Smokers’ misperceptions of risk over time

by

Luke van der Stelt

511776

Master thesis

Economics and Business

Health Economics

Erasmus School of Economics

23-07-2024

Rotterdam

Word count: 11088

Supervisor: Fanny Tallgren

Second assessor: Carlos Riumallo Herl

Abstract

This study examines whether smokers are associated with misperceptions of survival and the extent to which this association differs for differences in age, period and cohort. Whether smoking behavior is primarily driven by misperceptions of risk remains a focus of public debate and is often used as a rationale for implementing new anti-smoking policies. However, previous research on the association between smoking and misperceptions of risk has suffered from period and cohort bias. The examination of unique associations between smokers and their perceptions of risk for specific age groups, time periods and cohorts, provides valuable insights for policymakers on how best to develop anti-smoking policies.

To examine misperceptions of risk, this study compares the reported subjective probability of surviving to age 75 with the computed objective probability of surviving to age 75. The analyses are based on longitudinal data (2015-2022) with a sample of people aged 17 to 73. The negative association between smokers and their subjective probability of survival relative to those who have never smoked, shows that smokers do perceive additional health risks. However, smokers and particularly heavy smokers are relatively optimistic about their probability of surviving to age 75, compared to never smokers. This association differs for differences in age, period and cohort. Particularly younger smokers and smokers from recent cohorts appear to be relatively optimistic about their chances of surviving to age 75. Conversely, smokers aged 68-73 are found to be relatively pessimistic about their survival. There is no evidence of an effect of anti-smoking policies implemented between 2015 and 2022 on smokers' beliefs about longevity.

Future research could replicate the study over a longer period of time in order to be better able to distinguish between age and cohort effects. In addition, further research could examine the extent to which smokers’ relative optimism about survival is driven by misperceptions about the harm of smoking or by rosy thoughts about the future.

Table of contents

1. Introduction

For high-income countries, noncommunicable diseases (NCDs) have been the leading cause of mortality and morbidity over the last decades. However, many of the NCDs can be attributed to smoking behavior and are therefore preventable. To illustrate this problem in the Dutch context, more than 700.000 people in the Netherlands are currently diagnosed with coronary heart disease. In addition, cardiovascular diseases cause 400.000 hospitalizations and 80.000 cardiac interventions annually. This evidently has a substantial impact on the burden of the Dutch health care system (Leening et al., 2013). It is therefore in the interest of policymakers to develop strategies to discourage individuals from engaging in smoking behavior. Despite the health consequences, people may continue to smoke for a variety of reasons. An important question for policymakers to consider is whether smoking behavior is driven by a rational choice or by undesirable causes. For example, addiction or other internalities such as misperceptions of risk, time inconsistency and self-control problems. Whether this is because they misperceive the risks of smoking remains a focus of public debate and is often used as a motivation for implementing new tobacco control policies. While excise taxes may be justified and effective as compensation for the external costs, research has shown that these policies are much weaker in correcting for internalities (Kwaja et al., 2009). In addition, Viscusi and Kip (1990) demonstrated that changes in smokers' risk perception are estimated to have an equivalent effect on smoking behavior to changes in excise taxes.

For this reason, in addition to increasing the excise taxes, policymakers in recent years have focused more on how to influence the risk perception of smoking. This is reflected in the emergence of public information campaigns and compulsory messages on cigarette packages about the adverse effects of smoking. These policy measures suggest that smokers do misperceive the risk of smoking. Nevertheless, previous studies on smokers' risk perceptions demonstrated that, although not to the same extent as non-smokers, smokers do perceive increased health risks (Viscusi & Kip, 1990). However, it is difficult to quantify and compare the perceived costs of smoking-related diseases. This is because the perceived internal costs of lung disease or smoking-related morbidity issues may vary from person to person. Therefore, consistent with other literature, this study rather uses years of life lost to quantify the individual implications of these smoking-related diseases (Schoenbaum,1997; Kwaja et al.,2009; Sloan et al., 2011). In addition, the extent to which smokers misperceive the risks of smoking is examined by comparing reported subjective survival probabilities with their objective counterparts. While the association between smokers and their survival probability errors has been studied, no study has examined the extent to which this association varies for different age, period and cohort (APC) groups. Therefore, the research question of this study is as follows: *'To what extent does the relationship between smoking status and misperceptions of survival differ for different age, period and cohort groups?'*

