How Persistent Are Consumption Habits? Micro-Evidence From Russian Men*

Lorenz Kueng† Evgeny Yakovlev‡

Very preliminary - please do not cite or circulate

September, 2013

Abstract

Analyzing individual-level data on the alcohol consumption of Russian males we find evidence for a longstanding persistence of habits towards certain type of habit-forming goods. Males who grew up in the USSR are accustomed to vodka – the most popular liquor during the Soviet era – whereas those who entered their twenties in the post-Soviet period when the beer industry expanded prefer beer. This finding emphasizes the importance of policies targeted at young people as they form their habits. The second main finding is that habits and substitution effects outweigh "stepping stone" effects, both in the short and long run. Policy simulation shows that a 50% subsidy on beer and 30% tax on vodka will decrease male mortality from 1.41% to 0.95% in 10 years, cutting in half the gap between Russian and western-European mortality rates.

*We thank David Card, Irina Denisova, Sergey Guriev, Denis Nekipelov, Katya Zhuravskaya, and seminar participants at EEA meeting, NES and CEFIR for helpful discussions and comments.

†Northwestern University's Kellogg School of Management and NBER

‡New Economic School (NES). Corresponding author: Evgeny Yakovlev, New Economic School, 47 Nakhimovsky prospect, 117418 Moscow, Russia, e-mail: eyakovlev@nes.ru

1 Introduction

How persistent are habits? This paper provides evidence that state-dependence can be very persistent. The initial choice of a habit-forming good affects individual choices even decades later. Using a long panel data set on alcohol consumption of Russian males, we show that a person who starts consuming a certain type of alcohol at earlier ages forms strong habits for this type of beverage that last his entire lifetime. Importantly, the persistence of such habits applies to all levels of alcohol consumption and is not limited to heavy alcohol consumption or alcoholism. Figure 1 illustrates the persistence of consumption habits by showing strong cohort differences in alcohol consumption patterns. Although there is in a general a small positive trend towards a higher share of beer consumption as a percentage of total alcohol intake, preferences regarding beer and vodka have not changed significantly over the past ten years. Those born in the 1970s or earlier prefer vodka whereas younger generations prefer beer. In particular, vodka constitutes on average 57% of total alcohol intake for males born in the 1970s or earlier, but only 31% for those born in the 1980s, and 16% for those in the 1990s. In contrast, the share of beer in alcohol intake for these age groups constitutes 24%, 56%, and 68% respectively.

[Figure 1 about here]

We propose a simple explanation for this phenomenon. The vodka industry dominated the alcohol market during the era of the Soviet Union. Since 1992, however, the beer industry has expanded rapidly for reasons that are largely exogenous to these preference changes, such as the liberalization of the alcohol market after the collapse of the Soviet Union, a lower regulatory burden for the beer industry in particular compared to all other alcoholic beverages, and the entry of foreign beer companies into this new market. In 1991, before the collapse of the USSR, there were no foreign-owned beer breweries in Russia at all and no foreign brand was sold. However, already by 2009, less than 20 years later, the five leading foreign-owned companies – Carlsberg, Anheuser-Busch, SABMiller, Heineken, and Efes – produced combined more than 85% of the total beer sold in Russia. Opening markets lead to the introduction of new technologies. For example, beer sold in cans or in plastic bottles started to be produced only after the collapse of the Soviet Union. Brewing technologies have changed significantly as well,[1] and the assortment of beer has increased dramatically from only 20 varieties offered in 1991 to over 1,000 in 2009.[2] As a result, in the 10 years since 1991, beer sales have increased five-fold: sales in 2001 exceeded 10.7 billion liters compared to only 1.55 billion liters in 1991. In contrast, vodka sales have not followed the same trend as is evident from Figure

[1]See for example http://moepivo.narod.ru/about_beer/brewing-in-the-ussr.html and http://www.beerunion.ru/soc_otchet/2.html.

[2]The set of varieties available in 1991 was even more limited than this number suggests since one brand – Zhigulevskoe – dominated the entire industry.

2. Total annual sales of vodka were 1.59 billion liters in 2011, which is roughly the same level as during the Soviet era.[3]

These stark changes constitute a natural experiment for the study of habit formation. While Figure 1 shows that these changes altered the drinking patterns among the entire population, the most significant shift in tastes clearly occurred in younger generations. Males who started consuming alcohol during the Soviet period became accustomed to vodka and still prefer vodka today even after the Soviet Union collapsed. Younger generations, however, who spent their early adulthood in a time with easier access to beer than prevoius generations prefer beer over vodka.

[Figure 2 about here]

This paper provides clear evidence and formal identification of highly persistent habits in the consumption of habit-forming goods, ruling out various alternative explanations typically proposed in the literature to explain differences in consumption behavior across cohorts and over time, such as unobserved taste heterogeneity and stepping-stone effects. The persistence of consumption habits in turn has important implications for various fields in applied microeconomics such as health economics and consumer demand; see e.g. Becker and Murphy (1988), Becker et al (1994), Chaloupka (1991). Interpreted more broadly, habit formation has also been successfully applied in asset pricing and macroeconomics to explain several empirical puzzles within a unified framework.[4] However, most of this literature lacks convincing identification of habits in general and of the persistence of such habits in particular. The main difficulty is to statistically distinguish between persistent habits and unobserved taste heterogeneity. In this paper we exploit this natural experiment to identify such deep habits.

Although the literature on habit formation in general and rational addiction in particular does emphasize the importance of habits (see e.g. Becker and Murphy (1988) and Cook and Moore (2000)), there is little empirical evidence on how longstanding the state-dependence resulting from the initial choice of habit-forming goods might be. Our results provide some new answers by showing that state-dependence can act over a very long horizon. Moreover, our results also echo the literature on cohort differences in beliefs and preferences, and on preferences for redistribution and state intervention in former communist countries.[5]

---

[3]In the final twenty years of the USSR, 1970-1991, average annual sales of vodka were 1.62 billion liters, and annual sales of beer 3.02 billion liters. In terms of pure alcohol, these numbers correspond to 0.65 billion liters for vodka and only 0.15 billion for beer. In 2001, annual sales of vodka and beer in terms of pure alcohol these were 0.65 and 0.54 billion, respectively. We do not discuss market values here because there was 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 usually not the most significant factor.

[4]Examples of this literature include Eichenbaum et al (1988), Heien and Durham (1991), and Dynan (2000).

[5]See Guisio, Sapienza, and Zingales (2004, 2008), Alesina and Fuchs-Schuendeln (2007), and Malmendier and Nagel (2011) for examples of the former, and Denisova, Euler and Zhuravskaya (2010) for a discussion of the latter.

This research suggests that the cultural and political environment in which an individual grows up affects his preferences over an entire lifetime. In addition to identifying cohort differences in preferences in the data we also provide a model for the mechanism leading to the observed consumption behavior: individuals born with the same preferences but exposed to different initial conditions will form habits toward very different goods.

Our paper also ties to the literature on inter-temporal substitution between different kinds of habit-forming goods. Both the economic literature on addiction and current policy debates on the legalization of marijuana in several U.S. states as well as the taxation of beer in Russia (and an older debate on taxing alcoholic beverages in Scandinavian countries) raise several important questions regarding substitution patterns between light and hard alcoholic beverages and drugs. On the one hand, light alcoholic beverages or light drugs might serve as safer substitutes for harder drinks or drugs, thereby preventing people from consuming much unhealthier hard substances. Consumption of light alcoholic beverages or light drugs at younger ages may form habits for these goods and thus prevent a person from consuming harder substances in the future; see e.g. Becker and Murphy (1988), Cook and Moor (1995), and Williams (2005). On the other hand, it is possible that light alcoholic beverages or drugs might serve as a "stepping-stone" towards harder substances and thus might have negative long-run consequences for public health; see e.g. Mills and Noyes (1984), Van Ours (2003), and Deza (2012).[6] Although habit formation, stepping-stone effects, and contemporary substitution act simultaneously and are well known and widely studied, the current literature on addiction nonetheless lacks a joint discussion of these important points. In particular, there are few attempts to analyze which of these forces will prevail in the long run, and to quantify the cumulative long-run effects of regulation and taxation of light alcoholic beverages on public health and welfare. In this paper we perform such an analysis.

In this paper we analyze the interaction of such different consumption habits. We conclude that beer is indeed a substitute for vodka consumption as we find a significant positive cross-price elasticity. We also find little evidence for beer being a stepping-stone for hard alcohol. Instead, beer consumption during early adulthood forms habits for future beer consumption. Moreover, drinking beer at earlier ages results not only in higher beer consumption later in life, but also in lower consumption of hard drinks like vodka compared to both individuals who started to consume vodka early in life and even compared to abstainers. Finally, we show that drinking vodka early in life also forms habits for future vodka consumption showing that the persistence of the habit formation is qualitatively symmetric in both goods.

In order to study the implications of our results, we simulate the effects of different policies on mortality rates and welfare in a model that accounts for persistent habits. We find that even under the current set of policies and prevailing levels of relative prices of alcoholic beverages the mortality of Russian males will

[6] The more recent literature finds modest stepping-stone effects for marijuana and alcohol consumption towards harder drugs, although to the best of our knowledge we are the first to analyze stepping-stone effects within alcoholic beverages alone.

decrease by one-fifth within ten years. This will happen simply because new generations will be more accustomed to beer and will replace older generations with strong habits for vodka. Simulating the effects of different counter-factual policies on the consumption of different kinds of alcoholic beverages and on the hazard of death, we find that beer is a healthier drink compared to hard alcoholic beverages. Only the consumption of hard beverages but not of beer affects an individual's hazard of death. The most effective policy to decrease mortality rates is to tax vodka consumption and in turn subsidizing beer. . For example, a 50% subsidy of beer consumption payed for with a 30% tax on vodka will decrease male mortality by one-sixth in four years without decreasing consumer welfare. The long-run effect of such a policy is even higher – such a policy would decrease mortality rates by one-third over ten years. Moreover, we find that taxing beer alone will not decrease mortality rates but instead will decrease consumer welfare. In fact, subsidizing beer consumption will decrease mortality and will result in an increase in welfare. Taxing beer alone will have severe long-run consequences, creating a new generation that is accustomed to vodka and is therefore subject to much higher health risks in the future.

The paper is organized as follows. The next section describes our data and the variables employed in our analysis. Section 3 presents a dynamic model of habit formation. Section 4 estimates cohort differences in alcohol consumption and the persistence of habits independent of the model. Section 5combines the model and the estimated habit formation process to analyze the effect of alcohol consumption on the hazard of death and welfare, while section 6 uses the model to simulate counter-factual policy experiments. We conclude our analysis in section 7.

2 Data

We use data from the Russian Longitudinal Monitoring Survey (RLMS).[7] The RLMS is a nationally-representative annual survey panel that covers more than 4,000 households starting from 1992, which corresponds to about 9,000 individual respondents. We use rounds 5 through 20 of the RLMS covering the period from 1994 to 2011, but not including 1997 and 1999 when the survey was not conducted.[8] The data cover 33 regions including 31 oblasts (Russian republics), including two that are Muslim, plus Moscow and St. Petersburg. 75% of respondents live in an urban area and 43% are male. The percentage of male respondents decreases with age, from 49% for ages 13-20 to only 36% for ages above 50. The survey covers individuals starting at age 13 and has a low attrition rate due to low levels of labor mobility in Russia; see Andrienko and Guriev (2004) for more detail. Interview completion exceeds 84%, being lowest

[7]This survey is conducted by the Carolina Population Center at the University of Carolina at Chapel Hill and the High School of Economics in Moscow, and is publicly available from their website at http://www.cpc.unc.edu/projects/rlms-hse.

[8]We do not use data from rounds earlier than 5 because they were conducted by another institution, have a different methodology, and are generally considered to be of much lower quality.

in Moscow and St. Petersbug with 60% and highest in Western Siberia with 92%. The RLMS team provides a detailed analysis of attrition and does not finds any significant effect from it.[9]

[Table 1 about here]

Table 1 summarizes demographic characteristics as well as various measures of alcohol consumption for our population of interest consisting of all males between ages 16 and 65. Our primary measure of alcohol consumption are the shares of beer and vodka consumption in total alcohol intake, calculated in milliliters of pure alcohol.[10] Vodka and beer are the most popular alcohol drinks among Russian males, with an average share of total alcohol consumption across all years of 54% for vodka and 27% for beer, respectively. The share of beer for the average person increases and the share of vodka decreases during the time span of the survey. In 1994, the average share of vodka was 72.5% while beer had only a share of 10.6%. By 2009 these shares were already 46.7% and 39.5%, respectively.

[Figure 3 about here]

3 Suggestive Evidence of Persistent Consumption Habits

The aggregate trends in Figure 2 mask substantial heterogeneity in the changes of drinking behavior across the age distribution as shown in Figure 3a. Older males still overwhelmingly prefer vodka, whereas beer is the drink of choice for the younger generations. The share of beer consumption drops from 56% at age 18 to only 11% at age 65, while the share of vodka increases from 28% at age 18 to 61% at age 65.

This remarkable age profile can potentially be driven by within or between consumer variation, and both can be consistent with the aggregate time-series displayed in Figure 1. Within consumer variation implies that the shape of the age profile of alcohol shares over the life-cycle would look similar across cohorts.

[9] See http://www.cpc.unc.edu/projects/rlms-hse/project/samprep.

[10] To construct the variables we use the amount of various beverages consumed during the last month. We assume that beer contains 5% of pure alcohol and vodka contains 40% of pure alcohol, based on recommendations from the National Institutes of Health (NIH); see e.g. http://pubs.niaaa.nih.gov/publications/arh27-1/18-29.htm. 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 the 2000; see e.g. e.g. Nemstov (2000) and Bhattacharia et al AEJ 2013. We assume a constant share of 5% both for simplicity and to be conservative with respect to the growth rate of beer sales relative to vodka sales measured in pure alcohol shown in Figure 2. The assumption of a constant share of alcohol content in beer does not substantially change our results.

At the same time we should see the initial share of beer relative to vodka would from one cohort to the next, for instance due to relative price or income effects, giving raise to the aggregate time-series in Figure 1. That is, the intercept of the age profile of beer shares of younger cohorts would be higher than that of older cohorts, and vice versa for vodka shares. However, the overall shape of the profiles would look similar across cohorts, with beer serving as the 'stepping-stone' for the switch to hard alcohol later in life. In the case of between consumer variation, different cohorts would have relatively flat alcohol life-cycle profiles with very persistent drinking habits. The aggregate trend in this case results from a gradual increase of these persistent shares from one cohort to the next.

