Final draft -- to appear in Tobacco Control

Smoking and long-term labour market outcomes

Petri Böckerman*, Ari Hyytinen** and Jaakko Kaprio***

ABSTRACT

Objective To examine the long-term effects of smoking on labour market outcomes using twin data matched to register-based individual information on earnings. **Method** Twin data for Finnish men born 1945-1957 was used to remove the shared environmental and genetic factors. The results were subjected to extensive robustness testing. Lifetime cigarette consumption was measured by (cumulative) cigarette pack-years in early adulthood. The average of an individual's earnings (and, alternatively, taxable income) was measured over a subsequent 15-year period in later adulthood. **Results** Smokers have lower long-term income and earnings. For example, controlling for the shared environmental and genetic factors using the data on genetically identical twins, smoking is negatively associated with lifetime income (p-value = 0.015). The negative association was also robust to the use of various covariates, such as education, health indicators and extraversion. **Conclusions** Smoking is negatively related to long-term labour market outcomes. The provision of information about the indirect monetary costs of smoking may thus complement the policy efforts that aim at educating consumers about the health costs of smoking.

* Corresponding author. Labour Institute for Economic Research and IZA. Address: Pitkänsillanranta 3A, FI-00530 Helsinki, Finland. Phone: +358-9-2535 7330. Fax: +358-9-2535 7332. E-mail: petri.bockerman@labour.fi
** Jyväskylä University School of Business and Economics, Jyväskylä, Finland. E-mail: ari.t.hyytinen@jyu.fi
*** University of Helsinki, Hjelt Institute, Helsinki, Finland. E-mail: jaakko.kaprio@helsinki.fi

ABSTRACT

Objective To examine the long-term effects of smoking on labour market outcomes using twin data matched to register-based individual information on earnings. **Method** Twin data for Finnish men born 1945-1957 was used to remove the shared environmental and genetic factors. The results were subjected to extensive robustness testing. Lifetime cigarette consumption was measured by (cumulative) cigarette pack-years in early adulthood. The average of an individual's earnings (and, alternatively, taxable income) was measured over a subsequent 15-year period in later adulthood. **Results** Smokers have lower long-term income and earnings. For example, controlling for the shared environmental and genetic factors using the data on genetically identical twins, smoking is negatively associated with lifetime income (p-value = 0.015). The negative association was also robust to the use of various covariates, such as education, health indicators and extraversion. **Conclusions** Smoking is negatively related to long-term labour market outcomes. The provision of information about the indirect monetary costs of smoking may thus complement the policy efforts that aim at educating consumers about the health costs of smoking.

Cigarette smoking is among the three leading risk factors for the global disease burden[1] and one of the most important preventable causes of premature death.[2] And yet, despite the adverse health consequences of smoking, the literature is inconclusive on whether continued adult smoking reflects rational, imperfectly rational or irrational behaviour.[3-4] Rational smokers continue cigarette consumption because of its current benefits relative to the health risks and costs and/or because of the physiological and psychological costs of quitting. Imperfectly rational smokers may continue smoking, because they suffer from, e.g., biased beliefs about the harms of smoking, present-biased preferences or the inability to execute their quitting plans. The behaviour of irrational smokers is, in turn, driven by emotions, external cues and impulsive behaviour. It is successively harder to reconcile continued smoking with forward-looking rationality, the more evidence there is on the costs of smoking. If smoking turns out to have high indirect monetary costs in addition to the out-of-pocket costs of cigarette purchases and its adverse health impacts, there is less scope for smoking to be rational. In this paper, we therefore focus on documenting the consequences of smoking on *long-term* labour market outcomes.

According to the early US evidence, current smokers earn 1-7% less than those who do not smoke.[5-6] The cross-sectional wage gap was mostly driven by those who continue smoking.[7] Unobserved heterogeneity may also matter a lot for the results.[8] Using a cross-sectional survey from the Netherlands, a 10% wage gap was reported while taking into account unobserved heterogeneity.[9] A study using Canadian data, in turn, found that smokers earn 8% less than non-smokers.[10]

We contribute to the debate in several ways. First, identification of the effect of smoking is challenging, because there are unobservable factors that are correlated both with smoking and

the outcomes, such as earnings. This problem implies that the OLS estimation does not produce an unbiased effect of smoking on earnings. We address this problem by using data on twins.[11] It allows us to better control both for shared environmental factors, such as family background, neighborhood and peer effects,[12-14] and for genetic factors, which are determinants of time, risk and other preferences and personality traits. Using data on non-identical (dizygotic, DZ) twins is the same as controlling for sibling effects, because DZ twins originate from the same family and neighbourhood and share, on average, the same amount (50%) of segregating genes as ordinary siblings do. Using data on identical (monozygotic, MZ) twins allows us to further control for inherited traits and preferences, because two MZ twins are genetically identical at the sequence level.