Both Schoenbaum (1997) and Kwaja et al. (2009) compared the survival probability errors of smokers with those of people who had never smoked. The rationale behind this comparison is that if misperception causes smokers to continue smoking and to underestimate the health risks, we would expect smokers to be relatively optimistic about their probability of survival. However, in these studies, the derived subjective probabilities were cross-sectional and reported in a different time period than their objective counterparts. Therefore, the results are likely to be biased by period and cohort effects. In addition, both studies were conducted on a mature (ages 50-70) sample, which does not allow examination of the relationship between smoking status and misperceptions of survival at the time of smoking initiation. These limitations did not allow to study potential moderating APC effects. However, given that smoking prevalence and initiation have declined over time, and given that previous research demonstrated an association between smokers' risk perceptions and smoking behavior, these APC effects are to be expected (Viscusi and Kip, 1990; Marcon et al., 2018; Breton et al.,2021; Meza et al.,2023). Furthermore, at a given point in time, age effects on smokers' risk perceptions are likely to be due to younger people considering that the adverse effects will occur at a later point in time (Jeon et al., 2023). On the contrary, young people's perception of risk depends more on the sources of information presented to them. Whereas as people age, their perception of the risk of smoking depends more on information gained from experience (Viscusi & Kip, 1991). In addition, period effects on risk perceptions could be the result of policy measures in the Netherlands, such as the ban on the use of color in cigarette packaging in 2020. Examining unique associations between risk perception and smoking behavior for specific age groups, time periods and cohorts could provide valuable insights for policymakers on how best to develop tobacco control policies. In addition, policies that correct for internalities that drive smoking behavior result in both beneficial societal health outcomes and a higher societal utility (Ehrlich & Becker, 1972; Snow, 2010)

To test the hypothesis that the relationship between smoking status and (mis)perceptions of survival varies for different APC groups, this study uses data from the Longitudinal Internet studies for the Social Sciences (LISS) panel for the years 2015-2022. Furthermore, the sample includes individuals between the ages of 17 and 73. In contrast to previous studies, the longitudinal nature of the data permits to link (changes in) risk perceptions to actual smoking behavior. In addition, the dataset allows to control for other factors that may affect one's subjective survival probability or the accuracy of predicting survival such as, alcohol consumption, BMI level and educational attainment.

Firstly, this study tests the hypothesis that smokers perceive the increased health risks of smoking by comparing the subjective survival probabilities of smokers with those of non-smokers. Secondly, I test the hypothesis that smokers do update their subjective health differently than non-smokers. Thirdly, for both of these hypotheses, this study examines whether the relationship between smokers and the (updated) subjective probability of survival varies for heavy smokers. Fourthly, to tests the hypothesis that smokers misperceive the risk of smoking, this study compares smokers' subjective survival probabilities with their objective counterparts. The focus of this part is to examine whether smokers' errors indicate optimism or pessimism relative to never smokers, as all individuals may be inaccurate in predicting survival in general. Fifthly, to assess the accuracy of predicting survival in general, irrespective of the prediction being optimistic or pessimistic, I examine associations between smoking status and absolute differences in error. Sixthly, for both the error and the absolute error, I examine whether the association between smokers and the error is different for heavy smokers. Finally, to answer the research question, this study examines the extent to which the associations between smoking status and survival beliefs differ for different APC groups.

"The structure of this study is as follows: The background literature is covered in Section 2. Section 3 describes the data sources used and presents the descriptive statistics of the sample. Section 4 covers the methodology of this study. Section 5 tests the hypotheses and presents the results. Section 6 discusses the methodology and results of the study. Finally, conclusions regarding the implications of this study are drawn in Section 7.

