Online Appendix

The Long-Run Effects of a Public Policy on Alcohol Tastes and Mortality

Lorenz Kueng   Evgeny Yakovlev

A Data Appendix

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

A.1 National-Level Data

Population data is provided by The World Bank (2014) and mortality rates by gender are based on the Human Mortality Database (2014). National-level data on alcohol sales going back to 1970 are provided by the Federal State Statistics Service (1991, 2017c), FFFS or Rostat, and its predecessor Goskomstat.

Measuring Illegal Vodka Consumption 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 (Nemtsov 1998). The second approach uses data on violent and accidental deaths and deaths with unclear causes obtained from autopsy reports (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, including a discussion of the limitations of both approaches. We use the estimates summarized in his book and we extend the series back to 1970 using Tables 8-2 and 8-3 in Treml (1997).[32]

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 Panel (a) of Figure A.1 to decompose the total amount 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

[32] Kueng and Yakovlev (2020) provide an Excel spreadsheet containing this data.

shadow economy (Johnson, Kaufmann, Shleifer, Goldman, and Weitzman 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 prevalent than homemade beer. This is largely because homemade beer requires ingredients that do not grow naturally in Russia. Based on data from 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 (2001, 2017a,b). Second, regional data on alcohol consumption and mortality from 1980 to 1992, which covers the period before and during the anti-alcohol campaign, 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 (2014), 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 1989-2014. This information is based on official but previously unpublished data from the FSSS assembled by the CDR; see www.demogr.nes.ru/en/demogr_indicat/data for more details. Cause-specific mortality data is available for 5-year age groups while data on total mortality from all causes of death is available for 1-year age groups.

A.3 Individual-Level and Household-Level Data

Micro-level data comes from two sources. The main source is the Russian Longitudinal Monitoring Survey (2011), RLMS. We supplement this data with additional household expenditure data from the National Survey of Household Welfare and Program Participation (2003), NOBUS, which 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) plus the cities of Moscow and St. Petersburg. Two regions are predominately Muslim and hence contain fewer households that consume alcohol.

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 (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 (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.[33] We then document how our results change when we include this noisy measure of alcohol consumption. As expected, we find that the point estimate is smaller (attenuation bias) and 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.

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. Petersburg, Moscow City, and Moscow Oblast is over 88%.[34] 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. Petersburg sample. Because of high attrition the Moscow/St. Petersburg sample in round 10 was replaced with a new sample. And starting with 2001 the Moscow/St. Petersburg 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. Online Appendix C provides extensive robustness checks of our main findings to sample attrition (Table A.2, Panel B).

[33] 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.

[34] See www.cpc.unc.edu/projects/rlms-hse/project/samprep for a detailed discussion of attrition in the RLMS. Gorodnichenko, Sabirianova Peter, and Stolyarov (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 surveyed dwellings. This corresponds to an annual average attrition rate of 7%. Moscow and St. Petersburg however have a response rate of only 60%. We therefore perform robustness checks excluding these two sampling units from our analysis.

Comparison of Survey Consumption with Administrative Data Figure A.1 compares consumption data obtained from the RLMS with corresponding National Income and Product Account (NIPA) data and alcohol retail sales. Panel (b) is taken from Figure 2 of Gorodnichenko, Martinez-Vazquez, and Sabirianova Peter (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) compares the evolution of the ratio of vodka 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 shadow economy increased dramatically, with estimates around 40% of GDP (Johnson et al. 1997). The RLMS on the other hand measures total alcohol consumed, including alcohol produced in the informal economy. Starting in 2008 the two measures are almost identical.

Panel C of Table A.2 shows that our results are robust to reweighting the data to match the share of registered vodka in total retail sales of alcoholic beverages.

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 men in 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. As mentioned in Section 2, 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 (2003), 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 vs. Age Effects

In this section we leverage the survey's panel dimension to provide additional non-experimental evidence for the new mechanism proposed in this paper. A common hypothesis for heterogeneity in alcohol consumption put forward in the health literature are "steppingstone" or "gateway" effects of light drugs for the consumption of harder drugs later in life. In the case of alcohol, this means that beer might serve as a steppingstone 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.[35] To the best of our knowledge our study is the first to analyze the steppingstone effect of light alcohol towards harder alcoholic beverages.

We decompose both alcohol shares into unconditional age and cohort effects. A steppingstone effect of beer would generate within-consumer variation where younger consumers start out with

[35] For instance, Mills and Noyes (1984) and Deza (2015) find evidence for a modest steppingstone effect of marijuana and alcohol for the consumption of harder non-alcoholic drugs later on. Similarly, Beenstock and Rahav (2002) find a steppingstone 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 steppingstone effects can explain many patterns of drug consumption.

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 changes in alcohol shares are instead 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 left 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, steppingstone 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 steppingstone 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 steppingstone effects.

The average individual's slope shown in the middle panel could mask a steppingstone 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 about age 22.

C The Anti-Alcohol Campaign: Robustness

In this section we provide an extensive sensitivity analysis of our benchmark result for the anti-alcohol campaign in Column 4 of Table 2. These robustness checks are shown in Table A.2.

First, we perform robustness checks of our main results to concerns related to the use of the survey data. In particular, we analyze the effect of sample attrition (Panel A), of different definitions of rural consumers (Panel C), and we provide an alternative estimation that reweights the data to match the share of registered vodka in total retail sales of alcoholic beverages (Panel B). Finally, Panel D shows various additional robustness checks.

C.1 Sample Attrition

We assess the robustness of our main results to sample attrition in Panel A of Table A.2. In Column 1 we drop the three sampling units with the highest attrition rates, St. Petersburg, 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.[36] Column 2 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 3 we find no systematic difference in the 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 4 and 5 we collapse the data to a single cross-section. In Column 4 we assign an individual to the year it was first sampled while in Column 5 we assign it randomly to any year in which it responded. The results are similar as the baseline estimates, although with larger point estimates and standard errors. The results in Panel A therefore suggest that sample attrition does not substantially affect our main result in Table 2.

Finally, because attrition is higher before 2001 as discussed in Online Appendix A, we use survey years 2001-2011 as our baseline sample in Table 2. Column 15 of Table A.2 shows that our results are qualitatively robust to using the full sample 1994-2011.

C.2 Comparison with Administrative Retail Sales Data

Online Appendix A shows that starting in 2008, administrative retail sales and alcohol consumption measured in the RLMS match up well. In Panel B we therefore test the robustness of our results to potential mismeasurement of alcohol consumption in the RLMS relative to administrative retail sales. In Column 6 we 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 7 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 benchmark specification. Therefore, the results in Table 2 do not seem to be affected by potential underreporting in the survey.

[36] 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.

C.3 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 addition to recording current residence, the survey also asks about the respondents' birthplace.

In Panel C we construct various indicators for whether an individual became adolescent in a rural area. In Column 8 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 9 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 10 and 11 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 10) or the broader definition of a rural area (Column 11). Both estimates are quantitatively similar to our benchmark result. Finally, in Column 12 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 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).

C.4 Additional Robustness Checks

In Panel D we provide additional robustness checks of the main results of the anti-alcohol campaign on taste formation in the long run.

In light of the beer market expansion discussed in Section 5, one might be concerned that consumers in the control group that became adolescent after the end of the campaign faced different initial conditions than consumers that turned 17 before the campaign, and hence that the former do not form a proper control group for the analysis in Section 3. Here, we address this concern in two steps.

First, we drop households that turned 17 after 1995 when the beer market started to expand due to large inflows of imports and foreign direct investments. Although we cannot reject the hypothesis that the effect in Column 13 is the same as our benchmark result, the larger point estimate suggests that consumers that turned 17 before the campaign might be a more appropriate control group.

Second, we extend the difference-in-differences design of equation (2) to include two different sets of control groups, one containing men who turned 17 before 1986, and another with men who turned 17 between 1991 and 1995:

$$\begin{aligned} S_{it}^{vodka} = {} & \beta_{DD,1} \cdot \text{I(became adolescent before campaign)}_i \times \text{I(urban)}_i \\ & + \beta_{DD,2} \cdot \text{I(became adolescent after campaign)}_i \times \text{I(urban)}_i \end{aligned} \quad (6)$$

$$+ \beta_{D,1} \cdot \text{I(became adolescent before campaign)}_i$$
$$+ \beta_{D,2} \cdot \text{I(became adolescent after campaign)}_i + \lambda \cdot \text{I(urban)}_i + \gamma' x_{it} + \epsilon_{it}.$$

This specification, shown in Column 14, effectively decomposes the effect in Column 13, supporting the intuition that older individuals who turned 17 before the start of the campaign might form a more appropriate control group.[37]

In Column 15 we extend the baseline sample by including all available survey years from 1994 to 2011. While the coefficients are again not statistically different from the baseline results, the lower point estimates suggest that using the earlier part of the sample leads to a downward bias since individuals' consumption shares have not yet reached their steady state.

We assess this conjecture in Column 16 by restricting the sample to survey years after 2005, therefore estimating the effect of the campaign in the very long run, more than 17 years after the end of the campaign. We indeed find that the effect is larger, although we again cannot reject that it is statistically different from the baseline estimate because of the larger standard error due to the small sample size.

In Column 17 we use the statutory start date of the campaign instead of the estimated date based on Figure 2. Adding individuals that turned 17 in year 1985 does not affect the results. However, Section 4 shows that this does not mean that the power of the research design is low.

One might also be concerned that our results could be sensitive to heavy drinkers or alcoholism. In Column 18 we address this concern showing that the results are robust to dropping all consumers in the top quartile of the alcohol consumption distribution.

In Column 19 we include all men age 14 and above, the lowest age at which individuals are asked to complete the health module. Our main analysis restricts the data to males age 18 and above because we are concerned with underreporting by individuals younger than the legal drinking age. However, Column 19 shows that we obtain similar results when including minors.

Column 20 also controls for permanent income. We measure permanent income by forming income quintiles by 10-year age groups (i.e., five income bins within each cell of individuals age 18-24, 25-34, 35-44, 45-54, 55-65), assigning missing income values a separate income bin. We then add fixed effects for each of these 30 cells. Column 20 shows that we again obtain similar results as in our benchmark specification.

The main empirical analysis follows the previous literature and uses the share of vodka consumed as the dependent variable. However, these two-sided truncated models can be severely biased if estimated with OLS including fixed effects. We use the estimator developed in Honoré (1992) to explore this issue. This pairwise trimmed least-squares estimator for truncated models in panel data with fixed effects is consistent in this setting and does not require parametric assumptions for the disturbance terms. Column 21 shows that this alternative estimator yields

[37] We interact the policy with the indicator for turning 17 in an urban area, I(urban), which is the complement of I(rural), because this specification maintains the sign for the difference-in-differences coefficient and hence makes it easier to compare causal effects across columns.

very similar results as our benchmark specification.

C.5 Placebo Tests

We perform three placebo tests that assign 5-year treatment windows to periods other than 1986-90. Instead of randomly assigning treatment windows, we show all possible assignments over the sample period starting in 1970, the first year we have estimates of aggregate samogon consumption.

