Demand for Alcohol Consumption in Russia and Its Implication for Mortality: ONLINE APPENDIX

Evgeny Yakovlev*

January 12, 2017

*New Economic School. Address: Novaya str 100A, Skolkovo, Moscow, Russia. E-mail: eyakovlev@nes.ru

8 Online Appendix

0.1 Effect of tax policy on consumer welfare

In this section, I model the effect of tax policy on consumer welfare.

In both the forward-looking and myopic models presented above, consumers have bounded rationality: they do not take into account the effect of heavy drinking on hazard of death.[1] Within these models, the tax corrects a negative externality that appears from the bounded rationality of consumers. The welfare effect of the 50 percent tax is a 30 percent loss in consumer surplus.[2] At the same time, the tax saves 30,000-50,000 young male lives annually, which is 0.04-0.06 percent of the working-age population. The rough estimation of the value of their lives is the present value of the GDP that they generate. With a time discount of $\beta = 0.9$, the value of saved lives is equal to 0.4-0.6 percent of GDP, which equals the size of the whole alcohol industry in Russia (0.48 percent of GDP). This speculative calculation suggests that a 50 percent tax is actually likely to be smaller than the optimal one.[3]

Further, under certain assumptions about utilities, my model implies that the effect of a vodka tax on consumer surplus would be positive even for fully-rational consumers, forward-looking consumers who take into account the mortality risk associated with heavy drinking. The model implies that peer effects and the effect of habits are positive; all other things being constant, a consumer has higher utility if he or she drank within the previous period and if he or she has peers who are heavy drinkers. These forces, however, can also run a consumer's utility into the negative. First, quitting heavy drinking is costly. Second, a consumer who decides not to drink may suffer from the fact that peers are drinking – the consumer may experience peer pressure or the consumer may suffer if no peer wishes to participate in alternative activities, such as playing soccer or other sports.[4] Thus, in section 8.3.2 in the online appendix, I find that peer decisions matter for a consumer if he or she decides to do physical activities. These alternative assumptions about utilities, although barely distinguishable from the data, have different implications for the analysis of consumer welfare.[5] In this case, a 50 percent tax on vodka results in an increase in the

[1] I analyze the model where consumers do take into account the effect of drinking on hazard of death in the appendix (see Table A4 in appendix). Results are similar to those of the forward-looking model in the main body of text (with slightly higher magnitude).

[2] Consumer welfare is the expected (over realization of private utility shocks) present value of the flow of utilities. Under my model assumptions, $\triangle E(CS) = \frac{1}{\alpha_i}[ln(\sum(exp(V_{ij}))|tax - ln(\sum(exp(V_{ij}))|notax]$, where V_{ij} is choice-specific value function (for a consumer i and choice j), α_n is the marginal utility of income (negative coefficient with price).

[3] My model does not take into account the fact that the tax almost certainly saves other lives (children, females, the elderly), decreases crimes committed under alcohol intoxication, decreases car accidents, and so on.

[4] In this case, the consumer per-period choice specific utilities are as follows:
$\pi_{it}(0) = -\delta\overline{I(a_j = 1|S_{i,-i,t})} - \gamma a_{i,t-1}$, $\pi_{it}(1) = \Gamma' D_{it} + \Upsilon' G_{-it} + \rho_{mt}$

[5] In the "myopic" case, peer effects and peer pressure are not identified jointly. One can identify only the difference between them. In the "forward-looking" case, they are identified under additional assumptions. See proof of identification results in the appendix (Proof A2). In the online appendix, I provide results of estimation for the following model: $\pi_{it}(0) = \delta\overline{\sigma(a_j = 1|S_{i,-i,t})} + \gamma a_{i,t-1}$, $\pi_{it}(1) = \alpha\overline{\sigma(a_j = 1|S_{i,-i,t})}$. Point estimates of δ, γ and α are -1.373, -1.141, 0.114 correspondingly (see Table OA1).

consumer welfare of young males below age 40.[6] Figure OA1 in online appendix illustrates this point.

0.2 Hazard of death regressions: robustness checks

Table OA2 in the online appendix reports estimates of hazard of death by different causes of death. Unfortunately, the causes of death are reported in less than 60 percent of death cases. The RLMS recorded six causes of death,namely heart attack, stroke, accidents, poisoning, cancer, tuberculosis, and "other" causes. Heavy drinking results in higher rates of death due to accidents and poisoning and other reasons, and in a higher rate of death for which the cause was not reported.

