LEUPHANA
UNIVERSITÄT LÜNEBURG

A Note on Smoking Behavior and Health Risk Taking

Pfeifer, Christian

Published in:
Nordic Journal of Health Economics

DOI:
10.5617/njhe.195

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Link to publication

Citation for pulished version (APA):
Pfeifer, C. (2012). A Note on Smoking Behavior and Health Risk Taking. *Nordic Journal of Health Economics*, *1*(2), 135-151. https://doi.org/10.5617/njhe.195

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 16. Apr.. 2024

A Note on Smoking Behavior and Health Risk Taking

CHRISTIAN PFEIFER * [a) b)]

pfeifer@leuphana.de
[a)] Institute of Economics, Leuphana University Lueneburg, Germany
[b)] Forschungsinstitut zur Zukunft der Arbeit (IZA), Germany

Abstract: This research note presents two economic frameworks to describe the relationship between individual health risk aversion and smoking behavior. Using a large-scale representative data set (GSOEP), direct empirical evidence is found that individuals, who are more health risk taking, are more likely to be smokers and have a higher demand for cigarettes smoked per day. Non-linear specifications of risk taking reveal, however, that the risk effects are only significant for high risk takers. The estimated effects are robust with respect to the econometric methods (two-part vs. Tobit model) and hold also separately for men and women.

Keywords: cigarettes; health; risk aversion; smoking

JEL Classification: D11, D12, D81, I10, J17

1 Introduction

Smoking behavior has recently been analyzed in many studies (e.g., Chaloupka and Warner, 2000; Goel and Nelson, 2006). The interest arises from the high costs smoking imposes on health care systems. The impact of health risk taking preference has however received limited empirical attention and is mostly reduced to the impact of perceived smoking health care risks (e.g., lung cancer, mortality).[1] But especially the effect of risk aversion on smoking behavior, as an example for unhealthy behavior, is important for risk selections of health insurance companies, determination of risk premiums paid for health insurances, and health policy.

In this research note, I present two simple theoretical frameworks, which describe why risk takers might be more likely to be smokers and might have a higher demand for cigarettes than risk averse individuals, in order to derive the estimation framework (Section 2). The research hypotheses are tested with a large-scale representative data set of German households, which contains information about risk taking preferences as well as smoking behavior (Section 3). In the econometric part of this paper, results of the two-part model are compared with the results of the Tobit model. The paper concludes with a short summary of the main findings (Section 4).

[1] The lack in empirical studies might be reasoned by the fact that individual risk aversion cannot be observed in most survey data sets.

2 Theoretical Frameworks and Research Hypothesis

2.1 Smoking Probability

All individuals start naturally as non-smokers and have to decide if they start to smoke. Total utility U_S in equation (1) depends positively on utility from smoking ($u(S)$), health status ($u(H_S)$), and consumption of a numeraire composite normal good ($u(C_S)$). The smoking status S is binary and takes the value one for smokers and zero for non-smokers. The utility from non-smoking is hence $u(S=0)=0$ and from smoking $u(S=1)>0$. The health status of a non-smoker is better than the health status of a smoker and thus $u(H_0)>u(H_1)\geq 0$. Consumption C_S is total labor income Y_S minus income spent for smoking, which is the price of smoking P. Because labor income depends on productivity which is related to health status and health status is negatively affected by smoking, non-smokers earn higher income than smokers ($Y_0>Y_1$). Therefore $u(C_0)=u(Y_0)>0$ and $u(C_1)=u(Y_1-P)\geq 0$. The opportunity cost of smoking are consequently not only lower consumption due to money spent for smoking but also a lower maximum consumption level due to the negative productivity effect of smoking.

$$(1) \qquad U_S = u(S)+u(H_S)+u(C_S) = u(S)+u(H_S)+u(Y_S-PS)$$

A utility maximizing individual starts to smoke if his utility from smoking is larger than his utility from non-smoking.[2] As the individual is a non-smoker when the decision takes place, there are no risks from smoking and no uncertainty about the utility for the non-smoking status as presented in equation (2).

$$(2) \qquad U_0 = u(H_0)+u(Y_0)$$

The utility from smoking involves, on the other hand, health risks from smoking and uncertainties about these risks as well as about negative health and productivity effects from smoking. The expected health status when smoking ($E[H_1]$) can be written as in equation (3) and depends on non-smoking health status (H_0) minus the expected health reduction due to smoking. The expected health reduction is the probability d to get sick from smoking ($d>0$) which reduces health by D ($D>0$). I assume that the probability d and the health reduction D are both objective values (e.g., from scientific research in medicine), which are known and understood by all individuals to the same degree. Heterogeneity arises in this framework due to differences in individual risk taking

[2] Quite similar simple frameworks have been applied earlier by Viscusi (1990) and Liu and Hsieh (1995). Risk in these papers is related to perceived mortality risk from smoking, which is also a result of information and not only of individual risk aversion. The models for cigarette addiction by Chaloupka (1991) and Becker et al. (1994) do not explicitly take into account risk aversion and uncertainty of health and productivity effects of smoking. Since all individuals in my simple theoretical framework start as non-smokers, addiction does not play an explicit role in determining the decision whether to start smoking or not. Implicitly, addiction could be interpreted as a negative health effect from smoking.

preferences ($R \in [0,1]$), which affect subjective judgments of objective risks.[3] Therefore, the same objective health risks are perceived (weighted) differently by individuals with different risk taking preferences when forming their subjective expectations about health and income effects of smoking.[4] Extreme risk takers ($R = 1$) are assumed to perceive no health risks from smoking at all, i.e., health status and income are expected not to change when starting to smoke. Completely risk averse individuals ($R = 0$) perceive the health risks from smoking as extremely high or even infinite, i.e., they expect deadly diseases like cancer which would lead to zero utility from health status and to no labor income if insurance and social assistances are neglected.[5] Whereas the objective mean probability to get a health reduction (*d*) as well as the objective mean degree of the health reductions (*D*) are known to individuals, the individual actual risks and health reductions involve also uncertainties as they depend on exogenous unknown variables (e.g., genes, treatment quality). Therefore, unknown error terms are included for the probability (υ_d) and the health reduction (υ_D).

$$(3) \qquad E\left[H_1\right] = H_0 - (d / R - d + \upsilon_d)(D + \upsilon_D) \quad \text{with} \quad \frac{\partial E\left[H_1\right]}{\partial R} > 0$$

The expected income when smoking in equation (4) depends on the relative change in expected health status because I assume that productivity is the main determinant of income and changes proportionally to health.

$$(4) \qquad E\left[Y_1\right] = \frac{E\left[H_1\right]}{H_0} Y_0 = Y_0 - \frac{(d / R - d + \upsilon_d)(D + \upsilon_D)}{H_0} Y_0 \quad \text{with} \quad \frac{\partial E\left[Y_1\right]}{\partial R} > 0$$

Expected total utility from smoking is then as described in equation (5), which also allows for uncertainty aversion. Risk averse individuals are assumed to dislike uncertainties about the probability and the degree of health reductions due to smoking. This is reflected by the last term of the expected utility function as more risk averse individuals experience a larger expected utility loss if the variance of the error terms υ_d and υ_D is larger ($u(Var\left[\upsilon\right]) < 0$).