Panel (a) of Figure A.7 shows the design of this first placebo test. Specifically, we use 15-year rolling windows starting with consumers who turned 17 between 1960 and 1974 and ending with the sample of men who turned 17 before 2002.[38] Within each sample we estimate the same difference-in-differences specification as in equation (2), with a 5-year treatment window in the center.

According to our identification strategy we should not see any effect of the placebo treatment in years prior to the actual campaign. This prediction holds conditional on 17 being the sensitive age for alcohol taste formation based on Section 4. As the 15-year sample enters the campaign period 1986-90, we should initially see $\hat{\beta}_{DD}$ *decrease* as the true treatment group gets mistakenly assigned to the control group on the right. The coefficient should then gradually increase as the assigned treatment group more and more covers the actual treatment period, reaching its peak around the 5-year period from 1986-90. If we assign the 5-year treatment indicator to periods after 1990, then the outcome will depend on how quickly tastes form. In Section 4 we showed that tastes for hard alcoholic beverages form in a narrow interval centered at age 17. Hence, $\hat{\beta}_{DD}$ should decrease back to zero, before becoming negative again as we falsely assign the actual treatment group to the control group on the left. Finally, the coefficient should gradually increase back to zero. Our difference-in-differences identification strategy therefore predicts a W-shaped pattern for $\hat{\beta}_{DD}$, which is a stronger test than the typical placebo test which would just predict no effect.

Panel (b) plots the evolution of $\hat{\beta}_{DD}$ together with 95% confidence intervals. Consistent with our research design, we see this W-shaped pattern emerge. The peak response occurs when the treatment window reaches the actual treatment period from 1986-90.

The second placebo in Panel (c) uses the same research design shown in Panel (a) except that we assign individuals to the treatment window if they turned 30 during that 5-year window (instead of 17 as in Panel b). We choose age 30 because based on our hypothesis about taste formation, consumers have already formed most of their alcohol tastes by that age and hence their preferences are no longer malleable. Panel (c) shows that we indeed do not find any effect of the anti-alcohol campaign on these middle-aged men.

The third placebo test in Panel (d) implements the same research design as in Panel (b) but uses an outcome variable that should not be affected by the anti-alcohol campaign. This

[38] 2001 is the last year in which we have data on tea consumption used in the second placebo test discussed below.

approach mimics actual placebo tests used in clinical trials. Since we want to perform the same test as before, which is at the level of the individual rather than the household, we need to use data from the same health module of the RLMS. This module is the only place where we see individual consumption as opposed to household-level expenditures. Fortunately, until 2001 the health module asked respondents whether and how often they drink tea. We use the weekly frequency of tea consumption as the dependent variable in this second placebo test.

Panel (d) shows that the difference-in-differences estimator is never statistically significant even though it has similar precision as the first placebo test in Panel (b). In particular, we do not find any effect of the actual anti-alcohol campaign from 1986-90 on tea consumption, providing credibility to our research design.

C.6 Using Different Kernels for the Taste Age Function

This section shows that our estimates of the typical age at which men form their alcohol tastes in Section 4 are robust to the choice of weighting kernel. To show this, we use two alternative kernels shown in the right panels of Figure A.8. The first is an empirical kernel shown in Panel (b), which reflects the treatment intensity we estimated using the regional difference-in-difference estimate in (1). The other is a 5-year uniform kernel shown in Panel (d), covering the campaign's duration from 1986-90. This kernel only uses information about the length of the intervention but not the intensity of the campaign, similar to the triangular kernel used in Figure 3.

Panels (a) and (c) show that we obtain similar age profiles for taste formation for those two alternative kernels. Hence, our results are robust to alternative choices of weights assigned to the different years spanned by the anti-alcohol campaign.

C.7 Long-Run Means for Urban and Rural Consumers

Finally, in this section we provide another way to visualize the effect of both events (the anti-alcohol campaign and the expansion of the beer market), which complements Figure 1.

Panel (a) of Figure A.9 shows the effect of the anti-alcohol campaign on the long-run shares of beer consumed by men who turned 17 around that time, separately for urban and rural consumers following your suggestion. For robustness, in panels (b) to (e) we also plot the corresponding graphs where we use the other definitions of rural areas corresponding to the robustness checks in Table A.2, Panel C. The figure shows that the differences in the share of beer consumption between urban and rural youth is statistically significant only for cohorts which turned 17 years old during the anti-alcohol campaign.

Figure A.10 extends this analysis over the entire sample period, covering both experiments. Similar to the regression approach in the paper, we combine two consecutive birth cohorts (e.g., birth years 1960 and 1961) to avoid having cells with too few observations. (In the regression approach in Figures 4, 5 and A.7, we use rolling windows of 5-year birth cohorts.) The top figure shows the share of beer consumed. We see a significantly larger long-run share of beer consumption among two consecutive two-year birth cohorts of urban consumers who turned 17 during the campaign relative to rural consumers who also turned 17 during the campaign. For other birth cohorts, the two series of beer shares instead track each other fairly closely. Stepping back and looking at the entire pattern of beer consumption shows a kink in the time series (or rather "cohort series"), with a rapid increase in the long-run share of beer consumed among consumers who became adolescent during the 1990s and 2000s, when the overall beer market expanded most rapidly.

The bottom figure shows the same analysis for the long-run average of the log of total alcohol consumption. Consistent with Column 5 of Table 2, this series is noisier than the series of beer shares (i.e., the regression estimates in Table 2 are less precise for the log of total alcohol than for the share of vodka). Zooming in on the period around the anti-alcohol campaign, we see a decrease in alcohol consumption among urban consumers relative to rural consumers, who are not much affected by the campaign. Looking at the overall series (both urban and rural consumers), we see a downward trend in total alcohol consumption for younger birth cohorts who became adolescent after the end of the Soviet Union, consistent with the beer market expansion.

D Identification Using A Migrants Research Design

In this section we use a completely different research design based on migrants. This is the main approach taken by the previous literature on taste formation, including Bronnenberg et al. (2012) and Atkin (2016). We use three sets of movers to provide additional independent evidence for the mechanism.

First, we use migrants that moved from rural to urban areas in Russia to complement our difference-in-difference analysis of the anti-alcohol campaign. Table A.5 shows the results from this exercise. Consistent with taste changes and the fact that vodka consumption is more prevalent in rural areas, Columns 1 and 2 show that individuals who moved from a rural area to a city and thus had easier access to liquor during their taste-forming years consume a significantly larger share of vodka. This difference is relative to both consumers that moved between cities—the reference group—and to consumers that always lived in the same urban location, as shown by the difference between the two groups, i.e., (a)-(b). The average share of vodka among all urban consumers is 54 pp and is more than 11 pp higher for individuals that moved from a rural area to a city. More than 2 pp of this difference cannot be attributed to either age, year, income, or relative price effects, or any other observable characteristics.[39]

Second, we use information about the birth country for individuals who moved to Russia from another republic of the former Soviet Union.[40] While vodka and beer production was

[39] We find similar results if we use the much smaller set of migrants from urban to rural areas.

[40] Unfortunately, we do not have information on the country of origin for immigrants from non-Soviet countries.

relatively uniform across countries of the former Soviet Union (although different for rural and urban areas), production of wine was heavily concentrated in only two republics, Moldova and Georgia.[41] Columns 3 and 4 show that migrants from those wine-producing Soviet republics consume a significantly larger share of wine compared to all other consumers. This effect is also economically significant. The wine share of immigrants from wine-producing republics is twice as large as that of all other consumers. Of this 4 pp difference, 3 pp cannot be explained by other covariates, and this difference is robust to using consumers that never moved as the reference group.

Third, we use the leave-out mean wine share by country of origin to construct a continuous measure of market exposure during the taste-forming years. The leave-out mean is the average consumption share among all immigrants from a given republic, *excluding* other individuals living in the same location, such as a town or city (the survey's so-called secondary sampling units). Column 5 shows that this leave-out mean is a good predictor of individual consumption shares. However, it might potentially be affected by local unobservables, a point recently emphasized by Angrist (2014). To address this issue we use a second, noisier measure of the individual's initial market conditions: aggregate domestic consumption data from the World Health Organization for years between 1991 and 2010 for each of the fifteen countries of origin in the survey. These average shares range from 65% in Georgia to 5% in Kazakhstan, while Russia's share is just 9%. We use the noisier but arguably more exogenous country-of-origin shares to instrument for the less noisy but potentially endogenous leave-out means. The IV estimates are qualitatively similar to the OLS estimate. The fact that the IV estimate in Column 6 is larger than the OLS estimate indicates measurement error in the leave-out mean. Finally, Column 7 shows that the results are robust to controlling for age, year, real income, relative prices, and any other observable characteristic, most importantly city fixed effects. Column 8 reports the corresponding first stage regression.

In summary, this analysis provides additional evidence of persistent alcohol tastes that are shaped by the socio-economic environment during adolescence. Hence, these results are consistent with the findings from the anti-alcohol campaign and the beer market expansion even though they are based on a completely different research design.

E Taste Changes for Non-Alcoholic Goods

This section addresses the concern that our results might only apply to addictive substances. In order to identify changes in long-run tastes for other non-alcoholic goods we use the opening of many other markets in the 1990s.

Identifying such tastes is more challenging. Conceptually, the hypothesis that tastes are formed when consuming a new good regularly for the first time implies that food tastes are

[41] A part of Russia, Krasnodarskiy Kray, and a part of Ukraine, Crimea, also produced wine, but these two regions are small relative to the size of the corresponding republic.

formed during childhood. This creates a problem since children do not necessarily make their own consumption decisions. Hence, the effect of the exogenous changes in market conditions due to imports in the late 1990s will be dampened by the accumulated tastes of the parents who are making consumption decisions on behalf of their children.

In addition to this conceptual problem, there are several measurement issues that further complicate the clean identification of changing tastes for non-alcoholic goods. First, the parents' own consumption tastes obviously depend on their age. Unfortunately, we do not know the age of the survey respondents' parents. Second, when analyzing non-alcoholic goods, we must rely on household-level expenditure data instead of the individual-level consumption data available in the survey's health module. These expenditure data might be measured with substantially more error. Moreover, several individuals can decide on the consumption bundle in a multi-person household. Unfortunately, there are only few single households in the data which would mitigate this problem. Similarly, there are only few households where both spouses were born in the same or a similar cohort. Therefore, it is important to realize that household-level expenditures reflect complex, aggregated preferences which make a direct mapping from changes in market conditions to cohort differences in consumption patterns difficult.

With the exception of certain types of meat, the expenditure questionnaire of the RLMS does not provide sufficient details about those new, more "exotic" or "western" goods that became available only after the fall of the Soviet Union, such as pineapples and bananas for example.[42] We therefore turn to a second source of micro-level expenditure data that has more detailed, disaggregated expenditures allowing us to differentiate between those new goods and more traditional goods in the same category (i.e., close substitutes) that were also available during the Soviet Union. 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 identify seven expenditure groups for which we can classify the goods as either new or traditional. Listing the new goods first, these are subtropical fruits such as pineapples and bananas vs. apples, pears and plums; chocolate vs. jam and honey for desserts; yoghurt vs. cottage cheese for breakfast; long-lasting vs. short-lived milk; frozen and canned fruits vs. dried fruits; and chicken vs. pork and beef for meat. The availability of the new goods is mostly caused by two factors, the import of previously unavailable goods, such as subtropical fruits, and the inflow of modern technologies, such as new ways to preserve milk or new technologies to produce chicken at much lower cost. Table A.3 provides more detail about our classification of each good and Table A.4 contains the corresponding summary statistics.