Table OA5 in online appendix reports the hazard of death with different measures of heavy drinking. I check regression results where a heavy drinkers are defined as those who belong to a) the top 25% by alcohol intake; b) the top 25% by alcohol intake within 10-years age cohorts; c) the top 50% by alcohol intake; and d) the top 25% by days of alcohol consumption (per week). In all specifications, heavy drinking results in higher mortality rates, and the effect is highest for youngest age cohort.

0.3 Discussion of identification assumptions in peer effect estimation

This model relies on two assumptions, exclusion restriction and uniqueness of equilibria.

0.3.1 Exclusion restriction

The exclusion restriction requires that (i) the subset of demographic characteristics G_{-it} does not contain all of the set of demographic variables D_{it}, and that (ii) excluded demographic characteristics are independent of private utility shocks.

Although my estimates show that consumers do not have any preferences regarding G_{-it}, all coefficients in G_{-it} are insignificant, the assumption (i) on which the exclusion restriction is based is strong. One can argue that any demographic characteristics of peers may affect the utility of the consumer, and so should be included in G_{-it}; the utility of drinking may be greater for a consumer when he drinks with peers of the same marital status, peers with better health, etc. In addition, some of the excluded demographic variables may respond to alcohol consumption. Further, excluded demographic variables (and alcohol consumption) may be affected by unobservable shocks or the selection on unobservable characteristics (see Manski's "reflection problem" (1993), Moffit (2001)). As a result, excluded instruments may be econometrically endogenous. To verify the reliability of my model, I provide different robustness checks for the obtained results.

Recent literature emphasizes the importance of peers in making personal decisions, in particular whether to drink or not (see, for example, Akerlof and Kranton (2000), Card

[6] Determining this optimal tax rate is a question for my future research.

and Giuliano (2011), Cooley (2016), Gaviria and Raphael (2001), Krauth (2005), Kremer and Levy (2008), Moretti and Mas (2009)). The estimation of peer effects is a difficult task because it can be contaminated by common unobservable factors, non-random reference group selection, and the endogeneity of other group members' choices (Manski (1993), Moffit (2001)). Recent literature responds to this problem by using random assignments of peers in peer groups, or by using quasi-experiments (see for example Kremer and Levy (2008), Katz et al. (2001), Oreopoulos (2003)). However, as Card and Giuliano (2011) and Carrell, Sacerdote, and West (2011) argue, peer relationships that occur within randomly assigned groups may significantly differ from those occurring in natural environments where people grow up together and friendships naturally occur.

Further, studies that employ random assignment usually cover only relatively narrow groups within a population and a relatively short-run time horizon. In my paper, however, the task of quantifying the effect of government alcohol policy on the alcohol consumption, mortality, and welfare of all Russian males makes it particularly important to infer the consequences of alcohol consumption and peer interactions for a broad heterogeneous group of people and over a relatively long-run time horizon.

First, I employ a linear-in-means specification with the same set of instruments to test the endogeneity and relevance of instruments. The main regression specification is as follows:

$$I(heavy\ drinker)_{it} = \sum_k \delta_k I(age\ strata = k)\overline{I(heavy\ drinker)} + \gamma I(heavy\ drinker)_{it-1} + \Gamma' D_{it} + \Upsilon' G_{-it} + \rho_{mt} + e_{it} \quad (1)$$

where $\overline{I(heavy\,drinker)}$ is instrumented by average (across peers) demographic characteristics.[7]

Table OA4 in the online appendix presents IV regression results, as well as the results of different robustness checks. After correcting for the difference in the magnitude of coefficients of the logit and linear probability models, the results have the same magnitude as the myopic model.[8]

Column IV-1 of Table OA4 shows the results of regressions where the set of explanatory variables (D_{it}, G_{-it}, ρ_{mt}, $I_{it-1}(heavy\ drinker)$) is the same as in the main model discussed in the text.

[7] One can show that under the assumption that beliefs are linear, the structural model I describe in the main body of this paper can be rewritten as a 2SLS regression with average peer demographics used as instruments. To simplify the exposition of material, I do not follow structural specification. Within this structural framework, every particular set of instruments potentially changes the model itself. For example, I should add an additional game with fathers to the model if I wanted use paternal demographics as instrumental variables.