Second, a challenge that the earlier studies have not addressed is that self-reported annual earnings, or equivalent cross-sectional measures, are only poor proxies for lifetime earnings.[15-16] Our sample consists of twin pairs for whom we observe accurate administrative data on their prime working-age earnings. Unlike the prior work, we can use the average of an individual's taxable income and, alternatively, wage and salary earnings over the 15-year period as a measure for lifetime earnings. Using this average value reduces measurement error and it is not prone to non-response and reporting biases.

Third, many earlier studies have used self-reported information on current smoking status as the main explanatory variable. This approach is problematic for two reasons. The comparison group includes not only individuals who have never smoked but also former smokers, and the negative health effects of cigarette consumption may take a long time to develop.[17] We depart from earlier research and use a measure of cumulative cigarette consumption in early adulthood.

Fourth, we complement the literature on smoking and (short-term) absenteeism from work.[18] We examine whether the relationship between cigarette consumption and labour market activity continues to exist when a longer-term measure of individuals' labour market attachment is used.

METHODS

Data sources and the sample

Our twin sample data is based on the Older Finnish Twin Cohort Study (of the Department of Public Health in University of Helsinki), which we linked to the Finnish Longitudinal Employer-Employee Data (FLEED) of Statistics Finland. The twin cohort data and the linked data have been used previously,[19-20] so the prior studies can be consulted for details about, e.g., overall response rates and attrition.

The Finnish Cohort Study was initially compiled from the Central Population Registry of Finland. Initial twin candidates were persons born before 1958 with the same birth date, commune of birth, sex, and surname at birth.[20] A questionnaire was mailed to these candidates in 1975 to collect baseline data and to determine their zygosity. Two follow-up surveys were conducted in 1981 and 1990. We linked the twin data to FLEED using personal identifiers. FLEED includes information on individuals' labour market status, and salaries and other income, taken directly from tax and other administrative registers that are collected and/or maintained by Statistics Finland. Such data do not suffer from underreporting or recall error, nor is it top coded.

Our analysis focuses on men for two reasons. First, men are more strongly attached to the labour market. Moreover, male labour supply decisions are much less affected by family and fertility choices.[21] Second, the smoking rate has been much higher among men, especially among older age cohorts.[22]

To prevent early retirement from affecting our lifetime outcome measures, we further restricted the analysis to primary working-age persons. The estimating sample was therefore restricted to individuals who were born after 1944 but before 1958. Accordingly, the twins were aged 33-59 years over the measurement period of 1990-2004.

Measures

Our proxy for the lifetime income is the logarithm of the average of annual taxable income over the period of 1990-2004. It is a broad income concept, which includes annual wage and salary earnings, self-employment income and capital income (dividends, capital gains). It also includes income transfers and social security benefits, such as unemployment and parental leave benefits, which are often proportional to past wage and salary earnings. The proxy for the lifetime earnings is the logarithm of the average of annual wage and salary earnings over the period of 1990-2004. This income concept is narrower than our first measure, the lifetime income.

Our measure for smoking is self-reported retrospective cigarette pack-years, as measured in the 1981 twin survey. We point out three things about this measure. First, it is predetermined. This is useful, because otherwise there might be a problem of simultaneity between smoking and earnings due to the positive income elasticity of cigarette consumption.[4] Second, this

measure allows for the potential delay in the adverse effects of smoking. Third, cigarette pack-years capture the *cumulative* lifetime consumption of cigarettes, as they were calculated as follows: cigarette pack-years = average number of cigarettes smoked per day × person's age – age when the person started smoking. (Mean = 5.99, standard deviation = 7.31.) For example, a person has a 20 pack-year history of smoking if he has smoked one pack of cigarettes daily for 20 years. This information has been used in earlier research.[23] While not perfect, the medical literature has utilized cigarette pack-years and it is related to smoking-related diseases.[24] Because our response variables describe lifetime labour market outcomes, it is convenient to have a measure for the consumption of cigarettes that is capable of capturing an individual's cumulative smoking by his early adulthood (i.e., age 24-37).

