

**This is an electronic reprint of the original article.  
This reprint *may differ* from the original in pagination and typographic detail.**

**Author(s):** Böckerman, Petri; Hyytinen, Ari; Kaprio, Jaakko; Maczulskij, Terhi

**Title:** If you drink, don't smoke : Joint associations between risky health behaviors and labor market outcomes

**Year:** 2018

**Version:**

**Please cite the original version:**

Böckerman, P., Hyytinen, A., Kaprio, J., & Maczulskij, T. (2018). If you drink, don't smoke : Joint associations between risky health behaviors and labor market outcomes. *Social Science and Medicine*, 207, 55-63.  
<https://doi.org/10.1016/j.socscimed.2018.04.039>

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

# Accepted Manuscript

If you drink, don't smoke: Joint associations between risky health behaviors and labor market outcomes

Petri Böckerman, Ari Hyytinen, Jaakko Kaprio, Terhi Maczulskij



PII: S0277-9536(18)30206-5

DOI: [10.1016/j.socscimed.2018.04.039](https://doi.org/10.1016/j.socscimed.2018.04.039)

Reference: SSM 11719

To appear in: *Social Science & Medicine*

Received Date: 4 September 2017

Revised Date: 20 April 2018

Accepted Date: 24 April 2018

Please cite this article as: Böckerman, P., Hyytinen, A., Kaprio, J., Maczulskij, T., If you drink, don't smoke: Joint associations between risky health behaviors and labor market outcomes, *Social Science & Medicine* (2018), doi: 10.1016/j.socscimed.2018.04.039.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Manuscript ID SSM-D-17-02534R3

**If you drink, don't smoke: Joint associations  
between risky health behaviors and labor market outcomes**

Petri Böckerman<sup>1</sup>, Ari Hyytinen,<sup>2</sup> Jaakko Kaprio<sup>3</sup>, and Terhi Maczulskij<sup>4</sup>

<sup>1</sup> Corresponding author. University of Jyväskylä, School of Business and Economics, Labour Institute for Economic Research and IZA. Address: Pitkäsillanranta 3A, FI-00530 Helsinki, Finland. Phone: +358-9-2535 7330. Fax: +358-9-2535 7332. E-mail: [petri.boeckerman@labour.fi](mailto:petri.boeckerman@labour.fi)

<sup>2</sup> University of Jyväskylä, School of Business and Economics, Jyväskylä, Finland. E-mail: [ari.t.hyytinen@jyu.fi](mailto:ari.t.hyytinen@jyu.fi)

<sup>3</sup> University of Helsinki, Institute for Molecular Medicine (FIMM), P.O. Box 20 (Tukholmankatu 8), 00014 Helsinki, Finland

University of Helsinki, Department of Public Health, P.O.Box 20 (Tukholmankatu 8B), 00014 Helsinki, Finland. E-mail: [jaakko.kaprio@helsinki.fi](mailto:jaakko.kaprio@helsinki.fi)

<sup>4</sup> Labour Institute for Economic Research, Helsinki, Finland. E-mail: [terhi.maczulskij@labour.fi](mailto:terhi.maczulskij@labour.fi)

### **Acknowledgments**

This research has been supported by the Jenny and Antti Wihuri Foundation. All the data work of this paper was carried out at Statistics Finland, following its terms and conditions of confidentiality. Jaakko Kaprio is supported by the Academy of Finland (grants 265240, 263278) and the Sigrid Juselius Foundation. Petri Böckerman and Terhi Maczulskij acknowledge financial support from the Academy of Finland Strategic Research Council project Work, Inequality and Public Policy (number 293120).

**Abstract**

This paper examines the links between risky health behaviors and labor market success. We provide new evidence on the joint relationships between the most prominent forms of risky health behavior – alcohol consumption, smoking and physical inactivity – and long-term labor market outcomes. We use twin data for Finnish men and women linked to register-based individual information on earnings and labor market attachment. The twin data allow us to account for shared family and environmental factors and to measure risky health behaviors in 1975 and 1981. The long-term labor market outcomes were measured in adulthood as an average over the period 1990-2009. The sample sizes are 2156 and 2498 twins, for men and women, respectively. We find that being both a smoker and a heavy drinker in early adulthood is negatively related to long-term earnings and employment later in life, especially for men. We conclude that how and why risky health behaviors cluster and how that affects individual level outcomes call for more attention.

**Keywords:** Finland, risky health behaviors, alcohol consumption, smoking, physical activity, earnings, employment

## 1. Introduction

Health behavior consists of a bundle of choices, such as whether to consume alcohol, whether to smoke and whether to be physically active or not. People do not make decisions on health behaviors independently of each other (e.g., Van Ours, 2004). However, it is not well understood how health behaviors interact and whether and how they are jointly linked to long-term labor market outcomes.

The use of addictive substances is distinct from normal consumer choices. Smoking and, less so, alcohol use are addictive behaviors, and quitting either one – especially smoking – is difficult (National Cancer Institute, 2009). The initiation of smoking and alcohol use usually occurs in adolescence (Taioli and Wynder, 1991; Sartor *et al.* 2007), when the person is not mature enough to take long-term effects into account in decisions. Physical inactivity tracks moderately from adolescence onwards (Waller *et al.* 2017). There are motivational factors and societal barriers to maintaining an adequate level of physical activity (Aaltonen *et al.* 2014).

There is an extensive body of literature on the relationships between *specific* risky and protective health behaviors and labor market outcomes (Cawley and Ruhm, 2012). Smoking and heavy alcohol consumption are associated with weaker labor market attachment and lower earnings (French and Zarkin, 1995; MacDonald and Shields, 2001; Van Ours, 2004; Böckerman *et al.* 2015, 2017; Korhonen *et al.* 2015). Furthermore, there is a negative relationship between physical inactivity and subsequent labor market outcomes (Lechner, 2009; Hyytinen and Lahtonen, 2013). These negative correlations are consistent with risky health behavior and weak health eroding the capacity (e.g., owing to increased absence from work) and opportunities

(e.g., due to discrimination by employers, co-workers or consumers) to earn market income.

A crucial limitation of the empirical literature is that the effect of a specific health behavior has been analyzed in isolation, without considering the potential joint associations of health behaviors with labor market outcomes. An exception to this is Van Ours (2004), who examined the wage effects of alcohol use and smoking using survey data from the Netherlands. For men, Van Ours (2004) found that the association of wage with alcohol use was, a bit surprisingly, positive. The wage effect of smoking was opposite and of approximately equal (absolute) size as that of alcohol use. While insightful, Van Ours (2004) did not consider the (potential) *joint* relationships between alcohol use and smoking. Heavy alcohol consumption may reinforce the negative association of smoking with labor market outcomes.

This paper contributes to the literature by developing an empirical framework that treats risky health behaviors as a bundle and allows for a systematic examination of the joint associations (i.e., interactions) of health behaviors. We adopt the econometric approach of Carree et al. (2011), which was originally introduced to examine the joint effects (i.e., complementarities) between different production inputs. We modify their approach so that it allows a direct examination of whether, e.g., heavy alcohol consumption and smoking reinforce each other's links to labor market outcomes. In so doing, we address the recent call of Cawley (2015), who stressed the need to understand the interconnections among health behaviors.