[8] To compare coefficients in the logit model (Table 6) with those in the linear probability model (Table OA4), one needs to multiply the coefficients in Table OA4 on 5.3. To compare marginal effects of LPM and logit regression, one needs to divide the coefficients in LPM on $p(1-p)$, where p is the probability of being a heavy drinker. In our case $(p(1-p))^{-1} = 5.3$.

The P-value of the J-test for the exogeneity of instruments is 0.22, so based on this regression, one cannot reject the null hypothesis of exogeneity of the instruments. The F-statistic of the test for relevance of excluded instruments is 72 (with errors clustered on the municipality×year level), which shows that instruments are relevant. The J-statistic confirms a model assumption that demographic characteristics respond to alcohol consumption in a non-elastic way, and that after accounting for municipality×year fixed effects (that capture all shocks and all possible selection on municipality×year level) and individual demographics, average peer demographics are exogenous instruments for peer alcohol consumption. In addition, Column IV-2 of Table OA4 shows similar results in regressions when I use a subset of instruments. Column IV-3 of Table OA4 shows that the results are also similar when I change municipality×year fixed effects on individual fixed effects, which captures possible effects of omitted factors that are constant at the individual level.

0.3.2 Peer effects for Alternative Measures of Alcohol consumption and for Other Goods

I also employ alternative measures of alcohol-consumption frequency as a measure of alcohol consumption. I use a dummy (drinks two or more times per week, and thus is in the top 21% of drinkers) as an indicator for a heavy drinker, from which I get similar results with a slightly lower magnitude (see Table OA5 in the online appendix).

Further, I provide an additional test to check that the observed correlation between one's own consumption and peer consumption is driven by peer effects but not by common unobservable shocks. I check the model by applying a similar strategy to tea, coffee, and cigarette consumption, and to hours of physical activity (see Table OA5 in the online appendix). If the correlation is driven by common shocks (for example, local prices) I should find evidence of peer effects for every good. Indeed, I find no evidence that peers affect either tea or coffee consumption. At the same time, I find a positive and statistically significant (for younger groups) peer effect on the personal decision to play sports that are social activities (we play soccer or basketball in groups). The effect of peers on smoking is marginally significant for the two age strata.

0.3.3 Alternative Instruments: Military service and Father Characteristics

According to my paper, the strongest peer effects are observed for the younger generation of males (males of age 18-29). The assumption of exogeneity for subset demographic characteristics seems to be more reliable for this group because alcohol consumption does not have an immediate effect on demographic characteristics but rather manifests over the long-run.

Further, for this sub-population of males, I have the opportunity to provide additional tests of my results based on other peer characteristics used as instruments in the regression above.

First, I use the share of peers who returned from military service. Members of this group have a higher probability of being a heavy drinker. To control for selection bias, I include a share of peers who have served or will ever serve in the Russian army (according to my data) as a control for the regression. In addition, I include individual-level variables – I(served in army) and I(have served or will ever serve in army) – as control variables in the regression. The regression specification is as follows:

$$I(heavy\ drinker)_{it} = \delta\overline{I(heavy\ drinker)_{-it}} + \gamma I(heavy\ drinker)_{it-1} + \Gamma' D_{it} + \Upsilon' G_{-it} + \rho_{mt} + e_{it} \quad (2)$$

where D_{it} contains, in addition to a standard set of demographic characteristics, $I(already\ served\ in\ army)$, and $I(served\ or\ will\ ever\ serve\ in\ army)$, and G_{-it} contains $\overline{I(served\ or\ will\ ever\ serve\ in\ army)}$. Average peers alcohol consumption $\overline{I(heavy\ drinker)}$ is instrumented by the share of peers who came from the army, $\overline{I(served\ in\ army)}$.

Columns 8 and 9 in Table OA4 show the second and first stages of this regression. Column 9 indicates that those who were previously in the army drink more: the F-statistic for the first stage is 68. The second stage shows results similar to the IV regression discussed above.

Additionally, I verify the robustness of my results by estimating the IV regression on a sub-sample of respondents who had just returned from military service. These people are likely not to face shocks common to their peers. All estimates for this sub-sample have the same magnitude and statistically significant.

