NBER WORKING PAPER SERIES

THE LONG-RUN EFFECTS OF A PUBLIC POLICY ON ALCOHOL TASTES AND MORTALITY

Lorenz Kueng
Evgeny Yakovlev

Working Paper 20298
http://www.nber.org/papers/w20298

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2014, Revised February 2020

We thank Lucas Davis (the editor), two anonymous referees, Orley Ashenfelter, David Atkin, David Card, David Cutler, Irina Denisova, Yuriy Gorodnichenko, Sergey Guriev, Lee Lockwood, David Matsa, Brian Melzer, Michael Moore, Denis Nekipelov, Jonathan Parker, Paola Sapienza, Katya Zhuravskaya, and seminar participants at the NBER Summer Institute, Berkeley, Northwestern, Warwick, Pompeu Fabra, NES, CEFIR, HSE, and the EEA meetings for helpful discussions and comments. Evgeny Yakovlev gratefully acknowledges financial support from the Russian Science Foundation for the research project No. 18-18-00466. Two previous versions of this paper were titled "USSR babies: Who drinks vodka in Russia?" and "How Persistent Are Consumption Habits? Micro-Evidence from Russia's Alcohol Market." The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Long-Run Effects of a Public Policy on Alcohol Tastes and Mortality
Lorenz Kueng and Evgeny Yakovlev
NBER Working Paper No. 20298
July 2014, Revised February 2020
JEL No. D12,E21,G02,I10

ABSTRACT

We study the long-run effects of Russia's anti-alcohol campaign, which dramatically altered the relative supply of hard and light alcohol in the late 1980s. We find that this policy shifted young men's long-run preferences from hard to light alcohol decades later and we estimate the age at which consumers form their tastes. We show that the large beer market expansion in the late 1990s had similar effects on young consumers' tastes, while older consumers' tastes remained largely unchanged. We then link these long-run changes in alcohol consumption patterns to changes in male mortality. The shift from hard to light alcohol reduced incidences of binge drinking substantially, leading to fewer alcohol- related deaths. We conclude that the resulting large cohort differences in current alcohol consumption shares explain a significant part of the recent decrease in male mortality. Simulations suggest that mortality will continue to decrease by another 23% over the next twenty years due to persistent changes in consumer tastes. Program impact evaluations that focus only on contemporaneous effects can therefore severely underestimate the total effect of such public policies that change preferences for goods.

Lorenz Kueng
Northwestern University
Kellogg School of Management
Suite 4394
2211 Campus Drive
Evanston, IL 60208
and NBER
l-kueng@kellogg.northwestern.edu

Evgeny Yakovlev
100A Novaya str, Skolkovo,
Moscow, Russia
evgeny.yakovlev@gmail.com

1 Introduction

Most economic analysis assumes tastes are fixed, and many economists are understandably wary of allowing preferences to change. For instance, allowing tastes to change makes it more difficult to derive testable predictions and to identify causal links (Stigler and Becker 1977).[1] Nevertheless, a rapidly growing literature is accumulating evidence of taste changes using two main research designs to overcome this identification challenge: comparing behavior of migrants and natives, and of children and their parents (Carroll, Rhee, and Rhee 1994, Alesina and Fuchs-Schündeln 2007, Bronnenberg, Dubé, and Gentzkow 2012, Atkin 2016).

An important open question is whether *public policies* can affect tastes (in addition to the documented effects of migration and parenting), and in particular whether even *temporary* policies can change important long-run outcomes through this mechanism (Bernheim, Ray, and Yeltekin 2015).

To answer this question, we study the long-run effects on taste formation for different alcoholic beverages of two large changes to Russia's alcohol markets. We first document that the brief anti-alcohol campaign in the late 1980s still affects alcohol consumption choices today, decades after the policy ended. While the intervention sharply reduced the availability of all types of alcohol in urban areas, there was a significant increase in homemade vodka (moonshining), which was much more concentrated in rural areas. At the same time, homemade light alcohol did not significantly increase such that hard alcohol was relatively less rationed in rural areas during the campaign.

Because the policy affected the relative supply of particular types of alcohol differently in rural and urban areas, we can estimate its long-run effects on taste formation by comparing the alcohol consumption behavior of rural and urban consumer who started drinking during the campaign with consumers who started drinking before or after the campaign (i.e., a difference-in-differences research design).

Under the null hypothesis of exogenous tastes there should be no difference in consumption behavior between rural and urban consumers across cohorts more than twenty years after the campaign ended. We test this hypothesis against the alternative hypothesis that consumers form tastes when they start consuming a certain good regularly. In the case of alcohol, this happens during adolescence for most consumers.

We strongly reject the null hypothesis of exogenous tastes. Instead, we find that consumers who became adolescent in a rural area during the campaign consume today a much larger share of their alcohol in the form of vodka compared to both their urban counterparts and to other rural consumers who became adolescent shortly before or after the campaign. These consumption differences persist twenty years after the policy ended, even though the treated consumers, who are now in their late 30s, have access to the same product selection as the untreated consumers,

[1] Taste changes also complicate welfare comparisons; see e.g., Harsanyi (1953) for an early analysis and the survey by Bowles (1998). Dhami (2016, Ch. 22) reviews related topics in behavioral welfare economics.

and even after we control for relative prices, income, age, and other individual characteristics.

Our methodology also allows us to estimate the sensitive ages at which policies affect individuals' tastes, exploiting the fact that the anti-alcohol campaign only lasted for a couple of years. We non-parametrically estimate that the typical consumer forms tastes over distinct types of alcoholic drinks between ages 14 and 18, and most consumers reach their steady-state tastes by age 22. These tastes do not change much afterward even in response to large shocks. To the best of our knowledge, this is one of the first studies that causally identifies the age at which tastes form.

We then show that these taste changes have important consequences for one of the most pressing public policy concerns in Russia: the high mortality of working-age adults and the large gender gap in life expectancy.[2] Because estimating the link between alcohol tastes and mortality is challenging given the large number of potential confounds inherent with natural experiments, we establish this relationship in three steps.

First, we show that there is a quantitatively important association between alcohol sales and male mortality in Russia using national-level data going back to the 1970s. Changes in the market *share* of vodka are strongly positively correlated with changes in male mortality, even when we control for the level of alcohol sales. Surprisingly, as far as we know, this is the first study which shows that the effect of alcohol on mortality strongly depends on the type of drink consumed, in addition to the amount of total alcohol intake, which has been the focus of most previous work in epidemiology and health economics. While both channels affect mortality, we find that the type of drink consumed is quantitatively at least as important as the amount of alcohol consumed. This is because most alcohol-related deaths of working-age adults—fatal traffic accidents, homicides, suicides, and fatal alcohol poisoning—are a consequence of binge drinking.

Second, we leverage the anti-alcohol campaign as an instrument to estimate the causal long-run effect of changes in alcohol tastes on male mortality using detailed regional-level data on causes of death and alcohol sales by type of alcohol. We further highlight the underlying mechanism by focusing on the narrow subset of deaths due to alcohol poisoning. The strong effect of long-run taste changes on the rate of alcohol poising establishes a direct link between contemporaneous changes in the market share of hard alcohol and mortality resulting from binge drinking. We then show that these effects extend to external causes of death that are indirectly related to alcohol consumption (such as fatal traffic accidents, homicides, suicides, etc.) and hence generate externalities on third parties, but they do not extend to other causes of death that should not be closely related to contemporaneous alcohol consumption patterns (such as cancer, excluding liver cancer).

Third, we link alcohol consumption directly to mortality using individual-level data from

[2] Average male life expectancy in Russia from 2000 to 2009 was only 60 years, which is 15 years lower than in the U.S. and even 3 years lower than in North Korea. Moreover, the gap between female and male life expectancy is 13 years while only 5 years in the U.S. Both facts are widely attributed to alcohol consumption.

a nationally representative consumption survey. We find that the share of vodka consumed is an economically and statistically significant predictor of the mortality rate of individuals in the survey even after controlling for the amount of alcohol consumed and other observables.[3]

Having established the causal long-run effect of relative supply shocks on alcohol tastes and mortality using the anti-alcohol campaign allows us to study the second large change to Russia's alcohol market: the large expansion of the beer market in the late 1990s. Many goods that were not readily available during the Soviet Union—including high-quality beer—became suddenly accessible to the broader public after Russia opened its borders to trade and foreign investments.

While studying these changes is empirically more challenging because many other things also changed after the collapse of the Soviet Union, these shocks were even larger than the anti-alcohol campaign and were also permanent as seen in Panel (a) of Figure 1. Within only one decade, the share of beer sales increased from less than 10% to over 25%.

[Figure 1 about here]

Panel (b) shows that the alcohol tastes of different cohorts were affected very differently depending on whether they became adolescent before, during or after this expansion. The figure shows that most of the increase in beer consumption is driven by cohort effects—especially for the young post-Soviet cohorts—and changes in alcohol consumption patterns within cohorts over time are much smaller.[4]

We use our analysis to simulate two policy-relevant counterfactuals. First, we find that a significant part of the recent decline in male mortality since 2003 can be attributed to the long-run effect of the large beer market expansion that occurred a decade earlier. Second, we estimate a hazard model with individual-level micro data and use it to predict that going forward, male mortality will further decrease by a quarter over the next twenty years as a long-run consequence of these taste changes. This decrease will occur even under the current set of policies, current levels of relative prices, and current socio-demographic characteristics of the population, except for individuals' tastes. Mortality will decrease simply because new generations of consumers are more accustomed to light alcohol and will replace older generations who have stronger tastes for hard alcohol as seen in Figure 1.

The results of this paper have therefore important implications for public policies and economic research. For instance, they suggest that targeted interventions focused on young consumers might be more effective than broad-based policies such as excise taxes. Since many consequences of alcohol consumption such as traffic accidents or homicides generate externalities, government interventions in the alcohol market could improve welfare. Moreover, if consumers are not fully rational (for example because they are uninformed or have time-inconsistent preferences), then government intervention that address such "internalities" can further improve

[3]We leverage additional survey questions to document that binge drinking is more likely to occur when consuming a given amount of alcohol in the form of hard rather than light alcohol.

[4]Online Appendix B formally decomposes these consumption patterns into age (within) and cohort (between) effects and Section 6 estimates the effect of the beer market expansion on tastes.

welfare even for cases that do not affect third parties, such as alcohol poisoning (Gruber and Köszegi 2001).

Finally, our findings suggest that impact evaluations that focus only on the contemporaneous effects yield a severely biased estimate of the policy's full impact if the policy changes tastes. These long-run consequences can persist for decades after the intervention has ended and are often unintended when the policy is drafted.

The paper is organized as follows. Section 2 discusses the literature related to this paper. Section 3 describes the data. Section 4 estimates the causal effect of the anti-alcohol campaign on long-run alcohol tastes. Section 5 identifies the age at which tastes form. Section 6 studies the effects of the beer market expansion on taste formation. Section 7 estimates the causal effect of changes in alcohol tastes on mortality. Section 8 simulates two policy-relevant counterfactuals. Section 9 discusses alternative interpretations of our findings and Section 10 concludes.

2 Related Literature

In this section we briefly discuss related literature and how our research contributes to it. Our review is necessarily limited and does not include many important studies of taste changes and of the effects of alcohol consumption on health.

2.1 Taste Changes

Our paper is closely related to recent studies of long-run effects of temporary interventions.[5] Our study focuses on tastes that can change over an individual's lifetime including state-dependent preferences and social interactions (e.g., internal and external habits, peer effects, etc.).[6] One could also include culture and social norms if one was willing to assume that culture and norms can change quickly in response to a policy, and that these effects can be very local (e.g., only in rural areas). Most research on this topic however assumes that culture and norms change slowly. For instance, in their survey of this literature, Guiso, Sapienza, and Zingales (2006) state that to "claim a causal link from culture to economics, we restrict our attention to those cultural aspects that can largely be treated as invariant over an individual's lifetime."[7]

While different models of taste changes specify different primitives, they often have similar effects on outcomes policymakers care about, such as mortality. What matters for many policy-relevant questions is whether public policies can change tastes, and whether these changes then

[5] E.g. long-run effects on smoking cessation, Giné, Karlan, and Zinman (2010); on electricity consumption preferences, Costa and Gerard (2018); on physical exercise habits, Charness and Gneezy (2009); on commuting behavior, Larcom, Rauch, and Willems (2017); or on political participation, Fujiwara, Meng, and Vogl (2016).

[6] Note that if peer effects are uniform and spread across birth cohorts as much as within birth cohorts, the analysis of taste changes would not be able to identify any differences. Hence, social interactions such as peer effects outside one's birth cohort cannot be too strong.

[7] Recent research extends this analysis beyond tastes (e.g., endogenous risk preferences, Malmendier and Nagel 2011) and analyzes the transmission of preferences, with a focus on intergenerational transmission (Bisin and Verdier 2001, Doepke and Zilibotti 2017).

lead to persistent differences in behavior.[8]

We make three contributions to the growing empirical literature that studies taste changes. First, as discussed in the introduction, our approach uses variation from a public policy and from a large trade shock, while the previous literature mostly uses a migrants research design. Second, we estimate the typical ages at which such tastes form. Third, we show that the results from this literature apply much more broadly. For instance, the brand preferences identified by Bronnenberg et al. (2012) (e.g., Budweiser vs. Miller Light) extend to tastes over broader categories of goods, e.g., beer vs. vodka (or chicken vs. beef as in Online Appendix E). Moreover, we show that endogenizing these broader tastes has additional important consequences for individual welfare by affecting mortality and other health outcomes, in addition to the nutritional consequences highlighted by Atkin (2016). Hence, changing tastes are also relevant for other fields in economics such as health and public economics and are not limited to industrial organization, marketing, or trade.

2.2 Alcohol and Mortality

The second strand of related literature are studies of the health effects of alcohol consumption. Previous research has documented significant *contemporaneous* and medium-run effects of changes in alcohol supply on mortality rates in Russia. Brainerd and Cutler (2005) and Nemtsov (2011) provide comprehensive surveys of the evidence.

Yakovlev (2018) studies more recent changes in the Russian alcohol market and alcohol-related mortality. He analyzes the determinants of heavy drinking, focusing on peer effects and habit formation in addition to relative price effects. Our study differs in several important ways. First, we use a different research design and a different policy, focusing on the anti-alcohol campaign and the liberalization of the alcohol markets after the end of the Soviet Union. He uses a regression kink design resulting from the kink in the policy regime of the excise tax to identify price elasticities and he uses the clustered sampling structure of the survey data to identify peer effects. His analysis therefore focuses on different and more recent changes to Russia's alcohol markets. Second, he focuses on short-run habit formation while we document long-term effects on consumer tastes that last for decades. Third, our study documents persistent effects of policy changes not just on the level of total alcohol intake but also on the relative tastes for specific alcoholic beverages (e.g., beer vs. vodka). Our analysis therefore provides additional insights

[8]This point is related to the sufficient statistics approach to welfare analysis, which emphasizes the fact that several structural models can lead to the same set of reduced-form parameters necessary to conduct welfare analysis (Chetty 2009). In our context, several structural models of taste changes can lead to the same effect on observed consumption choices, which in our case explain differences in mortality over time and across regions. In Online Appendix F we provide an alternative model with multiple equilibria based on an extension of the Becker and Murphy (1988) model of rational habit formation, allowing for two habit-forming goods. In that model, persistent habits are formed when individuals start to consume a certain good regularly for the first time in their life. Individuals are born with the same initial tastes but are exposed to different initial market conditions and can therefore form long-run habits toward different goods. This model can rationalize our reduced-form results and generate multiple equilibria even without any heterogeneity in unobserved preferences.

for both academics and for policymakers. In particular, we show that even without a reduction in the level of alcohol consumption, changing the form in which alcohol is consumed can result in significant health benefits.

Bhattacharya, Gathmann, and Miller (2013) is another important recent analysis of the short- and medium-run effect of the anti-alcohol campaign on mortality rates during that period. The authors show that the anti-alcohol campaign significantly reduced contemporaneous alcohol-related deaths among working-age men between 1985 and 1990, either directly in the form of fatalities from alcohol poisoning and violent deaths or indirectly via heart attacks and strokes. They also document that mortality caught up after the campaign ended with above average mortality rates between 1991 and 1994 (i.e., "catch-up" mortality). Finally, they show that the anti-alcohol campaign had only small effects on deaths that are less related to alcohol consumption, such as respiratory and digestive diseases and cancer, consistent with a causal effect of alcohol consumption on mortality rates.