---

[3] This assumption is at least partly consistent with previous empirical findings that smokers have lower risk beliefs about the negative health effects of smoking (e.g., lung cancer and mortality risk) and that such risk beliefs negatively affect the probability to smoke (Viscusi, 1990; Liu and Hsieh, 1995; Hammar and Johansson-Stenman, 2004; Lundborg and Lindgren, 2004; Viscusi and Hakes, 2008; Lundborg and Andersson, 2008).

[4] Extreme examples would be that a completely risk averse individual perceives the risk of lung cancer due to smoking as so high that he expects to be dead for sure, whereas the extreme risk taker might not really care about this risk (e.g., a gamble to cheat death). From such different psychological valuations of objective risks follows a difference in expected health status from smoking and, consequently, a difference in the expected smoking effects on income.

[5] The specification of risk aversion *R* in equation (3) results into risk neutrality at $R = 0.5$ because the risk neutral individual perceives the probability *d* as it is. For values $R > 0.5$ individuals underestimate health risks (risk lovers) and for values $R < 0.5$ they overestimate health risks (risk haters).

$$(5)\quad \begin{aligned} E\left[U_1\right] &= u(S) + u(H_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)) \\ &+ u(Y_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)Y_0 / H_0 - P) \\ &+ (1/R - 1)\cdot(u(Var\left[\upsilon_d\right]) + u(Var\left[\upsilon_D\right])) \end{aligned} \quad \text{with} \quad \frac{\partial E\left[U_1\right]}{\partial R} > 0$$

The condition under which an individual starts to smoke is that expected total utility from smoking is larger than total utility from non-smoking as stated in (6).

$$(6)\quad \begin{aligned} &u(S) + u(H_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)) + u(Y_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)Y_0 / H_0 - P) \\ &\quad + (1/R - 1)\cdot(u(Var\left[\upsilon_d\right]) + u(Var\left[\upsilon_D\right])) > u(H_0) + u(Y_0) \end{aligned}$$

Since the primary interest of this research note is to estimate the impact of risk taking on smoking, the condition in (6) can be incorporated into a Probit estimation framework in equation (7), in which Φ denotes the standard normal cumulative density function, α the coefficient of risk taking parameter R, β a vector of coefficients for a vector of control variables X, and ε an idiosyncratic normally distributed error term. Even without further assumptions about the functional forms of $u(\cdot)$, the condition to start smoking is more likely to be met if risk taking is larger. The rationale behind this hypothesis is that risk takers expect lower utility losses from worse health status as well as lower income as consequences of smoking and are not strongly affected by uncertainty aversion.

$$(7)\quad \begin{aligned} \Pr\left[S = 1\right] &= \Pr\begin{bmatrix} u(S) + \varepsilon > u(H_0) - u(H_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)) \\ + u(Y_0) - u(Y_0 - (d/R - d + \upsilon_d)(D + \upsilon_D)Y_0 / H_0 - P) \\ - (1/R - 1)\cdot(u(Var\left[\upsilon_d\right]) + u(Var\left[\upsilon_D\right])) \end{bmatrix} \\ &= \Phi\left(\alpha R + \beta X\right) \end{aligned}$$

Hypothesis 1: $\dfrac{\partial \Pr\left[S = 1\right]}{\partial R} = \phi\left(\alpha R + \beta X\right)\alpha > 0.$

2.2 Demand for Cigarettes

The cost-benefit framework in the previous section does not include the demanded quantity of cigarettes, on which the model in this section has its focus. The consumption decision about the number of smoked cigarettes N is modeled within a simple rational consumer choice model.[6] For simplicity no additional utility from health status and no separate demand for health care is included in the model. Utility U in equation (8) thus depends only on consumption of cigarettes N and a composite good C. All goods are normal goods with decreasing marginal utilities.

[6] The model can also be seen as a simplified version of the health capital model by Grossman (1972). Grossman's basic model for the demand of medical health care does not include uncertainties, which have been later introduced by Dardanoni and Wagstaff (1990) in a static model and by Liljas (1998) in a dynamic model.

$$(8) \qquad U = U(N, C) \quad \text{with} \quad \frac{\partial U}{\partial N} > 0, \frac{\partial^2 U}{\partial N^2} < 0, \frac{\partial U}{\partial C} > 0, \frac{\partial^2 U}{\partial C^2} < 0$$

Let us now turn to the budget constraint in equation (9). The maximum budget is labor income Y, which depends on health as a major determinant of productivity. Because health is negatively affected by smoking, income is also negatively affected by the number of smoked cigarettes. As it does not matter for the subsequent analysis, I avoid further restrictions of the functional form of the effect of smoked cigarettes on health and income.[7] The budget can be spent on the composite good C with the price equal to one and on cigarettes N with the price P.

$$(9) \qquad Y(N) = C + PN \quad \text{with} \quad \frac{\partial Y}{\partial N} < 0$$

The optimal demand condition can be computed using the Lagrangian multiplier. Equation (10) yields the standard condition that the marginal rate of substitution equals the marginal costs. Marginal costs do not only comprise the cigarette price but also the negative health and productivity effect of smoking, i.e., the reduction of the earning capacity and the consequential lower consumption budget from smoking an additional cigarette. Therefore, the demand for cigarettes would be lower if negative productivity effects of smoking are larger.

$$(10) \qquad \frac{\partial U / \partial N}{\partial U / \partial C} = P - \frac{\partial Y}{\partial N}$$

As discussed in the previous section, negative health and productivity effects of smoking are likely to be perceived (weighted) differently between individuals who differ in risk aversion. Higher risk taking preferences are associated with lower subjective perceptions of objective health risks and, consequently, lead to lower expectations about the earnings reductions due to smoking. An extreme risk taker expects no negative effect at all ($E\left[\frac{\partial Y}{\partial N} | R = 1\right] = 0$), whereas the expected effect approaches infinity for a completely risk averse individual ($E\left[\frac{\partial Y}{\partial N} | R = 0\right] = -\infty$).[8] The expected marginal earnings reduction can be written as in equation (11).

$$(11) \qquad E\left[\frac{\partial Y}{\partial N}\right] = \left(\frac{1}{R} - 1\right)\frac{\partial Y}{\partial N} \quad \text{with} \quad E\left[\frac{\partial Y}{\partial N}\right] \leq 0, \frac{\partial E\left[\frac{\partial Y}{\partial N}\right]}{\partial R} > 0$$

The optimal condition in (10) is reformulated accordingly in (12), from which the econometric demand equation (13) is derived, where α denotes the coefficient of risk taking parameter R, β a vector of coefficients for a vector of control variables X, and ε an idiosyncratic normally distributed error term. It can be seen that the expected total costs of

[7] The functional form could be linear, concave, or convex.

[8] Note that this form of specifying R is consistent with the specification in the previous section.

smoking an additional cigarette are lower for risk takers, which increases their demand for cigarettes. Risk taking should thus be positively correlated with the number of smoked cigarettes.

$$(12) \qquad \frac{\partial U/\partial N}{\partial U/\partial C} = P - \left(\frac{1}{R} - 1\right)\frac{\partial Y}{\partial N}$$

$$(13) \qquad N = \alpha R + \beta X + \varepsilon$$

Hypothesis 2: $\frac{\partial N}{\partial R} = \alpha > 0$.