We examine the potential joint associations of health behaviors with *long-term* labor market outcomes. We use nationally representative twin data that are linked to register-based information on long-term labor market outcomes. The linked data have four major strengths. First, the data allow us to construct measures for risky health

behaviors (smoking, alcohol use, physical inactivity) that refer to the same points in time in the twins' early adulthood and that are predetermined relative to the outcome variables, measured later in life. Second, we measure individuals' smoking, alcohol use and physical inactivity over an extended period using two surveys, one conducted in 1975 and the other in 1981. Therefore, we are not forced to rely on cross-sectional measures of risky health behaviors, which may provide incomplete or noisy descriptions of what unhealthy lifestyle choices imply in the long term. Third, the register-based administrative data allow us to measure the average of an individual's annual earnings and employment over an extended 20-year period. Cross-sectional measures are poor proxies for individuals' lifetime labor market outcomes (Böhlmark and Lindquist, 2006). The use of register-based data also minimizes non-response and reporting biases that are typical in survey data. Fourth, the twin data allow distinguishing between monozygotic (MZ, genetically identical) and dizygotic (DZ, genetically full siblings) twins. We can therefore to account for both shared environmental factors, such as family background, neighborhood and shared peer effects, as well as for genetic factors, which are potential confounders. The data also allow us to control for a number of (predetermined) differences between the co-twins of twin pairs.

The prior medical and epidemiological literature suggests a number of reasons for why unhealthy behavioral patterns may cluster (i.e., why certain health behaviors are related to other health behaviors; see, e.g., Hale and Viner, 2016). For example, Rensburg *et al.* (2009) argue that exercising affects implicit motivational processes and promotes smoking cessation. Papathanasiou *et al.* (2012) note that the direction of causality may run from smoking to physical inactivity through reduced cardio-respiratory fitness. Using a twin design, Kujala *et al.* (2007) examine adolescent

physical activity and later smoking behavior. Smoking later in life is more prevalent among those who were initially persistently inactive. Evidence on whether exercise supports smoking cessation is inconclusive (Ussher *et al.* 2014). Using an experimental design, Savette *et al.* (2005) showed that alcohol use increased both the magnitude and emotional valence of cigarette cravings. Smoking may also enhance pleasure and reward from alcohol consumption (Lipperman-Kreda and Lee, 2011). Additionally, nicotine can promote the consumption of alcohol, for example, through different neurotransmitter expressions (Lajitha and Sershen, 2010). Finally, research based on twin data suggest that multiple risky health behaviors are correlated (see, e.g., Sudharsanan *et al.* 2016, and the references therein). For example, True *et al.* (1999) and Han *et al.* (1999) found that genetic factors contribute to the risk for dual dependence of both alcohol use and smoking (see also Madden and Heath, 2002). Further support for shared genetic liability comes from molecular genetic analyses of measured genetic variants across the genome; genetic correlations between smoking and alcohol use and abuse are high (Clarke *et al.* 2017). Liao *et al.* (2016) showed that smoking is negatively associated, and former drinking positively associated with BMI even when shared genetic and environmental factors were accounted for.

The clustering of smoking, heavy alcohol consumption and physical inactivity at the individual level may lead to weaker labor market attachment (e.g., increased absence from work) and to limited job market opportunities through numerous channels. While we cannot empirically pin down the exact mechanisms at work, there are two primary reasons why the clustering of adverse health behaviors can be expected to predict poor labor market outcomes. First, the clustering of addictive and irresponsible health behaviors may be related to a latent factor mirroring myopic time preferences or delayed discounting (impatience) of individuals. Such preferences



reduce a person's initial investments in human, social and health capital, making him/her less productive in the labor market. Second, the clustering of risky health behaviors may erode an individual's existing stocks of human, social and health capital. For example, the smoking-drinking interaction erodes employability and productivity at work: Heavy alcohol use leads to work absence (Norström, 2006) and can alone be a cause of severe adverse health conditions (e.g. due to accidents while drunk). When combined with smoking, alcohol use has a particularly negative effect on specific health measures (Antunes *et al.* 2013). For example, joint consumption of alcohol and cigarettes increases the risk of cancer and cardiovascular diseases and eventually mortality (e.g., Wang-Hong *et al.* 2007). Similarly, being a smoker who is physically inactive or being a heavy drinker who is physically inactive may, in the long-term, lead to reduced physical capacity and erode cognitive and non-cognitive skills, which are the fundamental determinants of earnings and employment capacity (Heckman *et al.* 2006).

## 2. Methods

### *Data sources and the sample*

We use the Older Finnish Twin Cohort Study of the Department of Public Health at the University of Helsinki. As in prior work (Hyytinen and Lahtonen 2013; Böckerman *et al.* 2015, 2017), the twin data have been linked to the Finnish Longitudinal Employer-Employee Data (FLEED) of Statistics Finland (SF) using personal identifiers. The record linkages of linked data comply the Data Protection

Act and have been approved by the ethical committee of the Department of Public Health, University of Helsinki and SF.

The Finnish Cohort Study was originally compiled from the Central Population Registry of Finland (Kaprio *et al.* 1979; Kaprio and Koskenvuo, 2002). Initial twin candidates were persons born before 1958 with the same birth date, commune of birth, sex, and surname at birth. A questionnaire was mailed to the candidates in 1975 to gather baseline data and to determine their zygosity. Two follow-up surveys were conducted in 1981 and 1990. The 1990 survey was sent only to those twins who were born after 1930. These twins were at least 33 years old in 1990. Therefore, we focus on the working-age population.

FLEED is an annual panel over the years 1990-2009 that covers the working-age population of Finland (see, e.g., Hyytinen and Lahtonen 2013; Böckerman *et al.* 2015, 2017). FLEED contains information, obtained from tax and other administrative registers, on individuals' labor market status, and salaries and other relevant sources of income.

The analysis focuses on twin pairs for whom we observe information on health behaviors in 1975 and 1981 and earnings and employment status from 1990 to 2009. Excluding those men and women who retired before 2009, the estimation sample includes 4,654 twin pairs, i.e., 9,308 individuals. The individuals were, on average 27, years old in 1975, 33 years old in 1981 and 42 years old at the time we start measuring their labor market outcomes in 1990.

### *Measures*

We use two outcome variables that capture poor labor market outcomes in the long run. The first outcome variable measures an individual's lifetime earnings (= wage + salary earnings + self-employment income). It is calculated as the gender-specific reverse rank order (100-1) of average lifetime earnings over the observation window 1990-2009. The variable is calculated using the percentiles of the distribution of lifetime earnings in our sample and describes how well an individual fares relative to other individuals in terms of his/her long-term earnings capacity. The second outcome variable measures an individual's long-term tendency to be frequently unemployed (i.e. his/her labor market attachment). It is calculated as the average number of unemployment months per year over the period 1990-2009.

We measure risky health behaviors by heavy alcohol consumption, smoking and physical inactivity. These behaviors have drawn considerable attention in the empirical literature in health economics (and related fields), but their labor market implications have rarely been explored in the same analysis (see Cawley and Ruhm, 2012). Heavy alcohol consumption and smoking are also the major contributors to differences in life-expectancy in both developed and developing countries (see Steel, 2017), and for example account for about one-half of the SES differences in mortality in Finland (Martikainen *et al.* 2014). They are also modifiable risk factors (no one needs to smoke or use alcohol), and physical activity levels can be increased also by societal action. For ease of interpretation, the variables for risky health behaviors are defined such that they reflect risky or undesirable health behavior.

To capture long-term risky health behaviors, the measures are based on self-reported information from the 1975 and 1981 twin surveys. We measure heavy

alcohol consumption by a persistent tendency to be a binge drinker. Binge drinking is associated with negative health consequences, such as increased risk of adverse cardiovascular events (stroke and myocardial infarction). It is a better measure of pattern of drinking than quantity alone (Sipilä *et al.* 2016; Rehm *et al.* 2017). Our measure is equal to one for those who in both 1975 and 1981 reported consuming at least once a month on the same occasion an amount of alcohol that corresponds to at least one bottle of wine. To capture persistent smoking, we use a binary indicator for current smoking status in 1975 and 1981 (i.e., whether the person reported smoking at the time of both surveys). To measure physical inactivity, we use the MET index (Metabolic Equivalent Tasks). For persons who are considered physically inactive, the mean MET index is ~1-2. This is roughly equivalent to the energy expenditure of sitting and of minor home activities (Ainsworth *et al.* 2000). We followed the World Health Organization (2010) and used an MET index threshold of being less than 1.5 both in 1975 and 1981 as a measure for persistent physical inactivity.

