Online Appendix

'Long-Run Consequences of Temporary Policies: Tastes and Mortality'

Lorenz Kueng Evgeny Yakovlev

A Data Appendix

We use data at three levels of aggregation: national-level, regional-level, and individual-level or household-level.

A.1 National-Level Data

Population data is provided by the World Bank (http://data.worldbank.org) and mortality rates by gender are based on the Human Mortality Database (www.mortality.org). National-level data on alcohol sales going back to 1970 are provided by the Federal State Statistics Service (FSSS or Rostat, www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/main) and its predecessor Goskomstat. There are two main approaches used in the literature to estimate samogon consumption. The first approach uses aggregate sales of sugar, which is one of the main ingredients in the production of samogon; e.g., Nemtsov (1998). The second approach uses data on violent and accidental deaths and deaths with unclear causes obtained from autopsy reports; e.g., Nemtsov (2002). For such death events there exist measures of alcohol concentration in the blood of the victim that can be used to estimate aggregate alcohol consumption. This approach gives similar estimates of samogon production as the first approach. Nemtsov (2011) provides a comprehensive survey of this literature and we use the estimates summarized in his book as our main source.

While samogon was by far the main source of illegal alcohol in the Soviet Union, much of the illegal alcohol consumed since 1992 comes from illegal imports as well as illegal production of unregistered alcohol by firms as a form of tax evasion. Unfortunately, estimates of samogon after 1991 do not distinguish between the production of home-made vodka (samogon) and other unregistered alcohol imported or produced by firms. From 1991 on we therefore follow the first approach and use changes in sugar sales per capita shown in the top panel of Figure A.1 to decompose the total amound of unregistered vodka into low-quality samogon and high-quality tax-evaded vodka produced by firms. Our estimates suggest that the latter accounts for about 35% of all unregistered vodka, consistent with independent estimates of the size of Russia's shaddow economy; e.g., Johnson et al. (1997). Finally, we note that these approaches cannot distinguish the type of alcoholic good that was produced at home, in particular whether it was homemade beer, wine, or samogon. Samogon, however, is much more popular than homemade beer. This is largely because homemade beer requires ingredients that do not grow naturally in Russia. Thus, according to the RLMS, for years 2008–2011 only 0.3% of male alcohol consumers consumed homemade beer compared to 6.2% who consumed samogon, with 2008 being the first year respondents were asked about their consumption of homemade beer.

A.2 Regional-Level Data

Regional-level data come from several sources. First, regional alcohol sales by type of alcohol, including beer, and regional mortality data by cause of death from 1998 to 2014 come from the Federal State Statistics Service (FSSS). Second, regional data on alcohol consumption and mortality from 1980 to 1992, which covers the period before and during the anti-alcohol compaign, is based on Bhattacharya et al. (2013). This dataset contains information of total alcohol consumed and of samogon production, but does not break out consumption by other types of alcohol, in particular beer. However, since home-produced beer was only a minor share of total alcohol consumption during the Soviet Union, this is only a minor limitation. Bhattacharya et al. (2013) provide an extensive discussion of this data in their online appendix. Third, we use the Russian Fertility and Mortality Database (RusFMD) of the Centre of Demographic Research at New Economic School (CDR NES), which contains detailed fertility and mortality indicators of Russia's regions. The database includes gender- and age-specific mortality indicators separately for urban and rural areas of 85 Russian regions, covering years 1989-2012. The information is based on official but previously unpublished data from the FSSS; see `www.demogr.nes.ru/en/demogr_indicat/data` for more details.

A.3 Individual-Level and Household-Level Data

Micro-level data comes from two sources. The main source is the Russian Longitudinal Monitoring Survey (RLMS). We supplement this data with additional household expenditure data from the National Survey of Household Welfare and Program Participation (NOBUS). NOBUS covers new or "western" goods better than the RLMS. These new goods became increasingly available only due to increased imports and foreign direct investments after the collapse of the Soviet Union in the second half of the 1990s.

A.3.1 Russian Longitudinal Monitoring Survey (RLMS)

The RLMS is a nationally representative survey conducted by the Carolina Population Center at the University of Carolina at Chapel Hill and the Higher School of Economics in Moscow and covers 33 regions (Russian oblasts), two of which are Muslim and hence contain fewer households that consume alcohol, plus the cities of Moscow and St. Petersburg.

In our analysis of alcohol consumption patterns, we take advantage of the detailed disaggregated responses by each individual household member age 18 and above provided in the health module of the RLMS ("Health Evaluation" section). Reported *household expenditures* on alcoholic beverages on the other hand are of much poorer quality than the *individual consumption* measures. For instance, 47% of males who report having consumed alcohol during the previous month report zero household expenditures on alcohol, and another 11% do not report their spending on alcohol at all. Individual consumption data on the other hand tend to be of much higher quality and have fewer nonresponses. This is most likely due to the fact that the health questions are asked in isolation without any other person being present except the interviewer in order to maintain full confidentiality. The average self-reported household budget share of alcohol in our sample is 5% for households reporting positive alcohol expenditures. This number is severely downward biased due to underreporting and more so than in other countries. Treml (1982), for example, shows that this level of underreporting already existed in earlier surveys, resulting in estimated alcohol expenditure shares of only 3%. Instead, we estimate the average share of alcohol in total retail sales based on official statistics to be 9% over our sample period; see Goskomstat, Statistical Yearbook, Table 20.16. While this measure of the alcohol budget share is conceptually close to the budget share in non-durable expenditures, the estimated magnitude is most likely understating the alcohol budget shares of the individuals in our sample. Many households do not consume any alcohol at all, either for religious, health, or other reasons and official sales do not include the consumption of illegally obtained or homemade alcohol. Hence, the typical household's expenditure share in our sample could be well above 10%.

Since there is no consistent aggregate price index, especially early in the sample and during the financial crisis of 1997-1998, we follow the literature and express real income by deflating it by the price of milk, which is stable over time and is measured at a geographical level which roughly corresponds to the area of a small city. The corresponding real series is then comparable across our sample period from 1994 to 2011. Moreover, by deflating income by the price of milk reported by the household, we implicitly also control for time-varying local effects. Inflation measured using the official aggregate consumer price index (CPI) is 320% in 1994 and 200% in 1995, and it jumps from 28% in 1998 to 85% in 1999; see http://stats.oecd.org. While this might be an accurate measure of inflation, using the CPI for our sample does not result in reasonable income figures across years; in particular, it appears to deflate income in later years too much relative to earlier rounds. Using nominal income or income deflated with the aggregate price index provided by Goskomstat and Rosstat instead does not affect any of our results because any difference induced by applying a different aggregate price index is fully absorbed by the period fixed effects. However, the summary statistics reported in Table 1 for real income would not be reasonable.

Measures of Alcohol Consumption We assume that beer contains 5% pure alcohol and vodka contains 40% pure alcohol, based on recommendations from the National Institutes of Health (NIH); see, e.g., Dawson (2003). Some researchers take into account the possibility that the percentage of alcohol contained in beer has increased from around 2.85% in the Soviet Union to around 5% in 2000; see, e.g., Nemtsov (2002) and Bhattacharya et al. (2013). We instead assume a constant share both for simplicity and to be conservative with respect to the growth rate of beer sales relative to vodka sales measured in pure alcohol. This assumption does not affect our results.

We then calculate consumption shares of total (pure) alcohol. We use the term "vodka" to include vodka and other hard liquor, but we exclude homemade liquor, i.e., samogon. The production of homemade liquor for personal consumption became legal only in 1997, and selling it remains illegal today. This variable is therefore measured very imprecisely and we do not include it in the main analysis.[1] We then document how our results change when we include this noisy measure of alcohol consumption. We find that the point estimate is very similar although much less precisely estimated. The term "beer" includes home-brewed beer in addition to purchased beer. The fraction of home-brewed beer however is negligible for the vast majority of households, and thus it was not asked separately in most rounds of the survey.

Next, we preform an extensive sensitivity analysis of our main results reported in Table 2 to concerns related to the use of the survey data. In particular, we analyze the effect of sample attrition, of different definitions of rural consumers, and we provide an alternative estimation that reweights the data to match the share of registered vodka in total retail sales of alcoholic beverages. These results are shown in Table A.2.