I then check the robustness of my results by using the demographic characteristics of the fathers of peers, rather than of the peers themselves, as instruments in my regression. The fathers of peers likely do not face common consumer shocks. Moreover, both the model estimates and the 2SLS estimates discussed above show no correlation between one's own and peer alcohol consumption for the old-age strata to which the fathers belong. This should not happen if common unobservable factors affect the behavior fathers or shape their demographics. Regression IV-8 shows the results of an IV regression that uses the demographics of all fathers as instruments. It shows the same magnitude, but the result is statistically-insignificant. Regression IV-9 shows the results of an IV regression where instead a subset of father characteristics with better predictive power is used as instruments, with results that are statistically-significant and similar in magnitude to peer effects.

Finally, in regression IV-10, I include both army IV and father IV in one regression. Again, the regressions show a similar magnitude and statistically-significant peer effects. The P-value of the J-test equals 0.57, indicating that the instruments are exogenous.

0.3.4 Important limitations

The important limitation on my study is that my data do not provide me with exact information on members of peer drinking groups. This may cause a bias in estimation of peer effects. First, within same-age and neighborhood groups of people, some people may indeed be peers and drink together while the others may not. Second, I do not have data on all peers in peer groups. As a reminder, the RLMS contains data on roughly the half of the population in peer groups. In this case, regression estimates of peer effects in our model suffer from attenuation bias even if I apply IV regression. Thus, I am likely under-estimate the magnitude of peer effects. Indeed, because I use data on half of the peers, the true effect is likely to be twice as high compared to my estimates. The short elaboration of this result is presented in Appendix, Note 2.

It is also possible that hard drinkers are less likely to respond to the RLMS survey. If hard drinkers react differently to changes in price of alcohol and/or if restricted alcohol consumption results in a higher drop in mortality rates among hard drinkers and/or magnitude of peer affects are different for them, then our estimates would be biased.

0.3.5 Uniqueness of Equilibria

The second identification assumption is that although multiple equilibria are possible, only one equilibrium plays out in the data. In the context of my model, this identification assumption states that equation [9]

$$\sigma_{it} = \frac{exp(\delta \sum_{-i} \sigma_{jt}/(N-1) + \gamma habit_{it} + \Gamma' D_{it} + \Upsilon' G_{-it} + \rho_{mt})}{\Sigma_j exp(\delta \sum_{-j} \sigma_{kt}/(N-1) + \gamma habit_{jt} + \Gamma' D_{jt} + \Upsilon' G_{-jt} + \rho_{mt})} = G(\sigma_{-it}, S_{it}, \Theta) \qquad (3)$$

has a unique equilibrium in the data. This assumption is commonly made in empirical studies of games with incomplete information. As was shown in Bajari, Hong, and Ryan (2009), in the case of monotone payoff, the expected number of equilibria decreases as the number of states $|S_{it}|$ goes to infinity. In my model setup, $S_{it} = U_{j\in\{i,-i\}}\{habit_{jt}, D_{jt}, G_{nt}, \rho_{mt}\}$, and thus $|S|$ is big enough to claim that the probability of observing multiple equilibria is small.

In addition, I ran the following experiment, which confirms the uniqueness of equilibria in my data (although does not prove it). Using an iteration procedure, I find the fixed points $\sigma*$ of equation $\sigma_{it} = G(\sigma_{-it}, S_{it}, \Theta)$, starting from different initial values of σ_{it}. First, I find the fixed point using an iteration procedure that starts from a zero level of alcohol consumption (the so-called "low-level equilibrium"). Then, I find the fixed point starting from the highest level of consumption (the high-level equilibrium). Further, I perform 1,000 simulations for which starting points are chosen randomly from the interval (0,1) for every agent in my data. I find that the estimated σ and utilities are essentially the same with only a small difference due to computational errors – estimated utility parameters differ only at the third decimal point. Finally, in the robustness section, I re-estimate my model using a linear probability model assumption, for which multiplicity of equilibria is not an issue. Results are robust to the choice of specification.

[9] Here S is a set of state variables; and Θ is a set of parameters. This equation comes from the expression of probability of choosing the heavy drinking option that comes from equation (6).

