level of beer consumption, rather than its change, that is related to cirrhosis mortality rates. Further research on the meaning and impact of these two approaches is certainly needed, and for the beer effect in this example, we lack clear evidence to support either or both results. As a last resort, the underlying data generating process and sensitivity to different model specifications might be better understood using simulation studies, presumably a complex process in the panel data setting.

When choosing the most appropriate model to specify and estimate alcohol consumption and mortality relationships, researchers need to take into consideration both substantive research questions and the model best summarizing underlying behavior of the data. When

modeling cross-panel heterogeneity and prediction of panel-specific effects is a major aim of the study, it is best modeled using multi-level technique with random slopes or ARIMA analysis when each panel unit is estimated separately. To generate predicted estimates of panel-specific effects, multi-level method is preferred to ARIMA for its usage of empirical Bayes estimator incorporating both overall and individual panel estimates. When a pooled estimate is the main interest, we regard GLS model a preferred choice in comparison to fixed effect and GEE, given its rich selection of error structure specifications. Because ARIMA is the most powerful and flexible approach to model and test error structure, these results can be used to inform the error structure specification to be used in the panel data analysis, e.g. whether a panel-specific AR is needed, or if a common AR term across panels is enough. When GLS is used, however, special attention should be paid when the estimate is different from the fixed effect estimate, which might be caused by varied within- and cross-panel effects described above, and group centering or adding panel dummies should be considered. This preference, however, does not rule out the necessary step of running other models to check robustness of findings and, when needed, reconciling differences between model estimates.

While accumulated alcohol effects are important when examining chronic disease mortality, this study found that the distributed lag transformation is most important when differenced series are used. Last, this study suggests caution when beverage specific volume is used as predictor of mortality rates in panel data. Compared with total alcohol consumption, entering beverages simultaneously into the model adds another complication to the estimation process given the generally high correlation between spirits, beer and wine volumes (correlation ranging from 0.45 to 0.67 in this study).

While various types of panel data models were examined in this study, our assessment was far from complete. One important model type not covered are dynamic panel models, in which a lagged outcome variable, together with other predictors, is used to predict the current period outcome, in order to “estimate behavioral relationships that are dynamic in character” (Hsiao, 2003). One problem in the estimation of this type of model is the endogeneity caused by correlation between lagged dependent variable and error term, typically addressed through approaches like instrumental variable estimation to obtain unbiased estimates (Baltagi, 2001). This method has been applied to model dynamic demand functions of alcohol and cigarettes (e.g. (Baltagi and Griffin, 2002; Baltagi and Levin, 1992)) to account for their potentially addictive nature. Although a common tool for health economists, these models have never been used in predicting alcohol-related mortality rates, to our knowledge, possibly due to the less dynamic nature of mortality rates.

Last, while this study focuses on methodology, it is worthwhile to discuss briefly the substantive findings. Our semi-log model of total alcohol consumption showed an effect size of 0.07–0.09 pooled across 47 US states, which translates to that on average there is a 7%–9% rise in cirrhosis mortality rates when per capita consumption of total alcohol increases by 1 liter of ethanol. In comparison, the pooled ARIMA method using semi-log models generated moderately stronger effects in Europe (0.13 for men and 0.08 for women) (Ramstedt, 2001b) and in Canada (0.17 for men and 0.12 for women) (Ramstedt, 2003). Using a semi-log model on differenced series for five English-speaking countries combined, Kerr, Fillmore and Marvy (Kerr et al., 2000) observed a coefficient of 0.094 when predicting cirrhosis mortality using current year total consumption and 0.066 for 1-year lag, and significant effect was only found for spirits consumption among beverages with positive but not significant effect from beer. Using a much shorter series (1975–86) of 50 US states, Gruenewald and Ponicki (1995) found a significant effect for spirits consumption while the beer coefficient was positive and marginally significant, in fixed effects models using untransformed level series. The coefficient estimates can not be directly compared to those

from this study, as they used a log-log model. Finally, examining the state-specific effects from ARIMA and multi-level models as shown in Figure 2, there is no evidence of strong regional patterns associated with the cirrhosis mortality and alcohol consumption relationship.

Acknowledgments

The authors thank the two anonymous reviewers for their insightful comments leading to substantial revision of the discussion section. We also thank Jason Bond and Meenakshi Subbaraman for their helpful comments on the revised manuscript.

