NBER WORKING PAPER SERIES

HOW PERSISTENT ARE CONSUMPTION HABITS? MICRO-EVIDENCE FROM RUSSIA'S ALCOHOL MARKET

Lorenz Kueng
Evgeny Yakovlev

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

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

We thank Orley Ashenfelter, David Card, Irina Denisova, Sergey Guriev, Denis Nekipelov, Gerard Roland, Katya Zhuravskaya, and seminar participants at UC Berkeley, Northwestern, the New Economic School and CEFIR, the Empirical Macro Workshop in New Orleans, and the EEA meeting for helpful discussions and comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

How Persistent Are Consumption Habits? Micro-Evidence from Russia's Alcohol Market
Lorenz Kueng and Evgeny Yakovlev
NBER Working Paper No. 20298
July 2014
JEL No. D12,E21,G02,I10

ABSTRACT

We use the Anti-Alcohol Campaign in 1986 and the rapid expansion of the beer market after the collapse of the Soviet Union as two quasi-natural experiments to identify highly persistent habit formation in alcohol consumption among Russian males. Importantly, these results apply to all levels of alcohol consumption and are not driven by heavy drinking or alcoholism. The two large shocks combined with persistent habits lead to large cohort differences in consumer behavior even decades later. We derive a basic model of habit formation with homogeneous preferences over two habit-forming goods, which is consistent with these facts. Using placebo tests as well as simple descriptive statistics, we show that habits are formed during early adulthood and remain largely unaffected afterward. The main alternative hypotheses such as income effects, unobserved taste heterogeneity, stepping-stone effects, and changes in culture or social norms are inconsistent with those patterns. Using the experiments as IVs, we estimate the first-order autoregressive coefficient to be 0.83, which is almost three times larger than its OLS estimate. Finally, our results suggest that male mortality in Russia will decrease by one quarter within twenty years even under current policies and prices due to the long-run consequences of the large changes in the alcohol market.

Lorenz Kueng
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
l-kueng@kellogg.northwestern.edu

Evgeny Yakovlev
Acumen LLC
500 Airport Blvd, Suite 365
Burlingame, CA, 94010
evgeny.yakovlev@gmail.com

1 Introduction

Consumption habits are notoriously difficult to measure in the data and are challenging to identify separately from other behavior. Nevertheless, they play an important role in many economic models ranging from applied microeconomics to macroeconomics and finance. These models usually have very different properties depending on the dynamics of the habit formation process as well as the mechanism that leads to habit formation. In this paper we provide strong evidence that state dependence can indeed be very persistent and we develop a simple framework to analyze such behavior. Our analysis shows that the initial choice of consuming a habit-forming good affects individual choices even decades later. Using two quasi-natural experiments and long panel data from Russia with consumption available at the individual level (as opposed to the household level), we show that males who start consuming one type of alcoholic beverage during early adulthood form strong habits toward this type of beverage. The habit formed during early adulthood lasts their entire lifetime, leading to very persistent consumption behavior, especially once the habit stock has fully accumulated, which we estimate to be around age 22. Importantly, this result applies to *all* levels of alcohol consumption and is not limited to heavy alcohol consumption or alcoholism.[1] As a matter of fact we find that individuals form habits at least as easily toward light alcohol as toward harder drinks.

[Figure 1 about here]

Figure 1 illustrates the persistence of consumption habits among male alcohol consumers, showing strong cohort differences in alcohol consumption patterns in each year from 1994 to 2011. Indeed, preferences regarding beer and vodka have not changed significantly during the entire sample period despite dramatic changes to the availability of different types of alcoholic beverages over the past two decades.[2] Males born in the 1960s or earlier who spent their early adulthood in the Soviet Union still prefer vodka today, whereas younger generations overwhelmingly prefer beer. These differences are also quantitatively large: vodka constitutes on average 60% of total alcohol intake for males born in the 1960s or earlier, but only 48% for those born in the 1970s who spent their adulthood in the transition period after the collapse of the Soviet Union, and 32% and 19% for those born in the 1980s and 1990s, respectively, who spent their early adulthood in the post-Soviet era. In contrast,

[1] First, we control for the level of total alcohol intake in all our specifications. Second, following the literature (e.g., Bronnenberg et al. (2012), Atkin (2013)) we use shares of alcohol consumed instead of levels in order to make the results robust to outliers. Finally, all our findings are robust to dropping the top quartile of the distribution of alcohol consumers, measured in terms of total alcohol intake, i.e., ethanol.

[2] We use the term "vodka" to include vodka and other hard liquor, but we exclude homemade liquor (i.e., samogon). The production of homemade liquor for personal consumption became legal only in 1997, and selling it remains illegal today. This variable is therefore measured very imprecisely, and we exclude it. Our results are robust to including samogon, although the standard errors increase. The term "beer" includes home-brewed beer in addition to purchased beer. However, the fraction of home-brewed beer is negligible for the vast majority of households, and thus it was not asked separately in most rounds of the survey.

the share of beer in total alcohol intake for these age groups constitutes 20%, 36%, 56%, and 68%, respectively.

Exploiting the long panel dimension of our data, we show that this pattern is driven almost entirely by variation in consumption between individuals rather than within individuals, i.e., variation over time following the same individual. In particular, we find little evidence for the hypothesis commonly formed based on cross-sectional data that differences in the consumption of light and hard alcohol at different ages are driven by a so-called "stepping-stone" or "gateway" effect of light alcohol, where individuals consume light alcoholic beverages when young before switching to harder alcohol later in life. We show that after controlling for individual fixed effects, the consumption-age profile is almost completely flat for consumers starting in the their mid 20s, while there is some modest stepping-stone effect at younger ages. More importantly, these small age effects are dominated by the cohort effects. The age effects can at best explain about one-fifth to a quarter of the unconditional age profile, while the rest is due to long-run effects of shocks to the alcohol market. These shocks that occurred in the distant past mostly affected the habits of consumers who were young at that time, and they manifest themselves as cohort effects in the survey data that were collected long after those shocks occurred.

In all our analyses we also control for household income to account for the possibility that post-Soviet cohorts might be richer than older cohorts and that this in turn might account for the pattern shown in Figure 1 if beer has a higher income elasticity than vodka. However, we do not find much evidence for such income effects. Moreover, Russia's gross domestic product (GDP) displayed a strong J-curve pattern after the collapse of the Soviet Union, which is typical for post-Soviet transition economies.[3] This pattern as well as the recession caused by the Russian financial crisis of 1998 do not show up in Figure 1, providing further indirect evidence against strong income effects. In our analysis we also control non-parametrically for contemporaneous relative prices as a potential alternative explanation using time and region fixed effects. However, in the absence of habit formation, contemporaneous relative prices cannot explain cohort differences since all cohorts face the same prices in a given period.[4]

Taking advantage of two quasi-natural experiments, we show that unobserved taste heterogeneity also does not explain a significant part of the observed consumption patterns. Similarly, slowly evolving unobserved factors such as culture or social norms are also not consistent with the individuals' responses to the two quasi-natural experiments.

[Figure 2 about here]

[3]See the International Monetary Fund's World Economic Outlook database, http://www.imf.org/external/pubs/ft/weo/.

[4]With habit formation, relative prices can affect the initial choice of young individuals of which habit-forming good to consume.

We propose a simple explanation for the observed consumption patterns based on persistent habits that are formed during early adulthood when individuals start to consume beer for the first time in their life. Individuals in our model are born with the same preferences but exposed to different initial alcohol-market conditions and will therefore form habits toward very different goods. Importantly, with two habit-forming goods we can observe multiple long-run equilibria even without any unobserved individual heterogeneity. To understand these cohort differences as a result of habit formation it is useful to note that the vodka industry dominated the alcohol market measured in terms of pure alcohol during the Soviet Union. Since 1992, however, the beer industry has expanded rapidly, as shown in Figure 2, 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 alcohol producers—and the entry of foreign competition into this new market.[5] For instance, in 1991, shortly before the collapse of the USSR, there were no foreign-owned beer breweries in Russia 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 the market for beer to foreign competition also 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 also changed significantly,[6] and the assortment of beer has increased dramatically from only 20 varieties offered in 1991 to over 1,000 in 2009.[7] As a result, from 1991 to 2011, the last year included in our analysis, beer sales have increased by a factor of four from 2.8 to 10.8 billion liters. In contrast, vodka sales have not followed the same trend. Total annual sales of vodka were 1.59 billion liters in 2011, which is roughly the same level as during the Soviet era.[8]

These stark changes form our first quasi-natural experiment for the study of habit formation. While Figure 2 clearly shows that these changes did alter the drinking patterns of the population as a whole, the most significant shift in tastes occurred in younger generations. Males who started

[5] We do not use official sales data in 1992 and 1993, since they are most likely wrong due to significant data-collection and reporting problems in the wake of the collapse of the Soviet Union. These numbers severely underreport all levels of alcohol production.

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

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

[8] In the final 20 years of the USSR, from 1970 to 1991, average annual sales of vodka totaled 1.66 billion liters, and annual sales of beer 3.09 billion liters. In terms of pure alcohol, these numbers correspond to 0.66 billion liters for vodka and only 0.15 billion for beer. We do not discuss values here because there were no formal market prices in the Soviet Union. Instead, the alcohol industry was monopolized by the state, and quantities produced were heavily regulated. As a result, it was difficult or even impossible to find many goods in stores, and prices were usually not the most significant factor as there was severe rationing.

consuming alcohol during the Soviet period became accustomed to vodka and still prefer vodka today even after the collapse of the Soviet Union and after the significant expansion of the beer industry. Younger generations, however, who spent their early adulthood in a time with easier access to beer than previous generations, prefer beer over vodka despite facing the same prices as those older cohorts, and even after controlling for income, demographics, and the level of total alcohol consumed.

In addition to using the rapid expansion of the beer market after the collapse of the Soviet Union, we also use the sharp drop in the production of both beer and especially vodka during Mikhail Gorbachev's anti-alcohol campaign, which affected the alcohol markets from 1986 to 1991, as our second quasi-natural experiment, also clearly visible in Figure 2. While this policy sharply reduced the *official* production of alcohol, it simultaneously also lead to a dramatic increase in the *illegal* production of homemade vodka called samogon. Crucial for our identification is the fact that the increase in home-produced samogon was much more prevalent in rural areas than in the densely populated urban areas for reasons we discuss in Section 5.3. We exploit this differential impact of Gorbachev's policy on young rural men relative to young urban men to identify the long-run impact of this policy on habit formation using a difference-in-difference design. Specifically, we analyze the persistent effects of this shock on the habit formation of urban males who spent their habit-forming years during the campaign relative to males reaching the same age before or after the anti-alcohol campaign.

Placebo tests for both experiments reveal that habits are indeed formed during early adulthood in a small window, which is roughly centered at age 18. Hence, the results from the experiments are consistent with the simple age profile obtained non-parametrically after controlling for individual fixed effects.

Exploiting the two quasi-natural experiments, we quantify the persistence of the habit formation process. We find that autoregressive specifications of habit formation typically estimated in the literature are severely *downward* biased when estimated by ordinary least squares (OLS), contrary to the upward bias implied by unobserved taste heterogeneity. This downward bias is a result of the attenuation induced by the substantial measurement error typically found in expenditure surveys. Once we instrument the lagged consumption share using the exogenous variation provided by the two quasi-natural experiments, the first-order autoregressive coefficient increases threefold, from 0.28 to 0.83 for beer and from 0.33 to 0.85 for vodka. The difference in estimates provides us with an estimated lower bound for the attenuation factor due to measurement error of two-thirds, with one corresponding to a complete attenuation to zero.

Finally, we study the implications of our results for male life expectancy—currently only 60 years compared with 75 years in the US—taking into account the persistent habits we uncovered. We estimate that male mortality in Russia will decrease by one quarter within twenty years even under the current set of policies and current levels of relative prices of alcoholic beverages. This will happen simply because new generations will be more accustomed to beer and will replace older generations with strong preferences for vodka. Since much of the gap in male life expectancy is due to occasional binge drinking, which is more likely to occur for males who prefer vodka, this shift in consumption habits toward beer in turn implies strong effects on morality. Hence, this reduction in the male mortality rate will be the result of changes that occurred several decades ago.

The persistence of the consumption habits as well as the underlying mechanism we document in this paper have important implications for various fields in economics such as health economics and consumer demand.[9] Interpreted more broadly, habit formation has also been successfully applied in asset pricing and macroeconomics to explain several empirical puzzles.[10] However, most of this literature lacks convincing identification of habits in general and of the persistence of such habits in particular. While the strong habit formation we document is certainly evidence in support of such models used in macroeconomics and finance, the microeconomic behavior we uncover is very different. In particular, we find that summarizing habits in terms of short lags of consumption, as is typically done in this literature when modeling aggregate data, is not appropriate for modeling individual behavior. Instead of lagged consumption, we find that *initial* consumption largely determines habits, and once those habits are formed, they are very persistent, at least for the goods we study. Hence, aggregate behavior changes as older cohorts who formed their habits in the distant past exit and younger cohorts enter and form different habits if the environment changes.[11] Our results therefore echo the literature on cohort differences in beliefs and preferences, such as beliefs about inflation and macroeconomic risk or preferences for redistribution and state intervention in former communist countries.[12] This research suggests that the cultural and political environment in which an individual grows up affects his behavior over his entire lifetime.

[9] Examples in this literature include Becker and Murphy (1988), Chaloupka (1991), Becker et al. (1994), Cook and Moore (2000), Williams (2005), Bronnenberg et al. (2012).

[10] Examples of this literature include Eichenbaum et al. (1988), Sundaresan (1989), Heien and Durham (1991), Campbell and Cochrane (1999), Dynan (2000), and Ravina (2007). This paper however follows the "deep habits" literature, which analyzes habit formation across goods rather than over a consumption aggregate; see, e.g., Ravn et al. (2006). A related paper is Atkin (2013), who uses a sample of Indian households to analyze the consequences of habit formation over basic foods such as rice in the context of a trade model.

[11] This finding is also consistent with the mechanism described in Atkin (2013).

[12] See Guiso et al. (2004, 2008), Alesina and Fuchs-Schündeln (2007), and Malmendier and Nagel (2011a,b) for examples of the former and Denisova et al. (2010) for a discussion of the latter.

2 Data

We use data from the Russian Longitudinal Monitoring Survey (RLMS), which is a nationally representative annual survey panel starting in 1992 that covers more than 4,000 households per year corresponding to about 9,000 individual respondents.[13] Our initial sample consists of rounds 5 through 20 of the RLMS spanning the period from 1994 to 2011, but not including 1997 and 1999 when the survey was not conducted.[14] The data cover 33 regions (Russian oblasts), two of which are Muslim and hence contain fewer households that consume alcohol, plus the cities of Moscow and St. Petersburg. The RLMS 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% and is lowest in Moscow and St. Petersbug at 60% and highest in Western Siberia at 92%. The RLMS team provides a detailed analysis of attrition and does not find any systematic relationship with observables.[15]

Importantly, the survey's health module, which includes questions about alcohol consumption, is completed by each individual household member age 14 and older. Hence, in contrast to the previous literature which is restricted to household-level data or credit-card accounts, our data has the individual consumer as the unit of analysis. Another distinct feature of the health module is the fact that these questions ask about quantities consumed instead of expenditure outlays. Our consumption measures therefore directly capture actual consumption and are not subject to timing issues that may lead to a wedge between expenditures and consumption.[16]

[Table 1 about here]

Table 1 summarizes the socioeconomic and demographic characteristics as well as various measures of alcohol consumption for our sample, which we restrict to all males age 18 and older, with 18 being the minimum legal drinking age in Russia. Our primary measures of alcohol consumption are the shares of beer and vodka consumption in total alcohol intake, calculated in milliliters of pure alcohol.[17] The measures are derived from the survey's health module, which covers each member

[13]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.

[14]We do not use data from rounds 1 to 4 because they were conducted by another institution, have a different methodology, and are generally considered to be of much lower quality.

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

[16]For instance, we do not have to assume that an individual consumed all the alcohol it purchased in the reference period.

[17]In all specifications we also include the level of alcohol consumption. To construct these variables we use the amount of all alcoholic beverages consumed during the last month. We assume that beer contains 5% pure alcohol and vodka contains 40% pure alcohol, based on recommendations from the National Institutes of Health (NIH); see, e.g., 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 2000; see, e.g., Nemtsov (2002) and Bhattacharya et al. (2013). We assume a constant share of

of the household, and relate to the member's consumption rather than the household's alcohol expenditures, which are typically the only available measures of consumption in most other household surveys. Specifically, we use the individual's reported quantity consumed in a typical day during the last 30 days, and we then transform the volume to grams of pure alcohol (e.g., grams of ethanol in beer). We do not use expenditures on alcohol because they apply to the entire household and because they are of poor quality.[18] For instance, the average self-reported household budget share of alcohol in our sample is 5% for households reporting positive alcohol expenditures. This number is severely downward biased due to underreporting and more so than in other countries. Treml (1982), for example, shows that this level of underreporting already existed in earlier surveys, resulting in estimated alcohol expenditure shares of only 3%. Instead, we estimate the average share of alcohol in total retail sales based on official statistics to be 9% over our sample period.[19] While this measure of the alcohol budget share is conceptually close to the budget share in non-durable expenditures, which is the consumption measure typically used in the literature on habit formation, the estimated magnitude is most likely understating the alcohol budget shares of the individuals in our sample. Many households do not consume any alcohol at all, either for religious, health, or other reasons and official sales do not include the consumption of illegally obtained or homemade alcohol. Hence, the typical household's expenditure share in our sample could be well above 10%.