3 Econometric Analysis

3.1 Data, Variables, and Methods

The German Socio-Economic Panel (GSOEP) is a panel survey of private households and persons in Germany (Dohmen et al., 2011). It contains a rather stable set of core questions asked every year (e.g., employment, education, income) and yearly topics with additional detailed questions. The 2004 wave includes questions about smoking behavior, from which the dependent variables are generated, and about health risk taking preferences. I focus on a sample of 13,636 Germans who are aged between 18 and 65 years and have no missing values in the variables used in this analysis. Smokers who are not pure cigarette smokers, i.e., persons who smoke cigars and pipes, are assigned with missing values, because their number of cigarettes cannot be interpreted as their total tobacco consumption.[9]

At first, a binary variable is generated, which takes the value one for a cigarette smoker and a zero for a non-smoker. About one third of all respondents are smokers and men are on average more likely to smoke than women. This binary smoking status variable is used to estimate the smoking probability equation (7) from Section 2.1. A second variable is generated, which comprises the number of cigarettes smoked per day. The average number of cigarettes smoked per day is about 5 for all respondents and about 16 for cigarette smokers. Men smoke on average more cigarettes than women. The number of cigarettes smoked per day is used to estimate the cigarette demand equation (13) from Section 2.2.

The respondents in the survey were also asked to rate their willingness to take risks with respect to health issues and in general. The questions are quite general and do not comprise specific health issues such as the risk of lung cancer from smoking. Risk taking preference questions of this kind have been recently used, for example, in labor economics, development economics, and economics in general (e.g., Dohmen et al., 2011). The questions and answer categories are worded as follows:

Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? (Likert scale, where the value 0 means "risk averse" and the value 10 means "fully prepared to take risks")

[9] This restriction leads to a reduction of the sample by 206 observations or 1.5 percent, respectively. The results for the probability of smoking are virtually identical without this restriction.

People can behave differently in different situations. How would you rate your willingness to take risks in the following areas? How is it with your health? (Likert scale, where the value 0 means "risk averse" and the value 10 means "fully prepared to take risks")

The original values from the eleven answer categories ranging from 0 to 10 have been divided by 10 for the empirical analysis. After this simple transformation the risk taking values lie in the unit interval ($R \in [0,1]$), which is consistent with the presented theoretical frameworks (R=0 for completely risk averse and R=1 for completely risk taking). The empirical results are however not affected by this transformation as it is only a rescaling of the variable values and of the estimated parameters. Average health risk taking is 0.32 in the total sample. Nearly 20 percent of the respondents are completely risk averse in health issues and only a low share is very risk taking. For example, only about one third rates their willingness to take health risks with more than 0.5. Women are on average less risk taking than men. In addition, the difference between health risk taking and general risk taking is computed. More than 60 percent are less risk taking in health issues than in general, while only about 20 percent are equally risk taking and less than 20 percent are more risk taking in health issues.[10]

Three different measures of health risk aversion are applied in the regressions for smoking behavior. In the first specification, I use the quasi-continuous measure of health risk taking lying in the unit interval ($R \in [0,1]$). Second, the difference between continuous health and general risk taking is used ($R^{health} - R^{general} \in [-1,+1]$). The advantage of this approach is that unobserved individual heterogeneity cancels out, which might be correlated with general risk taking. Larger values of this health risk measure indicate again higher willingness to take risks in health issues. The third specification uses a non-linear specification of health risk taking. The eleven answer categories are grouped into six new categories. The reference category are completely risk averse individuals, whereas the five other categories simply pool two original answer categories to one ($R_0 = 0 < R_1 \leq 0.2 < R_2 \leq 0.4 < R_3 \leq 0.6 < R_4 \leq 0.8 < R_5 \leq 1$). In all specifications a set of control variables is included (gender, age, secondary and vocational schooling, household income, employment status). Although I cannot control directly for differences in health knowledge and understanding of health risks from smoking, important variables are included which can be seen as proxies for such knowledge gathering and knowledge process capabilities (e.g., education). Descriptive statistics of all variables are provided together with the estimation results in the subsequent sections (see Table 1 for the total sample and Table 2 for the subsample of smokers).

In order to estimate the probability to be a cigarette smoker and the demand for cigarettes, two econometric models are estimated and compared. The first estimation strategy is the two-part model emphasized by Cragg (1971) as an alternative to the standard Tobit model (Tobit I). The two-part models are often used in the cigarette demand literature because they are more flexible than the Tobit model. A recent paper by Madden (2008) discusses the two-part model versus the sample selection model of the Heckman type (Tobit II) in the context of smoking and drinking. The two-part model estimates in its first part a Probit model for the probability to smoke in the total sample and in its second part a truncated regression for the number of smoked cigarettes using the truncated sample of smokers. Therefore, the two-part model allows to model and estimate separately the decision to smoke at all and the decision of how many cigarettes to smoke.

[10] For a detailed discussion and the validity of the risk measures in the used survey see Dohmen et al. (2011).

This brings the benefit of flexibility so that the determinants of the probability and the determinants of the intensity need not to have the same signs. The same logic applies for Heckman selection models, which need however a hard-to-find excluding restriction (instrument) to estimate the selection equation correctly.

The decision to smoke and the number of cigarettes demanded (on a typical day) might however be a joint decision instead of two separate decisions. Moreover, it can be expected that most determinants have the same signs in both equations; exceptions might be the effect of advertisements and income effects. In the case of risk taking preference, it seems reasonable to assume that higher risk taking preference increases the probability to smoke as well as the demand for cigarettes, as has been shown in the theoretical frameworks. In case of a joint decision the Tobit model is the correct choice from an econometric point of view. A comparison between the Tobit model and the two-part model is provided in the subsequent two sections.

3.2 Results from Two-Part Models

The first part of the two-part model is the binary Probit regression for the probability to smoke cigarettes. The marginal effects are presented in Table 1. The average smoking probability is 32.3 percent in the total sample. Individuals, who are 0.1 point more health risk taking, are on average 2.6 percentage points more likely to smoke. This result also holds in the second specification. Individuals, who are 0.1 point more risk taking in health issues than in general, are on average one percentage point more likely to smoke. The third specification allows to analyze non-linearity of the health risk taking effect. Individuals with low risk taking preference have no significant higher smoking probability than completely risk averse individuals. Individuals with more than average risk taking preference have however a statistical significant higher smoking probability. Risk takers, who rate their willingness to take health risks with 0.5 or 0.6, are about 10 percentage points more likely to smoke than individuals, who rate their risk preference with less than 0.5. Extreme risk takers are even up to 30 percentage points more likely to smoke than individuals who rate their willingness to take risks with less than 0.5.

Separate estimates for men and women reveal virtually same results (see also Table 1). Additional regressions with interaction terms between gender and risk taking show that the small differences between the two gender samples are not significant.[11] The control variables have the highly significant effects on the smoking probability given below: Women have a lower probability to smoke. The impact of age on the smoking probability is inverted u-shaped. Better educated individuals are less likely to smoke. Larger household income is correlated with a lower smoking probability despite of a theoretical income effect that should increase demand. Unemployed individuals are more likely to smoke than employed or non-employed. Employed individuals are more likely to smoke than non-employed.