Funding: This research was supported by research grant R01 AA014362 from the National Institute on Alcohol Abuse and Alcoholism.

References

- Baltagi, BH. *Econometric Analysis of Panel Data*. 2. John Wiley & Sons; New York, NY: 2001.
- Baltagi BH, Griffin JM. Rational addiction to alcohol: panel data analysis of liquor consumption. *Health Econ* 2002;11(6):485–491. [PubMed: 12203752]
- Baltagi BH, Levin D. Cigarette taxation: raising revenue and reducing consumption. *Structural Change and Economic Dynamics* 1992;3(2):321–335.
- EViews. *EViews 5 User's Guide*. Quantitative Micro Software; Irvine, CA: 2004.
- Gmel G, Rehm J, Frick U. Methodological approaches to conducting pooled cross-sectional time series analysis: the example of the association between all-cause mortality and per capita alcohol consumption for men in 15 European states. *Eur Addiction Res* 2001;7:128–137.
- Greene, WH. *Econometric Analysis*. 4. Prentice Hall; Upper Saddle River, NJ: 2000.
- Gruenewald PJ, Ponicki WR. The relationship of alcohol sales to cirrhosis mortality. *J Stud Alcohol* 1995;56(6):635–641. [PubMed: 8558895]
- Hedeker, D.; Gibbons, RD.; du Toit, M.; Cheng, Y. *SuperMix: Mixed Effect Models*, Version 1.0. Scientific Software International; Lincolnwood, IL: 2008.
- Hemström Ö. Per capita alcohol consumption and ischaemic heart disease mortality. *Addiction* 2001;96(1 Suppl 1):s93–s112. [PubMed: 11228081]
- Her M, Rehm J. Alcohol and all-cause mortality in Europe 1982–1990: a pooled cross-section time-series analysis. *Addiction* 1998;93(9):1335–1340. [PubMed: 9926539]
- Hsiao, C. *Analysis of Panel Data*. 2. Cambridge University Press; New York, NY: 2003.
- Kerr WC, Fillmore KM, Marvy P. Beverage-specific alcohol consumption and cirrhosis mortality in a group of English-speaking beer-drinking countries. *Addiction* 2000;95(3):339–346. [PubMed: 10795350]
- Kerr WC, Greenfield TK, Tujague J. Estimates of the mean alcohol concentration of the spirits, wine, and beer sold in the U.S. and per capita consumption: 1950–2002. *Alcohol Clin Exp Res* 2006;30(9):1583–1591. [PubMed: 16930221]
- National Center for Health Statistics. *Vital Statistics of the United States, 1959–1967*. U.S. Public Health Service; Washington, DC: 1961–1969.
- National Center for Health Statistics. *Compressed Mortality File 1968–1988*. National Center for Health Statistics; Hyattsville, MD: 2000.
- National Center for Health Statistics. *Compressed Mortality File 1989–1998*. National Center for Health Statistics; Hyattsville, MD: 2003.
- National Center for Health Statistics. *Compressed Mortality File 1999–2002*. National Center for Health Statistics; Hyattsville, MD: 2004.
- National Office of Vital Statistics. *Vital Statistics of the United States, 1950–1958*. U.S. Public Health Service; Washington, DC: 1954–1960.
- Norström T. The impact of per capital consumption on Swedish cirrhosis mortality. *Br J Addict* 1987;82:67–75. [PubMed: 3470044]