Our study extends this literature in three important ways. First, we document that these supply shocks also have large long-run effects by affecting tastes of young consumers that, to the best of our knowledge, have not been studied before. For instance, we find that individuals who became adolescent in a rural area during the anti-alcohol campaign formed tastes for hard alcoholic drinks, which increased their likelihood of dying due to binge drinking in later years relative to their urban counterparts. This negative long-run effect contrasts with the lower mortality during the campaign due to the successful short-run reduction in the total amount of alcohol consumed (see Panel B of Table A.5 in the Online Appendix).

Second, we show that the rapid expansion of the beer market in the second half of the 1990s had an even larger effect on male mortality than the anti-alcohol campaign. In contrast to the effect of the anti-alcohol campaign, this import shock had both a positive short-run and an unambiguously positive long-run effect on male life expectancy, both because it reduced contemporaneous instances of binge drinking and because it also changed younger consumers' tastes from hard to light alcoholic beverages and hence also reduced future instances of binge drinking.

Third, we show that the type of drink consumed has an important effect on mortality that so far has been overlooked by studies that focus only on the total amount of alcohol consumed. Consequently, policies and events that change *relative* tastes from hard to light alcohol substantially reduce the mortality of working-age men, even when holding fixed the total amount of alcohol consumed.

3 Data

We use data at the national, regional, and individual level and we provide a more detailed description in Online Appendix A. In this section, we instead focus on the most critical issues for our analysis.

We use *national-level* data of alcohol sales going back to the 1970s and estimates of home-made vodka (samogon) to show the aggregate effects of the supply shocks on the alcohol market.[9] We then use national-level male mortality rates to decompose the total effect of the supply shocks into (a) the effect of changes in the amount of alcohol consumed and (b) the effect of changes in the type of alcohol consumed, holding fixed the amount of alcohol consumed.

We use three sources of *regional-level* data. First, we use historical data provided by Bhattacharya et al. (2013) on the share of samogon in total alcohol consumed across Russian regions before and during the anti-alcohol campaign. We use these regional shares to show the differential effect of the policy on the supply of illegal vodka in rural and urban areas. Second, we use more recent regional-level data from the Russian Statistical Office (Rosstat) on alcohol sales by type of alcohol. Third, we use new epidemiological data from the Russian Fertility and Mortality Database (RusFMD) on mortality rates by year, gender, age, region and type of settlement (rural/urban), which allows us to estimate the differential impact of the anti-alcohol campaign on long-run male mortality rates and to identify the underlying mechanism that relates alcohol tastes to mortality. We explore both case-specific mortality data, which is available for 5-year age groups, and data on total mortality, which is available for one-year age groups.

Our main analysis of taste changes uses *individual-level* data from the Russian Longitudinal Monitoring Survey (RLMS), which is a nationally representative annual panel survey that covers more than 4,000 households per year corresponding to about 9,000 individual respondents. Our baseline sample consists of rounds 5 through 20 of the RLMS, spanning the period from 1994 to 2011, but not including 1997 and 1999 when the survey was not conducted.[10]

Our analysis leverages the rich individual consumption data for distinct types of alcoholic beverages. This data comes from the survey's health module and is completed separately by each adult. Hence, our data has the individual consumer as the unit of analysis compared to previous research, which is often limited to household-level expenditure data. Furthermore, the health module asks individuals about *quantities consumed* instead of expenditure outlays. Our consumption measures therefore capture individual consumption and are not subject to issues of timing and preference aggregation that may lead to a wedge between expenditures and consumption.[11] Since the health questions are confidential and asked of everyone separately without having other family members present, the answers are also less likely to be influenced by stigma.

Our primary measures of alcohol consumption are the shares of vodka and beer consumption in total alcohol intake, which is calculated in milliliters of pure alcohol. Specifically, we use the

[9]Note that we do *not* use these national-level data for our individual-level estimates of the effect of these shocks on taste formation. Hence, concerns about data quality of aggregate samogon consumption—which we discuss in Online Appendix A—do not affect our main results.

[10]We do not use data from rounds 1 to 4 because another institution conducted them using a different methodology. These rounds are of much lower quality according to the survey's website. We discuss the household-level expenditure data of non-alcoholic goods in more detail in Online Appendix E.

[11]In Online Appendix A we document that the quality of household-level alcohol expenditure data in the RLMS is lower than the individual-level data based on its health module.

individual's quantity consumed in a typical day during the last 30 days, which we express in grams of pure alcohol (ethanol).[12]

We restrict the sample to individuals age 18 and older, with 18 being the minimum legal drinking age in Russia, because one might be concerned with underreporting of underage drinking. Fortunately, we do not depend on survey responses from minors to identify taste formation because we study the long-run effects of policies that happened in the distant past. This contrasts with studies that estimate the contemporaneous impact of such policies.[13] Similarly, restricting the sample to consumers age 18 and above does not affect our estimates of the age at which consumers form alcohol tastes, because we measure consumption behavior decades after the shocks occurred using individuals who were adolescent during the policy intervention, but are now in their late 30s.

[Table 1 about here]

3.1 Descriptive Statistics

Table 1 summarizes various measures of alcohol consumption and individual characteristics. Our analysis focuses on men for two main reasons and we leave a detailed analysis of the effect on females for future research. First, the effect of alcohol consumption on mortality is much or pronounced for men than for women as mentioned in the introduction. Second, men mostly consume only two types of alcohol, vodka (62% of total alcohol) and beer (29%), which makes the substitution pattern much easier to study than for women, who also consume a significant share of wine (36%). Therefore, conditional on not becoming an abstainer, any behavioral response to a supply shock in the beer or vodka market implies a substitution to the other product for men.

To focus on the effect of *relative* tastes (i.e., taste for a specific type of alcoholic drink holding fixed the amount of alcohol consumed), we include the amount of alcohol consumed in most specifications and we analyze the effect of the shocks on the amount of alcohol consumed separately. We use the amount of all alcoholic beverages consumed during the previous month to construct these variables, using an alcohol content of 5% for beer and 40% for vodka based on data from the National Institutes of Health (NIH).

The average male alcohol consumer consumes almost four times more alcohol than the average female consumer does. This fact is crucial for understanding the large effect vodka consumption has on the gender gap in mortality in Russia, even after conditioning on the amount of alcohol consumed. The reason is that most alcohol-related deaths of individuals below the age of 65 are caused by occasional binge drinking, and stronger tastes for vodka makes binge

[12] The English translation of this survey question is "Now I'm going to list various alcoholic beverages, and you, tell me please, which of these you drank in the last 30 days and, for those you drank, how many grams you usually consumed in a day?"

[13] Online Appendix C shows that our findings indeed do not change if we include all individuals age 14 and above.

drinking much more likely. Hence, a higher share of vodka consumption increases mortality risk, even when comparing two individuals with the same average alcohol intake per month. While we do not observe binge drinking directly, we use additional questions on consumption frequency to relate the share of vodka consumed to binge drinking. Online Appendix Table A.1 shows that individuals with a higher share of vodka consume a given amount of alcohol over fewer days per month. Hence, their propensity to binge drink is higher than if they consumed this amount of alcohol in the form of beer.

Table 1 also provides summary statistics for the other covariates used in the analysis, both for the main samples of male alcohol consumers age 18 and above and for the sample of all men above age 18, including those who report not having consumed any alcohol during the previous month. We use the latter sample when we analyze the total effect of the shocks on changes in alcohol tastes and on mortality, including the extensive margin (i.e., abstaining).

3.2 Data Quality

Survey data have well-known measurement issues that can potentially bias the estimates, and the RLMS is no exception. We therefore provide a detailed analysis of these issues in Online Appendix A, including the effect of attrition on our results and a comparison with registered alcohol retail sales. We summarize the main results of these robustness checks where appropriate. For example, regarding sample attrition, the survey's website states that "the main effects [of attrition] are in the Moscow/St. Petersburg sample," while interview completion rates are above 88% in all other sampling units. We show that excluding individuals from Moscow and St. Petersburg in fact slightly strengthens the causal effects in the main analysis in Section 4, consistent with the hypothesis that data from these subsamples contain more measurement error than responses from other sampling units.

We also deal with issues related to goods that are potentially addictive. For example, our results could be sensitive to the behavior of a few individuals because alcohol consumption is known to be highly skewed to the right (Cook and Moore, 2000). To address this concern, we control for the amount of alcohol intake in most specifications and we follow the recent empirical literature by using consumption shares instead of levels to make the results robust to outliers.[14] Moreover, our findings are robust to dropping the top quartile of alcohol consumers.

Finally, we use the survey data to provide direct evidence of the link between alcohol consumption shares and mortality at the micro-level. While national-level time series data and regional-level panel data of mortality rates and alcohol sales by type of alcohol strongly suggest such a link, only the RLMS can provide direct evidence since it is the only dataset that simultaneously records alcohol consumption and mortality for the same individual. Previous research has used the RLMS to study mortality trends and we discuss the quality of the mortality data and its limitations more extensively in Online Appendix A. Brainerd and Cutler (2005) for

[14] Online Appendix C uses the estimator suggested by Honoré (1992) for two-sided truncated panel models with fixed effects to show that our results are not affected by this transformation.

instance use the RLMS to study mortality trends in post-Soviet Russia and describe its quality 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). [...] 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 percent 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.

4 Identifying Taste Changes

We use Russia's anti-alcohol campaign in the late 1980s that briefly but severely rationed distinct types of alcohol to estimate its persistent causal effect on tastes. In this section, we focus on individuals who were exposed to the policy during early adulthood and study their tastes for different types of alcoholic drinks later in life when most of them are in their late 30s. In the next section, we use this variation to estimate the age at which alcohol tastes typically form.

[Figure 2 about here]

4.1 Research Design and Identification

We begin by providing institutional background of this policy. In 1985, Mikhail Gorbachev introduced an anti-alcohol campaign that was designed to fight widespread alcoholism in the Soviet Union. The impact of the campaign on the alcohol market is shown in Figure 2. Regulated prices of vodka, beer and other types of alcohol were raised, sales were heavily restricted, and many additional regulations were put in place aimed at further curbing alcohol consumption.[15] The effect of the campaign on official alcohol sales (Panel a) was dramatic because the communist government directly controlled the production of any official alcohol. Beer sales dropped by 29% from 177 million liters of ethanol in 1984 to 125 million liters in 1987, and vodka sales dropped by 60% from 784 to 317 million liters. Although the campaign officially ended in 1988, Figure 2 shows that its impact lasted until 1990.

While the effect on official alcohol sales was dramatic, Panel (b) shows that the drop in vodka sales was partially offset by a substantial increase in the consumption of samogon, a low-quality

[15] The measures included, among other things, limiting the kinds of shops that were permitted to sell alcohol, closing vodka distilleries and destroying vineyards in the wine-producing republics, and banning the sale of alcohol in restaurants before 2p.m. White (1996) and Nemtsov (2011) provide a detailed account of this policy.

home-made vodka whose consumption remained illegal until 1997.[16] Home-made beer on the other hand was extremely rare at that time and remains rare today (Nemtsov, 2011). Table 1 for instance shows that the share of home-made beer in total alcohol is less than 0.1%. In fact, until 2008 the survey did not even have a question about home-made beer.

Important for our identification approach is the fact that the production of samogon was heavily concentrated in rural areas, as we show next. This happened mostly for technical reasons that are unrelated to changes in tastes. The production of samogon requires space, which is limited in urban areas, especially in Russian cities, which are very densely populated by international comparison, with most people living in large apartment buildings. Moreover, producing samogon causes smoke and a strong smell, which is at the same time very unpleasant and easy to detect by neighbors and law-enforcement, particularly in cities. Hence, the ban on the illegal production of samogon was more strictly enforced in urban areas. It was therefore much safer to produce samogon in single-unit homes, which are highly concentrated in rural areas. As a result, samogon was more readily available in rural areas than in urban areas.[17]

Next, we quantify the differential access of rural consumers to samogon before and during the campaign. We use regional-level data on shares of samogon in total alcohol consumption from 1980 to 1992 provided by Bhattacharya et al. (2013), since micro-level survey data is not available before 1994. We estimate a difference-in-differences specification by regressing region r's share of samogon in year t, $S_{rt}^{samogon}$, on the region's fraction of rural population in 1991, fully interacted with year dummies:

$$S_{rt}^{samogon} = \alpha_r + \sum_{t=1981}^{1992} \Big[\delta_{D,t} \cdot \text{I(year)}_t + \delta_{DD,t} \cdot \text{I(year)}_t \times \text{Rural Fraction}_r\Big] + \varepsilon_{rt}. \tag{1}$$

1981 is the first year in which reliable regional-level population data is available, and α_r is a full set of region fixed effects. Observations are weighted by the region's total population. Panel (c) plots the difference-in-differences coefficients, $\delta_{DD,t}$, documenting the differential impact of the campaign on the consumption of illicit hard alcohol in rural areas. While the rural share follows a parallel trend in the years leading up to the campaign, there is a jump in the share of samogon consumed in rural areas during the campaign, from 1986 to 1990.

4.2 Long-run Effects of the Anti-Alcohol Campaign on Tastes

If tastes can change and form during adolescence when an individual first consumes alcohol regularly, we would expect the campaign to affect tastes of young rural consumers differently

[16] Selling samogon remains illegal today even as the consumption of samogon has been legalized. We discuss the sources and methodologies used to construct these figures and their limitations in more detail in Online Appendix A.

[17] These alcohol restrictions also led to other kinds of substitution such as consuming industrial alcohol or even perfume (e.g., White 1996). However, the consumption of these goods was not as widespread as samogon (see Nemtsov, 2011, p.136). We therefore have no reason to believe that rural and urban consumers had different access to these substitutes.

both relative to rural consumers that were not adolescent during the campaign, but also relative to their urban counterparts. We implement a difference-in-differences approach using individual-level alcohol consumption data from the RLMS recorded decades after the end of the campaign to test this prediction. The treatment group is rural consumers that became adolescent during the campaign. We estimate the following regression:

$$S_{it}^{vodka} = \beta_{DD} \cdot \text{I(became adolescent during campaign)}_i \times \text{I(rural)}_i \\ + \beta_D \cdot \text{I(became adolescent during campaign)}_i + \lambda \cdot \text{I(rural)}_i + \gamma' x_{it} + \epsilon_{it}, \tag{2}$$

where S_{it}^{vodka} is individual i's share of vodka consumed in survey year t and x_{it} is a vector of controls. Anticipating the findings in the next section, we define adolescence as being 17 years old and define the campaign's impact to last from 1986 to 1990 based on Panel (c) of Figure 2. To focus on the long-run effects of the campaign we initially restrict our sample to survey rounds starting in 2001 in order to have at least one decade between the end of the campaign's impact on the alcohol market and the point at which we measure consumer tastes. We will later relax these assumptions, and in the next section, we estimate the age-profile of taste formation non-parametrically.

[Table 2 about here]

Table 2 summarizes the result of this analysis. Column 1 starts with a minimal specification including only region, age and survey year fixed effects, which flexibly control for life-cycle patterns, macroeconomic shocks and local differences. Columns 2 to 4 then gradually add more controls. The full set of age and year effects is identified because the policy affected rural and urban cohorts differentially.

Consistent with tastes forming during early adulthood, the intent-to-treat effect β_{DD} shows that individuals who became adolescent in a rural area during the campaign prefer vodka 5 percentage points (pp) more than their urban counterparts based on consumption choices recorded more than a decade after the end of the anti-alcohol campaign. Hence, tastes of young consumers can be persistently manipulated even by a temporary public policy. Furthermore, these treated individuals also use a 4 pp higher share of vodka relative to other rural consumers who became adolescent before or after the campaign (the sum of the difference and difference-in-differences coefficients, $\beta_D + \beta_{DD}$). This long-run increase is also economically significant: It corresponds to a 10% difference relative to the sample share of vodka of 50%. In comparison, urban consumers that turned 17 during the campaign have a 1 pp lower vodka share relative to younger and older urban consumers (the β_D coefficient), although this difference is not statistically significant.

Controlling for the level of total alcohol consumption in Column 2 shows that the campaign changed *relative* tastes, holding fixed the campaign's effect on the amount of alcohol consumed in the long run, which we show in Column 5. Column 3 adds changes in real income and in local relative prices to control for *contemporaneous* substitution patterns and differences in income elasticities. Column 4 adds a standard set of demographics such as personal health status, weight, education, and marital status. The point estimates are stable and become slightly more precise.

