Aalto University
School of Economics

The Relationship between Alcohol Use and Earnings

Economics

Master's thesis

Henri Syvänen

2011

Department of Economics
Aalto University
School of Economics

Acknowledgements

First, I would like to thank my supervisor, Professor Matti Liski, for his support and assistance throughout the years that I have spent conducting both my Bachelor's thesis and my Master's thesis on this same subject. Second, I would like to thank researcher Pia Mäkelä for her great help in acquiring the data and in helping me to determine the scope of my thesis.

Third, I would like to thank everyone I have discussed with about my thesis. I have been given valuable insights about the possible effects that drinking has on earnings. These insights have helped me to form the framework and crystallize all the forces that shape the relationship between alcohol use and earnings.

ABSTRACT

AALTO UNIVERSITY SCHOOL OF ECONOMICS November 1, 2011
Department of Economics
Master's Thesis
Henri Syvänen

THE RELATIONSHIP BETWEEN ALCOHOL USE AND EARNINGS

The link between individual earnings and alcohol use has never been investigated in Finland. The aim of this thesis is to add a missing piece into the existing alcohol use literature by providing the first study on this matter. The study is based on an empirical analysis that utilizes data from the Finnish Drinking Habits Survey 2008. Individual earnings information has been connected to the survey data from the national taxation register. The empirical analysis is conducted using multiple separate OLS-models to account for different aspects of the alcohol-earnings relationship.

According to the results of the analysis, alcohol use is strongly correlated with earnings and several interesting findings were made. First, drinkers earn more than abstainers. Second, the relationship between alcohol use and earnings follows an inverse u-shaped path, with peak earnings taking place at 2.6 drinks per day for men and at 1.2 drinks per day for women. In addition, the number of intoxication is negatively correlated with earnings for men, but the same correlation does not exist for women.

The results revealed that alcohol-earnings relationship also differs between socioeconomic positions and different drinker types. Wine drinkers earn 5 percent more than beer drinkers and 15 percent more of generic drinkers and some indication about the beneficial effects of drinking in a social manner was also found. Overall, the results indicate that those who drink moderately earn the most.

TIIVISTELMÄ

AALTO-YLIOPISTON KAUPPAKORKEAKOULU 1.11. 2011
Taloustieteen laitos
Pro gradu – tutkielma
Henri Syvänen

ALKOHOLIN KÄYTÖN JA TULOJEN VÄLINEN SUHDE

Alkoholin käytön ja tulojen välistä yhteyttä ei ole koskaan tutkittu Suomessa. Tämän tutkielman tavoitteena on laajentaa suomalaista tietämystä alkoholin vaikutuksista tutkimalla ensimmäistä kertaa tulojen ja alkoholin käytön välistä suhdetta. Tutkielma pohjautuu empiiriseen analyysiin, jossa käytetään vuoden 2008 Juomatapatutkimusta. Henkilökohtaiset tulotiedot on yhdistetty tutkimusaineistoon verorekisterin kautta. Empiirisessä analyysissä käytetään monia eri OLS-malleja, jotta alkoholin ja tulojen välisen suhteen eri näkökulmat voidaan ottaa huomioon.

Tutkielman tulosteen perusteella alkoholin käyttö on vahvasti korreloitunut tulojen kanssa ja tuloksista nousee esille useita kiinnostavia havaintoja. Ensinnäkin on selvää, että absolutistit ansaitsevat keskimäärin vähemmän kuin alkoholin käyttäjät. Toiseksi alkoholin käytön ja tulojen välinen suhde noudattaa käänteistä u-käyrää, jonka perusteella miesten tulot maksimoituvat 2,6 päivittäisen alkoholiannoksen kohdalla ja naisten tulot 1,2 päivittäisen annoksen kohdalla. Lisäksi humaltumiskertojen määrä on korreloitunut käänteisesti tulojen kanssa miesten osalta, mutta samaa korrelaatiota ei löydy naisten juomisessa.

Alkoholin käytön ja tulojen suhde myös vaihtelee eri sosioekonomisten asemien välillä. Viininjuojilla puolestaan on 5 prosenttia suuremmat tulot kuin oluenjuojilla ja 15 prosenttia suuremmat tulot kuin alkoholin yleiskäyttäjillä. Sosiaalisen juomisen positiivisesta tulovaikutuksesta löytyi myös todisteita. Kaiken kaikkiaan tutkielman tulokset osoittavat, että kohtuullisesti alkoholia juovilla on suurimmat tulot.

Table of contents

Figures and tables

1. Introduction

1.1. Background

Alcohol is the most used intoxicant in Finland and majority of the adult population consume alcohol. The average yearly amount in pure alcohol is over 10 liters per person (National institute for Health and Welfare, 2008). For most people, drinking is habitual and strongly related to social interaction as well as celebration, but at the same time alcohol is one of the leading causes of health problems and social problems. Alcohol use causes accidents, diseases and loss of productivity, which all result in social costs. Overall, the gains from alcohol taxes barely exceed the direct costs of alcohol-related harms (Mellin et al. 2003). Thus, it is clear that the government should have policies that aim to minimize the alcohol-related harms.

In addition to the governmental controls of alcohol, the labor market actors have multiple reasons for limiting alcohol use. Consumption of alcohol is normally restricted in workplaces and some working environments have a zero tolerance for alcohol. Heavy use and binge drinking cause absenteeism and direct losses of productivity. Furthermore, alcoholism can prohibit the participation in labor force or cause a permanent exit from it. But on the other hand, the health literature shows that moderate drinking may in fact be beneficial for physical and psychological well-being, which can lead to improved productivity. For example, moderate alcohol use has been proven to reduce risk of heart diseases (Hvidtfeldt et al. 2010). Thus, the productivity effects of drinking are not straightforward and they undoubtedly depend on drinking pattern.

Furthermore, the social side of drinking may have labor market consequences both in terms of productivity and job opportunities. Drinking might extend the contact network of and individual, making it easier to find employment. Moreover, socializing might also strengthen the bonds between co-workers and thus have a positive effect on productivity. Evidently, companies take the social side of drinking into consideration by serving alcohol to mark successful events and to promote the cohesion of staff in celebrations, while alcohol is prohibited in the normal operative functions.

Considering the effects that alcohol use has on productivity and individual employment opportunities, an interesting question arises: do the labor market outcomes differ according to the drinking pattern? For example, is there an earnings premium for those who drink moderately, or do the negative effects of drinking overcome possible positive consequences? In order to study the outcomes, the relationship between individual earnings and alcohol use needs to be investigated.

The relationship has received interest in the academic world and a number of studies have been published. The first study was made by Berger and Leigh in 1988. They used 1972 – 1973 Quality of Employment Survey and found a significant positive correlation between drinking and earnings. The male wage premium was highest for those who drink twice per day, while female wage premium peaked at category drink once or twice per week. Mullahy and Sindelar (1993) took a different approach by estimating the effects of alcoholism on labor market success. They found that alcoholism affects income mostly by restricting labor market participation. Kenkel and Ribar (1994) focused on the same issue, but their results rule out the negative effects on labor supply. Instead, they find a negative relationship between alcohol dependence and earnings.

French and Zarkin (1995) continued in this field by studying four different worksites based on a questionnaire that was constructed for their study purposes. Their results suggest an inverse u-shaped relationship which peaks at around 2 drinks per day. Hamilton and Hamilton (1997) had similar findings with Canadian data on prime-aged males. Also Heien (1996) came to the same conclusion by using medical findings to specify the relationship between alcohol consumption and earnings. Some researchers have even found a constant wage premium for drinkers regardless of the drinking pattern (Zarkin et al, 1998).

In the 21st century the scope of the studies has widened. The earlier studies were conducted on North American data, but during the past ten years there have been studies also on European datasets. For example Macdonald and Shields (2001) used the Health Survey for England and after controlling for the endogeneity of drinking in wage equation, they found significant returns to moderate drinking. Tekin (2004) conducted his study on Russian panel data, concluding that there exists an inverse U-shaped relation, but including individual fixed effects leads to small linear return to alcohol consumption. Van Ours (2004) studied the issue on Dutch data and includes smoking into his model. The results show that smoking leads to decreased earnings while drinking leads to increased earnings, regardless of the degree of use.

In addition to smoking, other aspects have been included into analysis in the latest studies. For instance, Bray (2005) developed a model that isolates the effects of alcohol use on human capital accumulation from the earnings equation. His results show that moderate drinking has a positive effect on both the returns to education and experience. Keng and Huffman (2007) studied binge drinking on young people and found that heavy binge drinking significantly lowers earnings. Ziebarth and Grabka (2008) used German data to test whether there are beverage-specific earnings effects and found them on regional level. Peters (2009) focused on US military personnel and argued that social capital is an important concept in explaining the wage premium.

Overall, there have been relatively many studies on the relationship between alcohol use and earnings and majority of them support the existence of a wage premium for those who drink moderately. Some researchers tend to rely on health aspects of alcohol use, while others focus on human capital. Although different hypotheses have been tested, there are necessarily problems with unobserved factors and causal links, making the estimation of true relationship difficult. It is however clear that extensive drinking and alcohol abuse cause negative outcomes in the labor market. Furthermore, Cook and Moore (2000) argued that the positive association between drinking and earnings is due to the effect of earnings on alcohol use, meaning that alcohol is a normal good and its consumption increases with income. The ambiguity about the relationship clearly calls for further research that takes multiple aspects into account.

1.2. Objectives and results

The consequences of alcohol use have been extensively studied in Finland, yet there has not been a single study on the relationship between alcohol use and earnings in Finland[1]. Therefore, this thesis is the first study on the effects of alcohol use on earnings in Finland and it adds a missing piece to the existing Finnish alcohol-related economic literature. The objective of the thesis is to examine the theoretical background of how alcohol use is related to earnings and to provide the first empirical research of the relationship with Finnish data.

The focus of the thesis is solely on individual effects and issues such as social costs associated with alcohol use are excluded from the analysis. Looking back at the existing literature on the subject, there has been a tendency to separate individuals merely into drinkers and non-drinkers. For sure,

[1]Johansson et al. (2007) have studied the association of alcohol dependency and employment probability in Finland. However, their study doesn't include variables for earnings or wages and the scope is different.

most of the studies have found that non-drinkers earn less than drinkers, but making the analysis in this manner does not provide a comprehensive view on the subject. Rather, the goal is to study multiple dimensions of drinking in respect to earnings and treat abstainers as a minor group which constitutes only a fraction of the population. Regarding the dependent variable of the analysis, earnings are observed only when an individual has positive income. Thus, the focus is on those individuals that *de facto* earn money and labor force participation is not in the scope of the study.

For the empirical part, I use Finnish Drinking Habits Survey from 2008. The survey has been conducted by the National Institute for Health and Welfare and it addresses volumes, intensities and beverage types in alcohol consumption as well as attitudes toward drinking. Register-based variables are linked to the survey data, which gives a broad range of measurements to control individual-level effects. The scope of the survey includes Finnish citizens aged between 15 and 69 with sample size of 3750 individuals. The total number of respondents was 2725 with male-female ratio close to 50%. Sample is adjusted for the needs of each hypothesis to be tested and the adjustments are explained in the empirical part.

The main research question is: *What is the relationship between alcohol use and earnings?* This leads to more specified aspects, such as looking at different drinking patterns separately and testing multiple hypothesis alternatives. Theories of labor and health economics form the backbone of the study, yet they alone are not sufficient for a careful examination of the research question. Since the problem is multidimensional, it is investigated with multiple hypotheses. In addition, earlier studies provide a wide array of tools and methods of how the theories can be applied in a practical way. These studies will be reviewed in the theoretical part and methods from these studies will be applied in the empirical analysis.

Methodology in the empirical part first utilizes basic statistical tools to provide descriptive information about the data and sample. After descriptive statistics, the analysis proceeds to econometric analysis in which the hypotheses are tested with ordinary least squares (OLS) estimation technique. In order to unravel the questions around the so-called alcohol-earnings puzzle and to get robust estimates with OLS, a multidimensional approach is employed. This means taking into account the dimensions of age, gender, socioeconomic stand, beverages and individual sociability. In this manner, it is possible to provide a broad view of the effects of alcohol use on earnings. Combining the multidimensional approach with the information from descriptive

characteristics of the sample, it is also possible to control the likely existent reciprocity and endogeneity.

The results of the thesis point out distinctive patterns in alcohol-earnings relationship. First, drinkers do earn more than abstainers, which is a consistent finding with earlier studies. Being a former drinker is associated with even larger negative earnings effect, which probably indicates excessive drinking in the past and problems in the current labor market situation. Turning into the continuous measures of alcohol use, an inverse u-shaped relationship was found for both men and women. For men, the peak point in earnings takes place at average consumption of 2.6 drinks per day, whereas for women the peak point is at around 1.2 drinks per day. In addition, the frequency of intoxications has a negative effect on the earnings, especially for males and individuals in the lower socioeconomic groups. Individuals in the lower socioeconomic groups also maximize their earnings at lower drinking quantities, indicating that it is beneficial to drink less than the average individual in these groups.

The beverage-specific model revealed that wine drinkers earn more than other drinkers, which of course is a very intuitive result and raises questions about cause and effect. However, a closer inspection showed that the same pattern holds when the sample is split at median income. The results from the sociability model indicate that people who can control their drinking and mainly drink in social manner are better off, but more explicit measures would have been needed to fully address the issue of sociability in alcohol-earnings relationship. A striking result across the models is the fact that alcohol use patterns are more significant explaining factors of earnings than any of the health-related measures.

This thesis is organized as follows. Chapter 2 presents the theoretical framework for the effects of alcohol use on earnings and lays the foundation for the rest of the thesis. Chapter 3 in turn describes the econometric methods of the analysis, presents the data and provides descriptive statistics on the sample. The results from the empirical analysis are presented in chapter 4 with each hypothesis having its own sub-chapter. Chapter 5 presents a brief discussion on the results and chapter 6 summarizes the thesis.

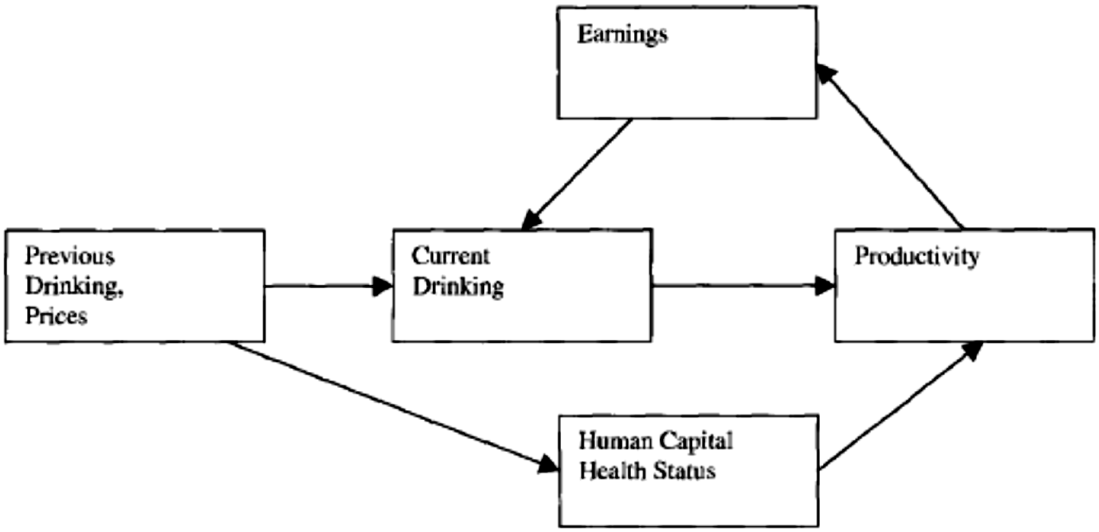
2. Theoretical framework

Governments use multiple methods to restrain alcohol consumption and beverage commerce. Taxation is the most important single factor for controlling the prices and affect demand for alcohol. Increases in alcohol taxes have been shown to directly affect public health figures (see kenkel 2005). In addition to excise taxes that affect the demand of alcohol, there are different constraints to curb the supply of alcohol. Age limitations prevent retailers from selling alcohol to minors. Also selling times and locations are regulated, as well as marketing of beverages. Government-held monopolies, such as Alko in Finland, have been established especially in the Nordic countries to organize the trade of wines and spirits.

In Finland, alcohol taxation has been a central issue in the public health policy discussions. In the past ten years, taxation has been adjusted multiple times. In 2004, quotas on tax free imports of alcohol from other EU countries were abolished and Estonia joined EU during the same year. Price difference in alcohol beverages between Finland and Estonia was significant and the Finnish government had reasonable fears about dramatic increase in alcohol consumption. Thus, alcohol taxes were reduced on average by 33% in March 2004 (Ministry of Finance). According to Mäkelä and Österberg (2009), alcohol consumption increased 10% in 2004 and it had severe public health consequences. The authors state that alcohol-related harms increased dramatically among the worst-off parts of the society. Recently, alcohol taxes have been raised in stages, the last increase being 10% in late 2009 (Ministry of Finance).

On individual level, alcohol consumption can be seen as a combination of several factors. In addition to alcohol prices, previous consumption affects current consumption level. Obviously, current drinking is determined partly by income. Cook and Moore (2000) constructed a comprehensive diagram of the various relationships, which is presented in figure 1 below.

Figure 1: Diagram on the relationships of alcohol use



Source: Cook and Moore, 2000

Looking at the diagram, current drinking status affects productivity which in turn affects earnings. As earnings have an effect on current drinking, there exists a reciprocal relationship between alcohol use and earnings. This means that an econometric model most likely suffers from endogeneity. Endogeneity may also arise from unobserved background factors that will be discussed in detail in the empirical part. Furthermore, previous drinking affects health status and the stock of human capital. This diagram forms the basis for theoretical framework in this thesis. Since the interest of the thesis is in earnings effects, the diagram could be turned into a simplified earnings function in the following way:

$$W_t = \beta_0 + \beta_1 A_t + \beta_2 A_{t-1} + \beta_3 H_t + \beta_4 K_t + \beta_5 X_t + \varepsilon_t \qquad (1)$$

In the equation, W_t is the measure of earnings, A_t and A_{t-1} are measures of alcohol use for the current and past use respectively, H_t is a health capital measure, K_t is a human capital measure, X_t is a measure all other exogenous covariates and ε_t is the error term. Productivity is not directly measured, but it is hypothesized that earnings reflect productivity. To take the reciprocal effect of earnings to current drinking into account, a simultaneous drinking equation could be formed in the following way:

$$A_t = \gamma_1 W_t + \gamma_2 X_t + \mu_t \quad (2)$$

Where A_t and W_t are defined as before and μ_t is the error term. In the next sections, the diagram is analyzed and also additional aspects are presented. For example, the relationship between sociability, drinking and earnings requires consideration for direct earnings effects that are not mediated through productivity. Yet, these are not observable in the diagram[2]. Overall, there are overlapping elements in the framework and many effects have joint outcomes. Keeping this in mind, the following sections are divided in a manner that maximizes clarity and understandability.

2.1. Alcohol prices and demand

As stated, taxation is the key method in controlling the prices of alcohol and changes in alcohol taxes have direct effects on alcohol demand (see Mäkelä and Österberg, 2009). In figure 1, current drinking reflects demand for alcohol. In order to study the relationship between drinking and alcohol prices, a measure for relative changes is needed. Thus, price elasticities of demand have to be investigated.

For policy makers, the elasticities tell about the effects of alcohol tax adjustments on consumption. To study the relationship between alcohol use and earnings, price elasticities may help to unwind some of the ambiguity arising from the endogenous nature of alcohol use in the equation (Dave and Kaestner, 2002). If alcohol is a normal good, it has negative price elasticity and an increase in price causes a drop in consumed quantity. On the other hand, normality also implies that alcohol demand increases with income. Not taking this demand effect into account would cause upward bias in the effects of alcohol use on earnings (Tekin, 2004). According to a comprehensive meta-analysis by Wagenaar et. al (2008), the average price elasticity including all beverage types is around -0.44, which indicates fairly little changes in demand as prices change. The authors also report that there are differences among beverages, beer being the most inelastic alcohol category.