Table 1 reports average lifetime income and earnings in euros, conditional on cigarette pack-years (Panel A) as well as on the current (i.e., at the time of the survey) smoking status in 1981 (Panel B) and in 1975 and 1981 (Panel C). Panel A reveals that persons with more than 10 cigarette pack-years earn, on average, less than those who have not smoked at all. Additionally, lifetime income is lower for smokers, but the difference between smokers and non-smokers is smaller. Panel B shows that when we condition on the smoking status in 1981, lifetime earnings and income are lowest for those who were current smokers then. Panel C reveals, in turn, that lifetime income and earnings are lowest for those who were smokers both in 1975 and 1981, as compared to the other groups. The null hypothesis of equal group means was rejected in all cases (p-value < 0.001).

[Table 1 here]

Statistical methods

We used four different types of regressions. First, we used OLS to regress our lifetime income and earnings measures on the cigarette pack-years in 1981 for a combined sample of DZ and MZ twin individuals. Second, we took twin differences and re-ran the same regression using the same combined sample. In this twin-differenced model, all factors that two twins share (i.e., the shared environmental factors, business cycle effects and age) are eliminated. Third, we repeated the previous within-twin pair regression using the (smaller) DZ sample. Finally, we ran the within-twin pair regression using the MZ sample. Both the shared environmental and genetic factors are differenced out in this twin-differenced model.

The baseline regression models do not include control variables. To assess the sensitivity of our baseline results, we estimated models that included controls for, e.g., education and health indicators; see the sub-section on robustness checks.

The prior medical and epidemiology literature[25-27] has established that smoking causes several health problems. The earlier results using the same Finnish twin data on which we build our analysis support this conclusion.[28-29] We confirmed that smoking was negatively associated with health status also in our *particular* estimating sample.

RESULTS

Main results – long-term income and earnings

The baseline estimates using the standard OLS specifications (Table 2, Panel A and B, column 1) show that smoking is negatively associated with lifetime income (p-value < 0.001) and earnings (p-value < 0.001). The coefficient of smoking is larger (in absolute value) when lifetime earnings are used, which is line with the view that smoking correlates with poorer health outcomes and that lifetime income includes elements of social insurance. The OLS results are consistent with the previous studies reporting the negative effects of smoking on earnings.[6, 9-10]

[Table 2 here]

The picture does not change much when we focus on the twin-differenced DZ-MZ model (Panel A and B, column 2) that controls for the shared environment. Even though the coefficients are slightly smaller in absolute value than the OLS estimates, the negative relationship between cigarette consumption and lifetime income (or earnings) remains statistically significant (with p-values < 0.001 and 0.002, respectively). The results for the smaller DZ sample (Panel A and B, column 3) confirm these findings. Finally, the within-MZ twin-pair regressions (Panel A and B, column 4) show that smoking is negatively associated with lifetime income (p-value = 0.015, 95% CIs [-0.025, -0.003]) and earnings (p-value = 0.058, 95% CIs [-0.038, 0.001]) even when the shared environmental and genetic factors are controlled for.

The quantitative magnitude of the within-MZ estimates for lifetime income and earnings is not negligible. For example, the estimates suggest that if one reduces smoking by an amount that parallels five pack-years, it would be associated with an income increase of ~7% (= 5 × 0.0138). Because the average annual income is ~24000 euros, this corresponds to an increase of ~1700 euros. Interestingly, this is roughly equivalent to an income increase owing to one more year of schooling.

Notably, the size of the estimated coefficients is smaller in absolute value in the DZ sample than in the MZ sample. This result indicates that further controlling for the genetic factors leads to a more negative estimate. There can be many explanations for the difference between the DZ and MZ estimates. One of them is that smoking and risk preferences are correlated[30] and that wage growth may be higher for individuals with a greater preference for risk-taking.[31] Smokers may also be more present-oriented,[3] which could lead to more short-sighted choices in the labour market. If the risk and/or time preferences are even partially genetically inherited, they are better differenced out in the MZ sample than in the DZ sample, leading to an upward bias in the DZ estimate. Biases such as this may explain why some earlier studies did not find robust negative effects of smoking on earnings.