In Figure 3b we visually assess the relative contribution of those two forces by showing the average drinking patterns after taking out individual means.[11] Explaining the aggregate trend with substantial within-consumer heterogeneity would imply that this de-meaned consumption profile should retain a significant slope, positive for vodka consumption and negative for beer. On the other hand, if the aggregate trend is driven by changes in persistent habits across cohorts, then these profiles should be flat. Figure 3b strongly supports the hypothesis that these aggregate trends are mainly driven by changes in persistent habits between cohorts, and there is little evidence for much change within cohorts over time. In fact, the slope of the de-meaned profiles even have the opposite sign than the general age profile in Figure 3b. This implies that consumers across all cohorts increase their consumption of beer and decrease that of vodka over their life-cycles. However, these changes over time for a given consumer are very small compared to the large changes in habits between consumers across cohorts.

4 A Model of Persistent Consumption Habits

Most of the empirical habit literature focuses on habit formation in the short run evidence habits. Relatively short expenditure panel data as well as the absence of large consumption shocks in most other countries prevent researchers from tracking changes in consumption behavior at the micro level over longer periods. Short-run studies, however, do not allow us to answer several questions of importance both for economists and for policy makers. For instance, what are the long-run effects of current regulations and taxation of habit-forming goods such as alcohol on welfare and life-expectancy? Will an increase in the price of hard alcohol force consumers to switch from vodka to healthier drinks, or do other equilibria with different levels of consumption exist?

The simple model derived in this section provides several key insights that help us tackle these questions. First, depending on initial conditions, different groups of individuals with identical preferences can end up consuming either vodka or beer. Second, policies aimed at increasing the relative price of one good may not induce everybody or even many to avoid consuming this good.

[11] To construct Figure 3b we only use individuals with more than one observation.

Instead, people who are accustomed to this particular good will still prefer it even after the policy change due to the stock of habits they already accumulated. Third, initial consumption choices can affect the pattern of consumption over an entire lifetime, and future changes in prices may have no effect on a individual's consumption behavior. The latter point in particular implies that policies aimed at influencing the initial choices of younger generations can have consequences over the entire lifetimes of these young people – intended or otherwise.

4.1 The Model

The model illustrates that in a situation wherein people consume two addictive goods, several steady-state consumption patterns are possible, both with high or low levels of vodka consumption. A person will end up conforming to a steady state depending solely on his or her initial consumption pattern.

For simplicity we assume that consumers spend all their budget on two addictive 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, and that income does not change over time.

Utility from drinking vodka and beer depends on the current consumption of vodka v_t and beer b_t as well as on the corresponding stocks of habit H_t^v and H_t^b, $u(v_t, b_t, H_t^v, H_t^b)$. The utility function has properties that are common in the literature, specifically that $u_g > 0$, $u_{gg} < 0$, $u_{bb}(.) < 0$, $u_{H_g H_g} < 0$, 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, the habit stocks evolve as

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

The budget constraint is

$$p_{v_t} v_t + b_t = y_t \tag{2}$$

and we also require that $u_g \to \infty\ as\ g \to 0$ in order to guarantee an interior solution.

The first-order condition of this optimization problem is

$$u_v(v_t, y_t - p_{v_t} v_t, H_t^v, H_t^b) - p_{v_t} u_b(v_t, y_t - p_{v_t} v_t, H_t^v, H_t^b) = 0. \tag{3}$$

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 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

$$H_t^g = H^g = [\delta/(1-\delta)]\, g. \tag{4}$$

The first-order condition in the steady state can then be rewritten as

$$\begin{aligned} u_v(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) - \\ p_v u_b(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) = \ 0\,, \end{aligned} \tag{5}$$

which is a non-monotonic in steady-state consumption v. Depending on the parametrization of the utility function u this equation may have a different number of solutions. As Figure 4 illustrates, for certain parametrizations there is a unique solution but for many other parametrizations several steady states exist, up to a continuum of solutions.[12] In a situation with several equilibria, the steady state in which a person ends up depends on the initial conditions. If this person initially consumes primarily vodka then he will also prefer vodka in steady state.

[Figure 4 about here]

5 Identifying Persistent Consumption Habits

In this section we analyze whether patterns of alcohol consumption differ among cohorts, and we also check for the presence of long-run and short-run state-dependence. Finally, we more formally test our conjecture from Section 2 that the change in aggregate alcohol consumption is mainly driven by cohort differences against various alternative hypotheses.

To test for the presence of cohort effects, we estimate the following regression of the share of alcohol S_{it}^g consumed by individual i in year t of alcohol of type $g \in \{\text{b: beer}, \text{v: vodka}\}$ by OLS,

$$S_{it}^g = \text{cohort}_i + \lambda \cdot \log(y_{it}) + \gamma' x_{it} + \alpha_t + \alpha_r + \epsilon_{it}. \tag{6}$$

cohort_i is a cohort fixed effect, and α_t and α_r are time and region fixed effects, respectively, controlling for relative price effects while simultaneously also holding log-income constant. x_{it} includes a standard set of demographic controls such as age, personal health status, weight, education, and marital status. Table 2 shows that even after controlling for time and regional fixed effects as well as demographic characteristics, younger generations still tend to consume more beer and less vodka. Column 3 shows that the average share of beer in total alcohol intake is 30 percentage points (pp) higher for males born in the 1990s than for those born before the 1970s.[13] Even those born in the 1980s have on average a 18 pp higher share of beer consumption than those born before the 1970s. Those born in the 1970s in turn have a 5 pp higher share of beer consumption than those born earlier, which is the reference group.

[Table 2 about here]

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

[13] The results for vodka shares are similar – although with opposite sign – since the sum of the shares of beer and vodka is close to one. In the following we therefore do not discuss the results for beer and vodka separately unless they provide additional insight, but we refer the interested reader to the online appendix.

Our hypothesis is that although people have similar tastes regarding alcoholic beverages, they differ in their initial choice of which habit-forming good to consume. This initial choice combined with the strong persistence of such habits explains the patterns observed in Figures 1 to 3. However, there are two main alternative hypotheses. First, individuals born in different time periods might have different preferences for certain types of alcohol because of cultural or other differences, and not because of different initial choices and subsequent habit formation. Second, these observed cohort differences may be the result of a stepping-stone effect. Young generations might start consuming beer or other light drinks, and then eventually switch to harder drinks later in life.

We begin our discussion by showing that drinking patterns do in fact demonstrate persistent habits. To test for such longstanding state dependence, we estimate the following dynamic extension of equation (6) first by OLS and then by IV

$$S^g_{it} = \text{cohort}_i + \rho_k \cdot S^g_{i,t-k} + \lambda \cdot \log(y_{it}) + \gamma' x_{it} + \alpha_t + \alpha_r + \epsilon_{it}. \quad (7)$$

In addition to the specification in equation (6) we include the lagged share of alcohol consumption, $S^g_{i,t-k}$. We chose two lag specifications, $k = 7$ to measure the long-run effects and $k = 1$ to capture short-run dynamics.[14] To deal with the effect of potentially auto-correlated unobserved taste shocks we instrument lagged shares of alcohol consumption with the year of birth as is standard in the literature. This specification estimates the effect of habits under the assumption that individuals born in different periods have the same preferences and differ only in the initial level of consumption, holding fixed income, relative prices and demographic characteristics. We discuss this assumption in more detail below.

[Table 3 about here]

Table 3 illustrates the results of these regressions. Both lagged shares significantly affect personal decisions regarding which good to drink. Thus, those who chose to drink only beer seven years ago have on average a 19 pp – or half a standard deviation – higher share of beer consumption, while those who drank only vodka seven years ago have on average a 16 pp higher share of vodka consumption.

The first alternative explanation for the observed heterogeneity is that individuals born in different times grew up in different cultural environments, and might therefore have different preferences regarding hard and light drinks. In the following we test of our hypothesis against this alternative using different sources of variation.

[14]The choice of the long-run period t=7 is driven by data availability, and we use the average of $S_{i,t-7}$ and $S_{i,t-8}$ to estimate the long-run effect instead of $S_{i,t-7}$ for three main reasons. First, the data are noisy and averaging helps reducing the measurement error. Second, some of the households we do not observe in interview 7 but we do observe them again in interview 8. Finally, as mentioned above, the survey was not conducted in 1997 and 1999. For cases where a data point is not available we choose the nearest available interview $I \geq 7$ for which we have data on the shares.

5.1 Evidence from the Soviet Youth

Our first natural experiment to identify long-run consumption habits is the large change in the Russian alcohol market that occurred in the wake of the collapse of the Soviet Union. Focusing on the relatively short period of time when the beer industry experienced rapid growth, we study the consumption behavior of individuals turned 18 years old that during this period, which is the minimum legal drinking age in Russia.[15] Since culture and institutions change only slowly (e.g. Roland (2004)), individuals born within a narrow interval face very similar initial conditions. Figure 5a illustrates the choice of timing in our analysis. We select males who were born within a narrow range around year 1980, i.e. those who turned 18 years old around year 1998, which is the midpoint of the rapid expansion of the beer industry. We then let the range of this interval successively increase from 5 years up to 21 years.

[Table 4 and Figure 5 about here]

The regression results summarized Table 4 and illustrated in Figure 5b show a strong effect of initial consumption choices on future consumption across for all sub-samples. The OLS specifications show that an increase in the 7-year lag of share of beer consumption by 10 pp increases the current share of beer consumption by 1.9 pp.[16] The IV estimates show an even larger effect, with an increase in the lagged share of beer consumption by 10 pp causing an increase in the share of current beer consumption by 3.6 pp. Although the statistical significance of our results naturally decreases with the contracting sample size, most of results are nonetheless statistically significant. The OLS estimates are statistically significant even for the smallest sample chosen, i.e. those born between 1978 and 1982, and the IV estimates are statistically significant for those born between 1976 and 1984 and for all larger sample sizes. Similar results obtain for the consumption of vodka and are provided in the appendix.

Figure 5b shows the coefficients from regressing the shares of beer on the year of birth for males who were born in intervals of increasing length centered around year 1980, i.e. for those who turned 18 in the intervals 1987-1989, 1986-1990 etc. It shows a positive correlation between the year of birth and the share of beer even for very narrow birth year intervals; the appendix documents a similar negative correlation for vodka. Moreover, the magnitude of the coefficients increases with shrinking intervals since we are selecting males that are more and more likely to have formed their consumption habits during the rapid expansion of the beer market.

[15] Since there is no discontinuity implied by the legal drinking age – both because of enforceability and because one cannot be forced to start consuming alcohol at 18 – and because habits do not necessarily form within a single year we cannot use a regression discontinuity. However, our identification approach closely mimics such a framework.

[16] For this analysis we only use data from 2001 to 2011 because we only have at least 8 years of data per individual starting in 2001.

[Table 5 and Figure 6 about here]

Second, we analyze whether the year of birth is correlated with the alcohol shares for age cohorts defined in a rolling window of a constant range of 10 years. Figure 6a again illustrates the choice of timing in our analysis. We choose a 10-year window in order to maintain sufficient power. Figure 6b, which plots the coefficients from these regressions, shows that the year of birth correlates with the alcohol shares only for those born after year 1975, i.e. only for those who who turned 18 after 1993 when the beer market took off.

Table 5a shows results from regressing the shares of goods on the year of birth, controlling for both age and time fixed effects. After controlling for age and time fixed effects, the correlation is much higher for those born after year 1980, i.e. for those who formed their consumption habits during the fundamental transformation of the alcohol markets.

As a final step we also explore regional variation in sales of different types of alcohol. Table 5b shows the results of regressing the shares of regional sales of alcohol fraction of regional sales of beer to alcohol, lagged by 7 years. Again, we find that the share of beer is positively correlated and that of vodka negatively with this lagged sales ratio, and the correlation is much higher for younger generation.

5.2 Evidence from Gorbachev's Anti-Alcohol Campaign

Our second natural experiment to identify consumption habits uses the so-called Gorbachev anti-alcohol campaign to test for long-run effects on preference. In 1985 Michael Gorbachev introduced an anti-alcohol campaign that was designed to fight widespread alcoholism in the Soviet Union. Prices of vodka, beer and wine were raised, their sales were heavily restricted, and many additional regulations were put in place aimed at further curbing alcohol consumption. The campaign officially ended in 1988, although research shows that high alcohol prices and sales restrictions continued until the collapse of Soviet Union at the end of 1991.[17]

[Figure 7 about here]

The effect of Gorbachev's anti-alcohol campaign on official sales of alcohol was dramatic. Official sales of beer dropped by 30%, from 177 millions liters in 1984 to 125 millions liters in 1987.[18] Official sales of vodka even dropped by 60% from 78 millions liters in 1984 to 32 millions liters in 1987, and wine sales experienced the most dramatic drop from 31 millions liters in 1985 down to only 12 millions liters in 1990. However, as shown in Figure 7a, the drop in official sales of vodka was compensated by increased production of samogon, ain illegal low-quality type of vodka home-produced by households. As a result, the effect

[17] See for example White (1996), Nemtsov (2000), and Bhattacharya etal. (2013).

[18] The volume of alcohol sales is again measured in terms of pure alcohol.

of the Gorbachev campaign on total vodka consumption, including samogon was small on average; see e.g. Tremlin (1997), Nemtsov (2000), and Bhattacharya et al (2013) for a discussion of the underlying data and methodology.[19] Indeed, after accounting for samogon production, the estimated volume of total alcohol consumed decreased by only 20% during the Gorbachev anti-alcohol campaign.

Even more important for our identification approach is the fact that the production of samogon was heavily concentrated in rural areas. This is due to several reasons related to the technology used to produce samogon. First, the production of samogon requires space which is limited in urban areas. Second, producing samogon causes a smoke and a strong smell which is at the same time very unpleasant and also easy to detect by neighbors and law enforcement agents, especially in cities. Third, the illegal production of samogon was more strictly enforced and punished in urban areas. As a result, it was much safer to produce samogon in ones own home, which are for Russia are particularly concentrate in rural areas, instead of producing in an apartment building. This geographical pattern of samogon consumption continues to the present even though the total samogon production has decreased dramatically since 1992. For instance, according to the RLMS data, males in rural areas drink 5.5 times more samogon, and the share of samogon in total alcohol intake is five times higher in rural areas, 13% for rural areas compared with only 2.4% in urban areas. Thus, although the effect of Gorbachev's anti-alcohol campaign on the average share of beer to total vodka consumption including samogon is modest as shown in Figure 7b, one can expect significant differences in the way the campaign affected habits of rural relative to urban males. Indeed, the fact that rural households still consume 5.5 times more samogon implies that the anti-alcohol campaign decrease the fraction of beer to hard drinks from 8% to 4.5% for those rural households, but only from 17.3% to 17.2% for the urban population, and from 12.8% to 9.3% for the total population.

