An effect of valuable skills on drinking patterns in contemporary Russia

A. S. Skorobogatov

National Research University Higher School of Economics
Saint-Petersburg, Sedova, 55/2, Russia
Tel: +7 921 320 68 27
E-mail: skorobogat@mail.ru

Abstract

The principal result of this research is that the valuable skills have a negative and significant effect on alcohol abuse. We found that a higher professional level is consistent with a stronger negative relationship between earnings and alcohol-related behaviors and problems. The explanation of the result is proposed in that the pecuniary losses imposed on an individual by alcohol-related physical inability are positively conditioned by the valuable skills.

Keywords: Russia; transitional economy; human capital; health behavior; drinking pattern; addictions

JEL classification: I1; I2; J2

INTRODUCTION

One of the recent papers on this topic begins by pointing to the ‘legendary’ alcohol consumption in Russia and the related health problems (Baltagi, Geishecker 2006), which reflect the well-known facts. Researchers of Russian drinking history note an abrupt jump in alcohol use since the beginning of the transitional period (Nemtsov 2000) which has been playing central role in mortality crisis in Russia among working men (Leon et al. 2009; Norstrom 2011). Allowing for the purely exogenous nature of the driving forces of this change, exploring the drinking patterns on data for this period could make it easier to cope with the endogeneity problem and the issues of parameters identification which it entails.

While the established theory in the field of health economics predicts an inverse relationship between drinking and human capital, the mentioned facts seem puzzling in light of the cross-country statistics according to which Russians, as a whole, have high educational attainments. This paper attempts to clarify the relationship between human capital and drinking patterns. The former we treat as valuable skills, i.e., as any abilities, congenital and acquired, which yield additional earnings to their owner.

The aim of this paper is to test the hypothesis which claims that the higher the value of an individual’s skills, the stronger their incentive to abstain from alcohol. The intuition behind the hypothesis is related to the opportunity cost of physical inability. As far as good physical condition is necessary for realizing valuable skills (the latter are the source of additional earnings) an owner of more valuable skills, other things being equal, has a stronger pecuniary incentive to support their physical ability.

This intuition can be elaborated on various directions one of which is related to alcohol. The latter is particularly relevant for checking the intuition owing to its immediate weakening effect on physical ability. When an individual is aware of this relationship, he/she abstains from alcohol depending on the return to their skills. Thus, valuable skills affect one’s drinking pattern through the expected effect of their health on the return to the skills.

BACKGROUND

Since the classical work of Grossman (1972) which, in particular, claimed positive link between health behavior and education a number of explanations of the link were developed by researchers. Cowell (2006) marks out three general classes of explanations of this positive correlation: efficiency mechanism which suggests that health is a smooth continuous function of education via productive efficiency as to health – efficiency for a given set of inputs – or allocative efficiency – efficiency at allocating inputs to the health production function; unobserved heterogeneity in which case unobserved variables, e.g. time preference (Fuchs 1982), are correlated with both the schooling and health decisions; and future opportunity costs of health-impairing behaviors which induce the educated to reduce unhealthy activities that might limit their earnings capacity by making them ill in the future.

Grossman’s work has given rise to, at least, the first and the third approaches. Health is treated in the Grossman’s model as a consumer and investment good. In its latter form its equilibrium quantity corresponds to equality between opportunity cost of health and its marginal utility (1972; Cawley, Ruhm 2012). Human capital raises productivity while health determines time available for utilization of human capital for earning income (Lye 2010).

Such a treatment of health has been elaborated in studies where opportunity costs of health-impairing behaviors were considered. One of them is opportunity cost of time (Grossman 1972; Dee 2001; Cowell 2006; Skorobogatov 2012). Cowell (2006) has developed the three period model of the future opportunity cost in which “the combined effect of the future health consequences of the unhealthy behavior on being alive in period three together with the effect of education on future wages” is analyzed (p. 127). For identification of the interest effect, Cowell considered discontinuous jumps in earnings as a result of a degree effect using discrete factor approximation (quasi maximum likelihood) estimator (Mroz 1999). The advantage of this estimator against the usual ML is weaker bias in case of incorrect distributional assumptions, though given the normality assumption is correct ML is strongly preferable with respect to both consistency and efficiency. Quasi ML was used by Cowell for solving the identification issues related to the unobserved heterogeneity, namely, for the sake of not imposing the normality assumption on the unobserved heterogeneity contained in an error term which allows one to control for this heterogeneity.

Using this estimator three dependent variables – smoking, binge drinking, and binge drinking frequency – were regressed by the interest degree variables and controls for effects implied by the alternative models – schooling and a number of personal and other controls to separate influence of unobserved heterogeneity. Many of the controls were instrumented so that as many as eight equations were estimated with the results that degree effects were significant for binge drinking and in still more for smoking. The main idea and conclusion were that people take care about their health depending on their long-term earnings perspective.

The hypothesis tested by Cowell has much in common with that tested in this paper. Common in research question is testing the mechanism relating human capital to health behavior; in interest mechanism – opportunity cost in terms of forgone earnings; in estimating technique – using instruments. The main difference in interest mechanism is opportunity cost in terms of current rather than future forgone earnings. It entails difference in prediction, namely, smoking is to be much less important factor than binge drinking.

The idea behind the hypothesis tested in this paper is that skills affect the drinking pattern via the earnings (Skorobogatov 2012). Control for the efficiency mechanism can be accomplished through inclusion of schooling, and control of unobserved heterogeneity at some extent can be made via inclusion of a number of personal and other controls; still more with this respect may be done using data with panel structure, but by the reasons explained below we use cross-section here.

The hypothesis expands upon some established ideas. One of these is the time allocation theory (Becker 1965) according to which leisure has an opportunity cost in units of forgone earnings. Others are related to human capital, namely that human capital increases the value of a human being (Schultz 1968) and their earnings (Mincer 1974). We adapt these ideas, treating physical inability as forced leisure which imposes on an individual opportunity cost depending on the value of their skills.

Our main prediction can be valid given various drinking measures do affect economic condition, including employment and wage. Bray (2005) using the same quasi ML estimator, for addressing sample selection issue, explored effect of alcohol use on the return to education and work experience. In his wage equation, Bray used three vectors related to demographics, human capital, and health status. Drinking was assumed to affect wage through health and human capital accumulation. The result was that heavy, rather than moderate, alcohol consumption adversely affects the return which suggests “different effect of alcohol use versus alcohol abuse on human capital accumulation and health” (p. 288). As a whole, his results support the inverted U-shape link: the return to education and experience is positively affected by moderate alcohol use, while heavier drinking reduces this gain.

The inverted U-shaped link was supported in a number of studies. Barrett (2002) having tested the relationship drinking–health–productivity–earnings inferred presence wage premium for drinkers and wage penalty for heavy drinkers. Heaviness of drinking was measured by Barrett via amount of drink at a single sitting as it was stated to be more strongly correlated with health effects than volume consumed per a period or frequency of drinking. Supporting an inverse U-shaped relationship between alcohol use and wages Barrett deduced optimal amount of drinks per day as well as threshold between moderate and heavy drinking. Using several indicator variables on the base of drinking frequency MacDonlad and Shield (2001) also supported the inverted U-shaped link. Similarly, in Srivastava (2010) frequent bingers experience reduced earnings whereas non-bingers and occasional bingers have wage premium over abstainers. On the data of RLMS the U-shaped link was supported in work of Kim and Roshin (2009).

Lye and Hirschberg (2010) in their literature review on the topic pointed to the wide-spread explanation of the inverted U-shaped curve according to which moderate alcohol use enhanced health and thereby increased time available for work. At the same time, they pointed to the alcohol-income puzzle consisting in that wage bonus was associated with moderate alcohol use while health effect related to moderate drinking was very little. It implies that there is to be some omitted variables associated both to alcohol use and human capital. They may include personality traits and sociability related to alcohol use. Positive impact of moderate drinking can partially explained by unobserved heterogeneity. Also ex-drinkers are misclassified as non-drinkers.

French et al. (2011) presents a review of alcohol-performance studies. Pathways to good and poor performance are distinguished as results of moderate drinking and alcohol misuse. The latter is described via behaviors like weekly or more frequent binge drinking and alcohol dependence. In this context, employment problems are considered such as being fired or laid off from a job, being unemployed, conflict with a supervisor/co-worker. Endogeneity and omitted variables problems are addressed by fixed effects models which require panel data. Significant negative impact of alcohol misuse on employment conditions is displayed.

The similar result is contained in Keng and Huffman (2007), namely, that problem drinking including binge one has a negative effect on earnings, while earnings have no significant effect on choice of binge drinking. As marked by the authors “binge drinking has the short-ran consequences of impaired judgment and physical coordination resulting from intoxications and hangovers” (p. 36). In this paper, as one of the personal controls is used religion affiliation on the base of its opposing alcohol use.

Long-term relationships between drinking and work performance are analyzed in Griffin et al. (2011). According to their results, abstaining from all drugs for 12 months in adolescence is associated with higher mean wage, specifically higher probability of a legitimate occupation, seven years later. In Sloan and Grossman (2011) the impact of alcohol on labor market outcomes is differed for blacks and whites, while being more favorable for the latter. What is negative for black men about drinking effect is not so for black women and whites which implies that the same drinking pattern may differently affect different races.