Robustness checks

Additional covariates

The baseline models of Table 2 did not include control variables, because the use of twin differences already controls for many potentially confounding factors. Our baseline results are nevertheless robust to the addition of various controls (not reported in tables): First, we added education years as a control. It obtained a positive and significant coefficient, but the results for smoking remained intact. Second, we added several indicators of health and health behaviour as controls. The measures were taken from the twin survey in 1981 and they included body mass index, self-reported poor health, and an indicator for heavy alcohol consumption. The last of these indicators is included as a control, because there is evidence that alcohol consumption and cigarette consumption are jointly determined.[9] While our earlier conclusions were supported, these results have to be treated with some caution, because these new controls are not as likely to be predetermined and may thus capture some of the effects of smoking on lifetime earnings/income. Third, we added the number of chronic diseases (as measured in the 1975 survey) to the set of controls to account for pre-existing health conditions. The number of different chronic diseases in 1975 is negatively associated with earnings and employment over the period of 1990-2004. However, the inclusion of chronic diseases did not change the effect of smoking on earnings.

As a final additional control, we considered extraversion, which is arguably correlated with smoking,[32] labour market outcomes[33] and risk-taking.[34] The relationship between lifetime income (or earnings) and smoking might therefore change when a measure of extraversion is added to our baseline models. Extraversion was measured using a short form of the Eysenck

Personality Inventory, the EPQ-E scale (containing 9 of the original items) in 1981.[35-36] The effect of cigarette pack-years in 1981 on both lifetime income and earnings remained negative and statistically significant at the 5% level or better in all models (Table 3).

[Table 3 here]

Business cycle effects

The effects of smoking on labour market outcomes may be contingent on the macroeconomic environment. We therefore experimented using the yearly incomes for 1990 (peak in the Finnish economic cycle) and 1993 (severe recession) as the dependent variables. The negative effect prevails during both years, but it seemed to be larger during the recession. This finding demonstrates the importance of averaging out the cyclical effects and provides a potential explanation for the variation in the previous estimates that have been estimated using shorter-term measures for earnings.

Auxiliary analysis – labour market attachment

The literature provides robust evidence that smoking is positively associated with (short-term) absenteeism from work.[18] Because labour market attachment is an important determinant of a person's lifetime earnings, it is of interest to explore whether there is also a relationship between smoking and long-term labour market attachment. We therefore studied how employment months, calculated as the average number of employment months per year over the sample period of 1990-2004, and employment years, calculated as the share of employment years over the sample period, are correlated with cigarette pack-years.

Panels A and B of Table 4 report the results from the specifications that correspond exactly to those of Table 2, but using the two employment variables as the response variables. The results show a negative and statistically significant association (p-values < 0.001) between smoking and labour market attachment in the standard OLS regressions (Panel A and B, column 1). This negative relationship can also be observed in the twin differenced data, even though in the smaller MZ sample, the standard errors are somewhat larger. These findings are consistent with the earlier Finnish evidence on short-term absenteeism[37-38] and suggest that the lower lifetime earnings of smokers may at least partly be due to their weaker labour market attachment.

[Table 4 here]

DISCUSSION

This paper used twin data on smoking linked to register-based individual earnings information to examine the long-term effects of smoking on lifetime labour market outcomes. We found that smokers have lower long-term income/earnings. The negative association between cigarette consumption and long-term income/earnings remained statistically significant when the shared environmental and genetic factors were controlled for. The result was also robust to the use of various covariates, such as education, heavy use of alcohol, health status, body mass index, and extraversion. We also found some tentative evidence that the effect may depend on the macroeconomic environment. This is an interesting direction for further research.

A possible limitation of our study is that there are two potentially worrying margins of sample selection. First, the heaviest smokers (with particularly poor labour market outcomes later in life) may have not responded to the 1981 survey. Second, severe diseases caused by intensive smoking could have increased the probability that a person was missing from our estimating sample and, specifically, that the outcome variables from FLEED referring to 1990-2004 would not have been available for him. The scope for the first type of selection was, however, limited, because the response rate to the 1981 twin survey was 84%. Prior analyses using the survey did not find significant selection either.[39] The second margin of selection was also limited, because, for a man not to be included in our analysis, it would have required that he did not earn anything over the period that covered his prime working age. These cases are most likely exceptions, as e.g. smoking-related severe morbidity and mortality among the studied cohort members should have been rare prior to 1990 (because of their age).

