Peer Effects and the Rise of Beer in Russia

Koen Deconinck[a1] and Johan Swinnen[b]

[a,b]: LICOS Centre for Institutions and Economic Performance & Department of Economics,
University of Leuven (KU Leuven)
Waaistraat 6 box 3511
B-3000 Leuven (Belgium)

Abstract

Between 1996 and 2007, per capita beer consumption in Russia more than quintupled, making beer the most important alcoholic drink in Russia today and transforming Russia into the world's third largest beer market. We study the role of peer effects in this rapid adoption of beer. In particular, we argue that the initial 'shocks' of improvements in quality and increased advertising were magnified through peer effects. Using detailed data from the Russian Longitudinal Monitoring Survey, we analyze the individual choice to drink beer, controlling for a wide range of potential factors. Our results indicate that peer effects played an important role as a 'transmission mechanism' in persuading consumers to start drinking beer. In contrast, changes in prices and incomes seem to have played only a minor role.

Keywords: peer effects, Russia, beer consumption

[1] Corresponding author: Koen Deconinck, koen.deconinck@ kuleuven.be, Phone: +32 16 326582, Fax: +32 16 326599, Mail: Waaistraat 6 box 3511, B-3000 Leuven (Belgium). We wish to thank Anna Salomons, Erik Schokkaert and Koen Decancq. All remaining errors are our own. We declare no conflict of interests. This research was financially supported by the KU Leuven Research Fund (Methusalem Fund).

Peer Effects and the Rise of Beer in Russia

1. Introduction

Up until the early 1990s, per capita beer consumption in Russia was negligible, fluctuating between 15 and 25 liters per capita. However, since the mid-1990s consumption has increased dramatically. Between 1996 and 2007, per capita beer consumption more than quintupled from 15 to 80 liters per capita, a level similar to the EU average. Thanks to this impressive growth, Russia has become the world's third largest beer market, after China and the US but ahead of Germany, the UK or Brazil (Colen and Swinnen 2011).

The rapid growth of beer consumption in Russia is remarkable given the traditional dominance of vodka. Using data from the Russian Longitudinal Monitoring Survey (RLMS), introduced in more detail below, Table 1 shows that as recently as 1994, 77% of Russians identified themselves as vodka-drinkers whereas only 28% reported drinking beer. However, by 2001, the proportion of beer drinkers (60%) had overtaken the share of vodka drinkers (57%). Between 1994 and 2007 the share of beer in total alcohol consumption increased from 51% to 79%, while the share of vodka decreased from 39% to 13% (Treml 1997; Euromonitor 2010b). Beer has also become the most popular drink in terms of money spent, occupying almost half of total sales in the alcoholic beverages market (Euromonitor 2010a). After its remarkable rise in past years, beer is now the dominant alcoholic drink in Russia.

This paper studies the role of peer effects in explaining the rapid adoption of beer in Russia. Since alcoholic beverages are usually consumed in social contexts, peer behavior is likely to be a major determinant of individual consumption behavior. Such peer effects can create a bandwagon effect, transmitting and amplifying the effect of initial exogenous 'shocks'. In particular, our argument is that the effects of improvements in quality and increased

advertising for beer were magnified through peer effects. Using detailed data from the Russian Longitudinal Monitoring Survey, we analyze the determinants of individuals' choice to drink beer or not, controlling for a wide range of potential factors. While our statistical analysis cannot directly shed light on the role of advertising and quality as initial 'shocks', our results indicate that peer effects played an important role as a 'transmission mechanism' in persuading consumers to start drinking beer. In contrast, changes in prices and incomes seem to have played only a minor role.

The existence of peer effects is important for policy-makers because it implies a social multiplier (Glaeser et al., 2003). An exogenous shock or a policy measure will not only change some individual's behavior directly, it will also indirectly affect the behavior of their peers. Therefore, knowledge of the existence and magnitude of peer effects is essential in order to evaluate the effects of exogenous shocks or policy interventions. The rapid adoption of beer in Russia is an important case to study given the major public health problems associated with alcohol consumption in Russia, where alcohol is thought to be the cause of approximately one-quarter of all deaths (Brainerd and Cutler, 2005; Nemtsov, 2002; Treisman, 2010). Studying the determinants of alcohol consumption patterns is necessary in order to arrive at effective policy interventions to limit the negative health impacts of alcohol abuse.[2]

The remainder of the paper is structured as follows. The next section briefly discusses other studies on alcohol consumption in Russia. Next, we present data on the extraordinary shift from vodka to beer consumption in Russia since the mid-1990s and show that the rise of beer is the result of larger numbers of beer drinkers, not of higher consumption per beer drinker. We then explain our conceptual framework in more detail in the fourth section and discuss the role of

[2] Interestingly, beer appears to have much less adverse health consequences than vodka; to the extent that the rapid increase in beer consumption is displacing vodka consumption, the adoption of beer as the dominant beverage may have positive health effects in the Russian context (Kueng and Yakovlev 2013).

advertising and quality improvements as initial 'shocks'. The fifth section discusses our methodology and dataset. Our empirical analysis is presented in the sixth section. We interpret the magnitude of our estimated peer effects in the seventh section, and the last section concludes.

2. Related Literature

Several other studies have investigated the determinants of alcohol consumption in Russia. For instance, Treisman (2010) studies the surge in alcohol-related deaths during the early years of transition in Russia and demonstrates that variations in mortality were driven by a strong decline in the real price of vodka, which fell by 77 per cent between 1990 and 1994. Moreover, several recent papers rely on the RLMS dataset used in this paper. Baltagi and Geishecker (2006) look at total alcohol intake to test rational addiction (RA) models, using RLMS data for 1994-2003, but do not find support for the RA hypothesis in Russia. Andrienko and Nemtsov (2006) use RLMS data for 1994-2002 and use Tobit and Heckman models to estimate a participation equation and a quantity decision equation for alcoholic drinks in general, as well as for different beverages separately. They recognize the importance of peer behavior for individual consumption decisions and implement this in their participation equation by including a dummy for whether there are other drinkers in the household. However, since households share a large number of unobservable characteristics, it is not clear which effect is picked up by this variable. Moreover, identification of the model depends crucially on the exogeneity of this parameter in the quantity decision equation, which seems questionable. Finally, although Andrienko and Nemtsov (2006) estimate both static and dynamic models based on the rational addiction literature, they do not fully exploit the panel data structure of the RLMS: their estimates do not control for individual fixed effects or year-specific shocks.[3] Herzfeld et al. (2014) look at food, alcohol and cigarette

[3] The analysis of Andrienko and Nemtsov (2006) is expanded to the 1994-2011 period by Roshchina (2013), using a similar methodology. The estimates provided by Roshchina (2013) again do not take into account individual fixed effects or year-specific shocks.

consumption in Russia during transition (1994-2005) using RLMS data. Their results indicate that habit formation, income, employment status and the price of alcohol play a role in explaining alcohol consumption. However, their analysis looks only at total alcohol intake without differentiating between vodka, beer or other beverages. Moreover, their estimates only look at quantity decisions, not at the decision to drink or not.

Recently, Goryakin et al. (2014) have used RLMS data for 1994-2009 to estimate price elasticities of demand for different alcoholic beverages, controlling for community-level fixed effects and a number of individual and community-level control variables. Their results show that while prices do seem to affect alcohol consumption, the magnitude is quite modest. Goryakin et al. (2014) study the quantity response of drinkers, while our own analysis focuses on the binary decision to drink beer or not. Our analysis is thus complementary to that of Goryakin et al. (2014): while their results on the demand elasticity of beer are conditional on the decision to drink beer, we explore the determinants of the individual decision to drink beer.

Our analysis is most closely related to that of Yakovlev (2013) and Kueng and Yakovlev (2013). Yakovlev (2013) looks at the relative impact of habits, prices and peer effects in explaining heavy drinking in Russia, defined as a pure alcohol intake in the top quarter (which therefore mostly coincides with heavy consumption of vodka). To study the determinants of heavy drinking, he models the decision to engage in heavy drinking as a game with imperfect information, where peers guess each other's behavior on the basis of demographic characteristics. These demographic characteristics act as instruments to identify the peer effect. Our paper differs from, and complements, Yakovlev (2013) in several ways. First, while Yakovlev (2013) looks at heavy drinking, we study the decision to drink beer. Second, while Yakovlev (2013) uses a structural model, our estimates are reduced-form. As in other contexts, structural and reduced-form approaches are best seen as complements rather than substitutes;

while structural models have the advantage of making counterfactual simulations easier, reduced-form models have the advantage of simplicity and less stringent assumptions. Third, identification in Yakovlev (2013) depends on the use of demographic characteristics of peers as instruments for peer behavior. This approach seems less appropriate in our context, however. Our baseline estimates show that demographic characteristics only have limited explanatory power in explaining the decision to drink beer, which casts doubt on their usefulness as instruments for our purposes. Despite these differences, the marginal effect of peer effects are comparable to the results obtained here, as can be seen from robustness checks using linear methods in the appendix to Yakovlev (2013).

Kueng and Yakovlev (2013) document the large differences in alcohol consumption patterns between different cohorts, demonstrating that younger cohorts are much more likely to drink beer while older cohorts are more likely to drink vodka. Their work rejects the 'stepping stone' argument that beer consumption leads to vodka consumption later on. Again, our work complements this analysis. While Kueng and Yakovlev (2013) analyze the different patterns between cohorts, we emphasize that even within cohorts over time there has been a strong increase in the fraction of people who drink beer, a finding which is true even for the older cohorts in our dataset. Moreover, while Kueng and Yakovlev (2013) focus on the share of beer and vodka in total alcohol consumption, we study the adoption of beer consumption more generally. Given the differing alcohol contents of beer and vodka, looking at the share of beer in total alcohol consumption may understate the drastic change in consumption habits in Russia.

3. Changes in Alcohol Consumption in Russia: From Vodka to Beer

For more than thirty years, from the 1960s through the 1980s and the early 1990s, per capita beer consumption in the Soviet Union and Russia fluctuated between 15 and 25 liters per capita –

considerably lower than in the rest of Europe. However, since the mid-1990s consumption has increased dramatically, as illustrated in Figure 1. Between 1996 and 2007 beer consumption grew from 15 liters per capita to around 80 liters. This increase in aggregate beer consumption is driven by growth in the number of beer drinkers rather than growth in individual consumption levels. Figure 2, based on Table 1, shows the fraction of beer drinkers and vodka drinkers over time.[4] The fraction of beer drinkers increased from roughly 30% in 1994 to more than 60% in 2008, while the fraction of vodka drinkers fell from more than 75% to less than 60%. On the other hand, Figure 3 shows self-reported monthly consumption volumes of beer over time, together with the median level of consumption reported.[5] Although these numbers should be treated with caution because of possible underreporting, there is no evidence of an upward trend in quantities consumed per drinker. The median reported intake remains constant at 2.5 liters of beer per month per beer drinker. Hence, the Russian beer boom seems to be driven by changes on the extensive margin (more people turning to beer) rather than changes on the intensive margin (beer drinkers consuming larger quantities).

The adoption of beer among Russian consumers is widespread and not limited to a specific gender, region or age group. Figure 4 shows the fraction of beer drinkers among men and women over time. While the fraction of female beer drinkers is always about 20 percentage points lower than the fraction of male beer drinkers, both show a similar trend over time. Thus, the increase in beer drinking is not limited to men only.