As to the explanation of the pathways from alcohol use to good work performance, Peters (2009) analyzed drinking as a way of investing in social capital in the American army. His result was that officers who drank had wage bonus comparing with non-drinkers. Their drinking premium was more than that of enlisted personnel. In a more general perspective, Cawley and Ruhm (2012) point out to the peer effects which may underlie the alcohol-performance pathways. They distinguish the three channels – common constraints, information spillovers and ‘bandwagon effect’.

Challenge to the literature claiming positive alcohol-performance link as well as solving the alcohol-income puzzle are contained in work of Cook and Peters (2005). For solving the paradox, they applied to exploring the link between prices and drinking. As a measure of drinking they used frequency of binge drinking, specifically 6 drinks or more on a single occasion. According to their results, prevalence of full-time work increases with alcohol prices that the authors explained via an inverse effect of alcohol use on labor supply. Hence a decrease of alcohol use due to the prices rise is to increase labor supply. Further, they revealed a positive association between alcohol prices and earnings of full-time workers which suggests an inverse effect of drinking use on not only employment but also on performance of the employed. Alcohol is treated by the authors as a normal good, and thereby positive association between its consumption and earnings can be explained by the income effect. With their conclusions the results of Stockwell et al. (2011) agree who explored effect of minimum pricing on alcohol use. Also there is match between their conclusion and positive income elasticity of alcohol demand supported by hundreds empirical studies (Cawley, Ruhm 2012).

Another strand in the literature on the prices-drinking link is related to the theory of rational addiction initiated with the pioneering work by Becker and Murphy (1988). This theory suggests that a rational addict allocates their budget to maximize their life-time utility given the addictive good pays depending on the stock of its past consumption. The most interesting 'counter-intuitive' implication hereof is that long-term price elasticity is positively linked with addictiveness of a good. In empirical models, consumption of an addictive good was regressed by its past consumption as well as past and expected prices. Its elaboration in the form of 'the two-stock model' was proposed in which two stocks of past consumption were introduced into the model – with the adjacent complementarity and substitutability (Cawley, Ruhm 2012). On Russian data the hypothesis based on the theory of rational addiction was tested in Baltagi and Geishecker (2006).

A strand in the literature which is in an immediate link with the hypothesis considered in this paper is related to impact of economic condition on drinking. An important paper here is that by Dee (2001). As measures of alcohol use Dee employed participation (alcohol use over the past 30 days), log number of drinks per month, chronic drinking (at least, 60 drinks per month), binge drinking (at least, one occasion in the last month of consuming five or more drinks in a row). An identification issue of purging confounding effect of the unobservable determinant was addressed by the author via the fixed effects. The result of the study is that alcohol abuse was induced by economic recessions. Binge drinking is strongly countercyclical. Economic recession induces drinking among both getting unemployed and remaining employed. Thus, income effect is dominated by other factors, namely opportunity cost of time and psychological stress. Implicit price of binge drinking falls during recessions so that this factor prevails if binge drinking is more among unemployed than those remaining employed. Overall alcohol consumption and binge drinking behave differently during the recessions – the former dropped while the latter rose.

Dee contrasted modeling different measures of alcohol use, namely consumption volume to abusive consumption. In addition, it makes sense with respect to the effect of alcohol use on alcohol-related problems as the results of the study of Danielsson et al. (2011) suggest. In particular, this study concludes that the bulk of the alcohol-related problems are accounted for by the minority of frequent heavy drinkers. Thus, alcohol-related problems can serve appropriate proxies of alcohol abuse and/or dependence.

Alcohol consumption of Russian-speaking people was investigated by Nemtsov (2000), Shiff et al. (2005), Pomerleau et al. (2005), Baltagi and Geishecker (2006), Kim and Roshin (2009), Kusmitch' and Roshin (2007) but without referring to the skills or associated variables. In the general research on drinking and its impact on mortality in Russia, Denisova (2010) discovered a trend, albeit a weak one, of the substitution of harder spirits with softer drinks for the more educated and well-paid. And according to the results of Leon et al. (2009), hazardous drinking is prevalent among those with low educational attainment and poor economic positions.

This paper starts with these results, examining the effect of the value of skills in terms of its pecuniary return in more detail. Unlike the well-known hypothesis, we analyze the effect of human capital on alcohol-related health behavior rather than that of health condition on earnings. At the same time, the latter hypothesis is adapted within our own hypothesis, as it suggests that, while that relationship is present, people are aware of it and accommodate it in their drinking decisions.

METHODS

DATA AND DATA SOURCES

Our research is based on data of the Russian Longitudinal Monitoring Survey (RLMS). This nationally representative survey has been performing annually since 1994 for the sake of analysis of the transitional processes in Russia. The survey was initiated by the Carolina Population Center at the University of North Carolina and untill recently was coordinated by the same Center as well as the RAS Sociology Institution (Moscow). Now the main coordinator of the survey is National Research University Higher School of Economics (Moscow).

We used data from the 18th round of the RLMS for the year 2009. The main reason of such a choice of the round is that it is only this round that contains the self-rated professional mastership most relevant for our purposes. The size of the representative sample is 9578 observations. All the rounds of RLMS contain rich information about various aspects of respondents' lifestyle, albeit subsequent rounds somewhat differs from each other with respect to the included questions. The same dataset was used by a number of researchers including those studying consumption of alcohol and other drugs in Russia (Baltagi, Geishecker 2006; Denosova 2010; Kim, Roshin 2009; Kuzmitch, Roshin 2007).

As for demographics we used data of Rosstat.

MEASURES

Labor earnings and professional mastership

As earnings variables we used logarithms of total self-reported labor earnings including those from the main job and the two possible secondary jobs of a respondent.

The 18th round of the RLMS contains a block of questions dedicated to various aspects of human capital. In particular, the respondents were asked to rate their professional mastership on the nine-level scale. This measure of human capital may be more appropriate than conventionally used years of schooling on the account of, on the one hand, ongoing decrease of the quality of Russian education so that it plays rather signaling function than that of production of human capital, on the other hand, the transitional processes owing to which many educated people are unemployed or work not by their specialty. It can explain surprisingly low correlation coefficient between years of schooling and self-rated professional mastership (0.33) (one respondent with zero schooling rated her professional mastership the highest level, while there were 16 respondents with completed higher education who rated their professional mastership the lowest level). Mean value of professional mastership for the sample of 7740 respondents who replied the question is 5.68, while median value is 6.

Also as proxies for the valuable skills, we used log earnings of those on different levels of self-rated professional mastership. These interaction variables were derived as the earnings variable multiplied by dummies for professional levels. In an altertanive model specification we used the interaction variable for earnings of higher median professional level.

Drinking pattern

For drinking pattern we used a number of dummy variables reflecting circumstances of alcohol intake.

On the streets. Respondents were asked if they drank on the streets. Among men about 13% of men and among women only 3,6% answered positively this question.

At workplace. Another question was whether a respondent drank on workplace. This was the only case women more often than men replied positively so that the corresponding values were 8.5% and 7.7%.

Without or before eating. Questions whether one drank without or before eating were replied positively 39% and 32% among men and 19% and 12% among women.

Self-reported alcohol-related problems

At work. Respondents were asked if they had alcohol-related problems at work. Slightly more than 2% of respondents had these problems. Among male respondents this value is about 4%, for female persons it is 0.4%. Apart from gender difference in alcohol use, this tenfold gender difference can be explained by significantly less responsibility of women in Russia for well-being of families.

At home. Also respondents were asked if they had alcohol-related problems in their families. More than 8% of respondents had this kind of problems. Among male respondents this value is more than 14.5%, while among women 2.4% reported existence of these problems.

In health. Respondents were asked if they had alcohol-related problems in health. About 8% of respondents replied positively. Among male persons this value is in excess of 12%, while among female ones it is slightly more than 4%.

Other problems. Finally, respondents answered if they had other problems associated to alcohol. In this case only about 0.9% of participants reported existence of these problems among which men amounted about 1.5% and women did 0.33%.

Our hypothesis implies that a rational individual is concerned about the immediate effect of alcohol intake on their health, and therefore on their work. Hence, from amongst the available data on drinking the self-reported effects of alcohol intake on a respondent's work should be considered as the most relevant for testing the hypothesis. The second most appropriate measure is the family problems, since the latter typically arise from the former through material deficiency made by alcohol consumption. Correlation between these two measures is 0.43 which, though not so much, is the highest comparing with correlations of the work problems with the rest problems (0.29 for health problems and 0.31 for other problems).

Thus, as our dependent variables, we used dummies for whether one has alcohol-related problems in terms of work, family, health and other matters.

Controls for personal characteristics

Gender. 42.3% of respondents reported their sex was male. This value drops to 40.9% for a sample restricted by those older 18 years in which age in Russia a person is allowed to buy alcoholic beverages.

Age and squared age. Mean age of the participants was 40.7 years (std. dev. equals to 22.17). At least, while Russia is concerned, age has a strong effect on drinking variables (see, e.g., Schiff et al. (2005)). Along with this measure we used squared age on the consideration of nonlinear impact of age on our dependent variable. Age typically positively affects alcohol consumption, but its marginal effect is inversely linked with it, i.e. additional year of age contributes decreasing positive effect on alcohol intakes.