Our estimates may be conservative for two reasons. First, retrospective cigarette pack-years may suffer from measurement error.[40] Having (classical) measurement error in an explanatory variable typically leads to a bias toward zero. This observation suggests that better measures, such as the duration of smoking, which appears to have a robust relationship with many smoking-related diseases, could provide stronger estimates.[41] Second, the adverse impacts of smoking on earnings may occur later in working life. However, extending our analysis to the older cohorts is not straightforward, because self-selection of employees to retirement may lead to a biased sample. This could be the case if persons with the highest earnings potential at the end of their working careers are more likely to remain in the labour force. Supporting this, an exploratory analysis with the older cohorts indicated that, in our data, the negative effects become stronger if they are included.

Our analysis does not imply causality, because we cannot conclusively rule out non-causal explanations for the negative association between smoking and income/earnings.[7] For example, a confounding psychological factor may induce one of the twins to smoke and this unmeasured characteristic may also be related to labour market performance. We can, however, conclude that the negative association was *not* driven by the shared environmental and genetic factors. Moreover, there seemed to be a negative association between cigarette consumption and long-term employment. This finding complements the prior evidence on the positive association between smoking and work absenteeism.[18] Interpreted from this perspective, our results support the causal explanations, such as weaker labour market attachment and lower productivity at work owing to the adverse health effects of smoking later in life (*not* captured by our controls for the past health status at the time smoking was measured). The negative association also bears on the debate about the potential beneficial

effects of smoking (nicotine) on cognitive functions.[42] If there are such effects, they appear not to lead to (substantive) positive earnings effects in the long term.

We have argued that it is successively harder to reconcile continued smoking with rationality if smoking is found to have high indirect monetary costs (i.e. lost earnings) *in addition to* its out-of-pocket costs and adverse health impacts. Interpreted from this perspective, our results are not easily reconciled with the view that prolonged smoking is rational, as that would call for relatively high compensating consumption utility (or other benefits) from cigarette usage. Given that most smokers are tobacco-dependent, their addiction hampers their ability to quit and thus act rationally.

Our findings suggest, but do not prove, that the provision of information about the indirect monetary costs of smoking may complement the efforts that aim at educating consumers about the health costs of smoking. This would not deter rational smokers from starting and continuing cigarette consumption, as the standard economic (Becker-Murphy) model[43] predicts that they are already fully aware of all the benefits and costs. However, we conjecture that such information provision might be useful for imperfectly rational (but nevertheless forward-looking) persons, as it is impossible to appreciate the full monetary consequences of continued smoking without having some information on its potentially adverse earnings effects.

Contributors

All authors participated in planning the study. PB and AH were responsible for the data analysis. All authors participated in analyzing the results and writing the paper.

Funding

This work was financially supported by the Academy of Finland (Grant No. 127796). JK also acknowledges support by the Academy of Finland (Grant No. 263278).

Competing interests None

Data sharing statement The data used in this study are confidential but other researchers can obtain access to it for replication purposes at the Research Laboratory of the Business Structures Unit of Statistics Finland. Obtaining access to the data requires approval by the administrators of the twin data (Department of Public Health, University of Helsinki) and by Statistics Finland.

REFERENCES

1. **Lim SS,** Vos T, Flaxman AD, *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2224-2260.

2. **Sloan FA,** Ostermann J, Picone G, *et al. The Price of Smoking.* Cambridge, MA: The MIT Press; 2004.

3. **Sloan FA,** Wang Y. Economic theory and evidence on smoking behaviour of adults. *Addiction* 2008; **103**: 1777-1785.

4. **Cawley J,** Ruhm C.J. 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; 2012, p. 95-199.

5. **Leigh JP,** Berger M. Effects of smoking and being overweight on current earnings. *Am J Prev Med* 1989; **5**: 8-14.