Figure 5 shows the fraction of beer drinkers in eight Russian regions over time. In a country as dispersed as Russia, it seems reasonable to expect large variations in consumption patterns between different regions. Surprisingly, however, the regional fraction of beer drinkers

[4] These fractions have been calculated using the sampling weights provided by the RLMS.
[5] This variable is computed by multiplying a survey question on the quantity normally consumed "per occasion" with a question on the normal number of days in a month the respondent drank alcohol.

follows the same pattern everywhere. Even though at any point in time there is considerable variation across regions, all regions experienced an upward trend in the number of beer drinkers. In 1994, for instance, the fraction of beer drinkers varied from somewhat less than 20% in the Volga-Vyatski and Volga region to around 35% in the Moscow and St. Petersburg region. In 2001, the regional fraction of beer drinkers varies between some 50% (Volga-Vyatski and Volga) and almost 70% (Ural). From 2001 onwards, regional averages remain within this band, although there is some variation over time.

Likewise, the spread of beer is not limited to a single age group. Figure 6 shows the fraction of beer drinkers by birth cohort over time, where we have divided the sample into those born before 1940, those born between 1941 and 1960, those born between 1961 and 1980, and those born after 1980.[6] The fraction of beer drinkers in each younger cohort is at least ten percentage points higher than in the preceding birth cohort. Yet, apart from this level effect, there is a strong increase in beer consumption among all cohorts. In the pre-1940 cohort, the fraction of beer drinkers increases from around 10% in the mid-1990s to more than 30% in 2003. Although there is some decline afterwards, the fraction of beer drinkers in the oldest cohort remains above 20%, or twice the level of the mid-1990s. A similar increase is present among the pre-1960 and pre-1980 cohorts.

[6] The results for the post-1980 cohort are only reported from 2001 onwards. The median age of this cohort increases steadily from 17 years in 2001 to around 21 years at the end of the period. Our analysis only includes respondents age 15 or older.

As these graphs make clear, the spectacular growth in beer consumption in Russia is the result of an increase in the number of beer drinkers, and not in their average intake levels. Moreover, the rising popularity of beer is not confined to men only, or to a specific region or age group. These findings suggest that an individual-level analysis of the decision to drink beer is needed in order to explain the rise of beer in Russia.

4. Conceptual Framework

4.1. Social Interactions and Cascades

The pattern of adoption of beer drinking in Russia documented in the previous section is reminiscent of the introduction of new products and technologies in general. The diffusion of such innovations is often said to follow an S-shaped curve, characterized by low initial levels of adoption, a period of strong growth, and eventual stabilization at a high level of adoption (Rogers, 2003).

The role of social learning and other social contagion effects in these processes has been well-established in the marketing literature (Peres et al., 2010) as well as in the literature on technology adoption (Rogers, 2003; Conley and Udry, 2010). After a new product is introduced, consumers who were initially skeptical may be convinced to use the product through word-of-mouth or by seeing other consumers using the product. In turn, this persuades other consumers to adopt the product.

Similar mechanisms have been described in other contexts. Frank et al. (2014) demonstrate that if consumers try to ‘keep up with the Joneses’, increased expenditures by some people induces those just below them in the income distribution to spend more as well. This gives rise to “expenditure cascades”. In the context of the outbreak of revolutions, Kuran (1989) has argued that under repressive regimes, individuals hide their negative sentiments about the

regime. This gives others the impression that the regime has widespread support, which in turn induces them to hide their own true sentiments. A revolutionary 'spark' may lead to a complete unraveling of this equilibrium, with an initially negligible opposition group receiving more and more support over time.

As these examples illustrate, the existence of social interactions can give rise to interesting dynamics. Initial events such as the adoption of a new technology by an early adopter, an increase in consumer spending by one category of consumers, or a revolutionary action by a small minority, can set in motion a cascade of similar behavior.

Our hypothesis is that the increase in beer consumption in Russia is best understood as such a process. In particular, quality improvements and increased advertising in the mid-1990s made beer a more attractive drink to Russian consumers. While these factors may have increased consumers' intrinsic preference for beer in the mid-1990s, they do not in themselves explain why the adoption of beer as the dominant drink in Russia was a gradual process that took place over a period of several years. We argue that peer effects played an important role in spreading beer consumption. Quality improvements and advertising directly persuaded some consumers to adopt beer, which in turn induced others to start drinking beer. In this way, peer effects transmitted and amplified the effect of advertising and quality improvements, leading to the observed pattern of diffusion.

4.2. Peer Effects

An extensive empirical literature has studied the impact of peers on individual behavior. Such peer effects have been demonstrated in a wide range of topics, including investment decisions for retirement among university professors (Duflo and Saez, 2002), crime rates (Glaeser et al., 1996), adolescent overweight (Blanchflower et al., 2009; Christakis and Fowler 2007; Trogdon

et al., 2008) and academic performance among university students (Sacerdote, 2001). In particular, much attention has been devoted to studying peer effects for cigarette and alcohol use (Ali and Dwyer, 2010; Christakis and Fowler, 2008, 2010; Kremer and Levy, 2008). The existence of such peer effects in a broad array of topics and across different social groups lends plausibility to our hypothesis that the decision to drink beer is influenced by the average choice in the peer group.[7]

4.3. Quality Improvements and Advertising

During the Soviet era, beer in Russia was 'unappetisingly cloudy and sour' (The Economist, 1999). Improving the quality of beer was one the main priorities of foreign investors entering the Central and Eastern European brewing industry. In their quest to improve the taste of beer, these companies quickly discovered that locally produced malt was of poor quality. At first, foreign breweries imported quality malt and barley from Western Europe. Later on, they adopted a strategy of vertical coordination in the supply chain to ensure a domestic supply of high-quality raw materials (Van Herck et al., 2012). Since foreign companies were more concerned about quality than local producers had been, the entry of multinational brewers in the Russian beer market quickly raised the quality of beer. When BBH gained control of Baltika in 1993, the company embarked on a large-scale investment program to modernize production, as part of the company's strategy to create a beer of European quality (Baltika 2008). In 1995, SUN Brewing (now Sun Inbev) invested $11.7 million in upgrading machinery to improve beer quality and taste since, according to then-chairman Shiv Khemka, 'the problem is not that [Russians] dislike beer (…) It's just that they don't have a first-rate national brand to choose from' (Russia Review 1996). When Carlsberg-Tuborg acquired the Vena brewery in St. Petersburg, the company

[7] For a popular introduction to the wide literature on peer effects and social interactions, see Christakis and Fowler (2009).

immediately invested $60 million to dramatically improve quality (Russian Journal, 1999b). Likewise, when Interbrew acquired the Rosar brewery in Omsk, they planned important investments to face an increasing demand for quality by Russian consumers (Russia Journal, 1999a). These quality improvements by foreign investors made beer a more attractive drink to Russian consumers.

Around the same time, a change in advertising regulations took place to restrict advertising for alcoholic drinks. However, when the first restrictions on alcohol advertising were adopted in 1995, these did not apply to beer (Osborne, 2011). As a result of the law, vodka commercials were banned from television, while beer commercials were still allowed (New York Times 2004). Around the time vodka commercials were banned (1995), BBH, owner of the Baltika brands, started a major advertising and distribution campaign. The strategy of SUN Brewers, the market leader in the 1990s, was to 'buy up a host of smaller breweries, raise the quality and unite the improved beer behind a well-marketed national brand' (Russia Review 1996). Given the conscious strategy of brewers to engage in an aggressive marketing campaign, advertising may have played a role in luring consumers towards beer.

The entry of foreign companies in the Russian beer market in the early 1990s thus led to improvements in quality and increased marketing efforts. At the same time, regulatory changes created an advantage for beer over vodka. These factors likely played a role in stimulating Russian beer consumption in the mid-1990s. However, as documented earlier, the adoption of beer as the dominant drink in Russia was a gradual process that took place over several years, whereas changes in quality and advertising appear to have been mostly concentrated in the mid-1990s. This suggests that quality improvements and advertising may have acted as initial 'shocks', which were then transmitted through peer effects. Ideally, our empirical analysis of the

rise of beer in Russia would explicitly take into account the improvements in quality and advertising efforts that happened throughout the period. Unfortunately, data on these factors is scarce. In our empirical analysis, we therefore focus on studying the transmission mechanism of peer effects.

5. Methodology and Data

Our empirical approach to study the role of peer effects in beer consumption is to estimate a linear probability model for beer drinking with a measure of peer behavior as one of the explanatory variables, while controlling for a wide array of possible confounding factors. Since the estimation of peer effects poses some well-known difficulties, we first discuss econometric issues related to the identification of peer effects and the strategy adopted here. Next, we discuss econometric issues related to the inclusion of a lagged dependent variable in order to control for possible habit formation. Finally, we present our dataset and the definition of variables used in our analysis.

5.1. Identification of Peer Effects

Identifying the role of social interactions on individual behavior is far from trivial. As emphasized by Manski (1993), an observed correlation between peer behavior and individual behavior may be due to three reasons (Table 2). First, the correlation may indeed be the result of peer *choices* influencing individual behavior (which Manski labeled *endogenous effects*); these constitute the peer effects we are most interested in. However, the correlation may also be due to common shocks (groups of consumers experiencing the same price shocks, for instance) or nonrandom selection of peers (people with similar preferences may seek each other's company, thus creating a false impression of peer effects). Manski referred to this as *correlated effects*. Third, the correlation may be the result of group *characteristics* (such as the average income or

education level in the peer group) influencing individual behavior, which Manski called *contextual effects*. Disentangling these three effects is a major concern in the literature on social interactions (Manski, 1993; Brock and Durlauf, 2007). The importance of addressing these issues is shown by Cohen-Cole and Fletcher (2008) who demonstrate that it is possible to find 'peer effects' for acne, height and headaches – phenomena where peer effects seem highly unlikely – if group characteristics are not included or selection issues are not properly taken into account.

Table 2 summarizes our approach for disentangling these three effects. We measure group behavior as the average fraction of beer drinkers in an individual's peer group (which we define below). To mitigate the problem of common shocks giving the appearance of a peer effect, we use a number of strategies. First, in all specifications we include time dummies to capture nation-wide trends. Second, in most specifications we include region-year dummies and extra covariates to capture local supply-side conditions and common shocks. Third, in one specification we use the lagged fraction of beer drinkers in the peer group instead of the simultaneous fraction of beer drinkers. To avoid the problem of non-random selection of peers (where people with strong preferences for drinking beer seek out each other's company, thus giving the impression of peer effects), our measure of peer behavior is based on an exogenous definition of the peer group. Finally, to disentangle the influence of group behavior versus group characteristics, we include group-level averages of individual covariates and additional group-level characteristics. Moreover, all our specifications include individual fixed effects (which also capture any group-level fixed effects).

Following Yakovlev (2013), we define an individual's peer group based on location, defined by the census districts included in the RLMS survey. The RLMS first selected around 170 communities throughout Russia and gathered detailed information about living standards

and prices in each community. In a next stage, a number of census districts within each community were selected for household surveys.[8] These census districts are narrowly defined administrative units. The whole of Russia is divided into 459,000 census districts, which implies an average population of around 300 people per census district. The small size of census districts makes it likely that our location-based definition of peers used here is a good proxy for the relevant peer group. Depending on the year, our dataset contains between 410 and 450 census districts from about 160 communities.