In addition to aggregate elasticity variations between beverage types, individual characteristics lead to different outcomes. This approach introduces also the possibility of more elastic demands. For example, differences between sexes or age groups may prove to have significant implications. Commonly younger people have smaller earnings and alcohol consumption takes relatively larger

[2] Alternative presentation is given in section 2.5.

share of a young person's income. Therefore even small price changes can have large effects on demanded qualities and own price elasticities for beverages are large. Cook and Peters (2005) studied drinker's bonus on young people and replaced drinking variables with alcohol price variable. The authors confirm the existence of drinker's bonus and conclude that higher alcohol prices are associated with better labor market outcomes, indicating reverse causation as main cause of drinker's bonus: at least for young people, increased income leads to higher alcohol consumption. Keng and Huffman (2007) looked at young people's binge drinking and price elasticity, stating that binge drinking is responsive to the price of alcohol and elasticities are relatively large.

Another way to look at the relationship between prices and consumption is to make a division between different drinking patterns. Manning et al. (1995) investigated price effects on light, moderate and heavy drinkers. The authors conclude that both light and heavy drinkers have significantly less elastic demand behavior than moderate drinkers. Furthermore, the upper tenth percentile consumes 51 percent of all alcohol and this percentile has almost perfectly price inelastic demand. A similar consumption for the upper tenth percentile is evident also in Finland (Mäkelä et al. 2010).

One interesting aspect of alcohol consumption and earnings relates to the actual price of alcohol: if moderate alcohol users earn "drinker's bonus", will the amount of money spent on buying alcohol eliminate the wage premium? This issue has been completely ignored in previous academic research, yet it is reasonable to assume that even for moderate users alcohol use may constitute a notable part of total spending. In order to include this aspect into the analysis, price measures for different beverage classes, such as average prices during the survey year, are required. Subsequently, these prices will be used to calculate the total cost of alcohol consumption, which is contrasted to earnings.

However, there are limitations to the aforementioned earnings-costs comparison. As a person succeeds in increasing his earnings, he will use more alcohol (assuming that it is a normal good). There will be changes also in consumption patterns because wealth effect shifts alcohol consumption towards better quality beverages. Thus, more expensive brands will be consumed and costs of alcohol use increase. This limits the comparison measures, since the quality of alcohol beverages can't be easily indexed in survey data. Nevertheless, it can also be assumed that the change in consumption habits happens between beverage classes, i.e. consumption of wine is

positively related to income. Thus, it may be sufficient to use the average prices for different beverage classes after all. Another limitation is the place of consumption, since prices in restaurants and bars are manifold in comparison to retail store prices. This problem can be overcome by having a variable that separates consumption at home and in restaurants[3].

2.2. Previous drinking

The study of price elasticities showed that the heaviest drinkers have almost perfectly inelastic demand behavior. On average, price elasticity estimations for alcohol yield small values, which mean relatively inelastic demand across drinker types and beverage classes. One driver behind this is the fact that alcohol has a habitual or even addictive nature.

Johansson et al. (2007) studied alcohol dependency and employment probability in Finland and in their sample 16% of men and 4% of women were classified as alcohol-dependent. Yet, there has been controversy about how alcohol-dependency is defined, some researchers saying that it is more of a disease than a complex set of health outcomes (see Mullahy and Sindelar, 1993). For the upper tenth percentile of users, alcohol-dependency may appear even in physical forms and they may suffer from alcohol withdrawal syndrome after prolonged periods of drinking. On the other hand, moderate users or even for heavy users most likely don't suffer from alcohol addiction, but rather face occasion-dependent habit formation. In certain social settings, most individuals find it hard to resist drinking alcohol as it works as a social lubricant. This kind of habit formation occurs even if the negative consequences are evident. For example, a normal quote after a binge drinking night is: "I'll never drink again". Usually this statement doesn't have long-lasting effects, emphasizing the role of sociability in habit formation context.

In modeling addictive consumption, two kinds of models have been normally used. In myopic models, past consumption stimulates current consumption and individuals do not take future consumption into consideration. In rational addiction models, on the other hand, also anticipated future consumption is taken into account. Becker and Murphy (1988) introduced rational addiction model in their seminal paper where they showed how the decision to use harmful substances can involve rational forward-looking utility maximization. It has become one of the most used frameworks to model the demand for addictive substances. Baltagi and Griffin (2002) followed the

[3] Peters and Stringham (2006) included a bar-hopping variable in their earnings equation, yet they didn't control the prices of alcohol.

rational addiction approach and specified the following model of consumer's lifetime utility maximization in case of alcohol

$$\sum_{t=1}^{\infty} \beta^{t-1} U(C_t, C_{t-1}, Y_t, e_t) \quad (3)$$

where $\beta = 1/(1+r)$, r is the discount rate that is equal to time preference, C_t is alcohol consumption in period t, Y_t is the consumption of a composite commodity and e_t is the impact of unmeasured life-cycle variables. The budget constraints for maximization are the following:

$$C_0 = C^0 \text{ and } \sum_{t=1}^{\infty} \beta^{t-1} (Y_t + P_t C_t) = A^0 \quad (4)$$

P_t is the price of alcohol and C^0 is the initial alcohol consumption level. A^0 is the present value of wealth. The author assumed a quadratic utility function and derived first-order conditions where current liquor consumption is a function of past and future consumption. The results of the study are consistent with the rational addiction theory, meaning that alcohol dependency can be modeled with traditional economic analysis.

Yet, it may prove difficult to treat alcohol dependency as exogenous in wage models. This can be the results of a specific kind of reverse causation; changes in labor market status may cause alcoholism. If an individual loses job, the risk of alcoholism is increased. Also risk attitudes may lead to biased estimation if alcohol dependency is thought to be exogenous. A strong appetite for risk can thus cause labor market outcomes and alcoholism simultaneously. (Johansson et al. 2007)

A further aspect related to addiction is tobacco use. Koksal and Wohlgenant (2011) estimated how demands correlate between alcohol beverages and cigarettes within the rational addiction framework. They found that alcohol is a complement for cigarette, but there is no complementary effect the other way around. The authors also state that alcohol is a gateway to cigarettes, especially in social sense: drinking e.g. in a bar may increase the likelihood of smoking. In alcohol-earnings puzzle literature, smoking has been included in models by some authors. Van Ours (2004) estimated the joint effect and discovered that alcohol use has a positive effect on wage while smoking has a similar size negative effect on wage. He also looked at the starting rates of the two substances and

found that they are perfectly correlated. His study supports the view of Koksal and Wohlgenant (2011) in the sense that tobacco use rarely happens without alcohol use.

Overall, theoretical framework on addiction implies that demand behavior for alcohol needs consideration for habit formation and alcoholism. As noted, moderate and heavy users may also suffer from addiction, at least in the social sense.

2.3. Direct effects on productivity

Figure 1 shows the effect of current alcohol use on earnings, which is mediated through productivity. In essence, earnings only indicate the underlying productivity if social factors are omitted. Clearly, it is extremely difficult to measure productivity directly because only a fraction of labor force is assigned to repetitive tasks that produce easily accountable tangible items. In contrast, knowledge intensive work focuses on analysis and development tasks on intangible issues, making direct productivity measures almost impossible.

There has been controversy over the correlation of productivity and wage in the academic literature. Medoff and Abraham (1980) made a pioneering study on individual level data, concluding that there is a strong positive relationship between experience and earnings, but performance and earnings are not clearly related. Only few studies have been made since then, all with mixed results. Thus, social and other indirect factors may explain more about the individual labor market success. This also shows that alcohol most likely works in an indirect way in this context.

Despite the ambiguity of measuring productivity, there are other related instruments that may be affected by alcohol consumption. For instance, hours worked has been used to measure labor market outcomes of drinking. Kenkel and Ribar (1994) stated that “the effect of alcohol on labor supply, in turn, can be decomposed into its effect on labor force participation and its effect on hours worked conditional on participation”. The authors didn’t find any negative supply effects of alcohol abuse and stated that lower earnings are entirely caused by decrease in wage rate. However, the sample consisted of young adults who don’t suffer from the negative health effects of long-time abuse. Over time, alcohol abuse creates cumulative health effects that impair productivity and affect labor supply decision, both in terms of participation and hours worked.

Studies on the supply decisions show the importance of time dimension. Most productivity effects develop gradually and become evident later in life. On the other hand, there are many short-term effects that may be used as instruments for alcohol consumption pattern. Of course, the most evident short-term productivity effect is working intoxicated, which is heavily penalized in most of the working environments. Yet, some professions "require" on-site drinking: salesmen often eat and drink together with customers, which is an essential part of their work and companies happily pay for the expenses of drinking. Maybe a more indicating measures of short-time productivity effects are absenteeism and impaired productivity. Impaired productivity may refer to hangover; it temporarily affects mental and physical capabilities, but if it occurs too often, it can lead to firing or suspension from work. Both absenteeism and impairment can also be results of alcohol related accidents.

One fundamental problem related to drinking's effect on productivity or earnings concerns sample selection, since some studies take into account only those individuals who are employed. Indeed, some researchers have addressed this by analyzing alcohol-dependent individuals' employment probability. Both Mullahy and Sindelar (1993) and Johansson et al. (2007) found that alcoholism is closely related to employment probability, but does not show up as decreased earnings. Both studies also noted that it is crucial to pay attention to age groups. Table 1 below presents an overview on the direct effects of alcohol that affect productivity.

Table 1: Direct effects of alcohol use

Short-run	**Long-run**
• Absenteeism	• Impaired productivity
• Intoxication	• Hours worked
• Hangover	• Labor force exit

Based on the existing literature, it seems that earnings are not directly affected by productivity, suggesting that alternative approaches are needed to link alcohol use and earnings. However, the fore mentioned measures, such as absenteeism and hangover, can be used to provide more explanatory power in alcohol-earnings models, but to disentangle the causal links, it is necessary to look at the indirect effects.

2.4. Indirect effects on productivity

2.4.1. Health status

The previous section indicated that numerous factors can affect productivity. Looking at table 1, all of these productivity effects have health-related dimensions. In this regard, productivity and health approaches overlap in many ways, but health status is more related to the long-run aspects. On the other hand, health stock and human capital formation are highly correlated and could be combined into a single model, but following Grossman's (1972) approach, they are analyzed separately.

Majority of studies on alcohol-earnings relationship have argued that the drivers behind a wage premium or penalty are health related. A common tool for formulating the demand for health has been Grossman's (1972) health capital model. In this model, health can be viewed as a durable stock of capital producing output in terms of healthy life. He argues that health capital differs from other forms of human capital because health determines the total amount spendable for working, while knowledge affects the total productivity. Cook and Moore (2000) took alcohol into account and presented the following health capital function:

$$H_t = H(M_t, A_t, H_{t-1}, v_H) \tag{5}$$

In this function, current health stock H_t is produced by medical care M_t, alcohol use A_t, previous period's health stock H_{t-1} and other determinants of health v_H. Medical care's effect on health is positive, while the effect of alcohol use is dependent on the quantity consumed. The functional form of health production depends on the hypothesized relationship, some studies including quadratic or cubic terms of alcohol use to take non-linearities into account. Grossman's model also includes depreciation of health stock with age and investments that can increase the stock. This has the following presentation:

$$H_{t+1} - H_t = I_t - \delta H_t \tag{6}$$

where the change in health stock is determined by gross investment I_t minus depreciation of health stock, δH_t. After the health capital function is explicitly defined, it can be used in the earnings function as one of the components that affect income. Mullahy and Sindelar (1993) constructed an earnings function in which health capital and human capital components are separately defined.

Borrowing from their work, a justification for health capital in earnings function can be derived in the following way. Supposing that earnings function constitutes of alcohol use and health capital, $Y = Y(A, H)$, and one is interested in how earnings vary with alcohol consumption, the total change is following:

$$dY/dA = Y_A + Y_H \frac{dH}{dA} \quad (7)$$

The total derivative divides alcohol's effects to direct effects Y_A and indirect effects $Y_H \frac{dH}{dA}$ through changes in health capital. Thus, equation 7 shows the theoretical importance of controlling for the changes that alcohol use causes in other components of earnings models[4].

In equation 5, the only explicitly given determinant of health stock is alcohol use. However, a variety of other determinants could be added to the equation. For example, chronic illnesses not only affect health status, but also have an effect on drinking. MacDonald and Shields (2001) included non-acute illnesses, such as asthma and diabetes, in their model equation. The authors hypothesized that these illnesses are not severe enough to negatively affect labor market outcomes, but they probably lead to reduced alcohol consumption. Their results proved the hypothesis, with all categorized non-acute illnesses leading to significantly reduced alcohol consumption, regardless of earnings. Evidently, chronic illnesses need to be controlled to determine the true effect of alcohol use, but alcohol use is also one of the main causes for chronic illnesses. Again, it is a matter of time perspective and alcohol-related illnesses need a long time before they may become evident.

The slowness that is affiliated with the negative health effects of alcohol use is also true to the opposite: positive health effects may take a long time before they start to appear. A recent study (Hvidtfeldt et al. 2010) shows that alcohol is associated with decreased risk of coronary heart disease, and the risk is relatively lower among the older people. The results also show that the risk decreases significantly relative to drinking amount. Other studies have found U-shaped relationship between alcohol consumption and heart diseases, as well as between alcohol consumption and total mortality (see Gaziano et al. 2000). It is intriguing that the U-shaped relationship coincidences with the inverse U-shaped relationship that has been frequently found between alcohol use and earnings.

[4] The same derivation applies to human capital, if it is also treated as its own entity. In this regard, health measure H could be changed to a measure of human capital, K.

Looking at equation 5 again, some health related attributes may be jointly determined with drinking status, such as obesity. However, obesity rarely appears in the studies of alcohol-earnings puzzle. Yet, one could consider that obesity is partly caused by alcohol consumption and obesity may also affect labor market outcomes. A recent Finnish study (Johansson et al. 2009) showed that obesity affects labor market participation, but wage effects are significant only for women. Dastan (2010) estimated the labor market effects of obesity, smoking and wages, contemplating that these factors might be interactive or additive. He found that if one fails to control all health behaviors, effects of obesity are underestimated while effects of binge drinking are overestimated. Overall, the effects of obesity appear to be ambiguous and significant only for women.

As discussed earlier, other addictive and risk-related behaviors, such as smoking, may need controlling and this applies also to the health capital model. Many authors have addressed the importance of risk preferences (e.g. Barret, 2002). Peters (2004) states that risk-loving nature is one of the variables not likely to change with age, emphasizing its validity as a measure that predicts the drinking behavior. Lower risk tolerance also explains partly the higher prevalence of drinking among males (Dave and Saffer, 2008). Drug use is another risk preference indicator that could be included in the health capital model and, consequently, in an earnings equation. However, this has been studied with mixed results and some studies have even found a positive labor market effect of drug use (MacDonald and Pudney, 2000).

Risk preferences may lead to different drinking patterns and health capital formation may in turn differ according to drinking pattern. For example, the depreciation rate in equation 6 could be faster for those individuals who binge drink frequently. Binge drinking can also mean higher medical care costs in equation 5. This can be the result of injuries sustained while intoxicated, such as broken bones and lost teeth. In the long-run, heavy drinking can result in chronic illnesses and high medical care costs.

An essential distinction in both risk preferences and drinking patterns is gender difference; this also affects health capital accumulation. Almost all studies have analyzed genders separately, often with conflicting results. For instance, Kenkel and Ribar (1994) found that women with alcohol problems earn 10% more than women without alcohol problems. For males, the opposite was true. Macdonald and Shields (2001) showed that the wage premium peaks in greater alcohol amounts for females than for males, while some researchers have not even found a wage gain for women and

results have been insignificant (Zarkin et al. 1998). Tekin (2004) argued that women drink less frequently and don't binge as much as males, which can lead to different labor market outcomes. Yet, some researchers have argued that health is not an important mechanism in the alcohol-earnings relationship and unobserved personal characteristics explain most of the variation (Auld, 2004).

2.4.2. Human capital

Becker (1962) stated the following in his seminal work: "This paper is concerned with activities that influence future real income through the imbedding of resources in people. This is called investing in human capital." In this regard, also drinking can be viewed as an investment in human capital. Moderate drinking may improve social or cognitive skills which lead to higher earnings. Obviously, too intensive and frequent drinking will have a negative effect on both. In this section, the focus is on cognitive skills, meaning the effects of alcohol use on human capital formation via schooling and other forms of knowledge accumulation.

Becker (1962) divided human capital investments to on-the-job training, schooling, other information and health. Since health investment is already discussed in the previous section, the focus is on the other forms of investment. According to Becker's view, on-the-job training includes all forms of learning while working. In alcohol-earnings studies, this has usually been indexed by labor market experience measure. Schooling on the other hand can be indexed by various ways and it is one of the key explanatory measures in the models. Educational variables often include a binary variable telling whether or not an individual has obtained high school or university degrees. In addition, Balsa et al. (2011) included grade point average as one of the schooling measures and found that increase in alcohol consumption resulted in small reductions in GPA. Other information is much vaguer concept and finding variables for it is extremely difficult.

The human capital function can be constructed in the same manner as the health capital function. The factors affecting human capital stock are the fore mentioned determinants and alcohol use. Analogically, equation 8 presents an example function for the determinants of human capital.

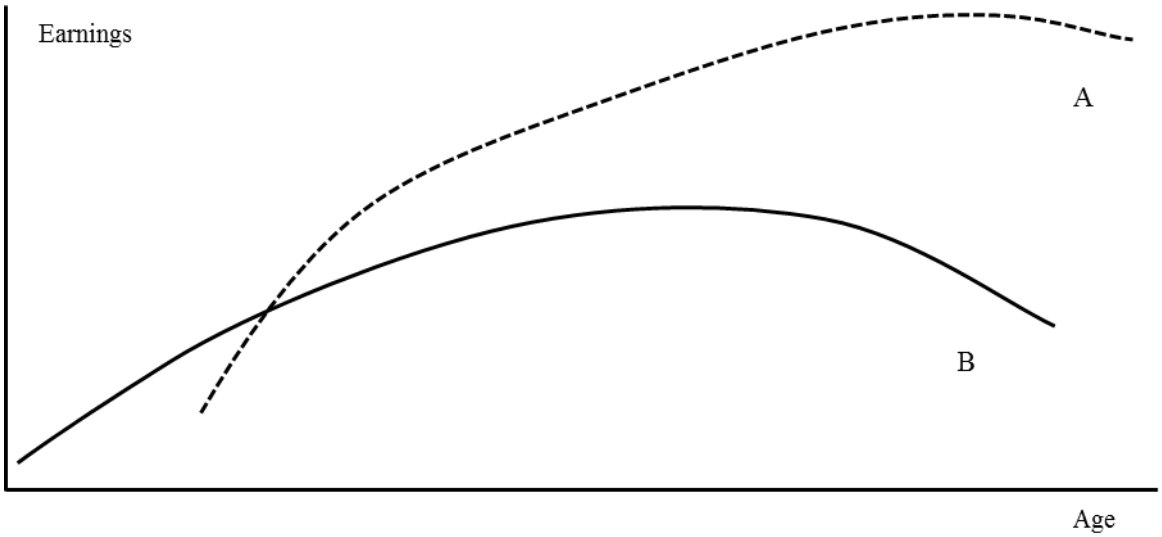
$$K_t = K(S_t, L_t, A_t, v_K) \tag{8}$$

In equation 8, human capital stock K_t is produced by schooling S_t, labor market experience L_t, alcohol use A_t and other determinants of human capital v_K. Replacing health capital with human capital in equation 5 yields similar outcome, too. Human capital depreciates with a constant rate without investments, such as schooling or relevant experience. Also the derivation of total change in earnings is similar as in equation 6. Alcohol-caused change in earnings is the sum of alcohol's direct effect and indirect effect through human capital.

Many authors have discussed the dimensions of alcohol use and human capital, yet Bray (2005) states that this approach has not been fully exploited. On the other hand, numerous authors have estimated the effects of drinking on schooling and found significant effects on human capital accumulation. Williams et al. (2003) found that alcohol use reduces GPA mainly via a reduction in the time spent on studying, concluding that heavy drinking reduces human capital formation and hence has a negative effect on future labor market outcomes. Renna (2008) displayed that binge drinking causes temporal dropouts, leading to prolonged completion of high school and poorer diploma. This reduction in human capital accumulation in turn reduces future earnings.