Next, we decompose the causal long-run effect into extensive margin (abstaining) and two components of the intensive margin: the amount of alcohol consumed and the substitution between different types of alcohol.

Column 5 shows that the anti-alcohol campaign significantly lowered total alcohol consumption in the long-run among urban consumers that became adolescent during the campaign (by 8.4 pp), but had a much smaller impact on rural adolescents (0.8 pp) presumably because of increased availability of samogon. The long-term effect of the anti-alcohol campaign on abstaining in Column 6 is imprecisely estimated and not statistically significant. Men that were adolescents during the campaign are about 5% more likely to abstain from alcohol both in rural and urban areas (from 28 pp to 28.4 pp).

Column 7 shows that the campaign has the opposite effect on the share of beer consumed, confirming that for male consumers the main substitution occurs between vodka and beer. This result is important to keep in mind when interpreting the long-run effects of this policy on male mortality in Section 7.

Finally, Column 8 assesses the campaign's total effect on tastes for hard alcoholic drinks, including cognac, fortified wine, and samogon. Although the effect is qualitatively similar to the main results suggesting that the campaign affected relative tastes for hard and light alcoholic beverages, the point estimate is smaller and less precise as one would expect when adding measurement error in reported samogon consumption (see Online Appendix A). Online Appendix C provides extensive robustness checks of these findings.

In summary, this section shows that the campaign significantly changed the consumption behavior in the long run given that most subjects in our sample are observed more than two decades after the end of the campaign. Moreover, the results highlight the differential impact the campaign had on consumers who were adolescent in rural areas compared to their urban counterparts. These consumers formed persistent tastes for distinct types of alcoholic beverages, and these taste differences are easily detectable in their current consumption behavior.

5 At What Age Do Alcohol Tastes Form?

In this section, we exploit the temporary nature of the anti-alcohol campaign to identify the age at which the average consumer forms tastes for different alcoholic drinks. We start by noting that the estimates in the previous section are intent-to-treat effects for three reasons: First, the legal drinking age is not strictly enforced; second, individuals above age 18 cannot be forced to drink alcohol; and third, tastes do not necessarily form within a single year but might take several years. Hence, to identify the ages at which tastes form we need to estimate how the campaign's long-run effect depends on age, still exploiting the differential impact of the campaign on rural and urban consumers.

[Figure 3 about here]

Figure 3 summarizes our analysis. We follow an approach that is related to the non-parametric estimation of an unknown density function. To obtain a "smooth" estimate of this unknown age function, we use a 5-year triangular kernel shown in Panel (b) that reflects the intensity of the campaign during the 5-year period from 1986-90 shown in Figure 2.[18] For a given age between 10 and 35 we then calculate each consumer's exposure to the campaign (i.e., treatment intensity) using the kernel weights and under the assumption that tastes form at that age. For example, suppose that we want to estimate the response to the campaign at age 17. Any person born before 1969 and hence turning 17 before 1986 and any person born after 1973 who turned 17 after 1990 receives a weight of zero. People born in 1969 or 1973 receive a weight of 1/9, people born in 1970 or 1972 a weight of 2/9, and people born in 1973 a weight of 1/3. We then estimate equation (2) using this measure of campaign exposure instead of the 5-year uniform indicator, I(became adolescent during campaign)$_i$.

Panel (a) shows the estimated age function. We see that tastes for vodka form in a short window between ages 14 and 18, with a peak at age 17 and statistically significant differences also at ages 15 and 16.

6 The Beer Market Expansion of the 1990s

The anti-alcohol campaign allows us to cleanly identify the causal long-term effect of temporary supply shocks on tastes because it affected urban and rural consumers differently and because the policy was short-lived. The temporary nature of the policy also helps us identifying the age at which consumers form tastes over distinct types of alcoholic drinks.

In this section, we test the external validity of our findings using large import shocks that occurred in the late 1990s. Many goods that were not readily available during the Soviet Union, including high-quality beer, became suddenly accessible to the broader public after Russia opened its borders to trade and foreign investments.

6.1 Background Information

Vodka dominated the alcohol market during the Soviet Union. Starting in 1995, however, the beer industry expanded rapidly for reasons that are largely exogenous to taste changes, such as market liberalization, a lower regulatory burden for the beer industry—compared to all other alcohol producers—and the entry of foreign competition and investments into this new market. Foreign competition also brought modern technologies. For example, beer sold in cans or in plastic bottles started to be produced only after the fall of the Soviet Union. Brewing

[18]Online Appendix C shows that these findings are robust to choosing an empirical or a uniform kernel instead.

technologies also changed significantly, and the assortment of beer increased dramatically.[19] Panel (b) of Figure 4 shows that beer sales increased by a factor of four between 1995 to 2011, from 2.8 to 10.8 billion liters. In contrast, vodka sales did not change much (see Figure 5). Total annual sales of vodka were 1.59 billion liters in 2011, which is roughly the same level as during the Soviet era.[20]

[Figure 4 about here]

We use this beer market expansion for two main purposes. First, we document a similar effect of this shock on long-run consumption tastes as in Section 4, although this time we study a *positive* supply shock which affected *light* alcohol. Of course, identification is much more challenging than with the anti-alcohol campaign, because many other things might have changed during this period and these other factors are difficult to fully control for, including social norms and culture.

Second, as shown in Figure 4, this positive supply shock is even larger than the negative supply shock of the anti-alcohol campaign. Section 7 below will show that changes in relative alcohol tastes play a key role in explaining changes in mortality, especially the recent decline in male mortality that started around 2003. Before turning to the effects of changes in alcohol tastes on mortality, this section shows that the beer market expansion in the 1990s caused younger cohorts to prefer beer to vodka, leading to fewer cases of binge drinking and hence fewer alcohol-related deaths today. This channel then plays a key role in the large decline of male mortality in recent years. Therefore, it is important to study this large positive supply shock even if identification is more challenging.

6.2 The Effect of the Beer Market Expansion on Long-Run Tastes

We study the long-run effects of the beer market expansion on relative alcohol tastes of young consumers by focusing on the relatively brief period when the beer industry experienced the most rapid growth. We implement two empirical strategies to identify the causal effect of the import shock on consumer tastes, shown in Panels (b) and (d) of Figure 4.

First, we estimate the differential impact of the beer market expansion on long-run alcohol tastes by comparing the consumption patterns of individuals who turned 17 in different years during the expansion. These consumers had different access to beer when they formed their tastes. We estimate the effect of the import shock on alcohol tastes by running the following

[19] For instance, in 1991 there were no foreign-owned beer breweries in Russia and no foreign brand was sold. By 2009, the five leading foreign-owned companies produced more than 85% of the total beer sold. Similarly, the number of beer brands increased from only 20 in 1991 to over 1,000 in 2009. The set of varieties available in 1991 was even more limited than this number suggests, since one brand—Zhigulevskoe—dominated the entire market; see www.beerunion.ru/soc_otchet/2.html.

[20] We measure sales in terms of quantities instead of values because there were no formal market prices in the Soviet Union. Instead, the alcohol industry was monopolized by the state, and quantities produced were heavily regulated. As a result, it was difficult or even impossible to find many goods in stores, and prices were often not the most significant factor as there was severe rationing.

regression:

$$S_{it}^{beer} = \phi \cdot \text{year-turned-17}_i + \gamma' x_{it} + \epsilon_{it}. \tag{3}$$

Panel (b) illustrates the research design. We start by estimating equation (3) on the sample of all men who turned 17 during the beer market expansion from 1995 to 2007. Since it is possible that other factors also changed during this period, which might have affected men differently depending on the year of their 17$^{\text{th}}$ birthday, we then let the sample window shrink on both sides until it only includes the three years from 2000 to 2002. Hence, as we shrink the sample window we identify the effect of the shock on alcohol tastes using consumers who grew up in a more and more similar environment (including similar culture, norms, and information), except that they face a different beer market when they turn 17. This is similar to shrinking the bandwidth parameter in a regression discontinuity design.

Panel (a) plots the estimates of ϕ for both beer and for vodka together with 95% confidence intervals. The effect of the beer market expansion on the shares consumed is remarkably stable, and it remains statistically significant despite the substantial gradual reduction in the sample size. The point estimates are also economically significant, implying that the average consumer who turned 17 during the expansion exhibits about a 4% higher long-run share of beer consumption compared with consumers who are only one year older. We use the term "long-run share" because we are estimating the individuals' consumption shares using data from 2001 to 2011. Hence, most of the individuals in our sample are substantially older than 17 when we measure their consumption choices.

Consistent with the results from the anti-alcohol campaign, we see the opposite effect on vodka, again confirming that for men, increases in the beer share mainly come at the expense of the vodka share. Hence, these results paint the same picture as the anti-alcohol campaign, providing external validity of these findings.

Second, we run a set of placebo tests shown in Panels (c) and (d).[21] We estimate equation (3) using a 5-year rolling window starting with men who turned 17 between 1970 and 1974 and ending with men who turned 17 between 2006 and 2010.[22] We should not see any significant effect of the year in which an individual turned 17 on the share of beer consumed for samples that do not include the expansion of the beer market *if* the results from the anti-alcohol campaign extend to this setting. As the 5-year sample window reaches the time at which the beer market expands rapidly, the estimate of ϕ in equation (3) should gradually increase, because men turning 17 at the end of the 5-year sample window have much easier access to beer than men who turned 17 at the beginning of the sample window. Finally, the beer market stabilizes around 2007 at a new long-run equilibrium shown in Figure 4. As the sample window starts to cover more and more of this new steady state, the regression coefficient should gradually decrease. Hence, the response should first be zero and then exhibit a hump-shaped pattern with a peak response

[21] Online Appendix C shows similar placebo tests for the anti-alcohol campaign.

[22] 1970 is the first year for which we observe aggregate sales by type of alcohol.

when the sample window fully covers the beer-market expansion period.

Panel (c) plots the estimates of ϕ from this research design, together with 95% confidence intervals. We indeed see this hump-shaped pattern emerge precisely as we would expect if tastes form in early adulthood. The coefficients are close to zero and not statistically significant for samples that only include consumers who turned 17 before the expansion of the beer market. The effect gradually increases when the import shock affects more and more individuals in the 5-year rolling window. The response reaches a peak for the window that ranges from 1997 to 2001, which coincides with the period that saw the most dramatic increase in the beer market over the entire 42-year period.

Age of Taste Formation Finally, we use the beer market expansion as an alternative source of variation to estimate the age profile of taste formation. We use a similar approach as in Section 5, using the empirical kernel shown in Panel (f) of Figure 4, which is derived in a straightforward way from the time-series of the market share of beer shown in Panel (d). We then apply this kernel to calculate each individual's exposure to the beer market expansion under the assumption that tastes form at a specific age between 10 and 35. We use this measure of exposure to replace the linear trend in equation (3); i.e., the kernel is a non-linear transformation of the time trend. Panel (e) shows the corresponding estimates $\hat{\phi}$ as a function of the potential age of taste formation. Consistent with the results from the anti-alcohol policy, we find that alcohol tastes form in early adulthood. The age function has a slightly less pronounced shape than in Figure 3 because the import shock is persistent and occurs over a span of 13 years and because we do not have a similar control group as in the difference-in-differences research design based on the anti-alcohol campaign.

Online Appendix B uses the panel dimension of the RLMS to decompose the changes in alcohol shares non-parametrically into age, cohort and time effects. We then test for steppingstone effects of light alcohol, i.e., the hypothesis that consuming light alcohol early in life is a steppingstone or gateway to consuming harder alcohol later in life. We find only modest support for the steppingstone hypothesis. While the share of beer increases from ages 18 to 22 (and decreases for vodka), it is completely flat thereafter. This is consistent with the estimated age profile shown in Figure 4 despite using a completely different source of variation: consumption changes within a person.

6.3 Extensions

Finally, we briefly mention two extensions of our analysis, which are discussed in more detail in Online Appendix D and E.

Alternative Identification using Migrants In Online Appendix D we use a completely different research design based on migrants. This is the main approach used in the previous literature on taste formation (e.g., Bronnenberg et al. 2012 or Atkin 2016). Consistent with this literature, we find that migrants bring their tastes with them.

Tastes Formation for Non-Alcoholic Goods Alcohol is of course special in many ways, including that it is potentially addictive. In Online Appendix E we use similar market expansions of non-alcoholic consumer goods during the 1990s to directly address this concern, in particular the question whether and to what extent our results on taste formation can be applied more broadly to other non-addictive goods. We find that the aggregate market share of these goods, measured at the time when consumers were young, is also a strong predictor of their relative tastes at a much later age, and holds even after controlling for relative prices, income, and individual characteristics. These results are therefore consistent with the causal estimates of long-run tastes from the anti-alcohol campaign and the beer market expansion.

7 Alcohol Tastes and Mortality

This section provides evidence that these changes in alcohol tastes have important consequences for one of the most pressing public policy concerns in Russia: the high mortality of working-age adults and in particular the large gap between male and female life expectancy. We build on previous research that establishes the role of alcohol in explaining the large changes in male mortality since the mid-1980s (Brainerd and Cutler 2005, Nemtsov 2011, Yakovlev 2018).

We extend this research by showing that the *type* of alcohol consumed has an important effect on mortality of working-age adults *in addition* to the amount of alcohol consumed, that is, even when we hold fixed the amount of alcohol consumed. We establish this link using the steps described in the introduction.

7.1 Background Information

Alcohol consumption has well-known long-term adverse effects on life expectancy and health outcomes (e.g., cirrhosis). Probably less well known is the fact that approximately 40% of all annual deaths among working-age adults in Russia are estimated to be related to alcohol consumption. Most of them are not due to long-run consequences of heavy drinking but due to the fact that alcohol is often consumed in large amounts over a short period of time, i.e., due to binge drinking. While Russia certainly has one of the highest levels of alcohol consumption per capita, other countries with elevated levels of alcohol consumption have a much lower number of alcohol-related deaths per capita, including many western European countries. This is because consumers in those countries tend to spread their alcohol intake more evenly over the year (Rehm and Shield 2013). The high level of alcohol consumption among Russian men is therefore widely believed to be a main contributing factor to the low male life expectancy and the large gender gap (Brainerd and Cutler 2005). Hence, while a high average level of alcohol intake can certainly be hazardous—especially for older individuals—it is mostly the occasional binge drinking that leads to high mortality rates across all age groups, and in particular among working-age adults.

Furthermore, since binge drinking is much less likely to occur when consuming beer rather than vodka, a natural hypothesis is that individuals who prefer beer to vodka have a lower alcohol-related probability of dying, even when holding fixed the amount of alcohol consumed.

[Figure 5 about here]

7.2 National Alcohol Sales and Mortality Rates

In the first step, we use data on aggregate sales by type of alcohol from 1970 to 2013 and calculate mortality rates of working-age males using data from the Human Mortality Database. Figure 5 shows the enormous changes in Russian male mortality over the past four decades.[23] For comparison, we also graph the evolution of the corresponding male mortality rate for the U.S. population, which declines much more gradually. For instance, the standard deviation of the Russian mortality rate is more than twice that of the U.S.

Figure 5 also shows that changes in mortality are closely associated with changes in alcohol sales per capita and in particular with vodka sales. Beer sales—which are also expressed in liters of pure alcohol and are therefore comparable with vodka sales—are much less related to male mortality, consistent with the hypothesis that the effect of alcohol on mortality of working-age adults mainly operates through binge drinking.

[Table 3 about here]

In Panel A of Table 3, we regress male mortality on the amount and the relative shares of aggregate alcohol consumed to quantify the relative importance of these two channels. To make the estimates comparable across specifications that cover periods with different baseline mortality rates, we report elasticities by taking the log of mortality (or equivalently the log of the number of deaths while controlling for log population).

Column 1 shows that the share of vodka consumption is strongly associated with male mortality. Column 2 decomposes this effect into the contribution of the amount of alcohol consumed and the contribution of the share of vodka consumed, holding fixed the amount consumed.[24] Consistent with previous medical research that documents that alcohol-related deaths are a major cause of the low life expectancy of Russian men, we find that total alcohol is an important predictor of Russian male mortality. However, Column 2 shows that controlling for total alcohol does not change the effect of the share of vodka on mortality, and both have similar explanatory power. A one standard deviation increase in the aggregate share of vodka (6.9 pp) increases male mortality by 10% while a one standard deviation increase in the log of total alcohol (7.3 pp) increase it by 10.4%.