Thus, our empirical strategy estimates the peer effect using an exogenous definition of the peer group controlling for peer group characteristics, unobserved individual (and hence also group-level) fixed effects, and proxies for common (nationwide and regional) shocks.

A separate problem highlighted by Manski (1993) is the 'reflection problem'. In a linear model, the explanatory variable of interest (the average behavior of the group) is itself a linear function of the other explanatory variables (group-level characteristics and the group-level average of individual characteristics). This causes the regressors to be linearly dependent, and no identification is possible. However, this particular problem only occurs with linear models (e.g. a regression of individual weight on group-level averages to measure contagion in obesity). With non-linear models, the average behavior of the group cannot be a linear function of the other regressors, and the reflection problem disappears (Blume and Durlauf, 2005). We note here that using a linear probability model to analyze a binary choice does not give rise to this reflection problem. What causes the reflection problem is the linear dependence of the average group *behavior* on the other covariates. In the case of binary choice, there is no such linear dependence, and the reflection problem does not arise. A linear probability model assumes a linear

[8] See http://www.cpc.unc.edu/projects/rlms-hse/project/sampling for a more detailed description of the procedure.

relationship between the *probability* and the covariates, but it does not imply a linear relationship between actual behavior and covariates, and thus avoids the reflection problem.[9]

5.2. Habit Formation

The individual decision to drink beer is likely to be influenced both by unobserved individual tastes and attitudes and by previous experiences. For this reason, we would like to control for past consumption behavior in our estimations. However, the combination of unobserved heterogeneity and state dependence leads to a number of well-known econometric difficulties. In order to avoid these issues, we opt to perform our empirical analysis using a linear probability model with a lagged dependent variable, which we estimate using standard dynamic panel data techniques (i.e. dynamic GMM). This approach has the advantage that it allows us to control for both individual fixed effects and state dependence in a computationally easy way and with a minimum of extra assumptions on the data-generating process.

In a dynamic linear probability model, the probability of a positive outcome is assumed to be a linear function of the lagged dependent variable, observable covariates and an individual-level (unobserved) fixed effect. The model is treated as a regular dynamic panel data model using dynamic GMM (Arellano and Bond, 1991; Arellano and Bover, 1995). By using a suitable transform (most commonly first-differencing) the fixed effect is removed. However, this step introduces endogeneity into the model, as the first-differenced lagged dependent variable will be correlated with the first-differenced error term. This endogeneity is then addressed by using lags of the dependent variable as instruments, and the model is estimated using GMM. In our specific case, we use the forward orthogonal deviation transform developed by Arellano and Bover

[9] To see this, suppose that for an individual some covariate (e.g. income) increases, thus increasing his individual probability from 0.2 to 0.4. Since this does not change his behavior, there is no linear dependence of group behavior on the group-level average of covariates. This example does not depend on the specific form of the probability model; thus, linear probability models also avoid the reflection problem.

(1995) instead of first differencing because of the unbalanced nature of our panel. In an unbalanced panel, first differencing leads to a cascade of missing values. By contrast, instead of subtracting the previous period values, the forward orthogonal deviation subtracts the average of all *available* future values. In a balanced panel, this approach is numerically identical to first differencing, but in an unbalanced panel this approach improves efficiency by avoiding the loss of information inherent in first differencing (Roodman, 2009). The dynamic linear probability approach has been used by Bernard and Jensen (2004) in their analysis of firm entry into exporting and by Chay and Hyslop (1998), who use it as a robustness check for their nonlinear models and conclude that this approach provides an attractive alternative to more complicated methods.

5.3. Data and Variables

The data used in our analysis is the Russian Longitudinal Monitoring Survey (RLMS).[10] The RLMS is a nationally representative survey organized annually since 1992 (except in 1997 and 1999), covering a large number of respondents (typically more than 8,000 per round) and providing valuable information on all aspects of life in Russia. Our analysis uses rounds 5 to 17, spanning the years 1994-2008. Following common practice, we do not use earlier rounds due to concerns about data quality. We focus on individuals age 15 or older. The dataset covers 400 census districts from 171 communities (or 'sites' in the RLMS nomenclature) across Russia, divided among eight regions (Moscow and St. Petersburg, Northern and Northwestern, Central and Central Black-Earth, Volga-Vyatski and Volga, North Caucasian, Ural, Western Siberian, and Eastern Siberian and Far Eastern).[11] Around three quarters of the sample lives in an urban

[10]The RLMS is organized by the Carolina Population Center at the University of Carolina (Chapel Hill) and by the Higher School of Economics in Moscow. More information regarding the survey can be found at www.cpc.unc.edu/projects/rlms-hse.

[11] A map of the sites and regions can be found at the RLMS website. A site coincides with a *raion*, which is a secondary administrative unit in the Russian federation (below the level of *oblasts*).

setting. Some 43% of our respondents are male. Depending on the specification used, our sample varies between 6,000 and 16,000 individuals and between 25,000 and 57,000 observations (N x T). Summary statistics of our variables are presented in Table 3.

Our outcome variable is an indicator equal to one if the individual drank beer in the 30 days preceding the interview.

Our main variable of interest is peer behavior, measured as the average fraction of beer drinkers living in the same census district as the respondent (with the exception of the respondent and the respondent's household). Depending on the specification used, our estimation relies on 260 to 400 peer groups.[12]

Our dataset includes two variables related to prices of alcoholic beverages. The real price of beer and the real price of vodka are based on the community-level price data included in the RLMS. In each community surveyed, the lowest and highest prices for several consumption goods were recorded. We use the lowest price recorded as our measure for the prices of beer and vodka.[13] Real prices are calculated using CPI data for Russia taken from the World Development Indicators (World Bank, 2012), using the CPI of 1995 as benchmark.[14]

As individual covariates, we use personal income, educational status, employment status, marital status, and television ownership. Our personal income variable is the log of real income. This variable is constructed using the real value of household expenditures (calculated by the

[12] Table 5 shows statistics about the peer groups used in the estimation for our first specification with peer effects (column 2 in Table 4).

[13] Alternatively, it is possible to compute prices from household level data on purchases of goods, which are both recorded in quantities and in total expenses. However, in general there is a lot of noise in these data since both quantities and expenses may be measured with error. We use the minimum price at the community level, since this variable has the best correlation with household level prices. The highest recorded price is less informative, because the most expensive brand may be an outlier. This is less of an issue if we use the lowest price since this price is bounded from below.

[14] We also ran specifications using the logarithm of the prices, and using the ratio of real prices. Using the real prices separately gave the best results.

RLMS). The value of household expenditures has been divided by the square root of household size to arrive at equivalent income at the individual level. Taking the logarithm of this value then gives us our income measure.[15] The RLMS contains several variables on educational status. Following Goryakin et al. (2014), we include dummies for whether the individual has a high school degree and whether the individual has a university degree. (Including separate indicators for different educational levels leaves the fundamental results unchanged while giving little information on the effects of different education levels). As our measure for employment status, we use an indicator equal to one if the individual is unemployed. Likewise, marital status is measured by an indicator equal to one if the individual is married. We include an indicator variable for TV ownership as a measure of exposure to commercials. Several 'standard' individual-level characteristics are not included since they are time-invariant (e.g. gender) or since they would be collinear (e.g. age, which is collinear given the presence of individual fixed effects and year dummies).

As discussed earlier, certain characteristics of the peer group may influence behavior directly (so-called 'contextual effects', see Table 2). For instance, the average unemployment rate may affect individual behavior even when the individual himself is not unemployed. For this reason, we include a number of group-level characteristics in our regressions. In particular, we include the log of average income, average age (linear and squared), the fraction of men, the fraction of people with a high school degree, the fraction of people with a university degree, the fraction of unemployed persons, and the fraction of married people in the peer group.[16]

[15] Alternatively, we ran our specifications using the income data provided by the RLMS (transformed into the log of real personal income using the same approach). However, our expenditure-based definition gave better results. To the extent that households engage in consumption smoothing, an expenditure-based definition is a better proxy of 'permanent' income.

[16] In other specifications (not reported here), we included similar variables at the (broader) community level, or included both sets; results were not affected.

Based on the community-level information included in the RLMS, we construct several indicator variables representing whether the community can receive TV signals, whether the community has access to cable TV, and whether there are movie theaters and video halls. We also include as variables the number of restaurants and cafeterias. In addition, we include community-level averages of individual covariates.

6. Empirical Results

6.1. Baseline Estimate

The first column of Table 4 presents a baseline estimate without peer effects. As covariates we use the full set of individual-level controls, the group-level averages of individual characteristics, the community-level covariates, and time dummies. All standard errors are clustered at the census district level.

The coefficients on the price variables are not significantly different from zero and small in absolute terms. For instance, the real price of beer varies between 1.2 and 22 rubles per liter. If the estimated coefficients were giving the correct population parameter value, changing the beer price from the lowest to the highest observation in the sample would thus have a negligible impact. There is evidence of a moderate income effect, with a 1% increase in income translating into an increase of around 1.7 percentage points in the probability of drinking beer. None of the other covariates seem to play an important role in explaining beer drinking. There is, however, strong evidence for a time trend. Using 1995 as the baseline year, the time dummies clearly pick up the upward trend in beer drinking, with individuals in 1996 having a three percentage points higher likelihood of drinking beer, all the way up to a 24% points higher probability near the end of our sample period. As this baseline estimate shows, the rapid increase in beer consumption in Russia cannot be explained solely by prices and income changes.

6.2. Peer Effects

In the second column of Table 4, we add our measure of peer effects (the fraction of other people in the census district drinking beer) as a covariate. This coefficient is strongly significant and relatively large: compared to a census district where nobody is drinking beer, living in a census district where everybody is drinking beer increases the individual likelihood of beer-drinking with about 25 percentage points. The coefficients of the other covariates do not change noticeably after including the peer effect: there is still evidence of a modest income effect but no evidence of an impact of prices. Compared to the baseline estimate, inclusion of our measure of peer behavior decreases the magnitude of the time dummies by around one-third, suggesting that peer effects explain a large part of the evolution of beer consumption over time.

These estimates control for individual fixed effects, nation-wide shocks, group-level characteristics, and local "supply-side" factors as proxied by the number of cafeterias and the number of restaurants in the community. Moreover, in order to capture possible effects of advertising, these estimates include indicator variables for whether there is a movie theatre or a video hall in the community, whether the community has TV reception, and whether it has cable access. We find a small negative effect of movie theatres and a small positive effect of video halls, but none of the other community-level variables play an important role.[17]

Although our community-level covariates should capture some relevant supply-side factors such as advertising or the availability of beer through cafeterias and restaurants, arguably there may be other unobserved supply factors and other 'common shocks' driving our results. In the third column, we include region x year fixed effects to address this issue. Although regions are quite large (the whole of Russia is divided in eight regions), including the possibility of

[17] Similar results (not reported here) are obtained if we include an indicator for whether there is a fast food restaurant in the community.

different yearly shocks in different regions goes some way to capture e.g. logistical issues related to the availability of beer or differences in regional regulation. Repeating our baseline estimate without peer effects after inclusion of these dummies leaves the baseline results practically unchanged (results not reported here). Compared to our previous peer effects estimates, including region x year fixed effects decreases our estimate of the peer effect from about 25% to 21%. Other coefficients do not change markedly compared to our earlier estimates.