We use the binary indicators to define risky health behaviors, because the binary indicators (as opposed to continuous measures) allow us to explore the clustering of risky health behavior in a transparent way. Every individual is assigned to a single category so that the categories are mutually exclusive. If, in 1975 and 1981, a person only had a history of consuming alcohol heavily, we set  $H = 1$  (and set  $H = 0$  otherwise); if he/she had only smoked, we set  $S = 1$  (and set  $S = 0$  otherwise); and if he/she was only physically inactive, we set  $P = 1$  (and set  $P = 0$  otherwise). Those who had heavy alcohol consumption and who smoked in 1975 and 1981 but who were not physically inactive are assigned to a category denoted as HS. The other categories, HP and SP, are defined similarly. Finally, we set HSP equal to one for those persons who had heavy alcohol consumption, who smoked and who were

physically inactive in 1975 and 1981. An individual cannot belong to more than one category from set {H, S, P, HS, HP, SP, HSP}, implying that the reference group consists of those who have no risky health behaviors.

We consider the robustness of the estimation results using different risky behavior measures for heavy drinkers, smokers and physically inactive persons. The description of these additional results is provided in the Online Supplementary Appendix. [INSERT LINK TO ONLINE FILES]

### *Econometric approach*

Inspired by Carree *et al.* (2011), we postulate the following econometric specification for the (poor) long-term labor market outcomes:

$$\begin{aligned}
 Y_{ij} = & \alpha_H H_{ij} + \alpha_S S_{ij} + \alpha_P P_{ij} + \alpha_{HS} HS_{ij} \\
 & + \alpha_{HP} HP_{ij} + \alpha_{SP} SP_{ij} + \alpha_{HSP} HSP_{ij} + X_{ij}'\beta + f_j + g_j + \varepsilon_{ij}
 \end{aligned} \tag{1}$$

where  $Y_{ij}$  is one of the measures for poor labor market outcomes for individual  $i$  from family (twinpair)  $j$ , binary indicators {H, S, P, HS, HP, SP, HSP} are mutually exclusive groups of health behaviors,  $X_{ij}$  is a vector of control variables,  $f_j$  is an unobserved shared family effect,  $g_j$  is an unobserved shared genetic effect and  $\varepsilon_{ij}$  is an i.i.d. error term.

To see how the model works, consider heavy drinking and smoking: There is evidence for an adverse interaction effect between them if the hypothesis  $\alpha_H + \alpha_S = \alpha_{HS}$  is rejected (i.e., if sum  $\alpha_H + \alpha_S$  is smaller than  $\alpha_{HS}$ ). Similar reasoning applies to

heavy drinkers and physically inactive ( $\alpha_H + \alpha_P \neq \alpha_{HP}$ ) and to smokers and physically inactive ( $\alpha_S + \alpha_P \neq \alpha_{SP}$ ).

We estimate model (1) in three ways. First, we treat the data as individuals (i.e., assume  $f_j \equiv 0$  and  $g_j \equiv 0$ ) and estimate the model parameters by ordinary least squares (OLS), using the sample that includes both MZ and DZ twins. Second, we take twin differences to remove  $f_j$  and run the regression using the same combined sample. Third, we further run the regression using the MZ sample only to remove both  $f_j$  and  $g_j$ . In the within-MZ model, all factors that the two twins share (e.g., shared environmental factors, shared genes, business cycle effects, cohort effects) are eliminated.

We include in  $X_{ij}$  an individual's age and a rich vector of control variables to account for potential confounders and within-twin pair heterogeneity. All control variables are taken from the 1975 twin survey and are thus predetermined. First, we account for heterogeneity in pre-existing health conditions that may be correlated with risky health behavior and labor market outcomes. To this end, we include measures for height, BMI, the number of chronic diseases, and reduced employability due to injuries (= 1 if a person had injuries causing weak employability). We also control for the possibility that individuals are physically inactive off work because they have a (presumably low-paid) physically demanding job (= 1 if work is physically demanding). Second, there may be unobserved psychological and psychopathological characteristics that affect both risky health behavior and labor market outcomes, such as mental stability. Mental stability is measured using the indicators of neuroticism and extraversion that originate from the 1975 survey. Neuroticism (extraversion) was assessed by 10 (9) items in the short form of the Eysenck Personality Inventory. We also add the use of tranquilizers as a covariate (= 1 if the twin reports using a positive

quantity of tranquilizers in 1975) to capture an aspect of mental ill-health. Finally, to account for potential reverse causality, we add controls for predetermined employment status and prior earnings from the 1975 survey. These variables address the concern that alcohol consumption, physical inactivity, or smoking may be the consequence of a poor labour market position early in life.

Regression model (1) includes a number of interaction terms between the variables that describe specific health behaviors. We first look at men and women separately. This analysis reveals that certain combinations of risky health behaviors are rare. To ensure that we have enough variation in the analysis sample (and thus to preserve statistical power), we confirm the results using pooled data consisting of both men and women.

### **3. Descriptive statistics**

#### *Mean values of risky health behaviors*

Table 1 reports the mean values of the mutually exclusive indicators of risky health behaviors by gender (not age-adjusted). Four key findings stand out. First, a large fraction of men (39%) and women (57%) are not persistently heavy users of alcohol, smokers or physically inactive. The gender difference is large in magnitude and statistically significant. Second, women are more likely just to smoke (S) or to be physically inactive (P) than men, whereas men are more likely just to be heavy drinkers (H). Third, men are more likely to have multiple types of risky health behaviors. Notably, men are from four to six times more likely than women to have heavy alcohol consumption and be smokers (HS), to have heavy alcohol consumption

and be physically inactive (HP) and to be heavy drinkers, smoker and physically inactive (HSP). Fourth, certain combinations of risky health behaviors are rare. In particular, only 1-2% of women both drink and smoke (HS), drink and are physically inactive (HP), or have all three risky behaviors (HSP).

#### *Association of risky health behaviors with labor market outcomes*

We analyze the mean values of the measures of the labor market outcomes by different categories of risky health behaviors and gender in Figures 1-2. The vertical lines of each bar display the 95% confidence intervals for the means. Figures 1-2 reveal two important patterns. First, those who are heavy drinkers, smokers *and* are physically inactive have much lower lifetime earnings and much more unemployment months in the long run, especially when they are compared to those who have no undesirable health behavior. Second, the lifetime earnings are also notably low and the number of unemployment months is high when smoking is combined with heavy alcohol use (HS), especially for men.

#### **4. Main estimation results**

The OLS and within-sibling estimates of equation (1) are reported in Table 2 (men) and Table 3 (women). The outcome variables are long-term earnings and unemployment. All models include {H, S, P, HS, HP, SP, HSP} as regressors, implying that the baseline comparison category refers to those who have no risky health behaviors. The rows in the lower part of the tables report tests for the model



coefficients. The rows labelled as T4, T7 and T10 report the relevant tests for the interaction effects between our three risky health behavior measures.

The OLS estimates tend to be more often statistically significant than the estimates of the within DZ-MZ and within-MZ regressions. The results based on twin design therefore suggest that it is important to control for unobserved family and genetic factors in estimating the joint associations of health behaviors with long-term labor market outcomes. In what follows, we focus only on those estimates that are statistically significant in the within DZ-MZ and, especially, within-MZ regressions. Using these criteria, four important findings stand out.