2. Theoretical background, models and prior research

2.1 Underlying theories and models

The changing demand for health investments and health within the lifecycle was first investigated and modelled by Michael Grossman (1972). In this model, health is considered as a durable capital stock which depreciates over time and can be improved by making health investments. The model predicts that the demand for health investments and health changes with age. These predictions are consistent with the study on lifecycle smoking behavior conducted by Carbone et al. (2005). This study, based on the Grossman model, examined how adverse effects from smoking affect incentives to invest in health at different points in one's life. Moreover, it found that the depreciating stock of damage caused by smoking provides an incentive to smoke more early in life and shift the demand for health investments to a later stage of life. Ehrlich and Becker (1972) modelled the chosen amount of effort an individual decides to invest on preventive health behavior to maximize expected utility.

Model of self-protection:

$$\max_{e} EU = \left(1 - p(e)\right)u(y - e) + p(e)u(y - L - e)$$

However, both Grossman's model and Ehrlich and Becker's self-protection model assume fully rational individuals, who use the models without distortions in their decision making and who know the probabilities of the different health states. On the contrary, these assumptions are found to be inconsistent with observed choices when individuals face conditions of uncertainty or ambiguity (Kahn and Sarin, 1988). Therefore, decision weights rather than the actual probabilities were included in the self-protection model. Evidence of probability distortions in decision making would hence result in suboptimal self-protective behavior (Kahneman and Tversky, 1979; Kahn and Sarin, 1988). This justifies encouraging preventive behavior (anti-smoking) to correct for internalities and not just to compensate for externalities.

Model of self-protection with weighted probabilities:

$$\max \left(1 - w\left(p(e)\right)\right)u(y - e) + w\left(p(e)\right)u(y - L - e)$$

2.2 Smoker's risk perception and survival probability errors

The role that perceptions of risk and probability distortions play in an individual's decision to smoke has primarily been demonstrated by Viscusi and Kip (1990). In their study, they examined the effect of lung cancer risk perceptions on the decision to smoke. In short, individuals generally overestimated the likelihood of lung cancer, which discourages the decision to smoke. According to the smoking decision model presented in their study, an individual decides to smoke if the benefits of smoking outweigh the costs of smoking. In this equation, pessimistic survival probability (s) errors discourage the decision to smoke. Whereas optimistic survival probability errors (s) would encourage the decision to smoke. Furthermore, differences in the internal costs of smoking (V) also affect the smoking decision. This is consistent with research on the internal cost of smoking conducted by Kwaja et al. (2009), which demonstrated that smokers do value a loss in health or longevity less than non-smokers. One plausible explanation is that this may reflect differences in time preferences. However, this study only focuses on the role that misperceptions of risk play in the decision to smoke.

Smoking decision model:

$$[U(\text{smoke}) - U(\text{don't})] + s[V - U(\text{don't})] > 0.$$

In addition, Viscusi and Kip (1991) examined the variation in smokers' risk perception by age and smoking status to create a risk perception model within a Bayesian learning framework. In summary, the differences in risk perception allowed to visualize how variations in risk perception are transmitted into the decision to smoke. In this model, which consists of three sources of information, (*p*) refers to the prior risk assessment, (*q*) indicates the risk information through experience and (*r*) refers to the risk information that has been presented to the individual. For younger birth cohorts, the accumulated risk perception is likely to be primarily driven by prior risk (*p*) and publicly presented risk information (*r*), whereas for earlier cohorts and older individuals, experience (*q*) is likely to play a larger role. Viscusi and Kip concluded that the accumulated perceived risk of smoking is greater for younger individuals and that this is negatively associated with smoking behavior. Furthermore, these younger individuals were found to substantially overestimate the risks of smoking. These findings are consistent with Liu and Hsieh (1995) & Lundborg and Lindgren (2004), who replicated this study approach in Taiwan and Sweden.