Since alcohol use appears to have effects on human capital accumulation, there can be long-lasting impacts on labor market outcomes. The development in earnings starts to differ as soon as individuals start to make schooling and labor market choices. This leads to different age-earnings profiles that can be explained by life-cycle hypothesis. The hypothesis has been discussed by many researchers in the study of alcohol-earnings relationship (Mullahy and Sindelar, 1993; French and Zarkin, 1995). According to explanation by French and Zarkin (1995), alcohol use starts to affect earnings profile already while at school. Heavy users may settle for less schooling or drop out of formal schooling earlier. Thus, they enter labor market at younger age and have already gained experience and wealth when the more schooled enter labor force. At that point, it may look as heavy drinkers have higher earnings, but over time those who drink less will have higher earnings level. This is due to better education and, on the other hand, because heavy drinkers will face negative consequences from their alcohol use that affect earnings. Utilizing the age-earnings profile framework of Polachek and Siebert (1993), figure 2 below presents age-earnings comparison of two different individuals.

Figure 2: Age-earnings profiles



In figure 2, the difference between starting age is the extra schooling that individual A receives. As A enters labor force, B has already gained working experience and earnings and is initially on a higher earnings level. However, A's earnings increase rapidly and the increase rate stays higher for the rest of the working career. The fore mentioned negative consequences of drinking lead to faster deterioration in human and health capital, causing decrease in the earnings and earlier exit from labor market. Hamilton and Hamilton (1997) found empirical evidence about the hypothesis, stating that heavy drinkers possess flatter age-earnings profiles. Furthermore, they receive lower returns to higher education than other drinker types and the beneficial earnings effect of being married also disappears.

Studies on the relationship between human capital and alcohol use have yielded contrasting results. It appears that alcohol reduces human capital accumulation by reducing the time spent on studying; this is fairly direct effect. On the other hand, the indirect effects may even enhance human capital accumulation. Bray (2005) found that alcohol has this kind of positive effects, contemplating that there are two possible ways how alcohol affects cognitive skills. First, moderate drinking may improve socialization, making human capital accumulation more efficient. Second, cognitive functioning may actually increase as a result of alcohol use.

2.5. Sociability and alcohol

Many authors have contemplated that the most significant underlying factors in alcohol-earnings relationship are related to social networking. Since networks involve multiple decision makers and are often implicitly defined, estimating the social effects in labor market is difficult. On the other hand, the concept of social capital has emerged for analytical examination of social effects. It can be considered as an extension of human capital, yet Coleman (1988) classifies social capital as being even less tangible than human capital, since it exists only in the relations of people. In turn, Glaeser et al. (2000) conducted a comprehensive analysis that enabled to analyze social capital on individual level, a viewpoint that the authors stated in the title "*The Economic Approach to Social Capital*".

Glaeser et al. (2000) defined individual social capital as a person's social characteristics that help him to get market and non-market returns as a result of social interactions. The authors included both intrinsic abilities (e.g. extrovert nature) and social capital investments (widening social network) in their framework. On basis of the empirical part of the study, the authors made multiple predictions on social capital formation. For example, social capital first rises with age and later declines if instrumented with organizational membership. Another result was complementarity of human capital and social capital, meaning that people who invest in human capital (education) also invest in social capital.

Evidently, individuals can invest in social capital and empirically tested hypotheses support common sense reasoning about sociability. However, it is not clear to what extent alcohol use affects social capital accumulation. Only a couple studies have tested sociability hypotheses empirically. Peters and Stringham (2006) used bar-hopping as an indicator for social drinking and concluded that those who drink in a social setting get an extra wage premium. Yet, the authors used a quite naïve model and their results should be interpreted with caution. In another study, Peters (2009) studied how earnings differ between officers and enlisted personnel. She found that officer get larger wage gain than enlisted personnel in each drinking category and claims that this is due to better promotion possibilities among drinking officers.

If social capital functions similarly to human and health capital, alcohol's effects on productivity are mediated through social capital and, consequently, changes in productivity lead to different

labor market outcomes. For example, a salesman may have acquired social skills that are partly due to his previous drinking. In a work related social setting, he can use these skills to meet new people or convince potential buyers. In addition, he can benefit from current drinking as confidence grows and confronting new people becomes easier. As a result, he has achieved a wider contact network and since contacting new people is his main duty, his productivity is increased. On the other hand, socializing and drinking with possible future colleagues can reduce information asymmetries on both sides of the table, leading to better selection of employees. Again, this can have productivity increasing implications.

A further aspect is provided by MacDonald and Shields (2001). They state that social time spent with colleagues may signal commitment and motivation in the eyes of superiors. Plausibly commitment is reciprocal, meaning that social drinking reinforces employer-employee relationship. Consequently, working environment becomes more stable and power distance may reduce, making it easier to share ideas. Social drinking may have beneficial productivity effects also in this sense. Nevertheless, drinking certainly has many adverse effects on social capital. Especially binge drinking may lead to situations where an individual acts incorrectly. In the worst case, inappropriate intoxication can cause losing a promotion opportunity or even firing.

The effects of alcohol use on social capital are not limited to the productivity effects and in this regard, it differs from human and health capital. In order to clarify this, it is necessary to consider sociability as means of improving one's labor market outcomes. MacDonald and Shields (2001) hypothesized that alcohol consumption may lead to additional social time spend with colleagues and associates, which gives possibilities to obtain information about promotion opportunities and free vacancies. This approach can also work outside a given organization and it can be applied to whole labor market.

However, this employment-seeking drinking behavior may also imply adverse selection for the firm. If a hiring decision is made through informal social networks, the employer may be unaware of the underlying true productivity of a possible employee. Then again, the overall productivity differences may be minimal and social ability is a key mean to differentiate oneself from other employment seekers. Companies often state that they look for "great personalities" and social drinking may help to distinguish those personalities from the rest. To sum up all this, I constructed an altered diagram about the factors in alcohol-earnings relationship, which is presented in figure 3 below.

Figure 3: The effects of alcohol use with social capital

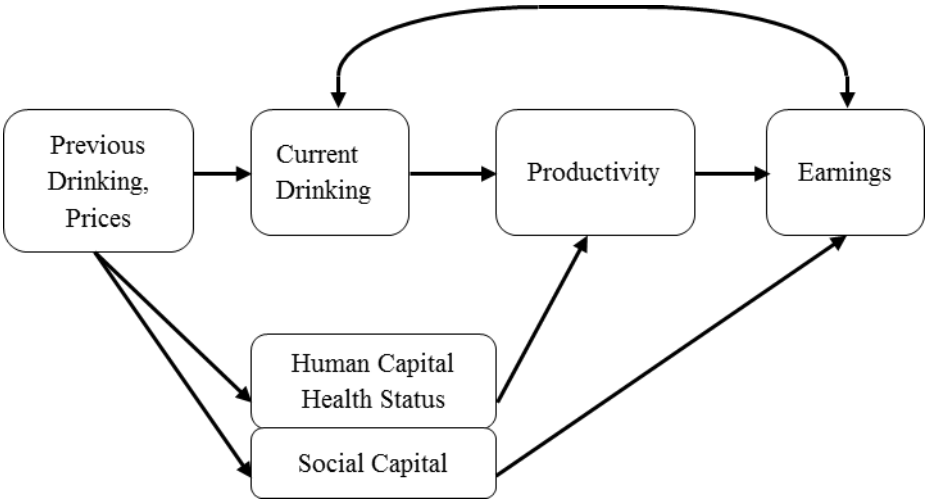
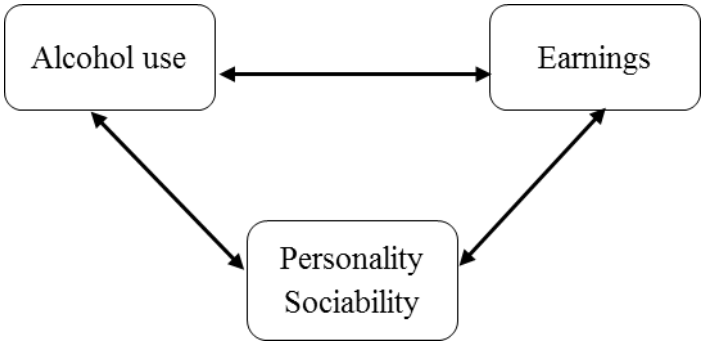


Figure 3 shows that social capital can have a direct effect on earnings, which is not mediated through productivity. On the other hand, the effect is depicted also in the reciprocal relationship between current drinking and earnings. This fundamental difference in the diagram is related to the opportunity-enhancing nature of social drinking (both current and past) in labor market. It is necessary to bear in mind that the distinction between social capital and other indirect effects is not as clear-cut as in figure 3, but the main purpose of the figure is to elucidate the differences between the effects.

Another essential issue regarding sociability-drinking relationship is the importance of personality traits. Alcohol use may not cause changes in social capital, but rather reflects latent dimensions of personality. Lye and Hirschberg (2010) made a retrospective study on the alcohol-earnings relationship in which they confirm the existence of drinker's bonus, regardless of the study method. The authors suggested that the reason for the bonus lies in a common set of personality traits which determine both drinking behavior and higher earnings. In addition, they stated that alcohol consumption may serve as an instrument for these traits, yet this is not true for all variables that are thought to affect human capital. For example, extroversion doesn't appear to affect earnings, yet it is related to alcohol consumption. To make the connections between personality, alcohol use and earnings clear, a visual presentation of the relationship is given in figure 4.

Figure 4: Alcohol use, Earnings and Personality traits



In figure 4, alcohol use and earnings have the fore mentioned reciprocal relationship, but personality traits affect them both. Although extroversion may not affect earnings, other personal traits may well affect both alcohol use and earnings as in figure 4. Lye and Hirschberg (2010) claim that when emotional stability is controlled, alcohol use doesn't appear to have effect on earnings. Thus, alcohol use only indicates the stability measure. Furthermore, the authors claim that it is important to control for ex-drinkers, since they are more likely to suffer from depression. Omitting the measure of former drinker may cause abstainers to look less mentally stable than they actually are. This means a serious bias in the alcohol-earnings estimation if personality is a major latent factor in the model.

Even though personality may simultaneously affect both drinking and earnings, a more complicated network of effects is most likely present. It can be seen from figure 4 that all three components have reciprocal nature, implied by two-way arrows. For example, people may use alcohol to reduce tension, which can help to maintain stable mood and personality. Lye and Hirschberg (2010) refer to a report by UK Mental Health Foundation, according to which a common reason for drinking is alleviation of anxiety and depression. Surely, this sort of mood stabilizing activity is depended on quantity of alcohol consumed and for example alcohol abusers may have started the use to alleviate depression, but over time alcohol itself becomes the problem.

Skogen et al. (2009) found a U-shaped association between alcohol consumption and the risk of anxiety and depression, showing that anxiety minimizes at the same level of consumption as earnings tend to maximize in the alcohol-earnings literature. This indicates that moderate drinking

may be beneficial also in terms of psychological well-being. Furthermore, the authors state that increased risk of anxiety among abstainers cannot be explained by illness, social activity or former abuse-drinking. In addition to anxiety and depression reduction, alcohol is commonly used to alleviate stress. For many individuals, relaxation activities are accompanied with drinking activities[5].

It is almost certain that personality-related attributes have an important role in alcohol-earnings relationship. While personality may explain a large share of alcohol-related earnings effects, the opposite may also be true: alcohol affects personality traits and the effect is mediated to earnings. Furthermore, it seems that former drinking of abstainers needs controlling (Lye and Hirschberg, 2010), but it may still not explain possible earnings differences from psychosocial point of view (Skogen et al. 2009). On the other hand, moderate alcohol use not only reduces anxiety, but also helps to build up confidence in social situations where networking may lead to better job opportunities[6].

2.6. Demographics

Getting back to figure 1 again, it was stated that previous drinking and prices influence the current drinking behavior. However, in equation 1 the measure X_t for exogenous covariates captures a large share of differences between individual drinking patterns. Such factors are related to family background and living conditions. They are not part of the causal relationship between alcohol use and earnings but need controlling in order to minimize variations in the model.

Clearly the most significant background factor is gender. Without controlling for gender difference, the estimation may suffer from serious bias and results would be ambiguous, but as stated, most researchers have estimated separate models for both genders. All other exogenous covariates are not gender-specific, yet their earnings effect may vary, for instance in the case of marriage. Keng and Huffman (2007) showed that the positive earnings effect of marriage is larger for men, which is a consistent finding with other studies from different fields. The authors also showed that presence of children causes a substantial decrease in female earnings, while male earnings are almost

[5] An extreme example of this might include binge drinking with the aim to "reboot one's brain", which is a common expression in Finland at least. It may happen after a long period of stressful work or other activities that have prohibited an individual from enjoying life.

[6] And why not to better mating opportunities? Many marriages are a result of meeting intoxicated in a bar for the first time. Marriage helps to stabilize life and a more stable life may imply better labor market outcomes.

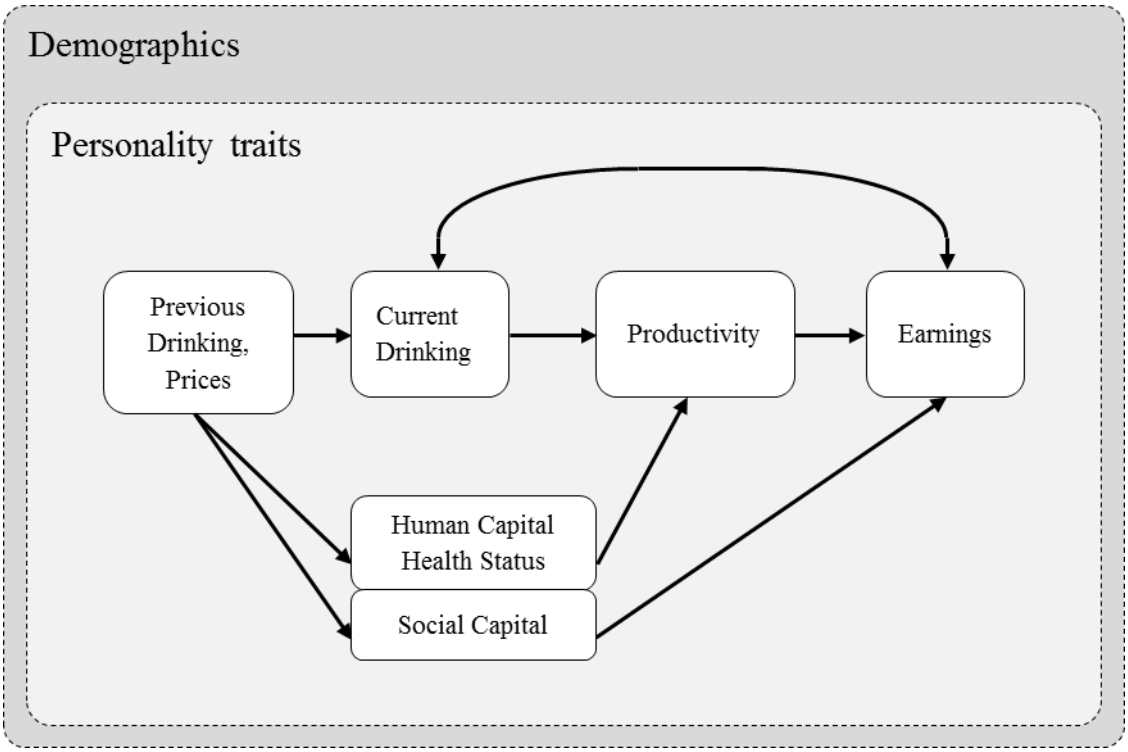
unaffected. These findings are intuitive: married men are less likely to spend their free time drinking, but having children mostly affects women's labor supply decisions.

Another major background factor is living area. Drinking behavior can alter greatly between different parts of a country, yet the key distinction in this regard is between urban and rural areas. For instance, Finland has only one metropolitan area and the rest of the country is made up of smaller cities and rural areas. Urban areas differ in labor structure as well as drinking culture. Ziebart and Grabka (2008) studied beverage-specific drinking effects and found that there are different wage gains on basis of living area: individuals in rural areas are found to gain most from beer drinking, while urban residents get the largest wage gain from cocktail drinking.

Other often used exogenous covariates include race and religion, yet in Finnish context such measures would probably be unnecessary since the population is fairly homogenous. Family background and area specific factors are still needed and they generally improve the explanatory power of alcohol-earnings models. Another important background factor is the socioeconomic stand (SES) of an individual. Yet, it cannot be taken as an exogenous covariate, as it correlates highly with earnings. Mäkelä (1999) studied the relationship between socioeconomic stand and alcohol-related mortality, concluding that alcohol-related mortality is higher in the lower socioeconomic groups. However, in this study SES is used as a control variable that separates the analysis between different SES groups.

To combine the demographic effects with personality traits and the framework presented in figure 3, a more comprehensive model is presented in figure 5 below. In this figure, the effects of drinking are the same as before, but the whole network of causes and consequences is affected by individual personality traits and the surrounding demographic conditions. This way, the total picture of alcohol-earnings puzzle becomes clearer and all the issues presented in the theoretical part can be combined into a single framework. Figure 5 also provides the essential background for the empirical part as it shows how the unobserved heterogeneity between individuals may affect each part of the alcohol-earnings puzzle. This framework also enables to reason for the cause and effect as all the forces that affect the relationship are known.

Figure 5: The alcohol-earnings framework



3. Empirical framework

3.1. Econometric methodology

The study of alcohol-earnings puzzle requires methods that can be applied for micro-level data. In the former studies, researchers have used a variety of econometric models to investigate the effect of alcohol use on earnings. Some researchers have relied on ordinary least squares (OLS), which is the basic method in majority of the studies. Others have used instrumental variable (IV) techniques, such as two-stage least squares (2SLS), or selection models. These can be called structural methods and they are used mainly to cope with problems related to endogeneity. On the other hand, some authors have used longitudinal data to handle the statistical problem and to get more information on the individual fixed effects of drinking. Table 2 summarizes the most relevant studies, their methods and results.

As can be seen from the main finding column, there is clear tendency towards drinking premium for moderate drinkers, regardless of the study method. Many authors have tested both OLS-models and structural models, with mixed results. OLS is basically used for linear estimation, but OLS-models can also include quadratic and cubic terms. However, for OLS to yield robust estimates the regressors need to be exogenous and multicollinearity should not exist. If the depended variable is continuous, the multivariate OLS model for alcohol-earnings estimation in its basic form is the following:

$$W_i = \beta_0 + \beta_1 A_i + \beta_2 X_i + \varepsilon_t \quad (9)$$

where W_i is the measure of earnings, A_i is the measure of alcohol use, X_i is the vector of all other covariates and ε_t is the error term. This a simplification of the earnings equation 1 in a sense that alcohol use is measured only for the current period and all other determinants are collapsed into a single vector that does not explicitly count for health and human capital effects. In alcohol-earnings literature the most common way to begin the analysis has been a model with single indicator variable for alcohol use.

Table 2: Earlier studies

Authors	Year	Main econometric model	Alcohol use variables	Main finding
Berger and Leigh	1988	Structural	Drinker dummy	Drinking increases earnings
Mullahy and Sindelar	1993	OLS	Alcoholism dummies	Alcoholism affects labor force participation
French and Zarkin	1995	OLS	Drinker categories / linear and quadratic term	Inverse u-shaped relationship
Hamilton and Hamilton	1997	Structural	Non-drinker / moderate / heavy	Moderate drinkers earn most
Zarkin et al.	1998	OLS	Many dummy categories	Constant drinker wage premium
MacDonald and Shields	2001	OLS / Structural	Weekly alc units/ dummy categories	Inverse u-shaped relationship
Tekin	2004	Longitudinal / fixed effects	Ethanol amount / dummy categories	Constant drinker wage premium
Saffer and Dave	2005	Structural	Average drinks per day	No positive earnings effect
Peters and Stringham	2006	OLS	Drinker dummy / bar hopper dummies	Social drinking increases earnings
Keng and Huffman	2007	Longitudinal / Structural	Binge drinking categorical variable	Binge drinking decreases earnings
Ziebart and Grabka	2008	OLS / Structural	Dummy categories / Beverage-specifc dummies	Earnings premium for moderate drinkers and winedrinkers
Peters	2009	OLS	Drinks per week dummies	Larger earnings premium for officers than enlisted personnel

Earlier studies have often chosen to study genders separately. Consequently, the following equation shows a more comprehensive model of the determinants of earnings separated by gender:

$$W_i^{male} = \beta_0 + \beta_1 A_i + \beta_2 A_i^2 + \beta_3 D_i + \beta_4 H_i + \beta_5 K_i + \beta_6 X_i + \varepsilon_i \qquad (10)$$

This follows a similar approach as French and Zarkin (1995) had in their study. In this specification, A_i is the continuous measure of alcohol use in centiliters. Quadratic alcohol use term is added in order to account for possible nonlinear relationships. D_i is the measure of other alcohol-related

variables, such as intoxication starting age and a dummy for ex-drinkers. H_i is a vector of health variables, K_i is a vector of human capital variables and X_i is a vector of demographic variables. To study different hypotheses, the selection of covariates differs on a case-by-case basis. Furthermore, the gender differences can be indexed only as a female dummy.