First, smoking *per se* is related to weaker earnings and employment (cf. Böckerman et al. 2015) when unobserved family factors are controlled for. The estimates become mostly statistically insignificant when also genetic factors are accounted for in the models. Second, earnings and employment are much lower for those men who have engaged in smoking *and* heavy alcohol use (HS) than for those who do not report any risky health behaviors. Third, consistent with Figures 1-2, those men and women who were heavy drinkers, smoked and were physically inactive in 1975/1981 (HSP) are more likely to have low lifetime earnings and experience more unemployment months than those who do not report any risky health behaviors.

Fourth, and most importantly, the F-tests in the lower parts of the tables reveal that there is a joint effect between smoking and heavy alcohol use for men when the reverse rank order of lifetime earnings and unemployment months are used as the outcome. In these cases, the null hypotheses  $\alpha_H + \alpha_S = \alpha_{HS}$  are rejected, especially in the MZ regression. For women, there is no evidence for the joint effects.

A problem with the above gender-specific analysis is that certain combinations of health behaviors are rare, especially among women. This data limitation may lead

to noisy estimates. To tackle the problem, we estimated model (1) in the twin-differenced form after pooling men and women together. This pooling results in larger estimating samples. The results are presented in Columns 1 and 3 (within DZ and MZ regression) and Columns 2 and 4 (within MZ regression) of Table 4. As Columns 2 and 4 show, we find evidence for the joint effects of smoking and heavy drinking when both shared environment and genetic factors are fully accounted for; for both the outcome measures the null hypothesis  $\alpha_H + \alpha_S = \alpha_{HS}$  is rejected (i.e. the tests T4 in the lower part of Table 4).

## 5. Robustness checks and auxiliary evidence

### *Alternative outcome variables*

We have re-run the within DZ-MZ and within MZ estimations for a combined sample of men and women reported in Table 4 using two alternative outcome variables. Both variables were chosen to reflect an individual's (potentially poor) long-term labor market position. First, we used the logarithm of average social income transfers over the period 1990-2009 as an outcome variable. Second, we used an indicator of being out of the labor force as an outcome variable. It was calculated as the average share of years that the individual was out of the labor force over the period 1990-2009, excluding retired persons, students and those who are in military service. The results are reported in Table A1 of Online Supplementary Appendix. [INSERT LINK TO ONLINE FILE] Consistent with our earlier results in Table 4, the within MZ estimations show that there is evidence for the joint effect between smoking and

heavy use of alcohol, as the null hypothesis  $\alpha_H + \alpha_S = \alpha_{HS}$  is rejected when social income transfers are used as the outcome variable.

Our findings suggest that those who heavily consume alcohol and are smokers and that those who are heavy drinkers, smokers and physically inactive are less successful in life. To check whether this pattern holds more broadly, we report further *descriptive* evidence in Table A2 of the Online Supplementary Appendix. [INSERT LINK TO ONLINE FILE] The table displays education years, marital status, whether a person has children and whether a person is a homeowner by risky health behavior and gender (not age-adjusted), as measured in 1990. We regard these variables as descriptive indicators of better socio-economic outcomes. The table shows that those men who report being smokers and heavy drinkers and those men who report all three risk factors are typically less educated, less likely married, less likely to have children and less likely homeowners than those who report none. The same findings largely apply to women. This auxiliary evidence supports the view conveyed by our baseline analysis that these two bundles of risky health behaviors {HS, HSP} are associated with particularly poor long-term outcomes. Therefore, the clustering of risky health behaviors is empirically related to a low level of human and social capital.

#### *Alternative measures for risky behaviors*

Identifying what amounts to risky health behavior and measuring it appropriately are notoriously difficult empirical tasks. We have therefore evaluated the robustness of the within DZ-MZ and within MZ estimation results of Table 4 using alternative measures for risky health behavior.

First, we used gender-specific thresholds for heavy alcohol use instead of the indicator for binge drinking. This extension is important, because all self-reported measures of alcohol consumption potentially contain some amount of measurement error. The twin surveys report alcohol use in grams per day, which has been converted into grams per week. We followed Dawson (2011) and used cut-off points that are 165 grams per week for men (14 units of alcohol) and 110 grams per week for women (9 units of alcohol). There are no official global limits for heavy alcohol use, and the thresholds used by the public health authorities vary considerably from one country to another and within a single country (Dawson, 2011 and Shield *et al.* 2017). We created an indicator variable that obtains the value of one if a person consumed more than the gender-specific cut-off amount of alcohol per week both in 1975 and 1981.

Second, we used a more restrictive indicator for smoking. The current smoking status is based on the subjective view of one's smoking behavior. Many light smokers could report themselves as smokers. Another measure for smoking captures the lifetime consumption of cigarettes: cigarette pack years = average number of cigarettes smoked per day  $\times$  years of smoking. The measure was constructed from questions about the use of tobacco products and frequency and the duration of smoking. For example, a person has only a 0.05 pack year history of smoking if (s)he has smoked one cigarette daily for one year. We used an indicator that is equal to one for those who were current smokers *and* ever smokers, defined as the persons who reported being current smokers in 1975 and 1981 and who had strictly positive pack-years both in 1975 and 1981.

Third, we used an alternative measure of leisure time physical activity. The measure was constructed from questions about the monthly frequency, average duration and average intensity of a person's physical activity sessions. Using

information on person's physical activity and activity level during the journey to and from work, leisure time physical activity was classified into categories: sedentary exercisers, occasional exercisers and conditioning exercisers (Kujala *et al.* 1998, p. 441; Hyytinen and Lahtonen, 2013). We use an indicator that is equal to one for those who are physically inactive, defined as those persons who reported not taking part in physical leisure activity (sedentary exercisers) in both 1975 and 1981.

Tables A3-A5 of the Online Supplementary Appendix report the robustness checks from the within DZ-MZ and within MZ samples that pool men and women together. [INSERT LINK TO ONLINE FILE] Importantly, there is evidence for the joint effects between heavy alcohol use and smoking, as the null hypothesis  $\alpha_H + \alpha_S = \alpha_{HS}$  is rejected in most specifications, especially when both shared environmental and genetic factor are accounted for. Finally, compared to those who do not have any risky health behaviors, individuals who smoke, drink heavily and are sedentary are more likely to have lower long-term earnings and experience more unemployment.

## 6. Conclusions

Health is a durable good that is determined by genetic predisposition to have good health and by complex health behaviors. Earlier research has examined the (causal) effects of *specific* risky health behaviors on labor market outcomes, such as earnings and (un)employment. This literature suggests that alcohol consumption, smoking and physical inactivity all lead *per se* to poor labor market outcomes. Our most important contribution is that we evaluated risky health behavior as a bundle and examined how (longer-term) measures of risky health behaviors are associated with long-term labor market outcomes (earnings and unemployment).

The key finding is that there is a degree of adverse complementarity between alcohol consumption and smoking, leading to worse long-term labor market outcomes. We also find evidence that those who smoke, drink heavily and are physically inactive experience much poorer long-term labor market outcomes than those individuals who are neither smokers, drinkers nor physically inactive.

Our findings are policy-relevant for two reasons. First, there is a debate over the extent to which health behavior is driven by rational decision-making (Becker and Murphy, 1988; Cawley and Ruhm, 2012). The case for rational health behaviors weakens if there are significant or complicated joint associations, because it is cognitively cumbersome for an individual to comprehend and correctly predict the importance of such effects, especially when young or for a long period of time. The addictive nature of smoking in particular and heavy and regular drinking may shape preferences, erode willpower, and/or undermine cognitive capacity needed for rational, forward-looking decision-making. Second, the trends in the health behavior indicators have varied during the past few decades and across countries. For example, countries are at different stages of the tobacco epidemic (Lopez *et al.* 1994; Reitsma *et al.* 2017). Obesity and physical inactivity at work and during commuting have become much more common in all industrialized countries. There has been a controversy over whether the long-term trends of smoking and obesity are related (Gruber and Frakes, 2006). Better awareness of the interconnections between health behaviors is needed to understand what mechanisms these trends reflect and how they affect labor markets.