We test the hypothesis that among people who live in rural area people who start consuming alcohol during peak of samogon production i.e. in years 1988-1992 (i.e. those who was born in 1970-1974) have stronger habits towards vodka than those born before 1970 and after year 1974.[20] We use this natural

---

[19]There are two main approaches used in the literature to estimate samogon consumption during and shortly after the Soviet Union. The first approach uses aggregate sales of sugar that is one of the main ingredients in the production of samogon; see e.g. Nemtsov (1998). This approach gives reliable estimates until the year 1986 when the production of sugar was rationed. The second approach uses data on violent and accidental deaths and deaths with unclear causes; Nemtsov (2000). For these death events there exist measures of alcohol concentration in the blood of the victim which can be used to estimate aggregate alcohol consumption. This approach gives similar estimates of samogon production as the first approach, but it cannot distinguish between the consumption of samogon and other illegal alcohol. While samogon was by far the main source of illegal alcohol in the Soviet Union, much of the illegal alcohol consumed since 1992, i.e. after the collapse of Soviet Union, comes from illegal imports as well as illegal production of unregistered alcohol by firms as a form of tax evasion.

[20]According to different expert estimates samogon production increased rapidly in the second half of 1980s; e.g Tremlin (1997), Nemtsov (2000), Bhattacharya et al (2013) and our own estimates based on the RLMS. After the collapse of the Soviet Union it has decreased rapidly because of the liberalization of alcohol market and the sharp decrease in the price and

experiment to test two hypothesis. First, we test whether males who lived in rural areas and turned 18 during the peak of the samogon production, i.e. in years 1988-1992 and hence were born between 1970 and 1974, have stronger habits toward vodka consumption that those living also in rural areas but that turned 18 either before or after this period. We implicitly also exploit the fact that labor mobility is very low in Russia compared to most other countries. Hence, the chance that the current residence of a survey respondent in our sample also identifies his location during the anti-alcohol campaign is very high. Similar to the identification in the first natural experiment in section 4.2 we rely on time-series variation in these tests.

Second, we test whether the same treatment group as above, i.e. males turning 18 during the campaign and living in rural areas, have formed stronger habits toward vodka consumption than their cohort peer who grew up in urban areas, and we test whether this difference in habits is strong that the same difference for cohort born before or after the period 1970-1974. In other words, we perform a difference-in-difference analysis exploiting the geographical incidence of the temporary shock caused by Gorbachev's anti-alcohol campaign on the habit formation of young men. To test this hypothesis we estimate the following regression

$$\begin{aligned} S_{it}^{v} = \; & \beta_1 \cdot I(\text{born before 1970})_i \times I(\text{born in city})_i + \beta_2 \cdot I(\text{born before 1970})_i + \\ & \beta_3 \cdot I(\text{born after 1974})_i \times I(\text{born in city})_i + \beta_4 \cdot I(\text{born after 1974})_i + \\ & \beta_5 \cdot I(\text{born in city})_i + \lambda \cdot \log(y_{it}) + \gamma' x_{it} + \alpha_t + \alpha_r + \epsilon_{it} \end{aligned} \tag{8}$$

Standard errors are clustered at the individual level, and the regressions are estimated on the sub-sample from 2005 to 2011 when the share of samogon consumption becomes low for all subgroups of the population.[21]

[Table 6 about here]

Table 6 present results. First, results show that among males who were born in rural areas those who born in 1970-1974 have 4.2 and 9.8 percentage points higher share of vodka compare to those who were born before 1970 and after 1974 correspondingly. Second, the gap in patterns of consumption between rural and urban population is greater for those who were born in 1970-1974: it is 6.4% higher that that for those who were born before 1970 and 5.1% higher than that for those who were born after 1974. Finally the effect of campaign on average among all subgroups of population is modest: those who born in 1970-1974 have 1.3 and 7.4 percentage points higher share of vodka compare to those who were born before 1970 and after 1974 correspondingly.

Second, we provide the following simulation experiment. We look on people from different 5-years periods of years of birth and compare first, how the share of vodka for people who were born in rural areas in the period higher than

increased availability of vodka.

[21] We choose 2005 as the starting point of the sub-sample because starting from this year the share of samogon consumption becomes stable and relatively low; see Figure 7. We provide results for regression using years 2001 to 2011 in the appendix.

that for those who were born in rural areas in other time. Second, we check whether the gap in patterns of consumption between rural and urban population is greater for those who were born in given period compare to that of those who were born in other time. The Figure 8 below illustrates the values and CI of coefficients β_1 and β_2 in regression below.

$$\begin{aligned} S^v_{it} = \quad & \beta_1 \cdot I(\text{born in five year peroid})_i \times I(\text{born in city})_i + \\ & \beta_2 \cdot I(\text{born in five year peroid})_i + \beta_3 \cdot I(\text{born in city})_i + \\ & \lambda \cdot \log(y_{it}) + \gamma' x_{it} + \alpha_t + \alpha_r + \epsilon_{it} \end{aligned} \tag{9}$$

We estimate regressions on two samples. First we look on look on all males; second we look only on males who were born in certain 20-years birth period interval: for given year X we compare those who born in years [X-2;X+2] to those who were born in years [X-10;X-2] or in years [X+2;X+10]. For all sample β_1 is negative and statistically significant only for those who born in 1969-1975; β_2 is positive and statistically significant only for those who born in years 1967-1976. For 20 years birth period sample β_1 is negative and statistically significant only for those who born in 1969-1974; β_2 is positive and statically significant only for those who born in 1969-1973.

[Figure 8 about here]

5.3 Testing for Stepping-Stone Effects in Alcohol

The second main alternative explanation for cohort difference put forward in the literature is a stepping-stone or "gateway" effect of consuming light drugs for the consumption of harder drugs later on. In the case of alcohol this means that beer might serve as a stepping-stone earlier in life for the consumption of harder alcoholic substances such as vodka later in life. According to this theory people would start out with beer but eventually switch to harder drinks. In this case, the observed cohort differences would just be an effect of aging.

This stepping-stone effect is widely studied in health economics and several studies have analyzed this theory in the context of different types of drugs and tested it against alternative explanations, in particular against unobserved individual heterogeneity in preferences. For instance, Mills and Noyes (1984) and Deza (2012) find evidence for a modest stepping-stone effect of marijuana and alcohol in general for the consumption of harder drugs later on. Similarly, Beenstock and Rahav (2002) find a stepping-stone effect in cigarette consumption leading to increase in the probability of smoking marijuana later on. Van Ours (2003) finds that unobserved individual heterogeneity and stepping-stone effects can explain many patterns of drug consumption.

However, to the best of our knowledge this study is the first to test for a stepping-stone effect of beer towards harder alcoholic beverages. We find strong evidence against beer being a gateway drug for hard alcohol. On the contrary, the habit formation we identified in the previous sections implies that beer consumption earlier in life forms consumption habits and therefore increased the probability of drinking beer instead of vodka later on, and a similar habit formation effect exists for vodka. To see this more clearly we can look at Figure 3b above which shows that for any particular person there is no evidence for an increase in the share of vodka consumption over the life-cycle of that person. The graph instead highlights that males at all ages tend to consume *less* vodka and more beer as they get older. The same conclusion is reached from looking at the age similarly flat age profiles for different cohorts shown in Figure 1.

Moreover, the simulations of a multinomial choice model in Section 6 below shows that both habit as well as cross-price substitution effects outweigh any stepping-stone effect for beer or vodka. A a decrease in the price of beer results in the substitution of beer for vodka, and this effect grows over time; see Figure 8 and Table 12 and Section 6 below.

[Table 7 about here]

Finally, Table 7 reports the probabilities of drinking vodka at age 25 and older, conditional on drinking different alcoholic substances as a teenager. The results show that those who drink only beer as teenagers have smaller chances of drinking vodka later in life compared to both those who consume vodka in their teens and even to those teenagers who abstain altogether. Quantitatively we find that the probability that a person drinks vodka after age 25 if he was an abstainer as a teenager is 0.66, whereas the probability of consuming vodka after drinking beer as a teenager is only 0.57. The probability of later drinking vodka for those who consumed vodka as a teenager is 0.81.

6 Implications for Life Expectancy and Welfare

Russian males are notorious for their hard drinking. The most notable example of the severe consequences of alcohol consumption is the male mortality crisis. Male life expectancy in Russia is only 60 years, which is eight years below the average in the remaining BRIC countries,[22] five years below the world average, and below even the life expectancy in Bangladesh, Yemen, and North Korea. High alcohol consumption is generally believed to be the main cause of this.[23] Approximately one-third of all deaths in Russia are related to alcohol consumption; Nemtsov (2002). Most of the burden lies on males of working age – more than half of all deaths in working-age men are caused by hazardous drinking; Leon et al. (2007), Zaridze et al. (2009).

In this section, we estimate the hazard of death as a function of not only overall alcohol consumption, but specifically the consumption of hard drinks

[22]The remaining BRIC countries besides Russia are Brazil, India and China.

[23]See for example Nemtsov (2002), Brainerd and Cutler (2005), Leon et al. (2007), Denisova(2010), Treisman (2010), Bhattacharya et al. (2011), and Yakovlev (2012).

(vodka) or beer. Beer is generally agreed to be a much safer drink than vodka, and presumably has less of an effect on mortality rates.

Large-scale studies in demographics literature (see for example Zaridze et al. (2009) study of 48 557 adult deaths) support the observation that the main cause of male death in Russia is so-called dose-related excess: a hazardous event occurring when the amount of pure alcohol consumed by a person is too high.[24] Table 11 shows that preferences towards beer are associated with lower level of alcohol intake whereas preferences towards hard alcohol drinks positively correlated with level of pure alcohol intake.[25]

Besides, most alcohol-related deaths in Russia are not due to diseases that result from long-time alcohol consumption (such as cirrhosis), but rather to (probably occasional) one-time hazardous drinking. First, 6% of all deaths of Russian males are caused by alcohol poisoning. The main cause of poisoning is not poor quality of alcohol, but rather drinking so much alcohol that the amount in the blood causes the heart to stop (see Zaridze et al 2009, Djoussé and Gaziano 2008). Thus, it takes binging with vodka only once to result in death. In contrast, beer consumption is safer – one must consume eight times more beer to produce the same amount of alcohol in the blood.

Second, another 35% of deaths are due to external causes – vehicular and other accidents, or homicides, for example – that occur largely under the effects of alcohol intoxication. Again, even with moderate average vodka consumption, it is enough to binge only once and get into an accident. However, beer consumption does not result in an increase of death hazard, and people who drink beer have a smaller chance of death compared to those who drink vodka, and to those who do not drink or drink beverages other than beer or vodka. The number of non-drinkers in Russia is very low (less than 10% of males reported that they did not drink in the previous month over three consecutive years), and there is possible negative selection to non-drinkers – non-drinkers have smaller incomes and lower levels of education, do not perform more physical training, and do not have lower rates of disease.

[Table 8 about here]

Table 8a estimates the effect of alcohol consumption on hazard of death for the following hazard specification:

$$\lambda(t,\, z_i) = exp(\phi' z_i)\, \lambda_0(t) \tag{10}$$

where$\lambda_0(t)$ is the baseline hazard, common for all units of population. We use a semi-parametric Cox specification of baseline hazard. The set of explanatory

[24] Zaridze et al. 2009 studied the death events of 48,557 residents aged 15-54 in three typical Russian cities. They found that alcohol-associated excess accounted for 59% of the deaths of males, and 33% of the deaths of females. This study also indicated that 8% of death are directly due to alcohol poisoning, and 37% are due to accidents and violence that primarily occurred during alcohol intoxication. See also Leon et all 2007, Denisova, 2012, Treisman 2010, and Yakovlev, 2012.

[25] The specification of regressions shown in Table 9a as follows:
$log(alcohol\, intake_{it}) = \beta_0 + \beta_1 share\, of\, beer_{it} + \beta_2 share\, of\, hard\, drinks_{it} + e_{it}$

variables includes alcohol consumption variables, log of family income, health status, weight, age, employment status, and educational level.

Table 8a shows that the probability of death is strongly positively-related with the consumption of vodka. As such, drinking vodka increases the hazard of death twice ($= exp(0.68)$). However, the hazard of death is high even for males who reported only moderate average monthly vodka consumption. This is because, even with moderate average consumption, a person can still die as the result of one-time hazardous binge drinking.

6.1 Persistent Habits and the Russian Mortality Crises

This subsection offers a speculative analysis of the relationship between the "Russian mortality crisis" and the consumption of vodka and beer. The mortality crisis refers to surge in male mortality in post-Soviet ' liberalization of the alcohol market (in 1992) is widely cited as the main cause of this change (see Figure A1 in the online appendix).

Table A3 in the online appendix shows yearXcountry-level regressions of mortality rates on sales of beer and vodka for the period of the Russian mortality crisis. This ("twenty-point") regression finds a positive correlation between sales of vodka and mortality rates, but no effect from beer sales on male mortality. Table A4 in the appendix shows yearXregion-level regressions of mortality rates on regional per-capita retail sales of beer and vodka for the period of 1997 to 2009.[26] Again, the regressions show a positive correlation between male mortality rates and sales of vodka, and again no (or negative) correlation between beer sales and male mortality rates.[27]

7 Simulating Counterfactual Policies

In this section, we simulate the effect of taxation on alcohol consumption, consumer welfare, and mortality rates. To do this, we estimate a dynamic model of consumer choice among the different kinds of alcoholic beverages.

In our models, agents are assumed to be myopic. Consumers have four choices: drink both vodka and beer $(1,1)$, drink only vodka $(1,0)$, drink only beer $(0,1)$, or drink neither beer nor vodka $(0,0)$. Indirect utilities of consumers are assumed to have linear parameterization:

$$\begin{aligned} U(k,\, j) = \quad & \alpha_{kj} + \gamma_{kj}(P_{it}^{beer}/P_{it}^{vodka}) + D_{cohort} + \beta_{vkj} I(vodka)_{i,t-1} \\ & + \beta_{bkj} I(beer)_{i,t-1} + \Gamma' D_{itjk} + \delta_{rkj} + \nu_i + e_{itkj} \end{aligned}$$

[26] The data for these regressions are collected from Rosstat (the Russian statistical agency, www.gks.ru). Regional-level data on retail sales of beer and vodka are available for the period of 1997-2009.

[27] The specifications of regional-level regressions are as follows

$$mortality\,rate_{rt} = \beta_0 + \beta_1 log(sales\,vodka)_{rt} + \beta_2 log(sales\,beer)_{rt} + \beta_3 I(Caucasus)_r + \rho_t + e_{it}$$

We use two specifications: unweighted regression, and regression weighted by regional population size.