Vodka and beer are the most popular alcoholic drinks among Russian males, with an average share of total alcohol consumption across all years of 53% for vodka and 29% 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 73%, while beer had only a share of 10%. By 2011 these shares were already 49% and 38%, respectively. This general trend is also apparent in Figure 2 and is caused by the expansion of the beer market in the early years of our sample.

Table 1 also provides summary statistics for the main control variables we use in our analysis, both for our main sample of male alcohol consumers age 18 and above as well as for the sample of all individuals above age 18, including those who report not having consumed any alcohol during the previous month. We will use the latter sample when we analyze the effect of the changed alcohol

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 as shown in Figure 2. The assumption of a constant share of alcohol content in beer does not change our results.

[18]For instance, 47% of males who report having consumed alcohol during the previous month report zero household expenditures on alcohol, and another 11% do not report their spending on alcohol at all. Individual consumption data on the other hand tend to be of much higher quality and have fewer nonresponses. This is most likely due to the fact that the health questions are asked in isolation without any other person being present except the interviewer in order to maintain full confidentiality.

[19]Source: Goskomstat, Statistical Yearbook, Table 20.16.

patterns on male mortality. Since there is no consistent aggregate price index, especially during the financial crisis of 1997–98 we express real income by deflating it by the price of milk, which is stable over time. The corresponding real series is then comparable across our sample period from 1994 to 2011.

3 Analyzing Stepping-Stone Effects

The first main alternative explanation for cohort differences put forward in the literature—which is often based on cross-sectional data—is a "stepping-stone" or "gateway" effect of light drugs for the consumption of harder drugs later on. In the case of alcohol, this means that beer might serve as a stepping stone earlier in life for the consumption of harder alcoholic substances such as vodka later in life. According to this theory, people would start out with beer but eventually switch to harder drinks, in which case the observed cohort differences in Figure 1 would just be the effect of aging. The stepping-stone hypothesis is widely studied in health economics. Several studies have analyzed it in the context of various types of drugs and tested it against alternative explanations, in particular against unobserved individual heterogeneity in preferences.[20] However, to the best of our knowledge our study is the first to analyze the stepping-stone effect of beer towards harder alcoholic beverages.

[Figure 3 about here]

The aggregate sales in Figure 2 mask substantial heterogeneity in the drinking behavior across the age distribution as shown in the top-left panel of Figure 3. The share of beer consumption drops from 68% at age 18 to only 17% at age 65, while the share of vodka increases from 19% at age 18 to 54% at age 65. This remarkable age profile can potentially be driven by within- or between-consumer variation. A stepping-stone effect of beer would generate within-consumer variation where younger consumers start out with beer before gradually substituting to harder alcohol as they become older. In the case of between-consumer variation, different cohorts would have relatively flat alcohol life-cycle profiles, implying very persistent drinking habits. The initial share of beer relative to vodka would increase from one cohort to the next, so that the intercept of the age profile of younger cohorts would be higher than that of older cohorts for beer consumption, and vice versa for the share of vodka.

The top-right panel of Figure 3 assesses the relative contribution of those two forces by showing the average drinking patterns after taking out individual means. Specifically, for each individual

[20] 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 an increase in the probability of smoking marijuana later on. Van Ours (2003) finds that unobserved individual heterogeneity and stepping-stone effects can explain many patterns of drug consumption.

we subtract his average share, and we normalize the average of the first observed share across all individuals to zero. Hence, this figure shows the average slope of the age profile over all individuals in the sample after controlling for individual fixed effects. Explaining the aggregate age profile in Figure 2 with substantial within-consumer heterogeneity would imply that this demeaned 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 relatively flat. The pattern shown in this figure 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.

The average individual's slope shown in the top-right panel of Figure 3 could mask a stepping-stone effect if habits are formed very quickly during early adulthood and then remain fairly constant. This could generate an age profile that is steep at the beginning and then flattens out quickly. In this case the average slope across all individuals would be small, since most individuals in our sample would be in the flat part of their life-cycle profile even though the age profile is steep at the beginning. In the bottom-left panel we assess this hypothesis plotting the demeaned age profile of individuals starting from age 18 and following them up to at most age 24. That is, we perform the same analysis as in the top-right panel of Figure 3 on this subsample, again controlling for individual fixed effects and normalizing the initial share to zero, which is now the share at age 18. The bottom-left panel of Figure 3 shows that there indeed is a steeper age profile from age 18 to about age 22.

The bottom-right panel of Figure 3 repeats this exercise, now following individuals starting at age 25 through at most age 29. We observe that the age profile already becomes flat when consumers are in their late 20s. In fact, the profiles are so flat that we cannot reject the hypothesis that the slope of the two age profiles for beer and vodka are the same. Figure A.1 in the online appendix performs the same analysis over the entire life-cycle, showing that the age profiles remain flat at all ages above age 22, such that the two slopes of the age profiles of beer and vodka shares are not statistically different from each other.

Hence, we find that within-consumer variation such as the stepping-stone effect cannot explain most of the average age profile shown in the top-left panel of Figure 3. Instead, most of the profile is driven by between-consumer variation such as cohort effects which is consistent with the flat cohort profiles shown in Figure 1. Moreover, this non-parametric analysis also reveals that consumer preferences form early in the life-cycle until the age of about 22, presumably at the beginning of an individual's consumption life-cycle. That is, habits form mostly in the first couple of years after the individual starts consuming alcohol for the first time. In all our subsequent analyses, we therefore control for age flexibly with fixed effects to account for potential life-cycle effects such as stepping-stone effects. We will use two quasi-natural experiments to formally test the hypothesis that habits form early in life and then remain relatively stable throughout the remaining part of the life-cycle. Most of these shocks occur well before our sample period, and our results are not affected when we exclude individuals from our sample und age 22, which is consistent with this hypothesis.

4 A Model of Persistent Consumption Habits

The previous analysis suggests that persistent habits might play an important role in explaining the observed consumer behavior given that individual fixed effects explain the vast majority of the age profile in alcohol shares. In this section we derive a basic model of habit formation that both is consistent with the consumption patterns observed so far and will guide our empirical analysis in the next section.

Most of the empirical literature focuses on habit formation in the short run. Relatively short expenditure panels 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. Studies of short-run effects, however, cannot answer several questions of interest for both researchers as well as policy makers. For instance, what are the long-run effects on life expectancy of current prices, regulations, and the taxation of such habit-forming goods? Will an increase in the price of hard alcohol induce consumers to switch to healthier drinks, or do other equilibria with different levels of consumption exist?

Our model illustrates that in a situation where people consume two habit-forming goods, several steady-state consumption patterns are possible even in the absence of any unobserved individual heterogeneity. Whether a person will end up conforming to a steady state depends solely on his initial consumption pattern. Moreover, once the stock of habit sufficiently accumulates, it is hard to change these consumption patterns even with very large shocks. Hence, policies aimed at increasing the relative price of one good may not induce everybody or even many to reduce the consumption of this good. Instead, due to the stock of habits already accumulated, people who are accustomed to this particular good will still prefer it even after the policy change. This implies that policies that influence the initial choices of younger generations can have consequences over their entire lifetime—intended or otherwise.

For simplicity we assume that consumers spend all of their budget on two habit-forming goods, beer and vodka. We also assume that consumers are myopic, i.e., that they maximize only current utility and do not save, that there are no outside goods, that income does not change over time, and that there is no uncertainty.[21]

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

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

The budget constraint is

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

Without loss of generality, we focus on interior solutions.[22] The first-order condition of this optimization problem is

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

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

$$u_v(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) - p_v u_b(v,\ y - p_v v,\ v/\delta,\ (y - p_v v)/\delta) = 0\ . \tag{4}$$

In general, this is a non-monotonic function in the steady-state vodka consumption v.[23] Depending on the parametrization of the utility function u, equation (4) may have a different number of solutions. Figure A.2 in the online appendix illustrates that. for certain parametrizations, there is a unique solution, but for many other parametrizations several steady states exist, up to a continuum of solutions.[24] These multiple equilibria are derived without any consumer heterogeneity except for

[21] In the online appendix, we reach the same qualitative conclusions if consumers are forward looking and solve a fully dynamic problem.

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

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

[24] See the appendix for a proof. Similar results are obtained for the model with forward-looking consumers because

differences in initial conditions. A person who initially consumes primarily beer will also prefer beer in the long-run steady state, and vice versa for vodka.

5 Identifying Persistent Consumption Habits

In this section we first verify that the patterns shown in Figure 1 are robust to including various controls. Using two quasi-natural experiments, we then formally test our conjecture that changes in alcohol consumption are mainly driven by cohort differences against various alternative hypotheses, and that those cohort differences in turn are caused by shocks to the initial conditions of young consumers when they start to form their habits. Finally, we use the quasi-experimental variation to relate our results to the previous literature that studies habit formation in the short run using micro data. To that end, we estimate a first-order autoregressive process as is standard in this literature. We show that the least-squares estimate is severely biased toward zero and that the underlying process is highly persistent. Comparing ordinary least-squares (OLS) and instrumental variable (IV) estimates, we provide a lower bound of this downward bias that can be informative in other settings where such exogenous variation is not readily available.

5.1 Accounting for Income and Relative Price Effects

To analyze the robustness of the cohort effects shown in Figure 1, we estimate the following reduced-form regression of the share of alcohol S^g_{it} consumed by individual i in year t of alcohol of type $g \in \{\text{b: beer}, \text{v: vodka}\}$ by OLS,

$$S^g_{it} = \text{10-year-cohort}_i + \gamma' x_{it} + \alpha_a + \alpha_t + \alpha_r + \epsilon_{it}, \quad g \in \{b, v\}. \tag{5}$$

10-year-cohort_i are the same ten-year cohort fixed effects as in Figure 1; α_a, α_t, and α_r are age, time and region fixed effects, respectively, controlling flexibly for life-cycle patterns and relative price effects. The vector of controls x_{it} includes most importantly the level of total alcohol intake and log-income as well as a standard set of demographics such as personal health status, weight, education, and marital status.[25]

[Table 2 about here]

In columns 1 and 7 of Table 2, we estimate the cohort effects shown in Figure 1 for beer and vodka consumption shares, respectively, using cohorts born before 1930 as the reference group. Consistent with the pattern shown in Figure 1, we find that cohorts turning 18 during the Soviet

the steady-state Euler equation is also non-monotonic in the consumption levels.

[25] Due to the restrictions we impose on the cohort effects, the model of age, time, and cohort effects is identified.

Union consume similar shares of vodka, which in turn are significantly higher than those of younger cohorts. Moreover, there is a clear difference in consumption behavior between individuals born in the 1970s—who spent their early adulthood during the economic transition and experienced the opening of Russian markets to foreign competition—and both older and younger cohorts. In turn, the consumption patterns of younger cohorts who turned 18 in the 21th century is again fairly similar and significantly different from all previous cohorts. The share of beer follows the opposite pattern, which increases with the birth year. These patterns are slightly stronger for vodka consumption since the beer market still expanded rapidly during the first half of our sample and hence the beer consumption pattern of many consumers we observe has not stabilized yet. To facilitate the interpretation of the results, we therefore use the cohorts who turned 18 during the Soviet Union as the reference group in all other columns. Columns 3 and 9 control for various socioeconomic demographics, most importantly household income and total alcohol intake. A higher level of total alcohol intake is achieved by consuming relatively more vodka, while having a higher share of beer is associated with a higher level of self-reported health.[26] Body weight does not affect the share of beer but is a significant predictor of the share of vodka. Both beer and vodka shares have a positive income effect due to the fact that homemade vodka (i.e., samogon) is a low-quality substitute for purchased vodka and hence is an inferior good.[27] The evidence in the literature on the relationship between alcohol consumption patterns and education and marital status is mixed and typically uses levels of total alcohol consumed instead of shares by types of alcohol; hence, these coefficients are difficult to interpret. Not surprisingly, adding time effects in columns 4 and 10 significantly changes the coefficients for real income and for having a college education, given that those two variables exhibit an upward trend during the sample period. Time and region fixed effects also flexibly control for other unobserved factors such as current relative prices and macroeconomic shocks. Columns 5 and 11 further add age effects in order to non-parametrically control for any stepping-stone effects. Finally, columns 6 and 12 assess the robustness of the results to dropping the top quartile of the distribution of total alcohol consumers. Overall, Table 2 clearly shows that younger generations tend to consume more beer and less vodka even after controlling for all those factors, and that the results are not driven by heavy drinkers. For instance, column 3 shows that the average share of beer in total alcohol intake is 44 percentage points (pp) higher for males born in the 1990s than for those born before 1970. Even those born in the 1980s have on average a 32 pp higher share of beer consumption than those born before 1970. Those born in the 1970s in turn have a 15 pp higher share of beer consumption than those born earlier.

[26]Recall that the subjective health-status variable equals 1 if the individual feels very healthy and 5 if he is in poor health.

[27]The results for samogon are provided in Table A.1 of the online appendix.

Our hypothesis is that while 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 can explain the patterns observed in Figures 1, 2, and 3. However, there are two main alternative hypotheses to consider that might also explain these patterns. First, individuals born in different time periods might have different preferences for certain types of alcohol because of slowly evolving unobserved factors such as culture or social norms 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 males might start consuming beer or other light drinks before eventually switching to harder drinks later in life. In this section we use two quasi-natural experiments to test our hypothesis against those alternatives, thereby providing more-direct evidence for the habit formation process modeled in Section 4 than the reduced-form results provided in Section 3. Furthermore, these quasi-experiments allow us to estimate the age window in which most individuals form their habits.

In the following we restrict our sample to years 2001–2011, since starting with year 2001 all cohort groups reach a new steady state as documented in Figure 1. The cohort profiles between 1994 and 2000 are compressed by the fact that individuals have only limited access to the beer market. As the beer market expands, all cohorts increase their average beer consumption and decrease their vodka consumption across the board, although the relative ranking of the shares is preserved even in those earlier years. Therefore, when analyzing the long-run effects of the quasi-natural experiments below, we need to restrict our analysis to the stable period after 2000. Otherwise, our analysis would be contaminated by the *current* evolution of the alcohol market instead of capturing only the long-run effects of these changes that occurred prior to our sample period. For instance, comparing the 1990s cohorts with the 1930s cohorts over the entire sample period from 1994 to 2011 would clearly overstate the pure cohort effect since it would attribute the fact that the 1930s cohorts did not have access to the same beer market in 1990s as in the 2000s to cohort rather than time effects.[28]

5.2 Evidence from the Collapse of the Soviet Union

The first alternative explanation for the observed heterogeneity is that individuals born at different times grow up in different cultural environments and might therefore have different preferences for hard and light drinks. To test our hypothesis against this alternative, we exploit the large change in the Russian alcohol market that occurred in the wake of the collapse of the Soviet Union as our first quasi-natural experiment. Focusing on the relatively short period of time when the beer industry experienced rapid growth, we study the long-run consumption behavior of individuals who turn 18

[28]We absorb part of this contamination effect non-parametrically with time fixed effects.

years old during this period, which is the minimum legal drinking age in Russia.[29] Since culture and institutions change only slowly (e.g., Roland (2004)), males who turn 18 during the beer-market expansion and hence according to our hypothesis form their habits in this period face a very similar cultural environment and similar social norms but very different access to beer compared with males who are only slightly older. We estimate the same regression equation as in (5) except that we replace the ten-year cohort effects with the year in which the individual turns 18 to analyze the effect of the rapid changes in the beer market for males of only slightly different ages,

$$S^g_{it} = \beta \cdot \text{year-turned-18}_i + \gamma' x_{it} + \alpha_a + \alpha_t + \alpha_r + \epsilon_{it}. \tag{6}$$

[Figure 4 about here]

The top panel of Figure 4 illustrates the design of the analysis. We start estimating equation (6) on the sample of all males who turn 18 during the expansion of the beer market, which we determine lasted from about 1994 to 2008 based on Figure 2. Since it is possible that other factors also changed during this period that may have affected males differentially depending on the year of their 18th birthday, we let the sample window, which is centered at year 2001, shrink until it only includes the three years from 2000 to 2002. Hence, as we shrink the sample window, we identify the effect of the expansion of the beer market on alcohol shares off of males who grow up in a more and more similar environment, except that they face a different beer market when they turn 18.

The bottom panel of Figure 4 plots the estimates of β for both types of goods together with 95% confidence intervals. The effect of the expansion of the beer market on the shares as a fraction of total alcohol intake is remarkably stable, and it remains statistically significant despite the substantial gradual reduction in the sample size. Moreover, consistent with our hypothesis, the magnitude of the coefficients increases (in absolute value) with shrinking sample periods since we are selecting males who are more and more likely to have formed their consumption habits during the rapid expansion of the beer market.[30] For instance, males who turn 18 in 2002 exhibit on average a 12% higher long-run share of beer consumption compared with males who are only two years older.[31]

[Figure 5 about here]

[29] Since there is no discontinuity implied by the legal drinking age—both because of limited enforceability of the minimum legal drinking age and because one cannot be forced to start consuming alcohol at 18—and also because habits do not necessarily form within a single year, we cannot use a regression discontinuity design. However, our identification approach closely mimics such a framework. Our results suggest that the majority of male consumers form their habits between ages 16 and 18. Since there is severe underreporting of underage drinking in the RLMS we cannot reliably estimate the drinking pattern of males below age 18. The literature surveyed in Koposov et al. (2002) suggests that the mean age at which minors started to binge drink was between 14 and 18 years in the Soviet Union and probably has not change much since then.

[30] The regression results underlying this and all other figures are available from the authors upon request.