A structural econometric model can provide better estimates than OLS when it comes to the evaluation of unobserved heterogeneity and when there is a need to fit models comprising of simultaneous equations. However, the effectiveness of instrumental variable approach crucially depends on the quality of the instruments (Dave and Kaestner, 2002). Furthermore, the instruments need to be correlated with the endogenous variables, while they should be uncorrelated with the error term of the earnings equation. This basically means that an instrument should be correlated with the drinking decision, while it should not affect earnings. In this sense, the most used instruments in alcohol-earnings studies are alcohol prices or taxes. However, to be able to use these measures, the data should contain observations from multiple periods or differences in taxes or prices between regions. Additionally, there must be at least as many instruments as there are endogenous variables.

Most authors acknowledge the problem of endogeneity. In addition to price or tax measures, some authors have used parents' or partner's drinking problems as instruments as they do not need to be controlled over time. Some authors have used relatively vague instruments, such as attendance at religious services (Hamilton and Hamilton, 1997). Surely, it may explain drinking patterns to a certain degree, but at least in the Finnish context the amount of individuals who restrain from drinking due to religion is small. Dave and Kaestner (2002) also emphasized the importance of observations in the context of instrumental variables: heavy drinkers' demand for alcohol is relatively inelastic and thus a large number of observations is needed to obtain robust estimates. Furthermore, many researchers (e.g. Tekin 2004, Mullahy and Sindelar 1993) found out that the instruments performed poorly in the first stage of 2SLS estimation, which led to rejecting the use of IV method. Tekin (2004) also stated that the estimated values for instruments are imprecise and implausibly large. Moreover, Ziebart and Grabka (2008) ran a test for endogeneity of drinking behavior and did not find evidence on endogenous relationship between drinking and earnings, concluding that OLS estimates should be used.

Overall, every econometric method has it problems. Due to the difficulty of finding good instruments and the fact that the data set in this study is cross-sectional, I use OLS as the estimation

method with a multidimensional approach. The justification for using OLS is in its efficiency and suitability to the dataset at hand. Furthermore, the role of descriptive statistics cannot be overemphasized. By studying the underlying characteristics of the sample at hand, one can learn many valuable insights about the possible causes and effects. This way, the results from econometric analysis using OLS are more robust and more comprehensively justified. It also helps in conducting a multidimensional econometric analysis, meaning separate regressions and control for certain background characteristics. For example, this can mean separate analyses for different socioeconomic groups.

3.2. Data

The data for the empirical analysis comes from Finnish Drinking Habits Survey 2008. Authors of the survey are Pia Mäkelä and Heli Mustonen from the National Institute for Health and Welfare. This cross-sectional survey is conducted in eight years interval beginning from 1968 and it aims to analyze the development of drinking habits as well as shed light on alcohol consumption and policy issues at hand. The 2008 survey assessed alcohol consumption and illegal drug use in Finland, type and volume of alcohol consumption, consequences of consumption and general attitudes towards alcohol.

Survey's data is based on face-to-face interviews and a paper questionnaire. The survey consists of several different parts, of which the "base material" section forms the frame of the survey. It includes all the individual level answers on vocal questions. Additionally, a section for measuring one alcohol consumption occasion and an extra section in paper form for measuring illicit drug use are included in the survey. Also register-based background and income variables from the registers of Statistics Finland are combined to the interview data. Face-to-face interviews were carried out during autumn 2008 with randomly selected sample of 3750 persons. The sample consisted of individuals aged between 15 to 69 years, excluding institutional population and individuals living in Åland islands. The total number of respondents in the data was 2750.

Altogether, the survey included 311 questions and register-based figures, of which the variables are formed. The selection of the variables aims at giving the most holistic view on the effects of alcohol use on earnings and allowing for testing of different relationship hypotheses on the basis of the theoretical part. In the following sections, the variables are presented in the same manner as described in equation 1, regarding earnings as the dependent variable. This way, variable definitions

provide easy approach to statistical analysis. Summary statistics are presented in chapter 3.3.4. and the survey questions that were used to form variables are presented in Appendix A.

3.3. Dependent variable

The dependent variable in the analysis is the measure of earnings, which is obtained from the national register and linked to the individual-level survey data. The information on the earnings is based on personal taxation figures from 2008. However, a brief discussion on different ways to measure earnings is needed before the dependent variable can be presented and justified.

The definition of income in the earnings equation has varied in the earlier studies. Some authors (e.g. Bray, 2005) have focused on hourly wage rate by dividing income by hours worked. On the other hand, Kenkel and Ribar (1994) used hours worked as the dependent variable in their study of problem drinking and labor force participation. Still, a number of studies use annual earnings measures. In some sense, the choice of earnings measure has been related to the hypothesized effect between earnings and alcohol use: if a researcher is more concerned with the direct effects on productivity, using hourly wage rate is justified. In this study, the main effects of alcohol use are hypothesized to be indirect and cumulated over time, thus the yearly earnings figure provides a more comprehensive view on the matter. Yet, in large alcohol consumption amounts there probably is a clear negative productivity effect that is a result of past and current drinking.

The measure for earnings in this study is based on yearly income. This means gross income from wages, pensions, other taxable welfare benefits and the proportion of entrepreneurial income defined as earned income. Capital income is not included into the analysis as its share of earnings is insignificant for most of the sample respondents. Furthermore, capital income is not as good predictor of productivity or labor market effects as gross income from wages. The data on yearly income is top coded at 150 000€ and all greater amounts are reported as that figure. However, this concerns only few survey respondents and thus truncation is not a major problem in the analysis. On the other hand, individuals who do not have any earnings are excluded from the sample as the focus of the study is on employed individuals with positive earnings, as was stated earlier. Finally, a natural logarithm of earnings is taken in usual wage equation manner to be able to study the percentage changes in earnings. The resulting amount of observations is 2601.

3.4. Explanatory variables

3.4.1. Alcohol use variables

Measures of alcohol use are the main explanatory variables in the analysis. All alcohol use variables are created by asking about alcohol use during the past year at the time of the interview. The most elementary alcohol use variable is an indicator (or a dummy variable) for whether an individual is a drinker or not, indexed as *Drinker*. As discussed earlier, the existence of ex-drinkers among abstainers should be accounted for and, thus, a dummy variable *Exdrinker* is also included[7]. In addition, a variable for the first intoxication age, *Startdrunk*, is included to enable a look at the age dimension on individual level.

The respondents were also asked about the intensity and frequency of their drinking. This was done for different beverages separately and the resulting number of drinks was used to convert the amounts into 100% alcohol in centiliters by assigning a multiplier for each beverage[8]. This way, the data provides a continuous measure for alcohol use, indexed as *Alc_cl*. To deal with long tails in the upper end of consumption spectrum in this thesis, a cut-off level was determined at 3000 centiliters of pure alcohol per year, which corresponds to over five drinks daily. To capture polynomial functional forms, a variable for quadratic term of alcohol use was created (Alc_cl^2). For measuring the intensity of drinking, a discrete variable *Freq8drinks* is included in the analysis. It measures the number of drinking occasions in which over 8 drinks were consumed, which can be considered as an intoxication amount. Again, to deal with large tail values, a maximum cut-off level was set at 100 intoxications per year.

Furthermore, an overall measure for the risk associated with individual's alcohol use was calculated on the basis of the responses. The risk is distributed into five classes, ranging from *Risk1* for abstainers to *Risk5* for heavy users. It classifies heavy users as men drinking over 40gr and women drinking over 20gr of pure alcohol daily, which is a common limit for heavy drinking in the health literature. These figures are the equivalent of 24 drinks per week for men and 12 drinks per week for women. *Risk2* contains individuals who consume some alcohol but stay under the amounts of heavy usage. Furthermore, individuals in this group never consume extensive amounts at one

[7] The frequencies of drinkers, abstainers and ex-drinkers can be seen from table 3 later in this chapter.
[8] See appendix A for further explanation about the beverage multipliers.

occasion. The limits for extensive amount at one occasion are 8 drinks for men and 5 drinks for women, which basically mean the limits of getting drunk. *Risk3* is formed of individuals who are drunk sometimes, meaning 1-11 times per year, but for whom the total consumption does not exceed the limit of heavy usage. Consequently, *Risk 4* contains individuals who drink extensive amounts at one occasion over 11 times per year, but for whom the total consumption also stays under the limit of heavy usage.

To study beverage-specific effects of drinking, three variables for different beverage drinkers were created from the data: *Winedrinker*, *Beerdrinker* and *Liquordrinker*. For example, an individual was classified as a beer drinker if he consumed beer more often than once a month and did not drink wine or liquor as often. The same classification applies also to wine and liquor drinkers. Together, these three drinker types account for over half of the survey respondents. The remaining individuals could not be assigned with a beverage-specific drinking class, thus they are the reference category.

3.4.2. Education, health and sociability variables

To test for the indirect effects of alcohol use and control for possible covariates in the earnings equation, variables for educational level and health status were included. Schooling is measured with two dummy variables, *Highschool* and *University*, which indicate whether and individual has completed high school or university, respectively.

In terms of health measures, two dummy variables are included in the analysis. Illness is an indicator for having a chronic illness, whereas Obese is an indicator for individuals with body mass index value exceeding 30. Furthermore, a dummy variable *Drug* was added to indicate whether an individual has taken any drugs in his life. The reason for adding drug use variable is to study if the addictive nature of substances plays any part in the earnings equation.

The sociability aspect is studied with a set of variables. *Lonely* is an indicator variable for those individuals who have felt lonely during the last year. *Mental1* is a dummy for those individuals who think that alcohol use has a positive effect on their mental health, whereas *Mental2* indicates whether an individual has used alcohol to alleviate depression or anxiety. *Restaurant* is an indicator for individuals who drink at restaurants or bars in a frequent manner but not in excessive amounts (between few times a week to once a month). *No_joy* is a dummy that gets value 1 when individual

thinks that drinking does not give real joy and *Drunk_fun* is a dummy that gets value 1 when individual thinks that being intoxicated is an innocent way to have fun.

3.4.3. Background variables

The last group of variables deals with individual background. *Female* is a dummy variable for females. *Age* is measured in years and Age^2 is the quadratic term for age. To measure the family relations, a dummy for marital status, *Married*, indicates those individuals who are married and *Children* is a dummy for having under aged children. Finally, five different area variables are included in the analysis. *Area1* is a dummy for individuals living in the metropolitan area, *Area2* for individuals living in other parts of Southern Finland, *Area3* for individuals living in Western Finland, *Area4* for individuals living in Eastern Finland and *Area5* for individuals living in Northern Finland.

3.4.4. Summary statistics of the variables

Table 3 below summarizes all the variables. For each variable, the table presents variable's name and definition as well as the number of observations, range of values, mean and standard deviation. The variables are classified in the same manner as in the earlier chapters. Looking at the age-related measures, it is interesting to see that the range for alcohol use starting age and first intoxication age is relatively large. However, the mean values are plausible and standard deviation is on a reasonable scale. The table is based on the full sample, which means that for example all socioeconomic classes are included in the variable statistics. However, some restrictions for the sample size are presented in the statistical analysis to improve the accuracy of the hypothesis at hand.

Table 3: Summary statistics

Variable	Definition	Obs.	Min	Max	Mean	S.D.
Earnings	Annual earnings	2725	0	150000	22464	17759
Logearn	Logarithm of annual earnings	2601	2.94	11.91	9.74	1.00
Drinking variables						
Alc_cl	Alcohol consumption in cl	2681	0	2960	337	471
Alc_cl^2/100	Alcohol consumption squared					
Drinker	1 if drinker, 0 else	2725	0	1	0.90	0.30
Exdrinker	1 if ex-drinker, 0 else	2725	0	1	0.95	0.22
Startdrunk	First intoxication age	2417	6	60	17.18	4.60
Freq8drinks	How many times >8 drinks	2679	0	94	6.85	14.69
Risk1	1 if abstainer, 0 else	2724	0	1	0.10	0.30
Risk2	1 if moderate drinker, 0 else	2724	0	1	0.31	0.46
Risk3	1 if drunk sometimes, 0 else	2724	0	1	0.34	0.47
Risk4	1 if drunk often, 0 else	2724	0	1	0.19	0.39
Risk5	1 if heavy user, 0 else	2724	0	1	0.06	0.24
Beerdrinker	1 if drinks mainly beer, 0 else	2725	0	1	0.31	0.46
Winedrinker	1 if drinks mainly wine, 0 else	2725	0	1	0.16	0.37
Liquordrinker	1 if drinks mainly liquor, 0 else	2725	0	1	0.15	0.36
Educational, health and sociability variables						
Highschool	1 if high school, 0 else	2725	0	1	0.38	0.48
University	1 if university or college, 0 else	2725	0	1	0.21	0.41
Illness	1if chronic illness, 0 else	2725	0	1	0.39	0.49
Obese	1 if obese (BMI>30), 0 else	2725	0	1	0.16	0.36
Drug	1 if tried drugs, 0 else	2498	0	1	0.13	0.34
Lonely	1 if felt lonely, 0 else	2719	0	1	0.25	0.43
Mental1	1 if alcohol has positive mental effect, 0 else	2511	0	1	0.36	0.48
Mental2	1 if uses alcohol to alleviate depression, 0 else	2565	0	1	0.15	0.35
Restaurant	1 if drinks often in restaurant, 0 else	2725	0	1	0.13	0.33
No_joy	1 if drinking gives no joy, 0 else	2725	0	1	0.41	0.49
Drunk_fun	1 if intoxication is an innocent way to have fun, 0 else	2725	0	1	0.35	0.48
Background variables						
Female	1 if female, 0 else	2725	0	1	0.52	0.50
Age	Age in years	2725	16	70	43.72	15.46
Age^2	Square of age					
Married	1 if married, 0 else	2725	0	1	0.48	0.50
Children	1 if has children, 0 else	2725	0	1	0.29	0.46
Area1	1 if metropolitan area, 0 else	2725	0	1	0.18	0.38
Area2	1 if Southern Finland, 0 else	2725	0	1	0.31	.046
Area3	1 if Western Finland, 0 else	2725	0	1	0.26	0.44
Area4	1 if Eastern Finland, 0 else	2725	0	1	0.13	0.33
Area5	1 if Northern Finland, 0 else	2725	0	1	0.11	0.32

3.5. Descriptive statistics

In terms of alcohol use, the data corresponds to average figures of the Finnish population. Of the full sample, abstainers make up 10% of the respondents, which is at the same level as the national percentage of abstainers (Statistics Finland, 2011). This percentage holds also when genders are investigated separately. Of the abstainers, however, exactly 50% are classified as former drinkers. Thus, it can be concluded that only 5 percent of the sample are life-time abstainers. Table 4 below presents the drinker status and percentage of each status in the sample data, sorted by genders.

Table 4: Drinker status by gender

Drinker status	Female	Male	Total
Never drunk	80 (5.6%)	56 (4.3%)	136 (5.0%)
Former drinker	63 (4.4%)	73 (5.6%)	136 (5.0%)
Current drinker	1285 (90.0%)	1168 (90.1%)	2453 (90.0%)
Total	1428	1297	2725

In earnings perspective, it is worthwhile to investigate whether the sample corresponds to national earnings averages. To do this, earnings are decomposed into deciles. Table 5 below presents the decomposition along with decile information from Statistics Finland's population-level data. It turns out that in the lower deciles the average earnings are much lower in the full sample, whereas in the upper deciles the earnings figures are close to the national statistics. However, if only those respondents that are actively working are taken into consideration, the lower decile averages converge towards the national statistics, as can be seen from the third column. On the other hand, the workers sample yields higher averages in the upper deciles, but the figures are still reasonably close to the national statistics. Overall, the earnings deciles of the sample provide a good comparison with national income distribution.

Table 5: Earnings deciles

Decile	Statistics Finland	Full sample	Workers sample
I	10148	818	8617
II	14287	5949	10084
III	17051	9954	13116
IV	19424	14073	17046
V	21820	18207	20563
VI	24084	22150	24320
VII	26674	25936	28301
VIII	29927	30256	33022
IX	34962	37313	40210
X	58719	60060	65678

In order to get further insight into the patterns of alcohol consumption and earnings, table 6 presents a comparison of drinking volumes and quantities in different earnings deciles. In addition, males and females are studied separately and average ages for each decile are included[9]. The measure for yearly consumption of pure alcohol in centiliters was adjusted so that individuals who consumed over 3000cl of alcohol were excluded from the sample. According to the survey design, the limit of heavy drinking is 24 drinks per week for men and 12 drinks per week for women. This is equal to 1856cl of pure alcohol for men and 923cl for women. Thus, the cutoff was done to minimize the effect of large tail values and altogether 44 individuals had consumption over this limit. A similar procedure was done to intoxication frequency, meaning that individuals who reported over 100 intoxications per year were excluded. This resulted in 46 individuals being excluded from the sample.

It can be seen from the table that there are major differences between genders at different earnings levels. Of course, the absolute drinking amounts are much smaller for women. For both genders, the peak in the frequency of intoxication happens at the second lowest earnings decile. However, the alcohol consumption in centiliters for men follows a quite steady pattern across the deciles, whereas for women there is a substantial increase in consumption in the highest earnings decile. Looking at the age dimension, it can be seen that the average age of respondents is much lower in the lowest

[9] The same earnings decile formation was applied for both men and women, but the results were similar if earnings deciles were calculated separately for the genders.

two earnings deciles, which is a logical result of short work experience. This may also explain part of the intoxication frequency in the second lowest decile, since it contains many young adults whose absolute alcohol consumption is not high but the consumption pattern is more about intense drinking as they have not "settled down" yet.

The consumption pattern of females is somewhat ambiguous. Probably the educational factors and labor market conditions of women with the highest income lead to similar consumption patterns as men have[10], which emphasizes the need of controlling education especially in the case of women. The male consumption pattern gives support for straightforward interpretation of alcohol-earnings relationship, meaning that endogeneity problems do not rise from reciprocity between income and alcohol consumption. However, another kind of endogeneity that originates from unobserved factors affecting both earnings and alcohol consumption is still a major concern in the statistical analysis.

Turning the approach to consumption patterns around, it is reasonable to investigate how mean earnings differ between different drinker types. In table 7, mean earnings are calculated for the five risk classes and for both genders separately. For men, there is a clear peak in earnings at risk group 3. For women, there are two peak points, in the third and in fifth risk groups. Thus, for women, the figures show similar kind of pattern between alcohol consumption and earnings as table 6 showed.

Looking at the tables below, certain patterns in the alcohol-earnings relationship seem to emerge. However, the age dimension has been treated only in terms of average age in table 6. As discussed earlier, researchers have found differences in age-earnings profiles of different drinker types. Hamilton and Hamilton (1997) discovered that heavy drinkers possess much flatter age-earnings profiles than other individuals in their sample. In order to study how earnings change over time, figure 6 presents a graph in which log earnings are plotted against age. The aggregate trend line in the graph exhibits a conventional shape of age-earnings profile. Yet, the graph only shows the average profile, regardless of drinking pattern.

[10] This is supported by comparing education level of women at the highest earnings decile. Alcohol consumption amount increases as earnings or education increases. A similar pattern for men does not exist. Furthermore, half of the women in the highest decile have university degree whereas only a quarter of men at that decile have a similar degree.