[23] We calculate standardized mortality rates (SMR) relative to the mid-year population using the U.S. standard population from 2000 provided by the NIH to avoid biases over time due to demographic changes.

[24] The share uses official registered sales of vodka. Total alcohol also includes sales of beer, wine, cognac, champagne and estimates of illicit samogon production and tax-evaded vodka. See Online Appendix A for a description of the data sources.

7.3 The Causal Long-Run Effect of the Campaign on Mortality

In Panel B of Table 3, we use regional-level data on alcohol sales by type and on cause-specific deaths to document that the channel works mainly through binge drinking. To establish the causal effect of the share of vodka on male mortality, we use the anti-alcohol campaign as an instrument. These long-run effects of the campaign have not been studied before even though the campaign's contemporaneous and medium-run effects on mortality have been well documented.

Mortality rates are available separately by gender, year, age, region, cause of death, and crucially also by type of settlement, thereby allowing us to identify rural and urban areas and apply our difference-in-differences research design from Section 4 to mortality related to binge drinking. Our research design requires age-specific alcohol consumption, which is only available from the RLMS. We therefore calculate average age-specific vodka shares and total alcohol consumption across 31 regions identified in the RLMS data covering years 1994-2011. We then merge this data with regional epidemiological data provided by the Center for Demographic Research at the New Economic School in Moscow. The epidemiological data provides causes of death for 5-year age groups and we compute corresponding alcohol consumption shares for those 5-year age groups by region and type of location (rural or urban). While this approach splits the data into many cells , the median and average number of people in each cell (i.e., 5-year age bin by region, location type and year) is 15, thereby providing a reasonably precise estimate of the share of alcohol consumption.

To isolate the causal long-run effect of changes in alcohol tastes on mortality, we use the variation generated by the anti-alcohol campaign to instrument for the current share of vodka consumption. Our instrumental variable for the regional share of vodka consumption is the fraction of people that turned 17 in a rural part of the region during the anti-alcohol campaign (calculated separately for each 5-year age bin). Specifically, the 1st-stage regression is

$$S^{vodka}_{a,r,l,t} = \gamma \cdot Pr(\text{became adolescent during campaign})_{a,t} \times I(\text{rural})_{a,r,l,t} + \nu' X_{a,r,l,t} + u_{a,r,l,t}. \quad (4)$$

$S^{vodka}_{a,r,l,t}$ is the average share of vodka consumption within the 5-year age group a (ages 20-24, 25-29, etc.) in region r and location type l (rural or urban) and year t. The variable $Pr(\text{became adolescent during campaign})_{a,t}$ is the fraction of males in 5-year age group a who turned 17 during the anti-alcohol campaign from 1986-90. For example, focusing on the age group of 30 to 34 year olds, in year 1999 20% of that group turned 17 during the anti-alcohol campaign (those age 30 who were born in 1970), and this fraction is 40% in 2000 (those age 30 or 31 who were born in 1970 or 1971), 60% in 2001, 80% in 2002, 100% in 2003, and 0% before 1999 and after 2003. The 2nd-stage regression is

$$\ln(\text{mortality})_{a,r,l,t} = \beta \hat{S}^{vodka}_{a,r,l,t} + \eta' X_{a,r,l,t} + \varepsilon_{a,r,l,t} \ . \quad (5)$$

$\hat{S}^{vodka}_{a,r,l,t}$ is the predicted share of vodka consumption from (4). The set of exogenous control variables $X_{a,r,l,t}$, which is common across the two stages, contains an indicator for rural localities; $Pr(\text{became adolescent during campaign})_{a,t}$; age, year, and region fixed effects; the log of average income; the log population of the region; and the average alcohol intake. Hence, we use the exogenous variation generated by the anti-alcohol campaign as an instrument for $S^{vodka}_{a,r,l,t}$ in (5), which is the interaction term $Pr(\text{became adolescent during campaign})_{a,t} \times I(\text{rural})_{a,r,l,t}$.

Columns 3 and 4 show the causal effect of changes in the regional share of vodka consumption on total mortality of working-age males in those regions and Column 8 shows the corresponding 1st-stage regression. Despite using only cross-sectional variation to identify the effect, we find similar magnitudes as in Panel A.

Columns 5-7 use additional age-specific information about the cause of death to identify the mechanism. Column 5 shows that an increase in the share of vodka significantly increases mortality due to alcohol poisoning among working-age men. Accidental poisoning constitutes about 9% of all deaths among working-age men. Not surprisingly, the elasticity in Column 5 is about three times larger than the one for all causes combined shown in Column 4.

Since many consequences of alcohol consumption such as traffic accidents or homicides generate externalities, government interventions in the alcohol market could improve welfare. Column 6 provides evidence of such indirect negative effects of binge drinking on third parties by studying the effect of changes in the share of vodka consumption on causes of death that are believed to be at least indirectly related to alcohol consumption. These include traffic accidents, homicides, suicides, deaths by accident caused by fire and electric current, accidental falls and injuries, and deaths caused by accidentally discharged firearms. They constitute about 30% of all deaths among working-age men. Column 6 shows that increases in the share of vodka consumption significantly increases these negative externalities.

Unfortunately, the estimated magnitude in Column 6 is difficult to interpret and most likely underestimates the size of the negative externalities caused by an increase in binge drinking. One reason is that about a quarter of all causes of death in the data are unspecified or ill-defined and there is recent evidence suggesting that these cases are more likely to be alcohol related, either directly or indirectly (e.g., Gavrilova, Semyonova, Dubrovina, Evdokushkina, Ivanova, and Gavrilov (2008), Ivanova, Sabgayda, Semenova, Zaporozhchenko, Zemlyanova, and Nikitina (2013), Danilova, Shkolnikov, Jdanov, Meslé, and Vallin (2016)). Moreover, the estimate only captures the effect on working-age males of events that resulted in fatalities. However, we would expect externalities such as alcohol-related traffic accidents to apply to the general population, including women and children. The estimate also does not include negative externalities resulting from events that are less severe, such as non-fatal injuries.

Moreover, if consumers are not fully rational (for example because they are uninformed or have time-inconsistent preferences), then government intervention that address such "internalities" can further improve welfare (Gruber and Köszegi 2001), even for cases that do not affect third parties, such as alcohol poisoning.

Finally, Column 7 provides a placebo test, showing that the instrumented share of vodka does not affect causes of death that we would ex-ante expect to be unrelated to short-term binge drinking, such as cancer unrelated to alcohol (i.e., excluding liver cancer).

Robustness The p-value of the coefficient on the instrument for the share of vodka in the 1st-stage regression in Column 8 is 0.001. However, the 1st-stage F-statistic of 10.2 raises questions about weak instruments. We deal with this concern it two ways. First, Table 5.2 in Stock and Yogo (2005) shows that if the observed p-value in the 2nd stage was 5%, a conservative upper bound of the unbiased size of the test would be 15% based on a 1st-stage F-statistic of 11. While the authors do not provide a similar table for a 2nd-stage test size of 1%, it seems reasonable to assume that 10% is a conservative upper bound for the size of the hypothesis tests in Columns 3-7.

Second, Panel A of Online Appendix Table A.5 uses mortality data from a different source, which has one-year age bins instead of the 5-year age bins but does not contain information about the cause of death. This allows us to repeat the analysis reported in Columns 3-4, but with a 1st-stage F-statistic of 16. We obtain quantitatively similar results and we cannot reject the hypothesis that the two coefficients are the same.

Online Appendix Table A.5, Panel B also provides the reduced-form estimates. Because the reduced form does not rely on survey data from the RLMS, it covers a longer period from 1989-2014. It shows that while the campaign has an unambiguously positive long-run health effect on urban consumers that turned 17 during the campaign (lowering their mortality rate relative to the comparison group by 7%), the effect is ambiguous for rural consumers and depends on the horizon at which we evaluate the policy. This in turn is the result of endogenous attrition of consumers, who form particularly strong tastes for hard alcohol, which is much more prevalent in rural areas.

Finally, we address 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 shown in Column 5, because vodka and samogon consumption are likely positively correlated. Using a separate instrument for endogenous samogon consumption, Online Appendix H shows that controlling for the effect of samogon consumption on mortality from alcohol poisoning does not decrease the effect of vodka consumption.

7.4 Individual-Level Analysis

In the last step, we use individual-level data. The RLMS is the only data set that has individual-level consumption data and thanks to its long-run panel dimension also records death events for members of multi-person households (see Section 3). We use the survey data to provide additional direct evidence of the mechanism linking alcohol consumption to mortality. The large effects of alcohol on male mortality makes it possible to study this mechanism despite observing relatively few death events in the RLMS sample (360 among working-age men and 600 in total).

We exclude individuals below age 22 since Sections 5 and 6 show that alcohol tastes of men below age 22 have not yet converged to their long-run equilibrium such that their observed consumption shares are not a good predictor of their future shares. This is important since the counterfactual simulation of the long-run effect of these changes in alcohol tastes on mortality in the next section crucially depends on the behavior of these young cohorts as they approximate the consumption behavior of the population in the new long-run steady state.

[Table 4 about here.]

Panel A of Table 4 summarizes the main results of the micro-level analysis. We estimate a standard semi-parametric Cox proportional hazard model to quantify the effect of alcohol tastes for distinct types of alcoholic drinks on the probability of dying, using the sample of male consumers.[25] We use a similar specification as in our analysis of the effect of changes in alcohol supply on long-run tastes in Sections 4, 5 and 6, with two modifications. First, we control for three additional variables: The first indicates whether the individual reports not drinking in a typical day during the previous month, the second is an indicator of whether the individual smokes, and the third is an indicator of whether the individual is a heavy drinker.[26] Second, we collapse the data to one observation per individual and we replace time-varying covariates with their mean. For individuals who report not consuming alcohol in an interview, we set their shares of beer and vodka to zero before collapsing the data.

Consistent with the aggregate data in Table 3, Columns 1 and 2 show a strong relationship between mortality and both the amount of alcohol and the share for vodka consumed. Column 2 decomposes the total effect into the two channels, i) the amount of alcohol and ii) the type of alcohol consumed, holding fixed the amount consumed. Like in the analysis with aggregate data, we find that the type of alcohol consumed has a large effect on mortality rates of working-age men in addition to effect of the total amount of alcohol consumed. A reduction in the share of vodka by one cross-sectional standard deviation (0.32) decreases the mortality hazard by 20%, while a one standard deviation reduction in the amount of alcohol consumed (0.09 kg) decreases mortality by 10%.

Column 3 shows that these results are robust to controlling for whether an individual is a heavy drinker, which is the top quartile of the alcohol consumption distribution. Because a high share of vodka at a low level of consumption might not have a substantial impact on health

[25]The model estimates $\lambda(a|x) = \exp(\gamma' x)\lambda_0(a)$, the conditional hazard of death, which approximates the instantaneous probability of dying at age a conditional on the covariates x. $\lambda_0(a)$ is the baseline hazard rate that is common across all individuals and can be estimated non-parametrically and independently of the parameter γ. λ_0 therefore controls for the (unconditional) effect of age on mortality.

[26]Following the definition in Yakovlev (2018), a heavy drinker is an alcohol consumer in the top quartile of the distribution of average daily alcohol consumption among males. This threshold corresponds to an average daily alcohol intake of 150 milliliters of pure alcohol.

outcomes, in Column 4 we test for such non-linearities by adding an interaction term. Indeed, the positive interaction term suggests that the share of vodka is more important at higher levels of alcohol. However, adding this term does not significantly change the effect of the share of vodka. Column 5 shows that we obtain comparable results if we include non-working age men, with the interaction term playing a bigger role.

In Column 6 we add the share of beer in addition to the share of vodka and the level of alcohol. We will use this specification in the next section when we simulate the counterfactual long-run effect of moving to a new population steady state with a lower population share of vodka and a higher share of beer.[27]

8 Simulating Two Counterfactuals

In this section we use the individual-level estimates from Section 7.4 to simulate two counterfactual scenarios. First, we decompose the substantial changes in male mortality in the past 20 years into the change that is due to changes in the amount of pure alcohol consumed if consumption was only in the form of light alcohol, and into the additional effect of consuming some of this alcohol in the form of vodka.

Second, we analyze what would happen going forward if there was no further change to the alcohol market or to public policies. This second counterfactual analysis estimates the new population steady state that would be reached because of the shocks in the 1980s and 1990s if no other shocks hit the economy and the alcohol market.

[Figure 5 about here]

8.1 Counterfactual 1: Alcohol and the Recent Decline in Mortality

Panel (a) of Figure 5 shows a substantial decline in male mortality since the mid-1990s, especially since 2003. We use our micro-level estimates of the mortality hazard in Table 4 and RLMS data from 1994-2011 to decompose this decline into the contribution of the shares of alcohol consumption, the amount of alcohol consumed, and all factors other than contemporaneous alcohol consumption. We rescale the predictions in each survey year to match the mortality rates based on official statistics. Based on this decomposition in Panel (b), the share of vodka, holding fixed the level of alcohol intake (i.e., the difference between the solid black line and the blue line with cross markers), explains half of the decline in male mortality.

8.2 Counterfactual 2: Forecasting Male Mortality Rates

Finally, we use the estimated hazard model to study the likely evolution of this downward trend over the next few decades as the economy converges to a new population steady state. To do so we simulate a counterfactual scenario that maintains the sample distribution of all

[27] Adding the share of beer affects the estimated coefficient of the vodka share because both measures are highly correlated (correlation coefficient of -0.63).

individual characteristics except for the shares of vodka and beer consumed. Specifically, we predict the shares of vodka and beer for each individual in our sample by regressing alcohol shares on a full set of cohort effects and the same set of controls used in Table 4, Panel A. To identify the model, we drop survey year fixed effects as these are quantitatively not very important. Using the estimated cohort effects, we then predict everyone's shares at different points in the future and in turn use the predicted shares together with the individual's characteristics to estimate its hazard of death. For example, to predict the hazard of death in 10 years of an individual born in 1970 we maintain its current characteristics, but we assign it the conditional cohort effect of individuals born in 1960. Integrating across the entire sample then provides us with an estimate of the evolution of male mortality as a consequence of the changes in relative alcohol tastes only. Online Appendix G describes this algorithm in detail.

Panel B of Table 4 shows the predicted population consumption shares and the annual mortality rate for the current population of males age 22 to 65 as well as for the corresponding counterfactual populations in 10, 20, and 55 years, with 55 years being the time at which the population reaches its new steady state.[28] Our results suggest that the mortality rate of males age 22 to 65 will decrease by 12% over the next 10 years (from 1.42 pp to 1.25 pp), by 23% over 20 years, and will be cut in half in the new long-run equilibrium. The predicted mortality rate of 1.42% in 2011 is only slightly lower than its official estimated average from 1994 to 2011, which is 1.55% (not adjusted to the standard population).[29] For comparison, the annual mortality rate in the U.S. is 0.5% and 0.4% in the U.K. and Germany. Hence, the counterfactual simulation predicts that the further increase in the population share of beer consumption (and decline in the vodka share) during the transition to the new steady state might further cut the gap between the Russian and U.S. male mortality in half over the next 55 years.

9 Alternative Explanations

In this paper we argue that changes in tastes due to changes in the supply of alcohol during an individual's formative ages changes its tastes for different types of alcohol persistently. An alternative explanation of our results that does not involve tastes is an information-based mechanism. While a policy that persistently changes consumers' information and causes persistent changes in tastes would have the same effect on health outcomes and hence be of similar value to a policymaker, here we discuss evidence that is inconsistent with such an information-based explanation.

First, such a brief information intervention would need to affect rural and urban consumers differently and only affect young consumers. Hence, to explain our causal estimates, this information intervention would have to successfully target young urban consumers that were about 17 years old during the campaign relatively more than their rural counterparts and permanently

[28] Online Appendix Figure A.4 graphs the entire path of both shares and the mortality rate.

[29] This 10% difference reflects the missing single-member households; see Section 3.

change their choices over different alcoholic drinks.[30] If this were indeed the case it would raise the question why policymakers would want to limit such a successful information campaign to urban areas and only to young consumers. However, we are not aware of any study that documents a targeted effort by the Soviet government to differentially inform young urban consumers during the campaign.

One information-based channel that could affect young rural consumers differentially and persistently is if these consumers learned how to moonshine vodka during the campaign because this skill was particularly valuable. However, we find that the campaign affects the long-run share of vodka consumed *excluding* illicit alcohol, which cannot be explained by learning about how to produce home-made samogon.