The second part of the two-part model consists of a truncated regression for the number of smoked cigarettes using the truncated sample of smokers. The marginal effects on the number of smoked cigarettes are presented in Table 2. An average smoker consumes about 16 cigarettes per day. The first specification shows that individuals, who are 0.1 point more health risk taking, consume on average about 0.33 cigarettes more per day. In the second specification, individuals consume on average about 0.22 cigarettes more per day, if they are 0.1 point more risk taking in health issues than in general. The third specification shows again that only individuals with more than average risk taking

[11] The results of these regressions are not presented in this note but can be requested from the author.

preferences have significant higher cigarette consumption level. As was the case in the Probit estimates, no significant differences between men and women with respect to the risk taking effect on smoking behavior is identified (see also Table 2).

The effects of the control variables have mostly the same direction in the truncated as in the Probit regressions, though not always statistical significant in the truncated regressions. An exception is the effect of the household income, which is negative in the selection equation for the probability to smoke and non-significant or even positive in the demand equation for cigarettes. An explanation for this findings might be the addictive nature of cigarettes that seems to prohibit a demand reduction to zero cigarettes in case of negative income shocks, despite of an income effect on the number of consumed cigarettes.

Table 1: Probit regressions for cigarette smoking probability (first part of two-part model)

	Total sample (n=13636)				Men (n=6524)				Women (n=7112)			
	Mean (sd)	(1)	(2)	(3)	Mean (sd)	(1)	(2)	(3)	Mean (sd)	(1)	(2)	(3)
Health risk taking [0,1]	0.3167	0.2631***			0.3556	0.2739***			0.2810	0.2547***		
	(0.2428)	(0.0172)			(0.2460)	(0.0250)			(0.2342)	(0.0235)		
Health risk taking minus	-0.1547		0.1040***		-0.1609		0.1348***		-0.1490		0.0774***	
general risk taking [-1,+1]	(0.2492)		(0.0167)		(0.2515)		(0.0245)		(0.2468)		(0.0229)	
Health risk taking (reference: 0.0)												
Health risk taking (0.1, 0.2)	0.2573			-0.0097	0.2301			0.0138	0.2823			-0.0266*
	(0.4372)			(0.0126)	(0.4209)			(0.0208)	(0.4502)			(0.0156)
Health risk taking (0.3, 0.4)	0.2266			0.0193	0.2333			0.0219	0.2205			0.0196
	(0.4187)			(0.0132)	(0.4230)			(0.0209)	(0.4146)			(0.0169)
Health risk taking (0.5, 0.6)	0.2243			0.1098***	0.2535			0.1133***	0.1974			0.1116***
	(0.4171)			(0.0137)	(0.4351)			(0.0210)	(0.3981)			(0.0182)
Health risk taking (0.7, 0.8)	0.0888			0.1885***	0.1160			0.1944***	0.0638			0.1881***
	(0.2845)			(0.0184)	(0.3203)			(0.0258)	(0.2445)			(0.0276)
Health risk taking (0.9,1.0)	0.0156			0.3050***	0.0201			0.3201***	0.0115			0.2906***
	(0.1240)			(0.0357)	(0.1403)			(0.0452)	(0.1068)			(0.0575)
Female (dummy)	0.5216	-0.0357***	-0.0549***	-0.0365***								
	(0.4996)	(0.0084)	(0.0083)	(0.0084)								
Age (years)	42.9528	0.0164***	0.0147***	0.0166***	43.1530	0.0165***	0.0148***	0.0169***	42.7691	0.0176***	0.0158***	0.0177***
	(12.7899)	(0.0024)	(0.0024)	(0.0024)	(12.8145)	(0.0036)	(0.0036)	(0.0036)	(12.7655)	(0.0032)	(0.0032)	(0.0032)
Age squared (/100)	20.0851	-0.0243***	-0.0231***	-0.0246***	20.2637	-0.0242***	-0.0231***	-0.0247***	19.9213	-0.0260***	-0.0244***	-0.0261***
	(10.9946)	(0.0028)	(0.0028)	(0.0028)	(11.0454)	(0.0043)	(0.0042)	(0.0043)	(10.9461)	(0.0038)	(0.0038)	(0.0038)
Secondary school degree (ref.: low)												
Medium schooling ("Realschule")	0.3764	-0.0807***	-0.0754***	-0.0784***	0.3331	-0.0744***	-0.0698***	-0.0713***	0.4162	-0.0835***	-0.0773***	-0.0829***
(dummy)	(0.4845)	(0.0099)	(0.0099)	(0.0100)	(0.4714)	(0.0145)	(0.0145)	(0.0146)	(0.4930)	(0.0136)	(0.0136)	(0.0136)
High schooling ("Abitur")	0.3145	-0.1429***	-0.1355***	-0.1394***	0.3394	-0.1421***	-0.1315***	-0.1389***	0.2918	-0.1392***	-0.1338***	-0.1357***
(dummy)	(0.4643)	(0.0123)	(0.0124)	(0.0124)	(0.4735)	(0.0185)	(0.0186)	(0.0186)	(0.4546)	(0.0161)	(0.0162)	(0.0162)
Apprenticeship degree	0.7023	-0.0406***	-0.0393***	-0.0401***	0.7045	-0.0173	-0.0127	-0.0168	0.7002	-0.0615***	-0.0622***	-0.0611***
(dummy)	(0.4573)	(0.0110)	(0.0109)	(0.0110)	(0.4563)	(0.0167)	(0.0166)	(0.0167)	(0.4582)	(0.0146)	(0.0145)	(0.0146)
University degree	0.2257	-0.1068***	-0.1006***	-0.1060***	0.2532	-0.1261***	-0.1211***	-0.1242***	0.2004	-0.0884***	-0.0812***	-0.0887***
(dummy)	(0.4180)	(0.0126)	(0.0127)	(0.0126)	(0.4349)	(0.0188)	(0.0189)	(0.0189)	(0.4003)	(0.0167)	(0.0169)	(0.0167)
Monthly net household income	2.9994	-0.0092**	-0.0081*	-0.0091**	3.0796	-0.0066	-0.0053	-0.0067	2.9259	-0.0133***	-0.0124***	-0.0129***
(in 1000 Euros)	(2.2645)	(0.0044)	(0.0041)	(0.0043)	(2.5360)	(0.0054)	(0.0050)	(0.0054)	(1.9803)	(0.0037)	(0.0037)	(0.0037)
Employment status (ref.: non)												
Unemployed	0.0847	0.1992***	0.2029***	0.1998***	0.0858	0.2173***	0.2208***	0.2180***	0.0837	0.1690***	0.1746***	0.1691***
(dummy)	(0.2784)	(0.0183)	(0.0182)	(0.0183)	(0.2801)	(0.0286)	(0.0285)	(0.0287)	(0.2769)	(0.0236)	(0.0235)	(0.0236)
Employed	0.7206	0.0521***	0.0602***	0.0544***	0.7915	0.0317	0.0433**	0.0342	0.6555	0.0623***	0.0686***	0.0648***
(dummy)	(0.4487)	(0.0111)	(0.0110)	(0.0111)	(0.4062)	(0.0207)	(0.0204)	(0.0207)	(0.4752)	(0.0130)	(0.0129)	(0.0130)
Mean smoking probability	0.3227	0.3227	0.3227	0.3227	0.3506	0.3506	0.3506	0.3506	0.2972	0.2972	0.2972	0.2972
Log Likelihood (Probit)		-7970.67	-8070.29	-7953.11		-3906.62	-3951.67	-3898.96		-4053.51	-4107.51	-4040.86

Note: ML-Probit regression (marginal effects Pr[N>0] at means of all covariates). Standard deviations and robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01.