[31] The term "long-run share" refers to the fact that we are estimating the individuals' consumption shares using data from 2001 to 2011. Hence, most of the individuals in our sample are (much) older than 18 when we measure their consumption shares.

To analyze the validity of this quasi-natural experiment, we run placebo tests as illustrated in the top panel of Figure 5. Specifically, we estimate equation (6) using a rolling window of 10 years, starting with males who turned 18 between 1970 and 1979 and ending with the sample of males who turned 18 between 2002 and 2011, with 1970 being the first year for which we have official aggregate sales data by type of alcohol. Once we reach the sample ranging from 2002 to 2011, we continue shrinking the window from the left until it only includes the five years from 2007 to 2011.

Under our hypothesis we should not see any significant effect of the year in which an individual turned 18 on the share of beer consumed for samples that do not include the expansion of the beer market. As the 10-year sample window reaches the time at which the beer market expands rapidly, the estimate of β in equation (6) should gradually increase, because men turning 18 at the end of the 10-year sample window have much easier access to beer than men who turned 18 at the beginning of the sample window. Finally, the beer market stabilizes around 2007 at a new long-run equilibrium shown in Figure 2. As the sample window starts to cover more and more of the new steady state, the coefficient should gradually decrease. For the shortest sample which includes only males who turn 18 in 2007 or later, the estimate should be zero, as all individuals in this subsample have access to a similarly developed beer market when they turn 18. To summarize, the response should first be zero and then exhibit a hump-shaped pattern with a peak response when the sample window fully covers the beer-market expansion period.

The bottom panel of Figure 5 plots the estimates of β together with 95% confidence intervals from these placebo regressions. We indeed see this hump-shaped pattern emerge from the data for beer consumption precisely as we would expect under our hypothesis. The coefficients are close to zero and not statistically significant for samples that only include males who turn 18 before the expansion of the beer market. The effect gradually increases when more and more individuals from the 10-year rolling sample are affected by this shock. The peak response is reached for the sample that ranges from 1998 to 2007, which corresponds to the 10-year period that indeed saw the most-dramatic increase in the beer market over the entire 42-year period shown in the top panel. Finally, as we let the sample shrink to include only males who turn 18 after the market stabilizes, we see the coefficient converge to zero, although the estimated precision naturally decreases with the smaller sample sizes. Figure A.3 in the online appendix adds the responses of vodka shares to this figure and shows that the beer-market expansion has the opposite effect on the share of vodka, echoing the findings shown in Figure 4. The response of the share of vodka is also not significantly different from zero for the samples that do not cover the beer-market expansion. Similarly, the vodka share's response peaks in absolute value for the sample that ranges from 1998 to 2007 before gradually converging back to zero. These results together with the evidence shown in Figure 4 suggest that the substitution toward beer mainly comes at the expense of vodka. This in turn has important health consequences to which we return in Section 6.

5.3 Evidence from Gorbachev's Anti-Alcohol Campaign

We use the so-called Gorbachev anti-alcohol campaign as a second quasi-natural experiment to identify consumption habits and to estimate the persistence of this shock to individual preferences. In 1985 Mikhail 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.[32] 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.[33]

Since the communist government directly controlled the production of any official alcohol in the Soviet Union, the effect of Gorbachev's anti-alcohol campaign on official sales of alcohol was dramatic as is evident from Figure 2. Sales of beer dropped by 29%, from 177 million liters of ethanol in 1984 to 125 million liters in 1987.[34] Official sales of vodka dropped by 60%, from 784 million liters in 1984 to 317 million liters in 1987, and wine sales experienced the most dramatic drop, from 292 million liters in 1985 down to only 108 million liters in 1990, a decrease of 63%. During the short period from 1984 to 1988 the ratio of official vodka sales to beer sales dropped by 43%, which in the absence of relative prices is our best approximation of the trade-off that individual consumers faced.[35]

[Figure 6 about here]

However, as shown in the top panel of Figure 6, the drop in official sales of vodka was partially compensated for by the increased production of samogon, a then-illegal low-quality home-produced vodka. As a result, the effect of the Gorbachev anti-alcohol campaign on total vodka consumption including samogon was small on average; see, e.g., Treml (1997), Nemtsov (2002), and Bhattacharya et al. (2013) for a discussion of the underlying data and methodology.[36] Indeed, after accounting

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

[33] See, for example, White (1996), Nemtsov (2002), Bhattacharya et al. (2013), and Figure 6.

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

[35] Similarly, the difference in log-changes between official vodka and beer sales is -57%, that is, vodka sales dropped by 57% more than beer sales from 1984 to 1988.

[36] 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, which is one of the main ingredients in the production of samogon; see, e.g., Nemtsov (1998). This approach gives reliable estimates until 1986 when the production of sugar was rationed. The second approach uses data on violent and accidental deaths and deaths with unclear causes obtained from autopsy reports; e.g., Nemtsov (2002). For such death events there exist measures of alcohol concentration in the blood of the victim that can be used to estimate aggregate alcohol consumption. This

for samogon production, the estimated volume of total alcohol consumed during the Gorbachev anti-alcohol campaign decreased by only 33%.

More important for our identification approach is the fact that the production of samogon was heavily concentrated in rural areas for reasons related to the technology used to produce samogon. First, the production of samogon requires space, which is limited in urban areas, especially in Russian cities, which are very densely populated by international comparison, with most people living in large apartment buildings. Second, producing samogon causes smoke and a strong smell, which is at the same time very unpleasant and also easy to detect by neighbors and law-enforcement agents, particularly 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 single-unit homes, which are highly concentrated in rural areas, than in apartment buildings, which are prevalent in cities. This geographical pattern of samogon production (and consumption) continues to the present even though total samogon production has decreased dramatically since 1992. For instance, males in rural areas drink 5.5 times more samogon and the share of samogon in total alcohol intake is five times higher than in rural areas—13% for rural areas compared with only 2.4% in urban areas according to the RLMS. The bottom panel of Figure 6 shows that accounting for samogon production dramatically changes the ratio of hard alcohol to beer available to consumers. Since rural males have much more access to samogon during the campaign, they see this ratio increase, while urban males face a relative decline in the availability of hard alcohol. One can therefore expect significant differences in the way the campaign affects the initial conditions of rural relative to urban males who turn 18 during the campaign and hence their habit formation in the long run.

This second experiment naturally leads to a difference-in-difference design, since rural men are affected differently by the campaign than urban men. The treatment group are rural men who turn 18 during the campaign. The peak impact of the campaign lasted from 1987 to 1991 as shown in the bottom panel of Figure 6. Based on the assumption that the elasticity of substitution between vodka and samogon is much higher than the elasticity of substitution between beer and either vodka and samogon, we conjecture that some rural males who would have formed habits toward beer in

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 comes from illegal imports as well as illegal production of unregistered alcohol by firms as a form of tax evasion. Unfortunately, both of these approaches cannot distinguish the type of alcoholic good that was produced at home, in particular whether it was homemade beer, wine, or samogon. Samogon, however, is much more popular than homemade beer. This is largely because homemade beer requires ingredients that do not grow naturally in Russia. Thus, according to the RLMS, for years 2008–201,1 only 0.3% of male alcohol consumers consumed homemade beer compared to 6.2% who consumed samogon, with 2008 being the first year respondents were asked about the consumption of homemade beer. In terms of pure alcohol, these numbers are even more striking: consumption of samogon is 69 times higher than homemade-beer consumption when measured in terms of pure alcohol.

the absence of the campaign substitute to samogon consumption, which is relatively abundant in rural areas during the campaign, and thus form a habit for hard alcohol. For urban males, samogon was much harder to obtain, and hence there were fewer who substituted beer with samogon during the campaign. Therefore, our model has two main predictions. First, it predicts that rural men who turned 18 during the campaign have a higher share of vodka consumption in our baseline sample from 2001 to 2011 relative to rural men who turned 18 both before or after the campaign.[37] Second, the difference between the vodka shares consumed by rural and urban males should be largest for those cohorts who turned 18 during the anti-alcohol campaign.

In this difference-in-difference approach, we implicitly exploit the fact that labor mobility is very low in Russia compared to most other countries; e.g., Andrienko and Guriev (2004). Hence, the chance that the birth place of a survey respondent in our sample also identifies his location when he turned 18—something we do not observe directly in the data—is very high.[38] To test the predictions of our theory for the long-run effect of the anti-alcohol campaign on the consumption shares, we estimate the following regression:

$$\begin{aligned} S_{it}^{g} = \; & \beta_{DD} \cdot I(\text{urban})_i \times I(\text{turned 18 in 1987-91})_i + \beta_D \cdot I(\text{turned 18 in 1987-91})_i \\ & + \lambda \cdot I(\text{urban})_i + \gamma' x_{it} + \alpha_a + \alpha_t + \alpha_r + \epsilon_{it}. \end{aligned} \tag{7a}$$

In our baseline specification we restrict the sample to individuals who turned 18 in 1970 or later because official data on aggregate sales by type of alcohol is available only starting from 1970. Ideally, we would also like to observe a sufficiently long period after the end of the Gorbachev anti-alcohol campaign in which there are no further disruptions to the alcohol market. Figure 2 shows that there was only a brief period between the end of the campaign's impact on the alcohol market and the beginning of the rapid expansions of the beer market after the collapse of the Soviet Union. In all our specifications we therefore restrict our sample to individuals who turned 18 before 1999 to avoid a contamination of this experiment with the shock to the beer market analyzed in the previous section. One might be worried that individuals turning 18 after the end of the campaign's impact face different initial conditions than individuals who turn 18 before the campaign, and that our approach does not fully address that issue. To address this concern we extend the difference-in-difference design of equation (7a) to include two different sets of "control groups," one containing

[37] According to different expert estimates, samogon production increased rapidly in the second half of the 1980s; e.g., Treml (1997), Nemtsov (2002), Bhattacharya et al. (2013), and our own estimates based on the RLMS. Since the collapse of the Soviet Union, samogon production has decreased rapidly because of the liberalization of the alcohol markets and the sharp decrease in the price and increased availability of vodka.

[38] Our proxy for whether an individual lived in a urban area when turning 18 combines the variables "birth place" and whether the individual currently lives in a major city. Specifically, we set the proxy equal to 1 if the birth place is a city instead of a town or village. We then use our measure of whether an individual currently lives in a big city to impute the remaining missing values. With the exception of Sochi and Tolyatti, all big cities in Russia correspond to the regional capital cities, and neither of these two exceptions is part of the RLMS sample frame.

individuals who turn 18 between 1970 and 1986 and another with individuals turning 18 between 1992 and 1998,

$$\begin{aligned} S^g_{it} = \; & \beta_{DD,1} \cdot I(\text{urban})_i I(\text{turned 18 before 1987})_i + \beta_{DD,2} \cdot I(\text{urban})_i I(\text{turned 18 after 1991})_i \\ & + \beta_{D,1} \cdot I(\text{turned 18 before 1987})_i + \beta_{D,2} \cdot I(\text{turned 18 after 1991})_i \\ & + \lambda \cdot I(\text{urban})_i + \gamma' x_{it} + \alpha_a + \alpha_t + \alpha_r + \epsilon_{it}. \end{aligned} \tag{7b}$$

[Table 3 about here]

Column 1 of Table 3, which estimates (7a) using the baseline sample, shows that the "treatment" indicator—i.e., whether an individual turned 18 during the campaign—predicts a 5 pp higher share of vodka consumption relative to rural males who turned 18 before or after the campaign, consistent with our hypothesis. Moreover, the difference in the vodka shares of individuals turning 18 in rural relative to those in urban areas is 7 pp larger in absolute values for individuals turning 18 during the campaign than for individuals turning 18 outside of this period. This reflects the differential impact the campaign had on rural and urban males, again consistent with the hypothesis of highly persistent habit formation combined with the differential impact of the campaign on the alcohol market shown in Figure 6. Column 2 estimates (7b) to assess whether using two comparison groups yields different results. While we cannot reject the hypothesis that the response is the same using the two control groups (again in absolute values), the larger point estimate for the group of males turning 18 before the campaign suggest that this group might be a more appropriate comparison group than the sample of males turning 18 after the campaign.

The coefficients on the control variables in all specifications are similar in sign and magnitude to those reported in Table 2. The positive coefficient of the urban fixed effect in specification (7a) reflects the fact that rural males consume a much larger share of samogon than their urban counterparts. This difference is large enough to turn the effect positive despite the fact that the share of total vodka consumption including samogon is on average 7 pp higher for rural than for urban males. The fact that the sign of this coefficient changes between columns 1 and 2 is driven by the large size of the difference-in-difference coefficient.[39] Columns 3 and 4, which drop the top quartile of total alcohol consumers, show that these results are not driven by heavy drinkers or alcoholism. Columns 5 and 6 extend the sample to include individuals who turned 18 before 1970 and hence before we have data on aggregate sales by type of alcohol. The results are again similar to the baseline specification. Columns 7 and 8 extend the baseline sample to include all years from 1994 to 2011. While the coefficients are again not statistically different from the baseline results, the lower point estimates suggest that using the earlier part of the sample leads to a downward bias of the coefficient, since the individuals' consumption shares have not reached a stable equilibrium

[39] To see this, note that λ of (7b) minus λ of (7a) should approximately equal β_{DD} of (7a).

yet due to the ongoing expansion of the beer market, as seen in Figure 1. Finally, columns 9 and 10 show that the campaign has the opposite effect on the share of beer consumed, again suggesting that the main substitution occurs between vodka and beer.

Overall, the results in Table 3 show that the campaign significantly changed long-run consumption behavior given that most subjects in our sample are observed two decades after the end of the campaign. Moreover, the results highlight the differential impact the campaign had on individuals who turned 18 in rural areas relative to their urban peers. Consistent with our hypothesis, these individuals formed persistent habits toward very different types of alcohol, and these differences in preferences are still highly visible in their consumption behavior today.

[Figure 7 about here]

Next we again analyze the validity of our quasi-natural experiment by running placebo tests similar to the approach taken in Figure 5 and which is illustrated in Figure A.4 in the online appendix. Specifically, we estimate equation (7a) on a rolling 15-year window starting with individuals who turned 18 between 1960 and 1974 and ending with the sample of males turning 18 between 1985 and 1999. The "treatment group" is the five-year window centered in this 15-year rolling sample, covering years 1975 to 1979 in the first sample window. Once we reach the sample ranging from 1985 to 1999 we again shrink the window from the left until it only includes the years from 1990 to 1999, implying that the control group turned 18 between 1990 and 1994 and the treatment group turned 18 between 1995 and 1999. Under our hypothesis we should not see any significant effects before the sample enters the anti-alcohol campaign. As the sample enters the campaign period, we should first see β_{DD} increase as the true treatment group gets mistakenly assigned to the control group. The coefficient should then gradually decrease as the assigned treatment group more and more overlaps with the actual treatment group, reaching its maximum (in absolute value) with the group turning 18 between 1987 and 1991. The coefficient should then increase back to zero, before becoming positive again as we falsely assign the actual treatment group to the control group. Finally, the coefficient should gradually decrease back to zero although it will not converge to zero completely under our hypothesis since we have to restrict our sample to individuals turning 18 before 1999. Hence, the pattern for β_{DD} should be W-shaped. β_D on the other hand should exhibit the opposite pattern, i.e., M-shaped.

Figure 7 plots the evolution of β_D and β_{DD} together with 95% confidence intervals for these placebo regressions. Consistent with the hypothesis of persistent habit formation around age 18, we indeed see these two patterns emerge, W-shaped for β_{DD} and M-shaped for β_D, although the shape of the latter is weaker. The peak response of both coefficients (in absolute value) occurs when the treatment window reaches the actual treatment period from 1987 to 1991.[40]

[40]Figure A.5 in the online appendix shows similar results with the opposite sign for the share of beer, although

5.4 Short-Run Dynamics of Habit Formation

In this section we use the two quasi-experiments as instruments to estimate the short-run dynamics of consumption behavior in the presence of habit formation. For this purpose we estimate the following first-order autoregressive process—which is the most common specification used in the literature—first by OLS and then by IV:

$$S_{it}^{g} = \rho \cdot S_{i,t-1}^{g} + \lambda \cdot I(\text{urban})_i + \gamma' x_{it} + \alpha_a + \alpha_t + \alpha_r + \epsilon_{it}. \tag{7}$$

To deal with the potential upward bias due to autocorrelated unobserved taste shocks, which might simultaneously drive current and lagged shares, and to address attenuation bias due to measurement error in the lagged consumption shares, we instrument the lagged share with the condition of the alcohol market that individuals faced when they first started to consume alcohol around age 18. For urban males, we use the ratio of aggregate sales of official vodka to beer when they turn 18 as a proxy of the market conditions they faced. For rural males, we instead use the ratio of all aggregate vodka to beer when they turn 18, including samogon. In the set of instruments we also include both aggregate ratios not interacted with the individual's geographical information when turning 18. This specification estimates the effect of habits under the assumption that individuals born in different periods have similar preferences and differ only in the initial level of consumption, holding fixed income, relative prices and demographic characteristics. Moreover, we assume that aggregate sales are uncorrelated with any unobserved individual taste heterogeneity.

[Table 4 about here]

Table 4 shows a severe downward bias of the OLS estimate of the autoregressive coefficient ρ relative to its IV estimate for both beer and vodka. The IV estimate for beer is three times larger than its OLS estimate, and the IV estimate for vodka is more than two and a half times larger than its OLS estimate. In the online appendix we use these results to provide a lower bound for the attenuation bias due to measurement error. The lower bound equals one minus the ratio of the OLS to the IV estimator and equals one if there is complete attenuation and 0 if there is no attenuation bias. We estimate this lower bound to be about two-thirds for both beer and vodka. Therefore, significant measurement error in expenditure surveys might explain why previous research has found little evidence of habit formation at the micro level when estimating specifications similar to (7); e.g., Dynan (2000).

with less precision.