Indexes k, $j \in \{0, 1\}$ stand for personal choices. The indirect utility of non-drinking is normalized to zero: $U(0,\, 0) = 0$; and γ is normalized to zero for the (0,1) choice, "drink only beer": $\gamma_{01} = 0$. In our model, we normalize price of vodka to 1. With this normalization, a change in beer price results in a change in P_{beer}/P_{vodka}.[28] In this modelβ_{b10} represents a stepping-stone effect for the choice "drink only vodka". β_{b11} captures both the stepping-stone effect of beer and beer habit formation for the choice "drink both vodka and beer." β_{b01} captures habit formation for the choice "drink only beer." Vodka habit formation effects are captured by β_{v10} andβ_{v11}. D_{it} is a set of demographic characteristics that affect utility. This set of demographic characteristics includes log(family income), health status, age, I(Muslim), I(college degree), and personal body weight. δ_{rkj} stands for (unobservable to the researcher, but observable to the individual) regional-specific factors that affect utility, such as official religion, temperature, and so on. e_{itkj} is a choice-specific utility component that is unobservable to a researcher, but observable to a consumer. We assume that e_{itkj} has logistic distribution.

Finally, ν_i stands for an individual-specific taste for alcohol, unobservable to the researcher but observable for the individual. This term captures unobserved personal heterogeneity in tastes for alcohol consumption that do not vary across time and kinds of alcohol. Further, we provide estimation of two utilities, with and without ($\nu_i = 0$) allowing for unobserved heterogeneity in tastes.

[Table 9 about here]

Estimates of utility parameters are shown in Tables 9a and 9b.[29] Tables 9a and 9b show strong habit-formation effects, but rather small (if at all) stepping-stone effects. In fact, Tables 9a and 9b show positive switching costs for changing patterns of drinking (from drinking only beer to drinking vodka), and a strong effect of habits over the same pattern of consumption. Drinking only beer in a previous period positively affects the utility of drinking only beer now, and negatively affects the utility of drinking vodka (with or without beer). Table 6 also shows that the relative price of beer has a negative effect on consumer utility specific to the choice of drinking beer.

Our first simulation exercise estimates the consequence on male mortality rates in 10 years if the prices of beer and vodka stay at their current levels. We find that current price policy will result in a decrease in mortality rates from 1.41% to 1.16% (that is, a decrease of about one-fifth). This decrease in mortality is driven by a new generation that prefers beer replacing an older generation that drinks vodka.

[28] See also Appendix for estimation of elasticity using hedonic regressions.

[29] Table 9b shows estimation results for following parameterization of indirect utility

$$\begin{aligned} U(k,j) = \quad & \alpha_{kj} + \gamma_{kj} P_{beer,it}/P_{vodka,it} + \beta_{vkj} I(drink\ only\ vodka)_{it-1} + \beta_{bkj} I(drink\ only\ beer)_{it-1} \\ & + \beta I(drink\ vodka\&beer)_{it-1} + \Gamma' D_{itjk} + \delta_{rkj} + \nu_i + e_{itkj} \end{aligned}$$

[Figure 8 and Table 10 about here]

Our next exercise simulates the effects of different government policies on consumer drinking patterns, consumer welfare, and mortality rates. We consider two policies: taxation and subsidization of beer consumption. Figure 8 and Table 10 demonstrate the effect of these policies over a four-year period.

[Table 11 about here]

Table 11 shows the effect of taxing (and subsidizing) beer, specifically the effect of a one-half decrease or two-times increase in the price of beer. The simulation shows that taxing only beer will not decrease mortality, and will result in a decrease in consumer surplus. Thus, the taxation of beer results in a decrease in the share of beer drinkers from 46% to 37%, and a corresponding increase in the share of vodka drinkers from 53 to 55%, and of those who drink neither beer nor vodka from 31% to 41%. This policy also results in a 21% decrease in consumer welfare, and an increase in mortality rates from 1.4% to 1.68%.[30]

Figure 10 shows the simulation results for simultaneously subsiding beer and raising taxes on vodka (specifically, halving the price of beer, and increasing the price of vodka by 30%). The mortality rate falls from 1.41% to 1.18% in four years. Moreover, in ten years the male mortality rate decreases from 1.41% to 0.95% – by fully one-third.

[Figure 9 about here]

8 Conclusion

Using individual-level data on the alcohol consumption of Russian males, this paper finds evidence of longstanding persistence of habits towards certain types of addictive goods. People who grew up in the USSR became accustomed to vodka (the most popular liquor during the Soviet era) and still prefer it; however, those who reached the age of twenty in the post-Soviet period (when the beer industry significantly expanded) prefer beer.

These findings demonstrate that the effect of policy may be very long-lasting, and so emphasize the importance of policy on young people when they form their habits: depending on the initial choice of alcohol beverage, a person may end up drinking that type of alcohol for the rest of his life. The paper also finds that habits and a substitution effect will outweigh any stepping-stone effect. For example, a decrease in the price of beer will result in decreased vodka consumption, not only in the short run but also for long-run horizons.

[30] See Train (2003) for a description of the estimation of model parameters, choice probabilities, and consumer welfare.

Policy simulations indicate that the government policy toward the substitution of vodka and other hard drinks with safer beer will result in significant reduction to male mortality rates. As a result, a 50% subsidy on beer and 30% tax on vodka will decrease male mortality from 1.41% to 0.95% in 10 years, halving the gap between Russian and western-European mortality rates. This policy will not decrease consumer surplus, and so might have a greater chance of being implemented by a populist government.

References

Alesina, Alberto, and Nicola Fuchs-Schündeln, "Good-bye Lenin (or Not?): The Effect of Communism on People's Preferences," American Economic Review, 97 (2007), 1507-1528.

Andrienko Y., and S. Guriev, 2004, "Determinants of inter-regional mobility in Russia," The Economics of Transition, vol. 12(1), pages 1-27, 03.

Becker, G. and Murphy, K., 1988. "A Theory of Rational Addiction", Journal of Political Economy, 96 (4), 675

Becker, G. S., Grossman, M., & Murphy, K. M., 1994. An empirical analysis of cigarette addiction. American Economic Review, 84(3), 396-418.

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

Bhattacharya, J., C. Gathmann, C., and G. Miller, 2011, "The Gorbachev Anti-Alcohol Campaign and Russia's Mortality Crisis"

Brainerd,E and D. Cutler, 2005. "Autopsy on an Empire: Understanding Mortality in Russia and the Former Soviet Union," Journal of Economic Perspectives, American Economic Association, vol. 19(1), pages 107-130, Winter

Chaloupka, F. 1991. Rational addictive behavior and cigarette smoking. Journal of Political Economy, 99 (4) (August), 722-742.

Cook, Philip J. and Moor, Michael J. 1995, "Habits and heterogeneity in the youthful demand for alcohol", NBER working paper #5152, 1995

Cook, Philip J. and Moore, Michael J. 2000. "Alcohol," Handbook of Health Economics, in: A. J. Culyer & J. P. Newhouse (ed.), Handbook of Health Economics, edition 1, volume 1, chapter 3.

Deza, Monica "Is There a Stepping-Stone Sect in Drug Use? Separating State Dependence from Unobserved Heterogeneity Within and Across Illicit Drugs", Job Market Paper, UC Berkeley, 2012

Denisova, Irina. 2010. "Adult mortality in Russia: a microanalysis", *Economics of Transition*, Vol. 18(2), 2010, 333-363.

Denisova, Irina, M. Eller and E.Zhuravskaya, What do Russians Think about Transition? Economics of Transition, 18(2), 2010, pp. 249–280

Djoussé L, and Gaziano JM, 2008. "Alcohol consumption and heart failure: a systematic review". Curr Atheroscler Rep 10 (2): 117–20

Dynan, Karen E. “Habit Formation in Consumer Preferences: Evidence from Panel”, The American Economic Review, Vol. 90, No. 3 (Jun., 2000), pp. 391-406

Fogarty, James “The Demand for Beer, Wine and Sperirts: A Survey of the literature”, Journal of Economic Surveys, Volume 24, Issue 3, pages 428–478, July 2010

Eichenbaum, Martin S.; Hansen, Lars Peter and Singleton, Kenneth J. "A Time Series Analysis of Representative Agent Models of Consumption and Leisure under Uncertainty." Quarterly Journal of Economics, February 1988, 103(1), pp. 51-78

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2004, “The Role of Social Capital in Financial Development,” American Economic Review, 94 , 526-556.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, 2008, ”Trusting the Stock Market,” Journal of Finance, 63, 2557-2600.

Heckman, J.J “Heterogeneity and state dependence”. Studies in Labor Markets, 1991, 31:91

Keane, M.P. “Modeling heterogeneity and state dependence in consumer choice behavior”, Journal of Business & Economic Statistics, 15(3):310-327, 1997.

Leon, D. A., Saburova L., Tomkins S., Andreev E., Kiryanov N., McKee M., Shkolnikov V., “Hazardous alcohol drinking and premature mortality in Russia: a population based case-control study”, Lancet 369, 2007, pp. 2001–2009.

Leung S. F., and Phelps, C. E. “My kingdom for a drink. . . ?” A review of estimates of the price sensitivity of demand for alcoholic beverages. In: Hilton, M. E. and Bloss, G., eds. Economics and the Prevention of Alcohol-Related Problems. NIAAA Research Monograph No. 25, NIH Pub. No. 93–3513. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism, 1993. pp. 1–32.

Malmendiers, U. and S. Nage, 2011, Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?. Quarterly Journal of Economics, February 2011, vol. 126(1), pp. 373-416

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

Nemtsov, Alexandr, “Alcohol-Related Human Losses in Russia in the 1980s and 1990s.” *Addiction*, 97, pp. 1413-25.

Roland, Gerard, 2004, Understanding Institutional Change: Fast-Moving and Slow-Moving Institutions, Studies in Comparative International Development, 2004.

Train, K., Discrete choice methods with simulation. Cambridge Univ Pr, 2003.

Treisman, Daniel. 2010., "Pricing Death: The Political Economy of Russia's Alcohol Crisis." *Economics of Transition*, Vol. 18, Issue 2, pp. 281-331, April 2010 .

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

Williams, 2005 "Habit formation and college students' demand for alcohol", Health Economics, 14: 119–134, 2005

Yakovlev, E, 2013, "Peers and Alcohol: Evidence from Russia", mimeo

Zaridze, D., P. Brennan, J. Boreham, A. Boroda, R. Karpov, A. Lazarev, I. Konobeevskaya, V. Igitov, T. Terechova, P. Boffetta, R. Peto, "Alcohol and cause-specific mortality in Russia: a retrospective case–control study of 48 557 adult deaths", The Lancet, Volume 373, Issue 9682, 27 June 2009-3 July 2009, Pages 2201-2214.

Appendix

1. Robustness

In addition to these regressions, we provide several robustness tests. We check whether people who grew up in rural areas tend to consume more vodka. As we discussed above, people in many rural areas in Russia, especially during the Soviet era, often use their own equipment to produce a home-made vodka called samogon. As such, a person who grew up in a rural area has a higher chance to start alcohol consumption with samogon, and so to be accustomed to vodka. Table 9a illustrates this point: after controlling for covariates, we see that those born in a village have 3 pp smaller share of beer consumption, and 3 pp higher of vodka. The results of these regressions hold for those who live in cities, and for those who have lived in their current location for at least the past seven years.[31] Moreover, cohort effect on patterns of consumption is stronger for those who born in village: Table 9b shows that, after controlling for covariates, the share of vodka is 10.4 pp higher for those who born before 1980s and grew up in a city, and 13.8 pp higher for those who born before 1980s and grew up in a rural area. The difference between these two groups is statistically significant. Besides, Table 9b shows that cohort effect is stronger for those who grew up in cold regions: these people also tend to have higher share of vodka.[32]

[Table 9 about here]

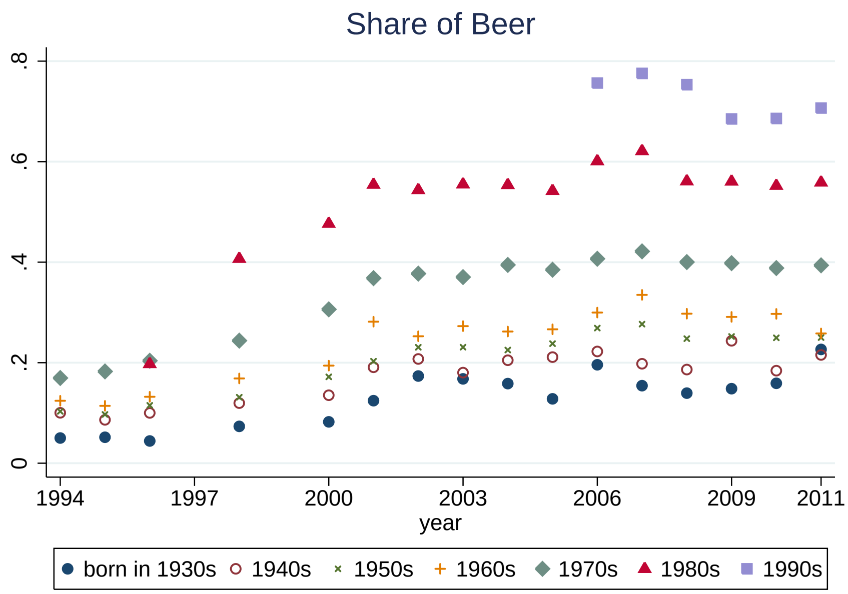
Second, we check whether people who grew up in the vine-making areas of the Soviet Union (Moldova, Ukraine, and the former Caucasus republic) now prefer wine. Table 10 shows that those born in these areas on average have a 3 pp higher share of wine consumption, although the statistical significance of results disappears when we restrict samples to those who have lived in their current location for the past seven years.

[Table 10 about here]

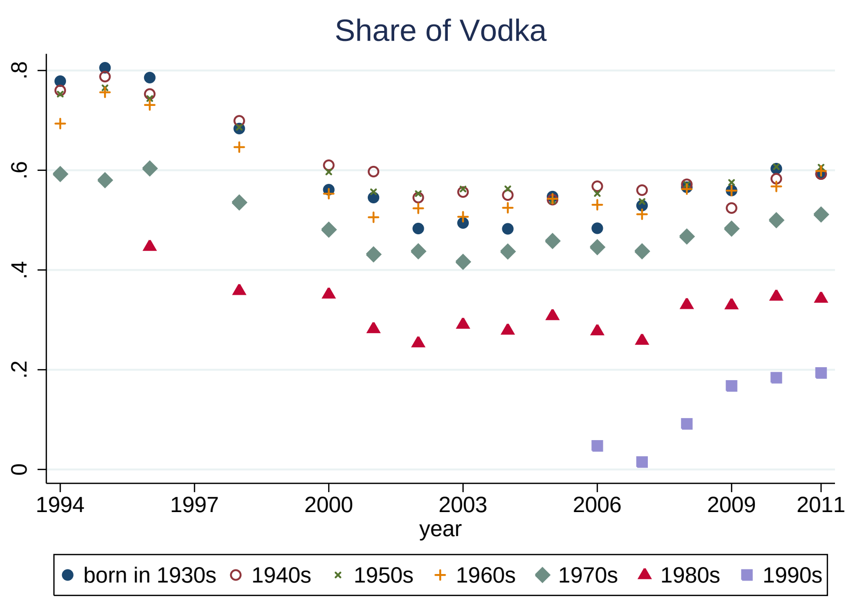
[31] The specification of regressions are as follows:
$share\,of\,good_{it} = \beta_0 + \beta_1 I(born\,in\,rural\,area)_i + \beta_2 controls_{it} + \rho_t + e_{it}$

[32] The specification of regressions are as follows: $share\,of\,good_{it} = \beta_0 + \beta_1 I(born\,before\,1980)_i + \beta_2[I(born\,in\,rural\,area)_i * \beta_2 I(born\,before\,1980)_i] + \beta_3 controls_{it} + \rho_r + \rho_t + e_{it}$, and
$share\,of\,good_{it} = \beta_0 + \beta_1 I(born\,before\,1980)_i + \beta_2[(annual\,temperature_r - \overline{annual\,temperature_r}) * I(born\,before\,1980)_i] + \beta_3 controls_{it} + \rho_r + \rho_t + e_{it}$

Figure 1: Shares of total alcohol intake for males across different age cohorts.