- Norström T. Per capita alcohol consumption and all-cause mortality in 14 European countries. *Addiction* 2001;96(Suppl 1):113–128.
- Norström T, Ramstedt M. Morality and population drinking: a review of the literature. *Drug and Alcohol Review* 2005;24(6):537–547. [PubMed: 16361210]
- Norström T, Skog O-J. Alcohol and mortality: methodological and analytical issues in aggregate analyses. *Addiction* 2001;96(Suppl 1):S5–S17. [PubMed: 11228078]
- Ponicki WR, Gruenewald PJ. The impact of alcohol taxation on liver cirrhosis mortality. *J Stud Alcohol* 2006;67(6):934–938. [PubMed: 17061012]
- Ramstedt M. Alcohol and suicide in 14 European countries. *Addiction* 2001a;96(Suppl 1):S59–S75. [PubMed: 11228079]
- Ramstedt M. Per capital alcohol consumption and liver cirrhosis mortality in 14 European countries. *Addiction* 2001b;96(Suppl 1):19–34.
- Ramstedt M. Alcohol consumption and liver cirrhosis mortality with and without the mention of alcohol: the case of Canada. *Addiction* 2003;98(9):1267–1276. [PubMed: 12930214]
- Rehm J. Aggregate time-series regression in the field of alcohol. *Addiction* 2001;96(7):945–954. [PubMed: 11440604]
- Roizen R, Kerr WC, Fillmore KM. Cirrhosis mortality and per capita consumption of distilled spirits, United States, 1949–1994: trend analysis. *British Medical Journal* 1999;319(7211):666–670. [PubMed: 10480821]
- Rossow I. Alcohol and homicide: a cross-cultural comparison of the relationship in 14 European countries. *Addiction* 2001;96(Suppl 1):S77–S92. [PubMed: 11228080]
- Skog O-J. The wetness of drinking cultures: a key variable in epidemiology of alcoholic liver cirrhosis. *Acta Medica Scandinavica* 1985;703:157–184. [PubMed: 3867240]
- Skog O-J. Testing causal hypotheses about correlated trends: Pitfalls and remedies. *Contemp Drug Prob* 1988;15:565–606.
- Skog O-J. Alcohol consumption and overall accident mortality in 14 European countries. *Addiction* 2001;96(Suppl 1):S35–S47. [PubMed: 11228076]
- Stata Corp. Stata Statistical Software: Release 10.0. Stata Corporation; College Station, TX: 2007.
- Sutton, AJ.; Abrams, KR.; Jones, DR.; Sheldon, TA.; Song, F. *Methods for Meta-Analysis in Medical Research*. John Wiley and Sons; Chichester, UK: 2000.
- Terris M. Epidemiology of cirrhosis of the liver: national mortality data. *Am J Public Health* 1967;57:2076–2088.

Figure 1A. Liver cirrhosis mortality rates (per 100,000) for US and five largest states (1950–2002)

Figure 1B. Liver cirrhosis mortality rates (per 100,000) for US and five smallest states (1950–2002)

Figure 1.

Figure 1A. Liver cirrhosis mortality rates (per 100,000) for US and five largest states (1950–2002)

Figure 1B. Liver cirrhosis mortality rates (per 100,000) for US and five smallest states (1950–2002)

Figure 2.
State-specific estimates between ARIMA and Multi-level model of coefficients of log cirrhosis mortality on total alcohol

Table 1

Pooled effect sizes from ARIMA modeling of log cirrhosis mortality on alcohol consumption of 47 states

	Total alcohol	Spirits	Beer	Wine
<i>Simple average</i>				
Mean	0.098**	0.261**	-0.010	-0.157*
se	0.008	0.024	0.025	0.078
<i>Meta-analytic Pooled effect</i>				
Fixed effect – Mean	0.083**	0.217**	-0.027	-0.095*
se	0.005	0.013	0.018	0.037
Random effect – Mean	0.085**	0.238**	-0.023	-0.099*
se	0.006	0.022	0.026	0.040
Test of Homogeneity – Q statistics	53.61 (df=46)	100.21 (df=45) ^I	75.67 (df=45) ^I	48.84(df=45) ^I
p	0.206	<0.01	<0.01	0.358

^I Nevada is not included in calculation of pooled effect for spirits, wine and beer due to poor model fitting.

**
p <0.01,

*
p <0.05

Table 2
Coefficients (se) of Panel data analysis of log cirrhosis mortality on total alcohol