6 Effects on Life Expectancy

Russian male life expectancy at birth was on average only 60 years between 2000 and 2009, which is 15 years lower than in the US, 7 years lower than in Bangladesh and even 4 years lower than in North Korea. Moreover, the gender gap in life expectancy over the same period was 13 years in Russia, but only 5 years in the US, one year in Bangladesh, and 7 years in North Korea.[41] One key difference between these countries and Russia is the prevalence of heavy drinking among men. The high level of alcohol consumption among Russian men is therefore widely believed to be a main contributing factor to the low male life expectancy and the large gender gap; see, e.g., Brainerd and Cutler (2005), Leon et al. (2007) and Yakovlev (2012). Approximately 40% of all annual deaths are estimated to be related to alcohol consumption, and most of them are not due to long-run consequences of heavy drinking such as cirrhosis but instead are due to so-called "dose-related excess," a hazardous event occurring when the amount of pure alcohol consumed is very high. Indeed, the literature finds that about 7% of deaths are due to alcohol poisoning, while over 30% are due to "external causes" related to alcohol intoxication, including vehicular and other accidents and homicides, and hence are unrelated to long-run consequences of alcohol consumption. Moreover, alcohol poisoning is typically not caused by the poor quality of the alcohol consumed but is rather caused by binge drinking.[42] Hence, while a high average level of alcohol intake can certainly be hazardous—in particular for older individuals—it is mostly the occasional binge drinking by men that leads to high mortality rates across all age groups. Furthermore, since binge drinking is much less likely to occur when consuming beer rather than vodka, we conjecture that individuals who prefer beer over vodka have a lower alcohol-related probability of dying, even holding fixed the average level of alcohol intake.

To test this conjecture and to quantify the effect of alcohol consumption habits toward different types of alcohol on the probability of dying, we estimate a semi-parametric Cox proportional hazard model as is standard in the literature.[43] We use a similar specification as in our previous analysis with two modifications. First, we add two additional explanatory variables to the vector x that

[41]Sources: The Human Mortality Database, `http://www.mortality.org/`, and The World Bank, `http://data.worldbank.org/`.

[42]Estimates of the effect of alcohol on mortality vary somewhat and are difficult to compare across studies due to differences in methodology and in the underlying data. However, most studies find similar magnitudes and broadly agree with official statistics, which attribute 4% of all deaths of Russian males to alcohol poisoning and an additional 36% to external causes related to alcohol intoxication; see Goskomstat, Demograficheskiy yezhegodnik RF, 2006. Similarly, in their sample of 48,557 residents of three typical Russian cities, Zaridze et al. (2009) find that 8% of deaths are directly due to alcohol poisoning, while another 37% are due to accidents and violent acts that are related to alcohol intoxication.

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

improve the fit of the model. The first indicates whether an individual reports not drinking in a typical day during the previous month ,while the second is an indicator of whether the individual smokes. Second, we collapse the data to one observation per individual, and we replace time-varying covariates with their mean. For individuals who report not consuming alcohol in a given interview, we set their shares of beer and vodka to zero before collapsing the data. We impose two additional sample restrictions relative to our previous analysis. First, our preferred specification focuses on working-age males. This excludes males older than 65 years due to selection bias arising from the fact that older individuals tend to consume a lower share of vodka (see the top-left panel of Figure 3), both because of endogenous attrition of heavy drinkers from the sample and because they might not be able to consume hard alcohol anymore due to medical conditions such as liver failure. Second, we exclude individuals below age 22 since our estimate of the long-run consequences of the changed consumption habits on male mortality crucially depend on them as they approximate the consumption behavior of the population in the new long-run steady state. The analysis in section 5.1 reveals that the habits of males below age 22 have not yet converged to their long-run equilibrium; hence, their observed consumption shares are not a good predictor of their future shares.

[Table 5 about here]

Table 5 shows that consuming a higher share of beer strongly decreases the hazard of death, while a higher share of vodka increases it, even conditional on the total level of alcohol consumed.[44] Hence, these results are consistent with the findings in the previous literature that a majority of alcohol-related deaths are due to alcohol poisoning or external causes in connection with binge drinking. The estimates are also economically significant: decreasing the share of vodka by 30 pp (which corresponds to one standard deviation in the sample) while simultaneously increasing the share of beer by 30 pp—holding fixed the level of total alcohol intake—decreases the hazard of death for males age 22 to 65 by 34%. The level of total alcohol intake on its own also increases the hazard of death. The fact that these estimates are larger and statistically more significant for the sample including older males points to the negative long-run consequences of alcohol consumption. The remaining controls have the expected sign with the exception of the indicator of whether an individual reported not drinking in all interviews. However, this effect is not stable across specifications and samples, is statistically not significant, and is likely affected by the aforementioned selection bias, particularly at older ages. Finally, columns 4 to 6 show that the main results are similar if we include these older individuals.

[44]The effect of both shares can be identified simultaneously as shown in columns 3 and 6 of Table 5, because the two shares are not perfectly collinear—the sample correlation is -0.42—due to the presence of nondrinkers and due to the fact that there is a small share of other types of alcohol consumed such as wine, which we exclude from the regressions. We obtain similar coefficients for the shares consumed when estimating the model separately for beer (columns 1 and 4) and vodka (columns 2 and 5) compared with the coefficients when using both shares jointly (columns 3 and 6), which we will use in the simulations below.

Our results seem to be in sharp contrast to the common belief that increased alcohol consumption after the end of the anti-alcohol campaign and after the liberalization of the alcohol market caused the surge in male mortality since 1991—the so-called Russian mortality crisis. Our results however apply to the shares of alcohol consumed, not the level of total alcohol intake. Our hypothesis is that because it is easier to binge drink with vodka than beer, forming a habit toward vodka increases an individuals' mortality risk, even holding fixed his level of total alcohol intake. Indeed, consistent with our individual-level results, we find that when we regress male mortality for ages 22 to 65 on aggregate sales per capita of beer and vodka—both expressed in pure alcohol—, vodka is highly predictive, while beer has no explanatory power.[45]

As a final step we study the implications of the changes in consumption patterns for the evolution of male mortality. To do so we simulate a counterfactual scenario that maintains the sample distribution of individual characteristics except for the shares of vodka and beer consumed. Specifically, we predict consumption shares of vodka and beer for each individual in our sample by running regressions similar to equation (5), except that we include a full set of cohort effects.[46] Using these cohort effects, we predict each individual's shares at different points in the future and in turn use the predicted shares together with the individual characteristics to estimate his hazard of death. For example, to predict the hazard of death in 10 years of an individual born in 1970, we maintain his current characteristics but we assign him the conditional cohort effect of individuals born in 1960. Integrating across the entire sample then provides us with an estimate of the evolution of male mortality as a consequence of the changes in consumption habits only.[47]

[Table 6 about here]

Table 6 provides the predicted population consumption shares and the annual rate of death for the current population of males age 22 to 65 as well as for the corresponding counterfactual populations in 10, 20, and 55 years, with 55 years being the time at which the population reaches its new steady state.[48] Our results suggest that the mortality of males age 22 to 65 will decrease by 12% from 1.48% to 1.31% over the next 10 years, by 23% over 20 years, and will be cut roughly in half in the long run. The predicted current rate of death of 1.48% is only slightly higher than its official estimated average from 1994 to 2010, which is 1.57%. For comparison, the annual rate of death is 0.5% in the US and 0.4% in the UK and Germany . Hence, the counterfactual simulation predicts that the increase in the share of beer consumption at the expense of vodka as suggested by the persistent habits we find in the data combined with the large changes to the alcohol market

[45] See Table A.2 and Figure A.6 in the online appendix.

[46] To identify the model, we exclude time effects.

[47] Online appendix A.3 provides more detail for this algorithm.

[48] Figure A.7 in the online appendix shows the path of both shares and the mortality rate.

that occurred in the past might further cut the gap between the Russian and US male mortality in half over the next 55 years.

References

Alesina, Alberto and Nicola Fuchs-Schündeln, "Good-Bye Lenin (or Not?): The Effect of Communism on People's Preferences," *American Economic Review*, 2007, *97* (4), 1507–1528.

Andrienko, Yuri and Sergei Guriev, "Determinants of interregional mobility in Russia," *Economics of transition*, 2004, *12* (1), 1–27.

Atkin, David, "Trade, Tastes, and Nutrition in India," *American Economic Review*, 2013, *103* (5), 1629–63.

Becker, Gary S and Kevin M Murphy, "A theory of rational addiction," *Journal of Political Economy*, 1988, pp. 675–700.

_ , Michael Grossman, and Kevin M Murphy, "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 1994, pp. 396–418.

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

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

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

Bronnenberg, Bart J, Jean-Pierre H Dubé, and Matthew Gentzkow, "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 2012, *102* (6), 2472–2508.

Campbell, John Y and John H Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107* (2), 205–251.

Chaloupka, Frank, "Rational Addictive Behavior and Cigarette Smoking," *Journal of Political Economy*, 1991, *99* (4).

Cook, Philip J. and Michael J. Moore, "Alcohol," in A. J. Culyer and J. P. Newhouse, eds., *Handbook of Health Economics*, Vol. 1, Elsevier, 2000, chapter 30, pp. 1629–1673.

Denisova, Irina, Markus Eller, and Ekaterina Zhuravskaya, "What do Russians think about transition? 1," *Economics of Transition*, 2010, *18* (2), 249–280.

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

Dynan, Karen E, "Habit formation in consumer preferences: Evidence from panel data," *American Economic Review*, 2000, pp. 391–406.

Eichenbaum, Martin S, Lars Peter Hansen, and Kenneth J Singleton, "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty," *Quarterly Journal of Economics*, 1988, *103* (1), 51–78.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales, "The Role of Social Capital in Financial Development," *American Economic Review*, 2004, *94* (3), 526–556.

_ , _ , and _ , "Trusting the stock market," *Journal of Finance*, 2008, *63* (6), 2557–2600.

Heien, Dale and Cathy Durham, "A test of the habit formation hypothesis using household data," *Review of Economics and Statistics*, 1991, pp. 189–199.

Koposov, Roman A, Vladislav V Ruchkin, Martin Eisemann, and Pavel I Sidorov, "Alcohol use in adolescents from northern Russia: the role of the social context," *Alcohol and alcoholism*, 2002, *37* (3), 297–303.

Leon, David A, Lyudmila Saburova, Susannah Tomkins, Evgueny Andreev, Nikolay Kiryanov, Martin McKee, and Vladimir M Shkolnikov, "Hazardous alcohol drinking and premature mortality in Russia: a population based case-control study," *The Lancet*, 2007, *369* (9578), 2001–2009.

Malmendier, U. and S. Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *Quarterly Journal of Economics*, April 2011, *126* (1), 373–416.

Malmendier, Ulrike and Stefan Nagel, "Learning from inflation experiences," *WorkingPaper*, 2011.

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

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

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

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

Ravina, Enrichetta, "Habit Formation and Keeping Up with the Joneses: Evidence from Micro Data," *Working Paper*, 2007.

Ravn, Morten, Stephanie Schmitt-Grohé, and Martin Uribe, "Deep habits," *Review of Economic Studies*, 2006, *73* (1), 195–218.

Roland, Gerard, "Understanding institutional change: fast-moving and slow-moving institutions," *Studies in Comparative International Development*, 2004, *38* (4), 109–131.

Sundaresan, Suresh M, "Intertemporally dependent preferences and the volatility of consumption and wealth," *Review of financial Studies*, 1989, *2* (1), 73–89.

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

_ , "Soviet and Russian statistics on alcohol consumption and abuse," *Premature death in the new independent states*, 1997, pp. 220–238.

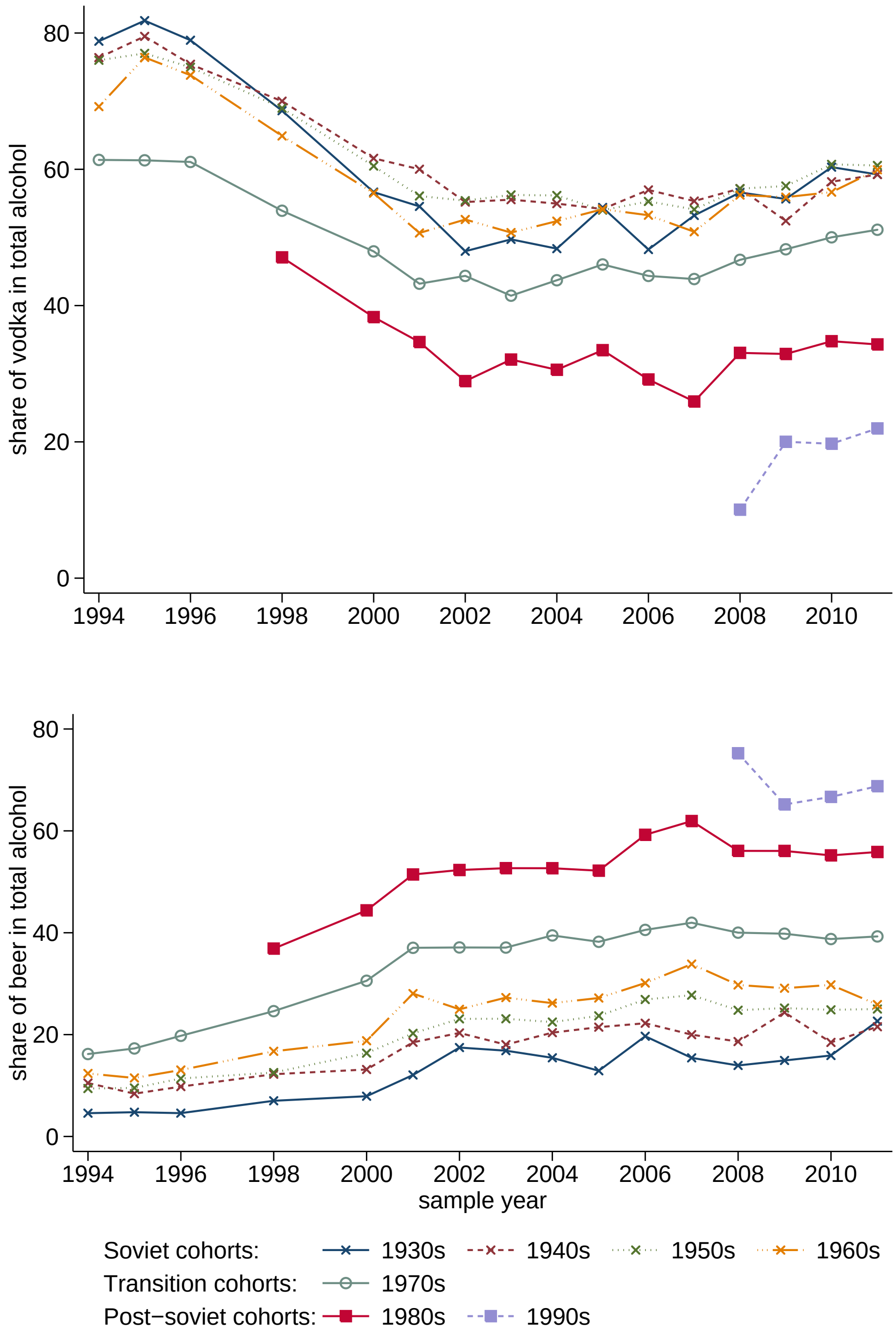
White, Stephen, *Russia goes dry: alcohol, state and society*, Cambridge University Press, 1996.

Williams, Jenny, "Habit formation and college students' demand for alcohol," *Health Economics*, 2005, *14* (2), 119–134.

Yakovlev, Evgeny, "Peers and Alcohol: Evidence from Russia," *Working Paper*, 2012.

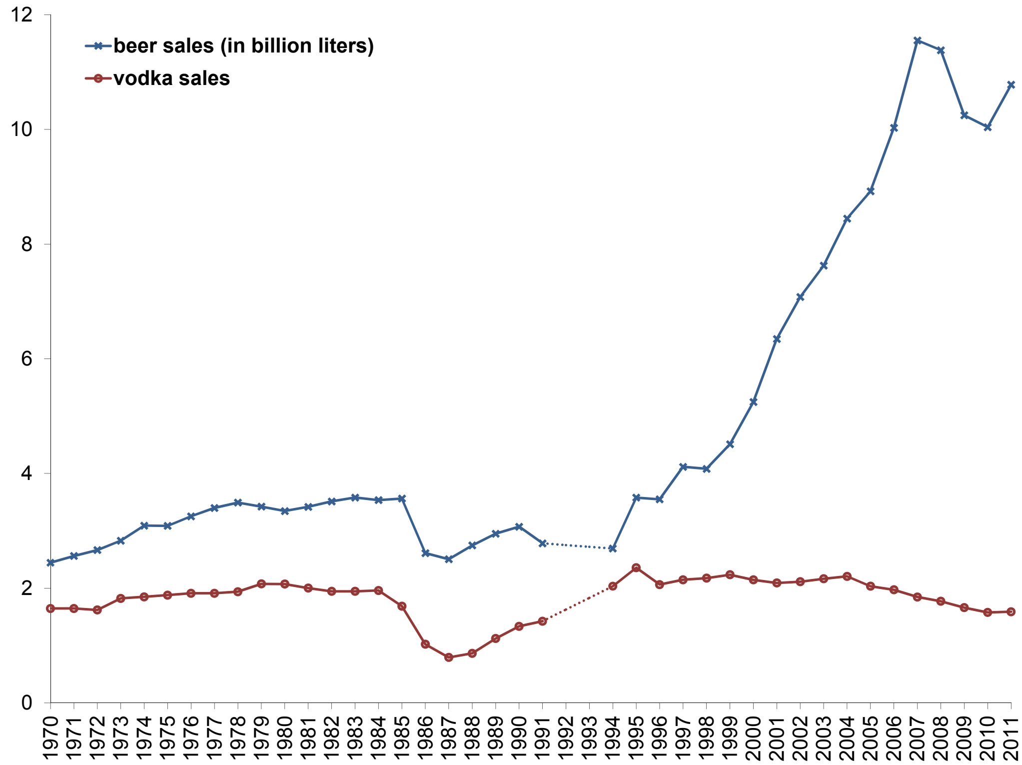
Zaridze, David, Paul Brennan, Jillian Boreham, Alex Boroda, Rostislav Karpov, Alexander Lazarev, Irina Konobeevskaya, Vladimir Igitov, Tatiana Terechova, Paolo Boffetta et al., "Alcohol and cause-specific mortality in Russia: a retrospective case–control study of 48 557 adult deaths," *The Lancet*, 2009, *373* (9682), 2201–2214.