Second, and more important, such an information campaign would presumably inform consumers about the relative harmfulness of hard alcohol. However, our results show that the campaign permanently *increased* the share of hard alcohol consumed by rural consumers that became adolescents during the campaign, both relative to their urban counterparts but importantly also relative to older as well as younger rural consumers. Hence, unless such a targeted information campaign had the perverse goal of (mis)informing young rural consumers that hard alcohol is less harmful than they thought, it cannot explain our results.

Third, the effects we identify are very persistent. Even if young rural consumers were initially misinformed by the campaign, decades later they would presumably have acquired the same information as most other consumers. However, our results are unchanged if we drop earlier survey rounds that contain individuals that experienced the anti-alcohol campaign more recently and hence might not yet have acquired the same amount of information as the rest.

10 Conclusion

This paper makes three main contributions. First, it documents how public policies, even temporary ones, can have significant long-run effects by shaping tastes of consumers during their sensitive ages. Second, it shows that the age at which most consumers form tastes varies across products and depends on when an individual starts to consume the product regularly. We find that alcohol tastes from around age 17 while tastes for basics foods form during early childhood. Shocks to product availability in the 1980s and 1990s significantly changed tastes of young consumers, and the resulting consumption differences are still large today and thus easily detectable in survey data decades later.

Third, this paper shows that the *type* of alcohol consumed—i.e., hard vs. light alcohol—has a significant effect on mortality *in addition* to the negative effect of the amount of alcohol consumed, which has been the main focus of previous research. We find that effect of changes in the share of alcohol consumed in the form of hard alcohol, holding fixed the amount of alcohol

[30]Similarly, changes in social norms and culture also need to affect consumers differently both in rural and urban areas and as a function of age in order to explain the difference-in-difference estimates.

consumed, on the mortality of working-age men in Russia is about the same as the effect of changes in total alcohol consumed. The reason for this large effect is the fact that a significant fraction of deaths among working-age men are related to alcohol, and most of these in turn are associated with binge drinking, such as alcohol poisoning, traffic accidents and homicides.

Combining these three contributions, we conclude that public policies targeted at young consumers can have significant effects on both contemporaneous and long-run health outcomes by persistently changing consumers' tastes.

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.

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.

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.

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.

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.

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.

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

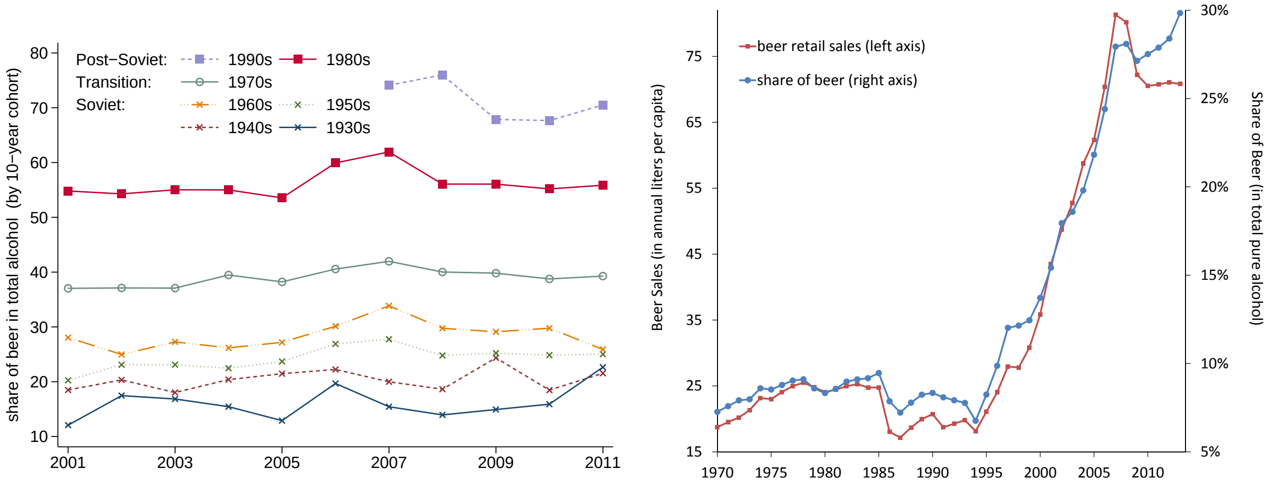
Treml, V. G. (1997). Soviet and russian statistics on alcohol consumption and abuse. Premature death in the new independent states, 220–238.

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 1 – Beer Market Expansion

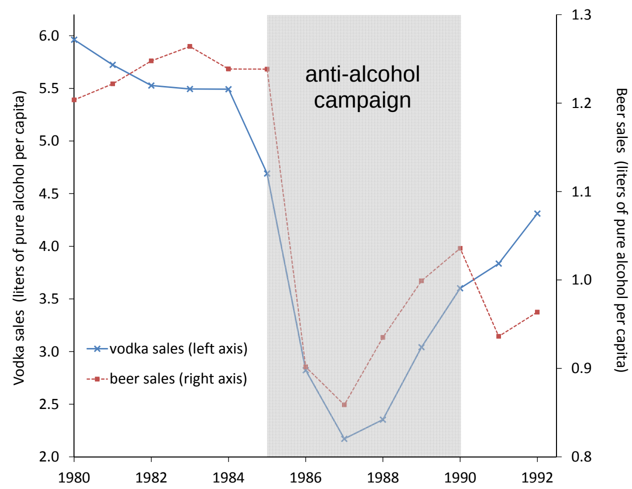


(a) beer consumption across cohorts of adolescents

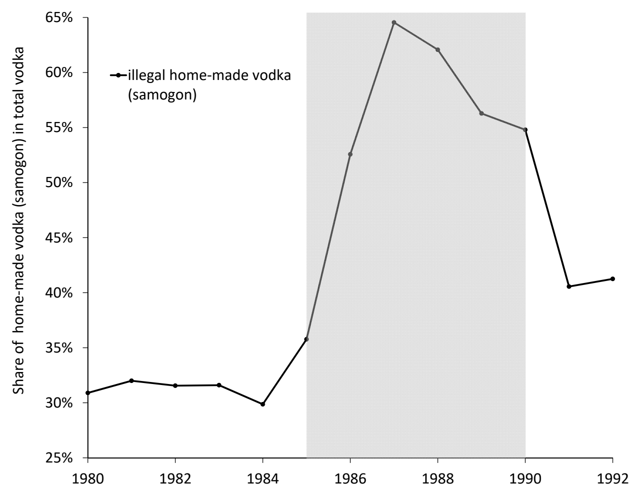
(b) aggregate beer sales

Notes: Panel (a) shows the share of beer consumption across RLMS sample years for 10-year birth cohorts. We classify cohorts as Post-soviet, transition, and Soviet based on the average year in which their members turned 17 (based on our findings of the typical age at which alcohol tastes form; see Section 5). Panel (b) shows the expansion of the aggregate beer market in levels and shares using aggregate data from the Federal State Statistics Service (Rosstat).

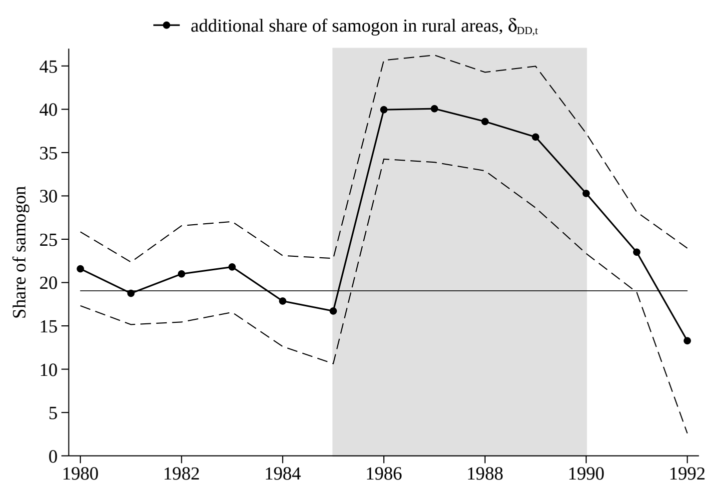
Figure 2 – Effect of the Anti-Alcohol Campaign on the Alcohol Market



(a) effect on legal alcohol



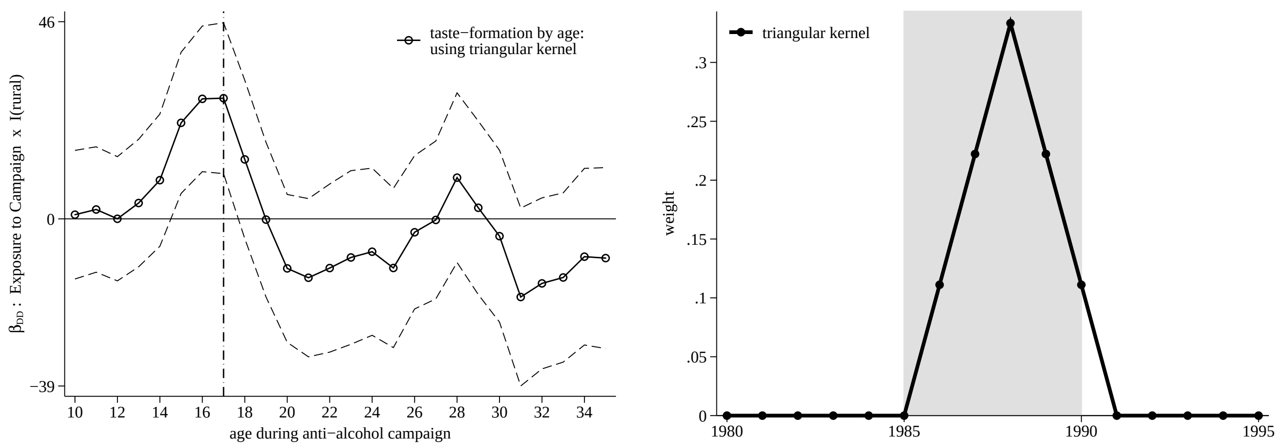
(b) effect on illegal vodka (moonshining)



(c) additional consumption of illegal vodka in rural areas

Notes: Panels (a) and (b) show the aggregate effect of the anti-alcohol campaign on alcohol consumption using *national*-level data. (a) shows the campaign's effect on official alcohol sold and (b) on the share of home-made vodka (samogon) in total vodka consumption. Panel (c) shows the marginal effect ($\beta_{DD,t}$) of the campaign on rural areas using difference-in-differences with *regional*-level data from 1980-1992: $S^v_{rt} = \alpha_r + \sum_{t=1981}^{1992} \left[\delta_{D,t} \cdot \mathrm{I(year)}_t + \delta_{DD,t} \cdot \mathrm{I(year)}_t \times \text{Rural Fraction}_r\right] + \varepsilon_{rt}$. Rural Fraction$_r$ is region r's fraction of rural population and S^v_{rt} is its share of samogon consumed in year t.

Figure 3 – Taste Formation as a Function of Age

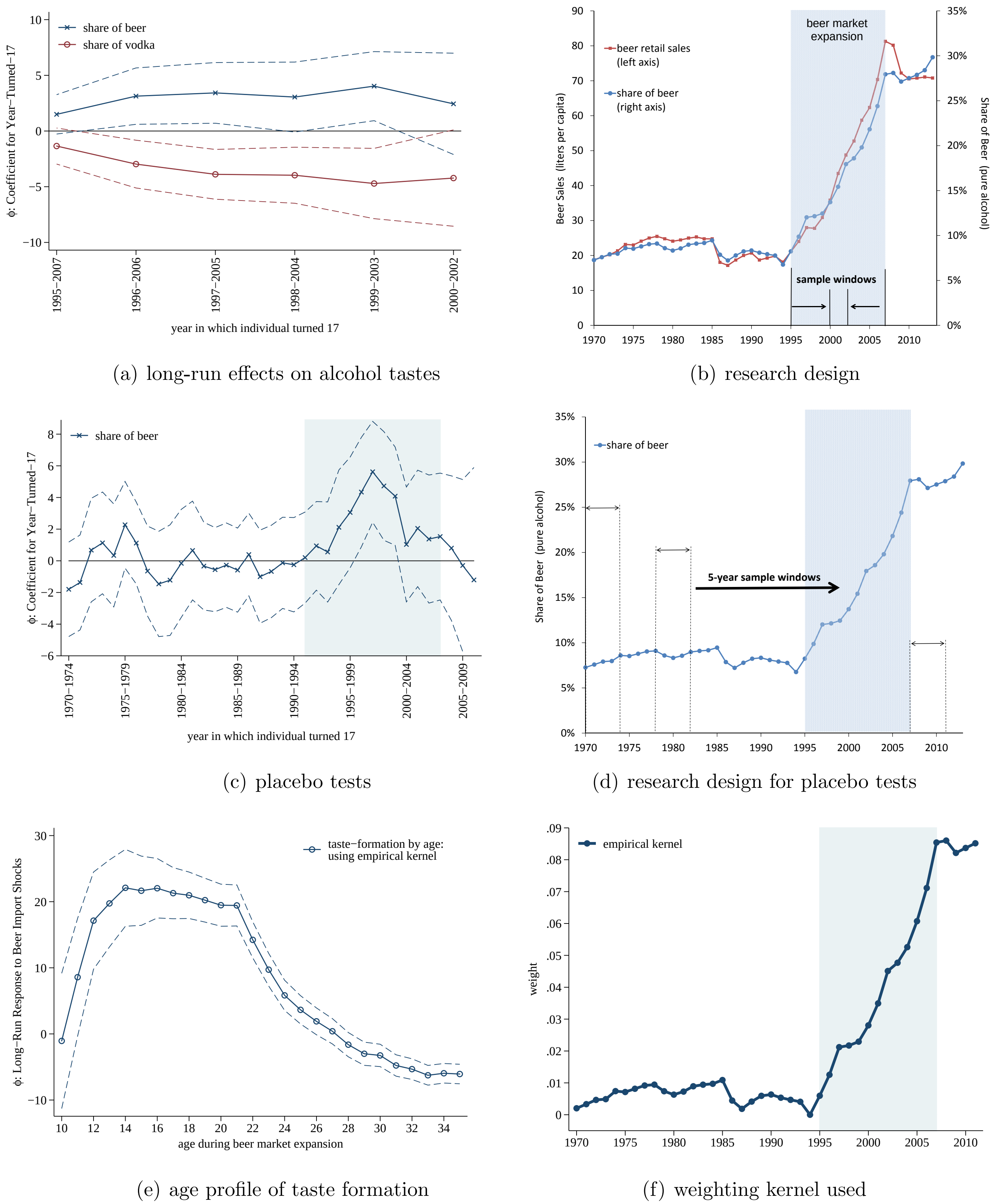


(a) response to anti-alcohol campaign by age

(b) 5-year triangular weighting kernel

Notes: Panel (a) shows the estimated age at which tastes for vodka form. Panel (b) plots the weighting kernel used. Dashed lines are two standard error confidence bands using robust standard errors clustered by individual. The anti-alcohol campaign is shaded in gray in Panel (b), and the maximum impact of the campaign on the age of taste formation is shown with a vertical dashed line in Panel (a).