Table 2: Truncated regressions for number of smoked cigarettes per day (second part of two-part model)

	Total sample (n=4401)				Men (n=2287)				Women (n=2114)			
	Mean (sd)	(1)	(2)	(3)	Mean (sd)	(1)	(2)	(3)	Mean (sd)	(1)	(2)	(3)
Health risk taking [0,1]	0.3667	3.2969***			0.4005	3.3072***			0.3301	3.2480***		
	(0.2573)	(0.5466)			(0.2568)	(0.8211)			(0.2529)	(0.7028)		
Health risk taking minus	-0.1342		2.1696***		-0.1368		1.2768		-0.1313		3.0026***	
general risk taking [-1,+1]	(0.2676)		(0.5362)		(0.2636)		(0.7918)		(0.2720)		(0.7078)	
Health risk taking (reference: 0.0)												
Health risk taking (0.1, 0.2)	0.2084			-0.8430**	0.1955			-1.3081**	0.2223			-0.5264
	(0.4062)			(0.4192)	(0.3966)			(0.6598)	(0.4159)			(0.5276)
Health risk taking (0.3, 0.4)	0.2045			-0.4873	0.2007			-1.1310*	0.2086			0.0732
	(0.4034)			(0.4217)	(0.4006)			(0.6555)	(0.4064)			(0.5369)
Health risk taking (0.5, 0.6)	0.2695			0.7976*	0.2886			0.4217	0.2488			1.0768**
	(0.4437)			(0.4275)	(0.4532)			(0.6617)	(0.4324)			(0.5451)
Health risk taking (0.7, 0.8)	0.1266			2.0106***	0.1535			1.5080*	0.0974			2.3812***
	(0.3325)			(0.5376)	(0.3605)			(0.7874)	(0.2966)			(0.7450)
Health risk taking (0.9,1.0)	0.0286			3.3027***	0.0354			2.8679**	0.0213			3.7489***
	(0.1668)			(0.9855)	(0.1849)			(1.3762)	(0.1444)			(1.2849)
Female (dummy)	0.4803	-3.3266***	-3.5604***	-3.3427***								
	(0.4997)	(0.2689)	(0.2669)	(0.2685)								
Age (years)	40.5115	0.6607***	0.6417***	0.6616***	40.7114	0.8425***	0.8311***	0.8511***	40.2952	0.4652***	0.4363***	0.4606***
	(12.0647)	(0.0754)	(0.0755)	(0.0755)	(12.1396)	(0.1130)	(0.1127)	(0.1131)	(11.9823)	(0.0987)	(0.0992)	(0.0990)
Age squared (/100)	17.8670	-0.6623***	-0.6530***	-0.6640***	18.0472	-0.8282***	-0.8290***	-0.8395***	17.6721	-0.4825***	-0.4602***	-0.4782***
	(9.9407)	(0.0913)	(0.0915)	(0.0914)	(10.0611)	(0.1361)	(0.1360)	(0.1362)	(9.8076)	(0.1208)	(0.1212)	(0.1211)
Secondary school degree (ref.: low)												
Medium schooling ("Realschule")	0.3960	-2.0327***	-1.9609***	-2.0173***	0.3629	-1.7917***	-1.7525***	-1.7216***	0.4319	-2.2577***	-2.1835***	-2.2972***
(dummy)	(0.4891)	(0.2927)	(0.2929)	(0.2927)	(0.4809)	(0.4233)	(0.4245)	(0.4236)	(0.4955)	(0.3995)	(0.3981)	(0.4025)
High schooling ("Abitur")	0.2290	-2.8757***	-2.7801***	-2.8475***	0.2352	-2.6418***	-2.4766***	-2.6147***	0.2223	-3.0954***	-3.0537***	-3.0841***
(dummy)	(0.4203)	(0.4168)	(0.4202)	(0.4167)	(0.4242)	(0.6392)	(0.6475)	(0.6403)	(0.4159)	(0.5212)	(0.5197)	(0.5208)
Apprenticeship degree	0.7267	-0.3483	-0.3191	-0.3247	0.7495	-0.4530	-0.4119	-0.4111	0.7020	-0.1954	-0.1431	-0.1816
(dummy)	(0.4457)	(0.3371)	(0.3380)	(0.3370)	(0.4334)	(0.5299)	(0.5310)	(0.5303)	(0.4575)	(0.4154)	(0.4153)	(0.4158)
University degree	0.1359	-1.4262***	-1.2901**	-1.4160***	0.1408	-2.0606***	-1.9594**	-1.9940***	0.1306	-0.7275	-0.5367	-0.7574
(dummy)	(0.3427)	(0.4988)	(0.5040)	(0.4953)	(0.3479)	(0.7602)	(0.7704)	(0.7552)	(0.3370)	(0.6235)	(0.6251)	(0.6192)
Monthly net household income	2.6944	0.0258	0.0403	0.0304	2.7357	0.0935**	0.1049**	0.0940**	2.6497	-0.1627	-0.1402	-0.1429
(in 1000 Euros)	(2.2641)	(0.0629)	(0.0591)	(0.0602)	(2.6824)	(0.0438)	(0.0424)	(0.0430)	(1.6988)	(0.1170)	(0.1170)	(0.1164)
Employment status (ref.: non)												
Unemployed	0.1316	1.1619**	1.2249**	1.1798**	0.1360	0.9271	0.9432	0.9337	0.1268	1.4600*	1.5783**	1.4765**
(dummy)	(0.3381)	(0.5479)	(0.5514)	(0.5469)	(0.3428)	(0.7688)	(0.7774)	(0.7655)	(0.3328)	(0.7567)	(0.7614)	(0.7520)
Employed	0.7344	0.3155	0.4165	0.3585	0.7814	0.6438	0.7660	0.6696	0.6835	0.0492	0.1292	0.1003
(dummy)	(0.4417)	(0.4111)	(0.4121)	(0.4100)	(0.4134)	(0.6538)	(0.6578)	(0.6506)	(0.4652)	(0.4884)	(0.4894)	(0.4863)
Mean smoking probability	15.8655	15.8655	15.8655	15.8655	17.6270	17.6270	17.6270	17.6270	13.9598	13.9598	13.9598	13.9598
Log Likelihood (Truncated)		-15453.18	-15463.16	-15442.77		-8192.80	-8200.53	-8185.31		-7247.61	-7247.47	-7243.88
Log Likelihood (Two-part)		-23423.85	-23533.45	-23395.88		-12099.42	-12152.20	-12084.27		-11301.12	-11354.98	-11284.74

Note: ML-Truncated regression (marginal effects E[N|N>0] at means of all covariates). Standard deviations and robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Log Likelihood for the two-part model is the sum of the log likelihoods for the truncated regression and the Probit regression in Table 1.

3.3 Results from Tobit Models

Whereas the two-part model assumes independent decisions of smoking status and smoking intensity, the Tobit model assumes a joint decision so that probability to smoke and the number of consumed cigarettes are jointly modeled and estimated. Thus, the Tobit results in this section can be seen as a robustness check. Table 3 contains the results of the Tobit regressions for daily cigarette consumption in the total sample. Risk taking has the expected significant coefficients throughout all regressions. For a quantitative interpretation marginal effects are computed for the probability of being a cigarette smoker (compared to being a non-smoker) ($\Pr[N > 0]$) and for the expected number of consumed cigarettes conditional on being a cigarette smoker ($E[N|N > 0]$) (McDonald and Moffitt, 1980).