Table 6: Earnings deciles and alcohol consumption

Earnings decile	Female			Male		
	Average age	Alcohol consumption, cl/year	How often over 8 drinks	Average age	Alcohol consumption, cl/year	How often over 8 drinks
I	24.7	174	4.1	26.1	393	10.9
II	36.9	189	4.3	39.4	476	14.6
III	46.9	169	3.0	45.8	494	11.3
IV	49.4	172	2.7	48.1	436	12.5
V	46.1	186	2.7	47.2	493	8.3
VI	46.1	226	2.9	46.8	544	13.2
VII	45.0	173	2.3	45.5	557	12.5
VIII	46.8	180	1.6	45.0	526	10.7
IX	47.2	226	2.8	45.5	516	10.9
X	48.5	277	2.7	48.2	532	9.3
All	43.5	191.6	2.9	44.0	501	11.3

Table 7: Mean earnings by risk class

Risk class	Female	Male	Total
1	14307	16598	15939
2	19213	25167	21618
3	20630	31747	25485
4	17677	25129	22300
5	22600	21965	22218
Total	19185	26079	22465

To emphasize the importance of drinking pattern, figure 6 presents age-earnings profiles for each risk group. As can be seen from the graph, individuals in risk group 1, i.e. abstainers, face a significantly lower profile, but it should be emphasized that the size of the group is small and it takes into account also previous drinkers. More interestingly, risk groups 2 and 3 have the steepest

earnings rise and the flattest end part of the profile. Individuals in these groups are those ones that drink moderately and are probably the best ones in controlling their drinking habits and the negative consequences of drinking.

Another interesting fact is that individuals in group 4 have higher initial earnings level, but the growth in earnings is slower and decreases rapidly after reaching maximum level. There are probably many reasons for this kind of a profile, but most likely it is related to the background factors that cause a certain drinking pattern. It may indicate the latent personal attributes and risk behavior of individuals belonging to that group. On the other hand, it may be an indicator of alcohol use behavior similar to what French and Zarkin (1995) described. In their view, individuals consuming large amounts of alcohol do not acquire extensive education and start working early in life, thus having earnings growth at early working years. However, the detrimental effects of alcohol use cause the profile to be flatter and the drop in earnings at later years will also be larger. This is supported be the graph and data[11].

The profile of risk group 5 is in line with the theoretical assumptions. Group 5 consists of problems drinkers that can be found in all socioeconomic and education level groups. This explains the relatively steep rise in earnings, but as the graph shows, the maximum earnings level is achieved much earlier than in other groups and the decrease in earnings is also much faster.

Overall, the graph gives further insight into the effects of drinking patterns, but any clear conclusions cannot be immediately drawn from it. For instance, the risk group of an individual may change multiple times during life and this effect cannot be captured without observing the alcohol consumption for a longer time period. Such a change in drinking habits may results from having children or, conversely, from losing a job. Still, the risk group profiles are highly indicative and justified by the theoretical consideration. One major finding is that the slope of earnings decrease becomes greater as risk-level increases[12].

[11] A cross-tabulation of risk groups and education levels shows that in group 4 the share of individuals with only basic education is the largest. See Appendix B for further explanation.
[12] This holds if group 1(abstainers) is excluded. Reasons for the profile of this group may lie in the small amount of respondents belonging to the group, in the existence of former drinkers in this group or in a joint effect of these two reasons.

Figure 6: Age-earnings plot

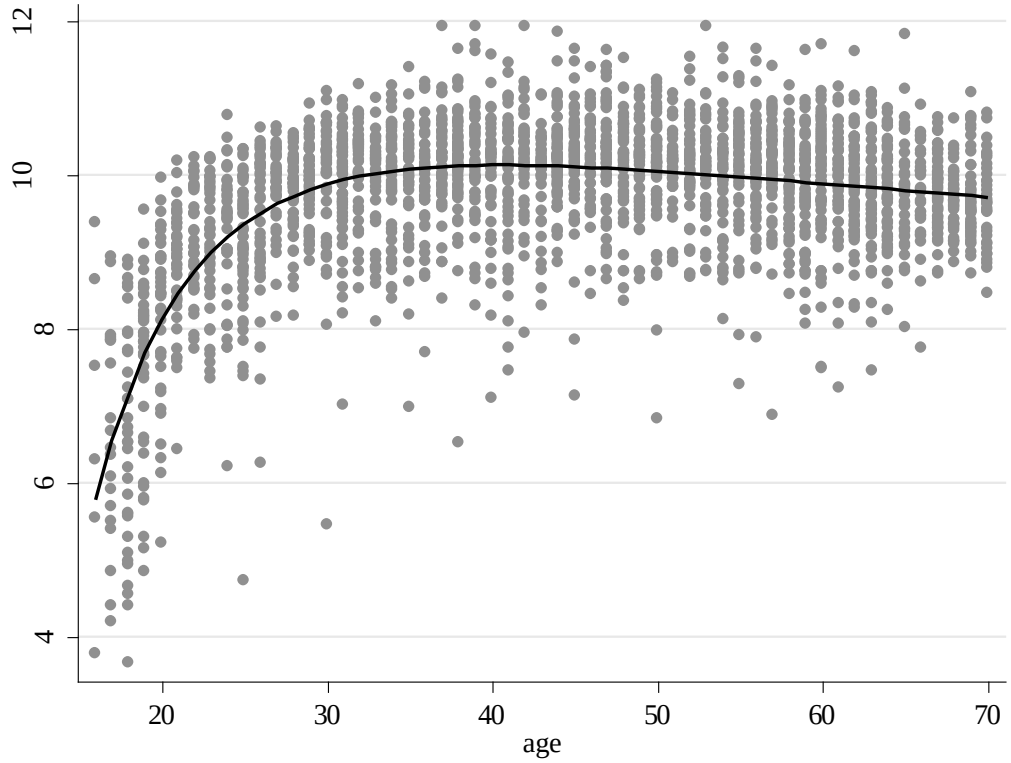
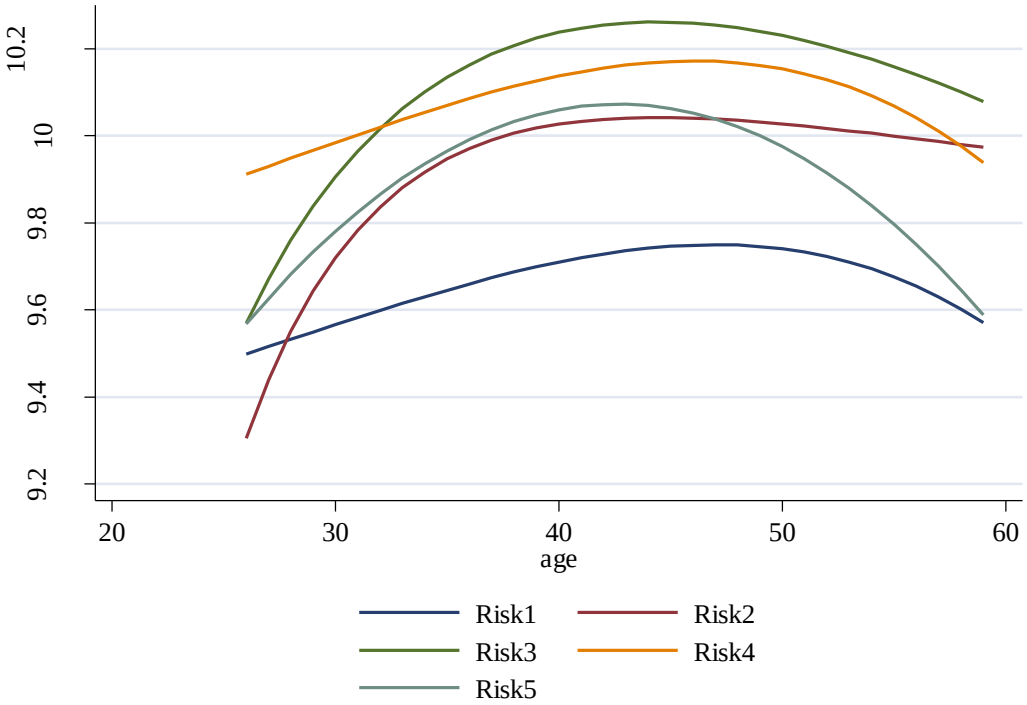


Figure 7: Risk groups and earnings profiles



4. Results

4.1. Basic OLS-specification

Equation 11 shows how the OLS-approach is applied in its initial form:

$$Logearn = \beta_0 + \beta_1 Drinker + \beta_2 Age + \beta_3 Age^2 + \beta_4 Female + \beta_5 Married + \beta_6 Highschool + \beta_7 Area1 + \varepsilon_t \quad (11)$$

In equation 11, the explanatory variables include the drinker dummy as the only alcohol use variable. Age and its square are included to capture the age-earnings effects shown in figures 5 and 6. Other variables are dummies for females, marital status, high school completion and people living in metropolitan area. These are included to capture some of the most obvious explanatory factors[13]. The reason for including the metropolitan area dummy lies in the population structure of Finland. Almost fifth of the Finnish population lives in the metropolitan area and it is the most diverse area in terms of cultural aspects that may influence alcohol consumption patterns. Furthermore, a large share of well-paid jobs is located in the metropolitan area. The results from this OLS regression are reported in table 8 below. Regression was executed using robust standard errors since heteroskedasticity test led to rejection of H_0 hypothesis[14].

As it turns out, all the variables are significant at 1 percent level. Furthermore, the coefficients reflect the expected behavior on basis of the theoretical consideration and descriptive analysis. Focusing on the drinker dummy, it can be seen that drinkers earn substantially more than abstainers and the earnings premium is almost 25 percent. However, this estimation was done for the whole sample with only few control variables. Regarding the control variables, the difference between female and male earnings is evident as well as the earnings gains of being married, high school graduate or a person living in the metropolitan area. The age variable reflects a typical age-earnings profile with the peak point in earnings at the age of 52[15].

[13] Getting married does not automatically alter one's alcohol consumption pattern but a likely indicator of some stability in life in comparison to single individuals of the same age.
[14] H_0: constant variance. Appendix B shows the results of heteroskedasticity test.
[15] This was done with a simple calculation of taking the first order derivative and determining the point where the derivate gets value 0.

Table 8: OLS with alcohol user dummy

Variable	Coefficient (robust standard errors in parentheses)
Constant	4.789*** (0.211)
Drinker	0.248*** (0.053)
Age	0.209*** (0.009)
Age^2	-0.002*** (0.000)
Female	-0.304*** (0.031)
Married	0.162*** (0.029)
Highschool	0.300*** (0.034)
Area1	0.267*** (0.039)
Observations	2601
R^2	0.404

*** Significant at 0.01 level

4.2. Gender-specific OLS-models

Descriptive statistics and the results from the initial OLS estimation implicate that there are major differences both in female alcohol consumption and earnings. Thus, in the gender-specific OLS-specification both genders have their own earnings equation in order to see if the returns on drinking differ considerably. Also a broader set of covariates needs to be included for controlling all the theorized effects in the alcohol-earnings relationship.

Multiple gender-specific models were tested and the results are presented in table 9 for men and in table 10 for women. Model 1 included only continuous alcohol use variables and the number of intoxications. Looking at table 9, the results obtained by using model 1 support the existence of an inverse u-shaped non-linear relationship between alcohol use and earnings for men. Furthermore, the number of intoxication is associated with reduced earnings and all coefficients are significant. However, the correlation coefficient shows that the model fit is very low, which of course is a logical outcome since only three variables are included in the model. For women, model 1 behaves in a similar manner and the coefficients have the same direction and level of significance.

In the second model, the drinking variables are kept the same while demographic control variables are added to the model, together with one educational variable, the dummy for high school completion. The result for men shows that all drinking variables express similar values as in model 1, only the degree of intoxication variable is slightly smaller. Other control variables behave in an expected manner and the correlation coefficient yields a reasonably large value. For women, the coefficients for continuous alcohol use variables have similar values as in model 1, but the intoxication variable yields a positive coefficient value which is not statistically significant. An interesting fact about the control variable coefficients is that marriage does not have statistical significance.

Turning into the third model, a broader set of covariates is included into the analysis. These include alcohol use variables that measure for starting age of alcohol use and the impact of being an ex-drinker. Also more controls for demographics, education and health are added. Looking at table 9, it can be seen that the coefficients for alcohol use again express similar values as in the previous models. Adding control variables only slightly reduces the significance of alcohol use variables and the degree of the coefficient values. However, adding more alcohol use variables provides further insights into the alcohol-earnings puzzle: the importance of controlling for ex-drinkers is evident, as the coefficient shows high negative payoff and it is significant at 1% level. On the other hand, interpretation of the ex-drinker variable is troublesome at least in graphical sense, since the variable only affects those who are not consuming any alcohol at the moment. Intoxication starting age in turn appears to have no effect on earnings. The third model for women behaves in a similar manner. The coefficient values for alcohol use and the significance of the variables are again of same degree as in model 2. Being an ex-drinker is associated with reduced earnings while the starting age of getting drunk does not affect earnings.

Control variable coefficients in model 3 for both men and women express anticipated results. Using metropolitan area as a reference group, the other geographic areas are associated with significantly lower earnings. Being married or having children leads to a smaller positive earnings effect for women than for men and the coefficients are not statistically significant. This probably indicates that men change their drinking habits more as a result of getting married or having children, which is a consistent finding also from the earlier studies. Including a university dummy into the model proves to have high explanatory power, but interestingly the high school dummy for men is no longer a significant determinant of earnings equation, while for women it still has a major impact on earnings. This is most likely an indication of differences between men and women in their labor

market positions. Looking at the other control variables, it can be seen that for both men and women the health measures are not significant even at 10% level[16]. Regarding the health measures, smoking was excluded from the analysis, since alcohol use is most likely complementary to tobacco use, whereas the opposite may not be true. This is supported by the study made by Koksal and Wohlgenant (2011). On the other hand, tobacco use could indicate an individual's risk preference, but a more concrete risk preference measure is the indicator of drug use, which appears to have no explanatory power. Interestingly though, the coefficient value for drug use is positive for women.

In the fourth model, the first risk class (abstainers) is excluded due to the small size and problems with the sample selection as mentioned before. Furthermore, the frequency of drinking over 8 drinks is dropped since the risk classes take intensity of drinking into account. Ex-drinker variable is omitted due to collinearity. The results for men show that belonging to the third risk group has the greatest positive payoff. Thus, the fourth model also supports the hypothesis that earnings are maximized at a certain drinking level. A new finding is that the inverse u-shape functions also with intensity, since those who get drunk often face lower earnings level than those who get drunk infrequently. In the second model, getting drunk was associated with a negative linear impact on the earnings. All demographic, health and human capital variables in the third model for men have similar coefficient values as in model 2.

For women, the fourth model exhibits slightly different results. The positive earnings effect is greatest for women belonging to risk group 4. Comparing this result with the earlier models, it seems that getting drunk does not have as negative effect for women as it has for men, since also in models 2 and 3 the coefficient value for intoxication frequency was positive for women. The reason for this might lie in more detrimental health and human capital effects of men getting drunk. For instance, men likely face drunken accidents more often and drink far greater amounts at one occasion, leading to negative health and productivity effects. On the other hand, the health determinants do not have significant explanatory power in any of the models. It may indicate about the irrelevance of health-related effects in the alcohol-earnings relationship, giving further evidence in favor of the social aspects of drinking. Yet, it was stated earlier that health effects cumulate over time which requires an age-group specific investigation. Furthermore, beverage-specific aspects as well as social dimension require a closer investigation, which is provided later in this chapter.

[16] Alternative specifications were tested with controls for different health effects, but they did not alter the results of the alcohol-earnings equation. Health effects on earnings were also tested by controlling for drinker risk class, but the health variables did not have explanatory power, indicating that health effects do not explain differences between different drinker types.

Overall, the four models exhibit similar effects for both males and females. For instance, the inverse u-shaped relationship is evident for both genders[17]. On the other hand, the risk class model shows that there is a turning point in the drinking's effects also for women, but it takes place on relatively higher level. To estimate the turning point for both genders, I calculated the maximum values of the earnings-alcohol use polynomial functions by determining when first order derivative equals zero. This yielded the following turning points (see table 11). For males, the wage premium peaks at 2.6 drinks per day, which is close to the figures that French and Zarkin (1995) found using similar functional form. For females, the peak happens at 1.2 to 1.4 drinks per day. On the basis of model 3, figures 8 and 9 show the estimated u-shaped curves of alcohol-earnings relationship for men and women respectively. These figures show the pure effect of alcohol use for an average man and woman after the effect of all other covariates is controlled.

At first, the peak in wage premium at 2.6 drinks per day sounds implausibly large. However, the definition of a heavy drinking male is 40g of pure alcohol per day, which corresponds to 3.2 drinks per day. For women the definition is 20g of pure alcohol per day, which is the equivalent of 1.6 drinks per day. Thus, both peak points are at a significantly lower level than the limit of heavy drinking. Furthermore, the national recommendations for moderate drinking amounts are 0-24 drinks per week for men and 0-16 drinks for women (National Institute for Health and Welfare, 2008) and thus the results support moderate drinking, both in terms of quantity and intensity.

The results for drinking amounts require some critical scrutiny. The distribution of alcohol consumption in the data shows that for men, majority of the respondents consume less than 1500cl per year and majority of women consume less than 800cl per year (see Appendix B). It may be that the most earning individuals are also the ones drinking the most, as the sample takes into account only those who are receiving income, but the amounts still cause some ambiguity. Thus, the absolute alcohol use amounts should be approached with some caution. The major contribution of the gender specific analysis is actually the shape of the relationship combined with the notion that the frequency of intoxications is associated with a negative earnings effect especially for men.

[17] Also a specification that included cubic alcohol use term was tested but the explanatory power of the model was substantially lower.