Figure 4 – Long-Run Effects of the Beer Market Expansion on Alcohol Tastes



(a) long-run effects on alcohol tastes

(b) research design

(c) placebo tests

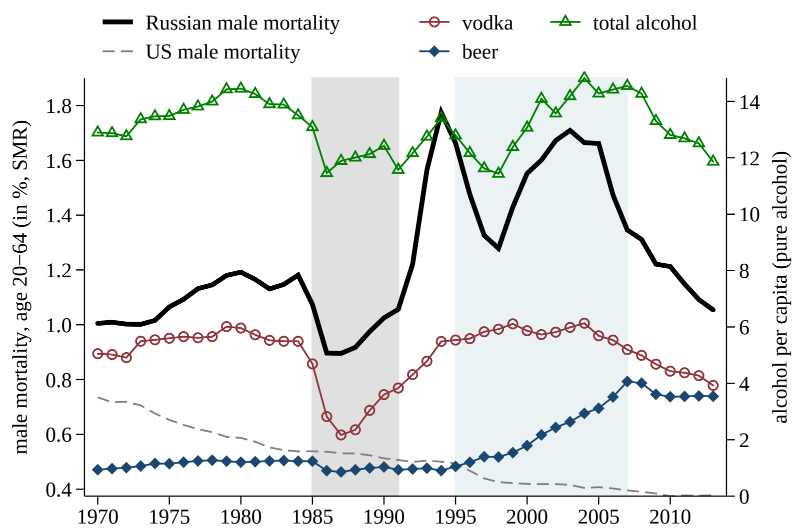
(d) research design for placebo tests

(e) age profile of taste formation

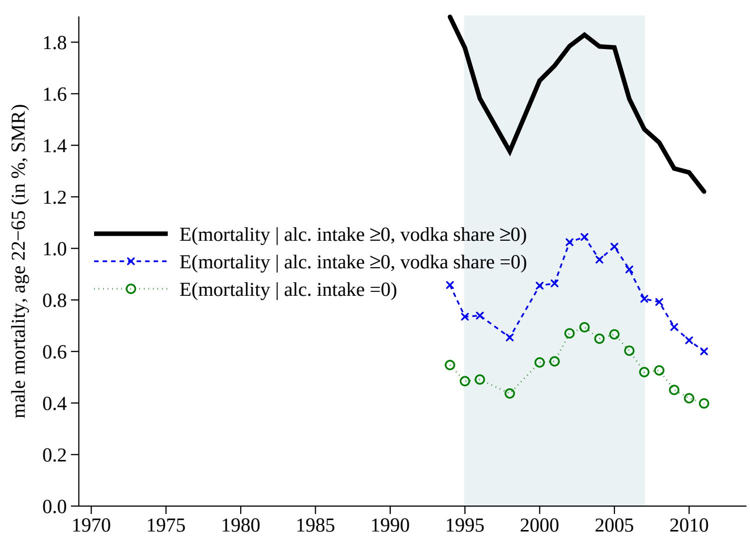
(f) weighting kernel used

Notes: The right panels show the research designs and the empirical kernel used to identify the effect of the beer market expansion on alcohol tastes, shown in the left panels (see Section 5 for details). The regressions control for the level of total alcohol intake, the log of real income, subjective health status, body weight, education, marital status, and a full set of year, age, and region fixed effects. The beer market expansion is shaded in light blue. Dashed lines show two standard error confidence bands using robust standard errors clustered by individual.

Figure 5 – Alcohol Tastes and Mortality



(a) types of alcohol and male mortality



(b) decomposition of the effect of alcohol on male mortality

Notes: This figure shows the effect of alcohol on mortality of working-age men. The anti-alcohol campaign is shaded in gray and the beer market expansion in light blue. Standardized mortality rates (SMR) use the U.S. standard population of 2000. Panel (b) decomposes the effect of alcohol on male mortality using the estimated hazard model in Table 4 and RLMS data from 1994-2011. The green line with circle markers shows the predicted mortality rate for abstainers calculated by setting the amount of alcohol consumed (and the share of vodka) to zero. The blue line with cross markers adds consumers who consume all alcohol in the form of beer (i.e. with positive alcohol intake but a share of vodka of zero). The solid black line shows the unconditional mortality rate, adding men who also consume vodka.

Table 1: Descriptive Statistics

	N	Mean	St.Dev.	p75	N	Mean	St.Dev.	p75
Alcohol Shares	Males				Females			
Share of beer	46,985	29.3	35.3	38.5	45,182	22.6	35.5	32.9
Share of home-brewed beer (starts in 2008)	14,363	0.1	1.5	0	14,837	0.0	0.6	0
Share of vodka	46,985	52.9	39.7	92.3	45,182	34.9	42.1	78.4
Share of home-produced vodka (samogon)	46,985	8.7	24.3	0	45,182	3.7	17.0	0
Share of wine	46,985	7.4	20.9	0	45,182	35.7	42.3	100
Share of other alcohol	46,985	1.8	10.9	0	45,182	3.1	15.2	0
Socio-Economic Demographics	Male Alcohol Consumers				All Males			
Age	46,985	41.4	15.4	52	68,350	42.5	16.4	54
Birth year	46,985	1962.4	16.4	1976	68,350	1961.3	17.3	1975
I(turned 17 in a rural area)	46,972	0.6	0.5	1	68,322	0.5	0.5	1
Daily alcohol intake when drinking (in g of ethanol)	46,985	144.7	133.8	200.0	68,350	101.4	132.0	146.0
I(no alcohol consumed in the past 30 days)	46,985	-	-	-	68,350	0.3	0.5	1
Total monthly real income	45,280	245.1	426.6	300.3	65,688	233.7	404.2	288.3
I(college degree)	46,950	0.4	0.5	1	68,290	0.4	0.5	1
Subjective health status (1=very good, 5=very bad)	46,884	2.7	0.7	3	68,186	2.7	0.7	3
Body weight (in kg)	44,180	76.7	13.7	85.0	64,114	76.5	13.6	85.0
I(married)	46,985	0.7	0.5	1.0	68,350	0.7	0.5	1

Notes: We use rounds 5 to 20 of the Russian Longitudinal Monitoring Survey (RLMS) corresponding to years 1994-2011 except for 1997 and 1999 when the survey was not conducted. All individuals are age 18 or above. To safe space, the indicators I(turned 17 in a rural area) and I(no alcohol consumed in the past 30 days) are referred to as I(rural) respectively I(abstainer) from here on.

Table 2: Long-Run Effect of the Anti-Alcohol Campaign on Alcohol Tastes

Dependent variable:	Share of vodka				Log(alcohol)	I(abstainer)	Share of beer	Share of hard alcohol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I(became adolescent during campaign) × I(rural)[(a)]	5.243***	5.049**	5.008**	5.232***	7.594*	0.280	-3.129*	3.027*
	[2.016]	[2.009]	[1.998]	[1.986]	[4.585]	[2.247]	[1.730]	[1.780]
I(became adolescent during campaign)[(b)]	-1.296	-0.930	-0.945	-1.038	-8.431**	1.145	-0.069	-0.893
	[1.514]	[1.497]	[1.490]	[1.483]	[3.532]	[1.679]	[1.370]	[1.385]
Sum of diff-in-diff and diff coefficients, (a)+(b)	3.947***	4.119***	4.063***	4.194***	-0.836	1.425	-3.198**	2.134
	[1.524]	[1.527]	[1.521]	[1.505]	[3.336]	[1.724]	[1.248]	[1.313]
Alcohol intake (in grams of ethanol)		0.063***	0.064***	0.065***			-0.104***	0.099***
		[0.003]	[0.003]	[0.003]			[0.003]	[0.004]
Year, age, region, rural FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Real income and relative price			Yes	Yes	Yes	Yes	Yes	Yes
Socio-economic demographics				Yes	Yes	Yes	Yes	Yes
Observations	29,083	29,083	29,083	29,083	29,083	40,507	29,083	29,083
R-squared	0.107	0.144	0.146	0.152	0.076	0.075	0.258	0.256
Sample mean of dependent variable	47.37	47.37	47.37	47.37	459.3	27.81	36.66	57.50

Note: 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. Adolescence is defined as being 17 years old based on the analysis in Section 5. Both of these assumptions are relaxed in Online Appendix C. Since the level of alcohol is highly skewed to the right, Column 5 winsorizes the dependent variable at the 95th percentile. For comparability with the alcohol shares, which are in percentages, I(abstainer) and Log(alcohol) are multiplied by 100; see the sample mean reported in the last row. We also report the sum of the coefficients of the difference-in-difference and the single difference estimators, (a)+(b), which captures the effect of the campaign on rural consumers that became adolescent during the campaign relative to rural consumers that became adolescent before or after the campaign. The main effect, I(rural), indicates the place of residence at age 17 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 3: Effect of Alcohol Tastes on Mortality

Dependent variable:	A. National Male Mortality Rate (in logs)		B. Regional Male Mortality Rate (in logs) by Cause of Death:					C. 1st Stage:
			all causes		alc. poisoning	external	cancer	vodka share
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Share of vodka	1.456***	1.454***	1.253***	1.271***	3.836**	1.230**	-0.190	
	[0.251]	[0.200]	[0.455]	[0.473]	[1.532]	[0.523]	[1.225]	
I(became adolescent during campaign) × I(rural)								6.234***
								[1.950]
Log(total alcohol)		1.423***		-0.032	-0.189	-0.017	0.005	0.059***
		[0.293]		[0.039]	[0.125]	[0.042]	[0.073]	[0.013]
I(became adolescent during campaign)			-5.281***	-5.440***	-10.331**	-6.455***	1.726	-1.965
			[1.488]	[1.524]	[4.934]	[1.674]	[3.813]	[1.298]
Log population	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quadratic time trend	Yes	Yes						
Year, age, region, rural FE and log income			Yes	Yes	Yes	Yes	Yes	Yes
Observations	44	44	1,343	1,343	1,327	1,343	1,273	1,343
R-squared	0.785	0.868	0.596	0.592	0.592	0.654	0.584	0.245
1st-stage F-statistic			10.55	10.22	10.22	10.22	8.744	
Estimator	OLS	OLS	IV	IV	IV	IV	IV	OLS

Notes: Columns 1-2 use the national male mortality rate and alcohol sales by type of alcohol from 1970-2013. Columns 3-8 use regional alcohol sales by year, region, age and type of settlement (urban/rural) from 1994-2011 and the total number of deaths (Columns 3-4) as well as cause-specific deaths (Columns 5-7), such as alcohol poisoning fatalities, external but alcohol-related causes of death, and cancer deaths unrelated to alcohol. The set of additional controls includes log populaton and a quadratic time trend in Columns 1-2 and log population, log regional GDP per capita and year, age, region and rural FE in Columns 3-8. Robust standard errors in squared brackets are clustered by region. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

Table 4: Individual-Level Mortality Analysis and Simulation of Counterfactuals

A. Cox Proportional Hazard Model	(1)	(2)	(3)	(4)	(5)	(6)
Share of vodka	0.619***	0.630***	0.631***	0.686***	0.534***	0.512**
	[0.190]	[0.192]	[0.192]	[0.196]	[0.158]	[0.205]
Total alcohol (kg of ethanol)		1.125**	1.022	-1.967	-3.064**	-2.116
		[0.526]	[0.642]	[1.527]	[1.426]	[1.529]
Share of vodka × Total alcohol				5.342**	7.245***	5.400**
				[2.207]	[1.876]	[2.172]
Share of beer						-0.842**
						[0.426]
Socio-economic demographics	Yes	Yes	Yes	Yes	Yes	Yes
Heavy-drinking indicator			Yes	Yes	Yes	Yes
Sample (age)	22-65	22-65	22-65	22-65	≥ 22	22-65
Observations	6,623	6,623	6,623	6,623	7,506	6,623
Number of deaths events	356	356	356	356	599	356
Pseudo R-squared	0.0852	0.0858	0.0858	0.0873	0.0736	0.0885

	population share of		male mortality,
B. Counterfactual Simulations	vodka	beer	age 22-65 (%)
current year	46.2	31.4	1.43
in 10 years	32.0	42.4	1.26
in 20 years	22.5	49.4	1.10
new long-run steady state	14.5	55.9	0.81

Notes: The dependent variable in Panel A are death events of family members in non-single households in the RLMS from 1994-2011. The independent variables have a sample mean (standard deviation) of 0.11 (0.09) liters for total alcohol, 0.46 (0.32) for the share of vodka, and 0.31 (0.29) for the share of beer. ***, **, * denote statistical significance at the 1%, 5%, and 10% level, respectively.

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 (http://data.worldbank.org) and mortality rates by gender are based on the Human Mortality Database (www.mortality.org). National-level data on alcohol sales going back to 1970 are provided by the Federal State Statistics Service (FSSS or Rosstat, www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/main) 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).

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 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 (FSSS). 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 (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. The information is based on official but previously unpublished data from the FSSS; see www.demogr.nes.ru/en/demogr_indicat/data for more details. 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 (RLMS). We supplement this data with additional household expenditure data from the National Survey of Household Welfare and Program Participation (NOBUS). NOBUS covers new or "western" goods better than the RLMS. These new goods became increasingly available only due to increased imports and foreign direct investments after the collapse of the Soviet Union in the second half of the 1990s.

A.3.1 Russian Longitudinal Monitoring Survey (RLMS)

The RLMS is a nationally representative survey conducted by the Carolina Population Center at the University of Carolina at Chapel Hill and the Higher School of Economics in Moscow and covers 33 regions (Russian oblasts) 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.[31] 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%.[32] 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).

[31] 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.

[32] 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 3, Brainerd and Cutler (2005) use the same data for their analysis of mortality trends in Russia and summarize the data as follows:

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

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

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

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

B Taste 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.[33] 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

[33] 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.[34] 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.

[34]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 6, 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 4. 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:

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

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

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.[35]

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 5 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

[35] 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.[36] 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 5. 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 5 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

[36] 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 5 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.6 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.[37]

Second, we use information about the birth country for individuals who moved to Russia from another republic of the former Soviet Union.[38] While vodka and beer production was

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

[38] 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.[39] 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

[39] 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.[40] 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

[40]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.7 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.[41]

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.[42] 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

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

[42]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

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

In general, this is a non-monotonic function in the steady-state vodka consumption v.[43] 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.[44] 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.

[43] 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.

[44] 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,

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

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

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

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

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.

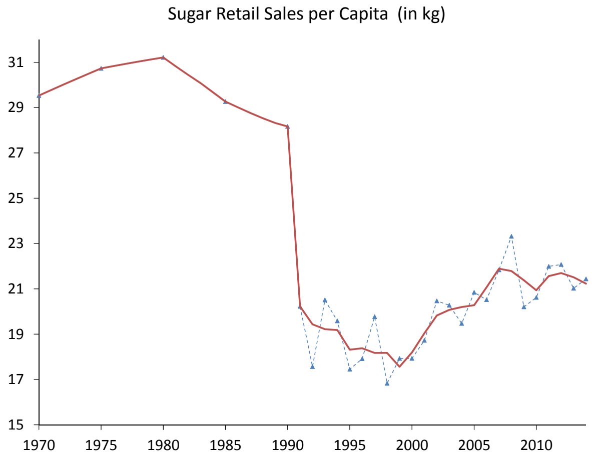
Recall that in Table 3 we use the difference-in-differences estimator from the anti-alcohol campaign to instrument for the share of vodka in the 2^{nd}-stage IV regression of mortality on alcohol poisoning. Unfortunately, as shown in Column 2 of Table A.8, we do not find statistically significant changes 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 2^{nd} stage and Columns 2 and 3 show the corresponding 1^{st} 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.

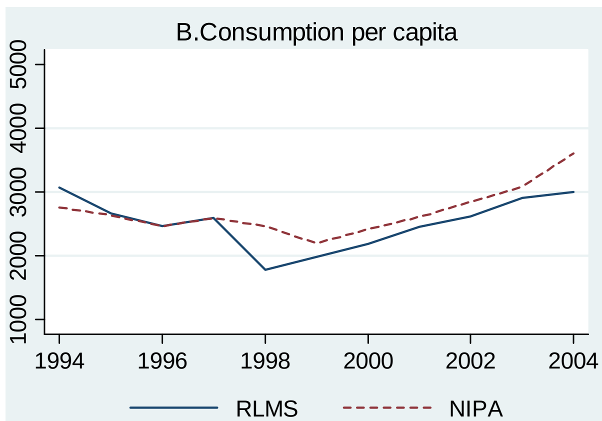
Both IV regressions suggest that the share of samogon consumption and the share of vodka consumption both increase mortality from alcohol poisoning. That is, controlling for the effect of samogon consumption on mortality does not decrease the effect of vodka consumption on mortality from alcohol poisoning. 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.