Specification one shows that individuals, who are 0.1 point more health risk taking, are on average 2.6 percentage points more likely to smoke cigarettes and smoke on average approximately 0.4 cigarettes more per day. The second specification shows that individuals, who are 0.1 point more risk taking in health issues than in general, are on average 1.1 percentage point more likely to be a smoker and smoke about 0.18 additional cigarettes. The third specification takes into account non-linearity of the risk taking effect. Individuals with low risk taking preference have neither significant higher smoking probability nor intensity than completely risk averse individuals. Individuals with more than average risk taking preference have however a statistical significant higher smoking probability and intensity. An extremely high risk taker smokes for example about 5 additional cigarettes per day compared to less than average risk takers. The results are robust in separate male and female samples (see Table 4 and Table 5).[12]

The overall findings for risk taking and most control variables are virtually the same as in the two-part model presented in the previous section. Thus, the choice between Tobit and two-part models seems largely arbitrary for this application. Household income however represents an example where Tobit and two-part models reveal different results, as income effects might indeed have separate effects on smoking probability and smoking intensity due to the addictive nature of cigarettes. From a statistical point of view, the two-part model performs better than the Tobit model with respect to the log-likelihoods in all used specifications and samples (see bottom of Tables 3, 4, and 5).

[12] Additional regressions with interaction terms between gender and risk taking show that the small differences between the two samples are not significant. The results of these regressions are not presented in this note but can be requested from the author.

Table 3: Tobit regressions for number of smoked cigarettes per day (total sample)

		Specification (1)			Specification (2)			Specification (3)		
	Mean (sd)	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]
Health risk taking [0,1]	0.3167	14.6868***	0.2588***	4.0459***						
	(0.2428)	(0.9127)	(0.0161)	(0.2492)						
Health risk taking minus	-0.1547				6.4521***	0.1125***	1.7823***			
general risk taking [-1,+1]	(0.2492)				(0.9090)	(0.0159)	(0.2509)			
Health risk taking (reference: 0.0)										
Health risk taking (0.1, 0.2)	0.2573							-0.8533	-0.0150	-0.2336
	(0.4372)							(0.6889)	(0.0121)	(0.1874)
Health risk taking (0.3, 0.4)	0.2266							0.7422	0.0132	0.2057
	(0.4187)							(0.7014)	(0.0125)	(0.1955)
Health risk taking (0.5, 0.6)	0.2243							5.6704***	0.1032***	1.6386***
	(0.4171)							(0.6849)	(0.0128)	(0.2071)
Health risk taking (0.7, 0.8)	0.0888							9.6952***	0.1823***	3.0185***
	(0.2845)							(0.8534)	(0.0166)	(0.2980)
Health risk taking (0.9,1.0)	0.0156							14.8412***	0.2835***	5.1034***
	(0.1240)							(1.5499)	(0.0294)	(0.6524)
Female (dummy)	0.5216	-3.4349***	-0.0606***	-0.9488***	-4.5331***	-0.0791***	-1.2568***	-3.4723***	-0.0614***	-0.9591***
	(0.4996)	(0.4436)	(0.0078)	(0.1230)	(0.4441)	(0.0077)	(0.1237)	(0.4426)	(0.0078)	(0.1228)
Age (years)	42.9528	1.2228***	0.0215***	0.3369***	1.1361***	0.0198***	0.3138***	1.2352***	0.0218***	0.3403***
	(12.7899)	(0.1235)	(0.0022)	(0.0339)	(0.1244)	(0.0022)	(0.0343)	(0.1234)	(0.0022)	(0.0339)
Age squared (/100)	20.0851	-1.6464***	-0.0290***	-0.4535***	-1.5922***	-0.0278***	-0.4398***	-1.6619***	-0.0294***	-0.4578***
	(10.9946)	(0.1476)	(0.0026)	(0.0405)	(0.1486)	(0.0026)	(0.0409)	(0.1474)	(0.0026)	(0.0404)
Secondary school degree (ref.: low)										
Medium schooling ("Realschule")	0.3764	-5.1280***	-0.0889***	-1.3877***	-4.8572***	-0.0834***	-1.3195***	-4.9851***	-0.0867***	-1.3495***
(dummy)	(0.4845)	(0.5362)	(0.0091)	(0.1425)	(0.5395)	(0.0091)	(0.1441)	(0.5356)	(0.0091)	(0.1425)
High schooling ("Abitur")	0.3145	-8.8615***	-0.1492***	-2.3319***	-8.4770***	-0.1417***	-2.2422***	-8.6224***	-0.1458***	-2.2712***
(dummy)	(0.4643)	(0.7256)	(0.0117)	(0.1824)	(0.7316)	(0.0117)	(0.1853)	(0.7253)	(0.0117)	(0.1827)
Apprenticeship degree	0.7023	-2.1454***	-0.0381***	-0.5988***	-2.0901***	-0.0368***	-0.5847***	-2.1085***	-0.0376***	-0.5884***
(dummy)	(0.4573)	(0.5643)	(0.0101)	(0.1595)	(0.5696)	(0.0101)	(0.1612)	(0.5626)	(0.0101)	(0.1589)
University degree	0.2257	-6.6933***	-0.1124***	-1.7490***	-6.3551***	-0.1060***	-1.6705***	-6.6173***	-0.1115***	-1.7298***
(dummy)	(0.4180)	(0.7713)	(0.0122)	(0.1906)	(0.7809)	(0.0123)	(0.1947)	(0.7688)	(0.0122)	(0.1901)
Monthly net household income	2.9994	-0.4642*	-0.0082*	-0.1279*	-0.4045	-0.0071	-0.1117	-0.4554*	-0.0081*	-0.1254*
(in 1000 Euros)	(2.2645)	(0.2618)	(0.0046)	(0.0720)	(0.2518)	(0.0044)	(0.0695)	(0.2576)	(0.0046)	(0.0709)
Employment status (ref.: non)										
Unemployed	0.0847	9.3885***	0.1758***	2.9137***	9.7345***	0.1805***	3.0385***	9.3879***	0.1764***	2.9146***
(dummy)	(0.2784)	(0.8327)	(0.0160)	(0.2887)	(0.8372)	(0.0160)	(0.2925)	(0.8286)	(0.0160)	(0.2875)
Employed	0.7206	2.7876***	0.0484***	0.7542***	3.3105***	0.0567***	0.8954***	2.9145***	0.0507***	0.7878***
(dummy)	(0.4487)	(0.6241)	(0.0107)	(0.1657)	(0.6277)	(0.0105)	(0.1660)	(0.6219)	(0.0106)	(0.1648)
Log Likelihood (Tobit)		-24020.03			-24127.41			-23996.47		
Log Likelihood (Two-part)		-23423.85			-23533.45			-23395.88		

Note: ML-Tobit regression (coefficients and marginal effects at means of all covariates) for total sample (n=13636). Standard deviations and robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Log Likelihood for the two-part model is the sum of the log likelihoods for the Probit and truncated regressions in Table 1 and Table 2.