As our region x year dummies are a relatively crude measure of common shocks such as supply conditions, the fourth column includes an extra variable to capture this effect. Exploiting the sampling structure of the RLMS, we calculated the average fraction of beer drinkers in *other* census districts in the same community. Arguably, supply conditions will be roughly similar for these different census districts, as they are all situated in the same community.[18] On the other hand, it seems plausible that individual behavior would depend more on the behavior of people living in close proximity. Including the fraction of beer drinkers in other census districts in the same community would thus capture all community-wide effects of supply-side conditions or other shocks. As shown in column 4, the effect of beer drinking in other census districts is small and not statistically significant, while its inclusion has only a small impact on our estimates of peer effects: our coefficient falls from 21% to 19%. Thus, our estimate of peer effects appears remarkably robust, even after including this fine-grained proxy for common shocks.

As a final attempt to control for common shocks, column 5 uses a lagged measure of peer behavior. The measure of peer behavior used so far is the 'simultaneous' fraction of beer drinkers in the peer group (excluding the respondent and household members of the respondent). However, even after controlling for common shocks as in column 5, this simultaneous measure

[18] In fact, the RLMS itself assumes that the community is the appropriate level at which to collect detailed information about infrastructure, public goods and living conditions.

might still pick up some unobserved shocks affecting all members of the peer group at the same time, giving the false impression of a peer effect or overstating the true extent of the peer effect. One way to avoid this problem is using a lagged measure of peer behavior. If any 'remaining' shocks (after controlling for region x year fixed effects and the fraction of beer drinkers in other districts) are transitory, and if habit formation does not play a major role, our lagged measure of peer behavior avoids contamination by common shocks. (We study the role of habits in the next section.) As our estimate in the final column of Table 4 shows, our lagged measure of peer effects is statistically significant and around 6% in magnitude, even after controlling for region x year fixed effects and the fraction of beer drinkers in other districts. While this estimate is much lower than that using the simultaneous measure, peer effects thus appear to be robust to several ways of controlling for common shocks. Moreover, this last estimate may understate the true size of the peer effect. In our dataset, based on yearly observations, the use of lagged peer behavior implicitly assumes that individuals update their expectations of peers' behavior with a one-year lag. It seems reasonable that belief updating is faster in reality, in which case our lagged measure is not an accurate proxy for individuals' beliefs about peer behavior. In addition, as the RLMS dataset does not contain data for 1997 and 1999, the lagged measures used as explanatory variables for 1998 and 2000 are based on behavior observed two years earlier. Our lagged measure thus captures peer behavior with considerable measurement error, which leads to attenuation bias.[19]

Although we cannot fully exclude the possibility that unobserved shocks explain the correlation between peer group behavior and individual behavior, the estimates presented so far

[19] Our simultaneous measure may in fact also understate true peer effects if measurement error is considerable and if this effect outweighs the effect of common shocks. Given the fact that we use individual fixed effects (which take away a lot of the variation in the data) and that our peer groups are defined narrowly (which leads to small numbers of observations on which to base the calculation of peer behavior), it is plausible that there is some measurement error in this variable. We thank an anonymous referee for pointing this out.

show a peer effect even after accounting for fixed effects at the individual level (and hence also at the regional or local level); proxies for exposure to advertising (e.g. TV signal reception); proxies for supply conditions (e.g. the number of restaurants; region x year dummies; and the fraction of beer drinkers in other census districts in the same site); and dummies to capture year-specific shocks (whether at the national or the regional level). The results so far suggest that the magnitude of this peer effect is between 19% and 25% for our 'simultaneous' measure of peer effects.

6.3. Habit Formation

In order to account for the possibility of habit formation, we next turn to including a lagged dependent variable in our regression. As discussed in the previous section, the combination of fixed effects and a lagged dependent variable introduces endogeneity into the model, biasing not only the lagged dependent variable but the estimates of the other coefficients as well. A dynamic GMM estimation is needed to arrive at consistent estimates. However, it is possible to get a first idea of the magnitude of state dependence using simpler techniques. If a lagged dependent variable is used in a panel context, pooled OLS has an upward bias, while a fixed effects estimation has a downward bias (Roodman, 2009). Thus, the correct coefficient on the lagged dependent variable must lie between the lower bound provided by the fixed-effects estimate (presented in the first column of Table 5) and the upper bound provided by the OLS estimate (second column of Table 5). All of our specifications again include the extra site-level covariates (e.g. the number of restaurants) as well as region x year dummies. Based on these estimates, we see that drinking beer in the previous period has an effect on beer-drinking today of somewhere between a decrease of the likelihood with 7 percentage points and an increase with 40 percentage points. Interestingly, our estimate of peer effects still gives coefficients of 0.19 to 0.21, in line with results in the previous section.

The remaining columns of Table 5 present our GMM estimates. As explained earlier, we use the Arellano-Bond estimator with a forward orthogonal deviations transform. All specifications are estimated using two-step GMM with Windmeijer-corrected standard errors (Windmeijer, 2005; Roodman, 2009). The different columns report specifications using different lags as instruments; thus, column (3) reports the specification using lag 1 and higher, while column (4) reports results using lag 2 and higher, and so on.

The Arellano-Bond estimator relies on the use of lagged levels of the endogenous variable as an instrument for the transformed endogenous variable. However, this depends on the assumption of no autocorrelation in the error term (apart from that induced by the presence of fixed effects). More specifically, there should be no autocorrelation of order 2 in the differenced errors.[20] The Arellano-Bond test for autocorrelation, assuming a null of no autocorrelation, is reported at the bottom of the table. The null is not rejected at conventional significance levels in the second and third column. Moreover, the Hansen test does not reject the null of jointly valid instruments at conventional significance levels for the last three columns.

The results in all columns indicate that after controlling for habit formation there is evidence of a peer effect. Interestingly, our coefficients on peer behavior are comparable to those obtained earlier and remarkably robust across specifications, varying between 0.19 and 0.21. Our GMM estimates thus show little evidence for habit formation while leaving our results on the role of peer effects and other covariates basically unchanged. Our coefficient on habit formation itself is positive but small in the first two GMM specifications. In the last three columns, estimated habit effects are strongly negative but outside the interval established by FE and OLS

[20] In difference GMM, the difference transform induces serial correlation of order 1 in the first-differenced errors. Hence, to check for serial correlation in *levels*, one needs to check for second-order serial correlation in *differences*. For the case of orthogonal deviations considered here, the transform induces correlation of all orders in the transformed errors. Hence, even for the orthogonal deviations case, the test is run on differenced residuals (Roodman, 2009).

estimates. Thus, we do not find convincing evidence for a strong habit effect; however, the FE and OLS estimates as well as our GMM estimates show that peer effects are robust.[21]

6.4. Robustness Checks

One concern with the previous specifications may be that the identification of the peer effect is entirely due to data from one part of our observation period driving the result. For instance, the estimated peer effect may be entirely due to some spurious correlation during the period of strong growth in beer consumption between 1995 and 2002. To address this, we split the sample into two time periods. Column (1) of Table 6 presents results for the period 1994-2002 (when the growth in beer consumption was strongest), while column (2) presents results for the period 2003-2008 (during which beer consumption remained roughly stable). Again, our specifications include individual fixed effects, individual covariates, peer group covariates (such as average age or the unemployment rate), community covariates (such as the number of restaurants), region x year dummies, and the fraction of beer drinkers in other districts in the same community. Our measure of peer effects is robust. Both during the strong growth in beer consumption (0.20) and during the more stable period (0.10), our estimate suggests peer effects play a role in explaining beer consumption in Russia, although the effect becomes only half as large during the second period. Moreover, the estimated income effect seems to have been larger during the period of strong growth in beer consumption than during the later years.

7. Discussion

Our empirical analysis points to the existence of a robust and important peer effect in beer consumption in Russia. In most of our specifications, the estimated peer effect is around 0.20. The lowest estimates found were our estimate using lagged peer effects (around 0.06) and our

[21] One concern with our approach here is that, since data for 1997 and 1999 are lacking, our lagged dependent variable is measured with error, biasing our results downwards. To check this, we estimated all our specifications for the years after 2000 but obtained similar results.

estimate for the second half of the time period (around 0.10). These estimates control for a wide variety of potential confounding factors at the individual level, the peer group level, and the community level, and include individual fixed effects and time fixed effects.

These estimates are in line with the empirical literature on peer effects. Ali and Dwyer (2010), studying peer effects in alcohol consumption among adolescents, find that the marginal effect of the peer group on the likelihood of drinking is between 0.16 and 0.45 in different specifications. The estimated peer effect on the intensity of drinking is between 0.24 and 0.41 across different specifications. Trogdon et al. (2008) look at peer effects on individual's body-mass index using a sample of high-school children around 16 years old. They find a marginal effect of mean peer weight of 0.30. Using an instrumental variable approach, the marginal effect increases to 0.52. Powell et al. (2005) find evidence of peer effects in cigarette use among high school students. He estimates a peer effect of 0.58. Similarly, Lundborg (2006) finds a marginal effect of peer behavior of 0.56 for binge drinking, 0.47 for smoking and 0.16 for illicit drug use among a sample of Swedish high-school students. In contrast with these rather high estimates of peer effects in alcohol and cigarette use, Gaviria and Raphael (2001) find marginal effects of 0.15 for alcohol drinking and 0.13 for smoking among a sample of high-school students age 16. They estimate a marginal effect of peer behavior for drug use of 0.24. Thus, estimates of the marginal effect of peer behavior on individual behavior range from 0.13 to 0.56 across different contexts and using different methods. Our own estimates are in line with these results.

Apart from having a reliable point estimate, we would like to know the relative importance of peer effects in explaining the rise of beer in Russia. Because of our binary dependent variable, traditional measures of the goodness-of-fit such as R^2 are not informative. One approach would be to compare the predicted behavior based on our specifications with

actual consumption behavior. However, because of the elimination of individual fixed effects in our empirical approach, it is not possible to derive meaningful predictions from these estimates.

To get some idea of the importance of peer effects, rather than comparing the statistical fit of models, we may look at the magnitude of the coefficients on the time dummies. Because of the strong growth of beer consumption over time, these dummies "soak up" much of the unexplained increase over time in our baseline model. In a sense, these dummies are a "black box" capturing the unexplained increase in beer consumption. The explanatory power of other models compared to our baseline model without peer effects can then be evaluated by the decline in magnitude of the time dummies. A large decline in magnitude of the time dummies relative to the baseline estimate would indicate that peer effects account for a large part of the observed increase in beer consumption over time.

Figure 7 plots the estimated coefficients of the time dummies of our baseline model and our first specification (columns (1) and (2) from Table 4). These time dummies give the relative increase in the probability of beer drinking as against the baseline year 1995. As can be seen from a comparison of Figure 7 with Figure 2, the time dummies in our baseline model closely mimic the evolution of the fraction of beer drinkers over time. For instance, in 1995 the fraction of beer drinkers was 26%, while in 2001 the fraction of beer drinkers was 60%, an increase of 34 percentage points. The coefficient of the 2001 dummy in our baseline model is 27%. Thus, the time dummies in our baseline model capture most of the trend in beer consumption over time.

