CSSR Research Proposal

Alex Sacharow & Torben Klausa

October 28, 2016

Contents

1 Introduction

Alcohol has a history of causing both joy and pain. In our research we want to focus on the latter: Alcohol, if consumend in high amounts and/or regularly over a longer period of time, does severe damage to individuals' health (see e.g. World Health Organization 2014; Miller et al. 2007). It is therefore a policy concern not only to identify reasons and motivation for excessive alcohol consumption but also to develop measures to tackle this problem.

It is the aim of this paper is to identify factors interacting with alcohol related health problems and to assess alcohol policy interventions against it. Using the example of all German federal states and time series data for different age groups from 2000 to 2014, we will [1] analyse the possible connection between medical diagnoses of alcohol misuse and possibly explanatory socioeconomic factors like gender, age, unemployment rate, education level, and regional economic performance. For this, we will take both short-term and long-term medical consequences of alcohol misuse into account by examining hospital health records on acute alcohol intoxication as well as on alcoholic liver disease. Besides the socio-economic factors mentioned above, we will [2] test the effect of recent policy measures on the German state level on alcohol consumption, namely the ban on alcohol night sales introduced in 2010 in the state of Baden-Wuerttemberg.

2 Related Literature

Our research proposal is inspired by the work of Marcus and Siedler (2015) who analysed *"The Effect of a Ban on Late-Night Off-premise Alcohol Sales on Alcohol-Related Hospital Stays in Germany."* They find that the introduction of the alcohol night sales ban reduces alcohol-related hospitalizations among adolescents and young adults by about seven percent. In our analysis, we will not only try to replicate their work with publicly available data sets. We are also going to build upon it – by differenciate between short-term and longer-term effects. The focus on socio-economic explanatory factors like economic performance for the alcohol abuse as such is based upon research done by e.g. Popovici and French (2013) and Ettner (1997). Focusing on the United States, the two studies found inconsistent results: While one sees “a positive and significant effect of unemployment on drinking behaviors and the findings are robust to numerous sensitivity tests” (Popovici and French 2013) the other argues that “non-employment significantly reduces both alcohol consumption and dependence symptoms, probably due to an income effect.” (Ettner 1997) For the German case, Henkel (2000) finds a negative correlation between economic situation and alcohol and nicotine consumption.

3 Data

Our research health data origins from the German hospital diagnosis statistics from 2000 to 2014, obtained via the Information System of the Federal Health Monitoring (GBE). The data is reported by hospitals to the statistical bureaus of the respective German states and then aggregated by the Statistische Bundesamt Destatis. The data contains aggregated numbers of hospital diagnoses for each German state by age group, gender and year. The diagnoses are published according to the WHO International Statistical Classification of Diseases and Related Health Problems (ICD-10).

The data is gathered via the online-interface of the GBE. Unfortunately, the data provider neither provides an API nor a web-scrappable interface. Therefore we download the base tables manually by searching for the respective ICD-10 code, using the malleable tables to gather as much information in a single table as possible and then export them.

To identify alcohol-related health problems we use the diagnose category *F10* (Mental and behavioural disorders due to use of alcohol), more specificly: its subdivision *F10.0* (Acute intoxication) and *F10.2* (Dependence syndrome), and *K70* (Alcoholic liver disease). The diagnoses category *F10.0* essentially captures short term effects of excessive drinking, and can be used as an indicator for binge drinking.[1] The categories *F10.2* and *K70* capture long-term effects of drinking as these diagnoses are consequences of regular drinking. The used data has been downloaded on 01 November 2016 and, due to the data platform’s table size limitations, consists of three csv files, one for each diagnose category, containing the German states, the former area of Western Germany as well as Eastern Germany as columns and years (2000-2014), gender (male/female/unknown), and age groups (less than 1 year; 1 to less than 100 years in 5-year groups; older than 100 years).

The advantages of this data is that they reflect a complete survey of all civilian hospitals in Germany for the time-frame of our investigation. Furthermore, the data is not self-reported by patients in interviews but instead consists of the professional diagnoses by a third party, i.e. doctors, which eliminates any problems of possible self-reporting biases. For the case study of Baden-Wuerttemberg’s night sale prohibition we do also benefit from the availability of a longer time series as compared to the prior work on the issue by Marcus and Siedler (2015). Finally, the state-year aggregation level does allows us to supplement to our analysis a wide variety of other freely available data, e.g. the level of alcohol sales, unemployment rates etc..