We restrict our analysis to households for which both head and spouse were born in the same 10-year cohort window to mitigate the preference aggregation issue. To have a sufficient sample

[42]For instance, in the RLMS we only have data on fresh fruits; dried fruits and berries; fresh berries; fruit and berry preserves; and melons and watermelons, including pickled and dried.

size, especially when estimating tastes good-by-good, we group the households into those born in the 1970s, the 1980s, and those born in the 1960s or earlier, which is the reference group. Because the survey was done in 2003 we do not have households born in the 1990s. Hence, the estimates in this analysis are likely lower bounds for the effect of the import shocks on long-run tastes since younger cohorts that are most responsive to the new market conditions have not formed their own households yet.

Table A.6 shows that consistent with tastes forming early in life, younger cohorts consume a significantly larger share of new "western" goods relative to traditional goods. This is true even after controlling for real income, family size as well as region respectively region-by-good fixed effects that capture relative price differences across regions. Column 1 uses all information in a pooled household-by-goods panel estimator, while Columns 2 to 8 show that the same pattern emerges good-by-good, although less precisely estimated.

Since NOBUS has only a single cross-section, we cannot separate cohort from age effects. We therefore turn again to the RLMS which contains sufficiently detailed data for one of the categories, chicken vs. beef and pork consumption. The RLMS also allows us to control for household age. Focusing on meat consumption has the additional advantage that we also have a long time-series of aggregate meat sales going back to 1970 to document these substantial changes. Figure A.5 shows similar rapid changes in the meat markets after the fall of the Soviet Union as in the alcohol markets. Columns 9 and 10 provide comparable estimates of the effects of the imports and foreign direct investments on the share of chicken consumed by younger cohorts in the RLMS as in the NOBUS data, even after we control for age and relative prices. The estimates are somewhat less precise due to the much smaller sample size of the RLMS.

Finally, we note a couple of potential shortcomings of this analysis. One potential confounding factor is learning. For example, if there are fixed costs to learning how to consume these new goods (e.g. how to prepare them), then younger consumers will benefit more from acquiring these skills over their life-cycle than older consumers even if they have the same preferences. Another potential confounding factor is status. Younger people may be more concerned with acquiring social status than older consumers. In that case, the larger consumption share of new, more expensive goods among young consumers may instead reflect signaling value, especially if the consumption of those goods is conspicuous.

F 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.

F.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.[43]

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

$$H_{t+1}^g = (1 - \delta)H_t^g + g_t,\ H_0^g \geq 0,\ \delta \in [0, 1]. \tag{7}$$

The budget constraint is $p_{v_t} v_t + b_t = y_t$. Without loss of generality, we focus on interior solutions.[44] The first-order condition of this optimization problem is

$$u_v(v_t, y_t - p_{v_t} v_t, H_t^v, H_t^b) - p_{v_t} u_b(v_t, y_t - p_{v_t} v_t, H_t^v, H_t^b) = 0\ , \tag{8}$$

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

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

[44]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.

be rewritten as

$$u_v(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) - p_v u_b(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) = 0\,. \tag{9}$$

In general, this is a non-monotonic function in the steady-state vodka consumption v.[45] Depending on the parametrization of the utility function u, equation (9) 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.[46] 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.

F.2 Model Properties and Extensions

This section shows that the model above with two habit forming goods can have any number of equilibria. 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.

F.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 $u_{gg} < 0$, $u_{H^g H^g} < 0$, 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.

[45] 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.

[46] 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.

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

F.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,

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

F.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,

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

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

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

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.

G 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

$$\hat{S}_i^g|_{H=1} = \hat{\varphi}_{1971} + \hat{\gamma}' x_i + \hat{\alpha}_{30}.$$

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 $H = 1$.

H Mortality and Long-Run Tastes for Vodka vs. Samogon

Table A.8 addresses the concern that if rural consumers consume more samogon during the anti-alcohol campaign, this could generate long-run tastes for samogon as much as for hard alcohol broadly, which in turn could explain the high rate of alcohol poisoning associated with greater vodka consumption in Column 5 of Table 3, because vodka and samogon consumption are likely positively correlated.

Recall that in Table 3 we use the difference-in-differences estimator from the anti-alcohol campaign to instrument for the share of vodka in the 2nd-stage IV regression of mortality on alcohol poisoning. Unfortunately, as shown in Column 2 of Table A.8, we do not find statistically significant changes in differences in the share of samogon consumed between urban and rural consumers for those who turned 17 during compared to those who turned 17 before or after the anti-alcohol campaign. Because of this, we cannot use this difference-in-differences estimator to instrument for the share of samogon, because it would be a weak instrument.

To address this concern, we therefore propose two other IV approaches. First, we note that there is a significant difference between rural and urban areas in the average shares of samogon consumed in the cross-section. We therefore use both the difference, I(rural), as well as the difference-in-differences, I(became adolescent during campaign) $\times$ I(rural), as instruments for the two endogenous variables in the regression of mortality on the share of vodka consumption and the share of samogon consumption. Column 1 of Table A.8 shows the 2nd stage and Columns 2 and 3 show the corresponding 1st stages for the share of samogon consumption (Column 2) and the share of vodka consumption (Column 3), respectively. Second, in addition to this IV regression shown in Columns 1 to 3, in Column 4 we also show another IV regression, where we only instrument for the share of samogon consumption using I(rural), which does not suffer from the weak instrument problem.

Both IV regressions suggest that the share of samogon consumption and the share of vodka consumption both increase mortality from alcohol poisoning. That is, controlling for the effect of samogon consumption on mortality does not decrease the effect of vodka consumption on mortality from alcohol poisoning. The point estimates suggest that the effect of vodka consumption on mortality is slightly larger, although the difference is not statistically significant.

Of course, these instrumental variables have their own drawbacks and are in our opinion less convincing that the analysis in the main text. Hence, this new IV approach only provides suggestive evidence that the effect of vodka consumption on alcohol poisoning is not driven by correlated changes in samogon consumption.

References

Alesina, A. and N. Fuchs-Schündeln (2007). Good-bye lenin (or not?): The effect of communism on people's preferences. American Economic Review 97(4), 1507–1528.

Angrist, J. D. (2014). The perils of peer effects. Labour Economics 30, 98–108.

Atkin, D. (2016). The caloric costs of culture: Evidence from indian migrants. American Economic Review 106(4), 1144–1181.

Becker, G. S. and K. M. Murphy (1988). A theory of rational addiction. Journal of Political Economy 96(4), 675–700.

Beenstock, M. and G. Rahav (2002). Testing gateway theory: do cigarette prices affect illicit drug use? Journal of Health Economics 21(4), 679–698.

Bernheim, B. D., D. Ray, and Ş. Yeltekin (2015). Poverty and self-control. Econometrica 83(5), 1877–1911.

Bhattacharya, J., C. Gathmann, and G. Miller (2013). The gorbachev anti-alcohol campaign and russia's mortality crisis. American Economic Journal: Applied Economics 5(2), 232–60.

Bisin, A. and T. Verdier (2001). The economics of cultural transmission and the dynamics of preferences. Journal of Economic Theory 97(2), 298–319.

Bowles, S. (1998). Endogenous Preferences: The Cultural Consequences of Markets and other Economic Institutions. Journal of Economic Literature 36(1), 75–111.

Brainerd, E. and D. M. Cutler (2005). Autopsy on an empire: Understanding mortality in russia and the former soviet union. Journal of Economic Perspectives 19(1), 107–130.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. American Economic Review 102(6), 2472–2508.

Carroll, C. D., B.-K. Rhee, and C. Rhee (1994). Are there cultural effects on saving? some cross-sectional evidence. Quarterly Journal of Economics 109(3), 685–699.

Charness, G. and U. Gneezy (2009). Incentives to exercise. Econometrica 77(3), 909–931.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. Annual Review of Economics 1(1), 451–488.

Cook, P. J. and M. J. Moore (2000). Alcohol. In A. J. Culyer and J. P. Newhouse (Eds.), Handbook of Health Economics, Volume 1, Chapter 30, pp. 1629–1673. Elsevier.

Costa, F. and F. Gerard (2018). Hysteresis and the welfare effect of corrective policies: Theory and evidence from an energy-saving program. NBER Working Paper No. 24608.

Danilova, I., V. M. Shkolnikov, D. A. Jdanov, F. Meslé, and J. Vallin (2016). Identifying potential differences in cause-of-death coding practices across russian regions. Population Health Metrics 14(1), 8.

Dawson, D. A. (2003). Methodological Issues in Measuring Alcohol Use. National Institute on Alcohol Abuse and Alcoholism.

Denisova, I. (2010). Adult mortality in russia. Economics of Transition 18(2), 333–363.

Deza, M. (2015). Is There a Stepping-Stone Effect in Drug Use? Separating State Dependence from Unobserved Heterogeneity Within and Across Illicit Drugs. Journal of Econometrics 184(1), 193–207.

Dhami, S. (2016). The Foundations of Behavioral Economic Analysis. Oxford University Press.

Doepke, M. and F. Zilibotti (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. Econometrica.

Federal State Statistics Service (1970–1991). Russian Statistical Yearbook. State Committee of the Russian Federation for Statistics.

Federal State Statistics Service (1992-2017c). Russia in figures. Rosstat, Moscow (Russia). available at https://www.gks.ru/bgd/regl/b14_12/IssWWW.exe/Stg/d02/21-10.htm (accessed September 4, 2018).

Federal State Statistics Service (1998–2001). Indicators of Social Development of the Russian Federation and its Regions. Goskomstat.

Federal State Statistics Service (2002-2017a). Regions of Russia. Socio-economic indicators. Rosstat, Moscow (Russia). available at http://old.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1138623506156 (accessed September 4, 2018).

Federal State Statistics Service (2002-2017b). Regions of Russia. The main characteristics of the constituent entities of the Russian Federation. Rosstat, Moscow (Russia). available at http://old.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1138625359016 (accessed September 4, 2018).

Fujiwara, T., K. Meng, and T. Vogl (2016). Habit formation in voting: Evidence from rainy elections. American Economic Journal: Applied Economics 8(4), 160–88.

Gavrilova, N. S., V. G. Semyonova, E. Dubrovina, G. N. Evdokushkina, A. E. Ivanova, and L. A. Gavrilov (2008). Russian mortality crisis and the quality of vital statistics. Population Research and Policy Review 27(5), 551.

Giné, X., D. Karlan, and J. Zinman (2010). Put your money where your butt is: a commitment contract for smoking cessation. American Economic Journal: Applied Economics 2(4), 213–35.

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

Gorodnichenko, Y., K. Sabirianova Peter, and D. Stolyarov (2010). Inequality and volatility moderation in russia: Evidence from micro-level panel data on consumption and income. Review of Economic Dynamics 13(1), 209–237.