Sample Attrition Although generally low in comparison to other expenditure survey panels, sample attrition could be a problem as with any other survey-based analysis. Average interview completion rate outside St. Petersbug, Moscow City, and Moscow Oblast is over 88%.[2] To deal with attrition, RLMS replenishes its sample on a regular basis, especially in the areas of high mobility and non-response rates such as Moscow and other large cities and concludes that "the main effects [of attrition] are in the Moscow/St. Petersbug sample. Because of high attrition the Moscow/St. Petersbug sample in round 10 was replaced with a new sample. And starting with 2001 the Moscow/St. Petersbug observations from 1994 sample are no longer a part of the cross-sectional RLMS sample"; see www.cpc.unc.edu/projects/rlms-hse/data/documentation/faq.

[1] Samogon consumption is much lower today than in the Soviet era. We exclude this variable from our main analysis because it is noisy, not because we think it is not important.

[2] See www.cpc.unc.edu/projects/rlms-hse/project/samprep for a detailed discussion of attrition in the RLMS. Gorodnichenko et al. (2010) and in particular Denisova (2010) provide a more in-depth analysis of sample attrition. During the 13 year period from 1994 to 2007 analyzed in Denisova (2010), 61% of individuals in the initial sample left it as their households moved out of the surveyd dwellings. This corresponds to an annual average attrition rate of 7%. Moscow and St. Petersbug however have a response rate of only 60%. We therefore perform robustness checks excluding these two sampling units from our analysis.

We deal with attrition in two ways. First, because attrition is higher before 2001 we use survey years 2001-2011 as our baseline sample in Table 2 and show that our results are qualitatively robust to using the full sample 1994-2011 (Table 2, Column 7). Second, we assess the robustness of our main results to sample attrition in Panel B of Table A.2. Panel A, Column 1 repeats our baseline baseline result for convenience (Table 2, Column 4). In Column 2 we drop the three sampling units with the highest attrition rates, St. Petersbug, Moscow City, and Moscow Oblast and find that the effect becomes slightly stronger, consistent with the hypothesis that data from these subsamples contain more measurement error.[3] Column 3 interacts the difference-in-difference variable with the survey year to assess whether the treatment effect changes depending on the survey years used. The interaction term is statistically insignificant and economically small. Similarly, in Column 4 we find no systematic difference in treatment effect when interacting it with number of years each respondent is in the sample. Hence, "survey fatigue" does not seem to affect our main results. Finally, in Columns 5 and 6 we collapse the data to a single cross-section. In Column 5 we assign an individual to the year it was first sampled while in Column 6 we assign it randomly to any year in which it responded. The results are similar as the baseline estimates in Column 1, althought with larger point estimates and standard errors. The results in Panel B therefore suggest that sample attrition does not substantially affect our main result in Table 2.

Comparison of Survey Consumption with Administartive Data Figure A.1 compares consumption data obtained from the RLMS with corresponding National Income and Product Account (NIPA) data and alcohol retail sales. The middle panel is taken from Figure 2 of Gorodnichenko et al. (2009). The authors state that "both RLMS and NIPA measures of consumption per capita include expenditures on durables but exclude imputed in-kind expenditures" and are deflated using the CPI. "The 1998 discrepancy in panel B can be explained by the fact that RLMS had been conducted right after the August financial crisis whereas NIPA's numbers are averaged over the year." Hence, consumption per capita in the RLMS matches NIPA personal consumption per capita reasonably well.

Panel C of Table A.2 compares the evolution of the ratio of vokda to beer, both measured in pure alcohol, between the RLMS and official retail sales. Although the two series have a similar trend, they diverge in early years. It is important to note that the two series are conceptually different (see the discussion above of estimated unregistered alcohol at the national level). Retail sales only measure official alcohol sales that were subject to a sales tax. However, with the privatization after the collapse of the Soviet Union, the size of the shaddow economy increased dramatically, with estimates around 40% of GDP; see e.g., Johnson et al. (1997). The RLMS on the other hand measures total alcohol consumed, including alcohol produced in the

[3]Ideally, one could directly estimate the treatment effect of the anti-alcohol campaign on survey exit. However, in our case the treatment causes 'natural' attrition under the null hypothesis since treated households have higher mortality rates *because* they formed long-run relative tastes for hard alcoholic drinks, a point we document in our analysis of the effect of relative alcohol tastes on mortality.

informal economy. Starting in 2008 the two measures are almost identical. In Column 7 we therefore restrict our sample to survey waves between 2008 and 2011 and find quantitatively similar results as in our baseline specification, although substantially less precisely estimated due to the smaller sample size. In Column 8 we instead reweight the RLMS data to match the annual share of vodka based on retail sales. Again, we find similar results as in our baseline specification in Column 1. Therefore, the results in Table 2 do not seem to be affected by potential underreporting in the survey.

Definition of Rural and Urban Consumers In our main analysis we take advantage of the detailed demographic information in the RLMS to measure the place an individual most likely lived in around age 17, i.e., during adolescence. The RLMS provides two measure that can be used to proxy for this unobserved variable. In addtion to recording current residence, the survey also asks about the respondents' birthplace.

In Panel D we construct various indicators for whether an individual became adolescent in a rural area. In Column 9 we start by only using the current place of residence and use a strict definition of rural, only including places with a population of less than 100,000. In Column 8 we relax this definition to include places with a population less than 250,000. Both definitions yield similar results and both are in line with the baseline estimates, which uses both the current place of residence and the self-reported place of birth. Columns 11 and 12 first use the self-reported place of birth and then impute the remaining missing data with the current place of residence, using both the strict (Column 11) or the broader definition of a rural area (Column 12). Both estimates are quantitatively similar to the baseline estimate in Column 1. Finally, in Column 13 we only use the place of birth for the subset of individuals that answer this question. While substantially less precise due to the much smaller sample size, the point estimate is quantitatively similar to the baseline estimate. We therefore conclude that our main results in Table 2 are robust to using different definitions of the difference-in-difference interaction variable, I(rural).

Mortality Hazards For our analysis of the long-run effects of changes in alcohol tastes on (male) life-expectancy it is important to know whether the RLMS gives an accurate representation of death events from life tables, although based on a much smaller sample. This issue has been studied by Denisova (2010), who concludes that "the attrition bias is likely to be rather limited" and that overall, "the RLMS is reasonably good in measuring adult mortality, while the richness of the individual-level information ... with the carefully measured household data makes it very attractive to study the determinants of mortality." Death events in the RLMS are inferred directly from survey responses. In the cases where some members of the household are absent in a given interview round, the interviewer asks for the reason, and one of the possible answers given is the member's death. Of course, this source of data has its limitations. For instance, we do not have information on death events for single households. To mitigate some of those shortcomings we restrict our sample to males age 22-65, which are also individuals for whom excessive drinking is a major problem. Furthermore, we exclude households that appear only once in the survey. 5.6% of male sin our initial sample died during the sample period before reaching age 65. Of those, 44% died before reaching age 50 and 18.4% before age 40. Brainerd and Cutler (2005) use the same data for their analysis of mortality trends in Russia and summarize the data as follows:

> "For families where there is at least one member surviving, the survey asks if anyone died during the time period. We are thus able to identify deaths among the vast majority of multiple-person households (about 85 percent of the population is in multiple-person households). Our analysis of mortality in subsequent sections is based on these multiple-person households. Trends in mortality in the RLMS match trends from the aggregate data, although the level of mortality in the RLMS is 10 - 20 percent lower than the national data." (p.113)

The 10-20% gap between the level of mortality measured in the RLMS and national-level mortality is due to the sample restrictions mentioned above, in particular the need to restrict the analysis to multi-person households.

A.3.2 National Survey of Household Welfare and Program Participation (NOBUS)

The National Survey of Household Welfare and Program Participation (NOBUS), which was collected in 2003 by Goskomstat in collaboration with the World Bank and includes about 45,000 households across 80 regions in Russia, contains detailed household-level expenditure data. We use this data to study the effect of import shocks to other non-alcoholic market goods on tastes. Table A.3 provides a detailed description and motivation for classifying the goods in the seven consumption categories into either "new" or "traditional" depending on whether the good became available mostly after the collapse of the Soviet Union ("new") or whether it was already available before the early 1990s ("traditional"). The goods in the seven categories are assumed to be close but imperfect substitutes.

B Taste versus Age Effects

In this section we take advantage of the survey's panel dimension to provide additional non-experimental evidence for our mechanism. A common hypothesis in the health literature for heterogeneity in alcohol consumption are "stepping-stone" or "gateway" effects of light drugs for the consumption of harder drugs later on. In the case of alcohol, this means that beer might serve as a stepping stone earlier in life for the consumption of harder alcoholic substances later in life. According to this theory, people would start out with beer but eventually switch to vodka. Several studies have analyzed this hypothesis in the context of various types of non-alcoholic drugs.[4] To the best of our knowledge our study is the first to analyze the stepping-stone effect of light alcohol towards harder alcoholic beverages.