Marital status. About a half of respondents (49.5%) reported that they were married. At the same time, because of relative deficincy of male people in Russia this percent is significantly more for male persons amounting 56.5% versus 44.4% for female ones so that a single woman is more typical phenomenon in Russia than a single man.

Health status. Respondents were asked to rate their health on the five-level scale from 'very good' to 'very bad'. For the convenience of using and interpreting this measure, we generated a new variable 'health' on the base of the latter one with the scale from 'bad' to 'good'. Mean value of this measure for the full sample is 3.23, while being slightly more for male persons (3.35) than for female ones (3.14).

Religious affiliation. Religious affiliation was not asked directly in this round, but affiliation with Islam could be credibly evaluated from the reported nationality, and it is Islam that is of most interest among religious confessions for our research because of its severe prohibitions on alcohol. We generated measure 'muslim' by the following self-reported nationalities: Tatar, Kabardinian, Azerbaijanian, Daghestan, Chechen, Tajik, Kazakh, and other lesser muslim national minorities. Share of 'muslims' in the sample was 7.4% which is generally conforms to estimates of quantities of muslims in Russia.

Respect status. As a measure of social status we took answers of respondents how they rated their position on the nine-level scale from the those whom people do not respect to very respected ones. Mean value of the measure is 6.18, which is quite close to median value, with a tiny difference between genders.

Controls for labor market conditions

Population size (log). As a control of labor market conditions we used log population size of a respondent's locality because this measure can serve as an important contributing factor of having well-paid job (Combes et al. 2008).

Average real income. Another measure of labor market conditions here is average real income which we calculated based on data of Rosstat. There we took data on ratio between per capita incomes and cost of minimal basket of foodstuffs across all the Russian regions (Prices 2010, p. 84-86) These data we used for characteristics of a respondent's locality.

Sample statistics of the drinking variables

Tables 1-4 contain some summary statistics of our dependent variables grouped by most important personal controls and interest variables. These tables display some statistical regularities. For males (Table 1) higher median group by earnings (hme) shows lower participation rates across all the drinking dummies. For females higher group by earnings, conversely, shows higher participation in most cases; only for two dummies, 3d and 6th, higher earnings female group displays tangibly less participation rate; only in one case (2d) females displayed higher participation, namely well-paid women more often drink at workplace than men with any fortunes. Thus, sign of the link between earnings

Table 1: Sample statistics of drinking patterns: by gender and median log earnings (indicator)

	male		female	
	lme	hme	lme	hme
a respondent drinks:				
on the streets	0.132	0.0975	0.0285	0.0281
	(0.338)	(0.297)	(0.167)	(0.165)
on workplace	0.113	0.0886*	0.107***	0.136***
	(0.316)	(0.284)	(0.309)	(0.343)
without eating	0.433***	0.334***	0.227***	0.130***
	(0.496)	(0.472)	(0.419)	(0.336)
before eating	0.348***	0.286***	0.116**	0.123***
	(0.477)	(0.452)	(0.320)	(0.329)
alcohol-related problems:				
on workplace	0.0594	0.0266	0.00300	0.00625
	(0.237)	(0.161)	(0.0548)	(0.0789)
at home	0.195***	0.0875	0.0270	0.0109
	(0.397)	(0.283)	(0.162)	(0.104)
with health	0.132	0.0676**	0.0315**	0.0375**
	(0.338)	(0.251)	(0.175)	(0.190)
the other	0.0191	0.0122	0.00300	0.00625
	(0.137)	(0.110)	(0.0548)	(0.0789)
Observations	471	903	666	640

standard deviations in parentheses; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

and alcohol use is rather different for genders. A preliminary account of such a difference may be related to what is the main effect of earnings with respect to drinking. For men it may be a financial opportunity cost of drinking so that higher earnings would be consistent with higher opportunity cost of alcohol use/abuse, whereas for women income effect may be of most importance. The cause for this difference may be the fact that it is men who as a rule bear main responsibility for a family welfare so that a man takes more care about his earnings than a woman.

In most cases (Table 2) age under 25 years displays more share of the positive response for various drinking variables in higher earnings group which is obviously related to the fact that they are not employed nor married. Only the most active age between 39 and 59 displays consistent inverse link between earnings and drinking participation rate. The latter age displays also more difference between the earnings groups by problems at workplace and in health.

There displays consistent drop in drinking participation along with the growth of educational attainment (Table 3). Difference in drinking problems in favor of lower earnings groups grows weaker from lower educational level to higher one.

Drinking participation tends to fall from lower median professional level to higher one (Table 4). Earnings tend to induce drinking participation in lower professional group and once more higher one, but alcohol problems are inversely related to earnings in both groups. Difference in drinking problems in favor of lower earnings groups grows weaker from lower professional level to higher one. Thus, contrary to the intuition and the hypothesis to be tested in this study, the mean values in Tables 3-4, as a whole, show stronger inverse association between earnings and drinking for less educated and professional workers. It could mean that either unskilled workers are more financially induced not to abuse alcohol or drinking entails more financial losses or, at last but not least, there is some mixture of these mutual relationships.

Table 2: Sample statistics of drinking patterns: by age and median log earnings (indicator)

	age<25		24<age<40		39<age<60		age>59	
	lme	hme	lme	hme	lme	hme	lme	hme
a respondent drinks:								
on the streets	0.209	0.239	0.0762	0.0797	0.0368	0.0260	0.0172	
	(0.408)	(0.428)	(0.266)	(0.271)	(0.188)	(0.159)	(0.131)	
on workplace	0.0798*	0.0755	0.103*	0.126***	0.122***	0.0984***	0.112*	0.121
	(0.272)	(0.265)	(0.304)	(0.332)	(0.327)	(0.298)	(0.317)	(0.329)
without eating	0.282**	0.333*	0.328***	0.262***	0.325***	0.213***	0.250*	0.318*
	(0.451)	(0.473)	(0.470)	(0.440)	(0.469)	(0.410)	(0.435)	(0.469)
before eating	0.362***	0.453***	0.246**	0.266***	0.164	0.133***	0.112	0.0909
	(0.482)	(0.499)	(0.432)	(0.442)	(0.371)	(0.340)	(0.317)	(0.290)
alcohol-related problems:								
on workplace	0.0123	0.0189	0.0264	0.0207	0.0348	0.0174	0.00862	
	(0.110)	(0.136)	(0.161)	(0.143)	(0.183)	(0.131)	(0.0928)	
at home	0.0613	0.0503	0.0997	0.0510	0.116	0.0637*	0.0517	0.0303
	(0.241)	(0.219)	(0.300)	(0.220)	(0.321)	(0.244)	(0.222)	(0.173)
with health	0.0307	0.0503	0.0616	0.0606	0.0909	0.0507*	0.0862	0.0606
	(0.173)	(0.219)	(0.241)	(0.239)	(0.288)	(0.219)	(0.282)	(0.240)
the other	0.00613	0.00629	0.00880	0.0128	0.0116	0.00868*	0.00862	
	(0.0783)	(0.0793)	(0.0935)	(0.112)	(0.107)	(0.0928)	(0.0928)	
Observations	163	159	341	627	517	691	116	66

standard deviations in parentheses; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 3: Sample statistics of drinking patterns: by education degrees and median log earnings (indicator)

	Primary school		High school		Higher school		Postgraduate degree	
	lme	hme	lme	hme	lme	hme	lme	hme
a respondent drinks:								
on the streets	0.134	0.161	0.0635	0.0691	0.0524	0.0450		
	(0.341)	(0.368)	(0.244)	(0.254)	(0.223)	(0.208)		
on workplace	0.102	0.124	0.104***	0.0895***	0.136***	0.139***		
	(0.303)	(0.331)	(0.306)	(0.286)	(0.344)	(0.346)		
without eating	0.344***	0.343	0.324***	0.279***	0.241***	0.176***		0.167
	(0.477)	(0.476)	(0.468)	(0.449)	(0.429)	(0.381)		(0.389)
before eating	0.287	0.372**	0.206***	0.230***	0.178*	0.160***		0.0833
	(0.454)	(0.485)	(0.405)	(0.421)	(0.384)	(0.367)		(0.289)
alcohol-related problems:								
on workplace	0.0764	0.0365	0.0203	0.0170	0.0105	0.0157		
	(0.267)	(0.188)	(0.141)	(0.129)	(0.102)	(0.124)		
at home	0.197	0.109	0.0877	0.0623*	0.0524	0.0313		
	(0.399)	(0.313)	(0.283)	(0.242)	(0.223)	(0.174)		
with health	0.146	0.102*	0.0673	0.0532	0.0366	0.0470		
	(0.355)	(0.304)	(0.251)	(0.225)	(0.188)	(0.212)		
the other	0.0127	0.0146	0.0102	0.0113	0.00524	0.00587		
	(0.113)	(0.120)	(0.100)	(0.106)	(0.0724)	(0.0765)		
Observations	157	137	787	883	191	511	2	12

standard deviations in parentheses; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 4: Sample statistics of drinking patterns: by median professional group (indicator) and median log earnings (indicator)

	lower median professional group		higher median professional group	
	lme	hme	lme	hme
a respondent drinks:				
on the streets	0.0983	0.103	0.0349	0.0431
	(0.298)	(0.305)	(0.184)	(0.203)
on workplace	0.119***	0.118***	0.0977***	0.103***
	(0.324)	(0.323)	(0.297)	(0.304)
without eating	0.315***	0.276***	0.307***	0.226***
	(0.465)	(0.447)	(0.462)	(0.418)
before eating	0.240***	0.273***	0.170*	0.179***
	(0.428)	(0.446)	(0.376)	(0.384)
alcohol-related problems:				
on workplace	0.0343	0.0213	0.0163	0.0147
	(0.182)	(0.145)	(0.127)	(0.121)
at home	0.103	0.0558	0.0814	0.0533*
	(0.304)	(0.230)	(0.274)	(0.225)
with health	0.0640	0.0542	0.0791	0.0544*
	(0.245)	(0.227)	(0.270)	(0.227)
the other	0.0140	0.00985	0.00465	0.0102
	(0.118)	(0.0988)	(0.0681)	(0.101)
Observations	641	609	430	882

standard deviations in parentheses; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