Gruber, J. and B. Köszegi (2001). Is addiction "rational"? theory and evidence. Quarterly Journal of Economics 116(4), 1261–1303.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does Culture Affect Economic Outcomes? Journal of Economic Perspectives 20(2), 23–48.

Harsanyi, J. C. (1953). Welfare Economics of Variable Tastes. Review of Economic Studies 21(3), 204–213.

Honoré, B. E. (1992). Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. Econometrica 60(3), 533–565.

Human Mortality Database (1959–2014). Russia, Life tables (period 1x1). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). available at https://www.mortality.org/cgi-bin/hmd/country.php?cntr=RUS&level=1 (accessed June 25, 2014).

Ivanova, A. E., T. Sabgayda, V. Semenova, V. Zaporozhchenko, E. Zemlyanova, and S. Y. Nikitina (2013). Factors distorting structure of death causes in working population in russia. Sotsial'nye aspekty zdorov'ia naseleniia.

Johnson, S., D. Kaufmann, A. Shleifer, M. I. Goldman, and M. L. Weitzman (1997). The Unofficial Economy in Transition. Brookings Papers on Economic Ectivity 1997(2), 159–239.

Kueng, L. and E. Yakovlev (2020). Data and code for: The Long-Run Effects of a Public Policy on Alcohol Tastes and Mortality. Inter-university Consortium for Political and Social Research. American Economic Association. http://doi.org/10.3886/E117443V1.

Larcom, S., F. Rauch, and T. Willems (2017). The benefits of forced experimentation: striking evidence from the london underground network. Quarterly Journal of Economics 132(4), 2019–2055.

Malmendier, U. and S. Nagel (2011). Depression babies: Do macroeconomic experiences affect risk taking? Quarterly Journal of Economics 126(1), 373–416.

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

National Survey of Household Welfare and Program Participation (2003). Federal State Statistics Service (Rosstat), Moscow (Russia), and The World Bank, Washington D.C. (USA).

available at https://www.hse.ru/org/hse/4432173/mathbase/databases/db_32 (accessed May 6, 2015).

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

Nemtsov, A. V. (2002). Alcohol-related human losses in russia in the 1980s and 1990s. Addiction 97(11), 1413–1425.

Nemtsov, A. V. (2011). A Contemporary History of Alcohol in Russia. Södertörns högskola.

Rehm, J. and K. D. Shield (2013). Alcohol and mortality: Global alcohol-attributable deaths from cancer, liver cirrhosis, and injury in 2010. Journal of the National Institute on Alcohol Abuse and Alcoholism 35(2), 174–183.

Russian Fertility and Mortality Database (1989–2014). Mortality rates by age and sex. Center for Demographic Research, New Economic School, Moscow (Russia). available at http://demogr.nes.ru/index.php/en/demogr_indicat (accessed June 17, 2017).

Russian Longitudinal Monitoring Survey (1994–2011). Carolina Population Center, University of North Carolina, Chapel Hill (USA), and Higher School of Economics, Moscow (Russia). available at https://www.hse.ru/en/rlms (accessed July 20, 2012).

Stigler, G. J. and G. S. Becker (1977). De gustibus non est disputandum. American Economic Review 67(2), 76–90.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear iv regression. Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg, 80–108.

The World Bank (1970–2014). World Development Indicators. available at https://www.mortality.org/cgi-bin/hmd/country.php?cntr=RUS&level=1 (accessed March 7, 2016).

Treml, V. G. (1982). Alcohol in the USSR: A statistical study. Duke University Press Durham, NC.

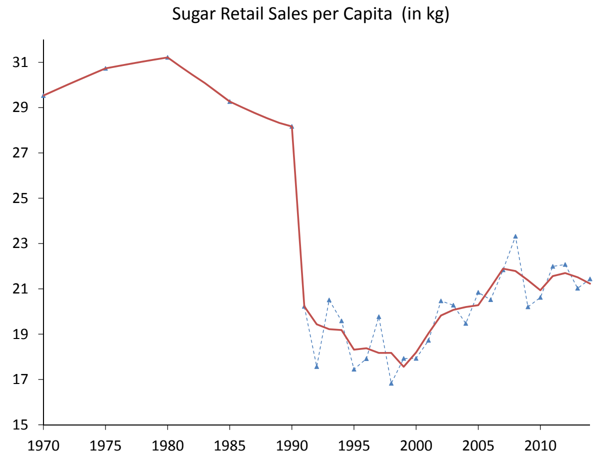
Treml, V. G. (1997). Soviet and russian statistics on alcohol consumption and abuse. Premature death in the new independent states, 220–238.

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

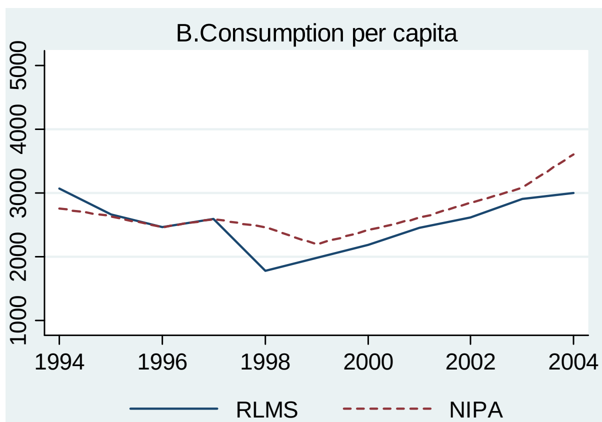
White, S. (1996). Russia Goes Dry: Alcohol, State and Society. Cambridge University Press.

Yakovlev, E. (2018). Demand for alcohol consumption in russia and its implication for mortality. American Economic Journal: Applied Economics 10(1), 106–149.

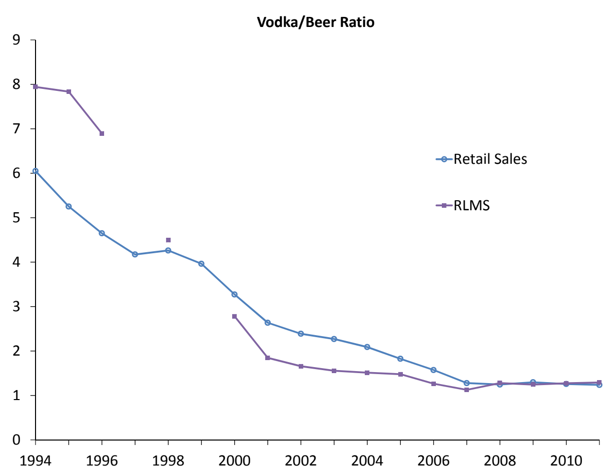
Figure A.1 – Comparison of RLMS Data with Official Statistics



(a) sugar consumption



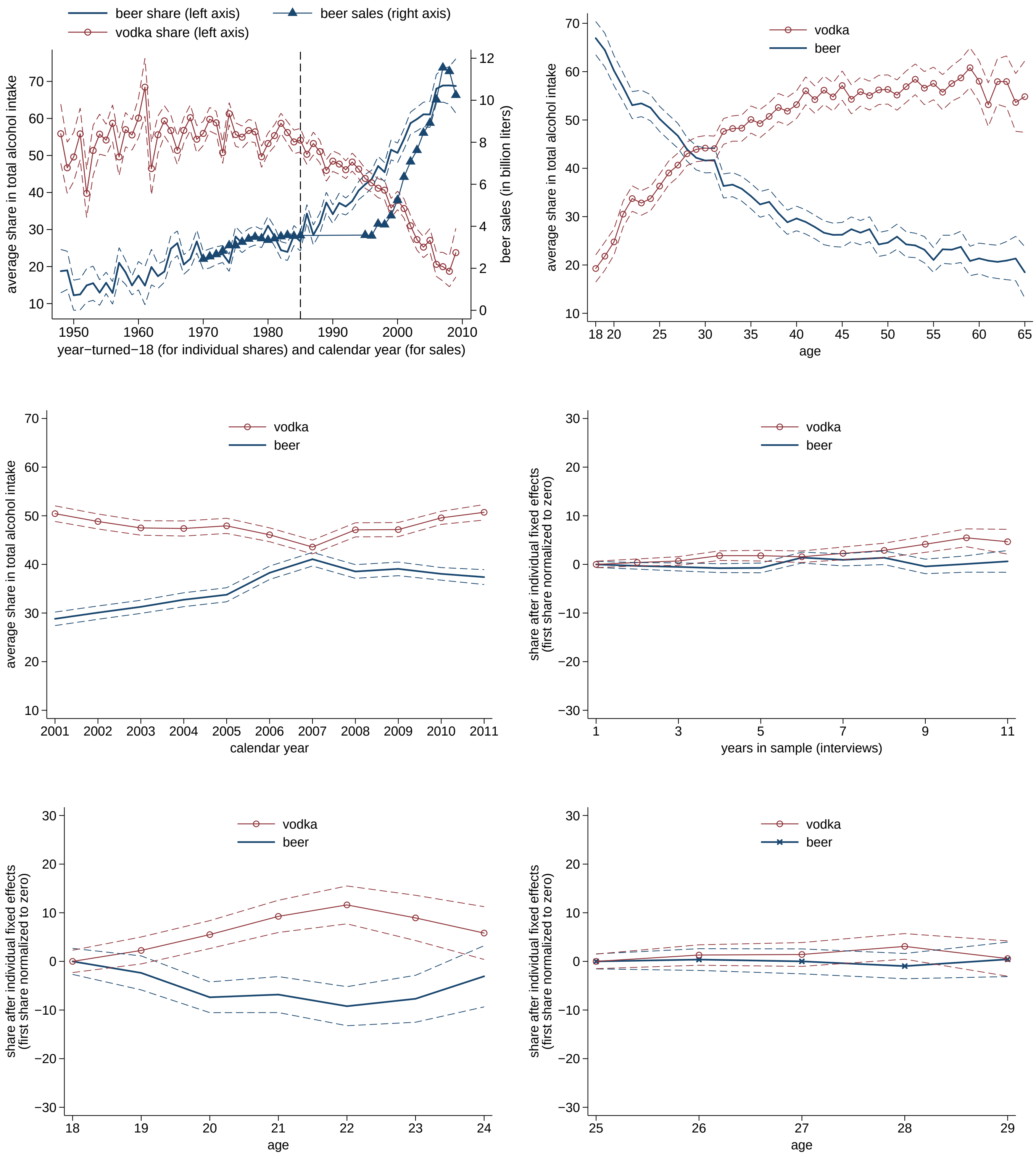
(b) comparison of RLMS and aggregate consumption