We decompose both alcohol shares into unconditional age and cohort effects. A stepping-stone effect of beer would generate within-consumer variation where younger consumers start out with beer before gradually substituting to harder alcohol as they become older. This would result in a downward sloping life-cycle profile of the beer share. If instead changes in alcohol shares are driven by persistent changes in tastes, then different cohorts would have relatively flat alcohol life-cycle profiles. The initial share of beer relative to vodka would increase from one cohort to the next, so that the intercept of the age profile of younger cohorts would be higher than that of older cohorts for beer consumption, and vice versa for the share of vodka.

The top panel of Figure A.2 shows the unconditional age and cohort profile of both alcohol shares. The pooled cross-sectional moments seem to support both mechanisms, stepping stone effects and changes in persistent tastes implied in the cohort effects. Survey year effects do not play a significant role as shown in the middle left panel.

Next, we exploit the panel dimension of the data to assess the relative contribution of those two forces in the middle right panel by showing the average drinking patterns after taking out individual means. Specifically, for each individual we subtract his average share, and we normalize the average of the first observed share across all individuals to zero. Hence, this figure shows the average slope of the age profile over all individuals in the sample after controlling for individual fixed effects. Under the stepping-stone hypothesis, this demeaned consumption profile should retain a significant slope, positive for vodka consumption and negative for beer. On the other hand, if changes in consumption shares are driven by changes in persistent tastes across cohorts, then these profiles should be relatively flat. The pattern shown in this figure strongly supports the latter, and there is little evidence for much change within cohorts over time and hence for stepping stone effects.

The average individual's slope shown in the middle panel could mask a stepping-stone effect if tastes form very quickly during early adulthood and then remain fairly constant. This could generate an age profile that is steep at the beginning and then flattens out quickly. In this case the average slope across all individuals would be small, since most individuals in our sample would be in the flat part of their life-cycle profile, even though the age profile is steep at the beginning. In the bottom-left panel we assess this hypothesis by plotting the demeaned age profile of individuals starting from age 18 and following them up to at most age 24. That is, we perform the same analysis as in the middle right panel on this subsample, again controlling for individual fixed effects and normalizing the initial share to zero, which is now the share at age 18. The bottom-left panel shows that there indeed is a steeper age profile from age 18 to

[4] For instance, Mills and Noyes (1984) and Deza (2012) find evidence for a modest stepping-stone effect of marijuana and alcohol for the consumption of harder non-alcoholic drugs later on. Similarly, Beenstock and Rahav (2002) find a stepping-stone effect in cigarette consumption leading to an increase in the probability of smoking marijuana later on. Van Ours (2003) finds that unobserved individual heterogeneity and stepping-stone effects can explain many patterns of drug consumption.

about age 22.

The bottom-right panel repeats this exercise, now following individuals starting at age 25 through at most age 29. We observe that the age profile already becomes flat when consumers are in their late 20s. In fact, the profiles are so flat that we cannot reject the hypothesis that the slope of the two age profiles for beer and vodka are the same. Figure A.3 shows the same analysis over the entire life-cycle, documenting that the age profile remains flat at all ages above age 22, such that the slopes of the age profiles of beer and vodka shares are not statistically different from each other.

In addition to supporting the quasi-experimental results of Section 4, this non-parametric analysis also supports the sensitive age function that we have estimated both using quasi-experimental variation in Section 5 and using import shocks in Section 6. Hence, tastes for alcohol form early in life and are fully accumulated already by the age of about 22, presumably at the beginning of an individual's consumption life-cycle when the individual starts consuming alcohol regularly for the first time.

C Algorithm for Predicting Male Mortality Rates

Let the forecast horizon $H = 0$ denote the current sample from 1994 to 2011. For simplicity, let us consider the example of an individual i that is 30 years old, was born in 1970, and has characteristics x_i. We then predict consumptions shares by running the linear regression

$$S_i^g = \varphi_c + \gamma' x_i + \alpha_a + u_i,$$

where φ_c are birth year effects, i.e., φ_{1970} and α_{30} for our individual. Similarly, we predict the mortality hazard by running the corresponding Cox regression,

$$\lambda(a|x_i, S_i^g) = \exp(\delta' S_i^g + \vartheta x_i)\lambda_0(a)\,.$$

Suppose we want to forecast the mortality rate in one year, i.e., at horizon $H = 1$. In order to do so we proceed with the following steps:

1. First, we predict the consumption shares by assuming that the same individual, with characteristics x_i and age 30, also represents a 30 year old next year, but with the consumption habit of a 30 year old *next year*, i.e., with φ_{1971} conditional on the covariates above, that is
$$\hat{S}_i^g|_{H=1} = \hat{\varphi}_{1971} + \hat{\gamma}' x_i + \hat{\alpha}_{30}.$$
Table A.4 provides the regression results for this step.

2. Next, we plug the predicted shares in the estimated mortality hazard,

$$\hat{\lambda}_i|_{H=1} = \lambda(a = 30|x_i, \hat{S}_i^g|_{H=1}; \hat{\delta}, \hat{\vartheta}).$$

3. Finally, doing this for all individuals in the sample and integrating over all individuals, we obtain the predicted male mortality rate at horizon $H = 1$.

D A Structural Model of Taste Changes

Several structural models can give rise to the persistent long-run effects of public policies we identified in the main paper. In this section we propose one particular structural model of taste changes under which even temporary policy interventions can lead to persistent effects in the long run. This basic model is consistent with the consumption patterns documented in the paper. The model extends the habit formation model by Becker and Murphy (1988) to allow for two habit-forming goods, illustrating that in this situation several steady-state consumption patterns are possible even in the absence of any unobserved individual heterogeneity. A person's consumption shares in steady state depend solely on his initial consumption pattern. Moreover, it is hard to change these consumption patterns even with very large shocks once the stock of habit is sufficiently large. Hence, policies aimed at increasing the relative price of one good may not induce everybody or even many to reduce the consumption of this good. Instead, due to the stock of habits already accumulated, people who are accustomed to this particular good will still prefer it even after the policy change. This implies that policies that influence the initial choices of younger generations can have long-run consequences over their entire life span—intended or otherwise.

D.1 Model Setup

For simplicity we assume that consumers spend all of their budget on two habit-forming goods, beer and vodka. We also assume that consumers are myopic, i.e., that they maximize only current utility and do not save, that there are no outside goods, that income does not change over time, and that there is no uncertainty.[5]

The individual derives flow utility $u(v_t, b_t, H_t^v, H_t^b)$ from consuming vodka v_t and beer b_t and also from the corresponding stocks of habit H_t^v and H_t^b. The utility function has properties that are common in the literature, specifically that $u_g > 0$, $u_{gg} < 0$, and $u_{gH_g} > 0$ with $g \in \{b, v\}$. These assumptions imply in particular that the marginal utilities of consuming beer or vodka are positive and increasing with the stock of habit of the corresponding good. Assuming a

[5] Below we reach the same qualitative conclusions if consumers are forward looking and solve a fully dynamic problem.

common rate of depreciation δ of the two habit stocks, they evolve as

$$H_{t+1}^g = (1-\delta)H_t^g + g_t,\ H_0^g \geq 0,\ \delta \in [0,1]. \tag{1}$$

The budget constraint is $p_{v_t}v_t + b_t = y_t$. Without loss of generality, we focus on interior solutions.[6] The first-order condition of this optimization problem is

$$u_v(v_t, y_t - p_{v_t}v_t, H_t^v, H_t^b) - p_{v_t}u_b(v_t, y_t - p_{v_t}v_t, H_t^v, H_t^b) = 0\ , \tag{2}$$

where u_v and u_b are the partial derivatives with respect to the first and second arguments, respectively. Since we are interested in the long-run effects of habit formation, we focus our analysis on the properties of the model's steady state. In the steady state where prices, income, and consumption are constant such that $p_{v_t} = p_v$, $y_t = y$, and $g_t = g$, the expression for the stocks of habit is g/δ. The first-order condition that implicitly defines the steady state can then be rewritten as

$$u_v(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) - p_v u_b(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) = 0\ . \tag{3}$$

In general, this is a non-monotonic function in the steady-state vodka consumption v.[7] Depending on the parametrization of the utility function u, equation (3) may have a different number of solutions. Figure A.6 illustrates that for certain parametrizations, there is a unique solution, but for many other parametrizations several steady states exist, up to a continuum of solutions.[8] These multiple equilibria are derived without any consumer heterogeneity except for differences in initial conditions. A person who initially consumes primarily beer will also prefer beer in the long-run steady state, and vice versa for vodka.