Table 4: Tobit regressions for number of smoked cigarettes per day (men)

		Specification (1)			Specification (2)			Specification (3)		
	Mean (sd)	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]
Health risk taking [0,1]	0.3556	15.5100***	0.2663***	4.5035***						
	(0.2460)	(1.3586)	(0.0234)	(0.3913)						
Health risk taking minus	-0.1609				7.7230***	0.1312***	2.2467***			
general risk taking [-1,+1]	(0.2515)				(1.3510)	(0.0230)	(0.3924)			
Health risk taking (reference: 0.0)										
Health risk taking (0.1, 0.2)	0.2301							0.0549	0.0009	0.0159
	(0.4209)							(1.1475)	(0.0198)	(0.3335)
Health risk taking (0.3, 0.4)	0.2333							0.5237	0.0090	0.1527
	(0.4230)							(1.1480)	(0.0199)	(0.3360)
Health risk taking (0.5, 0.6)	0.2535							5.7992***	0.1019***	1.7549***
	(0.4351)							(1.1023)	(0.0197)	(0.3470)
Health risk taking (0.7, 0.8)	0.1160							10.1266***	0.1819***	3.2902***
	(0.3203)							(1.2814)	(0.0234)	(0.4626)
Health risk taking (0.9,1.0)	0.0201							15.8309***	0.2857***	5.7270***
	(0.1403)							(2.1381)	(0.0371)	(0.9435)
Age (years)	43.1530	1.3639***	0.0234***	0.3960***	1.2782***	0.0217***	0.3718***	1.3900***	0.0240***	0.4036***
	(12.8145)	(0.1913)	(0.0033)	(0.0555)	(0.1921)	(0.0033)	(0.0558)	(0.1910)	(0.0033)	(0.0554)
Age squared (/100)	20.2637	-1.7750***	-0.0305***	-0.5154***	-1.7302***	-0.0294***	-0.5033***	-1.8062***	-0.0311***	-0.5245***
	(11.0454)	(0.2290)	(0.0039)	(0.0663)	(0.2300)	(0.0039)	(0.0667)	(0.2285)	(0.0039)	(0.0662)
Secondary school degree (ref.: low)										
Medium schooling ("Realschule")	0.3331	-4.7940***	-0.0810***	-1.3622***	-4.5714***	-0.0765***	-1.3030***	-4.5682***	-0.0775***	-1.2994***
(dummy)	(0.4714)	(0.7897)	(0.0131)	(0.2196)	(0.7965)	(0.0131)	(0.2224)	(0.7914)	(0.0132)	(0.2205)
High schooling ("Abitur")	0.3394	-8.7745***	-0.1460***	-2.4556***	-8.2069***	-0.1355***	-2.3070***	-8.5478***	-0.1428***	-2.3942***
(dummy)	(0.4735)	(1.0721)	(0.0173)	(0.2893)	(1.0885)	(0.0175)	(0.2958)	(1.0748)	(0.0175)	(0.2905)
Apprenticeship degree	0.7045	-1.1497	-0.0198	-0.3360	-0.9268	-0.0158	-0.2710	-1.1183	-0.0193	-0.3269
(dummy)	(0.4563)	(0.8912)	(0.0154)	(0.2621)	(0.8999)	(0.0154)	(0.2645)	(0.8888)	(0.0154)	(0.2614)
University degree	0.2532	-8.3678***	-0.1374***	-2.2987***	-8.1382***	-0.1325***	-2.2445***	-8.2003***	-0.1352***	-2.2549***
(dummy)	(0.4349)	(1.1816)	(0.0183)	(0.3062)	(1.2015)	(0.0186)	(0.3132)	(1.1777)	(0.0184)	(0.3057)
Monthly net household income	3.0796	-0.2908	-0.0050	-0.0844	-0.2187	-0.0037	-0.0636	-0.2903	-0.0050	-0.0843
(in 1000 Euros)	(2.5360)	(0.3177)	(0.0055)	(0.0922)	(0.3007)	(0.0051)	(0.0875)	(0.3160)	(0.0054)	(0.0917)
Employment status (ref.: non)										
Unemployed	0.0858	9.9976***	0.1794***	3.2694***	10.3726***	0.1841***	3.4088***	10.0124***	0.1803***	3.2764***
(dummy)	(0.2801)	(1.3355)	(0.0244)	(0.4883)	(1.3501)	(0.0243)	(0.4972)	(1.3336)	(0.0244)	(0.4882)
Employed	0.7915	1.8465	0.0314	0.5282	2.5597**	0.0429**	0.7296**	1.9832*	0.0338*	0.5667*
(dummy)	(0.4062)	(1.1590)	(0.0195)	(0.3266)	(1.1670)	(0.0193)	(0.3258)	(1.1553)	(0.0195)	(0.3248)
Log Likelihood (Tobit)		-12472.59			-12523.00			-12461.44		
Log Likelihood (Two-part)		-12099.42			-12152.20			-12084.27		

Note: ML- Tobit regression (coefficients and marginal effects at means of all covariates) for men (n=6524). Standard deviations and robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Log Likelihood for the two-part model is the sum of the log likelihoods for the Probit and truncated regressions in Table 1 and Table 2.

Table 5: Tobit regressions for number of smoked cigarettes per day (women)