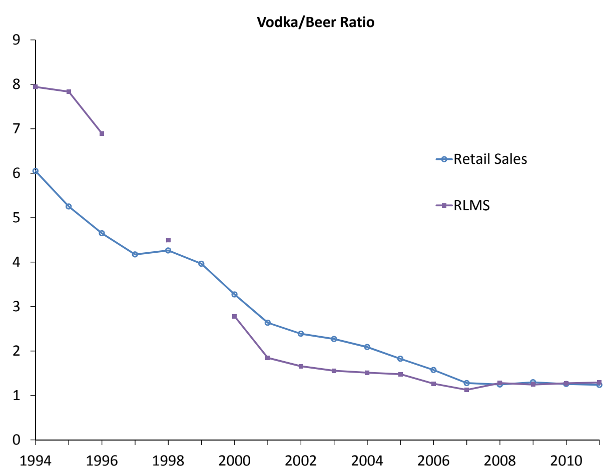
Figure A.1 – Comparison of RLMS Data with Official Statistics



(a) sugar consumption



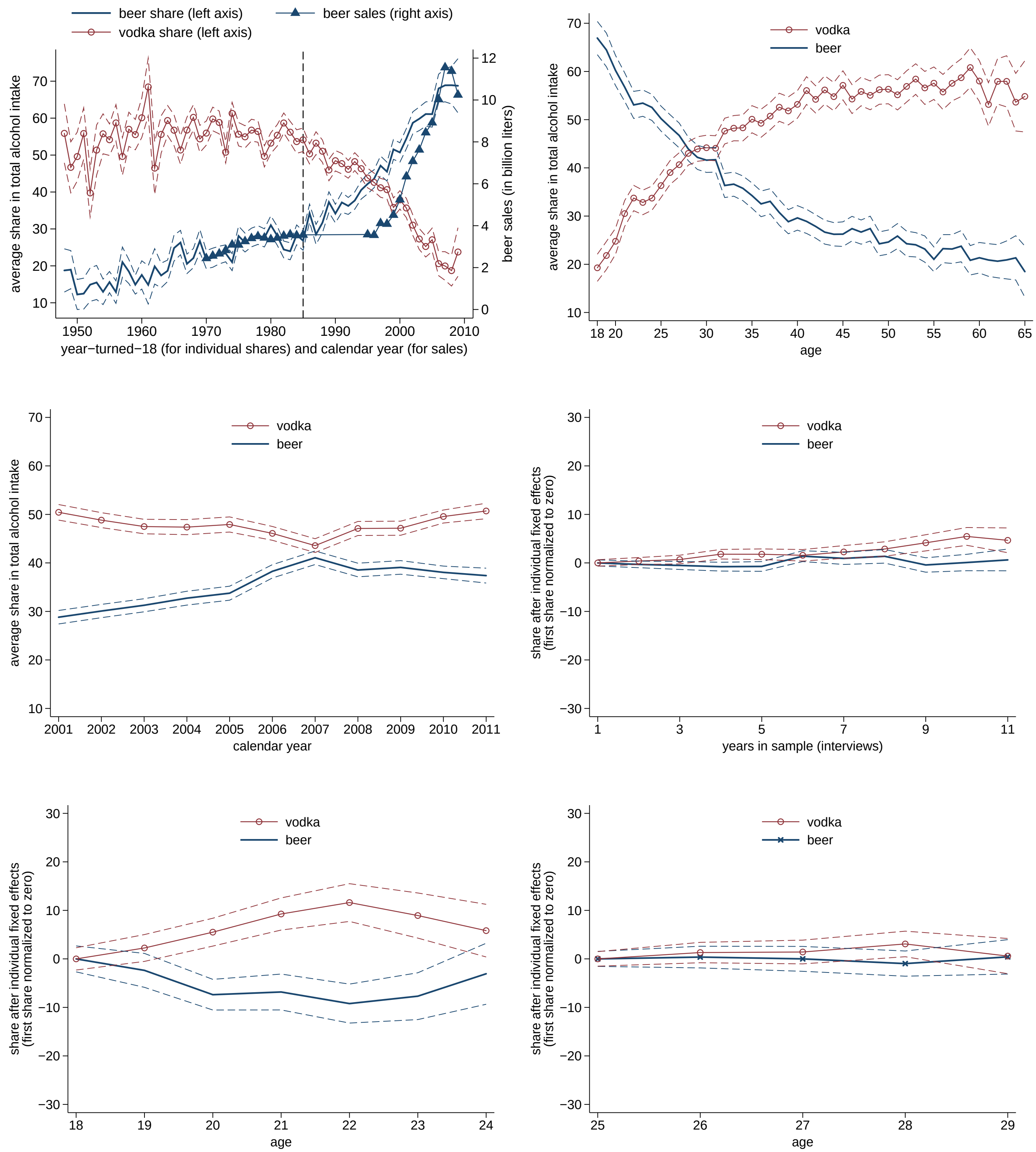
(b) comparison of RLMS and aggregate consumption



(c) comparison of RLMS and alcohol retail sales

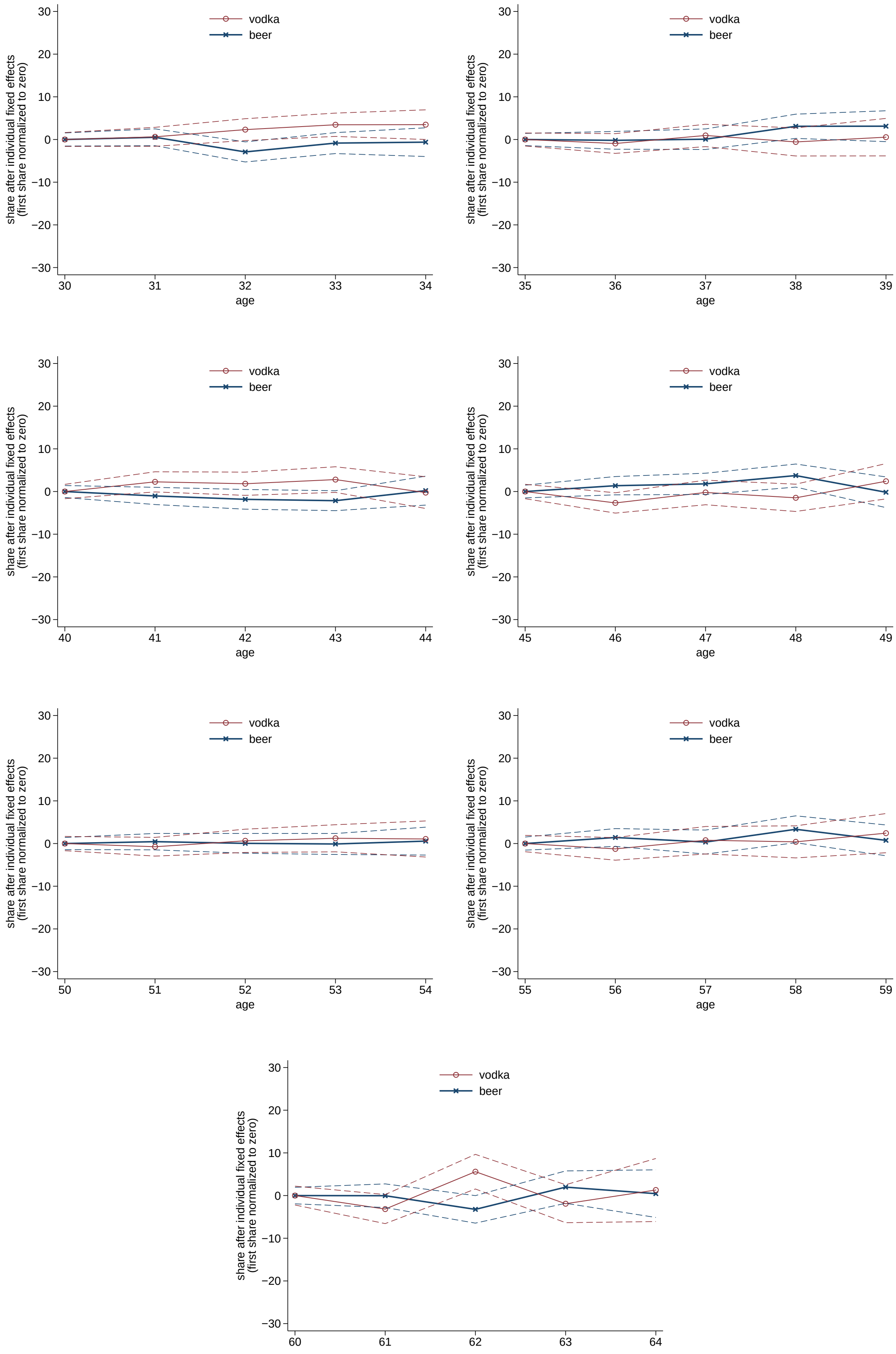
Notes: Panel (a) shows the time series of per capita retail sales of sugar from 1970-2014, an important ingredient of illicit vodka (samogon). The blue line is the raw data and the red line is the corresponding 2-year moving average. Panels (b) and (c) compare consumption data obtained from the RLMS 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



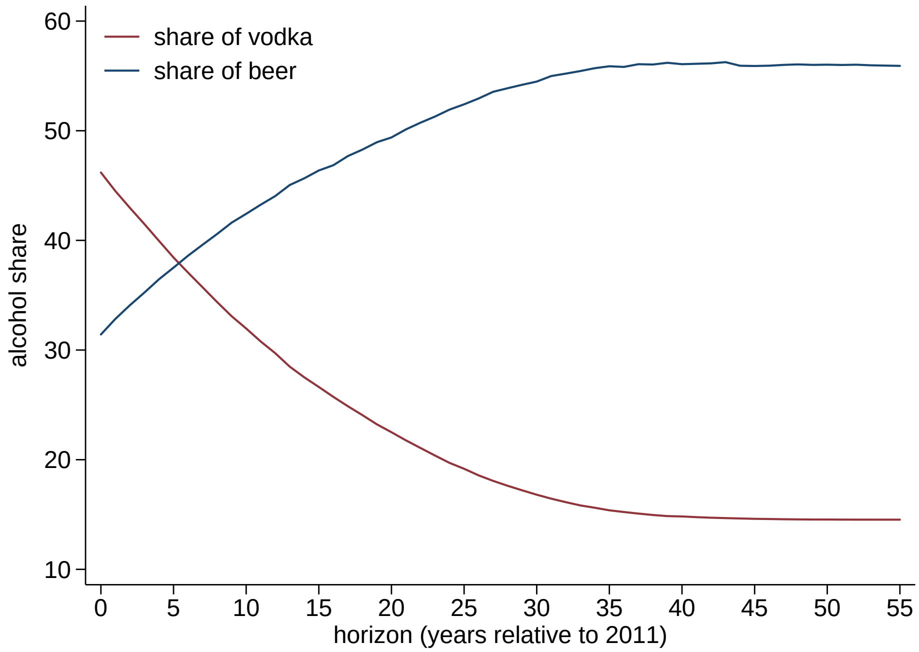
Notes: These figures show the profiles of the shares of beer and vodka consumed by men in the RLMS. The dashed lines represent two standard error confidence intervals. The top-left panel shows the age profile. The top-left panel shows the alcohol shares by cohorts measured by when and individual turned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

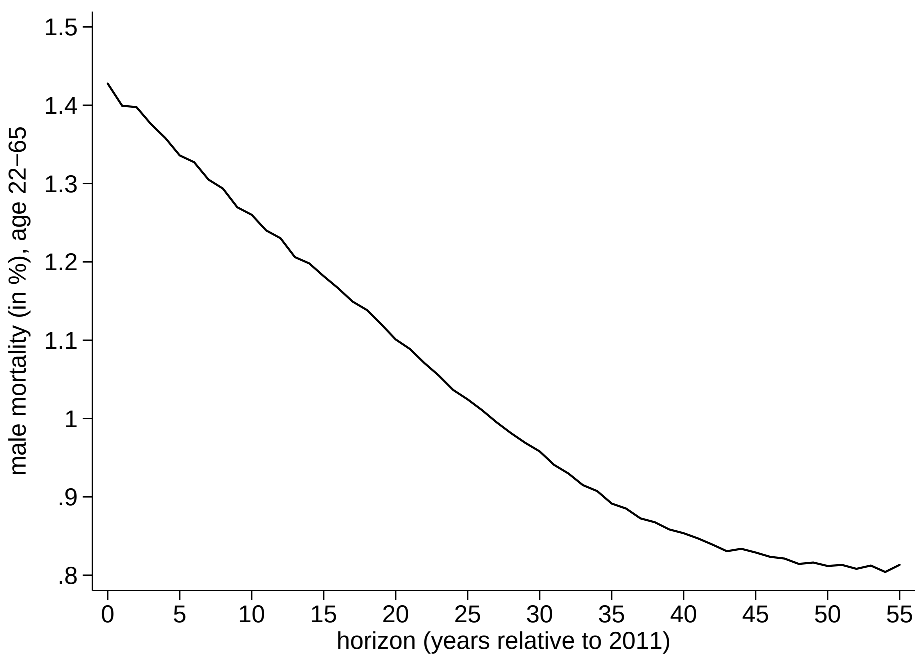


Notes: These figures provide the same analysis over the remaining part of the life-cycle as in Figure A.2.

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



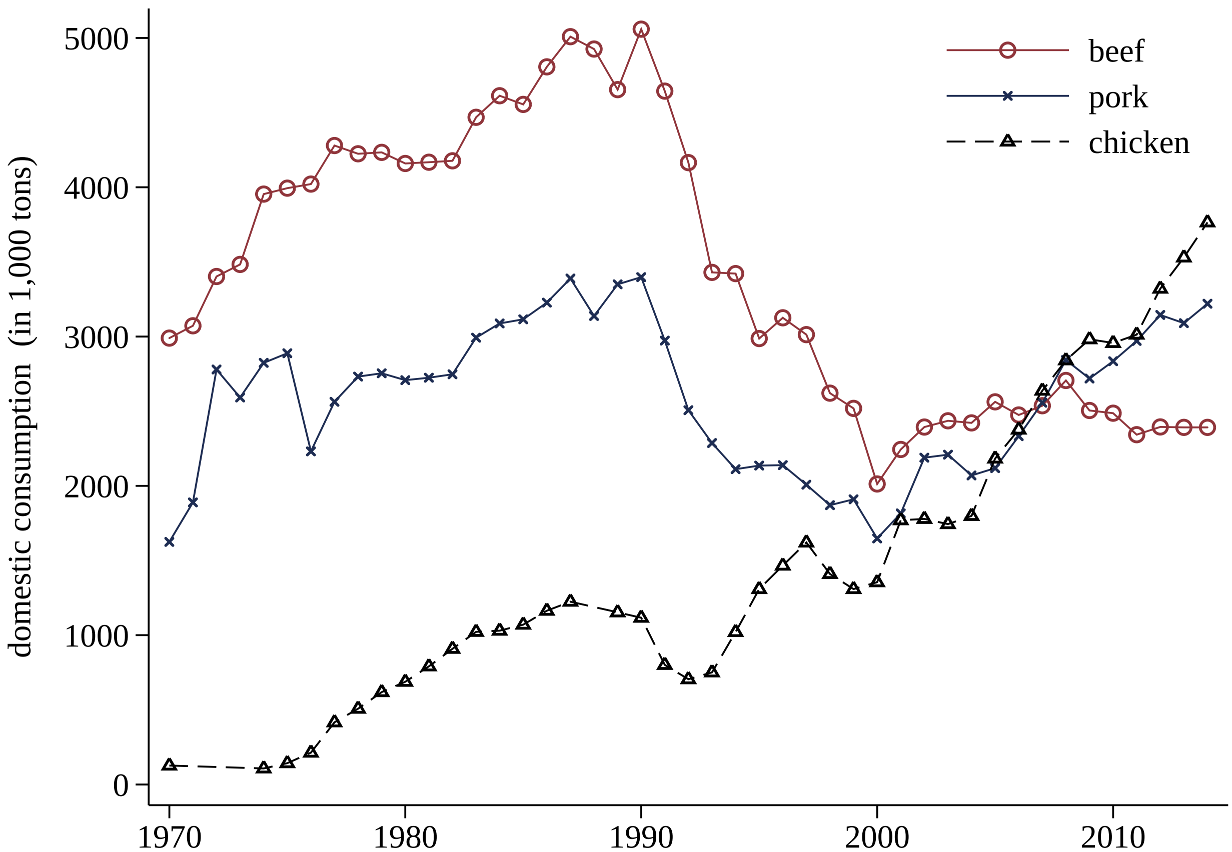
(a) predicted alcohol share



(b) predicted male mortality rate

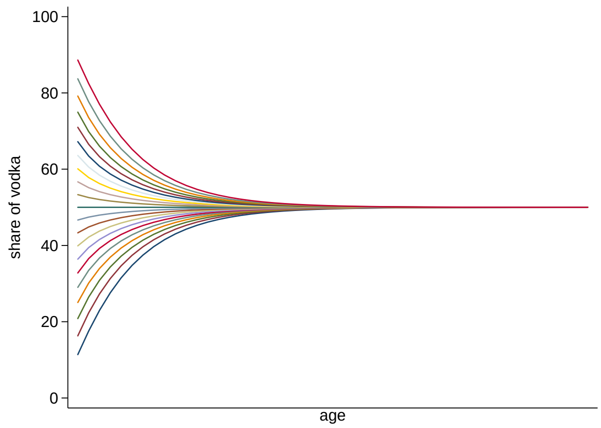
Notes: These figures shows the predicted consumption shares (Panel a) and implied mortality rates (Panel b) for males age 22 to 65 as a function of the forecast horizon in years.