Figure 1: Share of vodka and beer in total alcohol intake by birth cohorts



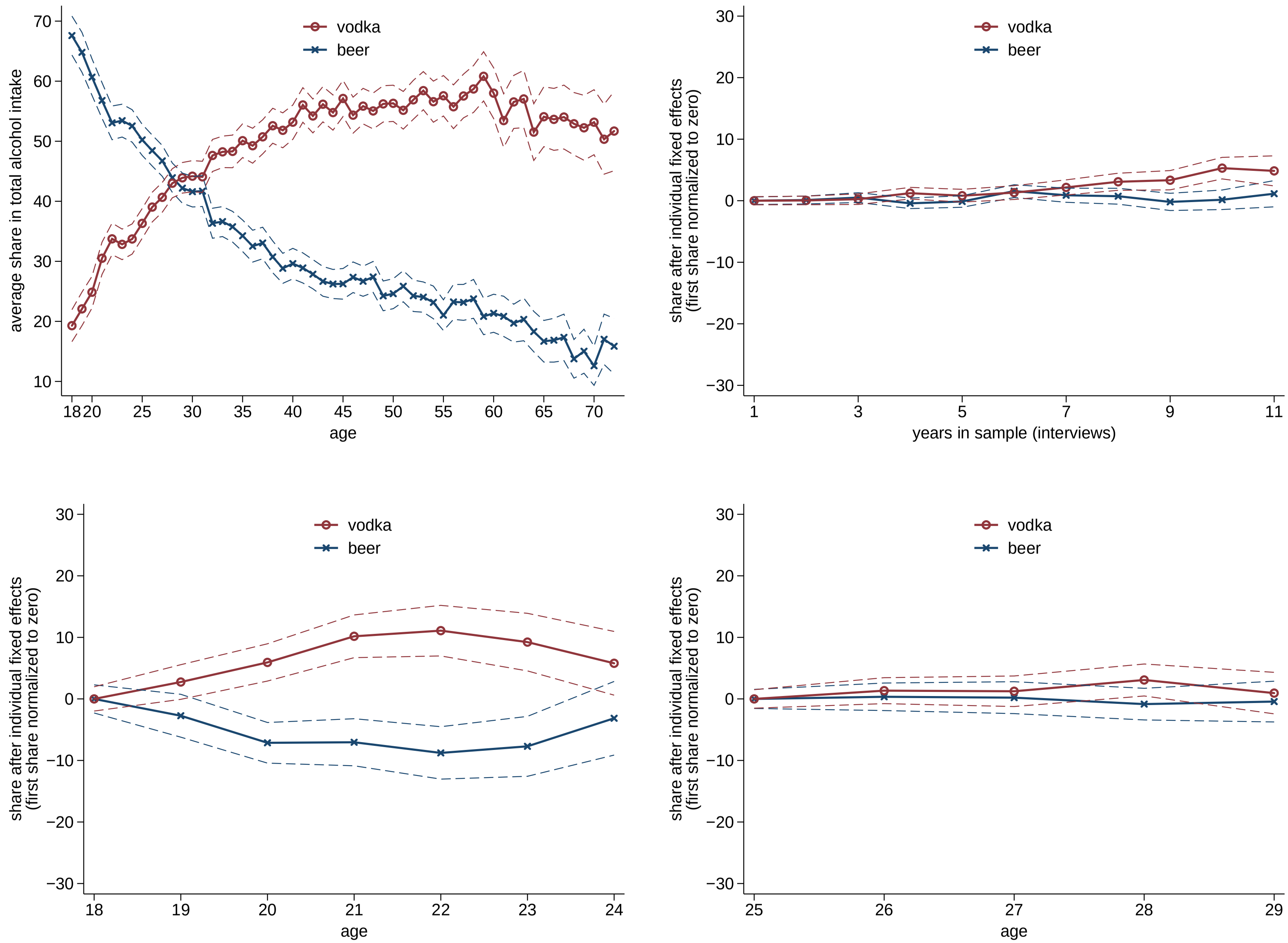
Notes: These figures show the share of vodka (top panel) and beer (bottom panel) as a function of the RLMS sample year for 10-year birth cohorts. Cohorts are classified based on the average time in which their members turned 18, which is our estimate of when people on average start to consume alcohol and to form habits under our hypothesis.

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



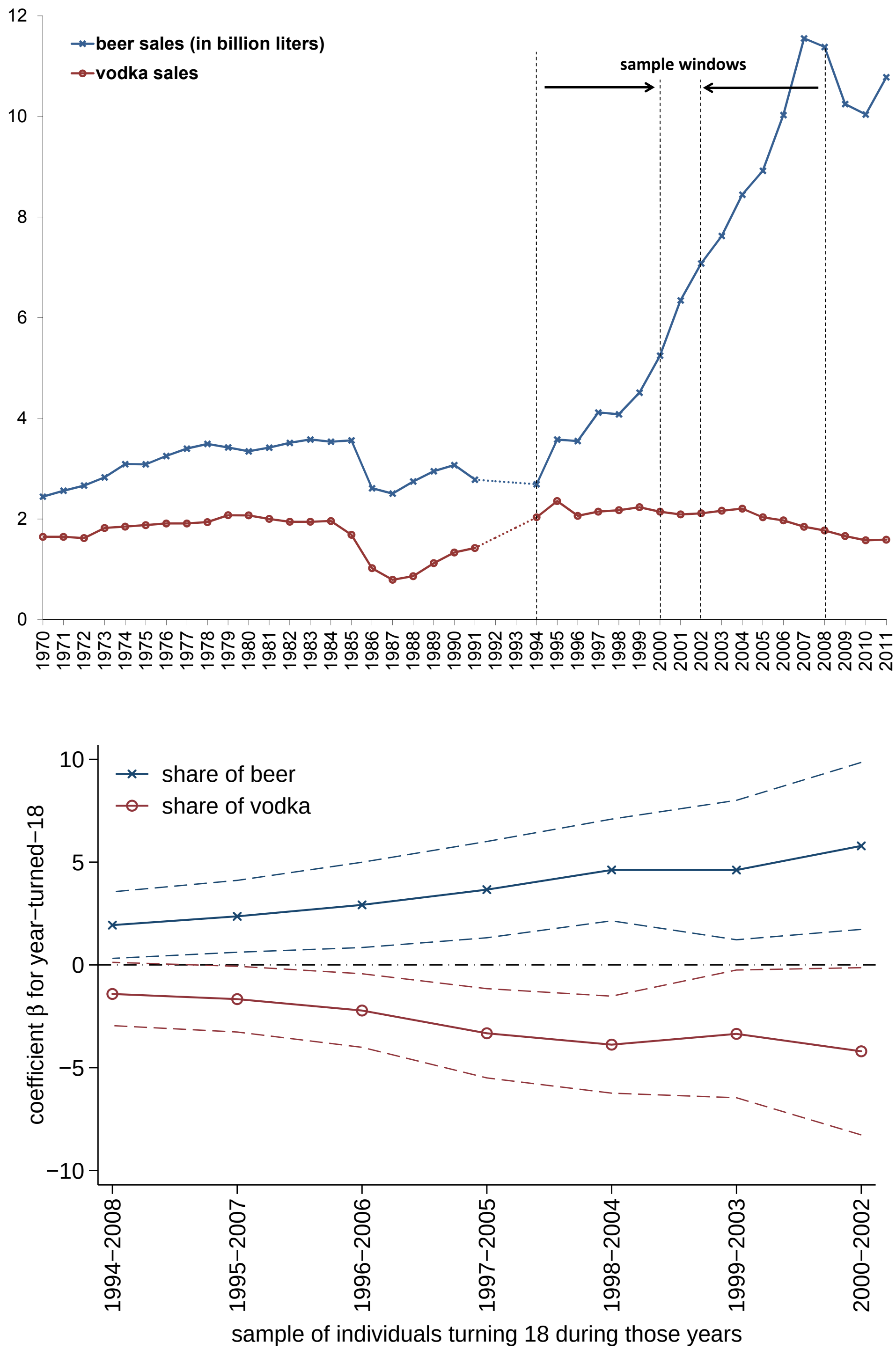
Notes: This figure shows the aggregate sales of beer and vodka in liters (not in ethanol). Data in 1992 and 1993 are not reliable and are excluded due to significant data-collection and reporting problems in the wake of the collapse of the Soviet Union, leading to severely underreported levels for all types of alcohol.

Figure 3: Shares by age and demeaned shares by age and years in sample



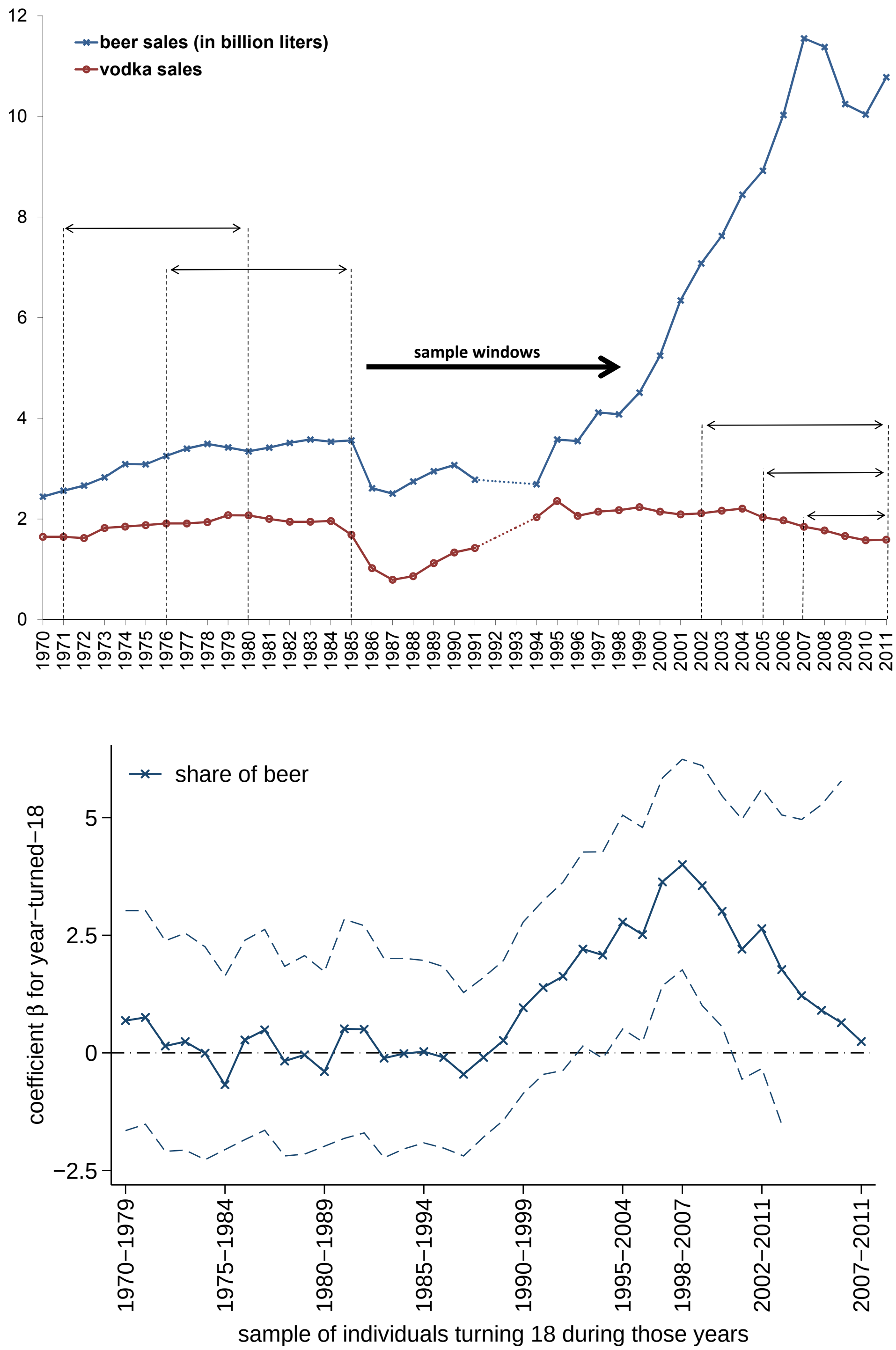
Notes: These figures show the life-cycle profiles of the shares of beer and vodka consumed. The dashed lines represent two standard error confidence intervals using data from 1994 to 2011. The top-left panel shows the sample age profile. The top-right panel shows the consumption shares against the number of years an individual is observed in the sample, after controlling for individual fixed effects. The two bottom panels show the age profile for the two subgroups of individuals age 18 to 24 and 25 to 29 as a function of age, again after controlling for individual fixed effects. Figure A.1 in the online appendix provides similar results for five-year age-sample intervals from age 30 to 74.

Figure 4: Effect of the beer-market expansion on consumption habits



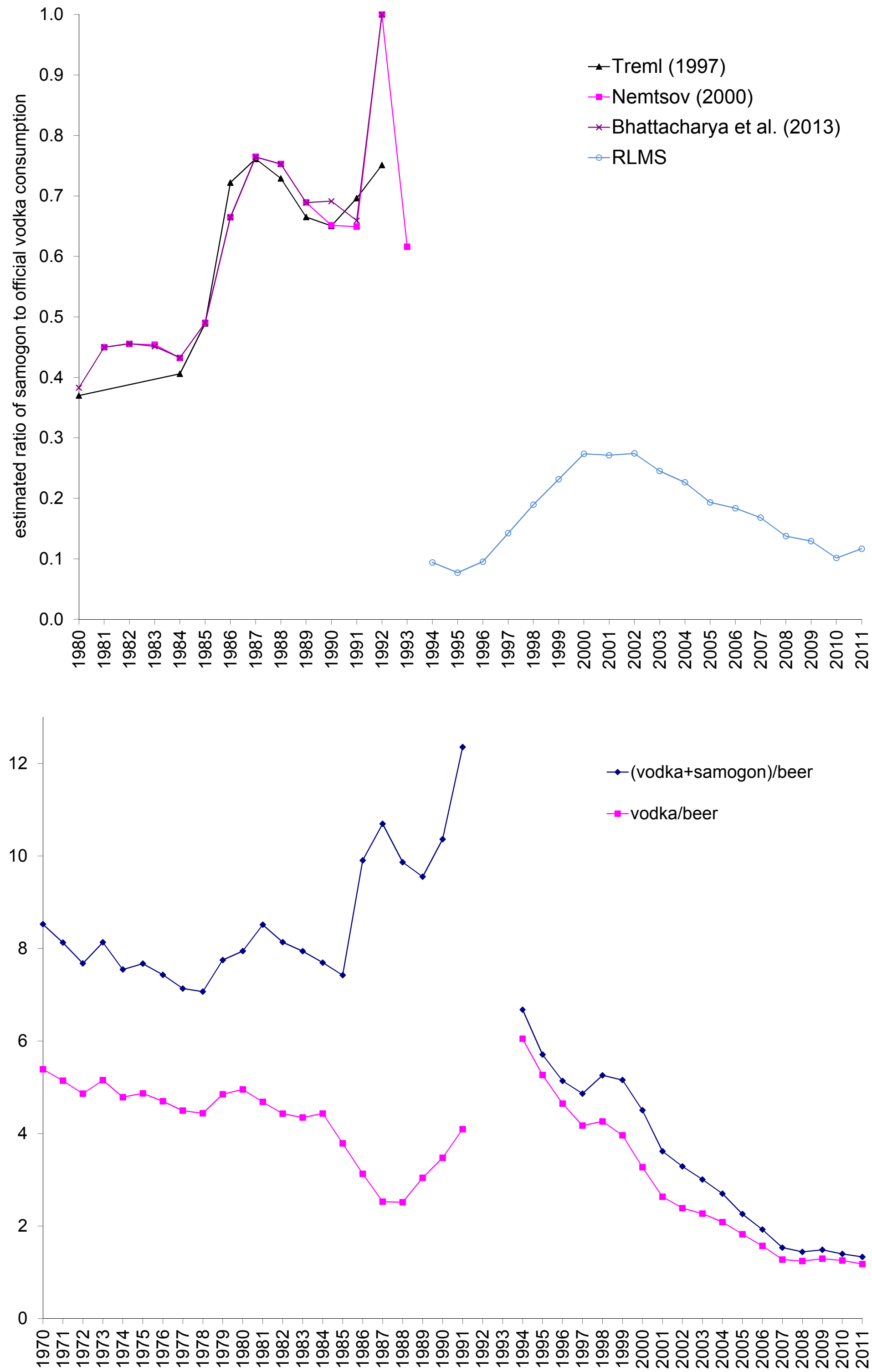
Notes: These figures show the regression design (top panel) and the regression results together with two standard error confidence bands (bottom panel) of the beer-market expansion analysis. The regressions control for the level of total alcohol intake, log of real income, subjective health status, body weight, education, marital status, and a full set of year, age, and region fixed effects. HAC standard errors are clustered by individual.