Compared to the baseline model, the coefficients on the time dummies decline by about one-third when our measure of peer effects is added in Table 4. Although there is still an upward trend in the time dummies between 1995 and 2002 (the period we are most interested in), peer effects clearly explain a large part of the increase in beer consumption over time. For 2001, the

coefficient on the time dummy is 18%, as compared to a baseline estimate of 27%. Including peer effects thus decreases the 2001 time dummy with about one-third. Similar results obtain for the other time dummies, and the decrease in the magnitude of the time dummies becomes even larger for the lagged measure of peer effects after 2001. Unfortunately, it is not possible to do a similar exercise for our other specifications (where, admittedly, the estimated coefficients on peer effects are somewhat smaller) because of the inclusion of region x year dummies. Hence, although these results need to be treated with caution, our results suggest that peer effects may account for up to one-third of the rise in beer consumption in Russia.

One interesting finding in our analysis is that prices of beer and vodka appear to play no role in explaining the likelihood of beer drinking in Russia. This result may seem counterintuitive at first, and seemingly at odds with some earlier studies which assigned a large role to prices in explaining Russian alcohol consumption (e.g. Treisman, 2010). However, in this specific case, the rise of beer consumption in Russia does seem to be too large to be plausibly explained by changes in relative prices. Our results support the alternative hypothesis that other factors (such as peer effects) may be a more important driver of alcohol consumption than purely 'economic' factors.

8. Conclusion

Between 1996 and 2007, Russian patterns of alcohol consumption underwent a dramatic change. While beer consumption had been practically non-existent before the mid-1990s, per capita beer consumption in Russia increased from 15 liters to 80 liters in the span of a decade. Understanding this transition is important for policy makers, given widespread problems of alcohol abuse and alcohol-related health problems in Russia.

This paper analyzed the determinants of beer consumption in Russia during the remarkable rise of beer. We documented how the increase in beer consumption is not driven by higher consumption levels per drinker, but rather by an increase in the fraction of beer drinkers. Our hypothesis in this paper was that the initial 'shocks' of advertising and quality improvements combined with peer effects to produce the observed diffusion pattern of beer consumption.

Relying on an extensive Russian panel dataset, our empirical analysis shows that beer consumption in Russia is indeed influenced by the behavior of an individual's peer group. In line with a large literature demonstrating the importance of the behavior of a peer group in shaping individual choices, our results show that the fraction of people drinking beer in an individual's census district positively influences the individual likelihood of drinking beer. Our empirical analysis tries to avoid the traditional pitfalls of estimating peer effects by relying on an exogenous definition of the peer group, and including controls for a wide array of individual characteristics, peer group characteristics, community characteristics, individual fixed effects (and thus also group-level fixed effects), time dummies and a proxy for common shocks. Most of our estimates vary between 0.19 and 0.25. We show that these estimates are robust to controlling for habit-formation (using dynamic panel data models) and splitting up the sample into two time periods. The estimates found here are in line with those in the literature, and we find that peer effects may account for up to one-third of the observed rise in beer consumption in Russia over time. Our results cast doubt on the importance of purely 'economic' factors in explaining alcohol consumption patterns, with potentially important implications for policy design. This issue is particularly relevant in Russia, where the negative health consequences of excessive alcohol consumption have been well documented.

References

Ali, M. and D. Dwyer (2010) "Social network effects in alcohol consumption among adolescents," Addictive Behaviors 35, pp.337-342

Andriendko, Y., and A. Nemtsov (2005). "Estimation of Individual Demand for Alcohol," Economics Education and Research ConsortiumWorking Paper 05/10.Moscow: EERC.

Arellano, M. and O. Bover (1995) "Another look at the instrumental variables estimation of error-components models," Journal of Econometrics 68, pp. 29-51.

Arellano, M. and S. Bond (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," Review of Economic Studies 58, pp. 277-297.

Baltagi, B. and I. Geishecker (2006) "Rational Alcohol Addiction: Evidence from the Russian Longitudinal Monitoring Survey," Health Economics 15, pp. 893-914

Baltika (2008). Baltika Annual Report 2007.

Bernard, A. and J. Jensen (2004) "Why some firms export," Review of Economics and Statistics 86(2), pp. 561-569.

Blanchflower, D., A. Oswald and B. Van Landeghem (2009) "Imitative obesity and relative utility," Journal of the European Economic Association 7, pp. 528-538.

Blume, L. and S. Durlauf (2005) "Identifying social interactions: A review," working paper, July 2005.

Brainerd, E. and D. Cutler (2005) "Autopsy on an empire: Understanding mortality in Russia and the former Soviet Union," Journal of Economic Perspectives 19(1), pp. 107-130.

Brock, W. and S. Durlauf (2001) "Discrete choice with social interactions," Review of Economic Studies 68, pp. 235-260.

Brock, W. and S. Durlauf (2007) "Identification of binary choice models with social interactions," Journal of Econometrics 140, pp. 52-75.

Chay, K. and D. Hyslop (1998) "Identification and estimation of dynamic binary response panel data models: Empirical evidence using alternative approaches," working paper, October 1998.

Christakis, N. and J. Fowler (2007) "The spread of obesity in a large social network over 32 years," New England Journal of Medicine 357(4), pp. 370-379.

Christakis, N. and J. Fowler (2008) "The collective dynamics of smoking in a large social network," New England Journal of Medicine 358(21), pp. 2249-2258.

Christakis, N. and J. Fowler (2009) *Connected. The Amazing Power of Social Networks and How They Shape Our Lives*, Harper Press, 338p.

Christakis, N. and J. Fowler (2010) “The spread of alcohol consumption behavior in a large social network,” Annals of Internal Medicine 152(7), pp. 426-433.

Cohen-Cole, E. and J. Fletcher (2008) “Is obesity contagious? Social networks vs. environmental factors in the obesity epidemic,” Journal of Health Economics 27(5), pp. 1382-1387.

Colen, L. and J. Swinnen (2011) “Beer-drinking nations: The determinants of global beer consumption,” in: J. Swinnen (ed.) *The Economics of Beer*, Oxford University Press, pp. 123-140.

Conley, T. and C. Udry (2010), “Learning About A New Technology: Pineapple in Ghana,” American Economic Review 100(1), pp. 35-69.

Duflo, E. and E. Saez (2002) “Participation and investment decisions in a retirement plan: the influence of colleagues’ choices,” Journal of Public Economics 85 pp. 121–148.

Euromonitor (2010a) “Alcoholic Drinks in Russia,” Country Market Insight Report.

Euromonitor (2010b) “Beer in Russia,” Country Sector Briefing.

Frank, R., A. Levine and O. Dijk (2014) “Expenditure Cascades,” Review of Behavioral Economics 1, pp. 55-73.

Gaviria, A. and S. Raphael (2001) “School-based peer effects and juvenile behavior,” Review of Economics and Statistics 83(2), p. 257-268.

Glaeser, E., B. Sacerdote and J. Scheinkman (1996) “Crime and social interactions,” Quarterly Journal of Economics 111(2), pp. 507-548

Glaeser, E., B. Sacerdote and J. Scheinkman (2003) “The social multiplier,” Journal of the European Economic Association, 1(2-3), pp. 345-353.

Goryakin, Y., B. Roberts and M. McKee (2014) “Price Elasticities of Alcohol Demand: Evidence from Russia,” European Journal of Health Economics, forthcoming.

Herzfeld, T., S. Huffman and M. Rizov (2014) “The Dynamics of Food, Alcohol and Cigarette Consumption in Russia During Transition,” Economics and Human Biology 13, pp. 128–143.

Honoré, B. and E. Kyriazidou (2000) “Panel data discrete choice models with lagged dependent variables,” Econometrica 68(4), pp. 839-874.

Kremer, M. and D. Levy (2008) “Peer effects and alcohol use among college students,” Journal of Economic Perspectives 22(3), pp.189-206.

Kueng, L. and E. Yakovlev (2013), “How Persistent Are Consumption Habits? Micro-Evidence From Russian Men,” working paper.

Kuran, T. (1989) “Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution,” Public Choice 61(1), pp. 41-74.

Lundborg, P. (2006) “Having the wrong friends? Peer effects in adolescent substance use,”

Journal of Health Economics 25, pp. 214-233.

Manski, C. (1993) “Identification of endogenous social effects: The reflection problem,” Review of Economic Studies 60(3), pp. 531-542.

Nemtsov, A. (2002) “Alcohol-related human losses in Russia in the 1980s and 1990s,” Addiction 97, pp. 1413-1425.

New York Times (2004). “Moscow Journal: As Russia Discovers Beer, Deputies Try to End

Osborne, A. (2011) “Beer Is No Longer Food For Russians,” British Medical Journal 342, p. 520.

Peres, R., E. Muller and V. Mahajan (2010) “Innovation diffusion and new product growth models: A critical review and research directions,” International Journal of Research in Marketing 27, pp. 91-106.

Powell, L., J. Tauras and H. Ross (2005) “The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior,” Journal of Health Economics 24, pp. 950-968.

Rogers, E.M. (2003) Diffusion of Innovations, 5th ed. New York: The Free Press

Roodman, D. (2009) “How to do xtabond2: An introduction to difference and system GMM in Stata,” Stata Journal 9(1), pp. 86-136.

Roshchina, Y. (2013) “To Drink or Not To Drink: The Microeconomic Analysis of Alcohol Consumption in Russia in 2006-2010,” National Research University – Higher School of Economics Working Papers WP BRP 20/SOC/2013.

Russia Journal (1999a). “Interbrew Tightens Grip on Beer Market,” available online at http://www.russiajournal.com/node/878 [Accessed April 16, 2014]

Russia Journal (1999b). “Changes Brew in Russian Drinking Habits,” available online at http://russiajournal.com/node/345 [Accessed April 16, 2014]

Russia Review (1996). ‘SUN’s GoldenBrew’. 21 October. SUN Group. Available at https://web.archive.org/web/20091223190925/http://www.sungroup-global.com/english/media_centre/selected_articles/183.asp [Accessed April 16, 2014]

Sacerdote, B. (2001) “Peer effects with random assignment: Results for Dartmouth roommates,” Quarterly Journal of Economics 116(2), pp. 681–704.

the Binge.” 6 August.

The Economist (1999). “Russian Beer: Vodka Chasers,” available online at http://www.economist.com/node/322940 [Accessed April 16, 2014]

Treisman, D. (2010) “Death and prices: The political economy of Russia’s alcohol crisis,” Economics of Transition 18(2), pp. 281-331.

Treml, V. (1997) “Soviet and Russian statistics on alcohol consumption and abuse,” in Bobadilla, J.L., C.A. Costello and F. Mitchell (eds.) *Premature Death in the New Independent States*, National Research Council, Washington.

Trogdon, J., J. Nonnemaker and J. Pais (2008), “Peer effects in adolescent overweight,” Journal of Health Economics 27, pp. 1388-1399.

Van Herck, K., J. Swinnen and K. Deconinck (2012) “How the East Was Won: Supply Chain Restructuring in the Eastern European Beer Market,” German Journal of Agricultural Economics 61(4), pp. 213-222.

Windmeijer, F. (2005) “A finite sample correction for the variance of linear efficient two-step GMM estimators,” Journal of Econometrics 126, pp. 25-51

Wooldridge, J. (2005) “Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity,” Journal of Applied Econometrics 20(1), pp. 39-54.

World Bank (2012) *World Development Indicators 2012*, online at http://data.worldbank.org/indicator

Yakovlev, E. (2013) “Peers and alcohol: evidence from Russia,” working paper, University of California – Berkeley.