Risk perception model:

$$RISK = \frac{\Psi_0 p + \gamma_0 q + \xi_0 r}{\Psi_0 + \gamma_0 + \xi_0} = \Psi p + \gamma q + \xi r,$$

However, Schoenbaum (1997) criticized the method used by Viscusi to measure perceptions of smoking risks. As the survey questions about lung cancer perception fail to quantify the internal costs to the individual of this disease. Moreover, these questions do not measure quantifiable perceived years of life lost due to lung cancer or other adverse health effects associated with smoking. In addition, Viscusi and Kip have been criticized for measuring the lung cancer risks of hypothetical smokers rather than the risk implications faced by the individual him-or-herself. While prior research has shown that when people are exposed to the same risk, they often perceive themselves to be at less risk than other people (Weinstein, 1989). As a result, questions measuring the risks of hypothetical smokers would not reflect an accurate assessment of the smoking risks perceived by the individuals themselves. Therefore, Schoenbaum examined the risk perception of survival using a mature (ages 50-62) sample of smokers, former smokers and heavy smokers by comparing the subjective probability of surviving to age 75 with the epidemiological objective probability of survival. The study concluded that only heavy smokers significantly overestimated their survival probability (optimistic survival probability errors). In 2009, Kwaja et al. (2009) replicated and extended this study approach with a 50- to 70-year-old sample and found that current smokers were relatively optimistic about their survival probability relative to never smokers. Furthermore, no significant differences were observed between the errors of former smokers and those who had never smoked. The study by Sloan et al. (2011) was the first to examine the relationship between smoking status and survival probability errors using a sample of adolescents (ages 12-18). This study concluded that young people are generally pessimistic about their probability of survival, and that this degree of pessimism increases after smoking initiation.

The increase in pessimism after smoking initiation is likely to be due to the source of risk information through experience (q) in the smoker's risk perception model (Viscusi and Kip, 1991). This theory is consistent with Hsieh (1998), who stated that negative changes in perceived health status are associated with smoking cessation. In addition, Smith et al. (2001) demonstrated that smokers and non-smokers differ in how information affects their subjective longevity expectations. The latter and the fact that smokers perceive risk through experience (q) suggests that smokers do update their subjective probability of survival differently than non-smokers. Under the assumption of no significant differences in non-smoking related health, the extent to which smokers update their subjective survival probabilities differently than non-smokers can be interpreted as the updated perceived risks of their smoking behavior.

In existing research, the methods and samples used to examine the association between smoking status and errors in the probability of survival have varied. Schoenbaum (1997) derived the subjective probabilities of surviving to age 75 from the 1992/1993 Health and Retirement Survey (HRS) with a sample between the ages of 50 and 62. The objective probabilities were derived from smoking-specific lifetables generated with data from the 1986 United States National Mortality Followback Survey (NMFS) and the 1985 and 1987 National Health Interview Surveys (NHIS). To examine the relationship between smoking status and survival probability errors, this study used sample means rather than measures of accuracy separately for each age. In the discussion section, Schoenbaum highlighted that the error may have been affected by reasons unrelated to smoking status, such as cohort and period differences between the datasets used for subjective and objective survival probabilities. On the other hand, Schoenbaum stated that differences between subjective and objective probabilities unrelated to smoking status may be due to the fact that individuals are generally unable to predict survival. This motivates the rationale for comparing the differences relative to never smokers, rather than focusing on the (optimistic or pessimistic) sign of the error by smoking status.

Unlike Schoenbaum and other prior research, instead of aggregate comparisons using sample means, Kwaja et al. (2009) examined the relationship between smoking status and errors by comparing individual-specific subjective and objective probabilities of surviving to age 75. This study was conducted using a sample with individuals between the ages of 50 and 70. In their paper, the individual-specific objective probabilities were estimated based on conditional mortality rates that were derived from seven waves of the HRS study (1992-2004). These were assumed to be an objective counterpart to the subjective survival probabilities derived from the 2004/2005 Survey of Smoking (SOS). In their analysis, compared to Schoenbaum, individual-specific objective probabilities likely improved the accuracy of the results. However, because the subjective and objective probabilities were derived from surveys conducted in different years, period and cohort effects are still likely to bias the relationship. Another limitation of these studies is that, both Schoenbaum and Kwaja et al. derived the subjective probability of survival from cross-sectional data. While using longitudinal data makes it possible to link (changes in) the perception of risk to actual smoking behavior. In addition, given the mature (ages 50-70) samples used, no conclusions can be drawn about survival beliefs at the point in life where smoking initiation occurs.