(c) comparison of RLMS and alcohol retail sales

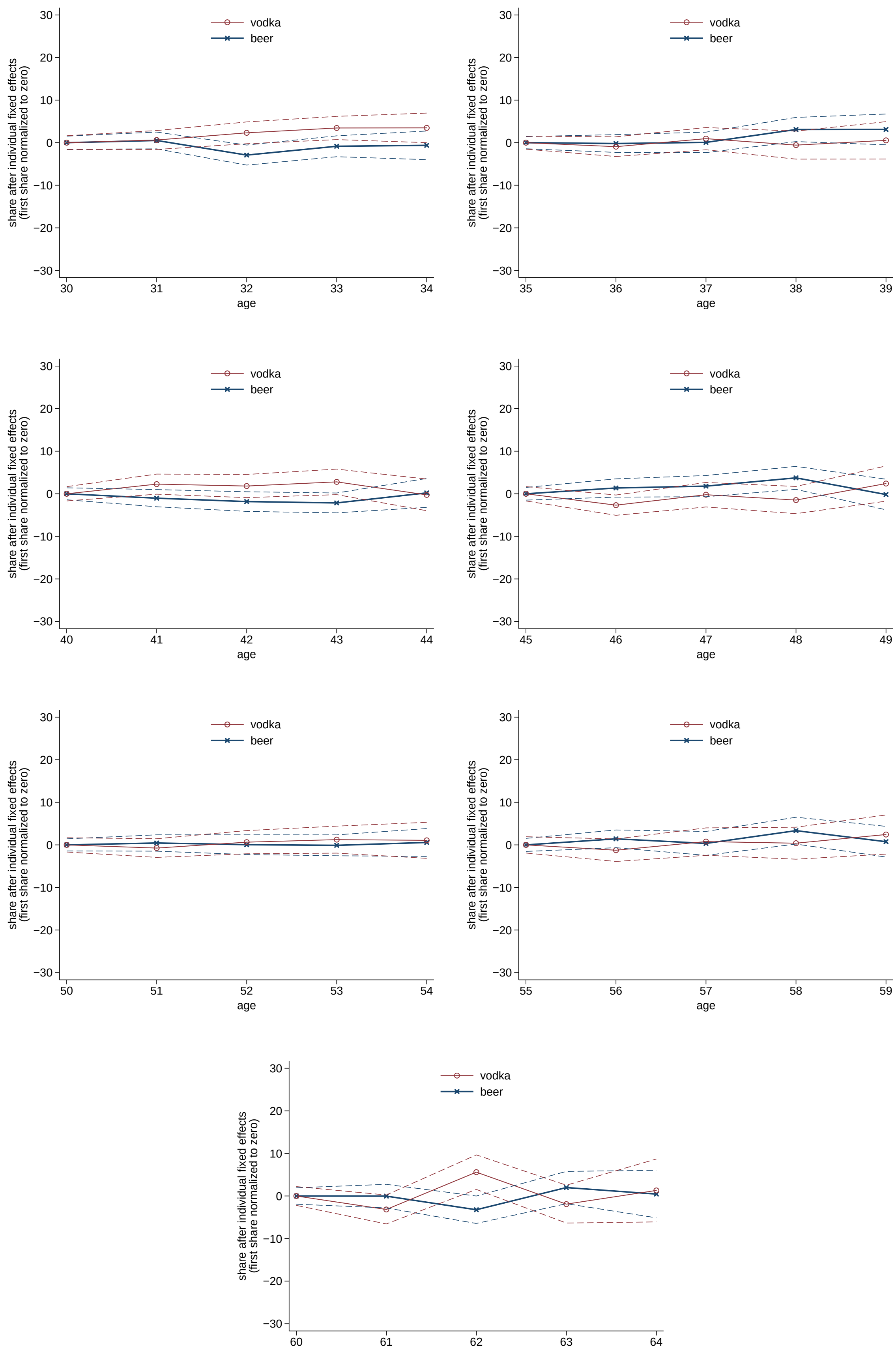
Notes: Panel (a) shows the time series of per capita retail sales of sugar from 1970-2014, an important ingredient of illicit vodka (samogon). The blue line is the raw data and the red line is the corresponding 2-year moving average. Panels (b) and (c) compare consumption data obtained from the RLMS with corresponding data from national accounts (NIPA) and alcohol retail sales. Panel (b) is taken from Gorodnichenko et al. (2009). Panel (c) compares the ratio of vodka to beer based on (official) registered alcohol retail sales and based on self-reported data in RLMS (both measured in pure alcohol).

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



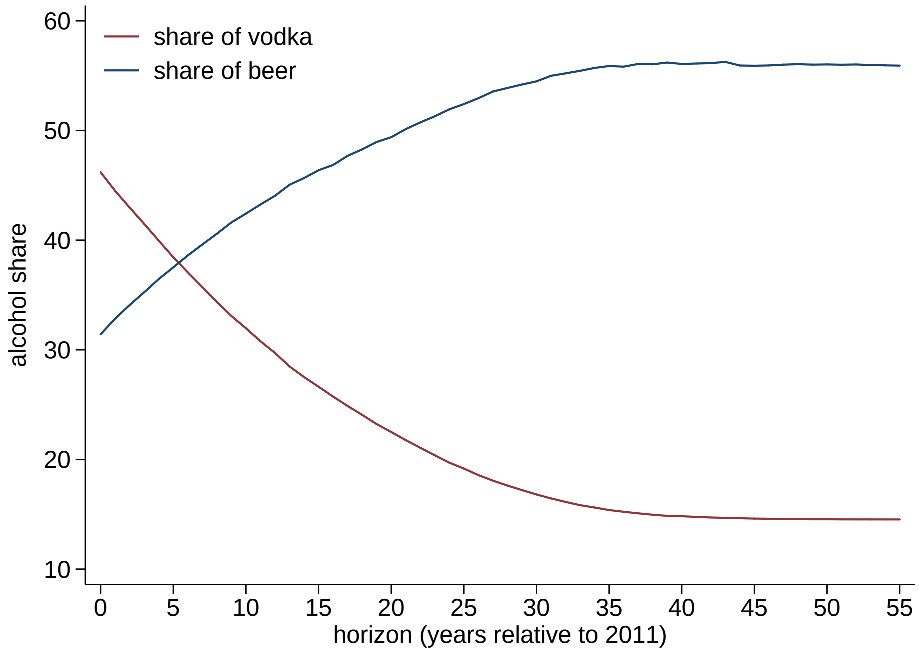
Notes: These figures show the profiles of the shares of beer and vodka consumed by men in the RLMS. The dashed lines represent two standard error confidence intervals. The top-left panel shows the age profile. The top-left panel shows the alcohol shares by cohorts measured by when and individual turned 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

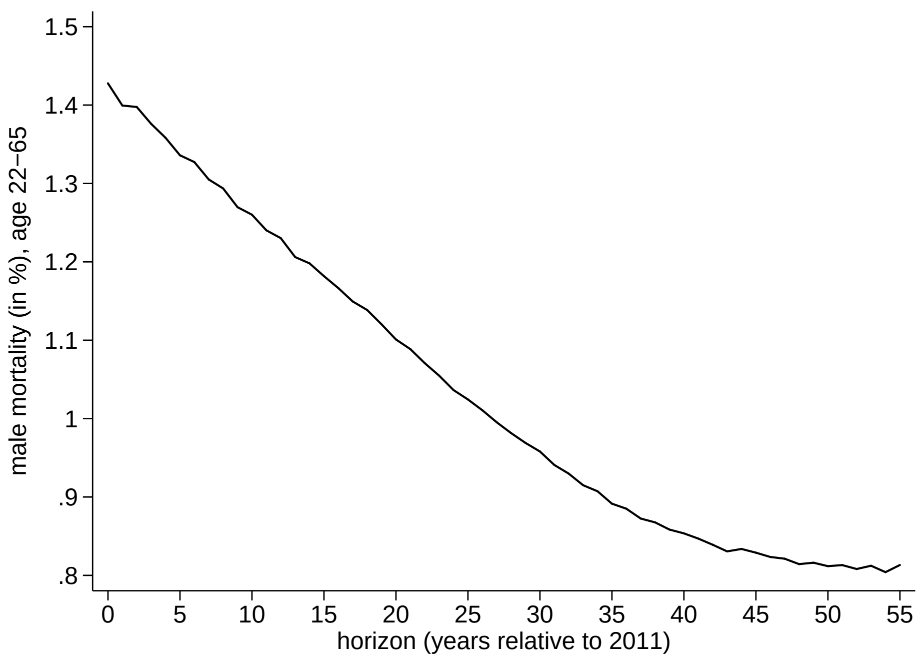


Notes: These figures provide the same analysis over the remaining part of the life-cycle as in Figure A.2.

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



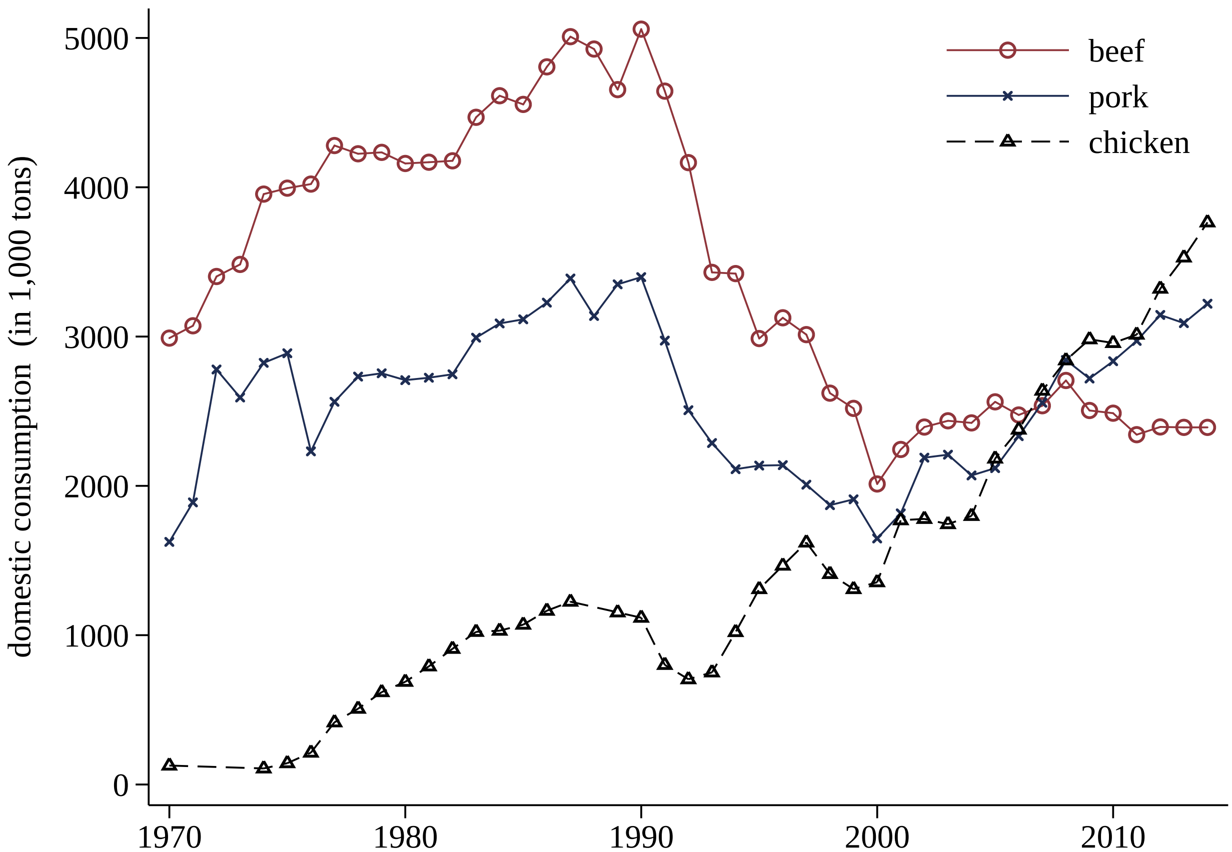
(a) predicted alcohol share



(b) predicted male mortality rate

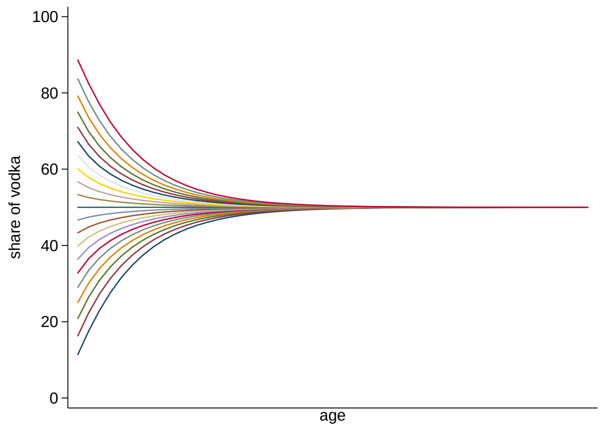
Notes: These figures shows the predicted consumption shares (Panel a) and implied mortality rates (Panel b) for males age 22 to 65 as a function of the forecast horizon in years.