EMPIRICAL ANALYSIS

Identification strategy

The hypothesis would be supported if the earnings in the upper levels were correlated with the drinking effect dummies in a different way than they are in the lower levels. In the first case, earnings are to be a less favorable factor of the positive outcomes, i.e., if a negative correlation is observed in both upper and lower levels, then in the upper ones it is to be stronger; if a positive correlation is observed for both, then in the upper ones it is to be weaker. The best result, with respect to fitting the hypothesis' prediction, would be if a negative correlation were observed in the upper levels, while a positive correlation or statistically insignificant relationship were observed in the lower levels.

As an initial version of our empirical model, we used a simple probit regression, in which logs of the total labor earnings for all levels of professional mastership, excluding the first (reference) level, were used as a vector of interest covariates, namely, eight ones (based on the none-level initial variable). In other specification of the simple probit model, we used only one interaction term for log labor earnings of higher median professional group. According to our hypothesis, the higher a respondent's professional level, the stronger the inverse relationship between their labor earnings and the dummy for alcohol effect.

Since we are interested in causation from the return to skills to a drinking pattern the conventional endogeneity issues arise, namely, those related to omitted variables, measurement errors, and a reverse causality. In the first case, an estimation bias arises if an omitted variable is correlated with both a dependent variable and one of the regressors. In Wooldridge (2002, p. 470) it is shown that in binary response models failure to include some variables would make estimates biased even given the neglected heterogeneity, i.e. if these omitted varables were not correlated with those included ones.

Such an omitted variable may contain the unobserved heterogeneity of respondents which implies that some of their unobserved characteristics affect both interest regressors and dependent variables. The well-known example of such unobserved heterogeneity is the intertemporal preferences which may impact educational decisions as well as health behavior (Fuchs 1982; Cowell 2006) so that the positive associations between the latter two behavioral characteristics may not reflect real link between them.

A measurement error may arise, in particular, when one deals with data based on surveys. Self-reported data are vulnerable to self-appraisal-related bias. Subjective nature of variables based on such self-reported values may cause the observations diffuse which entails, other things being equal, underestimation of the variables' coefficients. Data at our disposal is subject to such measurement errors, and it is most important in case of self-rated variables such as professional level or labor earnings. Here values of the variables depend on memory, incentive to give exact values, and various prejudices on the part of respondents.

At last, and it is likely to be the most important issue, as the relationship between drinking variables and the value of skills is concerned, the reverse causality problem occurs, meaning that an individual may form their attitude to alcohol and face its effects before making a decision about investment in human capital (Isabel, Molina 2007). In a more general sense, alcohol use may affect both human capital and labor earnings including premium for skills via its effect on physical or mental abilities as well as social interactions of an individual. If, for example, we estimate earnings' effect on alcohol use in a simple regression we will observe associations reflecting mutual impacts instead of wanted causal one-way effect.

A necessary identification condition facing a probit model is the normalization of the equations to make them treatable as based on standard normal distribution functions of error terms (Wooldridge 2002). Such a normalization determines functional form of the regressions used.

Also like a linear regression, our simultaneous probit model must meet the standard range and order conditions. As our model contains two endogenous regressors, the exact identification makes it to be necessary to use two excluded instruments. At the same time, when dealing with a linear model imposing overidentifying restrictions makes it possible to apply Hansen J test of the validity of the whole of instruments. For this sake, we will use three instruments.

The reliability of the test requires relevance and validity of at least of one of the instruments which is evaluated on proper theoretical foundation as well as common sense. In this context, our key identifying premise is that our instruments affect a drinking pattern variable only through our endogenous variables, i.e. log total earnings and professional level. It is based on the following considerations. First is that at least one of our excluded instruments, namely, regional real income is to affect the drinking pattern no other way than via individual earnings. Hardly ever there is other channel of the effect. The same is likely to be true for the work experience variables.

Secondly, among the included instruments there are a number of covariates of the excluded ones which helps to identify the interest causation from the excluded instrument to drinking pattern variables via endogenous ones. In particular, gender and all the more so age variables are close covariates of work experience and they control for gender and age effects on drinking isolating them from the interest effect of experience. Its other covariates are the respect status and log population size. The former is correlated with both experience via the recognition of people and drinking through psychological frustration the unrespected feel, on the one hand, and close social interactions of the respected, on the other hand. Population size affects both work experience and drinking through employment opportunities and opportunity cost of leisure. The latter two controls are close covariates of regional real income too. Correlation between such characteristics of a region as its population size and real income is implied in the models of the new economic geography (Combes et al. 2008) and was displayed in a number of empirical studies. As for respect status as well as other personal controls such as schooling, health status, and religious affiliation they are obviously affected by the regional characteristics.

Schooling displays close associations with regional real income via financial and locational opportunities of education. At the same time, it controls for the famous effect of knowledge of the health production function (Grossman 1972), and its inclusion separates our regional external instrument's effect from that related to schooling. Health also depends on financial opportunities and location of hospitals and other public health establishments. And it is obviously correlated with drinking via both direct causations related to drinking decision and reverse one from drinking to health status. Finally, religious affiliation may play role because muslims are unevenly distributed across Russian territory, and at the same time this religion is known by its strong prohibitive commandments as to alcohol. Thus, correlation of muslim affiliation is to take place both with geographical and drinking variables.

Some personal controls which are frequently included in earnings or health equations have not been included in our regressions, namely, smoking and body mass index (BMI). Though these controls are the close covariates of our dependent variables and, in a lesser extent, our interest covariates we have not included them since they may be determined endogenously with alcohol use. Similar strategy of conscious exclusion of some important controls was used by Dee (2001). We have not controlled for smoking because we are interested in the human capital variables' effect on alcohol use if an individual is a smoker or not, and the same is true for BMI.

Thus, our list of included instruments containing personal and labor marker controls include gender, age and squared age, schooling, marital, health, and respect statuses, and dummy for muslim religious affiliation. Our excluded instruments include real regional income, work experience and squared work experience. These excluded exogenous covariates are implied to affect the return to skills via a spectrum of both subjective and objective opportunities of employment and reward, while not being related to the error term in the second stage equations.

Within our instrumental probit model, we employed Wald test of exogeneity of endogenous variables on the base of the estimated covariance vector of the error terms from the first stage and the second stage equations. As a whole, p-value of the test in excess of 10 means our failure to reject the null so that we should prefer the simple probit, at least for given specification.

To test the relevance and validity of excluded instruments, we proceed from the tests from GMM linear probability models. Unlike nonlinear models including probit one, a linear probability model as far as it is estimated with the GMM estimator allows to use the relevance and validity tests of instruments, while not producing heteroskedasticity-related inconsistent standard errors (Baum et al. 2003). Such a testing strategy was used, in particular, in Dujardin and Goffette-Nagot (2009).

Specifically, we employed F statistic of the joint significance of the coefficients on the excluded instruments as well as the standard partial R-squared, and the Shea's partial R-squared which takes into account intercorrelations among the excluded instruments. In a model with one endogenous regressor the well-known rule of thumb (Staiger, Stock, 1997) claims that F statistic is to be no less than ten. To the best of our knowledge, there are not conventional rules for F statistic or other criteria as for model with multiple endogenous regressors. So, as a first approximation, we have been started from the rule of thumb stated in Baum et al. (2003) according to which difference between the standard partial R-squared and the Shea' partial R-squared is not to be much. Much difference between the two determination coefficients is indicative of the weakness of an instrument.

Another test from the linear model is Hansen J test of overidentifying conditions. This criterion allows to test the null on the orthogonality of the excluded instrument to an error term. The p-value of the test in excess of ten means failure to reject the null which allows us to consider the instruments to be valid.

Thus, we address the identification issues by using the instrumental model, control covariates, general statements about the validity of the excluded instruments as well as full set of the tests on the instruments. All of these are designed to help us to identify the interest effect from the return to skills to alcohol use.