Table 1. Fraction of Beer Drinkers and Vodka Drinkers in Russia

Year	Beer Drinkers	Vodka Drinkers
1994	28%	77%
1995	26%	77%
1996	30%	74%
1998	39%	70%
2000	51%	63%
2001	60%	57%
2002	59%	55%
2003	58%	54%
2004	57%	51%
2005	55%	52%
2006	61%	50%
2007	60%	49%
2008	59%	51%

Source: RLMS, authors' own calculations using sampling weights

Table 2. Three Explanations for Observed "Peer Effects"

Manski (1993)	Explanation	Empirical strategy adopted
Endogenous effects	Influence of group *behavior* on individual behavior (peer effect)	Measured by average behavior at group level (excluding members of respondent's household)
Correlated effects	Common shocks	- Time dummies - Region*Year dummies - Average behavior in other districts - Lagged peer effect
	Nonrandom selection	Exogenous definition (location-based)
Contextual effects	Influence of group *characteristics* on individual behavior	- Peer group covariates - Community covariates - Individual FE (control for group FE)

Table 3. Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Drinks Beer	72349	0.51	0.50	0.00	1.00
Real Price of Beer	109618	3.17	1.27	1.32	22.06
Real Price of Vodka	112904	13.62	4.27	4.12	42.38
% Beer Drinkers in Census District	131254	0.49	0.25	0.00	1.00
Individual Characteristics					
Age	131074	43.29	18.68	14.08	102.67
Male	131740	0.43	0.49	0.00	1.00
Log of Real Income	126327	8.35	0.89	0.00	12.56
High School Degree	132052	0.57	0.49	0	1
University Degree	132052	0.17	0.38	0	1
Unemployed	126350	0.04	0.20	0.00	1.00
Married	132052	0.60	0.49	0.00	1.00
Owns a Television	132052	0.92	0.28	0.00	1.00
Peer Group Covariates					
Average Age	132052	43.28	5.80	19.62	87.92
% of Men	132052	0.43	0.07	0.00	1.00
Log of Avg Income	132022	8.61	0.56	5.33	10.98
% High School Degree	132052	0.57	0.16	0.00	1.00
% University Degree	132052	0.17	0.13	0.00	1.00
Unemployment Rate	132022	0.04	0.05	0.00	0.67
% of Married People	132052	0.60	0.12	0.00	1.00
Community Covariates					
TV Signal Reception	120546	0.97	0.16	0.00	1.00
Cable Access	109901	0.71	0.45	0.00	1.00
Number of Cafeterias	109322	68.45	148.43	0.00	1300.00
Number of Restaurants	112692	18.47	61.28	0.00	770.00
Movie Theater	120526	0.80	0.40	0.00	1.00
Video Hall	109920	0.51	0.50	0.00	1.00

Table 4. Empirical Results

Dependent Variable: Respondent Drinks Beer (Linear Probability Model)

VARIABLES	(1) Baseline Regression	(2) Peer Effect	(3) Peer Effect Region*Year	(4) Peer Effect Region*Year Other Districts	(5) Lagged Peer Effect Region*Year Other Districts
% Beer Drinkers		0.250***	0.210***	0.187***	
		(0.020)	(0.021)	(0.025)	
Lagged % of Beer Drinkers					0.061**
					(0.024)
% in Other Districts				0.025	-0.055
				(0.052)	(0.067)
Real Price of Vodka	0.000	0.001	-0.001	0.001	0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Real Price of Beer	-0.005	-0.003	-0.006*	-0.005	-0.007
	(0.004)	(0.003)	(0.003)	(0.005)	(0.007)
Log of Real Income	0.017***	0.017***	0.016***	0.014***	0.015***
	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)
Individual Covariates					
High School	0.01	0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
University	-0.01	-0.01	-0.01	0.01	0.00
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Unemployed	-0.01	-0.01	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Married	0.02**	0.02**	0.02*	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Owns a Television	0.02	0.02	0.02	0.02	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Peer Group Covariates					
Log of Avg Income	0.006	0.003	0.006	0.005	-0.004
	(0.010)	(0.009)	(0.009)	(0.010)	(0.013)
Average Age	-0.008	-0.005	-0.002	-0.004	-0.001
	(0.008)	(0.007)	(0.007)	(0.009)	(0.012)
Average Age, squared	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
% of Men	-0.07	-0.09	-0.09	-0.14*	-0.17*
	(0.08)	(0.06)	(0.07)	(0.07)	(0.09)
% High School	-0.05	-0.04	-0.03	-0.00	0.01
	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
% University	-0.03	-0.01	-0.04	-0.08	-0.17**
	(0.07)	(0.06)	(0.06)	(0.06)	(0.08)
Unemployment Rate	-0.07	-0.06	-0.07	-0.10*	-0.07
	(0.06)	(0.05)	(0.05)	(0.06)	(0.08)
% of Married People	0.01	0.00	-0.00	0.01	-0.01
	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)

Community Covariates					
Number of Cafeterias	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Number of Restaurants	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Movie Theater	-0.03**	-0.02**	-0.01	-0.07***	-0.07***
	(0.02)	(0.01)	(0.01)	(0.02)	(0.03)
Video Hall	0.01	0.01*	0.02**	0.01	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
TV Signal Reception	0.03	0.03	0.04		
	(0.05)	(0.04)	(0.05)		
Cable Access	-0.01	-0.01	-0.01	0.00	-0.00
	(0.01)	(0.00752)	(0.01)	(0.02)	(0.02)
Year: 1996	0.01	0.00			
	(0.01)	(0.01)			
Year: 1998	0.10***	0.06***			
	(0.02)	(0.02)			
Year: 2000	0.20***	0.14***			
	(0.02)	(0.01)			
Year: 2001	0.27***	0.18***			
	(0.02)	(0.02)			
Year: 2002	0.26***	0.18***			
	(0.02)	(0.02)			
Year: 2003	0.25***	0.17***			
	(0.02)	(0.02)			
Year: 2004	0.23***	0.15***			
	(0.02)	(0.02)			
Year: 2005	0.22***	0.14***			
	(0.02)	(0.02)			
Year: 2006	0.24***	0.15***			
	(0.02)	(0.02)			
Year: 2007	0.24***	0.15***			
	(0.02)	(0.02)			
Year: 2008					
Constant	0.342	0.210	0.170	0.341	0.551*
	(0.211)	(0.184)	(0.195)	(0.241)	(0.323)
Individual FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	No	No	Yes	Yes	Yes
Observations	46,365	46,275	46,275	34,111	26,366
Nr of Individuals	14,505	14,483	14,483	10,541	7,915
Nr of Peer Groups	401	396	396	281	281
Overall R2	0.05	0.08	0.07	0.07	0.01

Robust standard errors in parentheses (clustered on census districts)
*** p<0.01, ** p<0.05, * p<0.1

Table 5. Overview of Peer Groups Used in Estimation

Year	Number of peer groups	Size	Number used in estimation
1994	259	24,3	7,6
1995	239	23,8	7,1
1996	230	24,5	7,6
1998	226	26,7	7,1
2000	321	29,4	11,9
2001	329	30,1	13,4
2002	313	32,1	13,6
2003	320	34,0	14,5
2004	333	34,0	13,9
2005	327	34,3	13,7
2006	301	39,5	16,1
2007	308	39,5	16,1

Note: Data in this table refers to the peer groups used in estimating column (2) of Table 4. For each year used in the estimation, this table shows the number of peer groups, the average size of the peer group, and the average number of individuals per peer group which were used in estimation (the difference between the last two columns being due to missing data for some individuals). Since data on some peer groups is not available in all rounds, the number of peer groups varies; in total, 396 distinct peer groups were used in estimation.

Table 6. Controlling for Habit Formation

Dependent Variable: Respondent Drinks Beer
Linear Probability Model, Dynamic GMM (Arellano-Bond using FOD)

VARIABLES	(1) Fixed Effects	(2) OLS	(3) GMM 1 lag	(4) GMM 2 lag	(5) GMM 3 lag	(6) GMM 4 lag	(7) GMM 5 lag
Drank Beer in Previous Period	-0.072***	0.395***	0.067***	0.085	-0.169	-0.333**	-0.394**
	(0.011)	(0.008)	(0.015)	(0.129)	(0.164)	(0.164)	(0.183)
% of Beer Drinkers	0.187***	0.211***	0.190***	0.187***	0.197***	0.204***	0.206***
	(0.030)	(0.018)	(0.023)	(0.024)	(0.024)	(0.025)	(0.026)
% in Other Districts	-0.023	0.120***	-0.045	-0.048	-0.014	0.013	0.020
	(0.065)	(0.043)	(0.064)	(0.066)	(0.068)	(0.070)	(0.073)
Real Price of Vodka	0.003**	0.000	0.003*	0.003*	0.004**	0.004**	0.004**
	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Real Price of Beer	-0.012*	-0.008	-0.011*	-0.011*	-0.010	-0.011	-0.010
	(0.007)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Log of Real Income	0.019***	0.027***	0.018***	0.018***	0.017***	0.017**	0.017**
	(0.007)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Individual Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer Group Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Community Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	Yes	No	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,708	20,708	14,199	14,199	14,199	14,199	14,199
Nr of Individuals	6,508	6,508	4,256	4,256	4,256	4,256	4,256
Nr of Peer Groups	281	281	280	280	280	280	280
AR(1)-test p-value			0.00	0.00	0.06	0.19	0.32
AR(2)-test p-value			0.06	0.51	0.40	0.08	0.06
Hansen test DF			54	44	35	27	20
Hansen test p-value			0.06	0.09	0.12	0.11	0.33

Robust standard errors in parentheses. Clustered on Census District for FE estimate and on individuals for OLS estimate; robust standard errors using Windmeijer's finite sample correction for GMM estimates.
*** p<0.01, ** p<0.05, * p<0.1

Table 7. Robustness Check

VARIABLES	(1) 1995-2002	(2) 2003-2008
% of Beer Drinkers	0.203***	0.098**
	(0.035)	(0.038)
% in Other Districts	-0.009	-0.048
	(0.085)	(0.084)
Real Price of Vodka	0.001	-0.003
	(0.001)	(0.002)
Real Price of Beer	-0.011*	0.010
	(0.006)	(0.009)
Log of Real Income	0.021***	0.010
	(0.007)	(0.008)
Constant	0.376	0.443
	(0.348)	(0.472)
Individual Covariates	Yes	Yes
Peer Group Covariates	Yes	Yes
Community Covariates	Yes	Yes
Individual FE	Yes	Yes
Region*Year FE	Yes	Yes
Observations	17,392	16,719
Nr of Individuals	7345	6869
Nr of Peer Groups	273	261
Overall R2	0.10	0.00

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 1. Beer Consumption Per Capita in Russia, 1960-2007