Figure A.5 – Expansion of the Meat Market

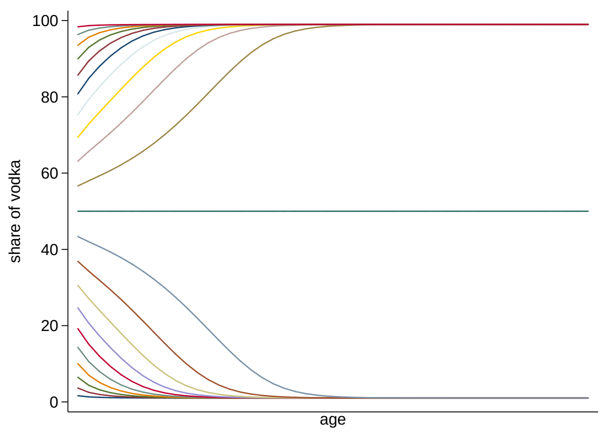


Notes: This figure shows the expansion of the meat market after the end of the Soviet Union.

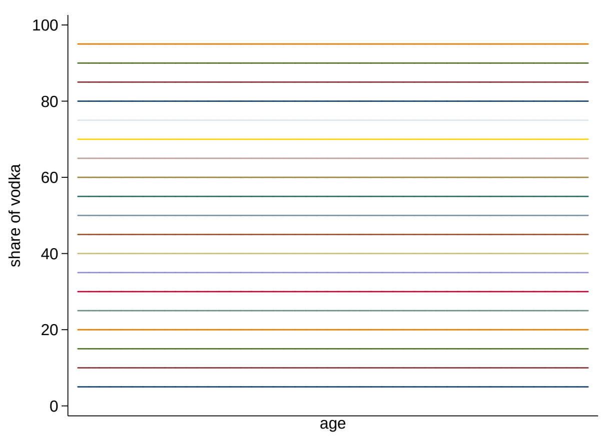
Figure A.6 – Potential Number of Steady States in the 2-Good Becker-Murphy Model



(a) one equilibrium



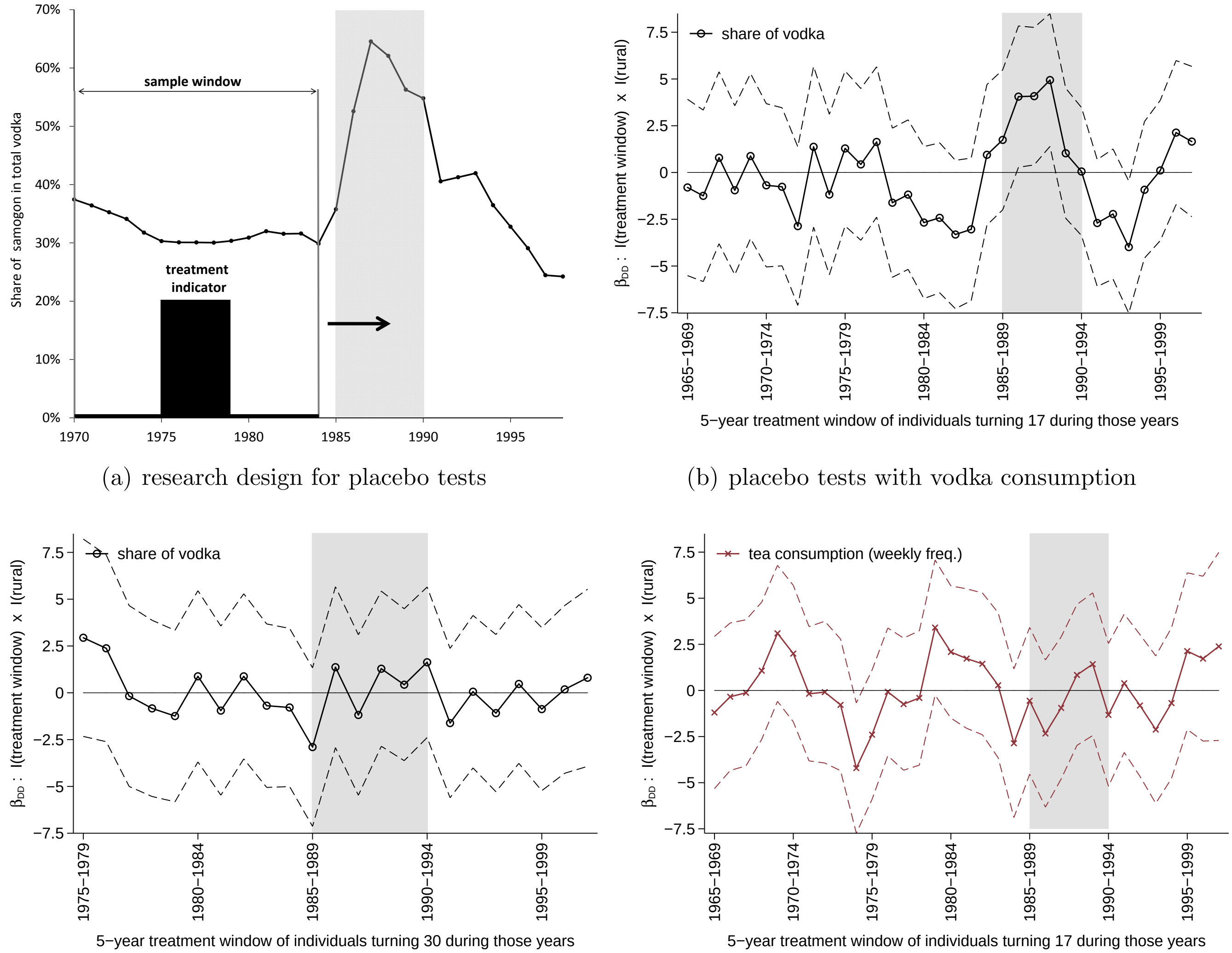
(b) three equilibria



(c) infinitely many equilibria

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. Panel (a) has one stable steady state, Panel (b) has three steady states, two stable and one unstable, and Panel (c) has an infinite number of steady states.

Figure A.7 – Placebo Tests for Anti-Alcohol Campaign



(a) research design for placebo tests

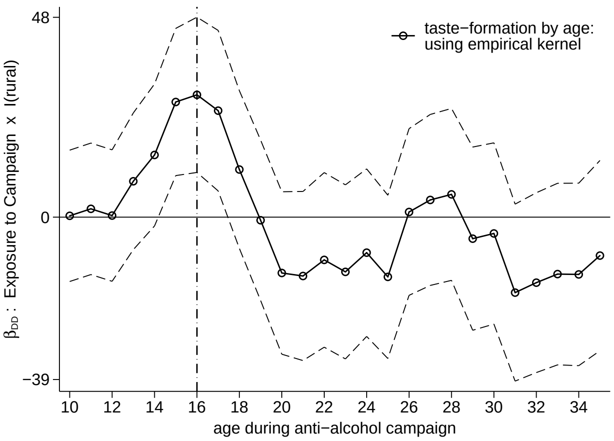
(b) placebo tests with vodka consumption

(c) placebo: (non)-response of 30-year olds

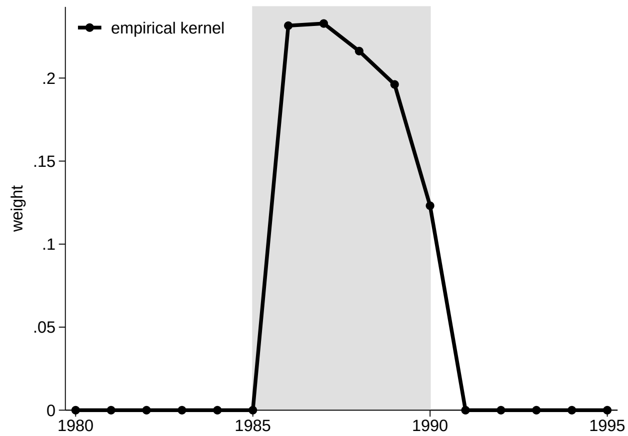
(d) placebo tests with tea consumption

Notes: This figure shows the design of the placebo tests (Panel a) together with the difference-in-difference estimates for vodka with men turning 17 in the 5-year treatment window (Panel b) respectively 30 (Panel c) and for tea consumption shares (Panel d). The anti-alcohol campaign is shaded in gray. Dashed lines are two standard error confidence bands using robust standard errors clustered by individual.