Our approach has important limitations. First, even though by using the twin design, the estimated relationships are purged from shared family and environmental effects, we cannot claim that the estimated relationships are causal. Additionally, we

cannot rule out that there are some other unaccounted (latent) person-specific attributes that have an impact on both risky health behavior and the long-term outcomes. For example, there may be relevant psychological and psychopathological differences between twins that have an impact on the estimates. We accounted for these psychological differences by controlling for within-twin pair differences using the measures of mental well-being in the regression model. Second, we estimated the empirical specifications in the Finnish setting. Because the prevalence and development over time of specific risky health behavior bundles is not identical in all industrialized countries, the potential joint effects on labor market outcomes may also differ across countries. However, our econometric approach can easily be applied to the data available for other countries in order to broaden the policy discussions about the consequences of risky health behaviors.

### **Declaration of interests**

None.

### **Acknowledgments**

This research has been supported by the Jenny and Antti Wihuri Foundation. All the data work of this paper was carried out at Statistics Finland, following its terms and conditions of confidentiality. Jaakko Kaprio is supported by the Academy of Finland (grants 265240, 263278) and the Sigrid Juselius Foundation. Petri Böckerman and Terhi Maczulskij acknowledge financial support from the Academy of Finland Strategic Research Council project Work, Inequality and Public Policy (number 293120).

**References**

Aaltonen, S., Kujala, U.M., Kaprio, J. 2014. Factors behind leisure-time physical activity behavior based on Finnish twin studies: The role of genetic and environmental influences and the role of motives. *BioMed Research International*, 931820. doi: 10.1155/2014/931820. Epub 2014 Apr 8. Review. PubMed PMID: 24809061; PubMed Central PMCID: PMC3997869.

Ainsworth, B.E., Haskell, W.L., Whitt, M.C., Irwin, M.L., Swartz, A.M., Strath, S.J., Leon, A.S. 2000. Compendium of physical activities: An update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise*, 32, S498–S504.

Antunes, J.L.F., Toporcov, T.N., Biazevic, M.G.H., Boing, A.F., Scully, C., Petti, S. 2013. Joint and independent effects of alcohol drinking and tobacco smoking on oral cancer: A large case-control study. *PLoS One*, 8(7), e68132.

Becker, G.S., Murphy, K.M. 1988. A theory of rational addiction. *Journal of Political Economy*, 96, 675–700.

Böckerman, P., Hyytinen, A., Kaprio, J. 2015. Smoking and long-term labour market outcomes. *Tobacco Control*, 24, 348–353.

Böckerman, P., Hyytinen, A., Maczulskij, T. 2017. Alcohol consumption and long-term labor market outcomes. *Health Economics*, 26, 275–291.



Böhlmark, A., Lindquist, M.J. 2006. Life-cycle variations in the association between current and lifetime income: Replication and extension for Sweden. *Journal of Labor Economics*, 24, 879–900.

Carree, M., Lokshin, B., Belderbos, R. 2011. A note on testing for complementarity and substitutability in the case of multiple practices. *Journal of Productivity Analysis*, 35, 263–269.

Cawley, J. 2015. An economy of scales: a selective review of obesity's economic causes, consequences, and solutions. *Journal of Health Economics*, 43, 244–268.

Cawley, J., Ruhm C.J. 2012. The economics of risky health behaviors. In: McGuire T. G., Pauly M. V., Barros P. P., editors. *Handbook of Health Economics*, Volume 2. New York: Elsevier, pp. 95–199.

Clarke, T.K., Adams, M.J., Davies, G., et al. 2017. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N= 112,117). *Molecular Psychiatry*, 22, 1376–1384.

Dawson, D. 2011. Defining risk drinking. *Alcohol Research & Health*, 34, 144–156.

French, M., Zarkin, G. 1995. Is moderate alcohol use related to wages? Evidence from four worksites. *Journal of Health Economics*, 14, 319–344.

Gruber, J., Frakes, M. 2006. Does falling smoking lead to rising obesity? *Journal of Health Economics*, 25, 183–197.

Hale, D.R., Viner, R.M. 2016. The correlates and course of multiple health risk behaviour in adolescence. *BMC Public Health*, 16(1), 458.

Han, C., McGue, M., Iacono, W. 1999. Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: Univariate and multivariate behavioral genetic analyses. *Addiction*, 94, 981–993.

Heckman, J.J., Stixrud, J., Urzua, S. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24, 411–482.

Hyytinen, A., Lahtonen, J. 2013. The effect of physical activity on long-term income. *Social Science and Medicine*, 96, 129–137.

Kaprio, J., Artimo, M., Sarna, S., Rantasalo, I. 1979. The Finnish twin registry: baseline characteristics. Section I: materials methods, representativeness and results for variables special to twin studies. Department of Public Health, Publications, M(47): Helsinki.

Kaprio, J., Koskenvuo, M. 2002. Genetic and environmental factors in complex diseases: The Older Finnish Twin Cohort. *Twin Research*, 5, 358–365.

Korhonen, T., Smeds, E., Silventoinen, K., Heikkilä, K., Kaprio, J. 2015. Cigarette smoking and alcohol use as predictors of disability retirement: A population-based cohort study. *Drug and alcohol dependence*, 155, 260–266.

Kujala, U., Kaprio, J., Rose, R. 2007. Physical activity in adolescence and smoking in young adulthood: A prospective twin cohort study. *Addiction*, 102, 1151–1157.

Kujala, U., Kaprio, J., Sarna, S., Koskenvuo, M. 1998. Relationship of leisure-time physical activity and mortality – the Finnish twin cohort. *Journal of the American Medical Association*, 279, 440–444.

Lajitha, A., Sershen, H. 2010. Nicotine: Alcohol reward interactions. *Neurochemical Research*, 35, 1248–1258.

Lechner, M. 2009. Long-run labour market and health effects of individual sports activities. *Journal of Health Economics*, 28, 839–854.

Lia, C., Gao, W., Cao, W., et al. 2016. The association of cigarette smoking and alcohol drinking with body mass index: A cross-sectional, population-based study among Chinese adult male twins. *BMC Public Health*, 16(1), 311.

Lipperman-Kreda, S., Lee, J. 2011. Boost your high: Cigarette smoking to enhance alcohol and drug effects among Southeast Asian American youth. *Journal of Drug Issues*, 41, 509–522.

Lopez, A., Collishaw, N., Piha, T. 1994. A descriptive model of the cigarette epidemic in developed countries. *Tobacco Control*, 3, 242–247.

MacDonald, Z., Shields, M.A. 2001. The impact of alcohol consumption on occupational attainment in England. *Economica*, 68, 427–453.

Madden, P., Heath, A. 2002. Shared genetic vulnerability in alcohol and cigarette use and dependence. *Alcoholism*, 26, 1919–1921.

Martikainen, P., Mäkelä, P., Peltonen, R., Myrskylä, M. 2014. Income differences in life expectancy: The changing contribution of harmful consumption of alcohol and smoking. *Epidemiology*, 25, 182–190.

National Cancer Institute. 2009. Phenotypes and Endophenotypes: Foundations for Genetic Studies of Nicotine Use and Dependence. *Tobacco Control Monograph No. 20*. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute. NIH Publication No. 09-6366, August 2009.