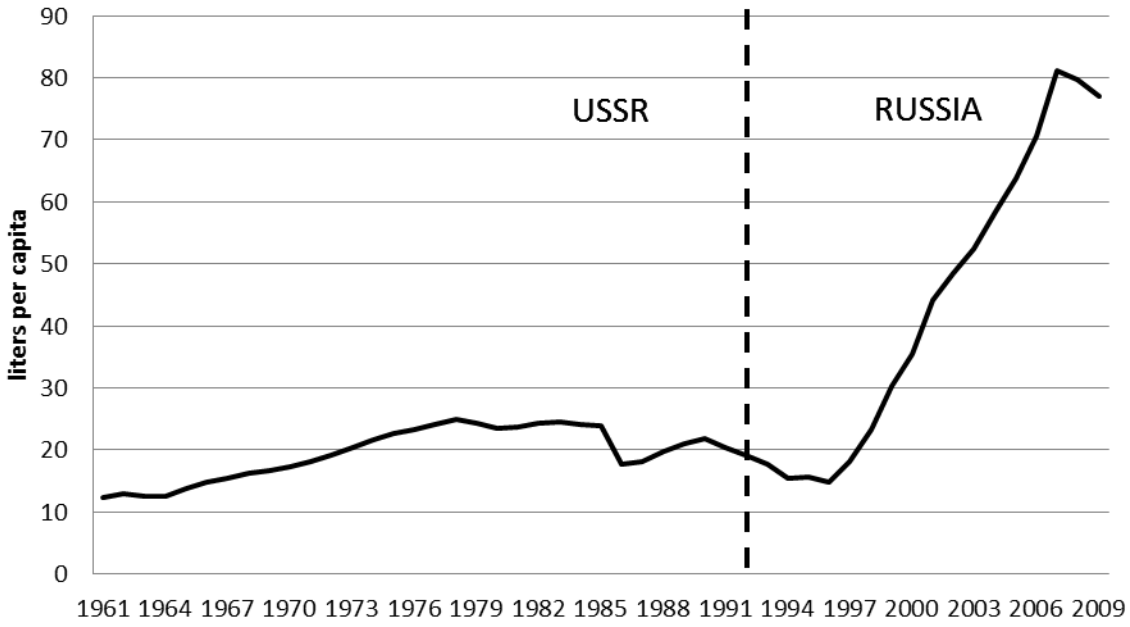

Figure 2. Fraction of Beer Drinkers and Vodka Drinkers over Time

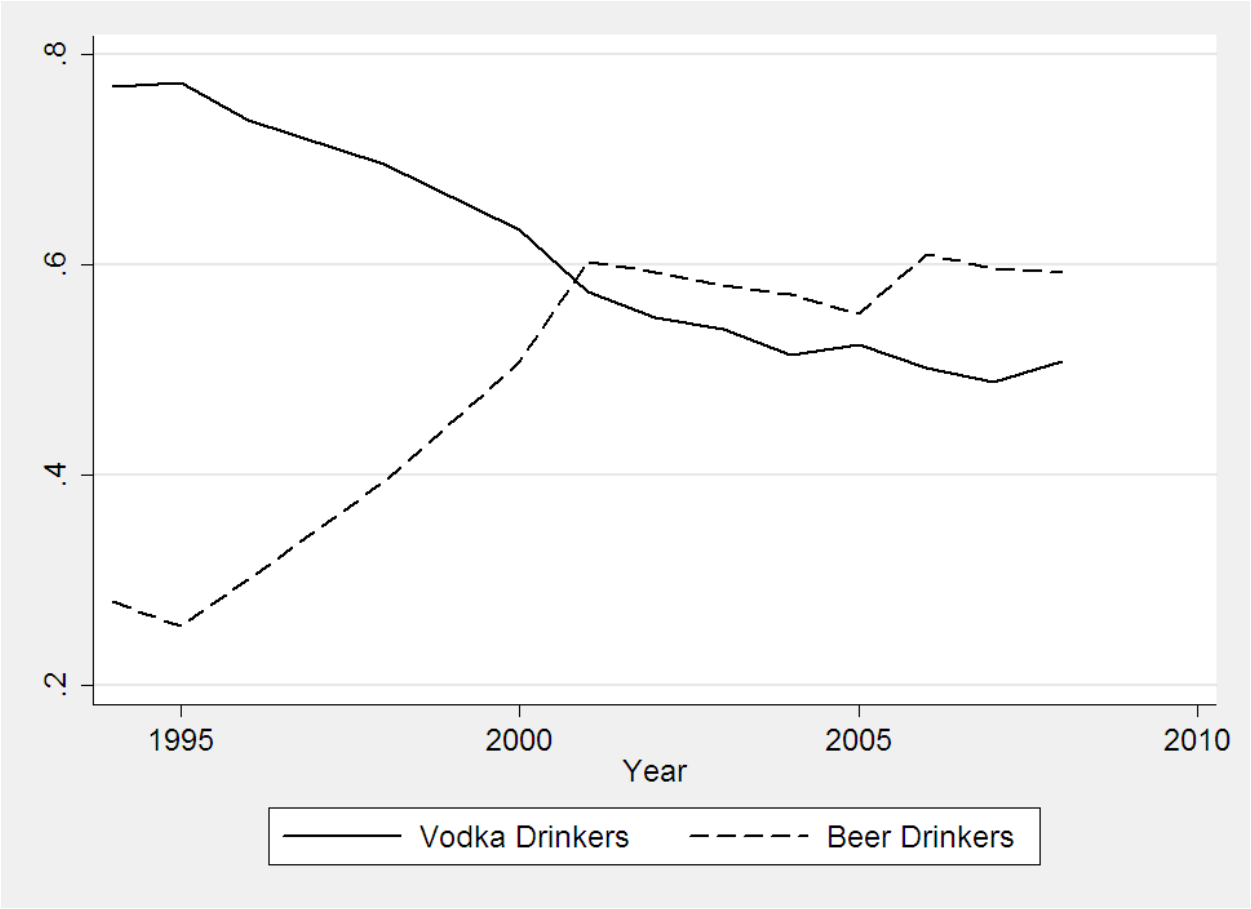

Figure 3. Monthly Volume of Beer Consumption

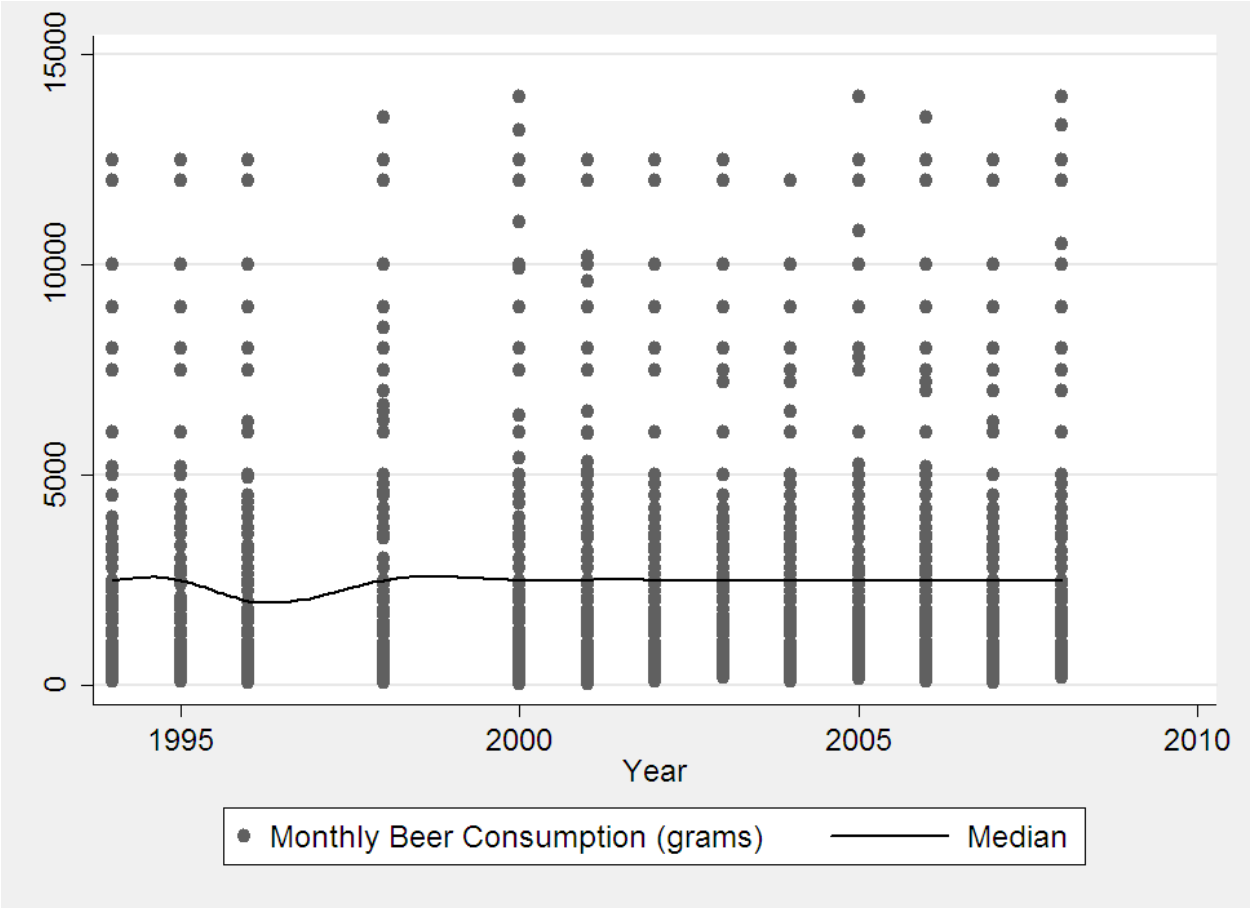

Figure 4. Gender Differences in Beer Consumption