0.4 Habits versus Unobserved Heterogeneity

To provide evidence that the observable correlation between the current and lagged level of consumption is driven not only by individual heterogeneity but also by habit formation, I estimate an instrumental variable regression:

$$I(heavy \quad drinker)_{it} \quad = \quad \alpha \ + \ \gamma I(heavy \quad drinker)_{it-1} \ + \ \Gamma' D_{it} \ + \ \rho_i \ + \ \delta_t \ + \ e_{it} \quad (4)$$

I use personal demographic characteristics (including current health status) to control for observed individual heterogeneity, and individual fixed effects to control for unobserved heterogeneity. I use lagged health status as an instrument for lagged $I(heavy\, drinker)$. The results of this regression are presented in Table OA6 in the online appendix.

Table OA6 shows the results of regressions with lagged $I(heavy\, drinker)$ as well as the results of regressions with an average across two and three lags of $I(heavy\, drinker)$. The regression results suggest that habits are important, with the same magnitude as elsewhere in my paper. Point estimates of the coefficient on lagged $I(heavy\, drinker)$ vary from 0.28 to 0.54. In two out of three specifications, the coefficient is statistically significant at the 5% level.

References for Online Appendix

Akerlof, George and Rachel Kranton. 2000. "Economics And Identity." The Quarterly Journal of Economics, MIT Press, vol. 115(3), pages 715-753, August.

Cattaneo, Matias, Sebastian Calonico and Rocio Titiunik, 2014, Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs, Econometrica 82(6): 2295–2326, November 2014

Card, David, and Laura Giuliano, 2013 "Peer Effects and Multiple Equilibria in the Risky Behavior of Friends", Review of Economics and Statistics, 2013.

Carrell, Scott E., Bruce I. Sacerdote and James E. West. 2011. "From Natural Variation to Optimal Policy? The Lucas Critique Meets Peer Effects." NBER Working Paper No. 16865, National Bureau of Economic Research, Inc.

Cooley, Jane. 2016, Can Achievement Peer Effect Estimates Inform Policy? A View from Inside the Black Box," forthcoming, Review of Economics and Statistics.

Gaviria, Alejandro, and Raphael, Steven. 2001. "School-Based Peer Effects And Juvenile Behavior." The Review of Economics and Statistics, MIT Press, vol. 83(2), pages 257-268, May.

Kremer, Michael, and Dan Levy. 2008. "Peer Effects and Alcohol Use among College Students. " Journal of Economic Perspectives, 22(3): 189–206.

Krauth, Brian. 2005. "Peer effects and selection effects on smoking among Canadian youth." Canadian Journal of Economics/Revue canadienne d'économique, Volume 38, Issue 3, pages 735–757, August 2005.

Manski, Charles. 1993. “Identification of Endogenous Social Effects: The Reflection Problem.” Review of Economic Studies, Vol. 60, No. 3, 1993, pp. 531-542.

Moffitt, R., 2001: “Policy Interventions, Low-Level Equilibria, and Social Interactions”, in Social Dynamics, edited by Steven Durlauf and Peyton Young, MIT press

Moretti, Entico and Alexandre Mas. 2009. "Peers at Work." American Economic Review, 99(1), 2009.

Oreopoulos, Philip. 2003. “The Long Run Consequences of Growing Up in a Poor Neighbourhood.” Quarterly Journal of Economics 118(4): 1533-1575.

Table OA1. Peer effects vs Peer pressure. Rust approach

	age 18-29
β=0.9	
Lag I(heavy drinker), γ	-1.373
Peer effect, α	0.114
Peer pressure, δ	-1.141
Log Likelihood	-3554.9

Note: Table reports point estimates of utility parameters from a model with forward-looking agents, where consumer's utility depends both on peer effects and peer pressure. In this case, a consumer per-period choice specific expected utilities are as follows: $\pi_{it}(0) = \delta\overline{\sigma(a_j = 1|S_{i,-i,t})} + \gamma a_{i,t-1}$, $\pi_{it}(1) = \alpha\overline{\sigma(a_j = 1|S_{i,-i,t})}$. $\overline{\hat{\sigma}_{jt}(a_{jt} = 1|S_{i,-i,t})}$ is discretized to set {0.2, 0.4, 0.6. 0.8, 1}.