6. **Levine PB,** Gustafson TA, Velenchik AD. More bad news for smokers? The effects of cigarette smoking on wages. *Ind Labour Relat Rev* 1997; **50**: 493-509.

7. **Grafova I,** Stafford FP. The wage effects of personal smoking history. *Ind Labour Relat Rev* 2009; **62**: 381-393.

8. **Neumann T.** The effect of drinking and smoking on the labour market outcomes of low-income young adults. *Appl Econ* 2013; **45**: 541-553.

9. **Van Ours JC.** A pint a day raises a man's pay; but smoking blows that gain away. *J Health Econ* 2004; **23**: 863-886.

10. **Auld MC.** Smoking, drinking, and income. *J Hum Resour* 2005; **40**: 505-518.

11. **van Dongen J,** Slagboom PE, Draisma HH, *et al.* The continuing value of twin studies in the omics era. *Nat. Rev. Genet* 2012; **13**: 640-653.

12. **Clark AE,** Loheac Y. "It wasn't me, it was them!" Social influence in risky behavior by adolescents. *J Health Econ* 2007; **26**: 763-784.

13. **Liang L,** Chaloupka F, Nichter M, *et al.* Prices, policies and youth smoking. *Addiction* 2003; **98**: 105-122.

14. **Harakeh Z,** Vollebergh WAM. Young adult smoking in peer groups: an experimental observational study. *Nicotine Tob Res* 2013; **15**: 656-661.

15. **Haider S,** Solon G. Life-cycle variation in the association between current and lifetime earnings. *Am Econ Rev* 2006; **96**: 1308-1320.

16. **Böhlmark A,** Lindquist MJ. Life-cycle variations in the association between current and lifetime income: replication and extension for Sweden. *J Labour Econ* 2006; **24**: 879-900.

17. **Kuh D,** Ben-Shlomo Y. *A Life Course Approach to Chronic Disease Epidemiology.* Oxford: Oxford University Press; 2004.

18. **Weng SF,** Ali S, Leonardi-Bee J. Smoking and absence from work: systematic review and meta-analysis of occupational studies. *Addiction* 2012; **108**: 307-319.

19. **Hyytinen A,** Ilmakunnas P, Toivanen O. The return-to-entrepreneurship puzzle. *Labour Econ* 2013; **20**: 57-67.

20. **Kaprio J,** Artimo M, Sarna S, *et al. The Finnish twin registry: Baseline characteristics. Section I: Materials methods, representativeness and results for variables special to twin studies.* Helsinki: Department of Public Health, Publications, M(47); 1979.

21. **Killingsworth MR,** Heckman JJ. Female labour supply: a survey. In: Ashenfelter O., Layard R., editors. *Handbook of Labour Economics, Volume 1.* Amsterdam: Elsevier; 1987, p. 103-204.

22. **Official Statistics of Finland.** *Tobacco Statistics 2010.* Helsinki: National Institute for Health and Welfare, Statistical Report No. 44; 2010.

23. **Huovinen E,** Kaprio J, Vesterinen E, *et al.* Mortality of adults with asthma: a prospective cohort study. *Thorax* 1997; **52**: 49-54.

24. **Lubin JH,** Alavanja MCR, Caporaso N, *et al.* Cigarette smoking and cancer risk: modeling total exposure and intensity. *Am J Epidemiol* 2007; **166**: 479-489.

25. **Shinton R,** Beevers G. Meta-analysis of relation between cigarette smoking and stroke. *BMJ* 1989; **298**: 789-794.

26. **Gandini S,** Botteri E, Iodice S, *et al.* Tobacco smoking and cancer: a meta-analysis. *Int J Cancer* 2008; **122**: 155-164.

27. **Shavelle RM,** Paculdo DR, Strauss DJ, *et al.* Smoking habit and mortality: a meta-analysis. *J Insur Med* 2008; **40**: 170-178.

28. **Kaprio J,** Koskenvuo M. Twins, smoking and mortality: a 12-year prospective study of smoking-discordant twin pairs. *Soc Sci Med* 1989; **29**: 1083-1089.

29. **Kujala UM,** Kaprio J, Koskenvuo M. Modifiable risk factors as predictors of all-cause mortality: the roles of genetics and childhood environment. *Am J Epidemiol* 2002; **156**: 985-993.