While using the instrument models we have to restrict our model as to the interest variables. Including as much as eight endogenous covariates involves the necessity to look for a lot of excluded instruments which entails both severe convergence difficulties and stronger bias due to inclusion of potentially weak instruments. So, instead of the eight interest covariates of our initial simple model, we have used log total labor earnings and the interaction term for the log total labor earnings and the higher median professional level. The latter is used as a proxy for valuable skills. In our alternative specification, professional level is used as a second endogenous variable. These endogenous variables are instrumented by covariates which are to impact the alcohol variables only through them. Our hypothesis would be supported given the coefficients on the skills endogenous variables were negative. While using the interaction skills endogenous variable it would mean that the skilled workers are more induced not to be involved in health-damaging drinking behaviors than the unskilled ones are. In case of using the professional level as the second endogenous variable the conclusion would be actually the same.

The simultaneous probit model and the conditional maximum likelihood estimator

Our structural model for a drinking pattern has a following form

$$\begin{aligned} D_i &= \mathbb{I}(D_i^* > 0) \\ D^* &= \mathbf{X}_1\boldsymbol{\alpha}_1 + \mathbf{Y}\boldsymbol{\beta} + \varepsilon \end{aligned} \tag{1}$$

where $\mathbb{I}(\cdot)$ is the indicator function

$$\mathbb{I}(D_i^* > 0) \equiv \begin{cases} 1 \text{ if } D_i^* > 0 \\ 0 \text{ if } D_i^* \leqslant 0 \end{cases}$$

and $\mathbf{Y}$ is given by a reduced form equation

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\alpha}_{21} + \mathbf{X}_2\boldsymbol{\alpha}_{22} + \boldsymbol{\eta} = \mathbf{X}\boldsymbol{\alpha}_2 + \boldsymbol{\eta} \tag{2}$$

In these equations D_i denotes our observed dependent variable, and D_i^* is a latent variable affecting the probability of the positive outcome of D_i; $\mathbf{X}_1$ is a $1 \times J$ vector of exogenous covariates including an intercept; $\mathbf{Y}$ is a $1 \times K$ vector of the endogenous regressors; ε denotes an error term in the second stage equation and $\varepsilon \sim N(0, 1)$ which implies the normalization of the model for the sake of its identification as a probit one; $\mathbf{X}_2$ is a $1 \times L$ vector of excluded instruments; $\boldsymbol{\eta}$ is a $1 \times K$ vector of error terms in the first stage equation for the endogenous regressors. In the previous subsubsection, we have described the nine included, the three excluded instruments, and the two endogenous variables so that $J = 10$, $K = 2$, and $L = 3$.

To fit a model (1), (2), we employed the conditional maximum likelihood estimator (Amemyia 1978; Hechman 1978; Wooldridge 2002). Joint normality of $(\varepsilon, \boldsymbol{\eta})$ given $\text{Var}(\varepsilon) = 1$ allows one to regress ε on $\boldsymbol{\eta}$ in the following way

$$\varepsilon = \boldsymbol{\eta}\boldsymbol{\Sigma}_{\mathbf{22}}^{\mathbf{-1}}\boldsymbol{\Sigma}_{\mathbf{21}} + e \tag{3}$$

where $e|\varepsilon, \boldsymbol{\eta} \sim N(0, 1 - \boldsymbol{\Sigma}_{21}^{\top}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21})$ and $\boldsymbol{\Sigma}_{ij}$ are defined from the error terms covariance matrix

$$\Omega \equiv \text{Var}(\varepsilon, \boldsymbol{\eta}) = \begin{pmatrix} 1 & \boldsymbol{\Sigma}_{21}^{\top} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}.$$

The normalization Var$(\varepsilon) = \mathbf{\Sigma}_{11} = 1$ is like that in a simple probit which is necessary for identifying the parameters of the model.

Then the structural model wherein the point probabilities for two possible outcomes for D_i are defined is given by

$$\begin{aligned} \mathbb{P}\{D_i = 1 \,|\mathbf{Y}, \mathbf{X}\} &= \Phi(w_i) \\ \mathbb{P}\{D_i = 0 \,|\mathbf{Y}, \mathbf{X}\} &= 1 - \Phi(w_i) \end{aligned} \tag{4}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, and w_i are the second equations from (1) where ε is defined from (3)

$$w_i = \frac{\mathbf{X}_1\boldsymbol{\alpha}_1 + \mathbf{Y}\boldsymbol{\beta} + \boldsymbol{\eta}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}}{\sqrt{1 - \mathbf{\Sigma}_{21}^{\top}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}}} \equiv \frac{\mathbf{X}_1\boldsymbol{\alpha}_1 + \mathbf{Y}\boldsymbol{\beta} + (\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha}_2)\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}}{\sqrt{1 - \mathbf{\Sigma}_{21}^{\top}\mathbf{\Sigma}_{22}^{-1}\mathbf{\Sigma}_{21}}}. \tag{5}$$

The same identification issue in a probit model requires normalization of (1) which is made in (5) where the denominator contains Var(e) =Var$(\varepsilon|\boldsymbol{\eta})$.

With the account that the joint density of our endogenous variables conditional on $\mathbf{X}$ is given by

$$f(D_i, \mathbf{Y}|\mathbf{X}) \equiv f(D_i|\mathbf{Y}, \mathbf{X})\, f(\mathbf{Y}|\mathbf{X}) \equiv \Phi(w_i) f(\mathbf{Y}|\mathbf{X}) \tag{6}$$

individual contributions to the likelihood are defined as

$$L_i = \Phi^{D_i}(w_i)\,\{1 - \Phi(w_i)\}^{1-D_i}\, f(\mathbf{Y}|\mathbf{X}) \equiv \begin{cases} \Phi(w_i) f(\mathbf{Y}|\mathbf{X}) & \text{if } D_i = 1 \\ (1 - \Phi(w_i)) f(\mathbf{Y}|\mathbf{X}) & \text{if } D_i = 0 \end{cases} \tag{7}$$

where the conditional density function is given by

$$f(\mathbf{Y}|\mathbf{X}) \equiv f(Y_i|Y_j, \mathbf{X})\, f(Y_j|\mathbf{X}) = \frac{1}{2\pi\sqrt{|\mathbf{\Sigma}_{22}|}} \exp\{-\frac{1}{2}(\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha}_2)\mathbf{\Sigma}_{22}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\alpha}_2)^{\top}\}. \tag{8}$$

Then for the log likelihood, the individual contributions are

$$\ln L_i = D_i \ln \Phi(w_i) + (1 - D_i) \ln\{1 - \Phi(w_i)\} + \ln f(\mathbf{Y}|\mathbf{X})$$

and log likelihood function is defined as

$$\ln L = \sum_{i:D_i=1} \ln \Phi(w_i) + \sum_{i:D_i=0} \ln\{1 - \Phi(w_i)\} + \sum_{i=1}^{N} \ln f(\mathbf{Y}|\mathbf{X}). \tag{9}$$

RESULTS

Table 5: Results for the interest regressors from simple probit regressions

	a respondent drinks				alcohol-related problems			
	on the streets	on workplace	without eating	before eating	at workplace	at home	in health	the other
log total labor earnings	−0.1740**	−0.0368	−0.1209**	−0.1497***	−0.1774*	−0.2583***	−0.0817	−1.1570
	(0.0541)	(0.0501)	(0.0385)	(0.0410)	(0.0740)	(0.0500)	(0.0504)	(70.2409)
	[−0.0192]	[−0.0067]	[−0.0378]	[−0.0384]	[−0.0078]	[−0.0284]	[−0.0095]	[−0.0224]
log total labor earnings∗$P2$	0.0100	0.0135	−0.0136	0.0355	0.0191	0.0232	−0.0325	0.8624
	(0.0269)	(0.0287)	(0.0210)	(0.0217)	(0.0400)	(0.0263)	(0.0284)	(70.2407)
	[0.0011]	[0.0025]	[−0.0042]	[0.0091]	[0.0008]	[0.0025]	[−0.0038]	[0.0167]
log total labor earnings∗$P3$	0.0184	0.0304	0.0090	0.0256	0.0144	0.0074	−0.0065	0.9067
	(0.0224)	(0.0218)	(0.0164)	(0.0179)	(0.0330)	(0.0219)	(0.0210)	(70.2407)
	[0.0020]	[0.0055]	[0.0028]	[0.0066]	[0.0006]	[0.0008]	[−0.0008]	[0.0176]
log total labor earnings∗$P4$	−0.0036	−0.0048	−0.0022	0.0326	0.0079	−0.0096	−0.0399	0.9033
	(0.0221)	(0.0221)	(0.0157)	(0.0171)	(0.0316)	(0.0212)	(0.0210)	(70.2407)
	[−0.0004]	[−0.0009]	[−0.0007]	[0.0084]	[0.0003]	[−0.0011]	[−0.0046]	[0.0175]
log total labor earnings∗$P5$	0.0209	0.0432*	−0.0055	0.0140	0.0102	−0.0297	−0.0508*	0.8802
	(0.0203)	(0.0194)	(0.0146)	(0.0163)	(0.0302)	(0.0207)	(0.0199)	(70.2407)
	[0.0023]	[0.0079]	[−0.0017]	[0.0036]	[0.0005]	[−0.0033]	[−0.0059]	[0.0170]
log total labor earnings∗$P6$	−0.0212	0.0357	−0.0093	0.0101	−0.0019	−0.0192	−0.0219	0.8539
	(0.0218)	(0.0196)	(0.0148)	(0.0164)	(0.0312)	(0.0205)	(0.0190)	(70.2407)
	[−0.0023]	[0.0065]	[−0.0029]	[0.0026]	[−0.0001]	[−0.0021]	[−0.0025]	[0.0165]
log total labor earnings∗$P7$	−0.0105	0.0106	−0.0033	0.0242	−0.0103	−0.0104	−0.0373*	0.8832
	(0.0208)	(0.0193)	(0.0141)	(0.0158)	(0.0306)	(0.0195)	(0.0185)	(70.2407)
	[−0.0012]	[0.0019]	[−0.0010]	[0.0062]	[−0.0005]	[−0.0011]	[−0.0043]	[0.0171]
log total labor earnings∗$P8$	−0.0047	0.0091	−0.0117	0.0139	−0.0116	−0.0176	−0.0170	0.8930
	(0.0219)	(0.0198)	(0.0146)	(0.0164)	(0.0317)	(0.0202)	(0.0185)	(70.2407)
	[−0.0005]	[0.0017]	[−0.0036]	[0.0036]	[−0.0005]	[−0.0019]	[−0.0020]	[0.0173]
log total labor earnings∗$P9$	−0.0227	0.0254	−0.0208	−0.0028	0.0113	−0.0155	−0.0033	0.9041
	(0.0264)	(0.0205)	(0.0157)	(0.0181)	(0.0325)	(0.0216)	(0.0193)	(70.2407)
	[−0.0025]	[0.0046]	[−0.0065]	[−0.0007]	[0.0005]	[−0.0017]	[−0.0004]	[0.0175]
Observations	2678	2675	2671	2673	3330	3324	3315	3304
Pseudo R-squared	0.1942	0.0207	0.0649	0.1214	0.1750	0.1813	0.1238	0.1669