Figure 5: Placebo regressions for the beer-market expansion experiment



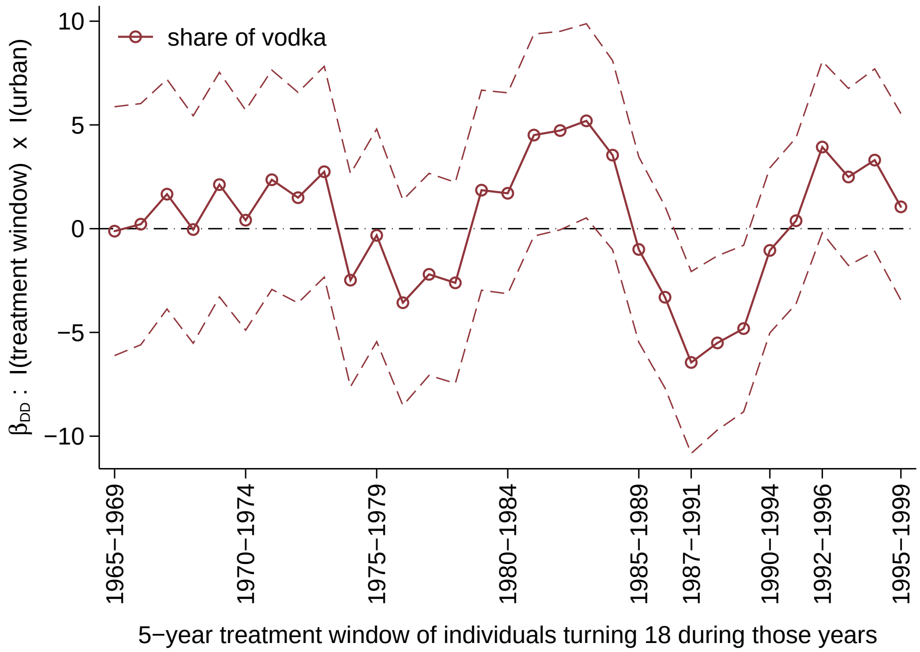
Notes: These figures show the design of the placebo regressions (top panel) and the regression results together with two standard error confidence bands (bottom panel). The regressions use the same set of controls as in Figure 4. HAC standard errors are clustered by individual.

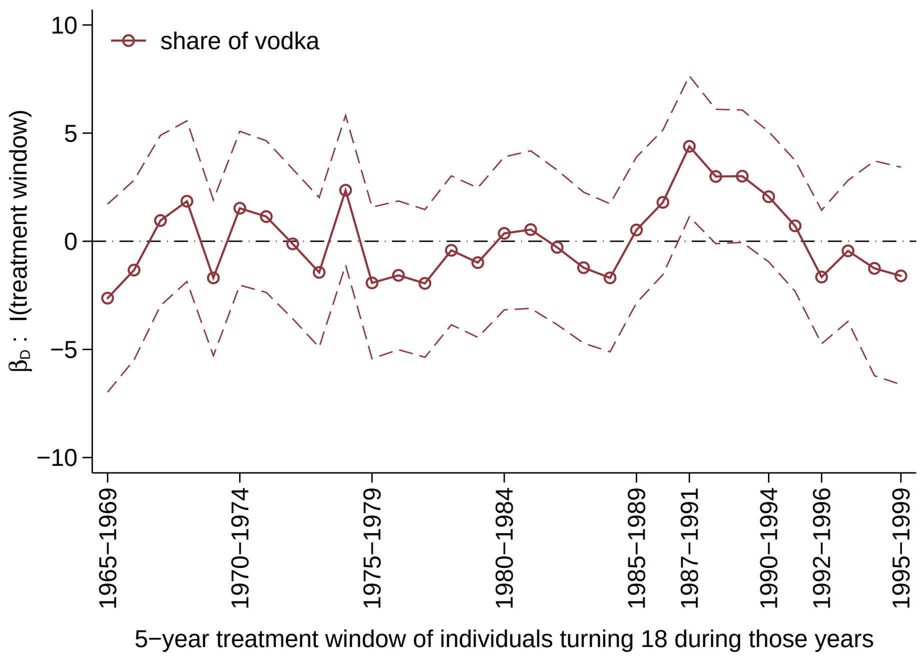
Figure 6: Relative alcohol production in Russia, 1970–2011



Notes: The top panel shows estimates of the ratio of illegally produced vodka (i.e., samogon) to officially produced vodka using four different data sources. The bottom panel displays the ratio of official vodka production to beer production and the ratio of total vodka to official beer production, all measured in grams of ethanol. Years 1992 and 1993 in the bottom panel are excluded; see footnote 5.

Figure 7: Placebo regressions for anti-alcohol campaign experiment





Notes: These figures show the results of the placebo regressions for the anti-alcohol campaign experiment together with two standard error confidence bands. The regressions use the same set of controls as in Figure 4. HAC standard errors are clustered by individual.

Table 1: Summary statistics

Males with positive alcohol intake	N	Mean	St.Dev.	p75
Share of beer	46985	29.30	35.34	38.46
Share of vodka	46985	52.86	39.70	92.31
Share of samogon	46985	8.66	24.29	0
Share of wine	46985	7.37	20.86	0
Share of other alcohol	46985	1.81	10.87	0
Share of home-brewed beer (starts in 2008)	14363	0.06	1.47	0
Age	46985	41.38	15.43	52
Birth year	46985	1962.41	16.35	1976
Total monthly real income (in liters of milk)	45280	245.06	426.55	300.28
Subjective health status (1=very good, 5=very bad)	46884	2.69	0.67	3
Body weight (in kg)	44180	76.67	13.68	85
I(married)	46985	0.68	0.46	1
I(college degree)	46950	0.40	0.49	1
Proxy for I(turned 18 in an urban area)	46972	0.45	0.50	1
Typical daily alcohol intake (in grams of ethanol)	46985	144.69	133.75	200
All males age 18 and above				
Age	68350	42.45	16.39	54
Birth year	68350	1961.33	17.27	1975
Total monthly real income (in liters of milk)	65688	233.70	404.16	288.28
Subjective health status (1=very good, 5=very bad)	68186	2.73	0.73	3
Body weight (in kg)	64114	76.47	13.62	85
I(married)	68350	0.67	0.47	1
I(college degree)	68290	0.38	0.49	1
Proxy for I(turned 18 in an urban area)	68322	0.44	0.50	1
Typical daily alcohol intake (in grams of ethanol)	68350	101.40	131.95	146

Notes: This table presents the number of observations, means, standard deviation, and the 75th percentile of key variables in the full sample of males above age 18 and for the subsample of males that report having consumed alcohol in the 30 days prior to the interview.

Table 2: Decomposition of the shares of vodka and beer

	(1) vodka	(2) vodka	(3) vodka	(4) vodka	(5) vodka	(6) vodka	(7) beer	(8) beer	(9) beer	(10) beer	(11) beer	(12) beer
I(born in 1990s)	-46.242***	-41.030***	-36.906***	-31.116***	-11.704***	-6.946***	57.876***	48.408***	44.026***	38.086***	20.720***	19.302***
	[2.386]	[1.470]	[1.493]	[1.616]	[2.342]	[2.544]	[2.026]	[1.778]	[1.720]	[1.743]	[2.266]	[2.659]
I(born in 1980s)	-33.561***	-28.350***	-25.764***	-21.173***	-7.597***	-7.401***	45.508***	36.041***	32.230***	27.502***	15.176***	17.173***
	[2.008]	[0.705]	[0.749]	[0.799]	[1.532]	[1.851]	[1.192]	[0.693]	[0.708]	[0.742]	[1.302]	[1.743]
I(born in 1970s)	-17.746***	-12.534***	-12.208***	-9.774***	-1.691	-3.027**	25.760***	16.292***	14.879***	13.125***	4.459***	6.693***
	[2.002]	[0.688]	[0.689]	[0.661]	[1.068]	[1.345]	[1.148]	[0.614]	[0.595]	[0.576]	[0.872]	[1.232]
I(born in 1960s)	-7.819***						14.496***					
	[2.022]						[1.118]					
I(born in 1950s)	-4.551**						10.571***					
	[2.048]						[1.125]					
I(born in 1940s)	-3.536						6.833***					
	[2.151]						[1.198]					
I(born in 1930s)	-4.408*						1.606					
	[2.254]						[1.191]					
Alcohol intake (in grams of ethanol)			0.040***	0.038***	0.039***	0.457***			-0.064***	-0.065***	-0.068***	-0.403***
			[0.002]	[0.002]	[0.002]	[0.007]			[0.002]	[0.002]	[0.002]	[0.007]
Log(real income)			0.103	0.493***	0.371***	0.146			1.264***	0.444***	0.615***	0.474***
			[0.115]	[0.124]	[0.126]	[0.157]			[0.091]	[0.102]	[0.102]	[0.142]
Health status			0.013	1.002***	0.081	0.401			-2.203***	-2.824***	-0.893***	-1.475***
			[0.380]	[0.356]	[0.366]	[0.441]			[0.288]	[0.281]	[0.286]	[0.389]
Body weight (in kg)			0.129***	0.131***	0.130***	0.089***			-0.000	0.000	-0.017	-0.009
			[0.021]	[0.020]	[0.020]	[0.025]			[0.017]	[0.016]	[0.016]	[0.022]
I(college degree)			-5.455***	0.732	0.677	2.646***			1.959***	-2.677***	-2.755***	-5.842***
			[0.466]	[0.673]	[0.672]	[0.756]			[0.396]	[0.596]	[0.588]	[0.732]
I(married)			1.010*	4.236***	3.521***	3.419***			-0.161	-2.068***	-1.931***	-2.891***
			[0.566]	[0.606]	[0.622]	[0.746]			[0.450]	[0.514]	[0.515]	[0.688]
Observations	46,985	46,985	44,066	44,029	44,029	27,353	46,985	46,985	44,066	44,029	44,029	27,353
R-squared	0.077	0.075	0.103	0.169	0.175	0.324	0.163	0.152	0.222	0.255	0.270	0.375
Region FE				YES	YES	YES				YES	YES	YES
Year FE				YES	YES	YES				YES	YES	YES
Age FE					YES	YES					YES	YES
Top alcohol quartile dropped						YES						YES

Notes: Heteroskedasticity- and autocorrelation-consistent (HAC) standard errors, clustered by individual, in parentheses.

Table 3: Effect of Gorbachev's anti-alcohol campaign on consumption habits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Baseline		Top quartile dropped		All turned 18 before 1991		Full sample 1994-2011		Baseline	
	vodka	vodka	vodka	vodka	vodka	vodka	vodka	vodka	beer	beer
I(urban)*I(turned 18 in 1987-1991)	-6.542***		-6.097***		-7.109***		-5.139***		3.861**	
	[2.065]		[2.093]		[2.031]		[1.742]		[1.787]	
I(turned 18 in 1987-1991)	4.779***		5.565***		5.071***		3.437***		-3.480***	
	[1.530]		[1.620]		[1.518]		[1.296]		[1.266]	
I(urban)*I(turned 18 before 1987)		7.357***		6.947***		7.847***		5.670***		-4.400**
		[2.200]		[2.255]		[2.097]		[1.781]		[1.871]
I(urban)*I(turned 18 after 1991)		5.226**		4.585**		5.408**		3.572*		-3.045
		[2.271]		[2.289]		[2.292]		[2.030]		[2.027]
I(turned 18 before 1987)		-4.957**		-6.892***		-4.374**		-2.414		3.232**
		[2.039]		[2.240]		[1.932]		[1.542]		[1.646]
I(turned 18 after 1991)		-4.365**		-4.178**		-5.126***		-3.688**		3.525**
		[1.900]		[1.992]		[1.861]		[1.543]		[1.658]
Alcohol intake (in grams of ethanol)	0.054***	0.054***	0.306***	0.306***	0.050***	0.050***	0.029***	0.029***	-0.093***	-0.093***
	[0.004]	[0.004]	[0.007]	[0.007]	[0.003]	[0.003]	[0.002]	[0.002]	[0.004]	[0.004]
Log(real income)	0.554***	0.549***	0.541**	0.537**	0.772***	0.768***	0.323**	0.318**	0.808***	0.813***
	[0.194]	[0.194]	[0.222]	[0.221]	[0.189]	[0.189]	[0.152]	[0.152]	[0.155]	[0.155]
Health status	-0.234	-0.234	-0.397	-0.394	-0.153	-0.152	0.205	0.210	-0.710	-0.709
	[0.562]	[0.562]	[0.614]	[0.613]	[0.503]	[0.503]	[0.414]	[0.414]	[0.469]	[0.468]
Body weight (in kg)	0.145***	0.145***	0.107***	0.106***	0.146***	0.146***	0.124***	0.124***	-0.021	-0.020
	[0.028]	[0.028]	[0.031]	[0.031]	[0.025]	[0.025]	[0.022]	[0.022]	[0.024]	[0.024]
I(college degree)	0.036	0.015	0.636	0.593	0.597	0.560	0.835	0.845	-2.646***	-2.635***
	[0.924]	[0.923]	[0.968]	[0.968]	[0.835]	[0.835]	[0.833]	[0.832]	[0.806]	[0.806]
I(married)	5.027***	5.020***	5.074***	5.088***	3.916***	3.891***	3.475***	3.418***	-2.744***	-2.731***
	[0.893]	[0.892]	[0.975]	[0.975]	[0.829]	[0.829]	[0.724]	[0.724]	[0.755]	[0.756]
I(urban)	2.624**	-3.936**	2.548**	-3.526*	1.791*	-5.382***	0.536	-4.657***	0.099	3.990**
	[1.043]	[1.981]	[1.115]	[2.009]	[0.930]	[1.978]	[0.770]	[1.695]	[0.842]	[1.712]
Observations	19,373	19,373	15,250	15,250	25,528	25,528	35,351	35,351	19,373	19,373
R-squared	0.100	0.101	0.235	0.235	0.099	0.099	0.128	0.128	0.179	0.179
Region FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Age FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Top quartile dropped			YES	YES						

Notes: HAC standard errors, clustered by individual, in parentheses.

Table 4: Short-run dynamics of consumption habits

	(1) OLS beer	(2) IV rural beer	(3) IV urban beer	(4) IVs combined beer	(5) 1st stage beer	(6) Full sample beer	(7) OLS vodka	(8) IV rural vodka	(9) IV urban vodka	(10) IVs combined vodka	(11) 1st stage vodka	(12) Full sample vodka
Lagged share of beer	0.280***	0.884***	0.807***	0.832***		0.815***						
	[0.008]	[0.143]	[0.189]	[0.122]		[0.067]						
Lagged share of vodka							0.327***	0.750***	0.874***	0.849***		0.950***
							[0.008]	[0.149]	[0.119]	[0.102]		[0.113]
I(urban) * (aggregate vodka/beer when 18)					0.250						2.208***	
					[0.610]						[0.691]	
Aggregate vodka/beer when 18					-1.271**						-0.035	
					[0.522]						[0.588]	
I(rural) * (aggregate vodka+samogon/beer when 18)					-0.500*						-0.164	
					[0.284]						[0.314]	
Aggregate vodka+samogon/beer when 18					-0.533**						0.724**	
					[0.265]						[0.294]	
Alcohol intake (in grams of ethanol)	-0.090***	-0.070***	-0.072***	-0.072***	-0.033***	-0.054***	0.054***	0.048***	0.046***	0.046***	0.014***	0.030***
	[0.003]	[0.006]	[0.007]	[0.005]	[0.002]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]	[0.003]
Log(real income)	0.304**	-0.171	-0.110	-0.130	0.793***	0.039	0.669***	0.389**	0.307*	0.324**	0.669***	0.313**
	[0.129]	[0.176]	[0.198]	[0.163]	[0.152]	[0.123]	[0.154]	[0.175]	[0.167]	[0.162]	[0.176]	[0.143]
Health status	-0.339	0.401	0.307	0.338	-1.219***	0.096	-0.367	-0.541	-0.592	-0.582	0.427	-0.287
	[0.358]	[0.403]	[0.419]	[0.384]	[0.417]	[0.303]	[0.422]	[0.385]	[0.393]	[0.390]	[0.500]	[0.361]
Body weight (in kg)	-0.015	-0.001	-0.003	-0.002	-0.024	-0.009	0.108***	0.047*	0.029	0.032	0.141***	0.030*
	[0.018]	[0.014]	[0.014]	[0.013]	[0.022]	[0.011]	[0.020]	[0.026]	[0.022]	[0.020]	[0.026]	[0.018]
I(college degree)	-1.998***	-1.318***	-1.405***	-1.376***	-1.182	-1.105**	0.226	0.166	0.148	0.152	0.300	-0.336
	[0.575]	[0.497]	[0.516]	[0.486]	[0.722]	[0.453]	[0.636]	[0.493]	[0.492]	[0.490]	[0.827]	[0.502]
I(married)	-1.759***	-0.690	-0.827	-0.782	-1.631**	-0.525	2.993***	1.277	0.775	0.878	4.060***	0.059
	[0.563]	[0.530]	[0.570]	[0.508]	[0.707]	[0.400]	[0.648]	[0.779]	[0.697]	[0.648]	[0.844]	[0.622]
I(urban)	0.592	-0.335	-0.216	-0.256	-3.313	-0.084	0.837	0.653	0.599	0.610	-10.470***	0.544
	[0.586]	[0.505]	[0.532]	[0.483]	[2.482]	[0.390]	[0.696]	[0.492]	[0.477]	[0.477]	[2.545]	[0.426]
Observations	23,801	23,801	23,801	23,801	23,801	29,357	23,792	23,792	23,792	23,792	23,792	29,347
R-squared	0.324	-0.005	0.068	0.045	0.193	0.030	0.230	0.033	-0.077	-0.052	0.125	-0.188
Region FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Age FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
F-test	.	17.61	9.197	12.57	.	38.91	.	11.73	19.70	13.27	.	11.80
J-test, p-value				0.746		0.564				0.532		0.129

Notes: This table shows the results from estimating equation (7) using RLMS sample years 2001 to 2011. Columns 1 and 7 provide OLS estimates. Columns 2 and 8 instrument the lagged consumption shares with the ratio of official aggregate vodka to beer sales when the individuals turned 18. Columns 3 and 8 use total vodka to beer including samogon as an instrument. Columns 4 and 10 use both instruments and report J-test, the p-value for over-identification. F-test reports the p-value for the joint significant of the first-stage coefficients. Columns 5 and 11 provide the first-stage coefficients using both IVs. Columns 6 and 12 use all sample years from 1994 to 2011 and both IVs. HAC standard errors, clustered by individual, in parentheses.