An analysis of the association between survival probability errors and smoking status among adolescents (ages 12-18) has been conducted by Sloan et al. (2011). In their study, subjective survival probabilities were derived from the (1997-2006) National Longitudinal survey of Youth (NLSY97). Since deaths were also reported in this survey, objective survival probabilities were calculated from parameter estimates that were obtained using logit. As a result, both observed and unobserved heterogeneity are common to the subjective and objective probabilities. However, this study examined the accuracy of the subjective probability of surviving to age 20, whereas most of the adverse effects of smoking are known to occur later in life.

3. Data

3.1 The LISS panel

The data used in this study is derived from the Longitudinal Internet studies for the Social Sciences (LISS) panel, which is managed by research institute Centerdata. The LISS panel consists of about 7500 individuals, who are distributed over approximately 5000 households. The two main sources of data from the panel archive used in this study are the monthly updated background variables and the annual conducted LISS Core Study. The background dataset provides panel member demographics. While the LISS Core Study provides repeated measures using questionnaires on a wide range of topics. For this study, the Health Questionnaire is used. This questionnaire measures a variety of health outcomes, health perceptions and health behaviors of the panel members. The main concepts of interest are subjective health, life expectancy and the use of tobacco products. For analysis, this study uses longitudinal data from waves (8-15) of the Health Questionnaire. These waves cover the years from 2015 to 2022. One of the advantages of using longitudinal data is that it allows for both within and between variation in smoking status. As a result, panel data allows to assess how individuals change their subjective health beliefs during smoking initiation and cessation. In addition, the Health Questionnaire allows to control for other factors that are likely to affect one's subjective survival probability, such as alcohol consumption and the body mass index (BMI).

The main question used as the dependent variable of interest is: "How would you rate your chance of living to be 75 years or older?". Panel members were asked to report a value between 0 and 10, with 0 indicating "no chance at all" and 10 indicating "absolutely certain". Similar to previous research**,** these values are converted to probabilities (Schoenbaum, 1997; Kwaja et al., 2009). These outcomes, where 10 has been converted to 100% (10*10), are used to examine associations between smoking status and the subjective probability of survival. Furthermore, to examine an association between smoking status and errors in survival beliefs. Age, period and cohort categories have been created to examine potential APC effects. All age groups have a width of five years, except for the first and the last group, which have a width of six years (Table 1). This decision was made to ensure sufficient observations for the minimum and the maximum ages. Similarly, cohort groups have a width of five years except for the first and the last two cohorts. The latter allows for the separate examination of more recent cohorts, which are likely to have had greater exposure to anti-smoking campaigns and evolving prior risk assessments. For the time periods, each survey wave (2015-2022) is included as a categorical variable. To distinguish between smoking status, two survey questions have been used: "Have you ever smoked?" and "Do you smoke now?". These questions allowed the categorization of never smokers, former smokers, and current smokers. In addition, current smokers were defined as heavy smokers if they reported smoking at least 20 cigarettes or 10 cigars per day. This is consistent with how the CBS (Centraal Bureau voor Statistiek) defines heavy smoking. The CBS definitions were also used to categorize origin and educational attainment. For origins, this study distinguishes between individuals with a Dutch background, Western immigrants and non-Western immigrants. Furthermore, educational attainment is classified as low, medium and high.

Longitudinal datasets, such as the LISS panel, deal with attrition and missing values. In that case, individuals in the sample do not have the same number of observations. Without adjustments, this unbalanced panel dataset would introduce omitted variable bias due to heterogeneity in the data. Balancing the data to ensure that each individual was observed in all sample years reduced the number of id observations from 48820 to 20620. However, only 13847 of these 20620 had reported their subjective probability of surviving to 75. After removing missing values for subjective survival probability and ensuring that all individuals are still observed each wave, the sample consisted of 10112 observations. In addition, cleaning of the variables in the dataset has been done to ensure accuracy, consistency and completeness. In this study, the number of cigarettes per day for current smokers was assumed to be constant when a missing value was observed (assumed to be <20 if never reported). Outliers due to data entry errors for weight and height have been corrected. Furthermore, never smokers who had reported being a smoker in previous years were reclassified as former smokers. After all, this resulted in a perfectly balanced full sample consisting of 9912 observations.