(a) share of beer



(b) share of vodka

Figure 2: Aggregate sales of beer and vodka 1970-2011, in billion liters.

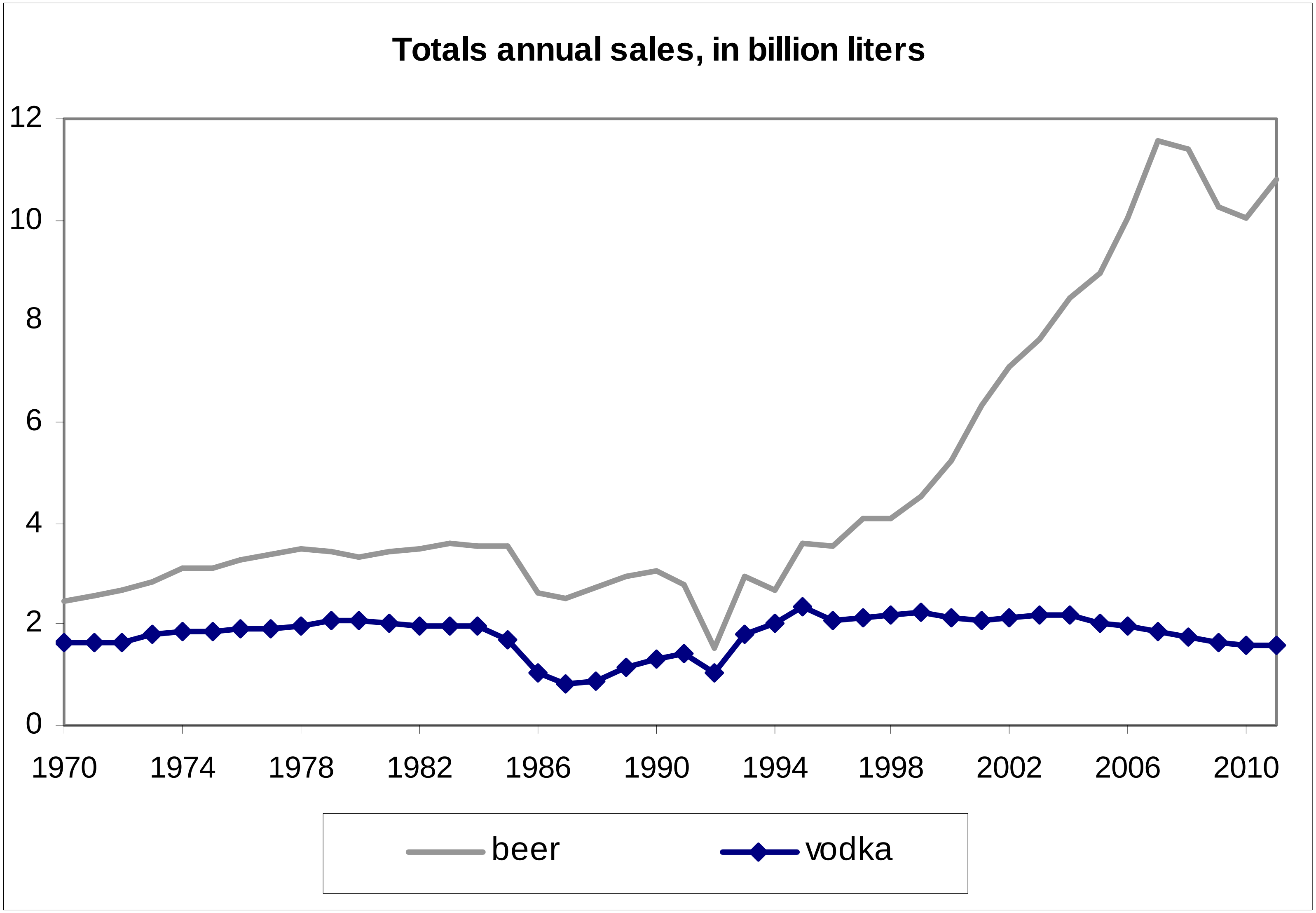


Figure 3: Share of beer and vodka consumption by age.

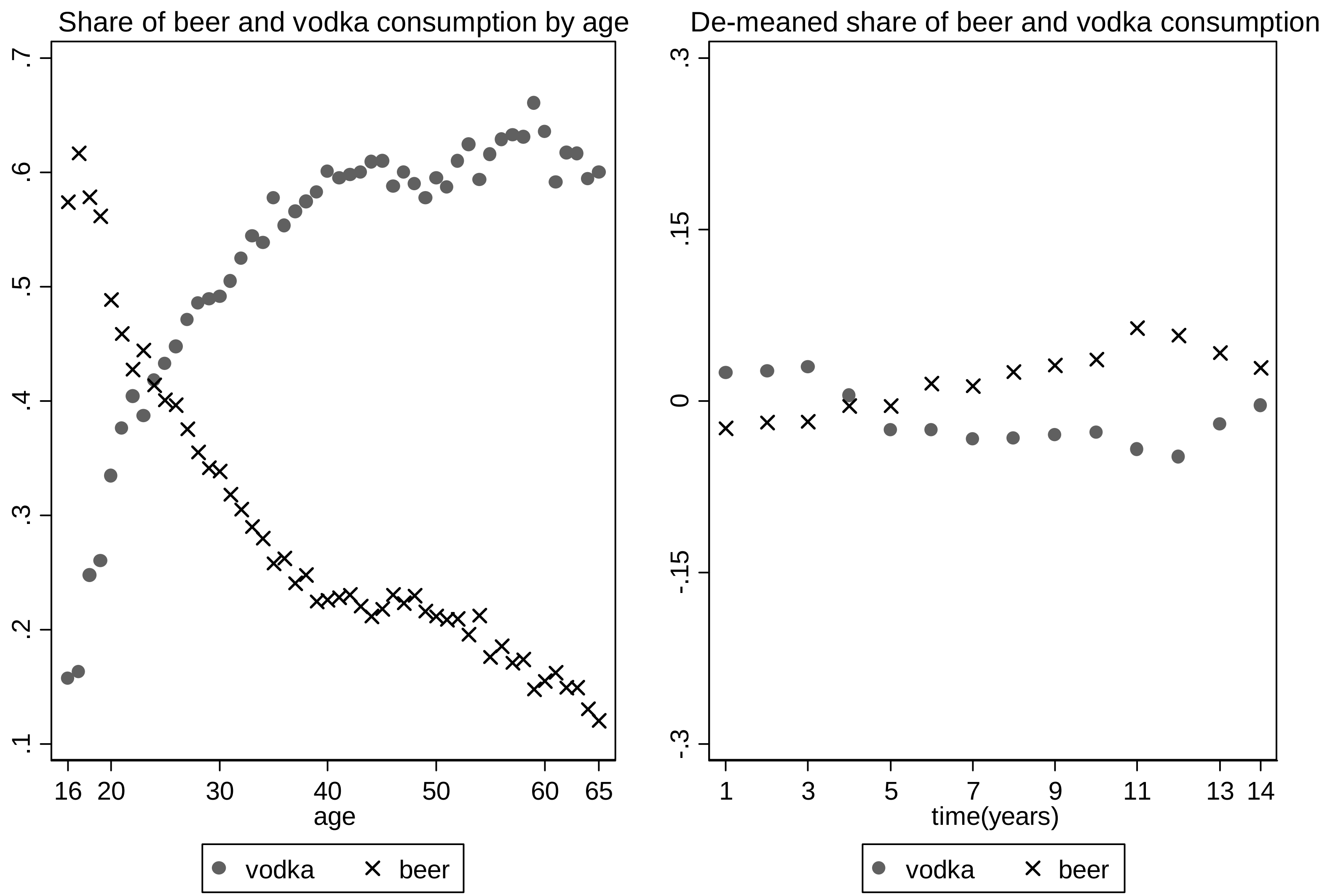
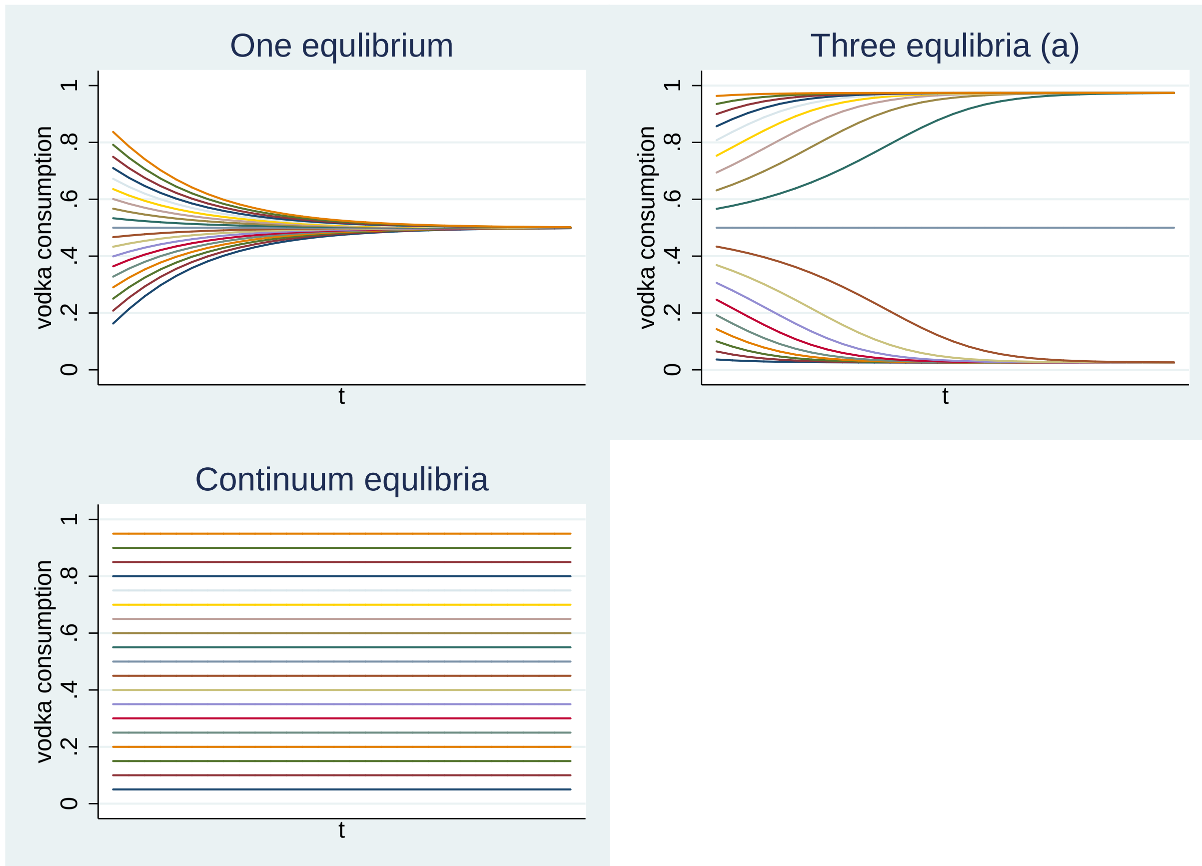
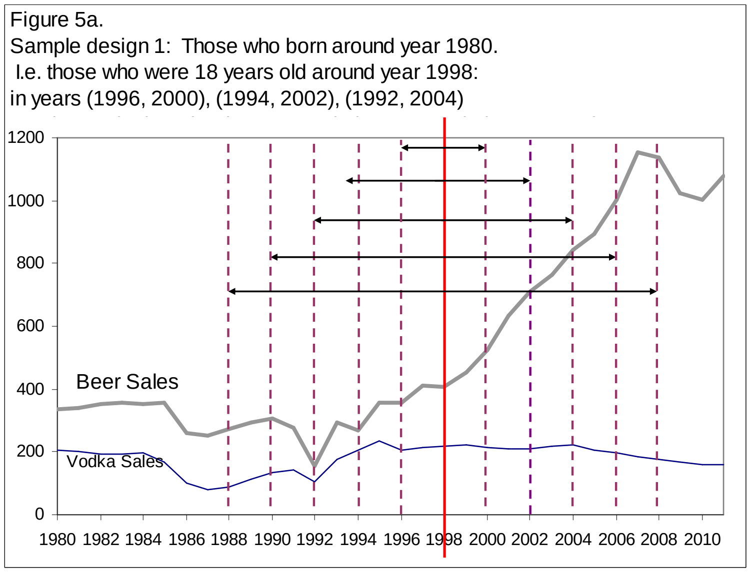


Figure 4: Illustration of different numbers of steady states

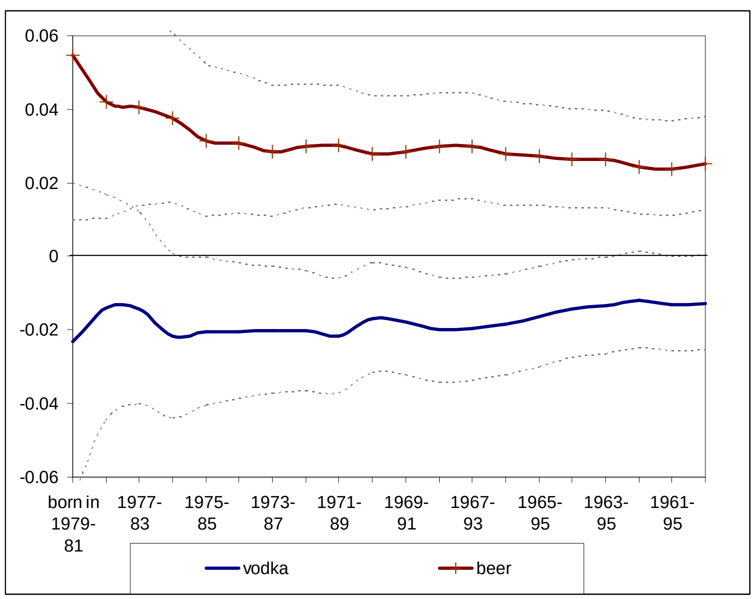


Notes: This figure shows how the share of vodka changes over time t depending on initial level at $t = 0$. The parametrization of the utilities in the three cases are as follows: $Y = 1$, $p_v = 1$,$\delta = 1/2$. (a): one equilibrium with $u = (v^{1/2} - 1)ln(Sv) + (b^{1/2} - 1)ln(Sb)$; (b): three equilibria with $u = (v^{1/2} - 1)ln(1.1 + Sv) + (b^{1/2} - 1)ln(1.1 + Sb)$; (c): continuum of of equilibria with $u = (v^{1/2} - 1)Sx^{1/2} + (y^{1/2} - 1)Sb^{1/2}$.