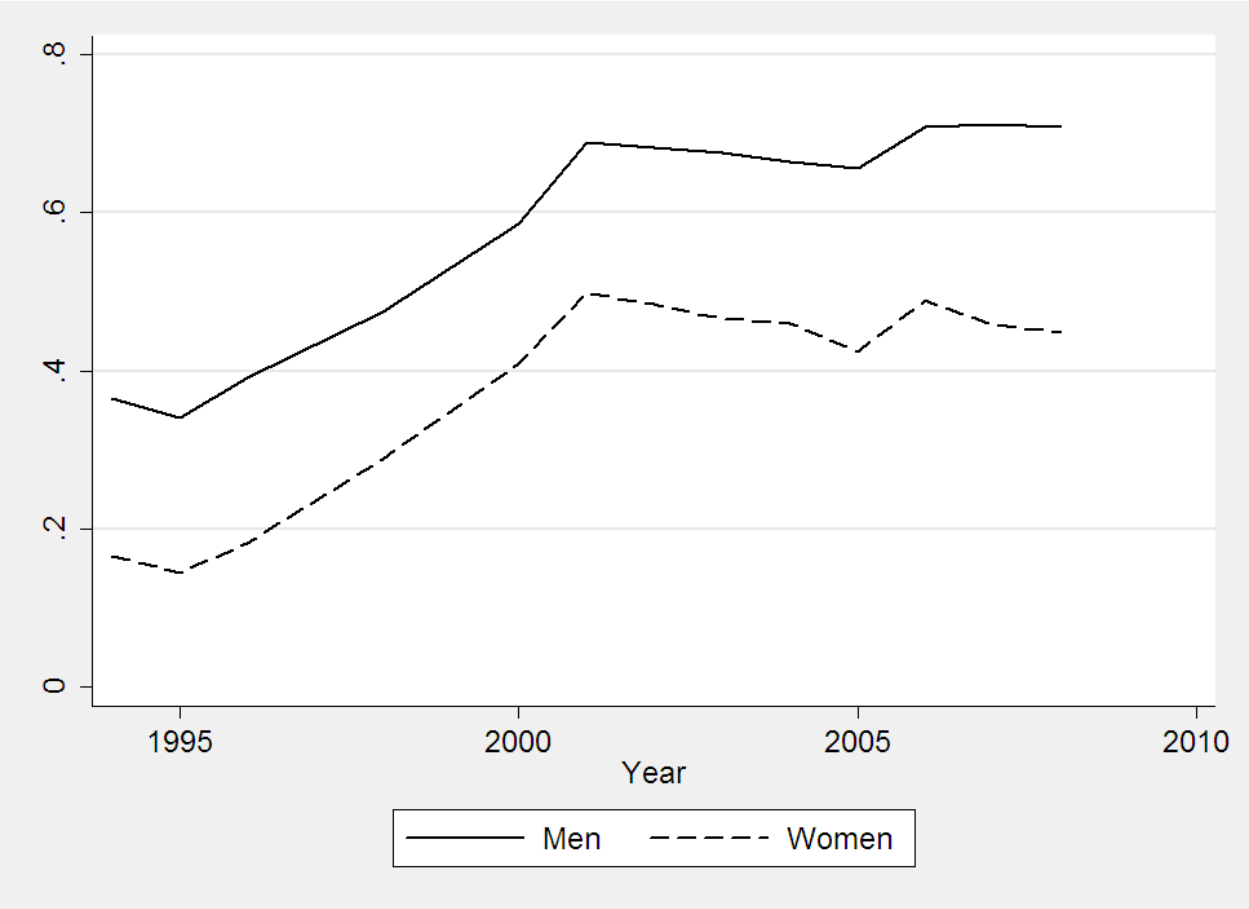

Figure 5. Fraction of Beer Drinkers per Region

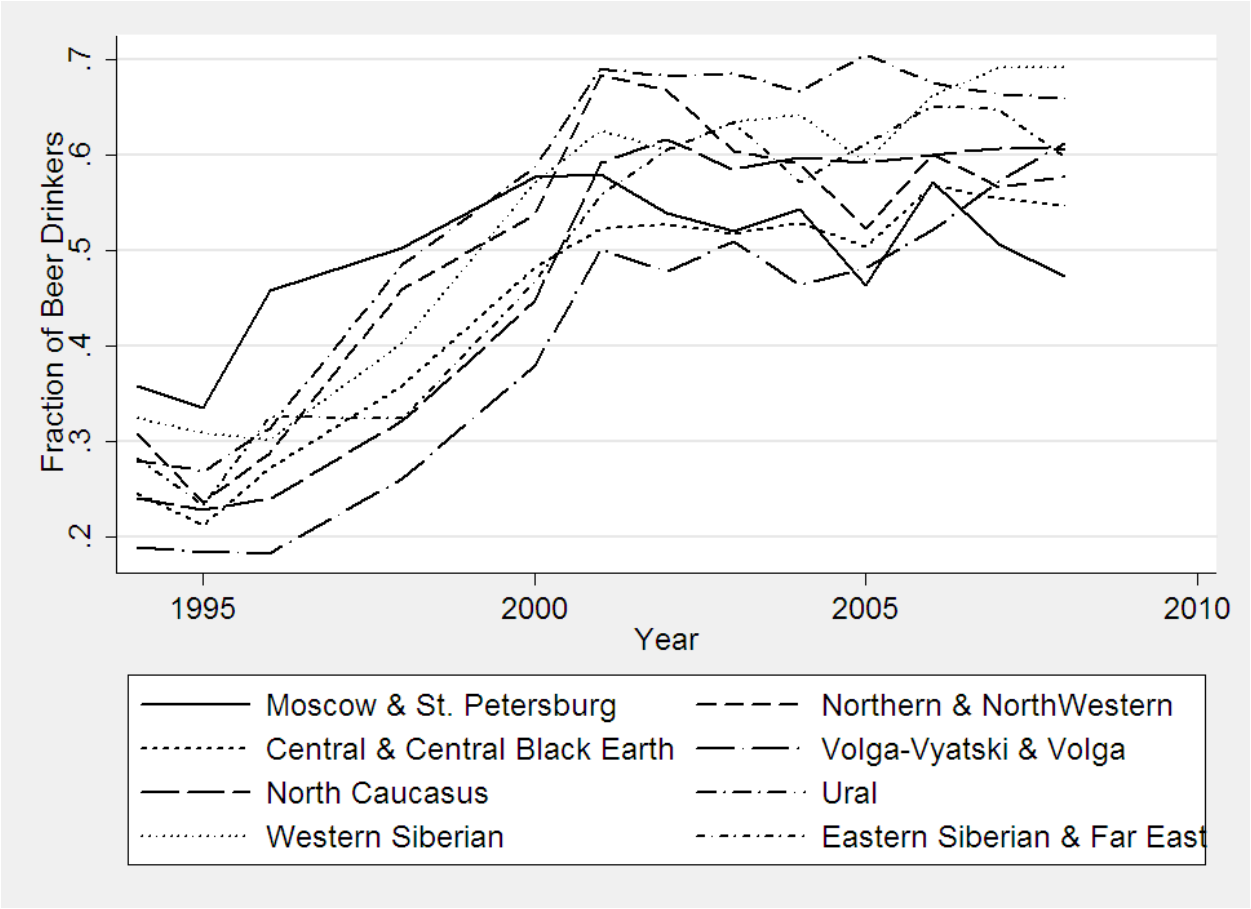

Figure 6. Fraction of Beer Drinkers per Cohort

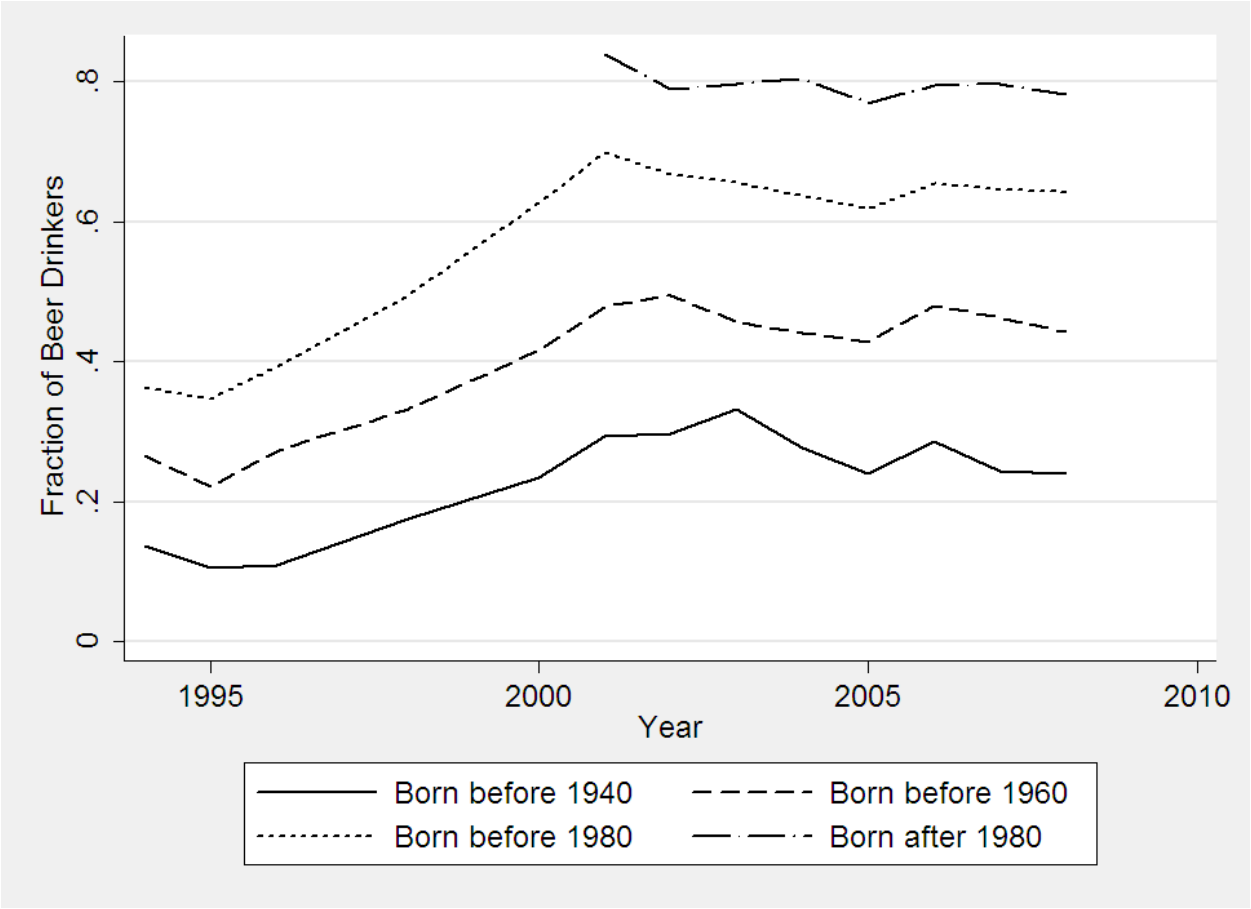
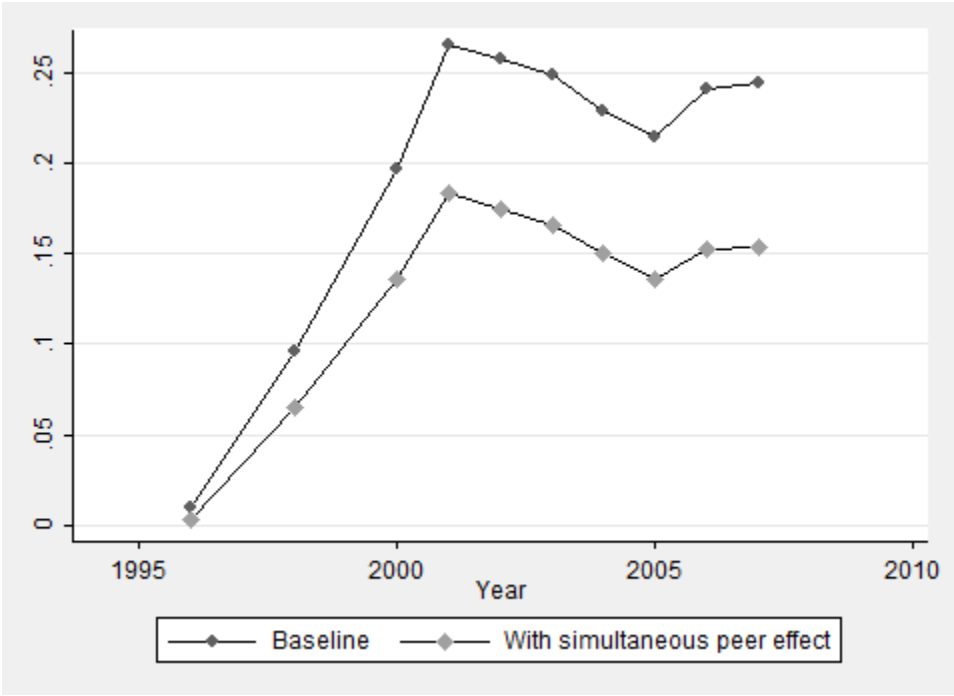
Figure 7. Pattern of Time Dummies With and Without Peer Effects