30. **Viscusi WK,** Hersch J. Cigarette smokers as job risk takers. *Rev Econ Stat* 2001; **83**: 269-280.

31. **Shaw KL.** An empirical analysis of risk aversion and income growth. *J Labour Econ* 1996; **14**: 626-653.

32. **Van Loon AJM,** Tijhuis M, Surtees PG, *et al.* Determinants of smoking status: cross-sectional data on smoking and initiation and cessation. *Eur J Public Health* 2005; **15**: 256-261.

33. **Fletcher J.M.** The effects of personality traits on adult labour market outcomes: evidence from siblings. *J Econ Behav Organ* 2013; **89**: 122-135.

34. **Lauriola M,** Levin IP. Personality traits and risky decision-making in a controlled experimental task: an exploratory study. *Pers Indiv Differ* 2001; **31**: 215-226.

35. **Eysenck HJ,** Eysenck SBG. *Manual of the Eysenck Personality Questionnaire.* London: Hodder and Stoughton; 1975.

36. **Rose RJ,** Koskenvuo M, Kaprio J, *et al.* Shared genes, shared experiences, and similarity of personality: data from 14,288 adult Finnish co-twins. *Pers Soc Psychol* 1988; **54**: 161-171.

37. **Kivimäki M,** Vahtera J, Thomson L, *et al.* Psychosocial factors predicting employee sickness absence during economic decline. *J Appl Soc Psychol* 1997; **82**: 858-872.

38. **Laaksonen M,** Piha K, Martikainen P, *et al.* Health-related behaviours and sickness absence from work. *Occup Environ Med* 2009; **66**: 840-847.

39. **Kaprio J.** The Finnish Twin Cohort Study: an update. *Twin Res Hum Genet* 2013; **8**: 1-6.

40. **Bernaards CM,** Twisk JW, Snel J, *et al.* Is calculating pack-years retrospectively a valid method to estimate life-time tobacco smoking? A comparison between prospectively calculated pack-years and retrospectively calculated pack-years. *Addiction* 2001; **96**: 1653-1661.

41. **Flanders WD,** Lally CA, Zhu BP, *et al.* Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: results from Cancer Prevention Study II. *Cancer Res.* 2003; **63**: 6556-62.

42. **Sakurai Y,** Kanazawa I. Acute effects of cigarettes in non-deprived smokers on memory, calculation and executive functions. *Hum Psychopharm Clin* 2002; **17**: 369-373.

43. **Becker GS,** Murphy KM. A theory of rational addiction. *J Polit Econ* 1988; **96**: 675-700.

Table 1 Smoking and lifetime earnings and income

	Panel A: Cigarette pack-years (1981)		
	%-Share	Lifetime income (€)	Lifetime earnings (€)
Cigarette pack-years = 0	29.23	26593.69	24157.24
10 > Cigarette pack-years > 0	35.62	25060.39	22402.88
Cigarette pack-years ≥ 10	35.15	21664.46	18268.51
F-test statistics		71.57	95.25
		(p-value < 0.001)	(p-value < 0.001)

	Panel B: Current smoking status (1981)		
	%-Share	Lifetime income (€)	Lifetime earnings (€)
Current smoker	35.23	23003.95	19909.78
Former smoker	25.25	26789.99	24040.82
Never smoker	35.53	26869.62	24505.84
Occasional smoker	3.99	28396.36	25856.01
F-test statistics		25.58	31.14
		(p-value < 0.001)	(p-value < 0.001)

	Panel C: Current smoking status (1975 and 1981)		
	%-Share	Lifetime income (€)	Lifetime earnings (€)
Smoker 1975, non-smoker 1981	12.87	25516.19	22649.36
Non-smoker 1975, smoker 1981	6.47	24418.69	21572.02
Smoker 1975 and 1981	28.24	22869.98	19735.33
Non-smoker 1975 and 1981	52.42	27285.47	24875.61
F-test statistics		24.53	31.22
		(p-value < 0.001)	(p-value < 0.001)

Notes: The data consist of twin men born 1945-1957. Income and earnings measures for the period 1990-2004 are deflated using the consumer price index (base year 2000). Heteroscedasticity-robust F-test statistics for the null hypothesis of equal group means is reported.