However, there are several serious limitations to the data we are using. First of all, the data set can only be exploited in limits with more sophisticated methods like panel data analysis. We only have access to data which is reported annually and on the state level.[2] Hence, for a panel data analysis our scope is limited to 14 years and 16 states, which makes a total of 224 data points for each combination of diagnoses, age group and gender. We still think it is meaningfull to work with the data at hand: On the one hand we can employ simpler but still insightful methods like multiple linear regression. On the other hand, our approach can easily be extended to more fine-grained data if such data becomes available to us. A second shortcoming is the lack of further information on the patients beyond their age and gender, e.g. socio-economic characteristics. Factors like available income and medical history will matter for alcohol-related health problems, but for good reasons (data protection) they are not recorded in the respective statistics.

[1] In contrast to Marcus and Siedler (2015) and Wicki and Gmel (2011) we do not include T51 (Intoxination due to alcohol) as its covers the consumption of pure ethanol and its number are relatively small.

[2] In principle there is more fine-grained data available, even down to the individual level. However, the access to the data requires using the paid services by forschungsdatenzentrum.de which is out of our students’ budget.

As already mentioned, our analysis will be supplemented by data from other sources. We are using indicators for the state level. They include the respecitve population, population density, unemployment rates of different age groups, the state GDP and beer sales by state. We are aware of the fact that other alcoholic beverages besides beer probably play a role in the cases under scrutiny. However, with these numbers not available, the beer consumption serves as a proxy for general alcohol consumption in the state. We are collecting the supplementary data from different statistics provided by *Destatis*.

4 Empirical Strategy

We want to make a two stage analysis of short and long-term alcohol-related hospitalizations (ARH)[3]. We will start with a multiple regression for a better understanding of factors correlated to high rates of hospitalizaton for alcohol-related reasons. This analysis shall show how socio-economic characteristics of states are correlated to the number of ARH. To keep it simple, the multiple regression model only includes observations of one year:

$$ARH_s = \alpha_0 + \alpha_1 \cdot GDP_s + \alpha_2 \cdot UR_s + \alpha_3 \cdot B_s + \alpha_4 \cdot PD_s + \epsilon_s \tag{1}$$

where GDP stands for the GDP level, UR for the unemployment rate, B for the beer consumption and PD for the population density. All variables are recorded on the state level as indicated by the indices. The regression will be conducted for the aggregated number for all age groups and than only for young people (15-25) which are often of particular interest for policy makers. We will run this first regression separately for the years 2000, 2007, and 2014.

Based on the idea that individuals adapt to their circumstances and are more reactive to changes than to actual levels we will then analysis ARH difference over time:

$$\Delta ARH_{s,t} = \beta_1 \cdot \Delta GDP_{s,t} + \beta_2 \cdot UR_{s,t} + \beta_3 \cdot \Delta B_{s,t} + \alpha_4 \cdot \Delta PD_s + \epsilon_{s,t} \tag{2}$$

In contrast to the multiple regression on levels, the difference approach highlights how *changes* in hospitalizations are correlated to *changes* in the socio-economic factors used as independend variables.

In the second step of our analysis, we are looking at the effect of the night sales ban on alcohol in Baden-Wuerttemberg. To measure the effect of the ban we use the difference-in-difference (DD) approach. As the ban was only introduced in Baden-Wuerttemberg, this state will be the treatment group. All other states are in the control group. This distinction can be justified by the fact that most alcohol regulation is done on the federal level with the exception of sales hour regulation and campaigns.[4] The ban was introduced in 2010, which is captured by the treatment dummy. We further assume there are no dynamic effects, which is reasonable as patients are registered when their treatment begins:[5]

$$ARH_{s,t} = \gamma_0 + \gamma_1 \cdot dBW_s + \gamma_2 \cdot dPOST_t + \gamma_4 \cdot dBAN_{s,t} + \epsilon_{s,t} \tag{3}$$

where dBW is a dummy variable to indicate the treatmeant group (here the state of Baden-Wuerttemberg), $dPOST$ a time dummy indicating the post-treatment period and $dBAN$ a dummy for the night-sale-prohibition (= 1 for Baden-Wuerttemberg from 2010 on). Since were are re-engineering Marcus and Siedler (2015) here, we will limit the DD application to ARH by young people (15-25). To check for further robustness of the

[3] ARH is normalized by the state population and denoted in hospitalizations per 100,000 inhabitants.

[4] The night-sale ban in Baden-Wuerttemberg has been the most prominent alcohol policy in 2010. There have also been general changes to opening hours as during this time the competency was transfered to the states and media campaigns. As we only have annual data we are not fully able to distinguish between these treatments, but we assume that they are outweigthed by the night-sales ban.

[5] Even if they were registered in consecutive reporting periods, the stay of patients for short-term ARH (*F10.0*) would be in average 2,1 days and for long-term ARH 10,7 days (*K70*) and 11,4 days (*F10.2*) in 2014, making the dynamic effect insignificantly small.

results, we are comparing short-term ARH (F10.0), which should show an effect according to Marcus and Siedler (2015), and long-term ARH (K70/F10.2), which should not be affected in the short term by the treatment.