In addition to the full sample, this study uses an adjusted sample to test the extent to which there is a moderating effect for APC groups on the relationship between smokers and (mis) perceptions of survival. This adjusted sample uses the balanced full sample, but only includes respondents that reported being a current smoker in the baseline year 2015. Consequently, never smokers are excluded, and respondents can only be former smokers from 2016 onwards. These adjustments resulted in an adjusted sample consisting of 1816 observations.

3.2 Data sources for the computation of objective probabilities

The input used to compute objective probabilities as a counterpart to the subjective probability is derived from two sources. First, the CBS Open data StatLine was used to derive gender-specific life tables for the Netherlands for the years 2015-2022. In Excel, these life tables have been transformed into tables showing the average probability of surviving to 75 for each age (see Appendix B). Second, the smoking-specific life tables generated by Swanson et al. (2020) were used to derive relative risks of surviving to age 75 by smoking status. In their paper, life tables for never smokers, former smokers, current smokers and heavy smokers have been generated with existing data from the Kaiser Permanente Smoking Study, United States (Friedman et al. 1997). Similarly, Excel was used to transform the life tables into tables showing the probability of surviving to age 75 for each age interval (see Appendix C). The approach used to make these tables representative of the Dutch population and the computation of the objective probabilities of survival will be explained in more detail (See Section 4.2).

3.3 Descriptive statistics

Table 1 presents the descriptive statistics for the two samples used in the analysis. The full sample (1) consist of 1239 unique individuals and 9912 observations, of which 48% are male and 52% female. As smoking status could be a time-varying variable, Table 1 shows that on average 46% of individuals are identified as never smokers. In addition, 40% reported being former smokers and 14% were current smokers, of whom 4% were classified as heavy smokers. The mean reported subjective probability of survival, which is discussed in more detail in Section 5.1, is 71.73% for male and 72.00% for female (Table 2). Furthermore, the distribution of the subjective probability of survival responses is shown in Figure 1. The sample distribution of CBS-based categorized educational attainment levels does not deviate far from the national distribution. The full sample consist of 24% lower educated, 39% medium educated and 36% higher educated individuals. In comparison, the national distribution in the survey years (2015-2022) was 30% to 25% lower, 39% to 38% medium and 30% to 36% higher educated individuals (Centraal Bureau voor de Statistiek, 2022). In addition, 87% in the sample are classified as having a Dutch background, 7% as having a Western background and 6% as having a non-Western background. Respondents are 51.57 years old on average during the sample period, with a standard deviation of 11.54 years. The minimum age at the baseline year 2015 is 17 years, while the maximum observed age in the final wave is 73 years. Thus, the outcome variables, which include someone's subjective probability of surviving to age 75, remains relevant for all respondents in the sample. For further analysis, respondents are categorized into the age group that corresponds to their age. As shown in Table 1, most observations (19%) belong to age group 9, which includes respondents between the ages of 58 and 62. Similarly, respondents are categorized by cohort group. The average year of birth for the full sample is

1967, with the oldest born in 1949 and the youngest born in 1997. Slightly more than one-fifth of the observations belong to the 1956-1960 cohort (21%). As the LISS panel is an annual survey, each survey wave contains 1239 observations.

In addition, the second column of Table 1 presents the descriptive statistics for the adjusted sample (2). As shown, 30% of the observations reported being a former smoker (possible from 2016 onwards) and 70% were current smokers, of whom 20% were classified as "heavy smokers". Male and female both comprise 50% of the sample observations. Since smoking is known to be associated with educational attainment, the adjusted sample contains more (31%) lower educated individuals and less (28%) higher educated individuals than observed in the full sample (1). Both the distribution of origins and the average year of birth (cohort) have not changed. The average age of the sample decreased slightly to 51.16 years. In addition, all periods (2015-2022) in the sample contain 227 observations (Table 1, col 2).

Table 1.