Table OA3. Hazard of Death regressions with different definitions of heavy drinker

	(1)	(2)	(3)	(4)
I(heavy drinker), age 18-29	2.304	2.247	1.678	3.576
	(0.467)	(0.456)	(0.650)	(0.440)
I(heavy drinker), age 30-39	1.704	1.912	1.359	1.905
	(0.353)	(0.357)	(0.473)	(0.334)
I(heavy drinker), age 40-49	0.588	0.470	0.819	1.133
	(0.315)	(0.336)	(0.403)	(0.248)
I(heavy drinker), age 50-65	-0.275	-0.317	0.001	-0.420
	(0.247)	(0.239)	(0.245)	(0.187)
I(Bad Health)	1.397	1.391	1.464	1.418
	(0.164)	(0.164)	(0.286)	(0.165)
Log (family income)	-0.415	-0.415	-0.105	-0.413
	(0.036)	(0.036)	(0.066)	(0.036)
I(smokes)	0.586	0.594	0.917	0.554
	(0.124)	(0.124)	(0.210)	(0.126)
I(college degree)	-0.073	-0.074	-0.248	-0.075
	(0.132)	(0.132)	(0.212)	(0.133)
Weight	-0.002	-0.002	0.002	-0.002
	(0.004)	(0.004)	(0.006)	(0.004)
I(work)	0.083	0.087	-0.086	0.003
	(0.146)	(0.146)	(0.242)	(0.147)
Observations	12,109	12,109	9,227	12,109
Heavy drinking definition	0	1	2	3

Notes: Table reports estimates of the effect of heavy drinking on the risk of death with different definitions of heavy drinking. Hazard of death regression specification as a following: $\lambda(t, X) = exp(X\beta)\lambda_0(t)$. A semi-parametric Cox specification of baseline hazard is used. Heavy drinking definitions: Model (0): Top 25% by alcohol intake; Model (1): Top 25% by alcohol intake within 10 years age cohorts; Model (2): Top 50% by alcohol intake; Model (3): Top 25% by days of alcohol consumption (per month). In model 3 data is available only for rounds 15-23 of RLMS survey. Standard errors are in parentheses.

Table OA2. Hazard of death by different causes of death

	(1) Heart attack	(2) Stroke	(3) Cancer	(4) Poisoning, accidents, injuries	(5) Tuberculosis	(6) Other	(7) Not reported
I(heavy drinker), age 18-29	0.366	1.725	1.806	2.263	-1.499	4.111	1.405
	(2.483)	(2.900)	(2.263)	(0.729)	(7.706)	(1.066)	(0.959)
I(heavy drinker), age 30-39	0.589	3.170	1.877	2.502	-3.837	1.805	0.732
	(1.281)	(1.339)	(1.449)	(0.593)	(5.034)	(0.860)	(0.730)
I(heavy drinker), age 40-49	1.136	0.541	-0.051	0.165	5.219	1.454	-0.513
	(0.713)	(0.947)	(1.248)	(0.778)	(2.134)	(0.721)	(0.661)
I(heavy drinker), age 50-65	-0.092	-0.320	-1.239	0.489	-1.874	-0.243	0.299
	(0.585)	(0.565)	(0.645)	(0.814)	(2.578)	(0.856)	(0.442)
Log (family income)	-0.113	-0.155	-0.116	-0.397	-0.514	-0.431	-0.771
	(0.100)	(0.105)	(0.110)	(0.086)	(0.318)	(0.101)	(0.065)
Bad health	0.968	0.951	1.665	-0.422	1.997	1.661	2.055
	(0.460)	(0.418)	(0.407)	(0.649)	(1.211)	(0.473)	(0.292)
I(smokes)	0.773	0.881	1.018	0.892	-0.759	1.031	-0.001
	(0.299)	(0.318)	(0.326)	(0.390)	(0.971)	(0.469)	(0.218)
I(college degree)	-0.160	0.226	0.029	-0.409	-587.986	-0.284	-0.098
	(0.310)	(0.309)	(0.325)	(0.390)	(0.000)	(0.453)	(0.262)
Body Weight	0.020	0.015	-0.002	-0.027	-0.081	-0.026	-0.017
	(0.008)	(0.008)	(0.009)	(0.010)	(0.041)	(0.012)	(0.007)
I(work)	0.266	-0.644	0.260	-0.221	-1.521	-0.209	0.666
	(0.384)	(0.377)	(0.397)	(0.360)	(1.375)	(0.416)	(0.275)
Observations	12,125	12,125	12,125	12,125	12,125	12,125	12,125

Notes: Table reports estimates of hazard of death by different causes of death. Table reports results of a hazard of death regression $\lambda(t,X) = exp(X\beta)\lambda_0(t)$ where $\lambda_0(t)$ is the baseline hazard. A semi-parametric Cox specification of baseline hazard is used. Standard errors are in parentheses.