As the DD approach crucially depends on the common trend assumption, it is advisable to include control variables to capture possible trends affecting treatment and control group differently. We control for youth unemployment (YUR) as an economic factor and B as a proxy variable for the general alcohol consumption.

$$ARH_{s,t} = \delta_0 + \delta_1 \cdot dBW_s + \delta_2 \cdot dPOST_t + \delta_3 \cdot dBAN_{s,t} + \delta_4 \cdot YUR_{s,t} + \delta_4 \cdot B_{s,t} + \epsilon_{s,t} \quad (4)$$

Finally, we are refining our DD approach further by turning it into a panel and splitting up the control group. This is the final version of the model, but we expect it to generate less insightful results as we (with all states seperately) have comparably few data points (due to comparably little money to invest in data...).

$$ARH_{s,t} = \theta_0 + \theta_1 \cdot dBAN_{s,t} + \theta_2 \cdot YUR_{s,t} + \theta_3 \cdot B_{s,t} + c_s + z_t + \epsilon_{s,t} \quad (5)$$

where c_s now is capturing unobservable and time-invariant heterogeneity of states. In addition, we are replacing $dPOST$ with z_t to capture time fixed effects affecting all states, e.g. a change in a federal law. This model can be estimated by using dummies for the time-periods and first-differencing or fixed-effects to address the unobserved heterogeneity of states.

5 First results

The first model is simple and its results are at maximum indicative. It does reveal to which state charateristics short-term alcoholrelated hospitalizations (ST-ARH) are correlated. We find that ST-ARH are higher in states with a lower GDP. This captures the basic intuition that binge dringing is more common in economically less well of regions. But this view is not confirmed by our second indicator for economic well-being: the unemployment rate. Rather contrary, states with higher unemployment rates have less cases of binge drinking. Finally, we find correlation between beer consumption (approximated by the beer tax revenue) and ST-ARH. These results are robust over all years of our investgation. But this is not surprising as the variables we are considering are no strongly fluctuating over the time-horizon of our investigation. Most of them follow more or less a trend. As the contradicting interpretation already hints at, we are most likey not seeing what is happing below the aggregated surface.

The major draw back of this regression is the low number of observations. With one observation per state the standard error are huge. In the second step we increased our data set by looking at difference over time, which increases the number of available observations to 224:

6 Robustness test

References

Ettner, Susan L. 1997. "Measuring the Human Cost of a Weak Economy: Does Unemployment Lead to Alcohol Abuse?" *Social Science & Medicine* 44 (2): 251–60.

Henkel, Dieter. 2000. "Zum Konsum von Alkohol, Tabak Und Psychoaktiven Medikamenten Bei Arbeitslosen Und Einkommensarmen." *Abhängigkeiten* 6: 26–43.

Marcus, Jan, and Thomas Siedler. 2015. "Reducing Binge Drinking? The Effect of a Ban on Late-Night Off-Premise Alcohol Sales on Alcohol-Related Hospital Stays in Germany." *Journal of Public Economics* 123:

Table 1: Regression results Model 1 for 2000, 2007 and 2014

	Dependent variable:		
	F100_p1000		
	(1)	(2)	(3)
GDP_P_C	-0.03^{***}	-0.02	-0.02^{*}
	(0.01)	(0.02)	(0.01)
UR.LF	-0.05^{***}	-0.05	-0.08
	(0.01)	(0.03)	(0.04)
BTAX_P_C	0.01^{**}	0.01	0.03^{*}
	(0.01)	(0.01)	(0.01)
PD	0.0001^{**}	-0.0001	-0.0001
	(0.0000)	(0.0001)	(0.0001)
Constant	1.76^{***}	2.32^{***}	2.57^{***}
	(0.26)	(0.65)	(0.54)
Observations	16	16	16
R^2	0.69	0.58	0.71
Adjusted R^2	0.57	0.42	0.61
Residual Std. Error (df = 11)	0.12	0.28	0.25
F Statistic (df = 4; 11)	5.99^{***}	3.74^{**}	6.76^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 2: Regression results for equation 2 with first differenced data

	Dependent variable:		
	F100.D	F102.D	K70.D
	(1)	(2)	(3)
GDP.D	0.02^{***}	-0.01	0.002
	(0.01)	(0.01)	(0.002)
UR.LF.D	-0.01	-0.04^{**}	0.002
	(0.01)	(0.01)	(0.003)
BTAX_P_C.D	-0.0001	0.02^{*}	-0.001
	(0.01)	(0.01)	(0.002)
Observations	224	224	224
R^2	0.09	0.04	0.01
Adjusted R^2	0.08	0.03	-0.01
Residual Std. Error (df = 221)	0.11	0.18	0.04
F Statistic (df = 3; 221)	7.22^{***}	3.06^{**}	0.40
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table 3: Robustness test model 1 for the years 2000-2004