Table 9: OLS specifications for males

	Coefficient (robust standard errors in parentheses)			
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	9.735*** (0.047)	5.053*** (0.300)	5.734*** (0.350)	5.413*** (0.337)
Alcohol use				
Alc_cl	$9.26x10^{-4}$*** (0.000)	$5.44x10^{-4}$*** (0.000)	$4.08x10^{-4}$*** (0.000)	
Alc_cl^2	$-3.13x10^{-7}$*** (0.001)	$-1.81x10^{-7}$*** (0.000)	$-1.32x10^{-7}$** (0.000)	
Risk2				0.205*** (0.095)
Risk3				0.366*** (0.091)
Risk4				0.335*** (0.096)
Risk5				0.075 (0.137)
Freq8drinks	-0.011*** (0.002)	$-4.11x10^{-3}$*** (0.001)	$-3.52x10^{-3}$** (0.002)	
Startdrunk			0.006 (0.005)	
Exdrinker			-0.278** (0.109)	
Demographics				
Age		0.208*** (0.013)	0.184*** (0.015)	0.189*** (0.015)
Age^2		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Married		0.247*** (0.044)	0.112*** (0.048)	0.148*** (0.047)
Children			0.220*** (0.054)	0.205*** (0.053)
Area1		0.284*** (0.059)		
Area2			-0.212*** (0.067)	-0.165** (0.069)
Area3			-0.291*** (0.070)	-0.259*** (0.072)
Area4			-0.377*** (0.079)	-0.418*** (0.084)
Area5			-0.310*** (0.107)	-0.332*** (0.111)
Education and health				
Highschool		0.255*** (0.049)	0.055 (0.057)	0.080 (0.060)
University			0.365*** (0.058)	0.376** (0.062)
Illness			-0.043 (0.046)	-0.049 (0.048)
Obese			-0.111 (0.072)	-0.111 (0.071)
Drug			-0.032 (0.078)	-0.048 (0.077)
Observations	1185	1185	998	1117
R^2	0.045	0.391	0.42	0.43

*** Significant at 0.01 level

**Significant at 0.05 level

*Significant at 0.10 level

Table 10: OLS specifications for females

Variable	Model 1	Coefficient (robust standard errors in parentheses) Model 2	Model 3	Model 4
Constant	9.541*** (0.034)	4.560*** (0.300)	4.740*** (0.388)	4.718*** (0.309)
Alcohol use				
Alc_cl	$7.42x10^{-4}$*** (0.000)	$5.79x10^{-4}$*** (0.000)	$4.12x10^{-4}$*** (0.000)	
Alc_cl^2	$-2.85x10^{-7}$*** (0.000)	$-3.48x10^{-7}$*** (0.000)	$-2.89x10^{-7}$** (0.000)	
Risk2				0.113 (0.075)
Risk3				0.234*** (0.080)
Risk4				0.345*** (0.095)
Risk5				0.142 (0.151)
Freq8drinks	-0.012*** (0.004)	$1.45x10^{-3}$ (0.003)	$4.22x10^{-3}$ (0.003)	
Startdrunk			0.007 (0.005)	
Exdrinker			-0.189* (0.108)	
Demographics				
Age		0.212*** (0.013)	0.214*** (0.016)	0.205*** (0.013)
Age^2		-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Married		0.049 (0.038)	0.052 (0.043)	0.052 (0.041)
Children			0.022 (0.050)	0.096* (0.049)
Area1		0.257*** (0.049)		
Area2			-0.213*** (0.067)	-0.214*** (0.061)
Area3			-0.163** (0.064)	-0.153** (0.063)
Area4			-0.371*** (0.078)	-0.354*** (0.075)
Area5			-0.246*** (0.073)	-0.250*** (0.082)
Education and health				
Highschool		0.336*** (0.044)	0.154*** (0.055)	0.182*** (0.054)
University			0.331*** (0.054)	0.320*** (0.056)
Illness			-0.038 (0.046)	-0.045 (0.043)
Obese			-0.060 (0.061)	-0.058 (0.058)
Drug			0.113 (0.076)	0.122 (0.309)
Observations	1348	1348	1082	1261
R^2	0.02	0.43	0.46	0.46

*** Significant at 0.01 level

**Significant at 0.05 level

*Significant at 0.10 level

Table 11: Turning points in drinking premium

	Male		Female	
Amount	**Model 2**	**Model 3**	**Model 2**	**Model 3**
cl / year	1503	1545	832	713
cl / day	4.12	4.23	2.28	1.95
Drinks / day	2.6	2.6	1.4	1.2

Figure 8: Alcohol-earnings curve for men

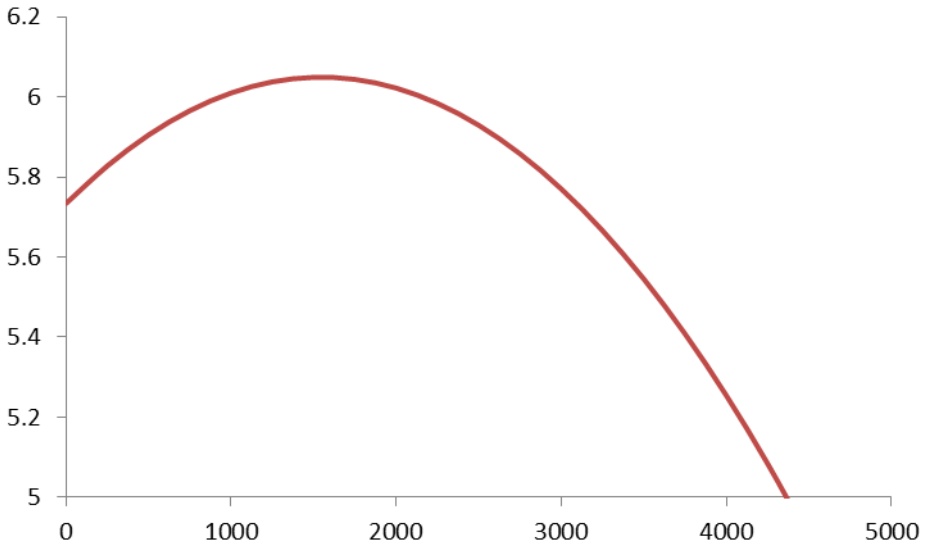
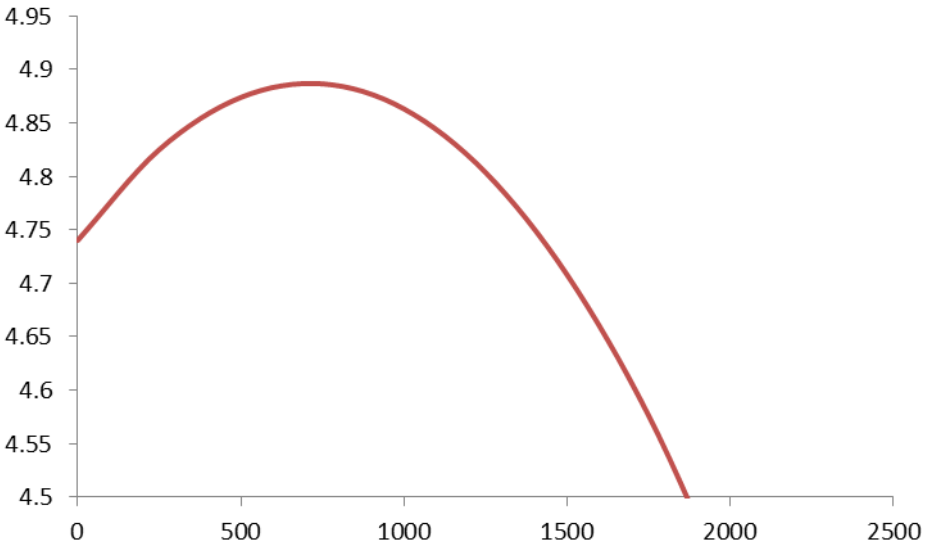


Figure 9: Alcohol-earnings curve for women



4.3. Alternative hypotheses

The analysis in the former section showed that the drinking behavior can be modeled in a similar manner for both men and women, yet there are differences in the drinking amounts and intensity-related behavior. The differences in other observable characteristics were evident but they are not central in the analysis. In the following estimations the goal was to test alternative hypotheses and the gender differences are treated with the female dummy. This was done also to keep the sample sizes on a reasonable level.

4.3.1. Socioeconomic status and drinking

Socioeconomic status (SES) is a good overall measure of an individual's position at a society. Obviously, socioeconomic status is closely related to individual earnings, but captures also a larger share of personal attributes that affect earnings. To study differences in alcohol-earnings relationship, individuals were classified in three groups on basis of their socioeconomic status. SES1 contains all individuals employed in managerial level positions. SES2 contains all individuals employed in clerical work positions. SES3 contains all blue-collar employees of the sample. Separate regressions were run for each socioeconomic class to control for differences in returns to drinking. Table 12 presents the results from this analysis.

The results show clear differences between socioeconomic groups. For instance, the constant shows that the basic earnings level is actually much higher for the lower socioeconomic groups, which arguably results from the importance of schooling and other human capital factors at the highest socioeconomic group. The groups also differ greatly with regards to alcohol use related variables. First, the turning point in inverse u-shaped relationship differs between socioeconomic groups. Looking at table 13, it can be seen that individuals who belong to SES1 maximize their earnings at 2.2 drinks per day, whereas the "optimal" amount for individuals in the other two groups is around 1.8 – 1.9 drinks per day. There are many possible explanations for this. The difference in turning points may indicate that individuals in the highest decile drink more, implying positive income elasticity. On the other hand, it may also indicate that drinking less than an average person in one's own socioeconomic group is especially beneficial in lower groups. Yet another explanation may be that the beneficial social and labor market effects of drinking cause the turning point to be at higher amount for SES1. Most likely, the networking effects of drinking are not as prevalent in the lower

socioeconomic groups as in the highest where promotion opportunities are highly related to social activity.

Considering the possible causes for the difference in the peak points, appendix B shows how drinking amounts differ between socioeconomic groups and sexes. For drinking to have any kind of an effect on earnings that can be compared across socioeconomic groups, the consumption should be fairly constant or at least significantly differ from zero. Thus, an arbitrary threshold level of 3 drinks per week was assigned for the tabulation, revealing that individuals at the lowest socioeconomic group consume significantly more alcohol than individuals in the highest group. Therefore, it is plausible to think that drinking less than others in one's own group is positively related to earnings level especially in the lower socioeconomic groups. This could be due to health effects of drinking. On the contrary, in the highest socioeconomic group drinking more that an average person of the group is beneficial due to the networking effects of drinking. These findings are supported also by the frequency of intoxications variable, as it is significant only in the lowest socioeconomic group.

Other covariates in the analysis exhibited similar behavior, at least in their signs. Interestingly, being female leads to larger negative payoff the lower the socioeconomic group. Having children is significant only in the lowest group, whereas marriage did not prove to have significance. Another interesting finding is the large and negative earnings effect of obesity in the lowest socioeconomic group. Obesity is probably closely intertwined with excessive drinking, which gives further evidence of drinking's negative effects in the lowest socioeconomic group. On the other hand, individuals in this group are often employed in manual labor positions in which being obese may lead to reduced productivity.

Table 12: OLS with socioeconomic status

	Coefficient (robust standard errors in parentheses)		
Variable	**SES1**	**SES2**	**SES3**
Constant	6.499*** (0.607)	7.503*** (0.339)	7.200*** (0.364)
Alcohol use			
Alc_cl	$4.25x10^{-4}$*** (0.000)	$4.08x10^{-4}$*** (0.000)	$3.04x10^{-4}$** (0.000)
Alc_cl^2	$-1.69x10^{-7}$*** (0.000)	$-1.88x10^{-7}$*** (0.000)	$-1.42x10^{-7}$* (0.000)
Freq8drinks	$-2.38x10^{-3}$ (0.002)	$-2.46x10^{-3}$ (0.002)	$-5.40x10^{-3}$**(0.002)
Startdrunk	0.006 (0.007)	0.007 (0.005)	-0.002 (0.005)
Exdrinker	-0.410** (0.190)	-0.311** (0.128)	-0.353*** (0.094)
Demographics			
Age	0.155*** (0.025)	0.117*** (0.014)	0.137*** (0.016)
Age^2	-0.002*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Female	-0.233*** (0.056)	-0.297*** (0.053)	-0.460*** (0.055)
Married	0.056 (0.061)	0.024 (0.042)	0.057 (0.047)
Children	0.058 (0.066)	0.022 (0.051)	0.144** (0.061)
Area2	-0.172*** (0.064)	-0.097* (0.054)	-0.252*** (0.079)
Area3	-0.158** (0.075)	-0.164*** (0.056)	-0.234*** (0.076)
Area4	-0.326*** (0.115)	-0.223*** (0.071)	-0.446*** (0.097)
Area5	-0.373*** (0.112)	-0.215*** (0.091)	-0.236** (0.106)
Education and health			
Highschool	0.192*** (0.072)	-0.005 (0.051)	-0.146 (0.105)
University	0.147*** (0.057)	0.223*** (0.060)	0.156 (0.169)
Illness	-0.056 (0.079)	-0.049 (0.042)	-0.034 (0.051)
Obese	0.103 (0.079)	-0.058 (0.064)	-0.216*** (0.076)
Drug	0.000 (0.088)	-0.081 (0.080)	-0.051 (0.109)
Observations	504	637	565
R^2	0.34	0.32	0.38

*** Significant at 0.01 level

**Significant at 0.05 level

*Significant at 0.10 level

Table 13: Turning points and SES

	Socioeconomic group		
Amount	**SES1**	**SES2**	**SES3**
cl / year	1257	1085	1069
cl / day	3.4	3.0	2.9
Drinks / day	2.2	1.9	1.8

4.3.2. Beverage-specific drinking

Following the approach of Ziebarth and Grabka (2008), beverage-specific effects of drinking were studied by dividing individuals into abstainers, beverage-specific drinkers and drinkers who do not have tendency to drink a particular beverage. In this manner, dummies for abstainers, wine drinkers, beer drinkers and liquor drinkers replaced other drinking variables in the OLS-function. As stated before, the rest of the individuals were categorized as non-specific drinkers who may consume many different beverages or change their drinking pattern at different times or occasions. Health variables were dropped from the analysis as they have proven not to have statistical significance in most of the cases. Looking at the first column of table 14, the results show clear differences between beverage-specific drinkers. Being a wine drinker is associated with a greater earnings premium than any other pattern. The coefficient for liquor drinking is insignificant, which probably indicates that being a liquor drinker is only a proxy for background factors that may also affect labor market outcomes[18].

It is plausible to think that the wine gain only indicates that better earning individuals tend to drink wine. A cross-tabulation reveals that the amount of both beer drinkers and wine drinkers increases towards the upper deciles, but for the most part of the earnings distribution, the percentages of each drinker type are relatively evenly distributed. Furthermore, the amount of beer drinkers is much greater than the amount of wine or liquor drinkers in each category (see Appendix B). Thus, it is difficult to draw direct conclusions from the cross-tabulation, but on the basis of regression analysis, it seems that drinking wine is associated with 5% higher earnings than drinking beer, which can be the result of earnings' effect on drinking habits or wine's beneficial health and psychological effects. On the other hand, the effect of regional, social and demographic factors may

[18] The regression was also conducted separately for those who earn less than median income and for those who earn more than median income, showing the same pattern. See appendix B for regression results.

be mediated through wine drinking into the earnings equation, meaning that wine drinker is a proxy for certain individual characteristics that are associated with higher earnings across the income distribution.

4.3.3. Sociability and drinking

The last hypothesis looks at how sociability related factors affect the estimation, especially those factors that can be directly linked to drinking. Drinking variables were kept the same as in the gender and socioeconomic group models and the coefficients exhibit similar behavior. Looking at the sociability factors, some interesting facts emerge from the analysis. First, however, it appears that some variables lack significance. For instance, individuals who have felt lonely fare worse than others, but the coefficient value is insignificant. The same is true for the two variables that measure opinions on drinking (*No-joy* and *Drunk_fun*).

The two mental state variables prove to have explanatory power. The positive coefficient on *Mental1* may indicate that individuals who think that alcohol use has a positive effect on their mental health have also better control over their alcohol use and might have greater propensity to drink in a social manner. Having control over own alcohol use is supported by a look at the average drinking volumes of individuals that answered yes only to one of these two questions. The average drinking amount of those who answered yes on *Mental1* is 40% lower than the average drinking amount of those who answered yes in *Mental2*. As can be seen from table 14, the coefficient value for *Mental2* is highly negative, meaning that using alcohol to alleviate depression or anxiety is associated with lower earnings, which probably tells about other mental problems. On the other hand, the framing of the questions comes to play here: if the individuals had been asked about stress alleviation instead of depression alleviation, the responses would probably have been much different.

The last sociability variable, *Restaurant*, has a positive coefficient value and it is significant at 5% level. If considered as an indicator of social effects in alcohol-earnings relationship, the variable shows that drinking in social settings is associated with large earnings premium. Of course, the causal link needs to be emphasized as people with higher earnings are likely to visit restaurants and bars more often. Yet, looking at restaurant drinking frequency and earnings for men (see Appendix B), it can be seen that those who drink infrequently in restaurants have the highest earnings. For

women, the same pattern holds in general. Thus, the distribution of earnings shows that being a restaurant drinker is not determined by the earnings level, whereas the earnings level may be partly determined by the propensity to drink in restaurant. In this sense, the analysis at least partly supports the social hypothesis.

Table 14: OLS with beverages and sociability

	Coefficient (robust standard errors in parentheses)	
Variable	**Beverage model**	**Sociability model**
Constant	5.524*** (0.217)	5.431*** (0.256)
Alcohol use		
Beerdrinker	0.107*** (0.035)	
Winedrinker	0.155*** (0.037)	
Liquordrinker	0.010 (0.043)	
Abstainer	-0.178*** (0.055)	
Alc_cl		$4.22x10^{-4}$*** (0.000)
Alc_cl^2		$-1.67x10^{-7}$*** (0.000)
Freq8drinks		$-2.19x10^{-3}$ (0.001)
Startdrunk		0.004 (0.003)
Exdrinker		-0.220*** (0.076)
Demographics		
Age	0.191*** (0.001)	0.196*** (0.010)
Age^2	-0.002*** (0.000)	-0.002*** (0.000)
Married	0.010*** (0.030)	0.061* (0.032)
Children	0.162*** (0.035)	0.133*** (0.037)
Female	-0.296*** (0.033)	-0.301*** (0.034)
Area2	-0.199*** (0.043)	-0.220*** (0.043)
Area3	-0.225*** (0.045)	-0.248*** (0.045)
Area4	-0.389*** (0.053)	-0.374*** (0.053)
Area5	-0.282*** (0.061)	-0.286*** (0.060)
Education and sociability		
Highschool	0.121*** (0.039)	0.117*** (0.039)
University	0.358*** (0.040)	0.354*** (0.039)
Lonely		-0.058 (0.038)
Mental1		0.056* (0.033)
Mental2		-0.176*** (0.048)
Restaurant		0.136** (0.060)
No_joy		0.050 (0.033)
Drunk_fun		-0.026 (0.033)
Observations	2601	2195
R^2	0.43	0.44

4.4. Limitations and problems

There are basically two kinds of problems that cause bias in this study and limit the robustness of the results: data problems and hypothesis problems. Data-specific problems are those that arise from the way the sample is formed, whereas hypothesis-specific problems concern the relationship between alcohol use and earnings. Obviously, these are common problems not only for alcohol-earnings studies but also for a variety of micro-level econometrical studies.

4.4.1. Data

The dataset contains many precisely defined variables and as the analysis showed, the results are highly consistent and statistically significant. In addition, the sample proved to be a good representation of the Finnish population. The raw sample contains 2750 individuals and this figure is in line with the earlier studies. Yet, the altered hypotheses cannot be studied with controlling for a subset of the sample, such as females. In this sense, the number of respondents should be higher to be able to conduct more elaborate partial analyses. This is especially the case with issues such as gender and age groups. On the other hand, some authors have stated that it is enough to treat gender differences with a dummy variable.

Another major issue concerns the cross-sectional nature of the data. This prohibits making a full-scale analysis on the basis of figure 1 as the consumption of past periods cannot be modeled. Furthermore, cross-sectional data does not allow for the study of individual fixed effects. According to Tekin (2004), the use of fixed effect helps to avoid potential bias arising from unobserved individual factors. Usually, studies that used longitudinal data have found smaller earnings gains for drinking compared to abstainers. On the other hand, longitudinal panel data makes the models more complicated and as Tekin (2004) acknowledges, the use of fixed effects may actually enlarge the bias caused by measurement errors.

The data in this study is basically a set of two different variable groups. Register based variables form the first group and they are the only variables that are not based on face-to-face interview. The information on these variables is reliable as it is obtained from the national registers. However, the problem with register based variables is the lack of some important labor market measures, such as

tenure and hours worked. Lack of these measures may cause a bias in the models, but on the other hand, the regression estimates are consistent also without tenure measures and hourly wage rate.

Problems with the other group of variables are related to question framing and interview method. In discussing the results from the sociability model, it was stated that the framing of depression alleviation question may significantly affect the way respondents answer. A more complicated issue is the face-to-face interviewing method: when questions about substance use are asked in this sort of a situation, the responses likely suffer from willful underreporting and faulty memory. Many authors have raised the issue of underreporting and Tekin (2004) stated that small coefficients in his fixed effect model may be due to reporting errors. Furthermore, Cook and Moore (2000) noted that a comparison between self-reported drinking and sales data suggest that surveys capture only 40 – 60% of real alcohol consumption. Yet, the absolute drinking amounts are not crucial for the analysis, assuming that all individuals included in the sample underreport their drinking in a somewhat similar manner. It may however be true that alcohol dependent individuals systematically underestimate and deny their drinking habits, but their share in the sample is relatively small and should not cause major problems.