Norström, T. 2006. Per capita alcohol consumption and sickness absence. *Addiction*, 101, 1421–1427.

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

Papathanasiou, G., Papandreou, M., Galanos, A., Kortianou, E., Tsepis, E., Kalfakakou, V., Evangelou, A. 2012. Smoking and physical activity interrelations in health science studies. Is smoking associated with physical activity in young adults? *Hellenic Journal of Cardiology*, 53, 17–25.

Reitsma, M.B., Fullman, N., Ng, M. et al. 2017. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: A systematic analysis from the Global Burden of Disease Study 2015. *The Lancet*, 389, 1885–1906.

Rensburg, K., Taylor, A., Hodgson, T. 2009. The effects of acute exercise on attentional bias towards smoking-related stimuli during temporary abstinence from smoking. *Addiction*, 104, 1910–1917.

Sartor, C. E., Lynskey, M. T., Heath, A. C., Jacob, T., True, W. 2007. The role of childhood risk factors in initiation of alcohol use and progression to alcohol dependence. *Addiction*, 102, 216–225.

Savette, M., Martin, C., Wertz, J., Perrott, M., Peters, A. 2005. The effect of alcohol on cigarette craving in heavy smokers and tobacco chippers. *Psychological Addiction Behavior*, 19, 263–270.

Shield, K.D., Gmel, G., Gmel, G., Mäkelä, P., Probst, C., Room, R., Rehm, J. 2017. Life-time risk of mortality due to different levels of alcohol consumption in seven European countries: Implications for low-risk drinking guidelines. *Addiction*, 112, 1535–1544.

Sipilä, P., Rose, R. J., Kaprio, J. 2016. Drinking and mortality: Long-term follow-up of drinking-discordant twin pairs. *Addiction*, 111, 245–254.

Steel, N. 2017. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390, 1151–1210.

Sudharsanan, N., Behrman, J., Kohler, H-P. 2016. Limited common origins of multiple adult health-related behaviors: Evidence from U.S. *Social Science and Medicine*, 171, 67–83.

Rehm, J., Gmel, G.E., Gmel, G., et al. 2017. The relationship between different dimensions of alcohol use and the burden of disease – an update. *Addiction*, 112, 968–1001.

Taioli, E., Wynder, E.L. 1991. Effect of the age at which smoking begins on frequency of smoking in adulthood. *New England Journal of Medicine*, 325, 968–969.

True, W., Xian, H., Scherrer, J., Madden, P., Bucholz, K., Heath, A., Eisen, S., Lyons, M., Goldberg, J., Tsuang, M. 1999. Common genetic vulnerability for nicotine and alcohol dependence in men. *Archives of General Psychiatry*, 56, 655–661.

Ussher, M., Taylor, A., Faulkner, G. 2014. Exercise interventions for smoking cessation. *Cochrane database of Systematic Reviews*. doi: 10.1002/14651858.CD002295.pub5

Waller, K., Vaha-Ypya, H., Tormakangas, T., Hautasaari, P., Lindgren, N., Iso-Markku, P., Heikkila, K., Rinne, J., Kaprio, J., Sievanen, H., Kujala, U.M. 2017. Long-term leisure-time physical activity and other health habits as predictors of objectively monitored late-life physical activity – a 40-year twin study. *BioRxiv* 205856; doi: <https://doi.org/10.1101/205856>

Wang-Hong, X., Xiang-Lan, Z., Yu-Tang, G., Young-Bing, X., Li-Feng, G., Wei, Z., Xiao-Ou, S. 2007. Joint effect of cigarette smoking and alcohol consumption on mortality. *Preventive Medicine*, 45, 313–319.

World Health Organization. 2010. *Global Recommendations on Physical Activity for Health*. Geneva: The WHO Press.

**Table 1.** Description of raw data: Measures of risky health behavior by gender.

	Men		Women		Difference	p-value
	% of total	N	% of total	N		
Heavy alcohol use (H)	0.13	533	0.02	82	0.11	<0.01
Smoking (S)	0.08	359	0.10	519	-0.02	<0.01
Physical inactivity (P)	0.15	634	0.22	1117	-0.07	<0.01
Heavy alcohol use and smoking (HS)	0.10	435	0.02	91	0.08	<0.01
Heavy alcohol use and physically inactive (HP)	0.04	192	0.01	28	0.03	<0.01
Smoking and physically inactive (SP)	0.05	218	0.05	253	0.00	0.99
All three (HSP)	0.06	244	0.01	47	0.05	<0.01
None of the above: No risky health behavior	0.39	1697	0.57	2859	-0.18	<0.01
Total N	1.00	4312	1.00	4996		

Notes: Mutually exclusive categories. Pooled data on MZ and DZ twins.



**Table 2.** Regression results: OLS and within DZ-MZ and MZ estimates for men.

	Reverse rank order of earnings			Unemployment months		
	(1) OLS	(2) DZ-MZ	(3) MZ	(4) OLS	(5) DZ-MZ	(6) MZ
Direct terms						
H	0.88 (1.35)	2.19 (1.76)	-0.35 (3.18)	0.26*** (0.10)	0.31** (0.14)	-0.31 (0.26)
S	7.16*** (1.54)	4.03* (2.09)	3.58 (3.57)	0.72*** (0.147)	0.53*** (0.19)	0.13 (0.30)
P	1.88 (1.27)	-0.42 (1.48)	3.54 (2.30)	-0.001 (0.084)	0.02 (0.11)	0.29* (0.17)
Two-way interactions						
HS	11.74*** (1.54)	12.12*** (2.13)	13.29*** (3.52)	1.23*** (0.15)	0.90*** (0.17)	0.81** (0.32)
HP	4.40** (1.91)	2.71 (2.73)	3.43 (5.66)	0.56*** (0.18)	0.45* (0.57)	-0.16 (0.54)
SP	7.69*** (1.84)	5.37** (2.39)	12.31*** (3.65)	0.30** (0.14)	0.28 (0.20)	0.37 (0.28)
Three-way interaction	12.27***	8.18***	12.38***	1.32***	1.11***	0.84**
HSP	(1.95)	(2.57)	(4.46)	(0.20)	(0.30)	(0.41)
Controls	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.18	0.07	0.07	0.06	0.04	0.04
N	4312	2156	684	4312	2156	684
Tests (F-test)						
	p-value	p-value	p-value	p-value	p-value	p-value
T1: $\alpha_H = \alpha_S = \alpha_P = 0$	<0.01	0.18	0.34	<0.01	<0.01	0.17
T2: $\alpha_H = \alpha_{HS}$	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
T3: $\alpha_S = \alpha_{HS}$	<0.05	<0.01	<0.05	<0.05	0.15	0.11
T4: $\alpha_H + \alpha_S = \alpha_{HS}$	0.11	<0.05	<0.05	0.25	0.82	<0.05
T5: $\alpha_H = \alpha_{HP}$	<0.10	0.84	0.55	0.12	0.57	0.80
T6: $\alpha_P = \alpha_{HP}$	0.23	0.28	0.98	<0.01	<0.10	0.39
T7: $\alpha_H + \alpha_P = \alpha_{HP}$	0.51	0.76	0.97	0.16	0.66	0.83
T8: $\alpha_S = \alpha_{SP}$	0.81	0.61	<0.10	<0.05	0.29	0.52
T9: $\alpha_P = \alpha_{SP}$	<0.01	<0.05	<0.05	<0.05	0.19	0.78
T10: $\alpha_S + \alpha_P = \alpha_{SP}$	0.58	0.55	0.27	<0.05	0.28	0.89
T11: $\alpha_H + \alpha_S + \alpha_P = \alpha_{HSP}$	0.44	0.52	0.39	<0.05	0.45	0.18

Notes: \*\*\* ( $p < 0.01$ ), \*\* ( $p < 0.05$ ), \* ( $p < 0.10$ ). The controls include BMI, height, the number of diagnosed chronic diseases, employment status, earnings, the use of tranquilizers, neuroticism, extraversion and indicators for work disability and physical work, as measured in 1975. OLS specifications also include age as an additional covariate because it makes the OLS estimates more comparable with the results obtained using twin differences (in which all age effects are eliminated).