Table 5: Cox proportional hazard model for male mortality

	(1)	(2)	(3)	(4)	(5)	(6)
	Males age 22-65			Males age 22 and older		
Share of beer (not in percentage)	-1.240***		-1.002***	-1.257***		-1.191***
	[0.359]		[0.379]	[0.289]		[0.296]
Share of vodka (not in percentage)		0.575***	0.378**		0.269**	0.125
		[0.160]	[0.166]		[0.106]	[0.106]
I(smokes)	0.435***	0.444***	0.440***	0.464***	0.478***	0.468***
	[0.131]	[0.131]	[0.131]	[0.091]	[0.091]	[0.091]
I(no alcohol consumed)	-0.133	0.137	0.002	-0.114	0.071	-0.061
	[0.192]	[0.192]	[0.193]	[0.140]	[0.140]	[0.140]
Alcohol intake (in kg of ethanol)	0.748	1.129**	0.862	1.195***	1.536***	1.236***
	[0.551]	[0.515]	[0.541]	[0.444]	[0.426]	[0.442]
Log(real income)	-0.462***	-0.495***	-0.471***	-0.508***	-0.537***	-0.511***
	[0.035]	[0.034]	[0.035]	[0.032]	[0.032]	[0.033]
Health status	0.431***	0.473***	0.449***	0.477***	0.510***	0.482***
	[0.109]	[0.110]	[0.109]	[0.073]	[0.073]	[0.073]
Body weight (in g)	4.298	3.862	4.077	4.515	3.902	4.361
	[4.127]	[4.138]	[4.125]	[3.030]	[3.071]	[3.044]
I(college degree)	-1.479***	-1.467***	-1.439***	-1.297***	-1.322***	-1.288***
	[0.240]	[0.246]	[0.242]	[0.184]	[0.187]	[0.185]
I(married)	-0.137	-0.189	-0.166	0.160	0.146	0.156
	[0.143]	[0.143]	[0.143]	[0.106]	[0.107]	[0.106]
I(urban)	0.017	-0.048	-0.015	0.178**	0.127	0.166*
	[0.111]	[0.110]	[0.112]	[0.085]	[0.086]	[0.086]
Observations	8,298	8,298	8,298	9,689	9,689	9,689

Notes: Heteroskedasticity-consistent standard errors in parentheses.

Table 6: Counterfactual simulations

	Poplulation shares of		Mortality rate of males
horizon	beer	vodka	age 22-65 (in %)
current	29.86	47.39	1.48
in 10 years	41.08	33.66	1.31
in 20 years	49.04	24.29	1.14
long run	57.07	16.34	0.81

Notes: This table provides predicted population consumption shares of beer and vodka and the corresponding annual death rates for males age 22 to 65 implied by those shares for forecast horizons 0 (i.e., the current state), 10, 20, and 55 years, which represents the new long-run steady state.

Online Appendix

'How Persistent Are Habits? Micro-Evidence from Russia's Alcohol Market'

Lorenz Kueng

Evgeny Yakovlev

A.1 Additional Results for the Model

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

A.1.1 Number of Equilibria in the Model with Myopic Consumers

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

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

Differentiating F with respect to v yields

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

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

A.1.2 Numerical Examples

A.1.2.1 One Equilibrium

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

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

Solving for v we obtain

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

A.1.2.2 Three Equilibria

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

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

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

A.1.2.3 Continuum of Equilibria

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

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

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

A.1.3 Expressing the Model Solutions in Terms of Shares

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

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

A.1.4 Allowing for Forward-Looking Consumers

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

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

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

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

The FOC for v_{t+1} is

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

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

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

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

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

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

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

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

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

A.2 A Lower Bound for the OLS Attenuation Bias

Let $b_{me} \geq 0$ denote the attenuation factor due to measurement error and $b_{uh} \geq 0$ the potential upward bias of the OLS estimator due to unobserved heterogeneity. Without loss of generality, we assume that the true coefficient is positive, i.e., $\rho > 0$. By definition, the total bias of the OLS estimator is

$$\rho_{OLS} = (1 - b_{me})(1 + b_{uh})\rho \;\geq\; (1 - b_{me})\rho \ .$$

Since the IV estimator is consistent and using the continuous mapping theorem, a consistent estimator $\underline{b_{me}}$ of the lower bound of b_{me} is given by

$$\widehat{\underline{b_{me}}} \equiv 1 - \frac{\hat{\rho}_{OLS}}{\hat{\rho}_{IV}} \;\xrightarrow{p}\; 1 - \frac{\rho_{OLS}}{\rho} \leq b_{me} \ .$$

A.3 Algorithm for Predicting Male Mortality Rates

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

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

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

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

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

1. First, we predict the consumption shares by assuming that the same individual, with characteristics x_i and age 30, also represents a 30 year old next year, but with the consumption habit of a 30 year old *next year*, i.e., with φ_{1970} conditional on the covariates above, that is

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

   Table A.3 provides the regression results for this step.

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

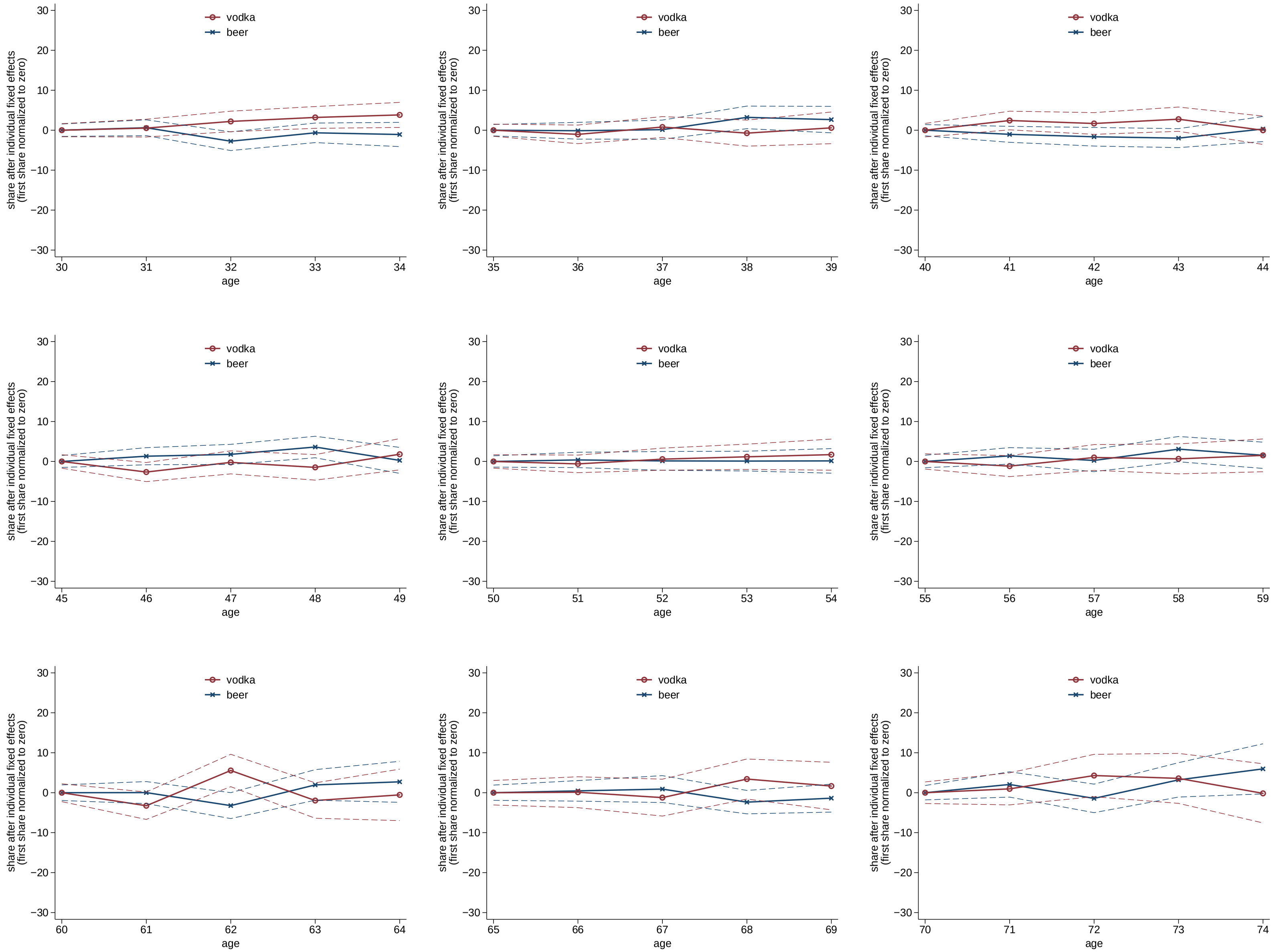
$$\hat{\lambda}_i|_{H=1} = \lambda(a = 30|x_i, \hat{S}_i^g|_{H=1}; \hat{\delta}, \hat{\vartheta}).$$

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

References

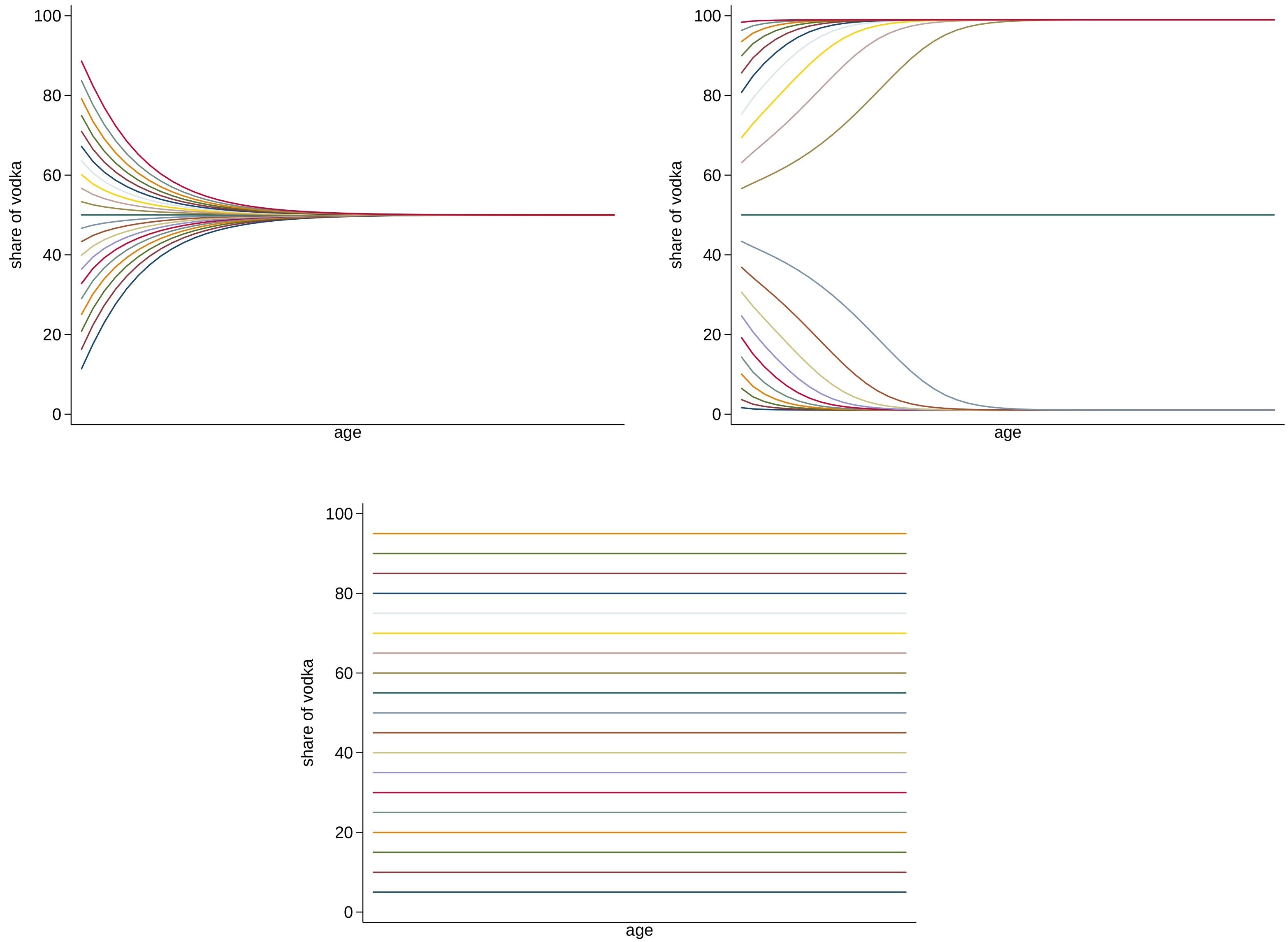
Gruber, Jonathan and Botond Köszegi, "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

Figure A.1: Demeaned shares over the life-cycle, ages 30-74



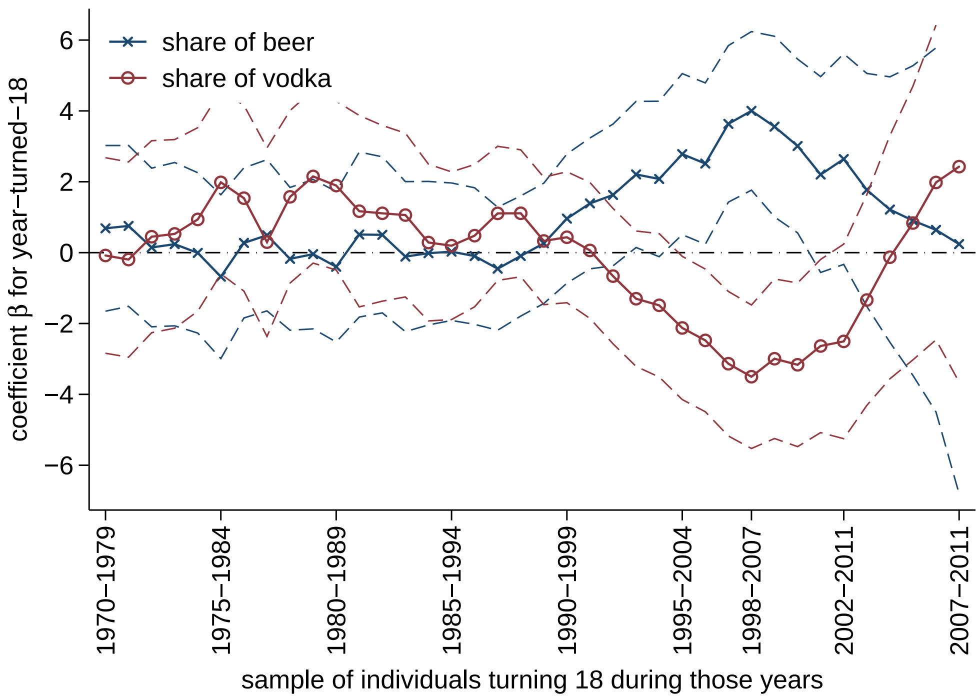
Notes: These figures complement the bottom panels of Figure 3 by providing the same analysis over the remaining part of the life-cycle.