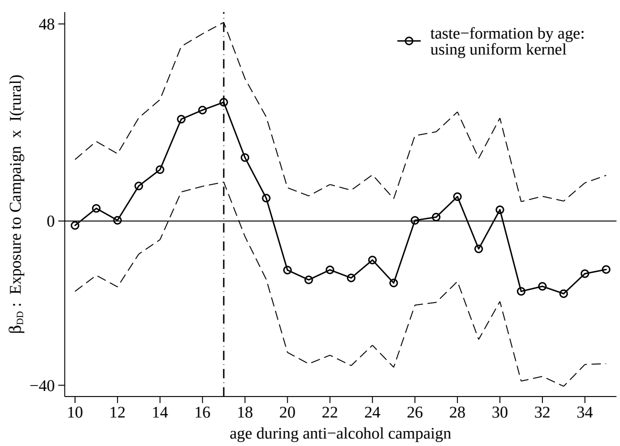
Figure A.8 – Taste Formation as a Function of Age – Robustness to Different Kernels



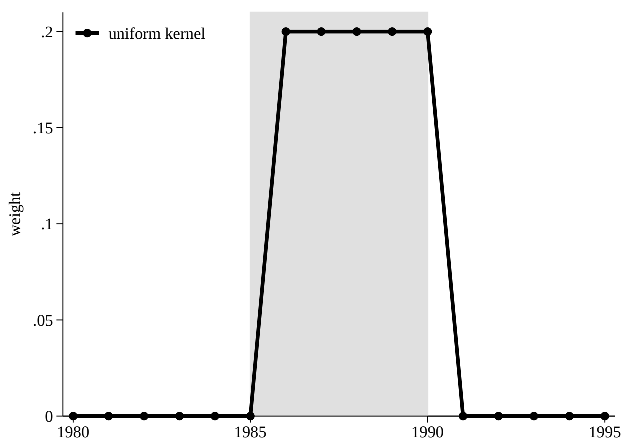
(a) response to anti-alcohol campaign by age



(b) empirical weighting kernel



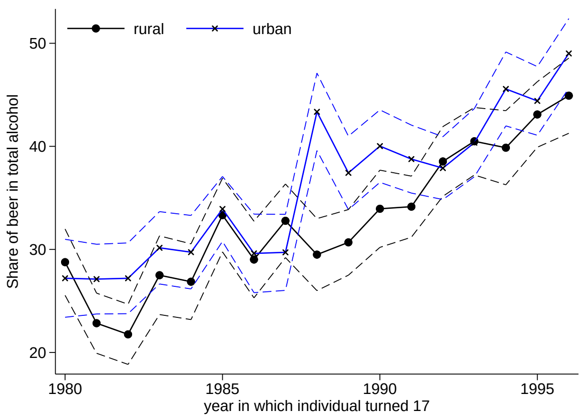
(c) response to anti-alcohol campaign by age



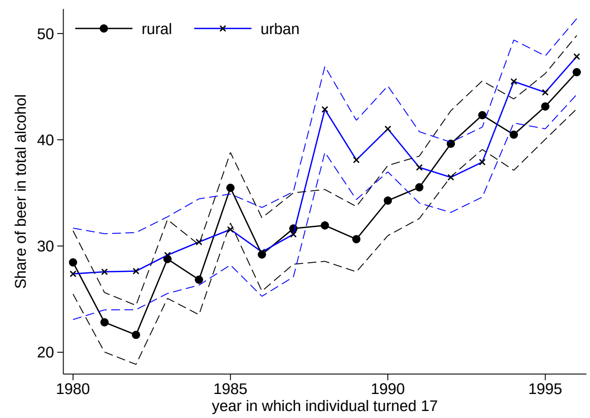
(d) 5-year uniform weighting kernel

Notes: This figure shows the robustness to using different weighting kernels (right panels) when estimating the alcohol consumption response to the anti-alcohol campaign by age (left panels). Panels (a) and (b) use a 5-year empirical kernel and Panels (c) and (d) a 5-year uniform kernel. Dashed lines are two standard error confidence bands using robust standard errors clustered by individual. The anti-alcohol campaign is shaded in gray in the right panels, and the maximum impact of the campaign on the age of taste formation is shown with a vertical dashed line in the left panels.

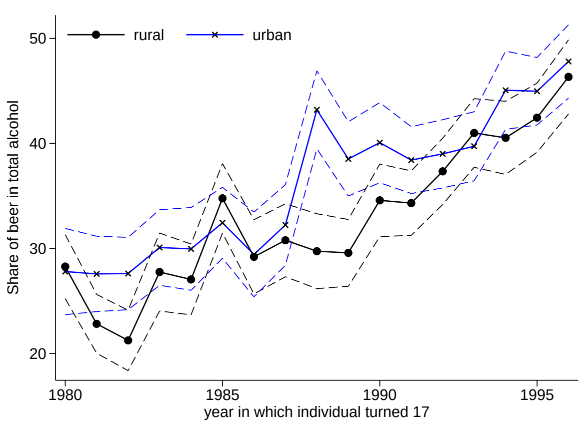
Figure A.9 – Long-Run Share of Beer Consumption of Urban and Rural Adolescents



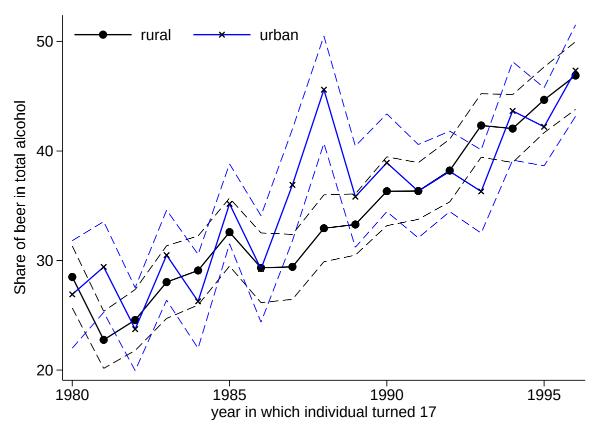
(a) main specification



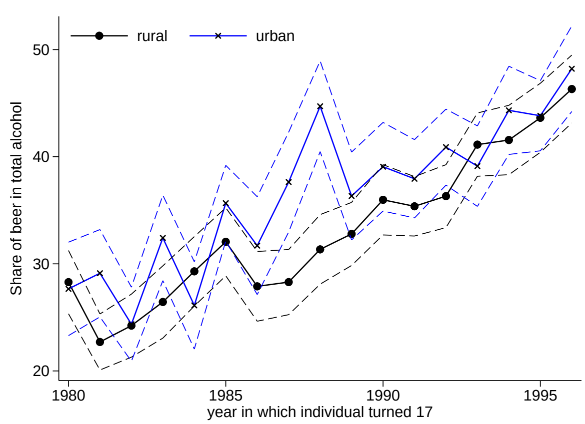
(b) using birth place or current population<250k



(c) using birth place or current population<100k



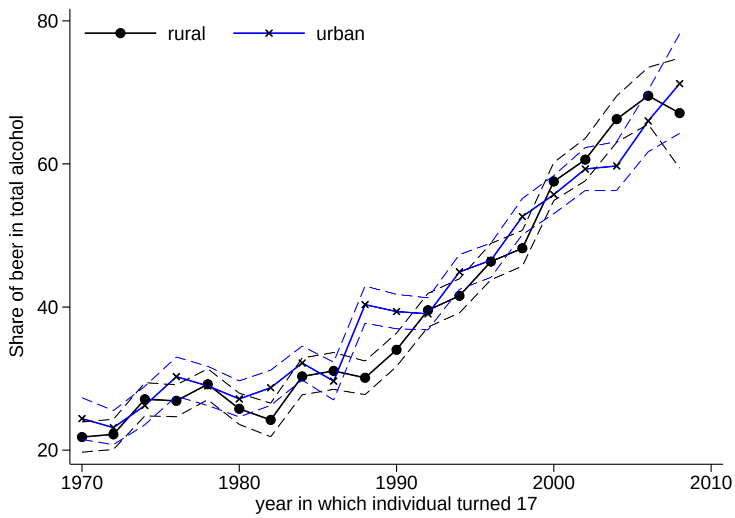
(d) using current population<100k



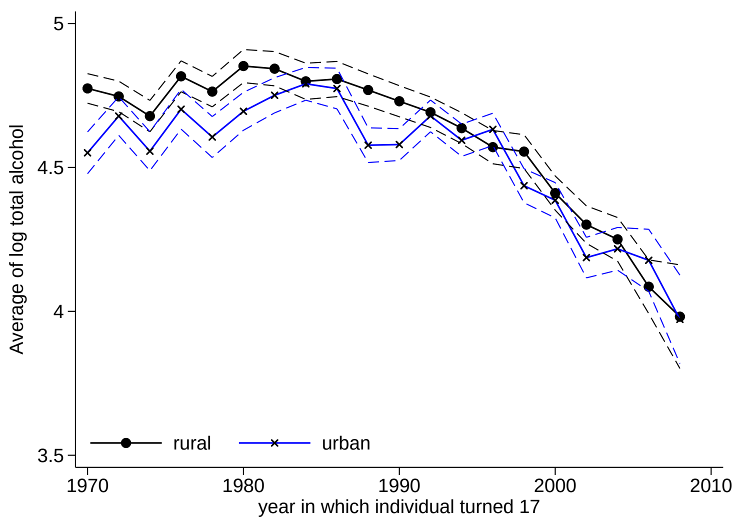
(e) using current population<250k

Notes: This figure shows the long-run shares of beer consumption of consumers who lived in a rural and urban area when they turned 17 as a function of the year in which when they turned 17, focusing on individuals that were adolescents in the years surrounding the anti-alcohol campaign. Panel (a) applies the main definition of what constitutes a rural area used throughout the main text. Panels (b) to (e) shows robustness to using different definitions of rural areas in line with the analysis in Table A.2, Panel C; see Section C.3. Dashed lines are two standard error confidence bands.

Figure A.10 – Long-Run Alcohol Consumption Patterns of Urban and Rural Adolescents



(a) share of beer consumption



(b) average log total alcohol

Notes: This figure shows—for the entire sample period—the long-run shares of beer consumption (panel a) respectively the long-run average log of total alcohol consumption (panel b) of consumers who lived in a rural and urban area when they turned 17 as a function of the year in which when they turned 17. To avoid having cells with too few observations, we combine two consecutive birth cohorts (e.g., birth years 1960 and 1961). Dashed lines are two standard error confidence bands.

Table A.1: Share of Vodka and Binge Drinking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of vodka (not in percent)	0.263*		-0.860***	-1.138***	-2.017***	-1.239***	-1.870***	-2.235***
	[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: The dependent variable is the number of days the respondents reports drinking alcohol. 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. RLMS Sample Attrition					B. Comparison with Retail Sales Data	
Specification:	drop Moscow and St. Petersburg	ATE by survey round	ATE by years in sample	using individual's first observation	using one observation per individual	using years 2008-2011 only	match retail sales vodka shares
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
I(became adolescent during campaign) × I(rural)	6.602***	5.909**	6.757**	6.935***	6.066**	7.356***	5.330**
	[2.139]	[2.483]	[2.876]	[2.591]	[2.572]	[2.589]	[2.283]
I(became adolescent during campaign) × I(rural)		-0.132					
		[0.378]					
I(became adolescent during campaign) × I(rural)			-0.202				
			[0.329]				
I(became adolescent during campaign)	-1.460	-1.100	-1.074	-1.460	-0.961	-2.376	-1.494
	[1.639]	[1.477]	[1.477]	[2.100]	[2.069]	[2.591]	[1.718]
Alcohol intake (in grams of ethanol)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year, age, region, rural FE	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	25,245	29,083	29,083	6,881	6,881	12,477	29,083
R-squared	0.154	0.153	0.153	0.128	0.133	0.192	0.167

	C. Using Different Definitions of I(rural)					D. Additional Robustness Checks	
Specification:	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	exclude beer expansion	use two control groups
	(8)	(9)	(10)	(11)	(12)	(13)	(14)
I(became adolescent during campaign) × I(rural)	5.995***	5.167**	5.812***	5.264***	3.911	6.668***	
	[1.965]	[2.042]	[1.964]	[1.982]	[2.968]	[2.067]	
I(became adolescent during campaign)	-2.177	-2.038	-1.428	-1.426	-0.810	-2.059	
	[1.566]	[1.723]	[1.488]	[1.571]	[2.423]	[1.549]	
I(became adolescent before campaign) × I(urban)							7.070***
							[2.124]
I(became adolescent after campaign) × I(urban)							5.307**
							[2.533]
I(became adolescent before campaign)							-4.156**
							[2.004]
I(became adolescent after campaign)							-4.485**
							[2.000]
Alcohol intake (in grams of ethanol)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year, age, region, rural FE	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	29,083	29,083	29,083	29,083	14,264	20,490	20,490
R-squared	0.153	0.153	0.152	0.152	0.125	0.088	0.088