Figure 5: Shares of total alcohol intake for males across different age cohorts.

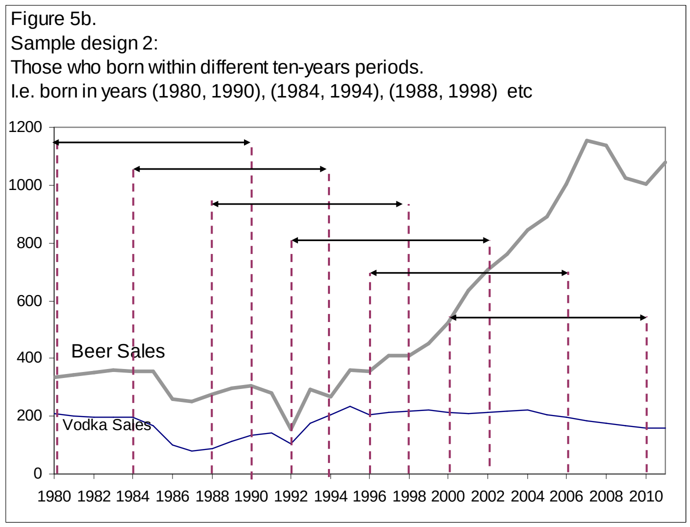


(a) sample design 1

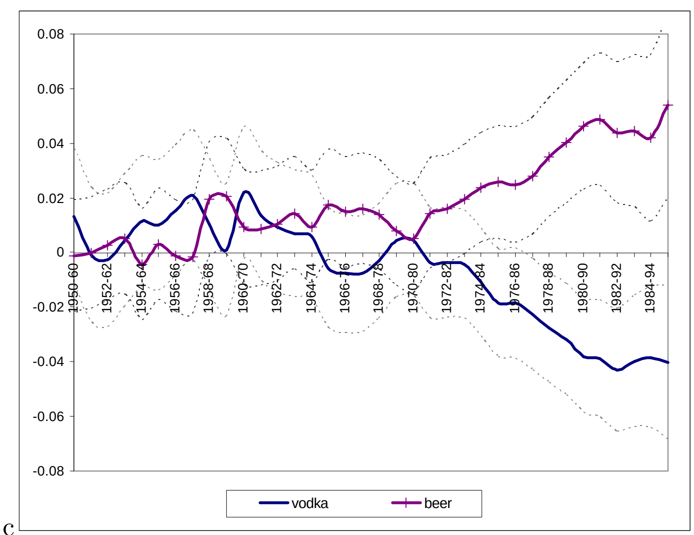


(b) sample design 2

Figure 6: Coefficients from regressions of shares of good on year of birth.



(a) regression coefficients



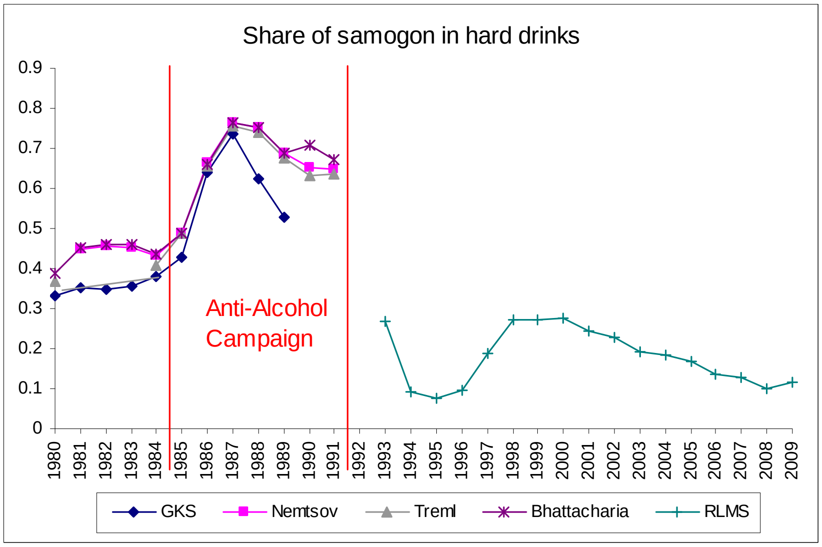
c

(b) regression coefficients

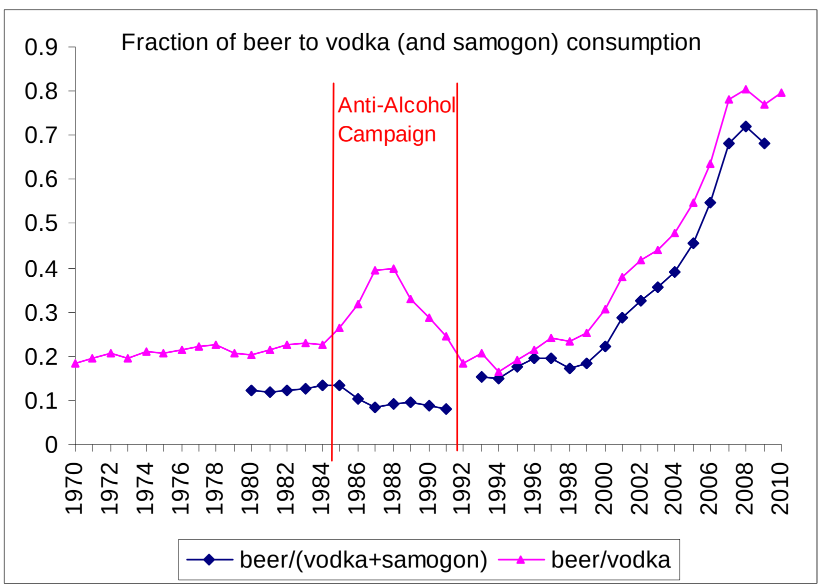
Notes: The figure on the left shows coefficients from regressions of shares of good on year of birth for people that were born in different time intervals around year 1980. Samples are constrained by length of periods of birth year are shown on horizontal axes.

The figure on the right shows coefficients from regressions of shares of good on year of birth for different age cohorts. Samples are constrained by ten-year periods of birth year are shown on horizontal axes.

Figure 7: Samogon to vodka production.

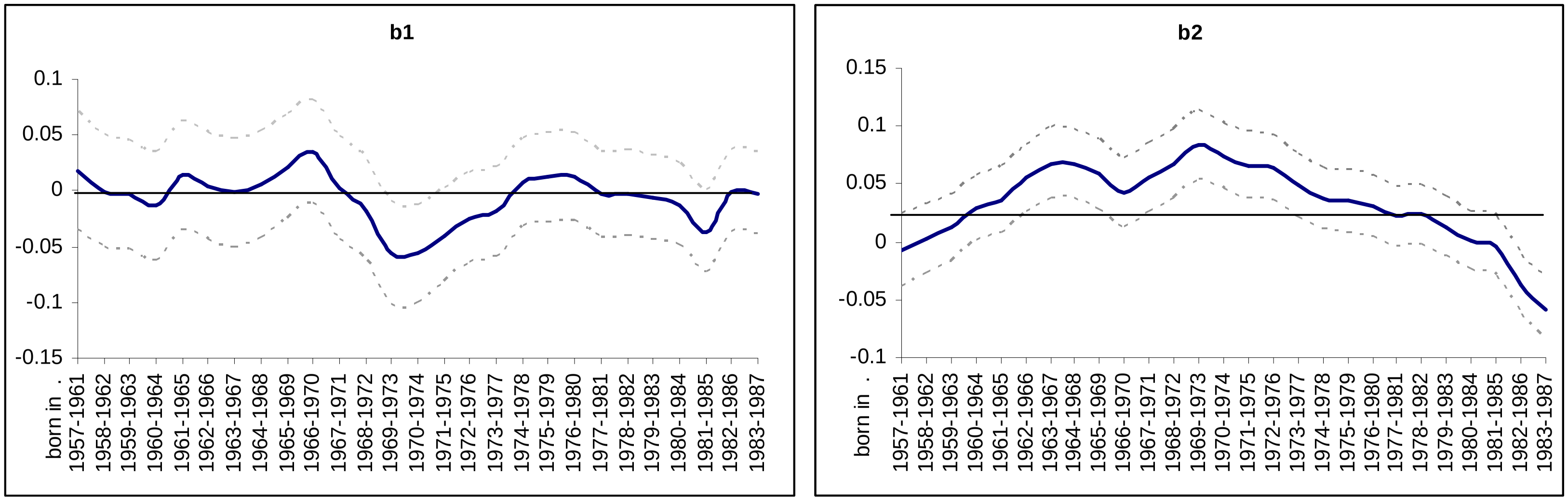


(a) share of samogon in hard drinks

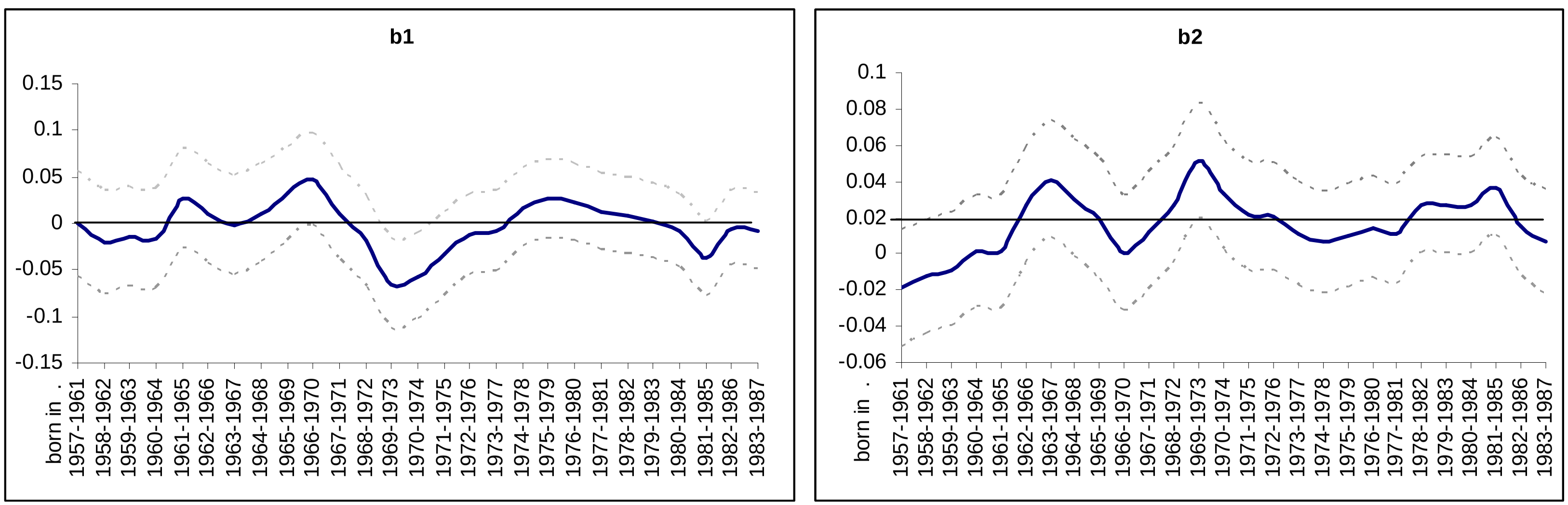


(b) fraction of beer to vodka (and samogon)

Figure 8: Gorbachev babies.

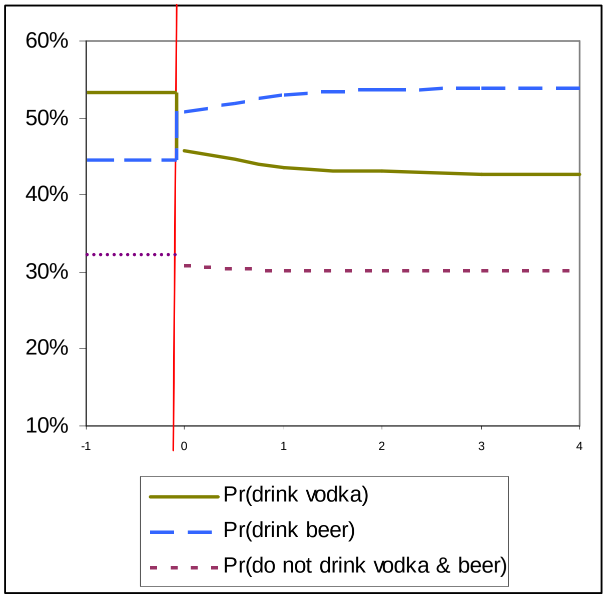


(a) Sample: All people

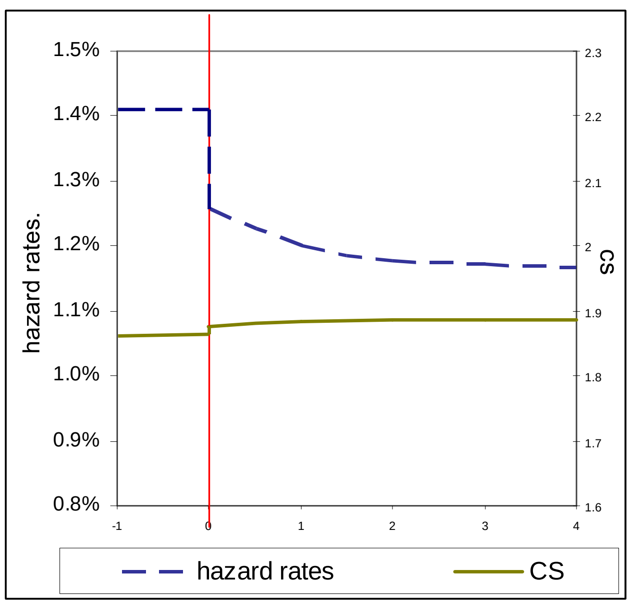


(b) Sample: those who born 20-years birth period interval

Figure 9: Effect of a 50% subsidy on beer and 30% tax on vodka.