Standard errors are in parenthesis; marginal effects are in square brackets; $^{*}p < 0.05,\ ^{**}p < 0.01,\ ^{***}p < 0.001$

Table 6: Results for log total labor earnings and professional level from simple probit and simultaneous probit regressions

	a respondent drinks							
	on the streets		on workplace		without eating		before eating	
	probit	ivprobit	probit	ivprobit	probit	ivprobit	probit	ivprobit
log total labor earnings	-0.1768^{***}	0.4469^{*}	-0.0204	-0.0789	-0.1312^{***}	-0.3113	-0.1219^{**}	-0.0426
	(0.0514)	(0.1826)	(0.0466)	(0.2287)	(0.0368)	(0.2107)	(0.0388)	(0.2212)
	[−0.0196]	[0.0896]	[−0.0038]	[−0.0189]	[−0.0411]	[−0.0978]	[−0.0315]	[−0.0116]
professional level	-0.0494^{*}	-0.4291^{***}	-0.0081	0.3309^{**}	-0.0140	0.0749	-0.0250	-0.2021
	(0.0237)	(0.0826)	(0.0192)	(0.1257)	(0.0156)	(0.1606)	(0.0168)	(0.1394)
	[−0.0055]	[−0.0860]	[−0.0015]	[0.0795]	[−0.0044]	[0.0235]	[−0.0065]	[−0.0553]
Observations	2576	2005	2574	2004	2571	2000	2573	2002
Pseudo R2	0.1854		0.0123		0.0607		0.1141	
Wald exogeneity test		32.0598		10.0823		0.9284		1.7345
(p-value)		0.0000		0.0015		0.3353		0.1878
Tests on instruments from								
GMM linear probability model								
1-st stage F-test (for log total labor earnings)		40.8110		40.8137		41.4983		40.8344
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2267		0.2267		0.2273		0.2268
Partial R squared		0.0563		0.0563		0.0571		0.0564
Shea's partial R squared		0.0418		0.0419		0.0427		0.0417
1-st stage F-test (for professional level)		9.5157		9.5171		9.4026		9.4179
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2596		0.2595		0.2583		0.2590
Partial R squared		0.0187		0.0187		0.0185		0.0185
Shea's partial R squared		0.0139		0.0139		0.0138		0.0137
2-nd stage J test		0.0000027		3.6013		5.5261		0.1881
(p-value)		0.9987		0.0577		0.0187		0.6645

Standard errors are in parenthesis; marginal effects are in square brackets; $^{*}p < 0.05,\ ^{**}p < 0.01,\ ^{***}p < 0.001$

Table 7: Continued

	alcohol-related problems							
	at workplace		at home		in health		the other	
	probit	ivprobit	probit	ivprobit	probit	ivprobit	probit	ivprobit
log total labor earnings	-0.2008^{**}	-0.0350	-0.2773^{***}	0.2266	-0.0816	0.2036	-0.2395^{**}	-0.2970
	(0.0704)	(0.3820)	(0.0485)	(0.2183)	(0.0492)	(0.2246)	(0.0859)	(0.5133)
	[−0.0088]	[−0.0040]	[−0.0302]	[0.0413]	[−0.0095]	[0.0382]	[−0.0049]	[−0.0328]
professional level	-0.0159	-0.3852^{***}	-0.0309	-0.3920^{***}	0.0150	-0.3880^{***}	0.0093	-0.3956^{**}
	(0.0323)	(0.1124)	(0.0214)	(0.0797)	(0.0206)	(0.0770)	(0.0447)	(0.1405)
	[−0.0007]	[−0.0440]	[−0.0034]	[−0.0714]	[0.0017]	[−0.0728]	[0.0002]	[−0.0437]
Observations	3210	2443	3205	2440	3198	2434	3186	2421
Pseudo R2	0.1823		0.1829		0.1134		0.1476	
Wald exogeneity test		11.6574		24.4652		26.9731		23.1085
(p-value)		0.0006		0.0000		0.0000		0.0000
Tests on instruments from								
GMM linear probability model								
1-st stage F-test (for log total labor earnings)		47.7660		47.8152		47.0951		48.6940
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2299		0.2284		0.2298		0.2270
Partial R squared		0.0532		0.0533		0.0525		0.0543
Shea's partial R squared		0.0449		0.0449		0.0449		0.0454
1-st stage F-test (for professional level)		12.9999		13.0320		12.9032		13.1057
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2595		0.2600		0.2618		0.2599
Partial R squared		0.0198		0.0199		0.0199		0.0201
Shea's partial R squared		0.0167		0.0168		0.0170		0.0168
2-nd stage J test		0.2298		0.0172		0.1541		0.0646
(p-value)		0.6317		0.8956		0.6946		0.7994

Standard errors are in parenthesis; marginal effects are in square brackets; $^{*}p < 0.05,\ ^{**}p < 0.01,\ ^{***}p < 0.001$

Table 8: Results for log total labor earnings and the interaction term from simple probit and simultaneous probit regressions

	a respondent drinks							
	on the streets		on workplace		without eating		before eating	
	probit	ivprobit	probit	ivprobit	probit	ivprobit	probit	ivprobit
log total labor earnings	-0.1705^{***}	0.3935^{*}	0.0035	-0.0924	-0.1279^{***}	-0.2829	-0.1377^{***}	-0.0458
	(0.0502)	(0.1688)	(0.0460)	(0.2175)	(0.0360)	(0.2158)	(0.0378)	(0.2221)
	[−0.0189]	[0.0890]	[0.0007]	[−0.0230]	[−0.0401]	[−0.0889]	[−0.0355]	[−0.0127]
log total labor earnings*higher median prof (lIS)	-0.0138	-0.1934^{***}	-0.0151	0.1471^{**}	-0.0053	0.0196	-0.0009	-0.1079
	(0.0103)	(0.0297)	(0.0078)	(0.0561)	(0.0064)	(0.0840)	(0.0069)	(0.0643)
	[−0.0015]	[−0.0437]	[−0.0028]	[0.0367]	[−0.0017]	[0.0062]	[−0.0002]	[−0.0299]
Observations	2678	2080	2675	2078	2671	2074	2673	2076
Pseudo R2	0.1880		0.0129		0.0633		0.1176	
Wald exogeneity test		47.4951		12.4909		0.5987		3.1090
(p-value)		0.0000		0.0004		0.4391		0.0779
Tests on instruments from								
GMM linear probability model								
1-st stage F-test (for log total labor earnings)		43.7347		44.0763		44.5599		43.8834
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2307		0.2313		0.2314		0.2309
Partial R squared		0.0572		0.0577		0.0582		0.0575
Shea's partial R squared		0.0396		0.0396		0.0398		0.0392
1-st stage F-test (for lIS)		7.5673		7.4928		7.6967		7.5550
(p-value)		0.000		0.0001		0.0000		0.0001
Adjusted R squared		0.1817		0.1810		0.1802		0.1809
Partial R squared		0.0123		0.0122		0.0125		0.0123
Shea's partial R squared		0.0085		0.0084		0.0086		0.0084
2-nd stage J test		0.2826		4.2930		3.7913		1.3142
(p-value)		0.5950		0.0383		0.0515		0.2516