Figure A.5 – Expansion of the Meat Market

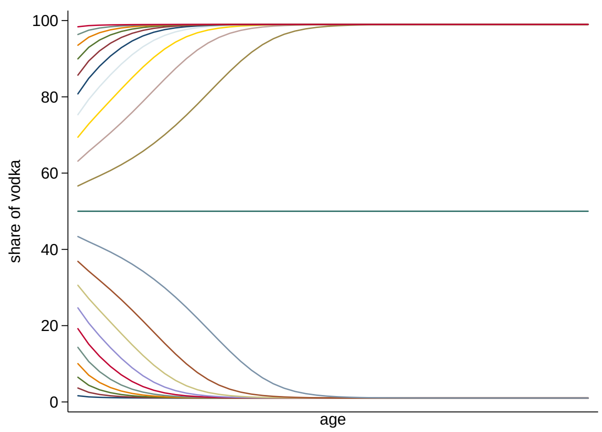


Notes: This figure shows the expansion of the meat market after the end of the Soviet Union.

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



(a) one equilibrium

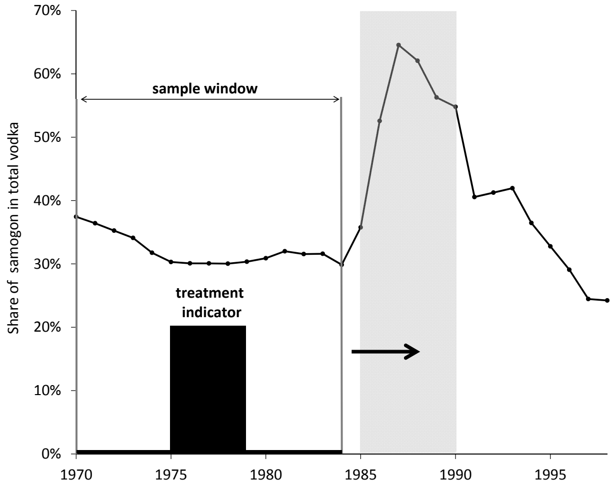


(b) three equilibria

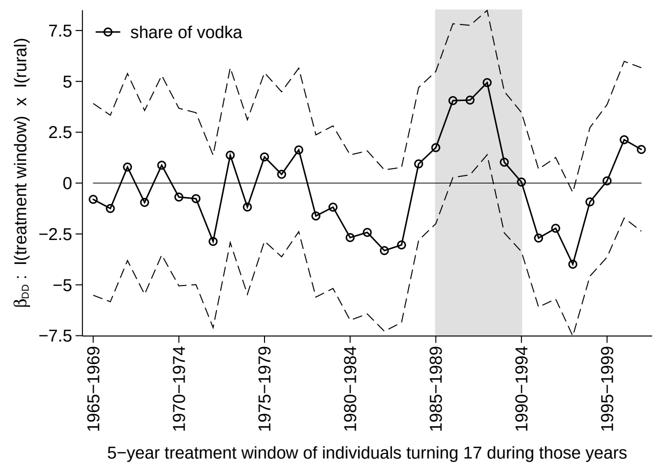
(c) infinitely many equilibria

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.

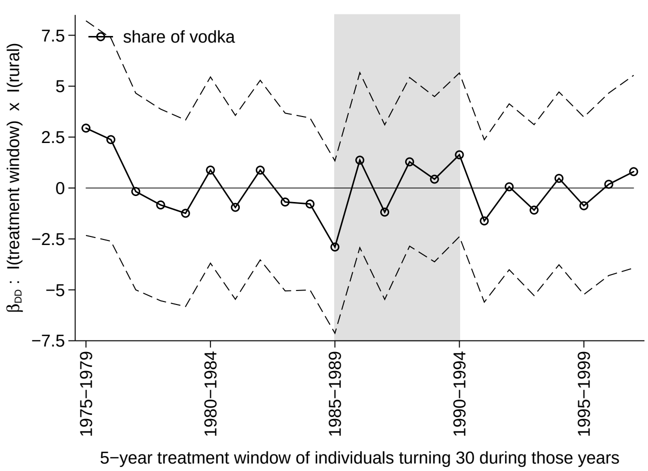
Figure A.7 – Placebo Tests for Anti-Alcohol Campaign



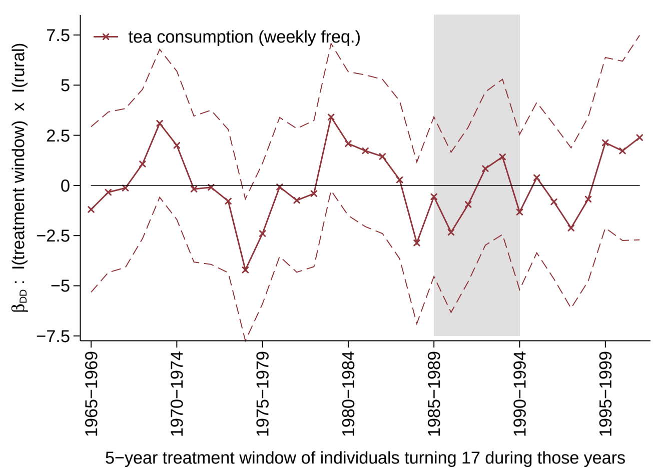
(a) research design for placebo tests



(b) placebo tests with vodka consumption



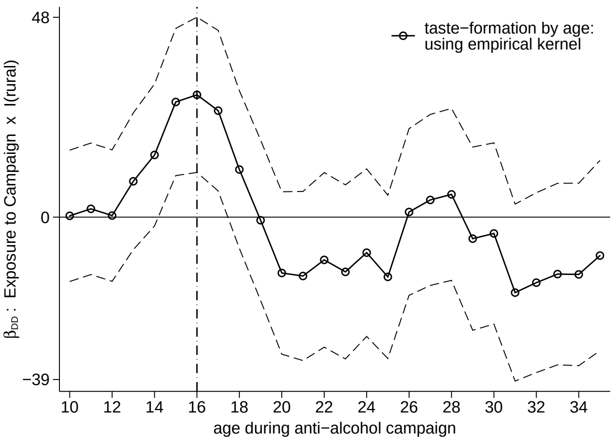
(c) placebo: (non)-response of 30-year olds



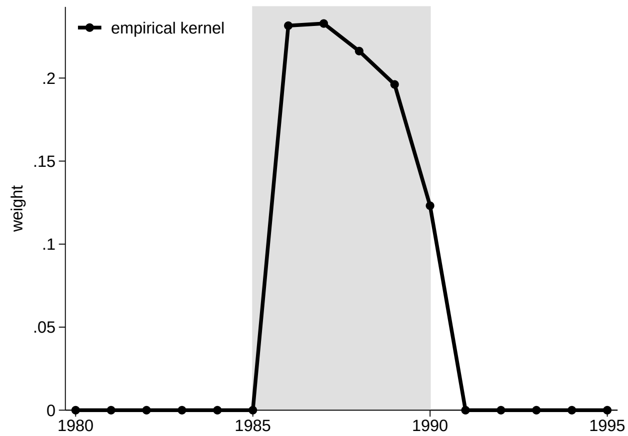
(d) placebo tests with tea consumption

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.

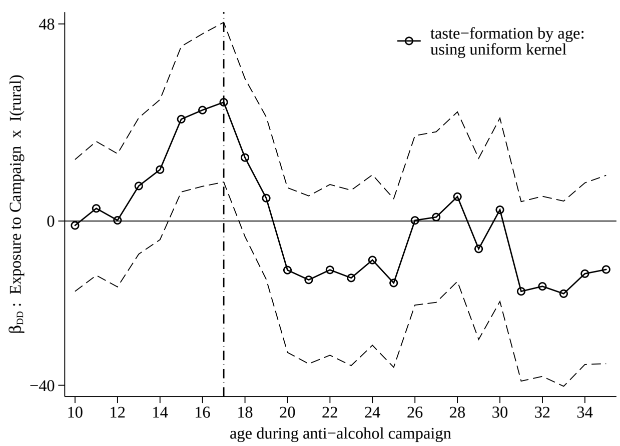
Figure A.8 – Taste Formation as a Function of Age – Robustness to Different Kernels



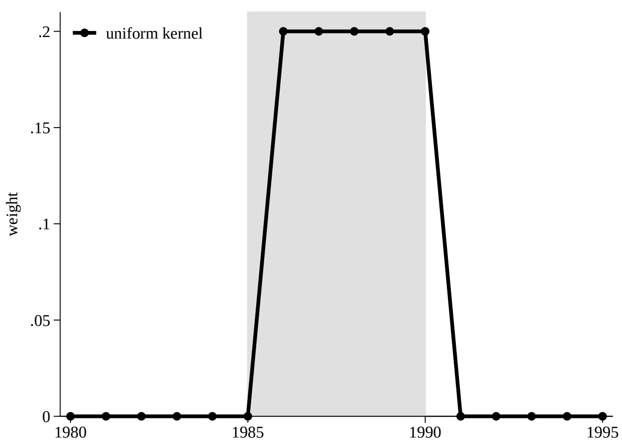
(a) response to anti-alcohol campaign by age



(b) empirical weighting kernel



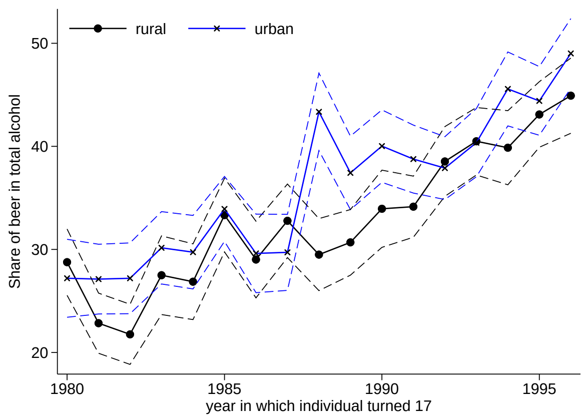
(c) response to anti-alcohol campaign by age



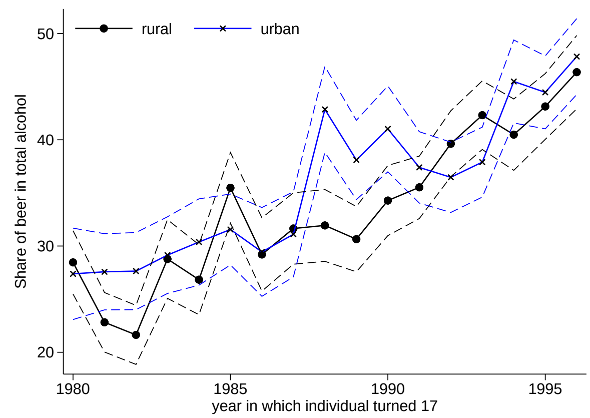
(d) 5-year uniform weighting kernel

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.

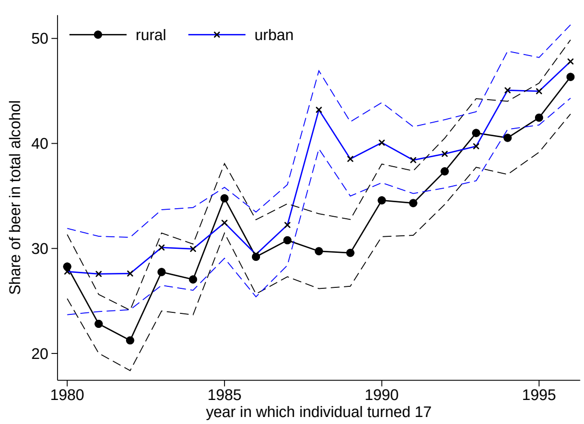
Figure A.9 – Long-Run Share of Beer Consumption of Urban and Rural Adolescents



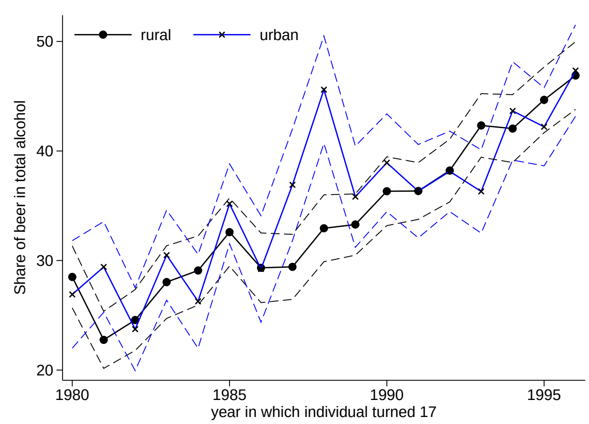
(a) main specification



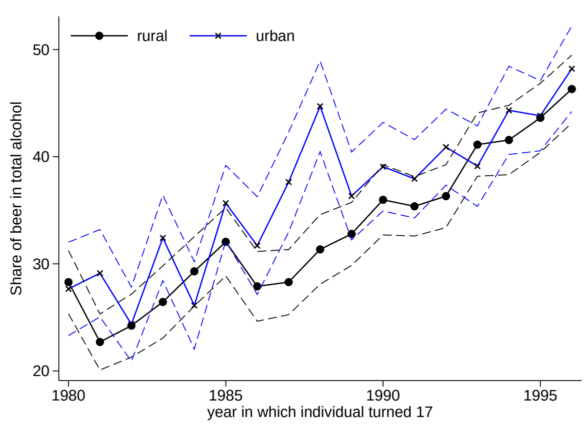
(b) using birth place or current population<250k



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



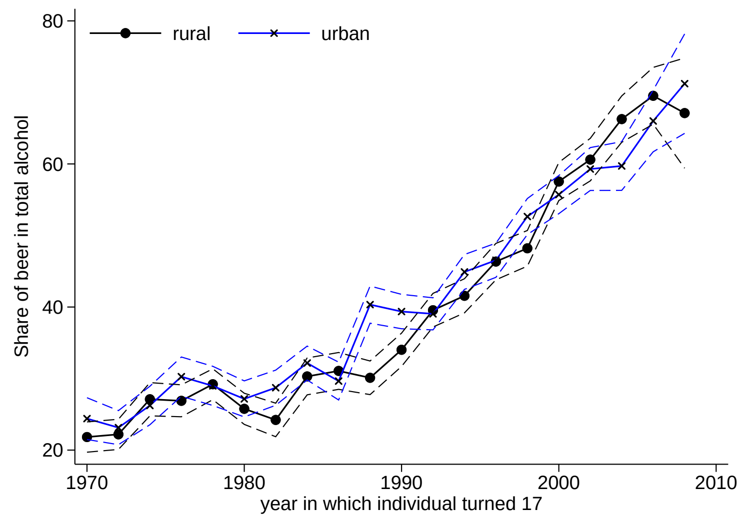
(d) using current population<100k



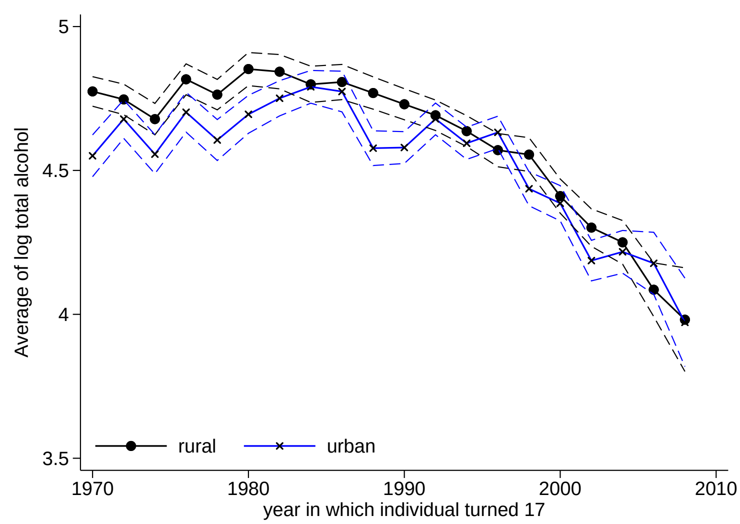
(e) using current population<250k

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

	2^{nd} Stage:	1^{st} Stage:		2^{nd} Stage:	1^{st} Stage:
Dependent variable:	alc. poisoning	samogon share	vodka share	alc. poisoning	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.