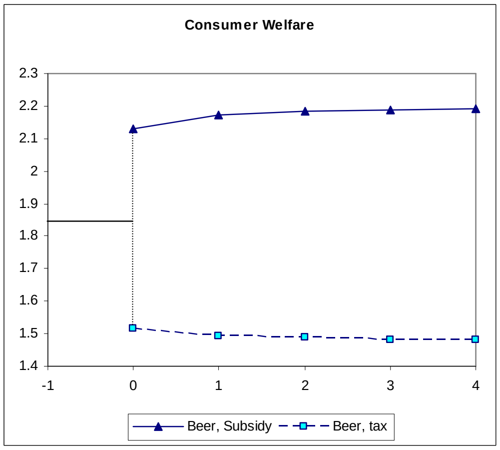


(a) probabilities

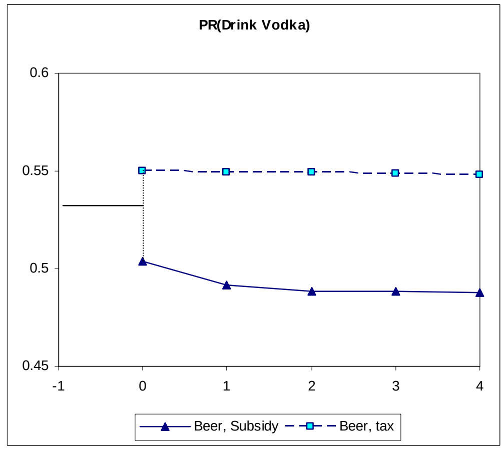


(b) hazard rates

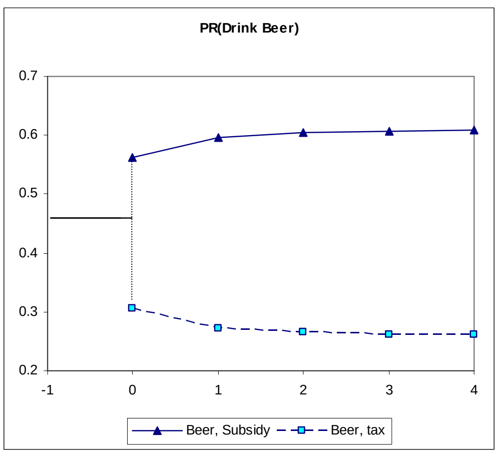
Figure 10: Effect of government policy on consumer welfare and on drinking patterns.



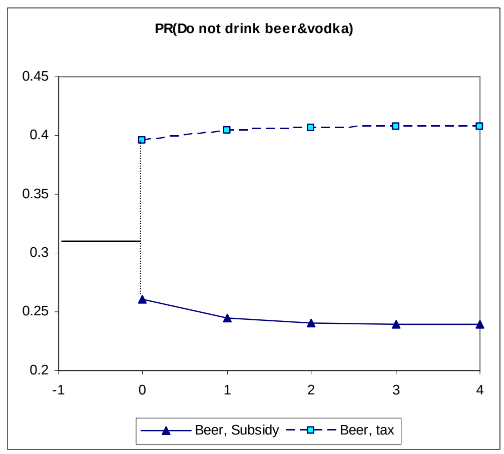
(a) consumer welfare



(b) probability of drinking vodka



(c) probability of drinking beer



(d) probability of not drinking

Table 1. Summary statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
Alcohol intake (in ml of pure spirit)	60549	107.50	134.83	0	2690
I(abstainer)	60325	0.30	0.46	0	1
vodka drunk last month (ml)	60089	158.54	236.52	0	5000
beer drunk last month (ml)	60249	448.72	682.86	0	10000
samogon drunk last month (ml)	60375	31.25	127.57	0	3000
vine drunk last month (ml)	60549	46.36	170.87	0	4000
share of vodka	43523	0.524	0.393	0	1
share of beer	43523	0.306	0.356	0	1
share of vine	43523	0.072	0.203	0	1
share of samogon	43523	0.079	0.230	0	1
I(drink beer)	35255	0.530	0.499	0	1
I(drink vodka)	60256	0.513	0.500	0	1
I(drink samogon)	59875	0.094	0.292	0	1
I(drink vine)	60092	0.129	0.335	0	1
lag7 share of beer	14130	0.188	0.264	0	1
lag7 share of vodka	14130	0.625	0.351	0	1
lag7 share of vine	14130	0.068	0.175	0	1
lag7 abstainer	9952	0.076	0.266	0	1
Age	60549	38.790	13.060	18	65
Body weight	56891	76.747	13.631	35	250
I(college)	60491	0.387	0.487	0	1
health status	60308	2.641	0.682	1	5
Log income	55351	9.170	1.855	2.08	16.59
birth year	60549	1965.0	14.2	1929	1993
born in rural area	60020	0.552	0.497	0	1
born in Georgia	60549	0.005	0.068	0	1
born in Moldova	60549	0.002	0.048	0	1
born in Ukraine	60549	0.030	0.172	0	1

Table 2. Cohort differences in patterns of alcohol consumption.

	(1)	(2)	(3)	(4)	(5)	(6)
	Share of beer			Share of vodka		
born in 1990s	0.557	0.161	0.312	-0.496	-0.152	-0.31
	[0.032]***	[0.049]***	[0.036]***	[0.025]***	[0.051]***	[0.030]***
born in 1980s	0.414	0.051	0.183	-0.338	-0.03	-0.169
	[0.009]***	[0.031]	[0.012]***	[0.010]***	[0.038]	[0.013]***
born in 1970s	0.211	-0.054	0.053	-0.177	0.052	-0.06
	[0.008]***	[0.025]**	[0.009]***	[0.010]***	[0.030]*	[0.010]***
born in 1960s	0.099	-0.081	omitted	-0.074	0.084	omitted
	[0.007]***	[0.018]***		[0.010]***	[0.022]***	
born in 1950s	0.06	-0.05	omitted	-0.04	0.056	omitted
	[0.007]***	[0.012]***		[0.011]***	[0.015]***	
born in 1940s	omitted	omitted	omitted	omitted	omitted	omitted
Year FE	NO	YES	YES	NO	YES	YES
Regional FE	NO	YES	YES	NO	YES	YES
Demographics	NO	YES	YES	NO	YES	YES
Observations	36206	30649	30649	36206	30649	30649
R-squared	0.15	0.19	0.19	0.08	0.15	0.15

Note: * significant at 10%; ** significant at 5%; *** significant at 1%. Robust standard errors clustered on individual level are in brackets. In regressions (2) and (5) pairwise tests of equality of cohort FE are rejected for all pairs except I(born in 1970s)=I(born in 1950s).

Table 3. State dependence in patterns of alcohol consumption.

			Share of beer			
Lag7(share of beer)	0.193	0.185	0.194			
	[0.024]***	[0.024]***	[0.024]***			
Lag1(share of beer)				0.316	0.305	0.316
				[0.010]***	[0.011]***	[0.010]***
Year of birth			-0.018			0
			[0.014]			[0.007]
born in 1950s		-0.022			-0.037	
		[0.024]			[0.014]***	
born in 1960s		-0.036			-0.069	
		[0.038]			[0.021]***	
born in 1970s		-0.002			-0.054	
		[0.055]			[0.031]*	
born in 1980s		0.047			-0.001	
		[0.072]			[0.039]	
born in 1990s		no data			0.105	
					[0.060]*	
Observations	7351	7351	7351	16178	16178	16178
R-squared	0.09	0.09	0.09	0.22	0.23	0.22
			Share of vodka			
Lag7(share of vodka)	0.159	0.155	0.159			
	[0.020]***	[0.020]***	[0.020]***			
Lag1(share of vodka)				0.328	0.32	0.328
				[0.010]***	[0.010]***	[0.010]***
Year of birth			0.021			0.005
			[0.016]			[0.008]
born in 1950s		0.018			0.046	
		[0.029]			[0.016]***	
born in 1960s		0.032			0.072	
		[0.045]			[0.024]***	
born in 1970s		-0.017			0.051	
		[0.065]			[0.034]	
born in 1980s		-0.08			0.004	
		[0.083]			[0.042]	
born in 1990s					-0.07	
					[0.058]	
Observations	7351	7351	7351	16178	16178	16178
R-squared	0.1	0.1	0.1	0.19	0.2	0.19

Table 4. "Pseudo-regression-discontinuity" regression results.

	Born in 1970-1990	Born in 1973-1987	Born in 1975-1985	Born in 1976-1984	Born in 1977-1983	Born in 1978-1982	
			Share of beer				
Lag7(share of beer)	0.198	0.174	0.195	0.174	0.147	0.196	OLS-1
	[0.037]***	[0.041]***	[0.047]***	[0.052]***	[0.058]**	[0.070]***	
Lag7(share of beer)	0.397	0.337	0.518	0.331	0.095	0.359	IV-1
	[0.119]***	[0.118]***	[0.142]***	[0.181]*	[0.236]	[0.355]	
F-test	156.45	156.14	81.89	46.1	27.52	12.29	
Year of birth	0.02	0.022	0.023	0.023	0.02	0.019	OLS-2
	[0.002]***	[0.002]***	[0.003]***	[0.003]***	[0.005]***	[0.007]***	
Lag7(share of beer)	0.181	0.154	0.161	0.162	0.15	0.19	OLS-3
	[0.039]***	[0.044]***	[0.051]***	[0.054]***	[0.060]**	[0.069]***	
Year of birth	0.009	0.01	0.021	0.01	-0.003	0.011	
	[0.005]*	[0.006]	[0.009]**	[0.011]	[0.016]	[0.024]	
			Share of vodka				
Lag7(share of vodka)	0.093	0.09	0.105	0.094	0.089	0.129	OLS-1
	[0.031]***	[0.036]**	[0.040]***	[0.049]*	[0.059]	[0.068]*	
Lag7(share of vodka)	0.193	0.204	0.271	0.159	-0.372	-0.034	IV-1
	[0.127]	[0.144]	[0.160]*	[0.207]	[0.258]	[0.381]	
F-test	117.16	86.67	56.82	34.24	24.22	9.93	
Year of birth	-0.013	-0.013	-0.014	-0.013	-0.008	-0.006	OLS-2
	[0.002]***	[0.002]***	[0.003]***	[0.003]***	[0.004]*	[0.007]	
Lag7(share of vodka)	0.086	0.083	0.093	0.09	0.115	0.135	OLS-3
	[0.033]***	[0.038]**	[0.043]**	[0.052]*	[0.061]*	[0.067]**	
Year of birth	-0.004	-0.005	-0.009	-0.004	0.029	0.01	
	[0.006]	[0.007]	[0.009]	[0.011]	[0.015]*	[0.023]	
# of obs.	1855	1288	826	633	462	331	OLS-1,3; IV-1
# of obs.	10224	8317	6396	5282	4160	3022	OLS-2

Table 5a. Share of good and year of birth.

	share of vodka			share of beer		
	all sample	born after 1980	born before 1980	all sample	born after 1980	born before 1980
Year of birth	-0.007	-0.033	-0.002	0.011	0.024	0.008
	[0.004]*	[0.009]***	[0.005]	[0.004]***	[0.010]**	[0.004]*
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25240	4391	20849	25240	4391	20849
R-squared	0.07	0.04	0.02	0.13	0.03	0.06

Robust standard errors clustered at individual level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 5b. Share of good and lagged regional sales.

	share of beer		share of vodka	
lag7 $\frac{regional\ sales\ beer}{regional\ sales\ vodka}$	0.01	0.006	-0.017	-0.014
	[0.003]***	[0.003]**	[0.003]***	[0.003]***
lag7 $\frac{regional\ sales\ beer}{regional\ sales\ vodka}$ *I$(age \leq 30)$		0.018		-0.014
		[0.004]***		[0.004]***
Observations	13730	13730	13730	13730
R-squared	0.05	0.05	0.02	0.02

Robust standard errors clustered at individual level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 6. Gorbahev babies.

	share of vodka	share of vodka
I(born after 1974)	-0.098	-0.074
	[0.017]***	[0.014]***
I(born before 1970)	-0.042	-0.013
	[0.018]**	[0.015]
I(born in city)*I(born before 1970)	0.064	
	[0.024]***	
I(born in city)*I(born after 1974)	0.051	
	[0.023]**	
I(born in city)	-0.065	
	[0.021]***	
Health evaluation	-0.003	-0.003
	[0.005]	[0.005]
I(college degree)	-0.011	-0.013
	[0.008]	[0.008]
Body weight	0.002	0.002
	[0.000]***	[0.000]***
Log(income)	0.006	0.006
	[0.001]***	[0.001]***
Age	0.006	0.006
	[0.001]***	[0.001]***
Regional FE	yes	yes
Time FE	yes	yes
Constant	0.402	0.36
	[0.039]***	[0.043]***
Observations	20601	20601
R-squared	0.11	0.11

Robust standard errors clustered at individual level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Second column shows results of following regression

$$\begin{aligned} share\ of\ vodka_{it} = \; & \beta_0 + \beta_1[I(born\ in\ 1970-1974)_i * I(born\ in\ city)] + \beta_2 I(born\ in\ 1970-1974)_i \\ & + \beta_3[I(born\ after\ 1974)_i * I(born\ in\ city)] + \beta_4 I(born\ after\ 1974)_i \\ & + \beta_5 I(born\ in\ city) + \Gamma' C_{it} + \rho_t + \rho_r + e_{it} \end{aligned}$$

Table 7. Estimators of $Pr(drink\, vodka|Y_{t-1})$

	Pr(drink vodka when 25 and older | .)				
	Obs	$\hat{Pr}()$	Std. Dev.	Min	Max
Pr(.|abstainer when teen)	182	0.659	0.475	0	1
Pr(.|try only beer when teen)	104	0.577	0.496	0	1
Pr(.|try vodka when teen)	301	0.817	0.387	0	1
	Summary Statistics				
Variable	Obs	Mean	Std. Dev.	Min	Max
Try beer when teen	2867	0.513	0.500	0	1
Try vodka when teen	2867	0.341	0.474	0	1
Abstainer when teen	2867	0.388	0.487	0	1
Try only beer when teen	2867	0.242	0.428	0	1

Table 8. Hazard of death estimates

	Hazard of death					
I(drink vodka)	0.68	0.82				
	[0.156]***	[0.165]***				
I(drink beer)		-0.516				
		[0.212]**				
I(heavy drinker: beer)			-0.503	-0.687		
			[0.280]*	[0.279]**		
I(heavy drinker: vodka)			0.555	0.935		
			[0.206]***	[0.224]***		
I(moderate drinker: beer)				-0.542		
				[0.290]*		
I(moderate drinker: vodka)				0.907		
				[0.183]***		
Log(beer consumption)					-0.14	
					[0.056]**	
Log(vodka consumption)					0.171	
					[0.034]***	
Share of vodka in alcohol consumption						0.301
						[0.080]***
Log(alcohol consumption)						-0.151
						[0.080]*
log(family income)	-0.323	-0.311	-0.314	-0.31	-0.31	-0.295
	[0.021]***	[0.022]***	[0.021]***	[0.022]***	[0.022]***	[0.023]***
Age	-0.278	-0.279	-0.272	-0.286	-0.279	-0.263
	[0.017]***	[0.017]***	[0.017]***	[0.018]***	[0.017]***	[0.017]***
Health evaluation	-0.576	-0.573	-0.562	-0.574	-0.579	-0.524
	[0.132]***	[0.132]***	[0.132]***	[0.132]***	[0.132]***	[0.140]***
I(smokes)	0.473	0.464	0.5	0.456	0.445	0.493
	[0.132]***	[0.132]***	[0.134]***	[0.133]***	[0.133]***	[0.146]***
I(college)	-1.836	-1.767	-1.841	-1.754	-1.753	-1.825
	[0.298]***	[0.300]***	[0.297]***	[0.301]***	[0.300]***	[0.308]***
Body weight	-0.002	-0.002	-0.003	-0.002	-0.002	-0.004
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]	[0.005]
I(employed)	-0.727	-0.696	-0.574	-0.683	-0.674	-0.559
	[0.145]***	[0.144]***	[0.144]***	[0.146]***	[0.144]***	[0.151]***
Observations	7069	7069	7069	7069	7069	6516

Standard errors in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%

Table 9a. Estimates of utility parameters. Multivariate logit. Main Specification.