D.2 Model Properties and Extensions

This section shows that the model above with two habit forming goods can have any number of equilibira. We then provide three numerical examples that generate, respectively, one, three, and an infinite number of equilibria. We also show how to map the steady state, which the model expresses in levels, to alcohol shares, which is the concept we use in our empirical analysis. Finally, we show that these insights from be basic myopic model extend to a model with forward-looking consumers.

---

[6] If there are corner solutions, there is always a symmetric specification with at least 3 equilibria where the two stable equilibria have a consumption share in each good of either 1 or 0.

[7] This condition can also be expressed as a function of the share of vodka, $S^v = \frac{v}{v+b}$, by using the fact that $v = \frac{y \cdot S^v}{1-(1-p_v)S^v}$; see below.

[8] See below for a proof. Similar results are obtained for the model with forward-looking consumers because the steady-state Euler equation is also non-monotonic in the consumption levels.

D.2.1 Number of Equilibria in the Model with Myopic Consumers

The steady state first-order condition (FOC) for myopic agents as a function of the level of vodka consumption, v, is

$$\begin{aligned} F &= u_v(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) \\ &\quad -p_v u_b(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) = 0. \end{aligned}$$

Differentiating F with respect to v yields

$$u_{vv} - p_v u_{vb} + \delta/(1-\delta)u_{vH^v} - p_v\delta/(1-\delta)u_{vH^b} - p_v[u_{bv} - p_v u_{bb} + \delta/(1-\delta)u_{bH^v} - p_v\delta/(1-\delta)u_{bH^b}].$$

Given the assumptions that $u_{gg} < 0$, $u_{H^gH^g} < 0$, and $u_{gH^g} > 0$, some terms in this expression are positive, e.g., $\delta/(1-\delta)u_{vH^v}$, $p_v^2\delta/(1-\delta)u_{bH^b}$, and some are negative, e.g., u_{vv}, $p_v^2 u_{bb}$. Therefore, the sign of the overall sum is ambiguous.

D.2.2 Numerical Examples

One Equilibrium Let the utility function be $u = ln(b)\cdot L_b + ln(v)\cdot L_v$—with $L_g = \ln(1.1 + H^g)$ for $g \in \{b, v\}$—so that the marginal utility is $u_g = \frac{L_g}{g}$. The FOC is

$$\begin{aligned} 0 &= u_v - p_v \cdot u_b \\ &= \frac{L_v}{v} - \frac{p_v L_b}{b} \\ &= \frac{L_v}{p_v v} - \frac{L_b}{b} \\ &= \frac{L_v}{p_v v} - \frac{L_b}{y - p_v v}. \end{aligned}$$

Solving for v we obtain

$$v = \frac{L_v}{L_v + L_b} \cdot \frac{y}{p_v}.$$

Three Equilibria Let the utility function be $u = \sqrt{b}\cdot L_b + \sqrt{v}\cdot L_v$—with $L_g = \ln(1.1 + H^g)$ for $g \in \{b, v\}$—so that the marginal utility is $u_x = \frac{L_g}{2\sqrt{g}}$. Solving for v we obtain

$$v = \frac{R \cdot y}{1 + R \cdot p_v},$$

with $R = \left(\frac{L_v}{p_v \cdot L_b}\right)^2$.

Continuum of Equilibria Let the utility function be $u = \sqrt{b \cdot H^b} + \sqrt{v \cdot H^v}$, so that the marginal utility is $u_g = \frac{\sqrt{H^g}}{2\sqrt{g}}$. Solving for v we obtain

$$v = \frac{R \cdot y}{1 + R \cdot p_v},$$

with $R = \frac{H^v}{p_v^2 \cdot H^b}$.

D.2.3 Expressing the Model Solutions in Terms of Shares

$S_g = \frac{g}{b+v}$, $S_b + S_v = 1$, $p_v v + b = y$, and $\frac{S_v}{S_b} = \frac{v}{b}$. Hence,

$$\begin{aligned} v &= \frac{S_v}{S_b} b = \frac{S_v}{1 - S_v}(y - p_v v) \\ &= \frac{y \cdot S_v}{1 - (1 - p_v) S_v}. \end{aligned}$$

D.2.4 Allowing for Forward-Looking Consumers

We now relax the assumption of myopic behavior. Forward looking agents maximize the present value of utility from consuming beer and vodka, $U = u(v_t, b_t, H_t^v, H_t^b) + \sum_{i=1}^{\infty} \beta^i [u(v_{t+i}, b_{t+i}, H_{t+i}^v, H_{t+i}^b)]$. To keep the model simple, we follow Gruber and Köszegi (2001) and assume no savings and that the stock of habits evolves as follows:

$$H_{t+1}^g = \delta(H_t^g + g_t).$$

The FOC for v_t, after substituting for b_t using the budget constraints, is

$$u_{v_t} - p_{v_t} u_{b_t} + \sum_{i=1}^{\infty} \beta^i \delta^i (u_{H_{t+i}^v} - p_{v_t} u_{H_{t+i}^b}) = 0.$$

The FOC for v_{t+1} is

$$u_{v_{t+1}} - p_{v_{t+1}} u_{b_{t+1}} + \sum_{i=1}^{\infty} \beta^i \delta^i (u_{H_{t+i+1}^v} - p_{v_{t+1}} u_{H_{t+i+1}^b}) = 0.$$

Combining the two FOCs and analyzing the steady state we obtain the following Euler equation:

$$\begin{aligned} 0 = \quad & u_v(v, y - p_v v, \tfrac{\delta}{1-\delta} v, \tfrac{\delta}{1-\delta}[y - p_v v]) - p_v u_b(v, y - p_v v, \tfrac{\delta}{1-\delta} v, \tfrac{\delta}{1-\delta}[y - p_v v]) \\ & + \tfrac{\beta\delta}{1-\beta\delta}[u_{H^v}(v, y - p_v v, \tfrac{\delta}{1-\delta} v, \tfrac{\delta}{1-\delta}[y - p_v v]) - p_v u_{H^b}((v, y - p_v v, \tfrac{\delta}{1-\delta} v, \tfrac{\delta}{1-\delta}[y - p_v v])]. \end{aligned}$$

Assuming that $u_g \to \infty$ *as* $g \to 0$ guarantees the existence of a steady state.

To check the possibility of multiple steady states, we can analyze the monotonicity of the right-hand side of the steady-state Euler equation by taking the first derivative with respect to v,

$$\begin{aligned} dRHS(v)/dv = \quad & u_{vv} - 2p_v u_{vb} + p_v^2 u_{bb} + \tfrac{\delta}{1-\delta}[u_{vH^v} - 2p_v u_{vH^b} + p_v^2 u_{bH^b}] \\ & + \tfrac{\beta\delta}{1-\beta\delta}[u_{vH^v} - p_v u_{bH^v} - p_v u_{vH^v} + p_v^2 u_{bH^b} + \tfrac{\delta}{1-\delta}[u_{H^vH^v} - 2p_v u_{H^vH^b} + p_v^2 u_{H^bH^b}]]. \end{aligned}$$

This expression can be both negative and positive. To see this, assume that the utility function is separable in the two goods and their stocks of habit. Then the expression above can be rewritten as

$$\begin{aligned} dRHS(v)/dv &= \left[u_{vv} + p_v^2 u_{bb} + \tfrac{\beta\delta}{1-\beta\delta}\tfrac{\delta}{1-\delta}(u_{H^vH^v} + p_v^2 u_{H^bH^b})\right] \\ &\quad + \left[(\tfrac{\delta}{1-\delta} + \tfrac{\beta\delta}{1-\beta\delta})(u_{vH^v} + p_v^2 u_{bH^b})\right]. \end{aligned}$$

The terms in the first square brackets are all negative, while the terms in the second square brackets are all positive. Thus, depending on the relative magnitude of these terms, the first derivative can be positive or negative. The following utility specifications provide two examples, one with a unique and stable steady state and one with three steady states, two of which are stable and one is unstable. We again set $p_v = y = 1$ so that the consumption levels correspond to shares, and for simplicity we assume that $\beta = 1$ and $\delta = 0.5$. Then the utility parametrization $u = \sqrt{g} + \sqrt{H^g} + gH^g$ results in a one equilibrium, while $u = \sqrt{g} + \sqrt{H^g} + 5gH^g$ yields three equilibria.