**Table 3.** Regression results: OLS and within DZ-MZ estimates for women.

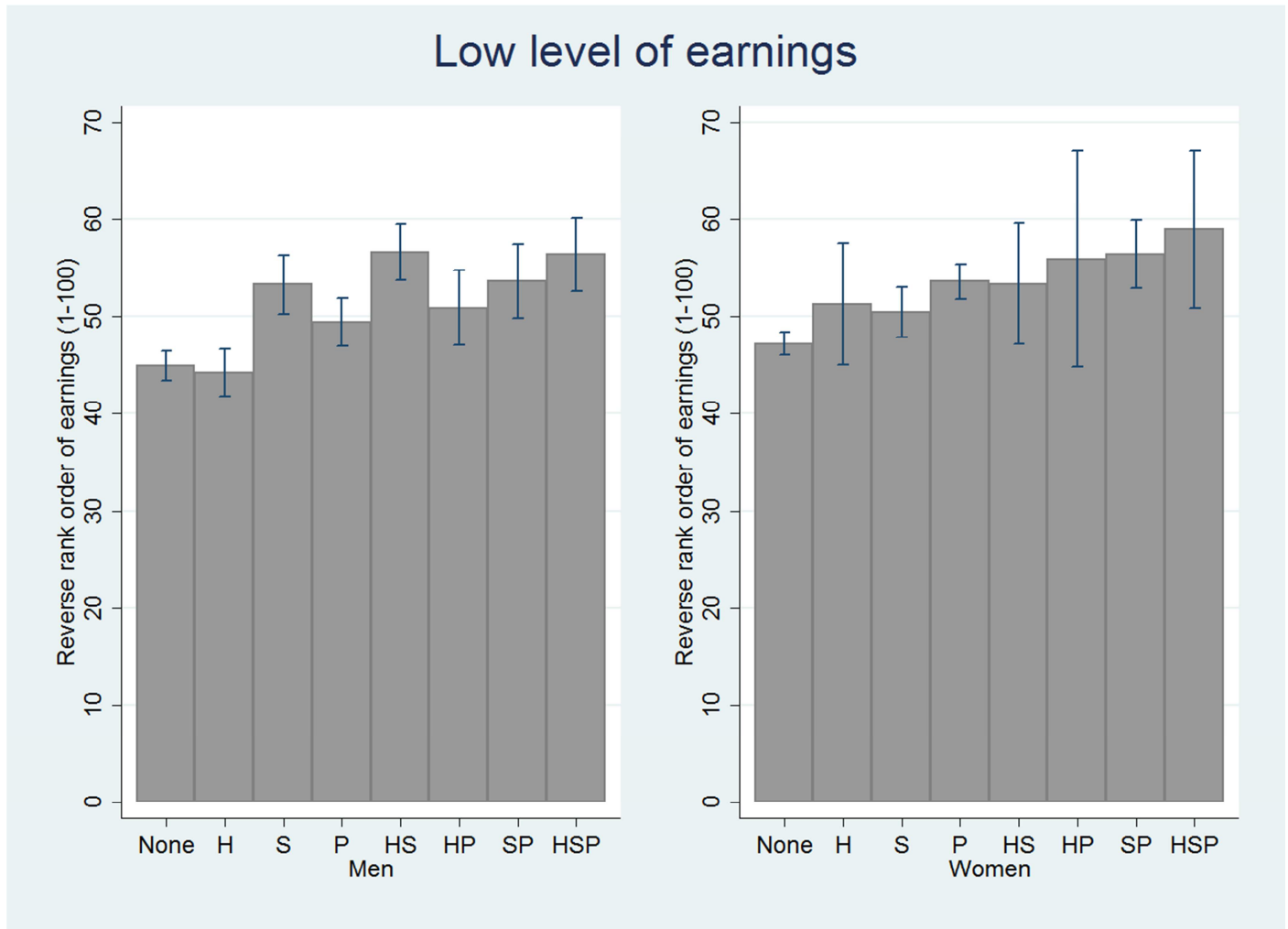
	Reverse rank order of earnings			Unemployment months		
	(1) OLS	(2) DZ-MZ	(3) MZ	(4) OLS	(5) DZ-MZ	(6) MZ
Direct terms						
H	5.76* (3.02)	5.84 (4.26)	5.05 (8.07)	0.58** (0.26)	0.68* (0.35)	-0.01 (0.61)
S	5.16*** (1.36)	3.34** (1.70)	5.07* (2.74)	0.41*** (0.10)	0.25 * (0.14)	0.35 (0.23)
P	5.37*** (0.99)	3.80*** (1.24)	4.03** (1.92)	0.11 (0.07)	0.11 (0.09)	0.20 (0.16)
Two-way interactions						
HS	7.91*** (3.01)	7.91** (3.81)	4.37 (6.64)	0.95*** (0.32)	0.81*** (0.35)	0.77 (0.69)
HP	6.09 (5.64)	4.27 (7.48)	9.88 (8.59)	1.02* (0.55)	0.32 (0.77)	0.17 (0.96)
SP	9.84*** (1.79)	7.46*** (2.29)	4.71 (4.50)	0.64*** (0.15)	0.39** (0.20)	-0.06 (0.30)
Three-way interaction HSP	11.67*** (3.90)	22.64*** (5.13)	15.34** (7.72)	1.17** (0.46)	1.70*** (0.49)	1.96** (0.89)
Controls	YES	YES	YES	YES	YES	YES
R <sup>2</sup>	0.14	0.05	0.05	0.04	0.02	0.04
N	4996	2498	880	4996	2498	880
Tests (F-test)						
T1: $\alpha_H = \alpha_S = \alpha_P = 0$	p-value <0.01	p-value <0.01	p-value <0.10	p-value <0.01	p-value <0.10	p-value 0.32
T2: $\alpha_H = \alpha_{HS}$	0.59	0.70	0.94	0.36	0.77	0.35
T3: $\alpha_S = \alpha_{HS}$	0.39	0.25	0.92	<0.10	0.13	0.57
T4: $\alpha_H + \alpha_S = \alpha_{HS}$	0.47	0.82	0.55	0.91	0.81	0.63
T5: $\alpha_H = \alpha_{HP}$	0.96	0.85	0.63	0.48	0.66	0.87
T6: $\alpha_P = \alpha_{HP}$	0.64	0.95	0.50	0.11	0.78	0.97
T7: $\alpha_H + \alpha_P = \alpha_{HP}$	0.90	0.52	0.94	0.59	0.57	0.98
T8: $\alpha_S = \alpha_{SP}$	<0.05	<0.10	0.94	0.18	0.53	0.21
T9: $\alpha_P = \alpha_{SP}$	<0.05	0.14	0.88	<0.01	0.18	0.42
T10: $\alpha_S + \alpha_P = \alpha_{SP}$	0.76	0.91	0.38	0.53	0.90	<0.10
T11: $\alpha_H + \alpha_S + \alpha_P = \alpha_{HSP}$	0.38	0.16	0.92	0.90	<0.10	0.22