Standard errors are in parenthesis; marginal effects are in square brackets; $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 9: Continued

	alcohol-related problems							
	at workplace		at home		in health		the other	
	probit	ivprobit	probit	ivprobit	probit	ivprobit	probit	ivprobit
log total labor earnings	−0.1688*	−0.0796	−0.2764***	0.1991	−0.1092*	0.2184	−0.2313**	−0.1830
	(0.0677)	(0.3761)	(0.0465)	(0.2007)	(0.0471)	(0.2026)	(0.0848)	(0.4721)
	[−0.0075]	[−0.0115]	[−0.0305]	[0.0428]	[−0.0128]	[0.0476]	[−0.0046]	[−0.0276]
log total labor earnings*higher median prof (lIS)	−0.0121	−0.1775***	−0.0010	−0.1804***	0.0061	−0.1789***	0.0122	−0.1763***
	(0.0135)	(0.0474)	(0.0091)	(0.0287)	(0.0087)	(0.0292)	(0.0191)	(0.0518)
	[−0.0005]	[−0.0256]	[−0.0001]	[−0.0388]	[0.0007]	[−0.0390]	[0.0002]	[−0.0266]
Observations	3330	2533	3324	2529	3315	2521	3304	2509
Pseudo R2	0.1724		0.1768		0.1141		0.1454	
Wald exogeneity test		16.5697		42.6704		42.8582		42.1210
(p-value)		0.0000		0.0000		0.0000		0.0000
Tests on instruments from								
GMM linear probability model								
1-st stage F-test (for log total labor earnings)		51.4848		51.3977		50.2515		52.2444
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.2326		0.2323		0.2324		0.2309
Partial R squared		0.0542		0.0542		0.0531		0.0552
Shea's partial R squared		0.0397		0.0396		0.0395		0.0405
1-st stage F-test (for lIS)		10.7079		10.7475		10.4573		10.8972
(p-value)		0.0000		0.0000		0.0000		0.0000
Adjusted R squared		0.1837		0.1847		0.1864		0.1852
Partial R squared		0.0137		0.0138		0.0135		0.0139
Shea's partial R squared		0.0100		0.0101		0.0101		0.0102
2-nd stage J test		0.5951		0.0783		0.4992		0.1495
(p-value)		0.4404		0.7796		0.4799		0.6990

Standard errors are in parenthesis; marginal effects are in square brackets; $^{*}p < 0.05,\ ^{**}p < 0.01,\ ^{***}p < 0.001$

SIMPLE PROBIT EQUATIONS

Table 5 contains the results of the ML estimation from the simple probit models. Amongst the eight equations for various drinking dependent variables five ones, 1st, 3d, 4th, 5th, and 6th, display significant negative associations between log total earnings and a drinking pattern. Thus, for the most drinking pattern variables a negative correlation is shown up. As for the interaction variables, a weak insignificant tendency is displayed for the higher professional level variables to have negative correlation with a drinking pattern comparing with those for lower professional levels. However, they fail to display a monotonous inverse trend across all the professional levels.

Dummy for alcohol intake on workplace tends to display a positive associations with the interest measures across the professional levels among which the fifth and the sixth display p-values under 0.1. Pseudo R-squared in this equation is the smallest among the equations which implies that drinking on workplace depends on earnings across all the professional levels and the specified range of the controls much less than the other drinking pattern variables.

Sample size is consistently more in case of the alcohol effects variables than those for drinking circumstances. And excluding the first equation the same holds for pseudo R-squared.

The control covariates (the results for them are dropped here) display, as a whole, the expected signs and significance. Except for the second equations, male gender control shows significant positive associations with drinking pattern variables. Age in most cases shows a positive associations with the drinking measures, and schooling is significantly inversely associated with, at least, half drinking pattern measures. As for the other controls they display a tendency for an inverse associations with a drinking pattern variables.

The results in Tables 6 and 7 from the probit regressions as to earnings variables are similar to those in Table 5. Except for the second equation, and in Table 6 the seventh too, all the equations display significant negative associations between log total labor earnings and a drinking pattern. At the same time, the other interest variable — professional level in Table 6 and the interaction term for log total labor earnings of higher median group by professional level in Table 6 — except for the first equation in Table 6 do not display any significant link with drinking variables. Yet it is worth noting that signs of the insignificant coefficients in most cases are negative.

SIMULTANEOUS PROBIT EQUATIONS

Tables 6 and 7 present results from the simultaneous models. As for log total labor earnings in Table 6 in all cases where it is admissible to have simultaneous probit estimates as consistent ones they are either positive (1st) or insignificant (2d, 5-8th), while a regular probit tends to give more significant and more intensively inversed associations between earnings and drinking pattern. Except for the second equation, in all cases in Table 6 where we can safely have simultaneous probit estimates as consistent ones they display significant and more intense inverse links between professional level and drinking patterns comparing with a simple probit. Similar results as to log total labor earnings and the interaction term are displayed in Table 7. Here all the equations show insignificant results for the earnings variable. The interaction term behaves generally the same way as the professional level variable does in Table 6.

Tables 6 and 7 contain also results of Wald tests of exogeneity. In two cases, in 3d and 4th equations in Table 6 and only in 3d equation in Table 7, the test strongly rejects the null of endogeneity so that here there is an unjustified loss of efficiency of the estimates.

Tables 6 and 7 also contain comprehensive results of the tests from the GMM linear probability model. In all cases p-values of F statistic are under 0.001. The values of F statistics and thereby partial

R-squareds for the earnings variable are as a whole 3-4 times exceed those for the skills variable. At the same time, the adjusted R-squareds are higher for the latter. The Shea's partial R-squareds are expectedly less than the standard partial one, and the difference between them tends to be the same for both earnings and skills variables.

As for the validity of the excluded instruments in both Table 6 and Table 7 Hansen J tests display p-values under 0.1 for the second equation and the third one. As for the latter it is worth noting that only for this equation we simultaneously fail to reject the null of endogeneity according to the Wald test from the instrumental probit model and do reject the null of orthogonality of the instruments to the error term according to the Hansen test from the linear model.

DISCUSSION OF THE RESULTS AND CONCLUDING REMARKS

Regarding the relevance issue if we start from the direct meaning of p-values of F tests their zero values are indicative of the relevance of the instruments in all cases. Also the threshold value of 10 implied by the rule of thumb (Staiger and Stock, 1997), though it is to be applied to one endogenous regressor models also generally support relevance of the instruments. Lastly, not much difference between the standard and the Shea's partial R-squareds allows us to infer the same.

The tests of overidentifying restrictions only in two cases reject the null of validity of the instruments so that there the simultaneous probit is to give inconsistent estimates. In the rest six equations the tests imply our instruments are valid.

Tests on the instruments as well as general considerations on their relevance and validity allows us to conclude that estimates obtained via the instrumental probit regressions are consistent and unbiased for which sake we have sacrificed the efficiency available in a simple probit. These estimates indicate lack of significant link between log total labor earnings and any drinking variables. At the same time, the skills variables have significantly inversely affect probabilities of drinking on the streets and of all the alcohol-related problems. One drinking variable, probability of drinking at workplace, is positively affected by the skills variables.

Insignificant effect of earnings presumably reflects its complex nature. On the one hand, earnings induce demand for alcohol given the latter is a normal good; on the other hand, higher earnings are consistent with more opportunity cost of drinking. Mixing these contrary effects may result in such insignificant estimates. That consistent estimate obtained in the instrumental probit is insignificant, whereas inconsistent estimates from the simple probit are significantly inverse can be explained by the reverse effect of alcohol abuse on earnings. As far as this effect is inverse one the inconsistent estimates containing the two bilateral effects are to be inverse ones.

Insignificant effect of the earnings takes place when we control for professional level so that this effect is for earnings unrelated to professional level, e.g. for people working not by their specialty, common labor, rent-takers etc.

Significant inverse effect of the skills variables is displayed in the presence of a number of controls so that this effect is not explained by a correlation of professional level with education, marital, health, or respect statuses etc. Using these personal controls and the result that the interaction term is displayed to significantly affect the alcohol variables mean that whereas unskilled workers do not respond to their earnings change as to their involvement in alcohol abuse skilled workers do respond, namely, their probability of alcohol abuse drops when their earnings rise. Such an inverse effect of the skilled variables can be explained by the opportunity cost of leisure. Skilled workers may respond to the earnings change other way than unskilled ones because of difference in their labor earnings per hour.

Positive effect of the skills variables on dummy for drinking at workplace can be explained by the widespread in Russia corporate culture assuming mini-parties on workplace at the end of or after working day. Interestingly, this dummy is the only drinking variable for which well-to-do women display more participation rate than men in the same earnings strata (Table 1). At the same time, it is for this variable that educated people display even more participation than less educated ones (Table 3). Thus, one can suggest that this drinking variable reflects involvement of an individual in social interactions, including those which pay, rather than some alcohol abuse.

Insignificance of the interaction variables for the earnings of all the professional levels (Table 5) actually means that associations between earnings and drinking patterns are the same across the professional levels. That the effect of earnings on the drinking variables do not depend on the professional mastership according to the inconsistent estimates and that this effect do depend on the skills variables according to consistent estimates can be explained by measurement error, namely, since professional level is a self-rated measure it is subject to self-estimation-related bias.

The same measurement error toward to underestimation is likely to take place in case of earnings, but in the latter case this measurement bias is overlapped by the reverse causality bias so that the earnings effect ultimately is overestimated in the simple probit.