References

Becker, Gary S. and Kevin M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy*, 1988, *96* (4), 675–700.

Beenstock, Michael and Giora Rahav, "Testing Gateway Theory: do cigarette prices affect illicit drug use?," *Journal of Health Economics*, 2002, *21* (4), 679–698.

Bhattacharya, Jay, Christina Gathmann, and Grant Miller, "The Gorbachev Anti-alcohol Campaign and Russia's Mortality Crisis," *American Economic Journal: Applied Economics*, 2013, *5* (2), 232–60.

Brainerd, Elizabeth and David M. Cutler, "Autopsy on an Empire: Understanding Mortality in Russia and the Former Soviet Union," *Journal of Economic Perspectives*, 2005, *19* (1), 107–130.

Dawson, Deborah A., "Methodological Issues in Measuring Alcohol Use," *National Institute on Alcohol Abuse and Alcoholism*, 2003.

Denisova, Irina, "Adult mortality in Russia," *Economics of Transition*, 2010, *18* (2), 333–363.

Deza, Monica, "Is There a Stepping-Stone Effect in Drug Use? Separating State Dependence from Unobserved Heterogeneity Within and Across Illicit Drugs," *Working Paper*, 2012.

Gorodnichenko, Yuriy, Jorge Martinez-Vazquez, and Klara Sabirianova Peter, "Myth and Reality of Flat Tax Reform: Micro Estimates of Tax Evasion Response and Welfare Effects in Russia," *Journal of Political Economy*, 2009, *117* (3), 504–554.

_ , Klara Sabirianova Peter, and Dmitriy Stolyarov, "Inequality and volatility moderation in Russia: Evidence from micro-level panel data on consumption and income," *Review of Economic Dynamics*, 2010, *13* (1), 209–237.

Gruber, Jonathan and Botond Köszegi, "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

Johnson, Simon, Daniel Kaufmann, Andrei Shleifer, Marshall I. Goldman, and Martin L. Weitzman, "The Unofficial Economy in Transition," *Brookings Papers on Economic Ectivity*, 1997, *1997* (2), 159–239.

Mills, Carol J and Harvey L Noyes, "Patterns and correlates of initial and subsequent drug use among adolescents," *Journal of Consulting and Clinical Psychology*, 1984, *52* (2), 231.

Nemtsov, Alexander V., "Alcohol-related harm and alcohol consumption in Moscow before, during and after a major anti-alcohol campaign," *Addiction*, 1998, *93* (10), 1501–1510.

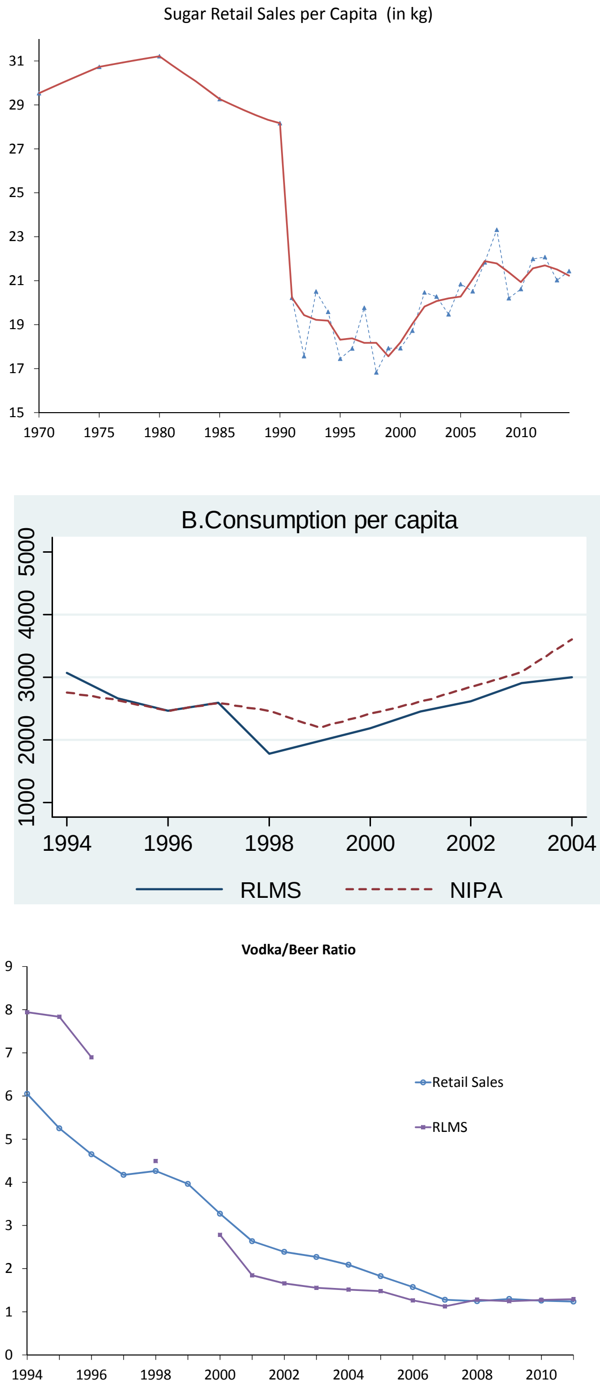
_ , "Alcohol-related human losses in Russia in the 1980s and 1990s," *Addiction*, 2002, *97* (11), 1413–1425.

Nemtsov, Alexandr V., *A Contemporary History of Alcohol in Russia*, Södertörns högskola, 2011.

Ours, Jan C Van, "Is cannabis a stepping-stone for cocaine?," *Journal of Health Economics*, 2003, *22* (4), 539–554.

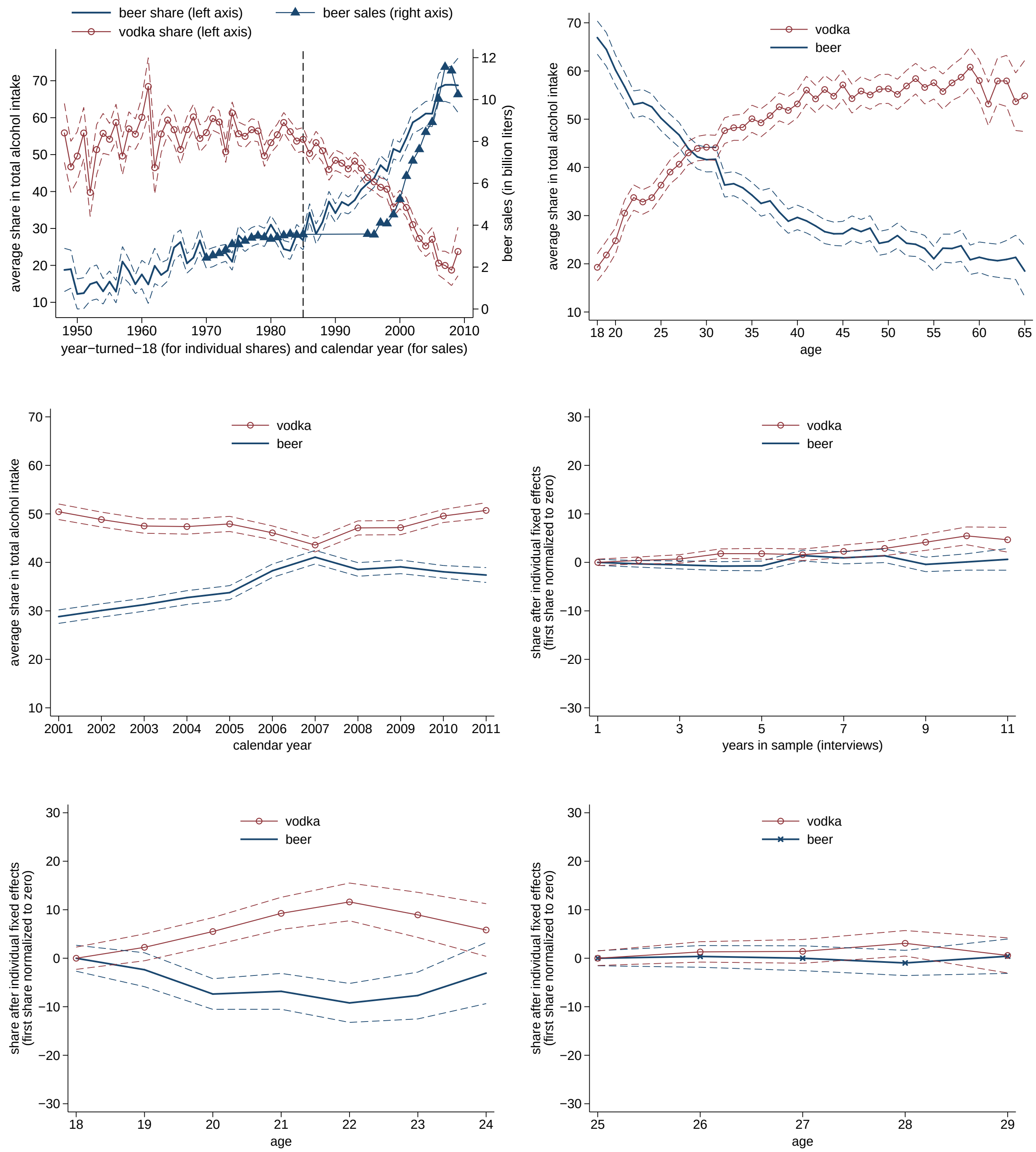
Treml, Vladimir G, *Alcohol in the USSR: A statistical study*, Duke University Press Durham, NC, 1982.