4.4.2. Endogeneity and unobserved heterogeneity

Endogeneity is a major problem in a variety of econometric analyses, especially in the case of cross-sectional data. According to Dave and Kaestner (2002), the endogeneity problems can be classified in two categories in alcohol-earnings studies: statistical and structural endogeneity. The problem with structural endogeneity refers to the reciprocal nature of alcohol use and earnings. MacDonald and Shields (2001) presented a formulation that shows how structural endogeneity biases an OLS estimation of an earnings equation. Following their approach, an earnings equation is first presented in a simplified manner:

$$W_i = \beta_1 A_i + \beta_2 X_i + \varepsilon_i \quad (12)$$

where A_i is the measure of alcohol use, X_i is the vector of other covariates and ε_i is a normally distributed error term. Assuming that alcohol consumption responsive to earnings, there is reciprocal causality and the effect of earnings on alcohol consumption can be presented in the following way:

$$A_i = \gamma_1 W_i + \gamma_2 X_i + \mu_i \quad (13)$$

where the terms are defined as before and now μ_i is the normally distributed error term. If the simultaneity of A_i and W_i is ignored, the estimated coefficient value for β_1 is biased. If the error terms are uncorrelated, the relationship between true measured value of β_1 and OLS-estimate β_1^{OLS} is following:

$$\beta_1^{OLS} = \beta_1 + \frac{1}{1 - \gamma_1 \beta_1} \frac{\sigma_\mu^2}{\sigma_A^2} \quad (14)$$

In this expression, σ_μ^2 is the variance of μ_i and σ_A^2 is the variance of A_i. If the coefficient value for wage in equation 13 and the coefficient value for alcohol use in equation 12 are positive, the OLS-estimation overstates the true impact of alcohol use on earnings.

Statistical endogeneity refers to personal traits and other factors that affect both labor market outcomes and alcohol use, which is the case depicted in figures 4 and 5. This type of endogeneity has also been called unobserved heterogeneity in the literature. Usually statistical endogeneity has been dealt with using instrumental variables models, other kinds of structural models or longitudinal data. Using an instrumental variable approach, there should be a covariate Z_i which is correlated with the drinking variable, but not with the error ε_i in equation 12. This way, the coefficient for alcohol use will be a consistent estimation and has the following quality:

$$plim\ \beta_1^{IV} = \beta_1 + \frac{cov(Z_i, \varepsilon_i)}{cov(Z_i, A_i)} = \beta_1 \quad (15)$$

However, the use of instrumental variable estimation technique, such as 2SLS, places additional requirements for the data. For instance, the fore mentioned quality requirement for instrumental variables makes it impossible to use this kind of a model in this study. For it to be possible, the data would need to have information on alcohol prices and taxes or other measures that are not correlated with earnings. Then again, there are statistical problems also with IV-methods as stated earlier.

5. Discussion

The analysis has shown that there are certain patterns in the alcohol-earnings relationship which prevail in the analysis. For instance, the inverse u-shaped dependence was evident in all models where quadratic term was included. Binge drinking was shown to have a clear negative effect on earnings both with the frequency of having 8 drinks and with the risk class dummies. The results are largely in line with the former studies, although the multidimensional approach of this study provided new findings, too.

Focusing in the theoretical part, the framework depicted in figure 5 gives a holistic look at the causes and effects of an individual's current drinking and it is also an effective tool for analytical investigation of the relationship. First, alcohol prices and taxes are central concepts in alcohol-earnings studies, yet their inclusion in the analysis is often troublesome. In a country such as Finland, where alcohol policy is centrally controlled for the whole country, the data would need to cover many years so that the changes in taxes and prices could be studied. Thus, a study with observations from multiple periods of Finnish Drinking Habits Survey could be a natural next step for the alcohol-earnings study in Finland. Yet, the survey is not collected as a panel that focuses each time on the same individual respondents, which means that the study method would need to be a pseudo panel. This way, the aggregated differences in the effects of alcohol use could be studied on a study year basis. Especially interesting would be to see whether the large tax increases between 2000 and 2008 explain the patterns in alcohol-earnings relationship.

To further emphasize the price aspect, it was stated earlier that alcohol demand is relatively inelastic in respect to price, especially for those who drink the most. Many authors have contemplated on the reciprocal effect of earnings on drinking, but no clear conclusions have been drawn. Surely, alcohol is relatively expensive consumer good and at least in the lowest earnings deciles extra income means possibility to buy more alcohol. The same may apply for young people and student, but to look at the big picture, the descriptive statistics showed that the drinking amounts do not change dramatically over the range of ten earnings deciles. Rather, a rise in earnings may shift the consumption towards more expensive alcohol beverages and the place of consumption may be a restaurant or a bar more often than home. If one looks at the relationship solely on economic terms, the price of alcohol should be included, as it may take the earnings gain away completely. For instance, the male peak in earnings was shown to take place at 2.6 drinks per day, which is the

equivalent of 950 drinks per year. Assuming constant beer buying process from a local store and constant price of one euro per beer bottle, a quick calculation means that an economic utility maximizing individual would spend 950 euros per year on alcohol. This amount is close to 80 euros on monthly basis, meaning that monthly salary would need to be at least 80 euros larger than the salary of an identical abstainer to really "profit from drinking".

In health perspective, there are a couple of interesting facts that came up in the theoretical part of the thesis. Both mortality and the risk of heart diseases have been shown to follow u-shaped path in relation to drinking and the minimum point coincidences with the earnings maximum point at alcohol-earnings studies, a finding that applies to this thesis, too. Curiously, the same relationship has been found also on the mental health side, as Skogen et al. (2009) showed that the relationship between anxiety level and drinking follows a similar u-shape and the level of anxiety minimizes at the same amount as earnings maximize. In this regard, it seems that drinking partly explains changes in health and stress level, which in turn may have an effect on earnings. On the other hand, the sociability model showed that those who have used alcohol to cope with anxiety or depression have lower earnings. Thus, moderate alcohol use may help to reduce the levels of anxiety and depression, while using alcohol directly to reduce these levels has a negative effect on earnings.

However, the empirical part did not find almost any health effects that are important determinants of earnings, whereas the alcohol use variables were constantly significant at 1% or at least 5% levels. Furthermore, the health variables did not affect the coefficient values of alcohol use variables. One single health factor that was significant was the dummy for obese individuals in the regression of lowest socioeconomic group. As already hypothesized, it may indicate that being obese reduces productivity especially in the case of manual labor individuals. Furthermore, obesity is tightly linked with heavy drinking and in this sense, the regression result may exhibit equation 7 in action. In equation 7, changes in earnings due to changes in alcohol use were decomposed into direct changes and changes mediated through changes in health status.

The effects of drinking on human capital are somewhat ambiguous. It is obvious that schooling is a crucial covariate in the analysis and the coefficient values of high school and university graduation exhibited anticipated values. On the other hand, the wage gain found in the beverage-specific model may indicate an intertwining relationship between schooling and drinking habits, as wine drinking arguably is more common to highly educated individuals. Yet, the wine gain cannot be explained

solely in terms of schooling, but the stock of human capital is likely correlated with the beverage selection.

The most interesting human capital related finding was made in the descriptive analysis. The age-earnings profiles turned out to be close to the hypothesized profiles shown in the theoretical part. The shape of the curves in figure 6 is highly indicative of age-related factors of alcohol use. The more an individual drinks, the flatter the profile, the shorter the rise period in earnings level and the steeper the decline in earnings at older age. Obviously, the curves do not take into account changes in drinking habits throughout an individual's life. Thus, the age-dimension should be taken further in future studies and the best result would be achieved with panel data that covers over 10 years to really capture changes in drinking habits and their lagged effect on earnings.

Arguably, the sociability aspect is the most crucial issue in studying the alcohol-earnings relationship. Figure 4 showed how sociability and personality traits affect both drinking and earnings, yet they are often difficult to measure. However, the sociability analysis showed that the measures of mental health capture part of the beneficial and detrimental effects of alcohol use. The finding from these measures is that individuals who think that alcohol use has a positive effect on their mental health and also drink less are better off in terms of earnings. Also the dummy for drinking in restaurants indicates the importance of the social side of drinking. In addition, the SES-analysis showed that individuals in the highest SES-group maximize their earnings at higher drinking amounts than other groups, while individuals in the highest group drink less on average than individuals in the other groups. It may indicate of similar kind of sociability factor as Peters (2009) studied, meaning that in the managerial level the sociability of an individual is related to promotion opportunities and alcohol is an important factor in social gatherings. Yet, this is a far stretched assumption and requires a closer inspection in the future studies.

Overall, the study provided many insights into the relationship of drinking and earnings, but there is a need for more explicit measures of sociability and personality traits to fully investigate the social side of drinking. A future study could take this aspect much further and for example use alcohol measures as instruments for sociability measures in an earnings equation.

6. Conclusion

The aim of this thesis was to determine the relationship between alcohol use and earnings as well as to investigate the possible relations and effects with a multidimensional empirical approach. The theoretical foundations of the thesis are based on labor economics and on the work done by other researchers during the past 23 years, whereas the empirical analysis utilized the Finnish Drinking Habits Survey from 2008.

The results of this thesis are interesting and they shed light on an issue that has never before been studied in Finland. Furthermore, majority of the results are supported by theoretical considerations and common understanding on the effects of alcohol use. The measures of alcohol use proved to have high explanatory power in all econometric models that were tested. Thus, the main research question can be answered from many angles.

The results show that on average, drinkers earn more than abstainers, but the share of abstainers in the sample is small. Furthermore, part of the poorer labor market success of abstainers is explained by the presence of former drinkers among them. When the relationship was studied with a continuous alcohol use variable, an inverse U-shaped relationship was found. In gender-specific models, male earnings peak at average consumption of 2.6 drinks per day on yearly basis, which falls into the category of moderate drinking. Female earnings peak at 1.2 drinks per day, which is also in the range of moderate drinking. A novel finding from the gender-specific model is the negative relationship between the amount of intoxications and earnings for men. A descriptive analysis on the age-earnings profiles also indicates the detrimental effects of bingeing and heavy drinking, as individuals with higher risk degree of alcohol use have flatter profiles and face earnings drop at an earlier age than moderate drinkers.

The relationship was also studied for different socioeconomic groups separately. The results from this analysis show that the peak point in earnings is found at different level according to the socioeconomic groups. Interestingly, for the lowest socioeconomic group the peak point in earnings is at a lower level of drinking than for the other groups, while the average drinking amount in this group is the highest. The reasons for this result are ambiguous, but may indicate the positive earnings effect of drinking relatively little alcohol in the lowest socioeconomic group. Additionally, the relationship was studied in terms of beverage-specific drinking and this analysis revealed that

wine drinkers earn 5 percent more than beer drinkers and 15 percent more than generic drinkers. Finally, the social side of drinking was analyzed with several indicator variables. The results imply that "social drinkers" earn more, whereas those who drink to alleviate depression earn significantly less. However, unobserved background factors are likely to affect these outcomes.

The results have a couple of limitations. First, the data is cross-sectional, making it impossible to study the earnings effects of changes in alcohol taxes or to account for individual fixed effects. Second, both structural and statistical endogeneity are likely present in the analysis, which may bias the results. On the other hand, the multidimensional OLS approach yielded consistent results and the separate effects were isolated with a broad set of control variables. Yet, the limitations show that there is a clear need for future studies that focus on longer time period and try to tackle the problem of endogeneity with more advanced statistical methods.

Overall, this thesis has analyzed the effects of alcohol use from a novel angle in the context of Finnish alcohol literature. The relationship between alcohol use and earnings truly is a puzzling one and numerous factors affect the outcomes of the analysis. However, the findings of the thesis are statistically significant and many consistent patterns in the relationship are evident. In conclusion, the results of my thesis suggest that it pays off to drink moderately, preferably wine and in a social manner.

References

Literature

Auld M. 2005. Smoking, Drinking, and Income. *Journal of Human Resources*, **40**(2), 505–518.

Balsa A, Giuliano L, French, M. 2011. The effects of alcohol use on academic achievement in high school. *Economics of Education Review,* **30**, 1–15.

Baltagi B, Griffin J. 2002. Rational addiction to alcohol: panel data analysis of liquor Consumption. *Health Economics*, **11**, 485–491.

Barret G. 2002. The Effect of Alcohol Consumption on Earnings. *Economic Record*, **78**(1), 79–96.

Becker G. 1962. Investment in Human Capital: A Theoretical Analysis. *The Journal of Political Economy*, **70**(5), 9-49.

Becker G, Murphy K. 1988. A Theory of Rational Addiction. *Journal of Political Economy*, **96**(4), 675–700.

Berger M, Leigh J. 1988. The Effect of Alcohol use on Wages. *Applied Economics*, **20**, 1343–1351.

Bray J. 2005. Alcohol Use, Human Capital, and Wages. *Journal of Labor Economics*, **23**(2), 279–312.

Bryant R. 1993. The Influence of Current and Past Alcohol Use on Earnings: Three Approaches to Estimation. *Journal of Applied Behavioral Science*, **29**(1), 9–31.

Coleman J. 1988. Social Capital in the Creation of Human Capital. *American Journal of Sociology*, **94**(S), 95–120.

Cook P, Moore M. 2001. *The Handbook of Health Economics*, Chapter 30: Alcohol. Edited by Culyer A, Newhouse J. Volume 1B, 1630–1666.

Cook P, Peters B. 2005. The Myth of the Drinker's Bonus. Working paper 11902, National Bureau of Economic Research, available at http://www.nber.org/papers/w11902.

Dave D, Kaestner R. 2002. Alcohol taxes and labor market outcomes. *Journal of Health Economics*, **21**, 357–371.

Dave D, Saffer H. 2008. Alcohol demand and risk preference. *Journal of Economic Psychology*, **29**, 810–831.

Dastan I. 2010. Labor Market Effects of Obesity, Smoking and Alcohol Use. Dissertation, Rutgers University. Available at http://mss3.libraries.rutgers.edu/dlr/showfed.php?pid=rutgers-lib:30031.

Deaton A. 1974. The Analysis of Consumer Demand in the United Kingdom, 1900–1970. *Econometrica*, **42**(2), 341–367.

French M, Zarkin G. 1995. Is Moderate Alcohol Use Related to Wages? Evidence from Four Worksites. *Journal of Health Economics*, **14**, 319–344.

Gaziano J, Gaziano T, Glynn R, Sesso H, Ajani U, Stampfer M, Manson J, Hennekens C, Buring J. 2000. Light-to-Moderate Alcohol Consumption and Mortality in the Physicians' Health Study Enrollment Cohort. *Journal of the American College of Cardiology*, **35**(1), 96–105.

Glaeser E, Laibson D, Sacerdote B. 2000. The Economic Approach to Social Capital. NBER Working Paper No. 7728, available at http://www.nber.org/papers/w7728.

Grossman M. 1972. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, **80**(2), 223–255.

Hamilton V, Hamilton B. 1997. Alcohol and Earnings: Does Drinking Yield a Wage Premium? *The Canadian Journal of Economics*, **30**(1), 135–151.

Hvidtfeldt U, Tolstrup J, Jakobsen M, Heitmann B, Gronbaek M, O'Reilly E, Bälter K, Goldbourt U, Hallmans G, Knekt P, Liu S, Pereira M, Pietinen P, Spiegelman D, Stevens J, Virtamo J, Willet

W, Rimm E, Ascherio A. 2010. Alcohol Intake and Risk of Coronary Heart Disease in Younger, Middle-Aged, and Older Adults. *Circulation*, **121**, 1589–1597.

Johansson E, Alho H, Kiiskinen U, Poikolainen K. 2007. The association of alcohol dependency with employment probability: evidence from the population survey 'Health 2000 in Finland'. *Health Economics*, **16**, 739–754.

Keng S, Huffman W. 2007. Binge drinking and labor market success: a longitudinal study on young people. *Journal of Popular Economics*, **20**, 35–54.

Kenkel D, Ribar D. 1994. Alcohol Consumption and Young Adults' Socioeconomic Status. *Brooking Papers on Microeconomic Activity*, 119–175.

Koksal A, Wohlgenant M. 2011. Alcoholic Beverages and Cigarettes: Complements or Substitutes? A Pseudo Panel Approach. Paper provided by Agricultural and Applied Economics Association, available at http://ageconsearch.umn.edu/bitstream/98788/2/saea%20paper.pdf.

Lye J, Hirschberg J. 2010. Alcohol consumption and human capital: a retrospective study of the literature. *Journal of Economic Surveys*, **24**(2), 309–338.

MacDonald Z, Pudney S. 2000. The Wages of Sin? Illegal Drug Use and the Labour Market. *Labour*, **14**(4), 657–673.

Macdonald Z, Shields M. 2001. The Impact of Alcohol Consumption on Occupational Attainment in England. *Economica*, **68**(271), 427–453.

Manning W, Blumberg L, Moulton L. 1995. The demand for alcohol: The differential response to price. *Journal of Health Economics*, **14**, 123–148.

Medoff J, Abraham K. 1980. Experience, Performance, and Earnings. *Quarterly Journal of Economics*, **95**(4), 703–736.

Mellin N, Vihmo J, Österberg E. 2006. Alkoholihaittojen yhteiskunnalliset kustannukset Suomessa vuonna 2003 (Social costs of alcohol-related harm in Finland in 2003). Stakes report, 10/2006

Mullahy J, Sindelar J. 1993. Alcoholism, Work, and Income. *Journal of Labor Economics*, **11**(3), 494–520.

Mäkelä P, Mustonen H, Tigerstedt C. 2010. *Suomi juo*. Helsinki University Press, Helsinki.

Mäkelä P, Österberg E. 2009. Weakening of one more alcohol control pillar: a review of the effects of the alcohol tax cuts in Finland in 2004. *Addiction*, **104**, 554–563.

Van Ours J. 2004. A pint a day raises a man's pay; but smoking blows that gain away. *Journal of Health Economics*, **23**, 863–886.

Peters B. 2004. Is there a wage bonus from drinking? Unobserved heterogeneity examined. *Applied Economics*, **36**, 2299–2315.

Peters B, Stringham E. 2006. No Booze? You May Lose: Why Drinkers Earn More Money Than Nondrinkers. *Journal of Labor Research*, **27**(3), 411–421.

Peters B. 2009. The drinkers' bonus in the military: officers versus enlisted personnel. *Applied Economics*, **41**, 2211–2220.

Polachek S, Siebert W. 1993. *The economics of earnings*. Cambridge University Press.

Renna F. 2008. Alcohol abuse, alcoholism, and labor market outcomes: looking for the missing link. *Industrial and Labor Relations Review*, **62**(1), 92–103.

Skogen J, Harvey S, Henderson M, Eystein S, Mykletun A. 2009. Anxiety and depression among abstainers and low-level alcohol consumers. The Nord-Trøndelag Health Study. *Addiction*, **104**(9), 1519–1529.

Tekin E. 2004. Employment, Wages, and Alcohol Consumption in Russia. *Southern Economic Journal*, **71**(2), 397–417.

Wagenaar A, Salois M, Komro K. 2008. Effects of beverage alcohol price and tax levels on drinking: a meta-analysis of 1003 estimates from 112 studies. *Addiction*, **104**, 179–190.

Williams J, Powell L, Wechsler H. 2003. Does alcohol consumption reduce human capital accumulation? Evidence from the College Alcohol Study. *Applied Economics*, 2003, **35**, 1227–1239.

Zarkin G, French M, Mroz T, Bray J. 1998. Alcohol use and wages: new results from the National Household Survey on Drug Abuse. *Journal of Health Economics*, **17**, 53–68.

Ziebarth N, Grabka M. 2009. In Vino Pecunia? The Association Between Beverage-Specific Drinking Behavior and Wages. *Journal of Labor Research*, **30**(3), 219–244.

Internet sources

Ministry of Finance, 2004. Alkoholiveron alennukset vaimentamaan viinarallia (Using alcohol tax reductions to curb 'liquor rally').
http://www.vm.fi/vm/fi/03_tiedotteet_ja_puheet/01_tiedotteet/2004/66323/name.jsp, 13.7.2011.

Ministry of Finance, 2008. Alkoholiveroa korotetaan kymmenellä prosentilla (Alcohol taxation will be raised by ten percent).
http://www.vm.fi/vm/fi/03_tiedotteet_ja_puheet/01_tiedotteet/20081118Alkoho/name.jsp, 13.7.2011.

National Institute for Health and Welfare, 2008. Alkoholi (Alcohol).
http://www.ktl.fi/portal/suomi/tietoa_terveydesta/elintavat/alkoholi, 23.10.2011.