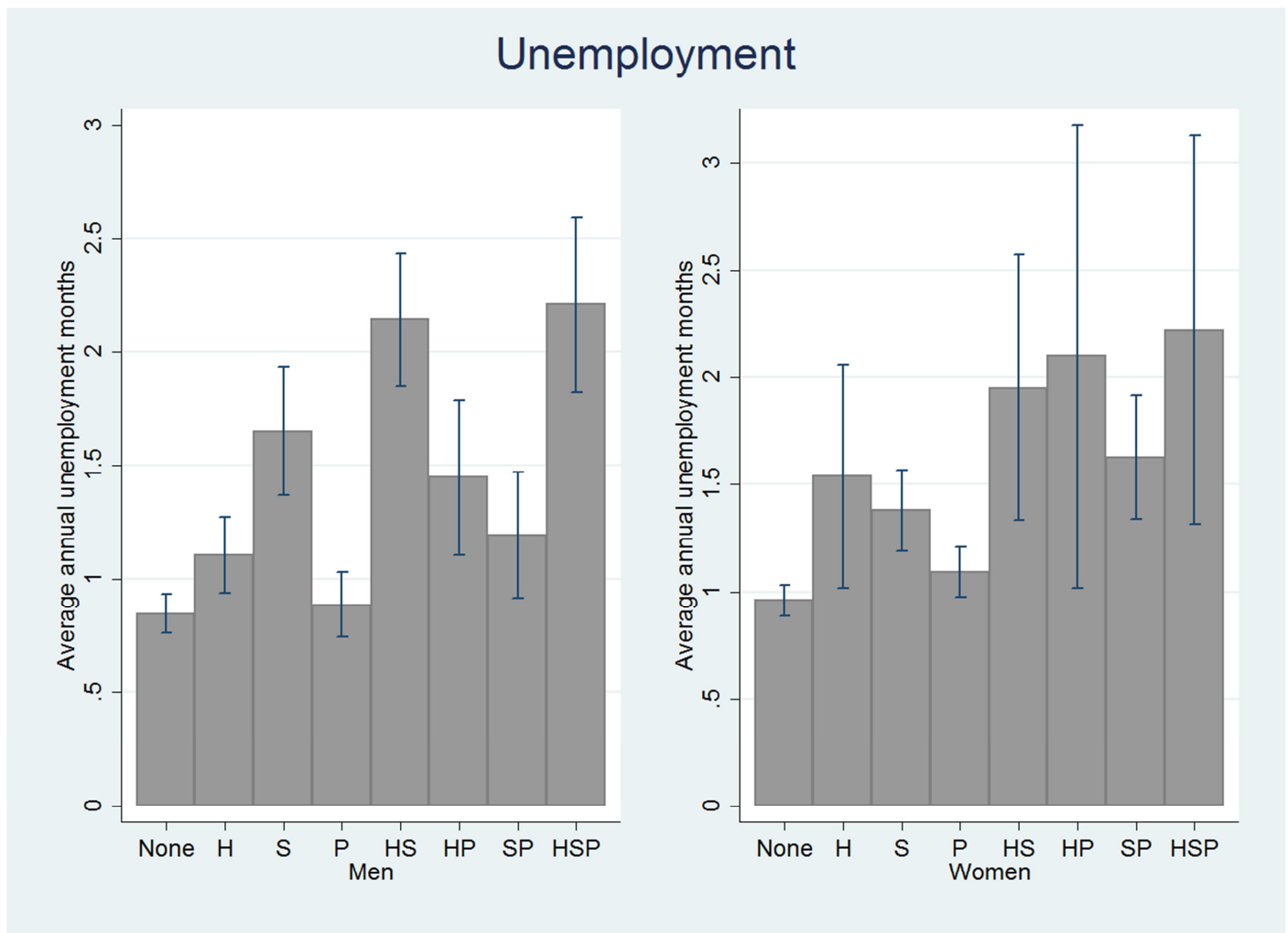
Notes: \*\*\* ( $p < 0.01$ ), \*\* ( $p < 0.05$ ), \* ( $p < 0.10$ ). The controls include BMI, height, the number of diagnosed chronic diseases, employment status, earnings, the use of tranquilizers, neuroticism, extraversion and indicators for work disability and physical work, as measured in 1975. OLS specifications also include age as an additional covariate because it makes the OLS estimates more comparable with the results obtained using twin differences (in which all age effects are eliminated).

**Table 4.** Regression results: Within DZ-MZ and within MZ estimates for the pooled sample of men and women.

	Reverse rank order of income		Unemployment months	
	Column (1) Within DZ-MZ	Column (2) Within MZ	Column (3) Within DZ-MZ	Column (4) Within MZ
Direct terms				
H	3.15 * (1.61)	-0.24 (2.99)	0.36 *** (0.13)	-0.29 (0.25)
S	3.78 *** (1.32)	4.53** (2.21)	0.37 *** (0.11)	0.26 (0.19)
P	2.30 ** (0.95)	3.71** (1.47)	0.07 (0.07)	0.21 * (0.12)
Two-way interactions				
HS	11.97 *** (1.84)	11.19 *** (3.08)	0.89 *** (0.18)	0.83*** (0.30)
HP	3.84 (2.58)	3.20 (5.20)	0.43 * (0.24)	-0.21 (0.51)
SP	6.64 *** (1.66)	8.32*** (2.93)	0.33 ** (0.14)	0.13 (0.21)
Three-way interaction				
HSP	11.50 *** (2.28)	12.14*** (3.81)	1.22 *** (0.21)	1.09*** (0.38)
Controls	YES	YES	YES	YES
R <sup>2</sup>	0.05	0.05	0.03	0.03
N	4654	1564	4654	1564
Tests (F-test)				
	p-value	p-value	p-value	p-value
T1: $\alpha_H = \alpha_S = \alpha_P = 0$	<0.01	<0.05	<0.01	<0.10
T2: $\alpha_H = \alpha_{HS}$	<0.01	<0.01	<0.01	<0.01
T3: $\alpha_S = \alpha_{HS}$	<0.01	<0.05	<0.05	0.11
T4: $\alpha_H + \alpha_S = \alpha_{HS}$	<0.05	<0.10	0.49	<0.05
T5: $\alpha_H = \alpha_{HP}$	0.81	0.55	0.76	0.89
T6: $\alpha_P = \alpha_{HP}$	0.56	0.92	0.14	0.39
T7: $\alpha_H + \alpha_P = \alpha_{HP}$	0.58	0.96	0.99	0.81
T8: $\alpha_S = \alpha_{SP}$	0.11	0.24	0.81	0.58
T9: $\alpha_P = \alpha_{SP}$	<0.05	0.13	<0.10	0.70
T10: $\alpha_S + \alpha_P = \alpha_{SP}$	0.78	0.98	0.51	0.19
T11: $\alpha_H + \alpha_S + \alpha_P = \alpha_{HSP}$	<0.01	0.43	0.27	0.11

Note: OLS estimates using twin-differenced data, using the MZ twins only. \*\*\* ( $p < 0.01$ ), \*\* ( $p < 0.05$ ) and \* ( $p < 0.10$ ). The controls include BMI, height, the number of diagnosed chronic diseases, employment status, earnings, the use of tranquilizers, neuroticism, extraversion and indicators for work disability and physical work, as measured in 1975.

**Figure 1:** Average reverse rank order of lifetime earnings by risky health behaviors.

**Figure 2.** Average unemployment months by risky health behaviors.

Notes (Figures 1-2): Mutually exclusive categories. The height of the bar displays the mean and the vertical lines are the 95% confidence intervals. The sample includes both MZ and DZ twins.

**Highlights**

We examine the joint effects of risky health behaviors using twin design

Health behaviors include alcohol use, smoking and physical inactivity

Long-term labor market outcomes are measured using administrative data

There are significant joint relationships between risky health behaviors

Smoking reinforces the negative effects of alcohol use on long-term earnings and employment