Table OA4. Linear-in-means peer effects. Robustness checks

	Dependent variable: I(heavy drinker)						
			Sample: males of age 18-65				
	IV-1	IV-2	IV-3	IV-4	OLS-1	OLS-2	
Peer effect, $\widehat{\delta}$:							
age 18-29	0.264	0.297	0.255	0.242	0.193	0.119	
	(0.04)	(0.05)	(0.09)	(0.04)	(0.03)	(0.02)	
age 30-39	0.194	0.218	0.16	0.181	0.17	0.111	
	(0.03)	(0.04)	(0.065)	(0.03)	(0.02)	(0.01)	
age 40-49	0.063	0.089	0.063	0.053	0.121	0.057	
	(0.030)	(0.037)	(0.059)	(0.031)	(0.02)	(0.01)	
age 50-65	-0.005	0.015	0.009	-0.022	0.088	0.03	
	(0.033)	(0.041)	(0.056)	(0.033)	(0.02)	(0.016)	
Munic×year FE	Yes	Yes		Yes		Yes	
Individual FE			Yes				
Year FE			Yes				
Muslim region excluded?				Yes			
Instruments	Peers 1	Peers 2	Peers 1	Peers 1			
Observations	29554	29554	29554	27400	29923	29923	
F-test	79.9	36.29	17.02	72.02			
J-test, p-value	0.22	0.13	0.02	0.26			
			Sample: males of age 18-29				
	IV-5	IV-6	IV-6, 1st st.	IV-7	IV-8	IV-9	IV-10
Peer effect, $\widehat{\delta}$:	0.211	0.25		0.359	0.197	0.225	0.298
	(0.09)	(0.079)		(0.180)	(0.136)	(0.14)	(0.125)
Munic×year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\overline{I(served\ in\ army)}$			0.211				
			(0.040)				
$\overline{I(served\ or\ will\ ever\ serve\ in\ army)}$			-0.128				
			(0.039)				
Just came from military service?				Yes			
Instruments	Peers 1	Army		Peers 1	Fathers 1	Fathers 2	Fathers 2 Army
Observations	7750	5629^{+}	5629^{+}	149	8152	8152	5629^{+}
F-test	34.24	61.05	61.05	6.85	16.52	28.97	12.9
J-test, p-value	0.06			0.17	0.4	0.86	0.58

Notes: Table reports estimates of peer effects from a reduced-form linear-in-means peer effects model. Different columns report estimates under different specifications. Instrument set: Peers: (1) average demographics (2) average demographics without lag I(heavy drinker). Instrument set: Peer fathers: (1) average demographics (2) average demographics-subset. Demographics controls are included in every regression.$^{+}$ The # of obs. in IV6 and IV10 is smaller because data on military service available only for subset of rounds. Standard errors clustered at municipality×year are in parentheses.

Table OA5. Linear-in-means peer effects. Peer effects for different products/activities

	Peer effects			
year	age 18-29	age 30-39	age 40-49	age 50-65
I(drink tea)	-0.016	-0.016	-0.003	-0.006
I(drink coffee)	0.02	0.055	0.055	0.057
I(smoking)	0.016	0.021	0.014	0.018
I(physical training)	0.14^{+}	0.127^{+}	0.141^{+}	0.073
I(Drink 2 days/week)	0.195^{+}	0.118^{+}	-0.014	0.009