		Specification (1)			Specification (2)			Specification (3)		
	Mean (sd)	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]	Coef.	Pr[N>0]	E[N|N>0]
Health risk taking [0,1]	0.2810	13.7950***	0.2530***	3.6209***						
	(0.2342)	(1.2035)	(0.0221)	(0.3129)						
Health risk taking minus	-0.1490				5.3045***	0.0964***	1.3974***			
general risk taking [-1,+1]	(0.2468)				(1.2035)	(0.0219)	(0.3170)			
Health risk taking (reference: 0.0)										
Health risk taking (0.1, 0.2)	0.2823							-1.5558*	-0.0283*	-0.4038*
	(0.4502)							(0.8355)	(0.0150)	(0.2145)
Health risk taking (0.3, 0.4)	0.2205							0.9861	0.0183	0.2610
	(0.4146)							(0.8608)	(0.0161)	(0.2299)
Health risk taking (0.5, 0.6)	0.1974							5.6286***	0.1080***	1.5623***
	(0.3981)							(0.8577)	(0.0170)	(0.2509)
Health risk taking (0.7, 0.8)	0.0638							9.2801***	0.1857***	2.7853***
	(0.2445)							(1.1611)	(0.0245)	(0.3954)
Health risk taking (0.9,1.0)	0.0115							13.6839***	0.2802***	4.4900***
	(0.1068)							(2.2621)	(0.0474)	(0.9116)
Age (years)	42.7691	1.1547***	0.0212***	0.3031***	1.0563***	0.0192***	0.2783***	1.1507***	0.0212***	0.3019***
	(12.7655)	(0.1613)	(0.0030)	(0.0422)	(0.1627)	(0.0030)	(0.0427)	(0.1613)	(0.0030)	(0.0422)
Age squared (/100)	19.9213	-1.5975***	-0.0293***	-0.4193***	-1.5211***	-0.0276***	-0.4007***	-1.5947***	-0.0293***	-0.4184***
	(10.9461)	(0.1933)	(0.0036)	(0.0504)	(0.1950)	(0.0036)	(0.0511)	(0.1933)	(0.0036)	(0.0504)
Secondary school degree (ref.: low)										
Medium schooling ("Realschule")	0.4162	-5.2299***	-0.0945***	-1.3552***	-4.8992***	-0.0878***	-1.2752***	-5.2064***	-0.0944***	-1.3485***
(dummy)	(0.4930)	(0.7119)	(0.0126)	(0.1817)	(0.7157)	(0.0126)	(0.1837)	(0.7103)	(0.0126)	(0.1813)
High schooling ("Abitur")	0.2918	-8.6688***	-0.1495***	-2.1548***	-8.4045***	-0.1441***	-2.1012***	-8.4493***	-0.1464***	-2.1015***
(dummy)	(0.4546)	(0.9325)	(0.0150)	(0.2188)	(0.9379)	(0.0150)	(0.2218)	(0.9342)	(0.0151)	(0.2197)
Apprenticeship degree	0.7002	-3.0408***	-0.0567***	-0.8143***	-3.0950***	-0.0571***	-0.8318***	-3.0077***	-0.0562***	-0.8049***
(dummy)	(0.4582)	(0.7096)	(0.0135)	(0.1936)	(0.7163)	(0.0135)	(0.1962)	(0.7085)	(0.0135)	(0.1931)
University degree	0.2004	-5.0113***	-0.0876***	-1.2550***	-4.5877***	-0.0799***	-1.1581***	-5.0125***	-0.0879***	-1.2544***
(dummy)	(0.4003)	(0.9703)	(0.0160)	(0.2311)	(0.9795)	(0.0162)	(0.2362)	(0.9688)	(0.0161)	(0.2305)
Monthly net household income	2.9259	-0.7266***	-0.0133***	-0.1907***	-0.6819***	-0.0124***	-0.1796***	-0.6959***	-0.0128***	-0.1826***
(in 1000 Euros)	(1.9803)	(0.1945)	(0.0036)	(0.0509)	(0.1947)	(0.0035)	(0.0512)	(0.1939)	(0.0036)	(0.0508)
Employment status (ref.: non)										
Unemployed	0.0837	8.1430***	0.1607***	2.3907***	8.5374***	0.1670***	2.5258***	8.1120***	0.1607***	2.3804***
(dummy)	(0.2769)	(1.0451)	(0.0214)	(0.3398)	(1.0533)	(0.0214)	(0.3462)	(1.0403)	(0.0214)	(0.3380)
Employed	0.6555	3.1618***	0.0571***	0.8173***	3.5708***	0.0638***	0.9248***	3.2967***	0.0597***	0.8511***
(dummy)	(0.4752)	(0.7142)	(0.0127)	(0.1814)	(0.7189)	(0.0126)	(0.1825)	(0.7127)	(0.0127)	(0.1806)
Log Likelihood (Tobit)		-11518.82			-11575.97			-11503.71		
Log Likelihood (Two-part)		-11301.12			-11354.98			-11284.74		

Note: ML- Tobit regression (coefficients and marginal effects at means of all covariates) for women (n=7112). Standard deviations and robust standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01. Log Likelihood for the two-part model is the sum of the log likelihoods for the Probit and truncated regressions in Table 1 and Table 2.

4 Conclusion

The main finding of this research note is that risk takers are significantly more likely to smoke and have a higher demand for cigarettes smoked per day, which holds for men and women. A comparison between the two-part model (Probit and truncated regressions) and the Tobit model does not reveal noteworthy differences, but the two-part model outperforms the Tobit model with respect to their log-likelihoods. One caveat of the empirical analysis in this paper is that the used cross-section data does not allow to analyze cigarettes as an addictive stock and that endogeneity of the risk taking preferences cannot be ruled out (e.g., Chaloupka, 1991). Moreover, risk taking cannot be perfectly separated from health knowledge, which includes collection of health information and understanding of health risks. But even if the effect of risk taking preferences on smoking behavior would not be directly causal, the significant correlations make risk taking preferences a useful signal. Future empirical research should nevertheless try to shed some light on these potential problems.

One reason for the effect of risk taking on smoking might be that risk takers expect lower negative consequences from smoking, which has been emphasized in the theoretical frameworks. Risk takers are hence more likely to be engaged in unhealthy behavior and to be more costly for health care systems and health insurances. It is therefore rational for health insurance companies to ask for risk premiums of more risk taking customers. The findings might also suggest that it is, on the one hand, especially important to provide information about health risks to low risk averse individuals to increase their subjectively perceived health risks. More risk averse individuals, on the other hand, tend to overestimate health risks and also have disutility from uncertainty about individual health risks and health reductions. The additional information costs for low risk averse individuals can hence be offset by reducing information to high risk averse individuals. Furthermore, these costs might be more than compensated by health costs reductions so that health insurance premiums and public subsidies for the health care sector could be reduced or spend otherwise.

* Correspondence to: Christian Pfeifer, Institute of Economics, Leuphana University Lueneburg, Scharnhorststr. 1, 21335 Lüneburg, Germany. Phone: +49-4131-6772301. Email: pfeifer@leuphana.de.

References

Becker, G.S., Grossman, M., Murphy, K.M., 1994. An empirical analysis of cigarette addiction. American Economic Review 84, 396-418.

Chaloupka, F., 1991. Rational addictive behavior and cigarette smoking. Journal of Political Economy 99, 722-742.

Chaloupka, F.J., Warner, K.E., 2000. The economics of smoking. Handbook of Health Economics 1(2), 1539-1627.

Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39, 829-844.

Dardanoni, V., Wagstaff, A., 1990. Uncertainty and the demand for medical care. Journal of Health Economics 9, 23-38.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual risk attitudes: measurement, determinants and behavioral consequences. Journal of the European Economic Association 9, 522-550.

Goel, R.K., Nelson, M.A., 2006. The effectiveness of anti-smoking legislation: a review. Journal of Economic Surveys 20, 325-355.

Grossman, M., 1972. On the concept of health capital and the demand for health. Journal of Political Economy 80, 223-255.

Hammar, H., Johansson-Stenman, O., 2004. The value of risk-free cigarettes - do smokers underestimate the risk? Health Economics 13, 59-71.

Liljas, B., 1998. The demand for health with uncertainty and insurance. Journal of Health Economics 17, 153-170.

Liu, J.-T., Hsieh, C.-R., 1995. Risk perception and smoking behavior: empirical evidence from Taiwan. Journal of Risk and Uncertainty 11, 139-157.

Lundborg, P., Andersson, H., 2008. Gender, risk perceptions, and smoking behavior. Journal of Health Economics 27, 1299-1311.

Lundborg, P., Lindgren, B., 2004. Do they know what they are doing? Risk perceptions and smoking behaviour among Swedish teenagers. Journal of Risk and Uncertainty 28, 261-286.

Madden, D., 2008. Sample selection versus two-part models revisited: the case of female smoking and drinking. Journal of Health Economics 27, 300-307.

McDonald, J.F., Moffitt, R.A., 1980. The uses of Tobit analysis. Review of Economics and Statistics 62, 318-321.

Viscusi, W.K., 1990. Do smokers underestimate risks? Journal of Political Economy 98, 1253-1269.

Viscusi, W.K., Hakes, J.K., 2008. Risk beliefs and smoking behavior. Economic Inquiry 46, 45-59.