	Dependent variable:				
	F100_p1000				
	(1)	(2)	(3)	(4)	(5)
GDP_P_C	-0.03^{***}	-0.03^{**}	-0.03^{**}	-0.03^{***}	-0.03^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
UR.LF	-0.05^{***}	-0.05^{***}	-0.06^{***}	-0.06^{***}	-0.05^{***}
	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
BTAX_P_C	0.01^{**}	0.02^{*}	0.02^{**}	0.02^{**}	0.01^{*}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PD	0.0001^{**}	0.0001	0.0000	0.0000	-0.0000
	(0.0000)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	1.76^{***}	1.82^{***}	1.89^{***}	2.14^{***}	2.30^{***}
	(0.26)	(0.38)	(0.34)	(0.33)	(0.36)
Observations	16	16	16	16	16
R^2	0.69	0.57	0.67	0.75	0.72
Adjusted R^2	0.57	0.41	0.55	0.65	0.62
Residual Std. Error (df = 11)	0.12	0.18	0.17	0.16	0.18
F Statistic (df = 4; 11)	5.99^{***}	3.62^{**}	5.51^{**}	8.05^{***}	7.05^{***}
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Table 4: Robustness test model 1 for the years 2005-2009

	Dependent variable:				
	F100_p1000				
	(1)	(2)	(3)	(4)	(5)
GDP_P_C	-0.03^{*}	-0.03^{*}	-0.02	-0.03^{*}	-0.03^{**}
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
UR.LF	-0.04	-0.05^{*}	-0.05	-0.06^{*}	-0.08^{**}
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
BTAX_P_C	0.01	0.02	0.01	0.01	0.02^{*}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PD	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.21^{***}	2.40^{***}	2.32^{***}	2.63^{***}	2.90^{***}
	(0.53)	(0.59)	(0.65)	(0.62)	(0.58)
Observations	16	16	16	16	16
R^2	0.58	0.57	0.58	0.63	0.72
Adjusted R^2	0.42	0.41	0.42	0.50	0.61
Residual Std. Error (df = 11)	0.23	0.26	0.28	0.28	0.24
F Statistic (df = 4; 11)	3.74^{**}	3.57^{**}	3.74^{**}	4.70^{**}	6.98^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01				

Table 5: Robustness test model 1 for the years 2010-2014

	Dependent variable:				
	F100_p1000				
	(1)	(2)	(3)	(4)	(5)
GDP_P_C	-0.03	-0.02^{*}	-0.02^{*}	-0.02	-0.02^{*}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
UR.LF	-0.09^{*}	-0.08^{*}	-0.08^{*}	-0.06	-0.08
	(0.04)	(0.04)	(0.04)	(0.05)	(0.04)
BTAX_P_C	0.03^{*}	0.02^{*}	0.02^{*}	0.02	0.03^{*}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
PD	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Constant	2.64^{***}	2.64^{***}	2.70^{***}	2.51^{***}	2.57^{***}
	(0.65)	(0.60)	(0.58)	(0.62)	(0.54)
Observations	16	16	16	16	16
R^2	0.66	0.70	0.71	0.64	0.71
Adjusted R^2	0.53	0.59	0.61	0.51	0.61
Residual Std. Error (df = 11)	0.26	0.24	0.24	0.27	0.25
F Statistic (df = 4; 11)	5.26^{**}	6.38^{***}	6.83^{***}	4.92^{**}	6.76^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

55–77. doi:10.1016/j.jpubeco.2014.12.010.

Miller, Jacqueline W, Timothy S Naimi, Robert D Brewer, and Sherry Everett Jones. 2007. "Binge Drinking and Associated Health Risk Behaviors Among High School Students." *Pediatrics* 119 (1). Am Acad Pediatrics: 76–85.

Popovici, Ioana, and Michael T. French. 2013. "Does Unemployment Lead to Greater Alcohol Consumption?" *Industrial Relations: A Journal of Economy and Society* 52 (2): 444–66.

Wicki, Matthias, and Gerhard Gmel. 2011. "Hospital Admission Rates for Alcoholic Intoxication After Policy Changes in the Canton of Geneva, Switzerland." *Drug and Alcohol Dependence* 118 (2-3): 209–15. doi:10.1016/j.drugalcdep.2011.03.020.

World Health Organization. 2014. *Global Status Report on Alcohol and Health.* World Health Organization.