As for bias related to the reverse causality, though it is to be displayed in estimates of effects of both earnings and professional level, in the former case this bias is to be much more than that in the latter case, since a time span for the revealing the reverse effect is much shorter. A hard-drinking person may immediately face their earnings fall, but it is not the case as to their professional level. There are people who decided to be drinkers long ago and their decision affected both their earnings and professional level. But there are also people who decided to be drinkers recently after obtaining some professional skills, and in this case their decision is to impact much their earnings, but no so their professionalism. In other words, professional level is a result of life-time experience which cannot be eliminated by current health-impairing behavior, while earnings is a result of both the life-time experience and current behavior and the latter being it wrong may deprive an individual of all the opportunities to realize their professional skills.

Our results indicate persistent inverse casual pathways of skilled workers' earnings on the measures of their drinking patterns. As a whole, such an effect means that higher earnings as far as they are subject to professionalism make people to conduct more sober lifestyle. Thus, the effects of the interest measures within the simultaneous probit fit the hypothesis tested. Extensive list of controls, which may correlate to a drinking pattern measure as well as to the instruments, makes our result more credible. In particular, it ensures that it is through our endogenous measure that professional skills affect drinking pattern measures. As the similar results are displayed by several models with distinct dependent drinking measures, our results are robust to various specifications.

The results have some implications with respect to theory and practice. Specifically, the fact that unskilled workers fail to adapt their drinking behavior to higher earnings, while skilled ones under the same effect do adapt their behavior towards more modest drinking supports the hypothesis tested in this paper and is in line with the approach based on the opportunity cost of leisure. From this viewpoint, this fact can be explained by the consideration that in the first case, higher earnings are related to more or harder work, while not being related to higher hour wage, and skilled workers do face higher hour wage. So, the same rise in earnings may involve higher opportunity cost of leisure for skilled workers, while not involving the same for unskilled ones.

This sheds light on possible causes of the explosive rise of alcohol consumption in Russia during the transitional period. Russian reforms have given rise to a wide scope of social changes. More than other transitional economies, Russia experienced income inequality increase, but in a lesser extent than the others due to human capital returns (Didenko 2012). In fact, the reforms frequently have entailed income redistribution away from skilled workers in a range of industries who were demanded in the Soviet economy to those who have managed to adapt to the new economic conditions. The total structural transformation has resulted in loss of jobs by skilled workers and made them work not by their specialty. So, their knowledge and work experience have become useless as regard to their job responsibilities and hour wages.

There are a number of pathways through which it may, in turn, induce change in the drinking behavior. In terms of the mentioned framework, it has decreased their opportunity cost of leisure and lowered the barrier to drinking. A policy implication hereof is that human potential and its realization may be a major force of, not only. economic growth but also, temperance of people with respect to alcohol and other drugs.

REFERENCES

[1] Amemiya T. (1978) The estimation of a simultaneous equation generalized probit model. *Econometrica* 46(5): 1193-1205.

[2] Baltagi B. H., Geishecker I. (2006) Rational alcohol addiction: evidence from the Russian longitudinal monitoring survey. *Health Economics* 15: 893-914.

[3] Barrett G. F. (2002) The effect of alcohol consumption on earnings. *The Economic Record* 78: 79-96.

[4] Baum C. F., Schaffer M. E., Stillman S. (2003) Instrumental variables and GMM: estimation and testing. *Stata Journal* 3: 1-31.

[5] Becker G. S. (1965) A theory of the allocation of time. *Economic Journal* 75: 493-517.

[6] Becker G. S., Murphy K. M. (1988) A theory of rational addiction. *Journal of Political Economy* 96(4): 675-700.

[7] Bray J. W. (2005) Alcohol use, human capital, and wages. *Journal of Labor Economics* 23: 279-312.

[8] Cawley J., Ruhm C. J. (2012) The economics of risky health behaviors. *Handbook of health economics* 2: 95-199.

[9] Combes P. P, Mayer T., Thisse J. F. (2008) *Economic geography. The integration of regions and nations*. Princeton: Princeton University Press.

[10] Cook P. J., Peters B. (2005) The myth of the drinker's bonus. *NBER WP* 11902. December.

[11] Danielsson A. K, Wennberg P., Hibell B., Romelsjo A. (2011) Alcohol use, heavy episodic drinking and subsequent problems among adolescents in 23 European countries: does the prevention paradox apply? *Addiction* 107: 71-80.

[12] Cowell A. J. (2006) The relationship between education and health behavior: some empirical evidence. *Health Economics* 15: 125-146.

[13] Dee T. S. (2001) Alcohol abuse and economic conditions: evidence from repeated cross-sections of individual-level data. *Health Economics* 10: 257-270.

[14] Denisova I. (2010) *Consumption of alcohol in Russia: an effect on health and mortality*. Moscow: CEFIR and New School of Economics (in Russian).

[15] Didenko D. V. (2012) Income inequality and systemic transformations: long-term trends of human capital private returns. *The Journal of Comparative Economic Studies* 7: 53-87.

[16] Dujardin C., Goffette-Nagot F. (2009) Does public housing occupancy increase unemployment? *Journal of Economic Geography* 9: 823-851.

[17] French M. T., Maclean J. C., Sindelar J. L., Fang H. (2011) The morning after: alcohol misuse and employment problems. *Applied Economics* 43: 2705-2720.

[18] Fuchs V. R. (1982) Time Preference and Health: an Exploratory Study *NBER WP* 539.

[19] Gibb S. J., Fergusson D. M., Horwood L. J. (2011) Working hours and alcohol problems in early adulthood. *Addiction* 107: 81-88.

[20] Griffin B. A., Ramchand R, Edelen M. O., McCaffrey D. F., Morral A. R. (2011) Associations between abstinence in adolescence and economic and educational outcomes seven years later among high-risk youth. *Drug and Alcohol Dependence* 133: 118-124.

[21] Heckman J. J. (1978) Dummy endogenous variables in a simultaneous equation system. *Econometrica* 46(4): 931-959.

[22] Isabel G., Molina A. (2007) Human development and alcohol abuse in adolescence. *Applied Economics* 39: 1315-1323.

[23] Keng S.-H., Huffman W. E. (2007) Binge drinking and labor market success: longitudinal study on young people. *Journal of Population Economics* 20: 35-54.

[24] Kim, V., Roshin, S. Y. (2009) An effect of alcohol consumption on wage. *HSE WP* 15/2009/01 (in Russian).

[25] Kusmitch', O. S., Roshin, S. Y. (2007) An effect of health on wage and employment: empirical estimates of the return to health. *HSE WP*15/2007/02 (in Russian).

[26] Leon D., Shkolnikov V. M., McKee M. (2009) Alcohol and Russian mortality: a continuing crisis. *Addiction* 104: 1630-1636.

[27] Lye J., Hirschberg J. (2010) Alcohol consumption and human capital: a retrospective study of the literature. *Journal of Economic Surveys* 24: 309-338.

[28] MacDonlad Z., Shield M. A. (2001) The impact of alcohol consumption on occupational attainment in England. *Economica* 68: 427-453.

[29] Mincer J. (1974) *Schooling, experience and earnings*. NY: Columbia University Press.

[30] Mroz T. A. (1999) Discrete factor approximations in simultaneous equation models: Estimating the impact of a dummy endogenous variable on a continuous outcome. *Journal of Econometrics* 92: 233-274.

[31] Nemtsov A. V. (2000) Estimates of total alcohol consumption in Russia, 1980-1994. *Drug and Alcohol Dependence* 58: 133-142.

[32] Norstrom T. (2011) The role of alcohol in the Russian mortality crisis. *Addiction* 106: 1957-1965.

[33] Peters B. L. (2009) The drinkers' bonus in the military: officers versus enlisted personnel. *Applied Economics* 41: 2211-2220.

[34] Pomerleau J., McKee M., Rose R., Haerpfer C. W., Rotman D., Tumanov S. (2005) Drinking in the Commonwealth of Independent States — evidence from eight countries. *Addiction* 100: 1647-1668.

[35] *Prices in Russia 2010. Collected statistics.* (2010) Moscow: Rosstat.

[36] Schiff M., Rahav G., Teichman M. (2005) Israel 2000: immigration and gender differences in alcohol consumption. *American Journal of Addiction* 14: 234-247.

[37] Schultz T. W. (1968) Institutions and the rising economic value of man. *American Journal of Agricultural Economics* 50: 1113-1122.

[38] Skorobogatov A. S. (2012) The value of human capital and health behavior. *Economics Bulletin* 32(2): 1785-1796.

[39] Sloan F. A., Grossman D. S. (2011) Alcohol consumption in early adulthood and schooling completed and labor market outcomes at midlife by race and gender. *American Journal of Public Health* 101: 2093-2100.

[40] Srivastava P. (2010) Does bingeing affect earnings? *Economic Record* 86: 578-595.

[41] Staiger D., Stock J. H. (1997) Instrumental variables regression with weak instruments. *Econo-*

metrica 65(3): 557-586.
[42] Stockwell T., Auld M. C, Zhao J., Martin G. (2011) Does minimum pricing reduce alcohol consumption? The experience of a Canadian province. *Addiction* 107: 912-920.
[43] Wooldridge J. M. (2002) *Econometric analysis of cross section and panel data*. The MIT Press: Cambridge Massachusetts.