Descriptive statistics of the full (1) and adjusted (2) sample

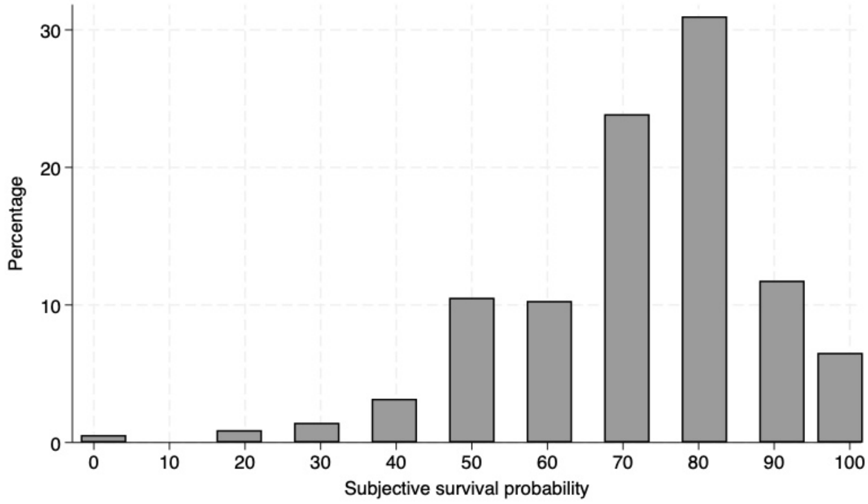
	Full Sample (1)		Adjusted sample (2)			
	n	Mean (*SD*)	*n*	Mean (*SD*)	*Min.*	*Max.*
Gender	9912		1816			
Male	4744	0.48 (0.50)	912	0.50 (0.50)		
Female	5168	0.52 (0.50)	904	0.50 (0.50)		
Smoking status	9912		1816			
Never smoker	4578	0.46 (0.50)	-	-		
Former smoker	3930	0.40 (0.49)	539	0.3 (0.46		
Current smoker	1404	0.14 (0.35)	1277	0.7 (0.46)		
Heavy smoker	373	0.04 (0.19)	366	0.20 (0.40)		
Education	9912		1816			
Lower educated	2408	0.24 (0.43)	568	0.31 (0.46)		
Medium educated	3904	0.39 (0.49)	744	0.41 (0.49)		
Higher educated	3600	0.36 (0.48)	504	0.28 (0.45)		
Origin	9912		1816			
Dutch Background	8608	0.87 (0.34)	1576	0.87 (0.34)		
Non-western	632	0.06 (0.24)	112	0.06 (0.24)		

Western	672	0.07 (0.25)	128	0.07 (0.26)		
<u>Age</u>	9912	51.57 (11.54)	1816	51.16 (11.57)	17	73
Age G1 (17-22)	111	0.01 (0.11)	22	0.01 (0.11)		
Age G2 (23–27)	311	0.03 (0.17)	65	0.04 (0.19)		
Age G3 (28–32)	456	0.05 (0.21)	106	0.06 (0.23)		
Age G4 (33–37)	483	0.05 (0.22)	71	0.04 (0.19)		
Age G5 (38-42)	676	0.07 (0.25)	115	0.06 (0.24)		
Age G6 (43-47)	1075	0.11 (0.31)	189	0.10 (0.31)		
Age G7 (48-52)	1414	0.14 (0.35)	263	0.14 (0.35)		
Age G8 (53-57)	1819	0.18 (0.39)	342	0.19 (0.39)		
Age G9 (58-62)	1910	0.19 (0.39)	379	0.21 (0.41)		
Age G10 (63-67)	1231	0.12 (0.33)	214	0.12 (0.32)		
Age G11 (68-73)	426	0.04 (0.20)	50	0.03 (0.16)		
<u>Cohort</u>	9912	1967(11.30)	1816	1967(11.32)	1949	1997
1949-1955	1424	0.14 (0.35)	192	0.11 (0.31)		
1956-1960	2064	0.21 (0.41)	464	0.26 (0.44)		
1961-1965	1856	0.19 (0.39)	320	0.18 (0.38)		
1966-1970	1400	0.14 (0.35)	264	0.15 (0.35)		
1971-1975	1136	0.11 (0.32)	176	0.10 (0.30)		
1976-1980	728	0.07 (0.26)	160	0.09 (0.28)		
1981-1985	384	0.04 (0.19)	32	0.02 (0.13)		
1986-1990	512	0.05 (0.22)	152	0.08 (0.28)		
1991-1994	240	0.02 (0.15)	16	0.01 (0.09)		
1995-1997	168	0.02 (0.13)	40	0.02 (0.15)		
Period	9912		1816		2015	2022
Total observations	9912		1816			

Note. Table provides observations (n), means, standard deviations (SD), minimum and maximum values.