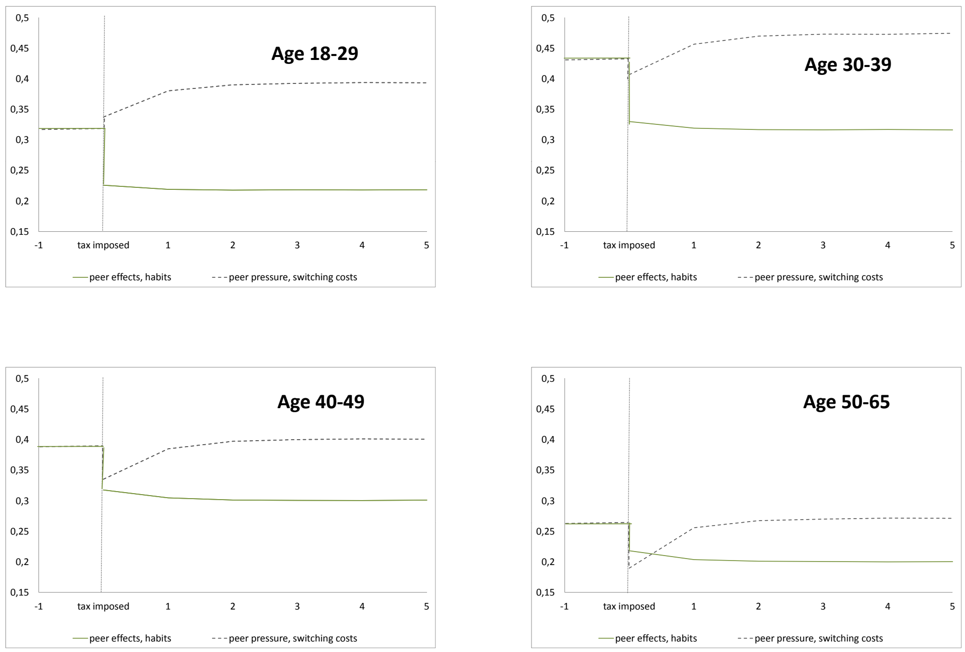
Notes: Table reports estimates of peer effects from a reduced-form linear-in-means peer effects model for different activities and for consumption of different goods.$^{+}$ indicates significance at 5%. Standard errors clustered at municipality×year are in parentheses.

Table OA6. Habits versus unobserved heterogeneity

	Y		
	I(heavy drinker)		
Mean(Lag Y, LagLag Y, LagLagLag Y)	0.516		
	(0.215)		
Mean(Lag Y, LagLag Y)		0.576	
		(0.225)	
Lag Y			0.363
			(0.250)
I(health problems)	-0.015	-0.013	-0.015
	(0.010)	(0.011)	(0.011)
Individual FE	Yes	Yes	Yes
Demographics	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	58,396	58,389	58,062
Number of individuals	9,666	9,665	9,634
F-test for instruments (with robust se)	21.21	20.77	15.93

Notes: Table reports results of IV regressions of current I(heavy drinker) on lagged I(heavy drinker). In all regressions I control for current health status and use lagged health status as an instrument for alcohol consumption. Instruments are Mean(Lag X, LagLag X, LagLagLag X), Mean(Lag X, LagLag X), and Lag X correspondingly, where X stands for I(health problems). Robust standard errors, clustered on individual level, are in parentheses.

Figure OA1. Effect of tax policy on consumer welfare



Notes: The figures show the simulated effect of 50 percent increase in the price of vodka on consumer welfare. Figures plot response under different assumptions on consumer utility function. Horizontal axis: years before and after imposing tax. Vertical axis: Consumer surplus.

Note 2. Peer effects estimation when data only on the part of peers is available

For simplicity, let's work with the example where every person has exactly two peers in his peer group, and the data contains information on only one peer.

The true model we want to estimate is

$Y_i = \phi_0 + \phi_1 \frac{X1+X2}{2}_i + e_i$,

but due to data availability we are restricted to estimate another model

$Y_i = \phi_0 + \phi_1 X1_i + e_i$.

The probability limit of OLS estimate of ϕ_1 in this case is

$plim\widehat{\phi}_{1OLS} = \frac{cov(\phi_1 \frac{X1+X2}{2}_i, x1_i)}{Var(x1_i)} = \phi_1(0.5 + 0.5cov(X2_i, X1_i))$

In case when$cov(X2_i, X1_i) < 1$, $\widehat{\phi_{1OLS}}$ has attenuation bias

Applying IV regression with instrument $Z1$ that is correlated with $X1$ and not correlated with $X2$ does not help to eliminate attenuation bias in this model:

The probability limit of IV estimate of ϕ_1 in this case is

$plim\widehat{\phi}_{1IV} = \frac{cov(\phi_1 \frac{X1+X2}{2}_i, Z1_i)}{civ(Z1_i, X1_i)} = \phi_1(0.5 + 0.5\frac{cov(X2_i, Z1_i)}{cov(X1_i, Z1_i)})$

If $cov(X2_i, Z1_i) = 0$, $\widehat{\phi_{1IV}}$ has attenuation bias again.

Indeed $plim\widehat{\phi}_{1IV} = 0.5\phi_1$ in this case.