	Raw (M1)	Time Adj. (M2)	Difference (M3)	Dist. Lag & Time Adj. (M4)	Dist. Lag & Difference (M5)
Generalized Estimating Equation (GEE)					
Exchangeable correlation	0.089** (0.002)	0.075** (0.002)	0.046** (0.007)	0.077** (0.003)	0.075** (0.010)
AR1 correlation	0.064** (0.005)	0.062** (0.005)	0.053** (0.006)	0.072** (0.006)	0.077** (0.008)
AR2 correlation	0.062** (0.005)	0.057** (0.005)	0.059** (0.006)	0.069** (0.006)	0.078** (0.007)
Generalized Least Square (GLS)					
Exchangeable correlation ¹	0.089** (0.002)	0.075** (0.002)	0.035** (0.008)	0.077** (0.003)	0.070** (0.012)
Exchangeable & AR1 ¹	0.073** (0.005)	0.066** (0.004)	0.013 (0.010)	0.075** (0.005)	0.049** (0.019)
Heteroskedastic	0.082** (0.001)	0.081** (0.002)	0.046** (0.006)	0.080** (0.002)	0.082** (0.009)
Heteroskedastic, common AR1	0.078** (0.004)	0.075** (0.003)	0.059** (0.005)	0.084** (0.004)	0.086** (0.007)
Heteroskedastic, panel specific AR1	0.069** (0.003)	0.063** (0.003)	0.054** (0.005)	0.068** (0.004)	0.085** (0.007)
Heter. corr panel, panel specific AR1	0.069** (0.001)	0.064** (0.001)	0.051** (0.002)	0.065** (0.001)	0.076** (0.002)
Fixed effect					
No AR	0.090** (0.002)	0.075** (0.002)	0.033** (0.008)	0.077** (0.003)	0.069** (0.012)
AR1	0.068** (0.006)	0.070** (0.005)	0.012 (0.010)	0.087** (0.006)	0.044* (0.020)
Multi-level model					
Random slope (alcohol), no AR	0.109** (0.009)	0.090** (0.007)	N/A ³	0.093** (0.007)	N/A ³
Random slope, AR1	N/A ²	0.074** (0.007)	N/A ³	0.086** (0.008)	N/A ³
Random slop (alcohol and time), no AR	⁴	0.073** (0.006)	⁴	0.068** (0.008)	⁴

** $p < 0.01$,

* $p < 0.05$

¹ Also called random effect model

² Time variable must be in the model when AR is specified in current SuperMix version

³ Can not converged using difference series

⁴Not applicable as time trend must be used as (random) predictors

Table 3

Coefficients (se) of Panel data analysis of log cirrhosis mortality on spirits, beer and wine, time adjusted in quadratic form

	Entered separately			Entered simultaneously		
	Spirits	Beer	Wine	Spirits	Beer	Wine
Generalized Estimating Equation (GEE)						
Exchangeable correlation	0.103** (0.004)	0.125** (0.005)	0.122** (0.016)	0.080** (0.004)	0.092** (0.005)	-0.031* (0.014)
AR1 correlation	0.095** (0.008)	0.055** (0.012)	0.182** (0.025)	0.071** (0.009)	0.031** (0.011)	0.119** (0.025)
AR2 correlation	0.088** (0.008)	0.061** (0.011)	0.116** (0.025)	0.073** (0.009)	0.032** (0.011)	0.067** (0.025)
Generalized Least Square (GLS)						
Exchangeable correlation ^I	0.103** (0.004)	0.125** (0.005)	0.155** (0.015)	0.080** (0.004)	0.090** (0.005)	-0.018 (0.014)
Exchangeable & AR1 ^I	0.095** (0.007)	0.092** (0.009)	0.168** (0.023)	0.077** (0.007)	0.063** (0.009)	0.029 (0.022)
Heteroskedastic	0.131** (0.003)	0.133** (0.003)	0.402** (0.009)	0.068** (0.004)	0.055** (0.004)	0.202** (0.011)
Heteroskedastic, common AR1	0.116** (0.006)	0.087** (0.007)	0.254** (0.018)	0.074** (0.007)	0.042** (0.007)	0.154** (0.019)
Heteroskedastic, panel specific AR1	0.111** (0.006)	0.093** (0.007)	0.187** (0.005)	0.083** (0.008)	0.043** (0.007)	0.066** (0.018)
Heter, corr panel, panel specific AR1	0.107** (0.002)	0.102** (0.002)	0.236** (0.007)	0.065** (0.002)	0.065** (0.002)	0.079** (0.005)
Fixed effect						
No AR	0.102** (0.004)	0.125** (0.005)	0.099** (0.016)	0.080** (0.004)	0.094** (0.005)	-0.043** (0.015)
AR1	0.093** (0.008)	0.099** (0.011)	0.055 (0.028)	0.079** (0.008)	0.076** (0.010)	-0.014 (0.025)
Multi-level model						
Random slope (alcohol), no AR	0.196** (0.013)	0.103** (0.015)	0.197** (0.067)	0.176** (0.016)	0.043** (0.012)	-0.114** (0.043)
Random slope(alcohol), AR1	0.128** (0.013)	0.085** (0.012)	0.134** (0.036)	0.123** (0.014)	0.047** (0.012)	-0.009 (0.032)

**
p < 0.01,

*
p < 0.05

^I Also called random effect model