Figure A.1: Comparison of RLMS Data with Official Statistics



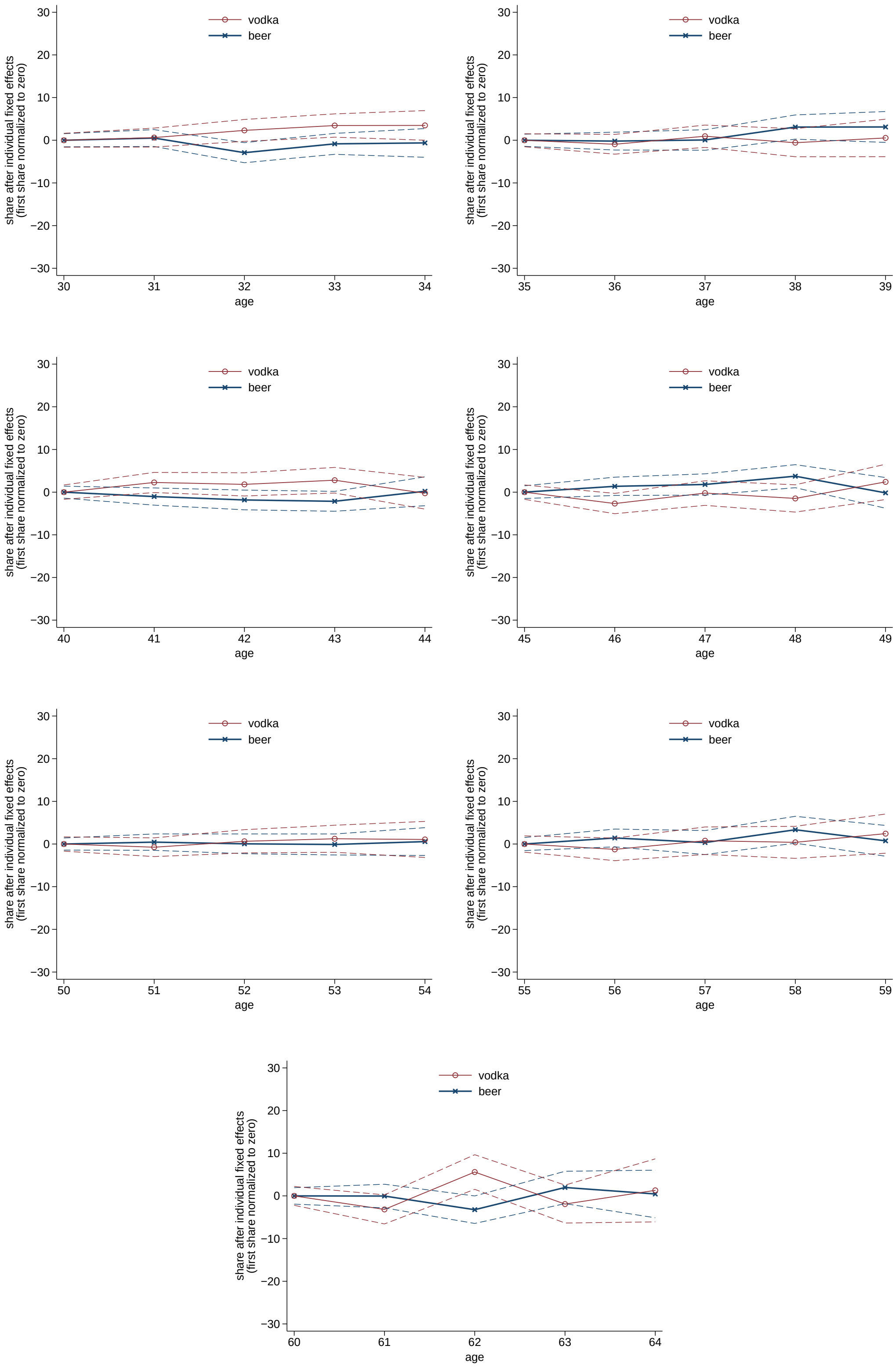
Notes: The top panel shows the time series of per capita retail sales from 1970-2014. The blue line is the raw data and the red line is the corresponding 2-year moving average. The middle and bottom panel compare consumption data obtained from the RLMS with corresponding data from national accounts (NIPA) and alcohol retail sales. The middle panel is taken from Gorodnichenko, Martinez-Vazquez, and Sabirianova Peter (2009). The bottom panel compares the ratio of vodka to beer based on (official) registered alcohol retail sales and based on self-reported data in RLMS (both in pure alcohol).

Figure A.2: Non-Parametric Decomposition of Alcohol Share Dynamics



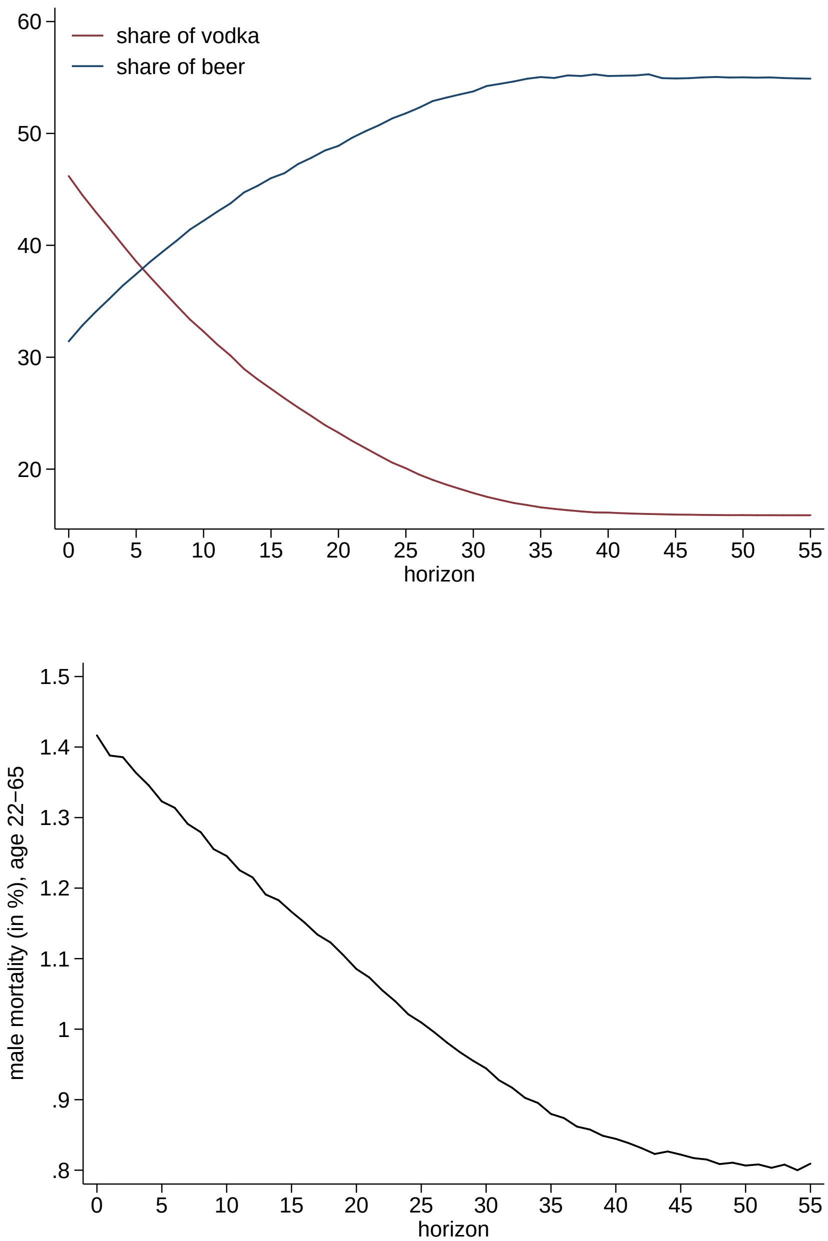
Notes: These figures show the profiles of the shares of beer and vodka consumed by men in the RLMS. The dashed lines represent two standard error confidence intervals. The top-left panel shows the age profile. The top-left panel shows the alcohol shares by cohorts measured by when and individual turned 18. We also add the volume of beer sold in the year. The vertical dashed line marks the start of the anti-alcohol campaign in 1985. The top-right panel shows the age profile for working-age men. The middle-left panel shows the average shares by survey year. The middle-right panel graphs the shares against the number of years an individual is observed in the sample, after controlling for individual fixed effects. The two bottom panels show the age profile for the two subgroups of individuals age 18 to 24 and 25 to 29 as a function of age, again after controlling for individual fixed effects. Figure A.3 provides similarly flat profiles for five-year age intervals from age 30 to 64.