	D. Additional Robustness Checks (continued)						
Specification:	all rounds, 1994-2011	very long run, 2006-11	campaign 1985-1990	drop heavy drinkers	include minors (age 14-17)	control for permanent income	Honoré (1992) estimator
	(15)	(16)	(17)	(18)	(19)	(20)	(21)
I(became adolescent during campaign) × I(rural)	4.661***	8.635***	5.172***	6.760***	5.065**	5.012**	6.094***
	[1.765]	[2.446]	[1.903]	[2.135]	[1.980]	[1.980]	[0.394]
I(became adolescent during campaign)	-1.898	-2.748	-1.639	-1.161	-0.958	-1.787	-0.490***
	[1.291]	[2.029]	[1.443]	[1.568]	[1.481]	[1.509]	[0.166]
Alcohol intake (in grams of ethanol)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year, age, region, rural FE	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
Income deciles x 10-year age group FEs						Yes	
Observations	29,276	11,734	20,490	16,291	30,017	29,083	29,083
R-squared	0.121	0.099	0.088	0.194	0.177	0.157	-

Notes: The dependent variable is the share of vodka consumption (in %). Socio-economic demographics include education, marital 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 2 and adolescence is defined as being 17 years old based on Section 5. 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: Descriptive Statistics of Other Non-Alcoholic Goods

Share of subtropical fruits in fresh fruits	5028	18.9	33.0	26.3
Share of chocolate in desserts	3350	83.1	35.4	100.0
Share of long-lived milk in milk	7488	5.5	21.1	0.0
Share of frozen fruits in preserved fruits	680	20.7	39.7	0.0
Share of yogurt in breakfast	5914	54.5	40.1	100.0
Share of salmon in salted fish	3650	25.8	41.8	60.4
Share of chicken in meat (NOBUS)	9492	51.4	43.6	100.0
Share of chicken in meat (RLMS)	6513	59.6	42.0	100.0

Notes: Data from the National Survey of Household Welfare and Program Participation (NOBUS), which was conducted in 2003 by the World Bank in collaboration with the Russian Statistical Office. Expenditures are measured at the household level.

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

	A. Using 1-Year Age Bins of Regional Mortality				B. Reduced-Form: Medium- and Long-Run			C. Reduced-Form: Cause-Specific		
Dependent variable :	log(male mortality), age 20-65			vodka share	log(male mortality), age 20-65			alc. poisoning	external	cancer
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Share of vodka	0.934***	0.960***	0.177***							
	[0.289]	[0.307]	[0.043]							
I(became adolescent during campaign) × I(rural)[(a)]				4.435***	4.011***	7.332***	3.004***	23.991***	7.666***	-1.129
				[0.935]	[0.517]	[1.569]	[0.475]	[5.575]	[2.421]	[7.240]
Log(total alcohol)		-0.491	0.211	0.360				0.036	0.055***	-0.005
		[0.376]	[0.229]	[0.224]				[0.048]	[0.020]	[0.032]
I(became adolescent during campaign)[(b)]	-5.888***	-6.183***	-5.198***	-7.506***	-6.999***	-7.753***	-6.392***	-17.508***	-8.872***	2.124
	[0.817]	[0.866]	[0.599]	[0.725]	[0.312]	[0.808]	[0.386]	[3.907]	[1.603]	[4.734]
Sum of diff-in-diff and diff coefficients, (a)+(b)					-2.988***	-0.422	-3.388***	6.482*	-1.206	0.995
					[0.407]	[1.360]	[0.389]	[3.658]	[1.739]	[5.811]
Log population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year, age, region, rural FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,400	1,400	1,400	1,400	52,345	11,412	40,933	1,327	1,343	1,273
R-squared	0.679	0.675	0.755	0.752	0.573	0.501	0.435	0.804	0.764	0.585
1st-stage F-statistic	16.57	15.74								
P-value of test that coefficient equals value in Table 3	0.27	0.31		0.05						
Sample period	1998-2011				1989-2014	1989-1997	1998-2014	1998-2011		
Estimator	IV	IV	OLS	OLS	OLS	OLS	OLS	OLS	OLS	OLS

Notes: Table uses male alcohol consumption data and mortality rates by year, region, age and type of settlement (urban/rural). Panel B covers additional years 2012-2014, which are not contained in the RLMS sample used to measure the age-specific alcohol consumption in Panel A. Reduced-form regressions in Panel B are weighted by population because they do not use data from the RLMS, which is not representative at the sub-national level. Panel C reports the reduced-form results corresponding to the IV estimates reported in Panel B of Table 3. Robust standard errors in parentheses are clustered by region; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.6: Identification of Long-Run Tastes for Alcoholic Drinks using Migrants

	A. Migrants to Cities		B. Immigrants from other Soviet Republics					
Dependent variable:	Share of vodka		Share of wine					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I(born in a rural now living in an urban area) [(a)]	11.232***	2.647**						
	[0.984]	[1.062]						
I(immigrated from Georgia or Moldova)			4.010***	2.924**				
			[1.514]	[1.443]				
Share of wine by country of origin (leave-out mean)					0.394***	0.988*	0.609	
					[0.138]	[0.507]	[0.406]	
Share of wine in aggregate alcohol sales of country of origin								0.056***
								[0.006]
I(always lived in the same urban location) [(b)]		0.270		0.291			-0.095	0.711***
		[0.901]		[0.222]			[0.357]	[0.013]
Alcohol intake (in grams of ethanol)		0.046***		-0.012***			-0.012***	-0.000
		[0.004]		[0.001]			[0.001]	[0.000]
Year, age, geography FE		Yes		Yes			Yes	Yes
Socio-economic demographics		Yes		Yes			Yes	Yes
Real income and relative price		Yes		Yes			Yes	Yes
Observations	19,107	19,107	44,028	44,028	43,819	43,819	43,819	43,819
R-squared	0.016	0.181	0.000	0.051	0.000	0.000	0.015	0.352
Difference (a)-(b)		2.377**		2.632*			0.704	
		[1.007]		[1.448]			[0.719]	
Weak-IV F-statistic (Kleibergen-Paap)						54.2	74.8	
Sample mean of dependent variable	54.4	54.4	4.2	4.2	4.2	4.2	4.2	4.2
Specification	OLS	OLS	OLS	OLS	OLS	IV	IV	1^{st} stage

Notes: Socio-economic demographics include education, marital status, body weight, and subjective health status. Geographical FE are region FE in columns 1-4 and city-level FE in columns 5-8. Robust standard errors in parentheses are clustered by individual; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.7: Identification of Long-Run Tastes for Non-Alcoholic Goods using 1990s Market Expansions

	A. Share of New "Western" Goods (NOBUS)								B. RLMS	
Dependent variable (in shares):	all new goods	subtropical fruits	chocolate	yogurt	long-lasting milk	frozen fruits	salted salmon	chicken	chicken	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
I(born in 1990s)									16.232***	9.408*
									[4.465]	[4.945]
I(born in 1980s)	11.930***	8.621***	9.157***	26.248***	2.909	30.807**	15.196***	10.737***	11.214***	7.005***
	[1.152]	[2.557]	[1.728]	[2.297]	[1.769]	[13.912]	[4.253]	[3.067]	[1.457]	[2.669]
I(born in 1970s)	7.173***	5.551***	6.814***	19.252***	1.764	27.584***	15.885***	-2.302	6.655***	4.952**
	[0.743]	[1.589]	[1.530]	[1.830]	[1.106]	[6.890]	[2.542]	[1.905]	[1.704]	[2.359]
Log(real income)	-0.032	0.211*	-0.165	-0.166	-0.002	-0.621*	-0.048	-0.000		-0.164
	[0.048]	[0.113]	[0.138]	[0.127]	[0.059]	[0.357]	[0.162]	[0.122]		[0.187]
Family size	-0.201	0.873	5.574***	2.869***	-0.357	2.672	1.813	-7.309***		-7.333***
	[0.350]	[0.813]	[0.995]	[0.912]	[0.432]	[2.678]	[1.190]	[0.890]		[0.846]
Age										-0.103
										[0.074]
Relative price of chicken to pork										4.524
										[5.027]
Relative price of chicken to beef										-0.440
										[4.765]
Region x product FE	Yes									
Region FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes		Yes
Year FE										Yes
Observations	44,186	6,576	4,584	7,504	10,075	845	5,110	9,492	6,513	6,513
R-squared	0.365	0.052	0.061	0.102	0.059	0.196	0.043	0.067	0.011	0.094

Notes: Robust standard errors in parentheses are clustered by individual; ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.8: Effect of Vodka and Samogon on Mortality from Alcohol Poisoning

Dependent variable:	2nd Stage: alc. poisoning	1st Stage: samogon share	1st Stage: vodka share	2nd Stage: alc. poisoning	1st Stage: samogon share
	(1)	(2)	(3)	(4)	(5)
Share of vodka	4.140***				
	[1.488]				
Share of samogon	2.671*			2.964***	
	[1.363]			[1.039]	
I(became adolescent during campaign) × I(rural)		-0.711	6.254***	26.098***	-0.711
		[1.124]	[1.957]	[6.343]	[1.124]
I(rural)		5.077***	0.360		5.077***
		[0.739]	[1.319]		[0.739]
Log(total alcohol)	-0.221*	0.005	0.059***	0.020	0.005
	[0.121]	[0.008]	[0.013]	[0.054]	[0.008]
I(became adolescent during campaign)	-10.021*	0.096	-1.871	-17.794***	0.096
	[5.278]	[0.600]	[1.302]	[4.513]	[0.600]
Log population	Yes	Yes	Yes	Yes	Yes
Year, age, region, rural FE and log income	Yes	Yes	Yes	Yes	Yes
Observations	1,327	1,327	1,327	1,327	1,327
R-squared	0.599	0.293	0.238	0.737	0.293
1st-stage F-statistic	52.05	40.32	9.770	47.18	47.18
Estimator	IV	OLS	OLS	IV	OLS

Notes: This table extends the analysis shown in Table 3 using regional alcohol sales by year, region, age and type of settlement (urban/rural) from 1994-2011 and alcohol poisoning fatalities. Note that while Columns 2 and 5 report the same coefficients, the F-statistics are different. Column 2 reports the F-statistic for joint significance of two instruments, I(rural) and I(became adolescent during campaign) × I(rural), with the corresponding two 1st-stage regressions shown in Columns 2 and 3. Column 5 reports the F-statistic for significance of one instrument, I(rural), with the corresponding single 1st-stage regression shown in Column 5. The set of additional controls includes log population, log regional GDP per capita and year, age, region and rural FE. Robust standard errors in squared brackets are clustered by region. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.