Table 2 Smoking and long-term income and earnings

Panel A: Lifetime income

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Log income	Log income	Log income	Log income
Cigarette pack-years	-0.0099***	-0.0097***	-0.0087***	-0.0138**
	(0.0014)	(0.0024)	(0.0027)	(0.0057)
	[-0.0126, -0.0072]	[-0.0144, -0.0049]	[-0.0140, -0.0034]	[-0.0250, -0.0027]
N	3914	1957	1350	607

Panel B: Lifetime earnings

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Log earnings	Log earnings	Log earnings	Log earnings
Cigarette pack-years	-0.0216***	-0.0141***	-0.0130***	-0.0185*
	(0.0027)	(0.0044)	(0.0050)	(0.0097)
	[-0.0268, -0.0163]	[-0.0228, -0.0054]	[-0.0228, -0.0032]	[-0.0375, 0.0006]
N	3914	1957	1350	607

Notes: The data consist of twin men born 1945-1957. Lifetime income and earnings are measured as the logarithm of the average over the period of 1990-2004. Cigarette pack-years are measured in 1981. Heteroscedasticity-robust standard errors are reported in parentheses: *statistically significant at the .10 level; **at the .05 level; ***at the .01 level. The 95% confidence intervals for the parameter estimates are reported in square brackets.

Table 3 Smoking and long-term income and earnings with extraversion as a covariate

Panel A: Lifetime income

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Log income	Log income	Log income	Log income
Cigarette pack-years	-0.0100*** (0.0014) [-0.0127, -0.0073]	-0.0095*** (0.0024) [-0.0143, -0.0047]	-0.0084*** (0.0027) [-0.0137, -0.0032]	-0.0139** (0.0057) [-0.0251, -0.0027]
Extraversion	YES	YES	YES	YES
N	3898	1949	1344	605

Panel B: Lifetime earnings

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Log earnings	Log earnings	Log earnings	Log earnings
Cigarette pack-years	-0.0216*** (0.0027) [-0.0268, -0.0163]	-0.0138*** (0.0044) [-0.0226, -0.0051]	-0.0126** (0.0050) [-0.0223, -0.0028]	-0.0188* (0.0097) [-0.0378, 0.0003]
Extraversion	YES	YES	YES	YES
N	3898	1949	1344	605

Notes: The data consist of twin men born 1945-1957. Lifetime income and earnings are measured as the logarithm of the average over the period of 1990-2004. Cigarette pack-years are measured in 1981. Extraversion is measured using a short form of the Eysenck Personality Inventory, the EPQ-E scale (containing 9 of the original items) in 1981. Heteroscedasticity-robust standard errors are reported in parentheses: *statistically significant at the .10 level; **at the .05 level; ***at the .01 level. The 95% confidence intervals for the parameter estimates are reported in square brackets.

Table 4 Smoking and long-term employment

Panel A: Employment months

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Emp. month	Emp. month	Emp. month	Emp. month
Cigarette pack-years	-0.0740***	-0.0532***	-0.0537***	-0.0515
	(0.0082)	(0.0147)	(0.0167)	(0.0316)
	[-0.0901, -0.0579]	[-0.0822, -0.0242]	[-0.0864, -0.0210]	[-0.1135, 0.0106]
N	3914	1957	1350	607

Panel B: Employment years

Sample	DZ and MZ	DZ and MZ	DZ	MZ
Estimation method	OLS	Twin differences	Twin differences	Twin differences
Dependent variable	Emp. year	Emp. year	Emp. year	Emp. year
Cigarette pack-years	-0.0071***	-0.0050***	-0.0050***	-0.0051*
	(0.0007)	(0.0012)	(0.0014)	(0.0026)
	[-0.0084, -0.0057]	[-0.0074, -0.0026]	[-0.0077, -0.0022]	[-0.0103, 0.0001]
N	3914	1957	1350	607

Notes: The data consist of twin men born 1945-1957. Employment months and employment years are measured as the average over the period of 1990-2004. Employment months refer to the number of employment months during each year (as recorded in the state-run pension registers of all legal employment contracts) and employment years to the exact labour market status during the last week of each year. Cigarette pack-years are measured in 1981. Heteroscedasticity-robust standard errors are reported in parentheses: *statistically significant at the .10 level; **at the .05 level; ***at the .01 level. The 95% confidence intervals for the parameter estimates are reported in square brackets.