Figure A.3: Demeaned Alcohol Shares over the Life-Cycle, Ages 30-64



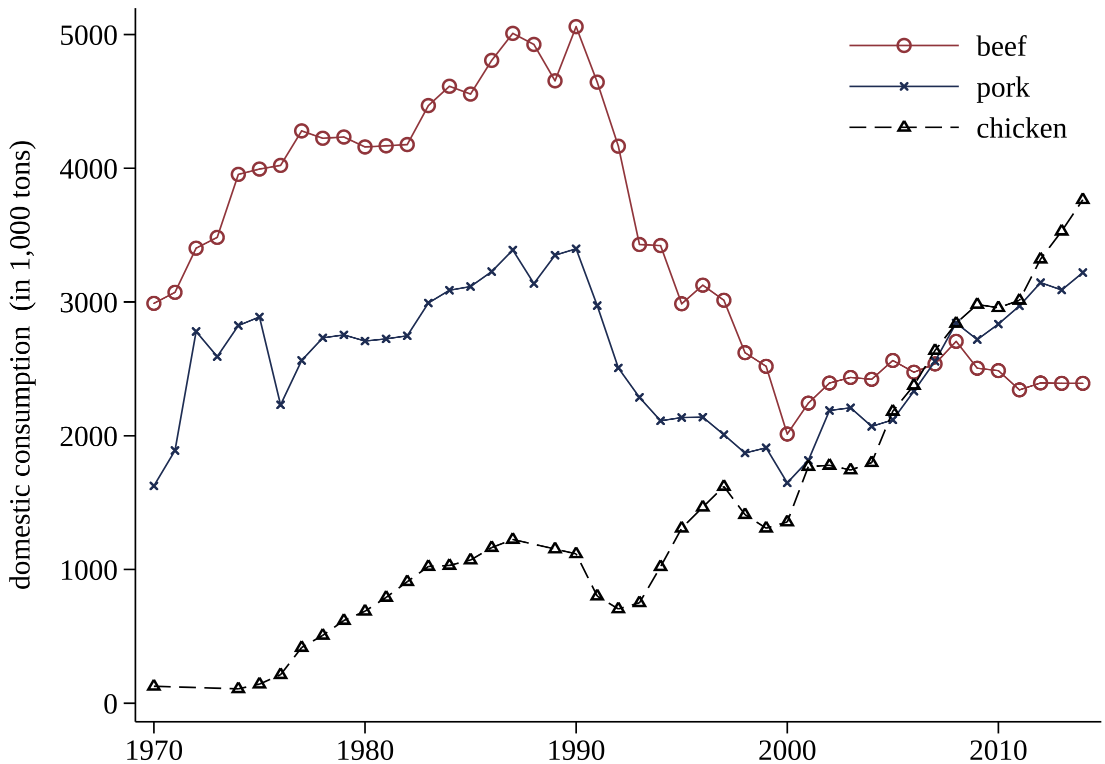
Notes: These figures provide the same analysis over the remaining part of the life-cycle as in the bottom panel of Figure A.2.

Figure A.4: Simulated Dynamics of Alcohol Shares and Mortality



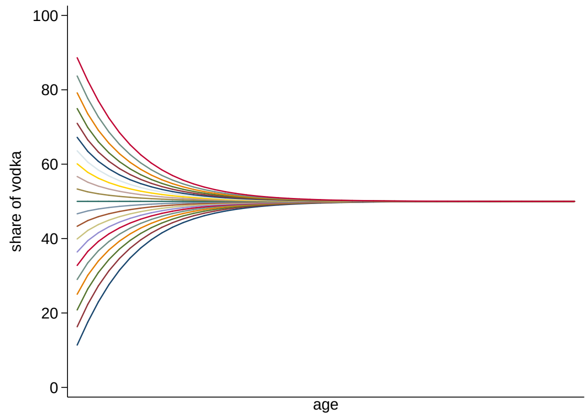
Notes: These figures shows the predicted consumption shares (top panel) and implied mortality rates (bottom panel) for males age 22 to 65 as a function of the forecast horizon in years.

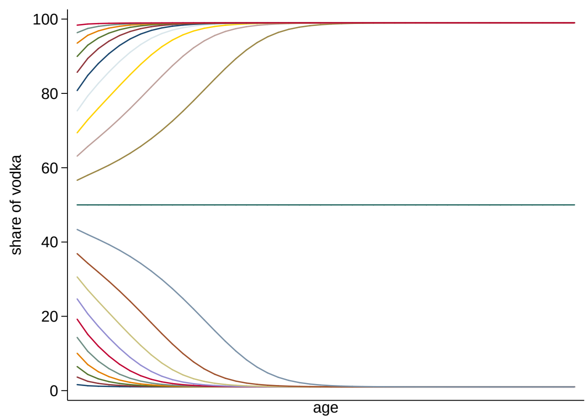
Figure A.5: Expansion of the Meat Market

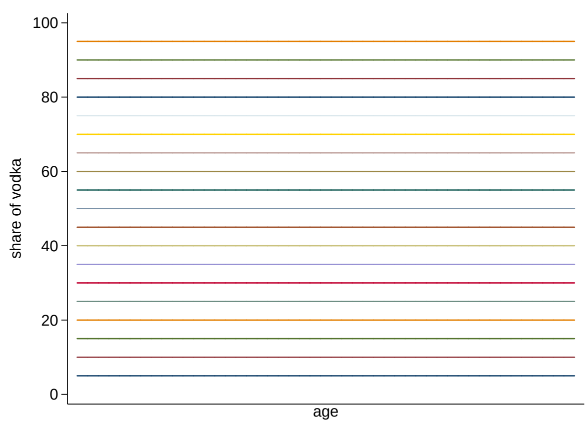


Notes: This figure shows the expansion of the meat market after the end of the Soviet Union.

Figure A.6: Potential Number of Steady States in the 2-Good Becker-Murphy Model







Notes: These figures show the dynamic behavior of the share of vodka in the two-good habit formation model, starting from different initial conditions, i.e., different initial consumption shares. The three figures correspond to the three parametrizations specified in the text. The top panel has one stable steady state, the middle panel has three steady states, two stable and one unstable, and the bottom panel has an infinite number of steady states.