Figure A.2: Illustration of the potential number of long-run steady states in the model



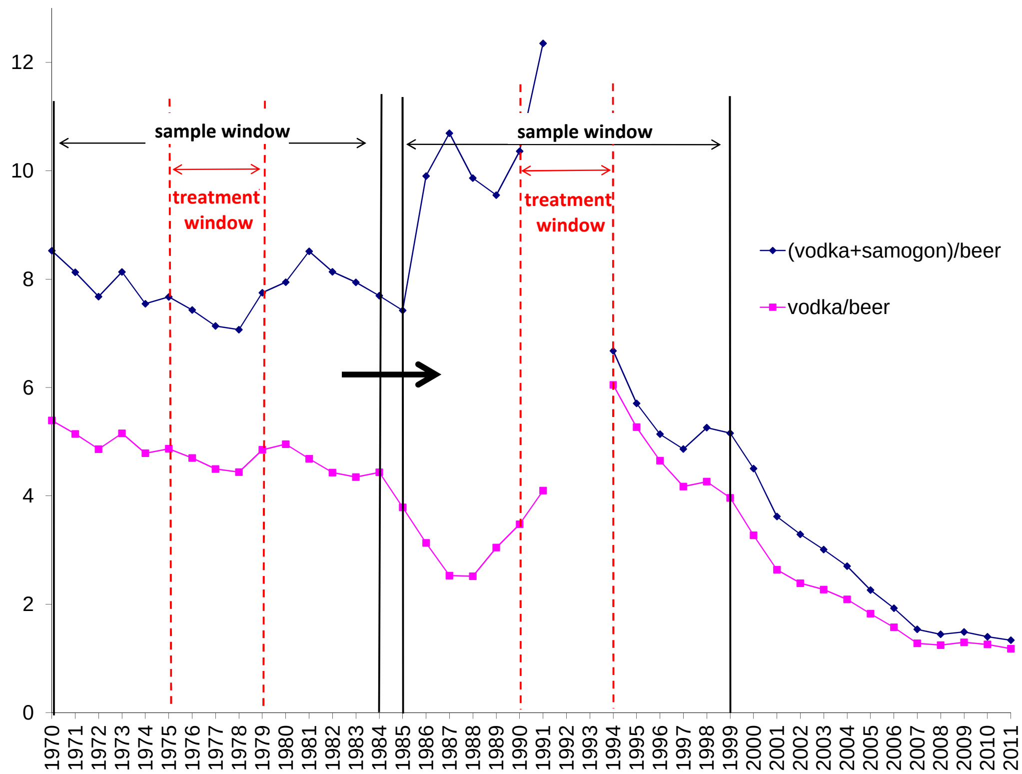
Notes: These figures show the dynamic behavior of the share of vodka starting from different initial conditions, i.e., different initial consumption shares. The three figures correspond to the three parametrizations in section A.1.2. The top left panel has one stable steady state, the top right panel has three steady states, two stable and one unstable, and the bottom panel has an infinite number of steady states.

Figure A.3: Placebo regressions for the beer market expansion experiment



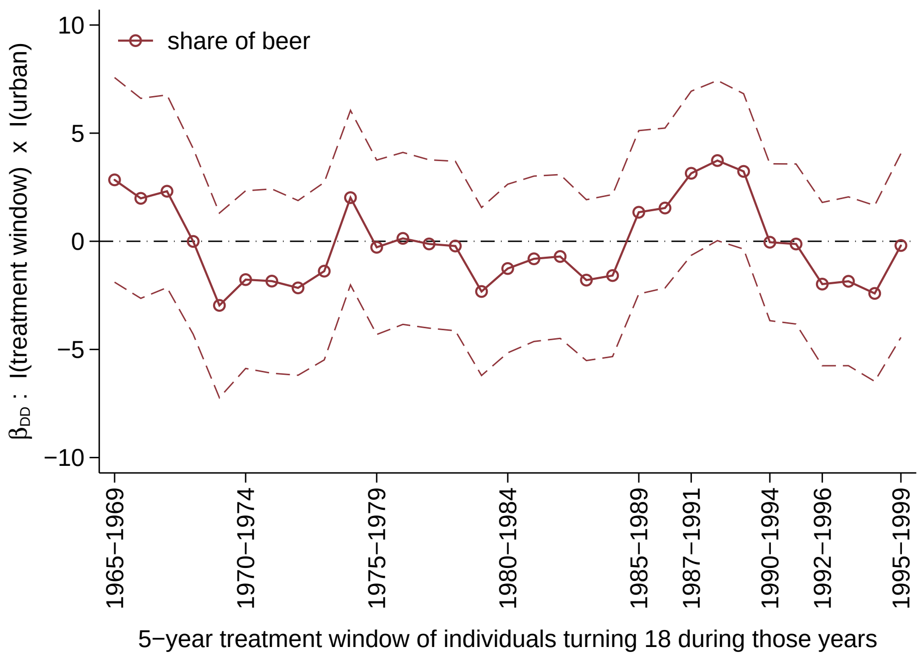
Notes: This figure complements the bottom panel of Figure 5 showing that consumers mostly substitute toward beer at the expense of vodka in response to the expansion of the beer market.

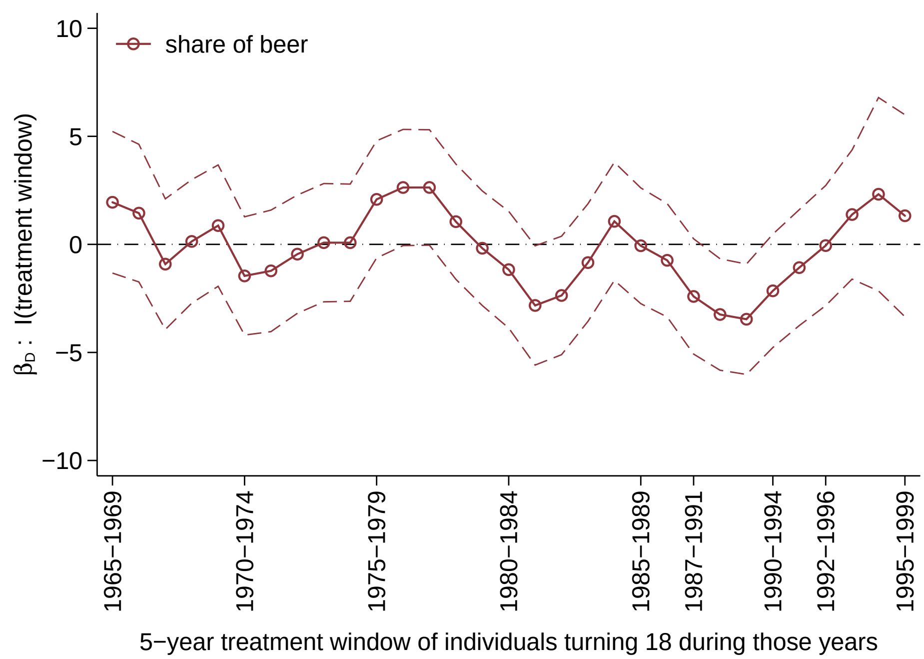
Figure A.4: Placebo regression design for anti-alcohol campaign experiment



Notes: This figure shows the design of the placebo regressions discussed in Section 5.3.

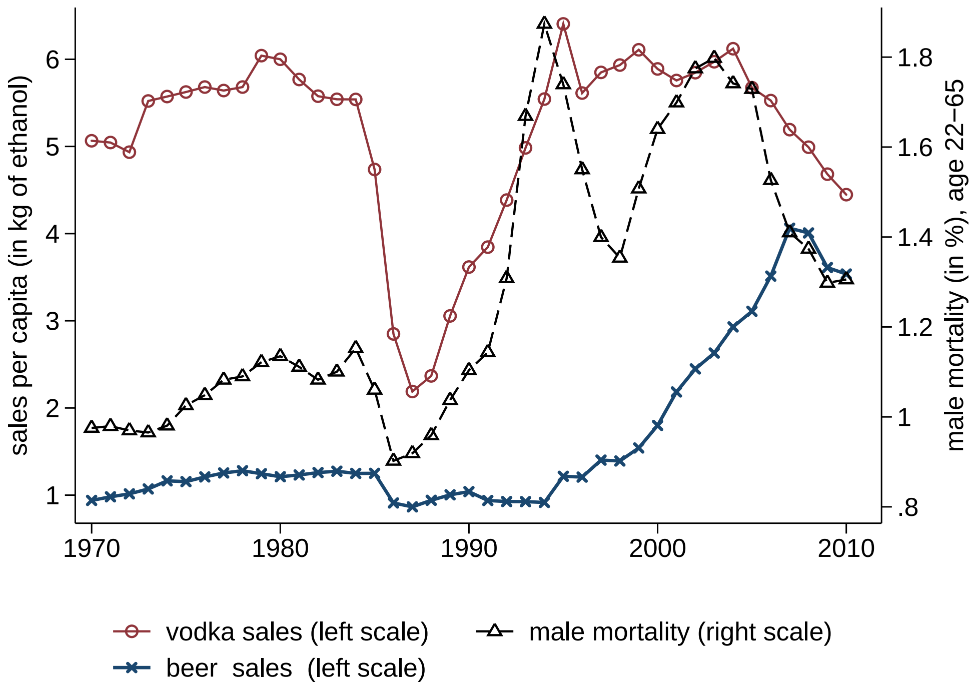
Figure A.5: Placebo regressions for anti-alcohol campaign experiment: beer





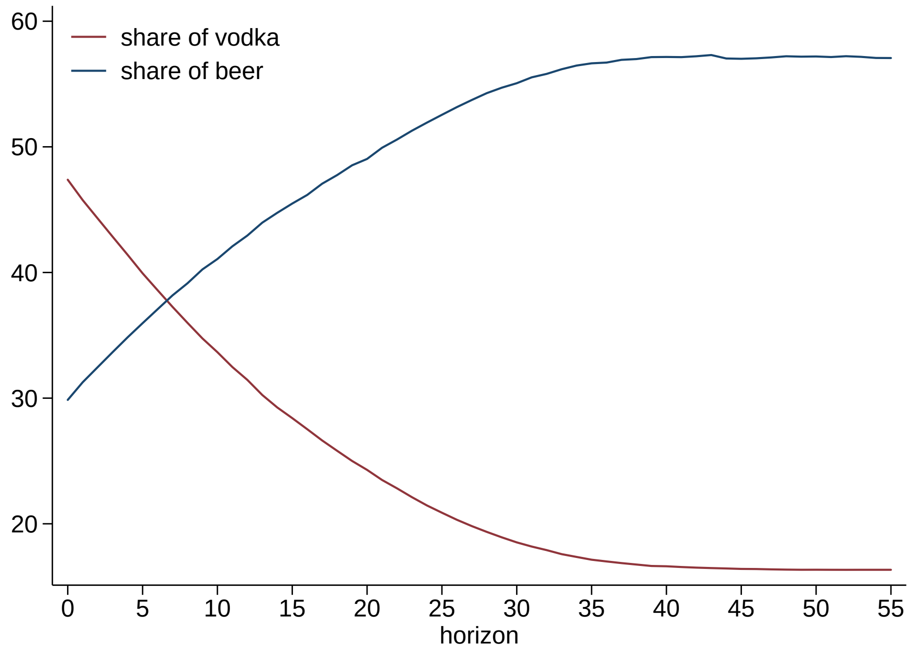
Notes: These figures complement Figure 7 showing that consumers mostly substitute between vodka and beer consumption in response to the anti-alcohol campaign.

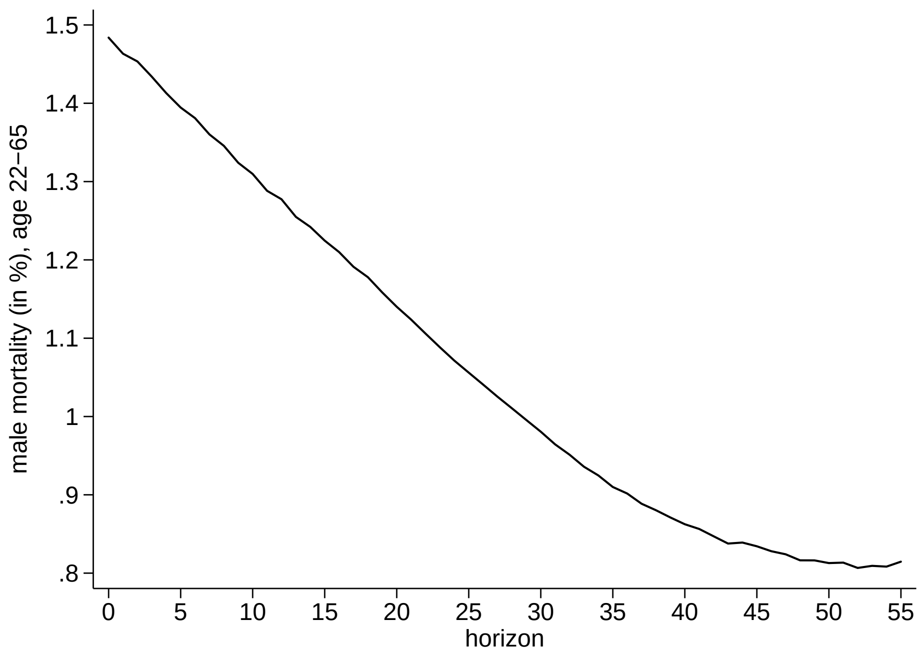
Figure A.6: Placebo regression design for anti-alcohol campaign experiment



Notes: This figure complements Table A.2 showing the co-movement of the mortality rate of males age 22 to 65 with aggregate sales per capita of vodka and beer, both measured in pure alcohol (i.e., ethanol).

Figure A.7: Simulated dynamics of shares and annual rate of death





Notes: These figures complement Table 6 showing the predicted consumption shares (top panel) and implied mortality rates (bottom panel) for males age 22 to 65 as a function of the forecast horizon in years.

Table A.1: Decomposition of the share of samogon

	(1) samogon	(2) samogon	(3) samogon	(4) samogon	(5) samogon	(6) samogon
I(born in 1990s)	-10.265***	-9.943***	-7.504***	-8.030***	-1.576	-1.163
	[1.629]	[0.616]	[0.669]	[0.737]	[1.157]	[1.163]
I(born in 1980s)	-8.558***	-8.236***	-5.645***	-5.975***	-2.504***	-1.617*
	[1.571]	[0.440]	[0.426]	[0.425]	[0.843]	[0.888]
I(born in 1970s)	-4.498***	-4.176***	-2.739***	-3.865***	-1.767***	-1.305*
	[1.598]	[0.529]	[0.465]	[0.403]	[0.647]	[0.670]
I(born in 1960s)	-2.366					
	[1.635]					
I(born in 1950s)	-1.279					
	[1.657]					
I(born in 1940s)	-0.014					
	[1.756]					
I(born in 1930s)	6.209***					
	[1.945]					
Alcohol intake (in grams of ethanol)			0.025***	0.026***	0.028***	0.057***
			[0.001]	[0.001]	[0.001]	[0.004]
Log(real income)			-1.690***	-1.087***	-1.194***	-0.874***
			[0.091]	[0.083]	[0.084]	[0.093]
Health status			1.374***	1.024***	0.224	0.525**
			[0.270]	[0.227]	[0.230]	[0.241]
Body weight (in kg)			-0.107***	-0.087***	-0.079***	-0.054***
			[0.014]	[0.012]	[0.012]	[0.013]
I(college degree)			0.774***	-1.920***	-1.907***	-1.423***
			[0.292]	[0.361]	[0.354]	[0.386]
I(married)			1.958***	-0.500	-0.810**	-0.313
			[0.407]	[0.378]	[0.397]	[0.424]
Observations	46,985	46,985	44,066	44,029	44,029	27,353
R-squared	0.025	0.017	0.069	0.256	0.265	0.238
Region FE				YES	YES	YES
Year FE				YES	YES	YES
Age FE					YES	YES
Top alcohol quartile dropped						YES

Notes: This table supplements Table 2 decomposing the share of illegally consumed samogon. Note the strong economically and statistically significant negative income effect of samogon, suggesting that samogon is indeed an inferior good, especially compared to beer and official vodka. HAC standard errors, clustered by individual, in parentheses.

Table A.2: Regressing male mortality on aggregate alcohol sales

	(1)	(2)	(3)	(4)	(5)	(6)
	1991-2010		1980-2010		1970-2010	
log(vodka)	1.126***		0.613***		0.391**	
	[0.263]		[0.155]		[0.160]	
log(beer)	-0.033		0.110		0.256***	
	[0.065]		[0.085]		[0.085]	
vodka sales per capita		0.215***		0.160***		0.117***
		[0.053]		[0.035]		[0.038]
beer sales per capita		-0.010		0.056		0.119***
		[0.032]		[0.039]		[0.041]
Observations	20	20	31	31	41	41
R-squared	0.520	0.503	0.480	0.503	0.369	0.377

Notes: This table provides the results of regressing the mortality rate of males age 22-65 on aggregate sales per capita of beer and vodka – measured in pure alcohol –, both in logs and in levels. Period 1991-2010 corresponds to the so-called "Russian Mortality Crisis", period 1980-2010 includes Gorbachev's anti-alcohol campaign, and period 1970-2010 is the longest sample for which we have aggregate sales of beer and vodka and fairly reliable data on mortality. See also Figure A.6 for the graphical evidence of the strong relationship between vodka per capita and male mortality.

Table A.3: Predicted alcohol shares

	(1) Share of beer	(2) Share of vodka
Alcohol intake (in kg of ethanol)	-0.884***	0.390***
	[0.037]	[0.043]
Log(real income)	0.005**	0.003
	[0.002]	[0.003]
Health status	0.000	-0.026***
	[0.008]	[0.009]
Body weight (in g)	-0.215	1.370***
	[0.223]	[0.259]
I(college degree)	-0.004	-0.020***
	[0.006]	[0.007]
I(married)	0.002	0.029***
	[0.008]	[0.009]
I(urban)	0.020***	-0.006
	[0.006]	[0.007]
I(no alcohol consumed)	-0.309***	-0.255***
	[0.010]	[0.012]
I(smokes)	0.012*	0.035***
	[0.007]	[0.008]
Observations	8,298	8,298
R-squared	0.309	0.262
Cohort FE	YES	YES
Age FE	YES	YES

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