National Institute for Health and Welfare, 2008. Suurkulutus ja kohtuukäyttö (Heavy use and moderate use).
http://www.ktl.fi/portal/suomi/tietoa_terveydesta/elintavat/alkoholi/suurkulutus_ja_kohtuukaytto, 14.9.2011

Appendix A: Survey questions and specifications

Register-based information that was used to create variables:

Muuttuja: SVATKP
Puhdas ansiotulo kunnallisverotuksessa vuonna 2007, vuoden 2008 tieto saadaan vuonna 2010
(Tieto on poimittu suoraan Verohallituksen henkilöverorekisteristä.)
Henkilön puhdas ansiotulo kunnallisverotuksessa henkilöveroaineistossa. Sisältää vain veronalaiset tulot. (150 000 = tulot 150 000 euroa tai yli)

Muuttuja: SOSE1
Huoltajan sosioekonominen asema
Muuttuja: SOSE2
Kohdehenkilön sosioekonominen asema
(Muuttujat perustuvat kysymyksiin huoltajan (A11b) tai omasta (A12) ammatista. Ammattitieto on poistettu aineistosta.)
21 = maatalousyrittäjät
22 = muut yksityisyrittäjät (esim. vähittäis- ja tukkukauppiaat)
31 = johtotehtävissä toimivat ylemmät toimihenkilöt
32 = suunnittelu- ja tutkimustehtävissä toimivat ylemmät toimihenkilöt
33 = opetustehtävissä toimivat ylemmät toimihenkilöt
34 = muut ylemmät toimihenkilöt
41 = työnjohtotehtävissä toimivat alemmat toimihenkilöt
42 = itsenäistä toimistotyötä tekevät alemmat toimihenkilöt
43 = epäitsenäistä toimistotyötä tekevät alemmat toimihenkilöt
44 = muut alemmat toimihenkilöt
55 = erikoistumattomat työntekijät
56 = erikoistuneet työntekijät
90 = sosioekonominen asema tuntematon
98 = ei ammattia

Muuttuja: SUURAL2, U_ SUURAL2
Suuralue + pääkaupunkiseutu
(Aluemuuttuja perustuu väestörekisterin tietoon marraskuun puolivälissä 2008. Vastaavassa U_-alkuisessa aluemuuttujassa on haastatteluhetken korjattu tieto.)
0 = pääkaupunkiseutu
1 = Etelä-Suomi – pääkaupunkiseutu
2 = Länsi-Suomi
3 = Itä-Suomi
4 = Pohjois-Suomi

Muuttuja: SUKUP
1 = Nainen
2 = Mies

Muuttuja: SYNTV
Syntymävuosi

Base information questions that were used to create variables:

Minkälaiset ovat perhesuhteenne?
Oletteko …

1 naimaton
2 avoliitossa
3 naimisissa
4 asumuserossa
5 eronnut
6 leski?

Kuuluuko asuinkumppaneihinne …
Teidän alaikäisiä lapsianne?

1 Kyllä
2 Ei

Millainen peruskoulutus Teillä on?
Oletteko suorittanut:

1 ylioppilastutkinnon
2 keskikoulun
3 peruskoulun vai
4 kansakoulun, kansalaiskoulun tai vähemmän

Millainen ammatillinen koulutus
Teillä on? Oletteko suorittanut:

1 korkeakoulututkinnon
2 ammattikorkeakoulututkinnon
3 ammatillisen opistotutkinnon
4 ammattikoulututkinnon
5 ammatillisen kurssin?
6 Ei ammatillista koulutusta

Seuraavaksi esitän alkoholin käyttöä koskevia väitteitä. Kertokaa kortin 1 vaihtoehdoilla, mitä mieltä olette kustakin väitteestä

	Täysin samaa mieltä	Osittain samaa mieltä	Vaikea sanoa	Osittain eri mieltä	Täysin eri mieltä
Alkoholista ei ole kenellekään mitään todellista iloa.	1	2	3	4	5
Humaltuminen on vain viaton tapa pitää hauskaa.	1	2	3	4	5

Kortilla 3 on erilaisia alkoholijuomia. Mitä niistä olette käyttänyt viimeksi kuluneiden 12 kuukauden aikana? Mainitkaa kaikki käyttämänne alkoholijuomat, vaikka olisitte nauttinut niitä vain pieniä määriä, esim. puoli pulloa keskiolutta.

1 Olutta tai sahtia

2 Siideriä, long drink-juomia, ruokakaupan Viiniä

3 Viiniä, kuohuviiniä tai väkevää viiniä

4 Väkeviä alkoholijuomia, esim. viskiä, konjakkia, votkaa tai pirtua

5 Kotivalmisteisia alkoholijuomia, kuten kotiviiniä, kiljua tai pontikkaa

6 Muita alkoholijuomia, esimerkiksi juomasekoituksia

1 Alkoholin käyttäjä
2 Raitis

Oletteko joskus aikaisemmin käyttänyt alkoholia?

1 On käyttänyt aikaisemmin
2 Ei käyttänyt aikaisemmin

Minkä ikäisenä olitte ensi kertaa humalassa?

Ikä: (kaksi merkkiä)

Kuinka usein viimeksi kuluneiden 12 kuukauden aikana

joitte yhteensä 18 annosta tai enemmän alkoholia yhtenä päivänä?

joitte yhteensä vähintään 13 annosta, mutta enintään17 annosta alkoholia yhtenä päivänä?

joitte yhteensä vähintään 8 annosta, mutta enintään12 annosta alkoholia yhtenä päivänä?

joitte yhteensä 5, 6, tai 7 annosta alkoholia yhtenä päivänä?

joitte yhteensä 3 tai 4 annosta alkoholia
yhtenä päivänä?

joitte korkeintaan kaksi annosta alkoholia
yhtenä päivänä?

Tiheyskysymysten luokat:
Päivittäin
4-5 kertaa viikossa
2-3 kertaa viikossa
Kerran viikossa
2-3 kertaa kuukaudessa
Noin kerran kuukaudessa
Noin kerran parissa kuukaudessa
3-4 kertaa vuodessa
Kerran pari vuodessa
Harvemmin kuin kerran vuodessa
Ei koskaan

kuinka usein nautitte tavallisesti...

olutta?
siideriä, Long drink -juomia tai vastaavia?
viiniä?
väkeviä juomia?

Luokat:

Päivittäin
4-5 kertaa viikossa
2-3 kertaa viikossa
Kerran viikossa
2-3 kertaa kuukaudessa
Noin kerran kuukaudessa
Noin kerran parissa kuukaudessa
3-4 kertaa vuodessa
Kerran pari vuodessa
Harvemmin kuin kerran vuodessa
Ei koskaan **TAI KORKEINTAAN MAISTANUT**

Kuinka paljon suunnilleen juotte tavallisesti kerralla olutta? Valitkaa yksi kortin 7 riveistä ja ilmoittakaa sen rivin numero.

	Pieniä pulloja (0,33 l)	**Puolen litran pulloja, tölkkejä tai tuoppeja (0,5 l)**	**Litraa (noin)**
1	alle 1	= alle puolikas	= alle 0.33 ltr
2	1 pullo	= vajaa 1	= 0.33 ltr
3	1,5 pulloa	= 1 kpl	= 0.5 ltr
4	2 pulloa	= 1.5 kpl	= 0.66 ltr
5	3 pulloa	= 2 kpl	= 1 ltr
6	4–5 pulloa	= 3 kpl	=1.3–1.7 ltr
7	6–9 pulloa	= 4–6 kpl	= 2-3 ltr
8	10–14 pulloa	= 7–9 kpl	= 3.5–4.5 ltr
9	15+ pulloa	= 10+ kpl	= 5+ ltr

Kuinka paljon suunnilleen juotte tavallisesti kerralla viiniä?

1	Puoli pientä lasillista (alle 12 cl)
2	Pienen lasillisen (12 cl)
3	Ison lasillisen tai kaksi pientä lasillista (24 cl)
4	Puoli pulloa (37,5 cl)
5	Hieman vähemmän kuin pullon (50 - 60 cl)
6	Yhden pullon (75 cl)
7	Puolitoista pulloa (n. litra)
8	Kaksi pullollista tai enemmän (1,5 litraa tai enemmän)

Kuinka paljon suunnilleen juotte tavallisesti kerralla väkeviä juomia?

1	1 ravintola-annos (4 cl)
2	2 ravintola-annosta (8 cl)
3	3 ravintola-annosta (12 cl)
4	4 ravintola-annosta (16 cl)
5	5-6 ravintola-annosta tai puoli pullollista (20–25 cl)
6	7-8 ravintola-annosta tai vähän yli puoli pullollista (noin 30 cl)
7	9-10 ravintola-annosta tai vähän vajaa pullo (noin 40 cl)
8	Puolen litran pullo (50 cl)
9	Puolitoista puolen litran pulloa tai enemmän (yli 75 cl)

Kuinka usein kaiken kaikkiaan nautitte alkoholia, olipa juomalaji mikä tahansa? Laskekaa mukaan myös ne kerrat, jolloin nautitte hyvin pieniä määriä alkoholia, esimerkiksi vain puoli pulloa keskiolutta.

1	Päivittäin
2	4-5 kertaa viikossa
3	2-3 kertaa viikossa
4	Kerran viikossa
5	2-3 kertaa kuukaudessa
6	Noin kerran kuukaudessa

7 Noin kerran parissa kuukaudessa
8 3-4 kertaa vuodessa
9 Kerran pari vuodessa
10 Harvemmin kuin kerran vuodessa
11 Ei koskaan

Kuinka usein viimeisten 12 kuukauden aikana olette juonut alkoholia baarissa tai ravintolassa?

1 Päivittäin
2 4-5 kertaa viikossa
3 2-3 kertaa viikossa
4 Kerran viikossa
5 2-3 kertaa kuukaudessa
6 Noin kerran kuukaudessa
7 Noin kerran parissa kuukaudessa
8 3-4 kertaa vuodessa
9 Kerran pari vuodessa
10 Harvemmin kuin kerran vuodessa
11 Ei koskaan

Onko Teillä jokin haittaava pitkäaikaissairaus tai vamma? Vastaukseksi riittää kyllä tai ei.

1 Kyllä
2 Ei

Alkoholinkäyttöni vaikuttaa myönteisesti mielenterveyteeni:

1 Täysin samaa mieltä
2 Osittain samaa mieltä
3 Osittain eri mieltä
4 Täysin eri mieltä

Käytän toisinaan alkoholia lievittääkseni masennusta tai ahdistusta:

1 Täysin samaa mieltä
2 Osittain samaa mieltä
3 Osittain eri mieltä
4 Täysin eri mieltä

Tunnetteko itsenne yksinäiseksi…

1 jatkuvasti
2 melko usein
3 joskus
4 hyvin harvoin
5 vai ette koskaan?

Kuinka pitkä olette?

Sentteinä: ___kolme merkkiä

Kuinka paljon painatte?

Kiloa: ____kolme merkki

Oletteko joskus kokeillut tai käyttänyt jotain huumausainetta (kuten hasista, marihuanaa, amfetamiinia, heroiinia tai muita vastaavia aineita)?

1 Kyllä
2 Ei

Alkoholijuomat muutettiin sataprosenttiseksi alkoholiksi alla olevilla kertoimilla. Niiden muodostamisessa käytettiin Päihdetilastollinen vuosikirja 2008:n tietoja alkoholijuomien tilastoidusta kulutuksesta vuodelta 2007. Kertoimet saatiin suhteuttamalla juomalajin kulutus litroina sataprosenttista alkoholia kulutukseen kyseistä alkoholijuomaa litroina.

olut 0.0459
siideri 0.0469
viini 0.1276
väkevät 0.3508

Appendix B: Statistical background information

Education, earnings and average drinking volumes in centiliters (number of individuals in parentheses)

Women

Earnings decile	Educational level					
	Total 11 – 12 years	Total 13 – 14 years	3 – 4 years of college	5 – 6 years of college or university	Doctor level	Total
I	158 (33)	43 (8)	572 (2)	59 (5)	- (-)	146 (48)
II	244 (77)	63 (8)	202 (7)	7 (1)	417 (1)	225 (94)
III	202 (82)	268 (8)	68 (15)	104 (5)	- (-)	184 (110)
IV	193 (85)	152 (15)	212 (8)	142 (7)	133 (1)	185 (116)
V	200 (86)	164 (28)	133 (13)	107 (5)	- (-)	182 (132)
VI	290 (86)	169 (23)	150 (22)	113 (11)	22 (1)	234 (143)
VII	193 (60)	187 (39)	126 (27)	148 (7)	- (-)	175 (133)
VIII	188 (29)	225 (36)	120 (27)	187 (22)	6 (1)	182 (115)
IX	245 (24)	272 (35)	110 (16)	270 (25)	66 (3)	234 (103)
X	170 (10)	287 (10)	347 (9)	362 (27)	1028 (8)	402 (64)
Total	216 (572)	198 (210)	150 (146)	219 (115)	600 (15)	209 (1058)

Men

Earnings decile	Educational level					
	Total 11 – 12 years	Total 13 – 14 years	3 – 4 years of college	5 – 6 years of college or university	Doctor level	Total
I	665 (28)	217 (4)	564 (3)	25 (1)	0 (1)	573 (37)
II	1259 (67)	380 (8)	1541 (2)	555 (3)	- (-)	1152 (80)
III	758 (42)	245 (4)	420 (4)	1247 (3)	- (-)	722 (53)
IV	497 (49)	522 (7)	956 (6)	452 (3)	- (-)	543 (65)
V	552 (55)	326 (10)	352 (4)	1392 (2)	- (-)	533 (71)
VI	624 (59)	584 (10)	335 (5)	568 (4)	- (-)	597 (78)
VII	673 (70)	726 (16)	481 (9)	254 (3)	- (-)	651 (98)
VIII	579 (83)	460 (13)	346 (15)	798 (9)	- (-)	554 (120)
IX	513 (69)	639 (24)	482 (32)	632 (21)	- (-)	544 (146)
X	559 (52)	761 (36)	588 (37)	556 (54)	399 (9)	595 (188)
Total	672 (574)	591 (132)	531 (117)	613 (103)	359 (10)	633 (936)

Risk groups and educational levels

Risk group	Educational level					
	Total 11 – 12 years	Total 13 – 14 years	3 – 4 years of college	5 – 6 years of college or university	Doctor level	Total
1	86 (61%)	29 (21%)	14 (10%)	10 (7%)	2 (1%)	141 (100%)
2	306 (51%)	134 (22%)	72 (12%)	77 (13%)	11 (2%)	600 (100%)
3	399 (54%)	116 (16%)	129 (17%)	87 (12%)	9 (1%)	740 (100%)
4	273 (70%)	47 (12%)	38 (10%)	31 (8%)	1 (0%)	390 (100%)
5	82 (67%)	16 (13%)	10 (8%)	13 (11%)	2 (2%)	123 (100%)
Total	1146 (57%)	342 (17%)	263 (13%)	218 (11%)	25 (1%)	1994 (100%)

Heteroskedasticity test for basic OLS-estimation

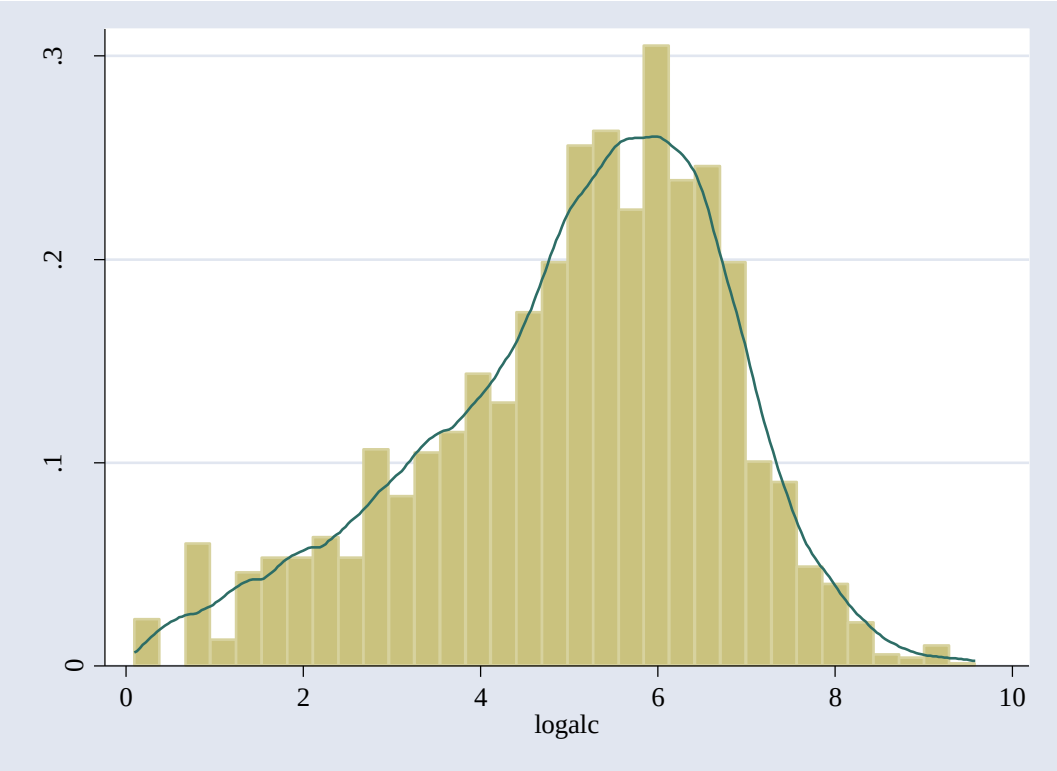
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity

H_o: Constant variance

Variables: fitted values of Logearn

χ^2: 618.70

Histogram on drinking distribution



Socioeconomic groups, gender and yearly drinking amounts in centiliters

	Drinking quantities in centiliters		
Socioeconomic group	Men	Women	Total
SES1	900	659	835
SES2	1066	624	823
SES3	1218	753	1126
Total	1051	655	929

Earnings deciles and beverage classes (number of individuals in each decile)

Earnings decile	Beverage type		
	Beer drinkers	Wine drinkers	Liquor drinkers
I	63	25	21
II	68	26	40
III	61	37	45
IV	84	45	34
V	83	36	34
VI	73	61	37
VII	97	41	44
VIII	102	49	42
IX	109	50	55
X	100	67	70
Total	840	437	422

Separate beverage-specific OLS-models for under and over median earnings

Variable	Coefficient (robust standard errors in parentheses) Under median earnings	Over median earnings
Constant	5.860*** (0.273)	9.373*** (0.108)
Alcohol use		
Beerdrinker	0.121** (0.057)	0.004 (0.019)
Winedrinker	0.172*** (0.063)	0.046** (0.024)
Liqourdrinker	-0.053 (0.064)	0.033 (0.025)
Abstainer	-0.053 (0.064)	0.010 (0.033)
Demographics		
Age	0.136*** (0.001)	0.044*** (0.001)
Age^2	-0.001*** (0.000)	$-4.33x10^{-3}$*** (0.000)
Married	0.034 (0.041)	0.017 (0.019)
Children	0.203*** (0.056)	0.051** (0.021)
Female	-0.017 (0.052)	-0.239*** (0.019)
Area2	-0.096 (0.077)	-0.087*** (0.025)
Area3	-0.062 (0.080)	-0.103*** (0.026)
Area4	-0.246*** (0.089)	-0.146*** (0.030)
Area5	-0.183* (0.102)	-0.108*** (0.030)
Education		
Highschool	0.178** (0.071)	0.124*** (0.019)
University	-0.070 (0.080)	0.196*** (0.022)
Observations	1222	1379
R^2	0.33	0.29

Earnings and the frequency of drinking in restaurants (number of individuals in parentheses)

Restaurant drinking frequency	Men	Women	Total
Daily	25826 (2)	- (-)	25826 (2)
4 – 5 times / week	29951 (4)	37990 (1)	31558 (5)
2 – 3 times / week	26554 (34)	24888 (5)	26340 (39)
Once a week	26191 (64)	11242 (24)	22114 (88)
2 – 3 times / month	27480 (131)	17830 (83)	23737 (214)
Once a month	30728 (142)	22249 (129)	26692 (271)
Once every second month	31829 (134)	22559 (154)	26872 (288)
3 – 4 times / year	30342 (199)	22786 (257)	26084 (456)
1 – 2 times / year	28047 (216)	20083 (316)	23317 (532)
Less than once a year	20667 (84)	18411 (92)	19489 (176)
Never	18140 (158)	14280 (223)	15881 (381)
Total	27122 (1168)	19733 (1284)	23252 (2452)