	U(only beer)	U(only vodka)	U(vodka&beer)
$P_{beer,it}/P_{vodka,it}$	**-0.925**		**-0.388**
	0.202		0.152
$I(drink\,vodka)_{it-1}$	**-0.423**	**0.898**	**0.818**
	0.049	0.043	0.042
$I(drink\,beer)_{it-1}$	**0.972**	**-0.271**	**0.980**
	0.050	0.048	0.044
Income	**0.001**	**-0.001**	**0.001**
	0.000	0.000	0.000
Born in 1980s	**1.748**	**-2.386**	**-0.076**
	0.233	0.200	0.193
Born in 1970s	**1.503**	**-1.291**	**0.413**
	0.194	0.154	0.158
Born in 1960s	**0.876**	**-0.677**	**0.466**
	0.143	0.111	0.115
Born in 1950s	**0.528**	**-0.402**	**0.208**
	0.103	0.076	0.082
age	**0.018**	**-0.024**	**-0.010**
	0.006	0.005	0.005
health status	**0.221**	**0.131**	**0.164**
	0.054	0.046	0.046
I(college degree)	**0.262**	**-0.144**	**0.147**
	0.046	0.043	0.041
weight	**-0.008**	**-0.003**	**0.002**
	0.002	0.002	0.001
I(Muslim)	**-0.266**	**-0.258**	**-0.419**
	0.112	0.097	0.097

Note: standard errors in italic.

Table 9b. Estimates of utility parameters. Multivariate logit. Alternative Specifications.

	U(only beer)	U(only vodka)	U(vodka&beer)
$P_{beer,it}/P_{vodka,it}$	**-0.801**		**-0.259**
	0.204		*0.153*
$I(drink\ only\ vodka)_{it-1}$	**-0.812**	**0.316**	**-1.032**
	0.077	*0.060*	*0.058*
$I(drink\ beer\&vodka)_{it-1}$	**0.551**	**0.634**	**1.729**
	0.063	*0.058*	*0.055*
$I(drink\ only\ beer)_{it-1}$	**0.502**	**-0.718**	**-0.906**
	0.072	*0.085*	*0.069*
Born in 1980s	**2.588**	**-2.688**	**0.577**
	0.289	*0.239*	*0.237*
Born in 1970s	**2.283**	**-1.564**	**1.004**
	0.252	*0.194*	*0.203*
Born in 1960s	**1.575**	**-0.905**	**0.984**
	0.206	*0.148*	*0.161*
Born in 1950s	**1.149**	**-0.586**	**0.667**
	0.170	*0.111*	*0.129*
Born in 1940s	**0.713**	**-0.193**	**0.512**
	0.155	*0.088*	*0.112*
Born in 1930s	*omitted*	*omitted*	*omitted*
log(income)	**0.001**	**0.000**	**0.001**
	0.000	*0.000*	*0.000*
age	**0.026**	**-0.029**	**-0.003**
	0.006	*0.005*	*0.005*
health status	**0.222**	**0.123**	**0.172**
	0.055	*0.046*	*0.046*
I(college degree)	**0.241**	**-0.124**	**0.124**
	0.047	*0.043*	*0.041*
weight	**-0.008**	**-0.002**	**0.001**
	0.002	*0.002*	*0.001*
I(Muslim)	**-0.268**	**-0.254**	**-0.418**
	0.112	*0.097*	*0.097*

Note: standard errors in italic.

Table 11. Share of drinkers and mortality rates under alternative regulatory policies

	No tax/subsidy	Subsidy on beer	Tax on beer
Drinking patterns:			
Do not drink beer/vodka	0.31	0.24	0.41
Drink only beer	0.15	0.27	0.04
Drink only vodka	0.22	0.15	0.33
Drink beer&vodka	0.31	0.34	0.22
Hazard of death (1)	0.0137	0.0113	0.0168
Hazard of death (2)	0.0152	0.0149	0.0154
CS	1.87	2.19	1.48

Table 9a. Beer and vodka consumption for those who born in rural/urban areas.

	whole sample	live in this place≥7 yrs	live in city	live in city, live in this place≥7 yrs
	Share of beer			
I(born in village)	-0.02	-0.022	-0.025	-0.038
	[0.007]***	[0.010]**	[0.011]**	[0.016]**
Observations	16469	9779	7013	3718
R-squared	0.16	0.13	0.15	0.12
	Share of vodka			
I(born in village)	0.028	0.036	0.043	0.057
	[0.009]***	[0.012]***	[0.012]***	[0.017]***
Observations	16469	9779	7013	3718
R-squared	0.14	0.13	0.12	0.11

Table 9b. Regional differences in cohort effect.

		share of vodka	
I(born before 1980)	0.106	0.117	0.104
	[0.011]***	[0.012]***	[0.018]***
I(born before 1980) X		-0.002	
average temperature in 1997		[0.001]**	
I(born before 1980) X			0.034
I(born in village)			[0.012]***
Health evaluation	-0.004	-0.003	-0.002
	[0.005]	[0.005]	[0.008]
I(college degree)	-0.008	-0.008	-0.018
	[0.008]	[0.008]	[0.012]
Weight	0.002	0.002	0.002
	[0.000]***	[0.000]***	[0.000]***
Age	0.005	0.005	0.004
	[0.000]***	[0.000]***	[0.001]***
Log(income)	0.016	0.016	0.02
	[0.004]***	[0.004]***	[0.006]***
Year FE	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes
Observations	21284	21284	10754
R-squared	0.1	0.1	0.09

Standard errors clustered at individual level in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 10. Beer and vodka consumption for those who born in rural/urban areas.

	share of wine	
	whole sample	live in current place$\geq$7 yrs
Born in Moldova	0.039*	0.004
Born in Georgia	0.036**	0.027
Born in Ukraine	0.013**	0.002

Table 11. Preferences towards beer are negatively correlated with alcohol intake

	alcohol intake	log (alcohol intake)
Share of beer in alcohol intake	-89.18	-0.512
	[3.565]***	[0.020]***
Share of hard drinks in alcohol intake	30.642	0.652
	[3.264]***	[0.018]***
Constant	158.721	4.439
	[2.995]***	[0.017]***
Observations	32637	32637
R-squared	0.08	0.22

Standard errors in brackets

* significant at 10%; ** significant at 5%; *** significant at 1%

Online Appendix

'How Persistent Are Habits? Micro-Evidence From Russian Men'

Lorenz Kueng

Evgeny Yakovlev

O1. Demand Elasticity from Hedonic Regressions

This section presents the results of hedonic regressions for the price of alcohol. Table 1 presents OLS and tobit estimates for own and cross-price elasticities for different measures of vodka and beer consumption, as well as for total alcohol intake measures, for the following hedonic regressions:

(1) $Y_{it} = \alpha + \gamma_b Log(P_{beer})_{it} + \gamma_v Log(P_{vodka})_{it} + \Gamma' D_{it} + \delta_r + e_{it}$

(2) $Y_{it} = \alpha + \gamma_{bv} P_{beer}/P_{vodka} + \Gamma' D_{it} + \delta_r + e_{it}$

Y_{it} stands for alcohol consumption, D_{it} is a set of demographic characteristics, and δ_r is the regional fixed effects. The set of demographic characteristics includes log(family income), health status, age, I(Muslim), I(college degree), and personal body weight.

Table 2 below and Table 2A at the end of the section illustrate significant negative own and positive cross-price elasticities. According to tobit estimates, a decrease in the price of beer by 10% results in an increase in consumption of beer by 6% and a decrease in the consumption of vodka by 7%. Similarly, a decrease in the price of vodka by 10% will result in an increase in consumption of vodka by 9% and a decrease in the consumption of beer by 10%.

Estimated own-price elasticities (-0.6 for beer and -0.9 spirits) are within the range of those obtained in other studies. Leung and Phelps (1993) and Fogarty (2010) survey estimates of price sensitivity of demand for alcoholic beverages. Leung and Phelps (1993) find that average estimates for the elasticity of beer and spirits are -0.3 and -1.5 correspondingly. Fogarty (2010) finds that average own elasticities are -0.44 for beer and -0.73 for spirits.

O2. Proof

Steady state FOC for myopic agents as a function of v is

$F = u_v(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) - p_v u_b(v,\ y - p_v v,\ [\delta/(1-\delta)]v,\ [\delta/(1-\delta)][y - p_v v]) = 0.$

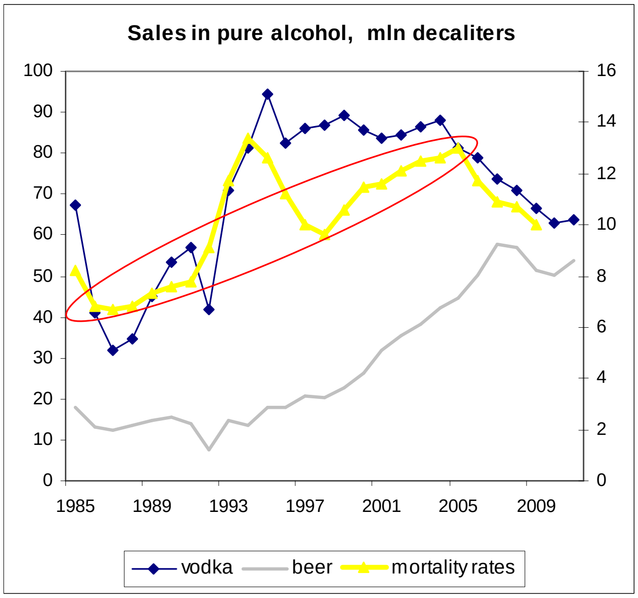
Differentiation with respect to the level of vodka we obtain
$dF/dv = u_{vv} - p_v u_{vb} + \delta/(1-\delta)u_{vH_v} - p_v\delta/(1-\delta)u_{vH_b} - p_v[u_{bv} - p_v u_{bb} + \delta/(1-\delta)u_{bH_b} - p_v\delta/(1-\delta)u_{bH_b}].$

We know that $u_{vv} < 0$, $u_{bb} < 0$, $u_{s_v s_v} < 0$, $u_{H_b H_b} < 0$, $u_{vH_v} > 0$, $u_{bH_b}(.) > 0$, and thus some terms in this expression are positive, e.g. $\delta/(1-\delta)u_{vSv}$, $p_v^2\delta/(1-\delta)u_{bSb}]$, and some are negative (u_{vv}, $p_v^2 u_{bb}$,). Therefore the sign of the overall sum is unknown.

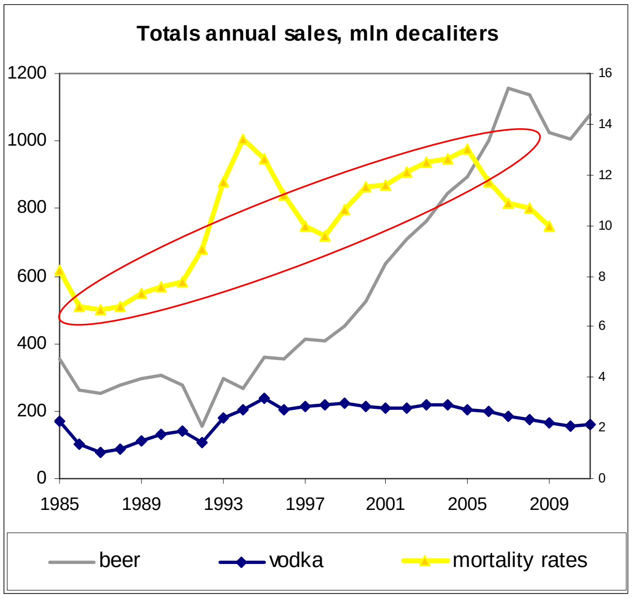
References in Online Appendix

TBA

Figure 2: Sales of vodka and beer (in billion liters), and mortality rates.



(a) total sales of beverages



(b) total sales of pure alcohol

Table A1. Share of home-made vodka (samogon) in total alcohol intake.

Variable: share of samogon	Obs	Mean	Std.Dev	Min	Max
sample: small-size cities and rural areas	23344	0.143	0.305	0	1
sample: middle-size and big cities	17790	0.025	0.124	0	1

Table A2. Demand elasticities: price hedonic regression.

	Log(beer consumption)		Log(vodka consumption)	
	OLS	Tobit	OLS	Tobit
$Log(P_{beer})$	-0.33	-0.582	0.358	0.673
	[0.085]	[0.195]	[0.110]	[0.208]
$Log(P_{vodka})$	0.5	1.029	-0.48	-0.894
	[0.068]	[0.155]	[0.089]	[0.168]

Standard errors are in brackets

Table A3. Sales of vodka and beer; and mortality rates: countryXyear-level regressions

	1985-2009	1992-2009
	Male mortality	
Vodka sales	0.035***	0.02***
Beer sales	0.001	0.0003

Table A4. Sales of vodka and beer; and mortality rates: regional-level regressions

	# of deaths per 1000 of working age males	
Log Vodka sales	0.809	0.554
per capita	[0.175]***	[0.254]**
Log Beer sales	-0.147	-0.66
per capita	[0.134]	[0.171]***
I(Caucasus)	-4.935	-6.341
	[0.296]***	[0.527]***
Year FE	YES	YES
Weighted?	NO	YES
Constant	13.99	10.658
	[0.299]***	[1.888]***
Observations	949	949
R-squared	0.49	0.37

Standard errors in brackets; * significant at 10%;
** significant at 5%; *** significant at 1%