Table A.1: Share of Vodka and Binge Drinking

Dependent variable: Number of days drinking alcohol	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of vodka (not in percent)	0.263*		-0.860***	-1.142***	-2.022***	-1.239***	-1.872***	-2.240***
	[0.151]		[0.119]	[0.124]	[0.115]	[0.162]	[0.172]	[0.247]
Monthly alcohol intake (in grams of alcohol)		0.006***	0.006***	0.006***				
		[0.000]	[0.000]	[0.000]				
Daily alcohol intake (in grams of alcohol)						0.021***	0.021***	
						[0.001]	[0.001]	
Age, region, and year FEs				Yes	Yes		Yes	Yes
Real income and relative price				Yes	Yes		Yes	Yes
Socio-economic demographics				Yes	Yes		Yes	Yes
Monthly alcohol intake FE					Yes			
Daily alcohol intake FE								Yes
Observations	19,781	19,781	19,781	19,781	19,781	19,781	19,781	19,781
R-squared	0.000	0.487	0.489	0.511	0.734	0.082	0.128	0.182

Notes: Columns 1 and 2 show that individuals with a larger share of vodka and with a higher level of alcohol intake both consume alcohol more frequently. However, Columns 3 to 8 show that individuals who consume the same amount of alcohol (per month or per day) but use a larger share of vodka consume less frequently. Hence, conditional on the level of alcohol intake, consumers with a larger share of vodka consume more alcohol when drinking and hence have a higher propensity to binge drink. Columns 3 and 6 uses a minimal specification without any controls. Columns 4 and 7 adds a full set of controls. Socio-economic demographics include education, marital status, body weight, and subjective health status. Columns 5 and 8 control for the level of alcohol non-parametrically. Robust standard errors in parentheses are clustered by individual; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.2: Sensitivity Analysis of the Anti-Alcohol Campaign

	A. Benchmark	B. Attrition				
Dependent variable: Share of vodka (in %)		drop Moscow and St.Petersburg	ATE by survey round	ATE by years in sample	using individual's first observation	using one observation per individual
	(1)	(2)	(3)	(4)	(5)	(6)
I(became adolescent during campaign) × I(rural)	5.243***	6.602***	5.909**	6.757**	6.935***	6.486**
	[2.016]	[2.139]	[2.483]	[2.876]	[2.591]	[2.583]
I(became adolescent during campaign) × I(rural) × (survey round)			-0.132			
			[0.378]			
I(became adolescent during campaign) × I(rural) × (years in sample)				-0.202		
				[0.329]		
I(became adolescent during campaign)	-1.038	-1.460	-1.109	-1.074	-1.460	-0.748
	[1.483]	[1.639]	[1.477]	[1.477]	[2.100]	[2.079]
Alcohol intake (in grams of ethanol)	Yes	Yes	Yes	Yes	Yes	Yes
Age, region, and year FEs	Yes	Yes	Yes	Yes	Yes	Yes
Real income and relative price	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	29,083	25,245	29,083	29,083	6,881	6,881
R-squared	0.152	0.154	0.153	0.153	0.128	0.133

	C. Comparison with Retail Sales Data		D. Definition of I(rural)				
	using years 2008-2011 only	match retail sales vodka shares	using current population < 100k	using current population < 250k	using birth place or current pop. < 100k	using birth place or current pop. < 250k	using birth place only
	(7)	(8)	(9)	(10)	(11)	(12)	(13)
I(became adolescent during campaign) × I(rural)	7.356***	5.330**	5.995***	5.167**	5.812***	5.264***	3.911
	[2.589]	[2.283]	[1.965]	[2.042]	[1.964]	[1.982]	[2.968]
I(became adolescent during campaign)	-2.376	-1.494	-2.177	-2.038	-1.428	-1.426	-0.810
	[2.591]	[1.718]	[1.566]	[1.723]	[1.488]	[1.571]	[2.423]
Alcohol intake (in grams of ethanol)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Age, region, and year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Real income and relative price	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic demographics	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,477	29,083	29,083	29,083	29,083	29,083	14,264
R-squared	0.192	0.167	0.153	0.153	0.152	0.152	0.125

Notes: Socio-economic demographics include education, martial status, body weight, and subjective health status. The length of the anti-alcohol campaign is defined to last from 1986 to 1990 based on Figure 1 and adolescence is defined as being 17 years old. The main effect I(rural) indicates the place of residence when becoming adolescent and is included in all specifications. Robust standard errors in parentheses are clustered by individual; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.3: Classification of Non-Alcoholic Goods into Traditional and New "Western" Goods

New Goods	*Traditional Goods*	*Classification*
chicken	pork and beef	After the collapse of the Soviet Union, chicken started to be produced on special chicken farms that used new technologies which more efficiently dealt with the cold weather and significantly lowered production costs. These changes lead chicken sales to exceed that of more traditional meets such as pork or beef within less than two decades.
yogurt	cottage cheese	Cottage cheese was a popular type of breakfast in the Soviet Union. After the collapse of the Soviet Union, the import of new technologies by foreign companies made mass production and storage of yogurt viable so that it became the most popular type of breakfast nowadays.
subtropical fruits	local fruits	Apples, pears, plums are locally grown fruits, while subtropical fruits such as bananas pineapples, or mango do not grow in Russia or any of the fifteen former Soviet republics. Therefore, subtropical fruits were barely available to consumers in the Soviet Union, but imports rose sharply after the collapse of the Soviet Union making them a popular and inexpensive alternative.
chocolate	jam and honey	Chocolate existed in the Soviet Union but was very expensive since cocoa beans do not grow locally. Therefore, many desserts were based on jam and honey, which are local. Today, chocolate is a significant part of Russian imports, and the relative price of chocolate has decreased dramatically.
frozen fruits	dried fruits	The technology to mass produce frozen fruits was introduced only after the collapse of the Soviet Union. Drying was the main technology for storing fruits over longer periods in the Soviet Union.
long-lasting milk	short-lived milk	Ultra-heat treated (UHT) and ultra-pasteurized milk as well as the Tetra Pak technology were introduced only after the collapse of the Soviet Union and contributed to making long-lasting milk popular. Before that, fresh milk or short-lived milk based on high-temperature, short-time (HTST) pasteurization was the only type of milk available for purchase.
salted salmon	salted herring	Salted Salmon started to be imported only after collapse of the Soviet Union, mostly from Norway. During the Soviet Union, herring was the main salted fish available.

Table A.4: Predicted Alcohol Shares in Model Simulation

	Share of beer (1)	Share of vodka (2)
Alcohol intake (in kg of ethanol)	-0.922***	0.381***
	[0.042]	[0.047]
Log(real income)	0.007***	0.005
	[0.002]	[0.003]
Health status	-0.000	-0.019***
	[0.007]	[0.007]
Body weight (in g)	-0.300	1.083***
	[0.254]	[0.290]
I(college degree)	0.005	-0.024**
	[0.009]	[0.010]
I(married)	-0.009	0.034***
	[0.009]	[0.010]
I(urban)	0.016**	-0.003
	[0.007]	[0.008]
I(no alcohol consumed)	-0.318***	-0.245***
	[0.012]	[0.013]
I(smokes)	0.015**	0.036***
	[0.008]	[0.009]
Birth year FE	Yes	Yes
Age FE	Yes	Yes
Observations	6,623	6,623
R-squared	0.321	0.270

Notes: This table reports the results from the regressions used to predict the shares of vodka and beer consumed for all individuals age 22 to 65. These predicted shares conditional on the covariates are then used to predict individual hazard rates which can be integrated across the sample population to predict the evolution of male mortality.

Table A.5: Effect of Alcohol Tastes on Mortality - Extensions

	A. Long-Run Effect of Beer-Market Expansion on Mortality			
Dependent variable :	male mortality, age 22-65 (%)			vodka share
	reduced form	instrumental variable		1st stage
	(1)	(2)	(3)	(4)
Year-turned-17	-0.044***			-1.512***
	[0.001]			[0.511]
Share of vodka consumption (in %)		0.014***	0.014***	
		[0.005]	[0.004]	
Total alcohol (liters of alc. per capita)		-0.002*		
		[0.001]		
Year, region, rural and age FE	Yes	Yes	Yes	Yes
Log(gross regional product per capita)	Yes	Yes	Yes	Yes
Observations	11,427	1,741	1,741	1,741
R-squared	0.56			0.20
1st stage F statistic		10.74	21.03	
Sample mean of dependent variable	0.44	0.40	0.40	27.32

	B. Medium- and Long-Run Effect of Campaign on Mortality	
Dependent variable :	male mortality, age 22-65 (%)	
	reduced-form effects	
	medium run 1989-1997	long run 1998-2012
	(5)	(6)
I(became adolescent during campaign) × I(rural)[(a)]	0.052***	0.024***
	[0.007]	[0.003]
I(became adolescent during campaign)[(b)]	-0.032***	-0.050***
	[0.003]	[0.002]
Sum of diff-in-diff and diff coefficients, (a)+(b)	0.019**	-0.039***
	[0.006]	[0.003]
Year, region, rural and age FE	Yes	Yes
Log(gross regional product per capita)	Yes	Yes
Observations	11,494	43,734
R-squared	0.43	0.51
Sample mean of dependent variable	0.46	0.80

Notes: Table uses male alcohol consumption data and mortality rates by year, region, age and type of settlement (urban/rural). Regressions are weighted by total population of the region and robust standard errors in parentheses are clustered by region; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.