Figure 1.

Distribution of the subjective probability of surviving to age 75



Note: x-axis is the subjective probability of surviving to age 75 (percentage).

4. Methodology

4.1 Aim study and rationale for the outcome variables

This study has three main objectives. First of all, to examine whether smokers report lower subjective survival probabilities than non-smokers and thus perceive the health risks of smoking. Secondly, to examine whether they are relatively optimistic about their chances of survival compared to non-smokers. Thirdly, it examines the extent to which these associations are affected by differences in age, time-period and cohort. Individual perceptions about the health risks of smoking are assessed by comparing smokers and non-smokers their subjective probability of surviving to age 75. Because existing research has highlighted the difficulty of quantifying and comparing the perceived internal costs of smoking-related diseases, this study rather uses subjective years of life lost (Schoenbaum, 1997).

In addition, the annual measures of subjective survival probabilities allow the assessment of changes in subjective survival over time (1). This is used to examine whether smokers change their subjective probability of survival differently than non-smokers. According to Viscusi and Kip (1991), the extent to which a smoker updates his or her subjective beliefs differently than those who have never smoked could be induced by two sources of information. First, smokers could have been presented with new information about their smoking behavior between period t and t-1. Second, new information about risks may have emerged from personal adverse experiences, such as a health shock.

$$\text{Subjective change}_{it} = \text{subjective survival probability}_{it} - \text{subjective survival probability}_{(t-1)} \quad \textbf{(1)}$$

In order to examine smokers' relative optimism about their chances of survival, errors are calculated and used as a dependent variable. These errors indicate the difference between the reported subjective probability of survival and the computed objective probability of survival (2). Thus, a positive error (0>) would indicate optimism about the probability of survival. Whereas a negative error (<0) would indicate pessimism about the probability of survival. Being (relatively) optimistic about the probability of survival could encourage smokers to keep smoking and non-smokers to start smoking. Conversely, strong pessimism about survival could also affect smoking initiation as the adverse health effects of smoking tend to occur in a later stage of life. Although this study distinguishes between current, former, heavy and never smokers, the analysis is focused on the comparison between current and never smokers. In line with prior research, former smokers are included but of less interest as most data does not provide information on how long and how much someone has smoked before quitting, making it difficult to interpret their computed survival probability error.

$$\text{Error}_{it} = \text{subjective survival probability}_{it} - \text{objective survival probability}_{it} \quad \textbf{(2)}$$

Another point of consideration is that associations between smoking status and errors do not necessarily indicate relative optimism or pessimism about their survival. Moreover, smokers may be unusually misinformed in predicting survival in either direction. In this case, larger errors for smokers could reflect the inability to predict survival instead of smokers being highly optimistic or pessimistic about their survival. It is therefore also important to assess and compare the overall accuracy of survival predictions. To examine this, absolute differences (how much does the error deviate from 0, irrespective of whether it is negative or positive) between subjective and objective probabilities are calculated and compared by smoking status.

4.2 Computation of objective survival probabilities

In an ideal study, the association between (relatively optimistic) survival beliefs and smokers would be examined by comparing reported subjective probabilities of survival with sample-specific objective counterparts. However, the LISS panel and other accessible datasets that measure health beliefs do not report information about deceased panel members. Therefore, this study uses computed objective survival probabilities as a counterpart. While all objective risk measures are subject to specification error, this study uses formulations that are much more detailed than used in existing research. This because the objective survival probabilities in this study are unique for each age, gender, survey year and smoking status, while previous studies have ignored period and cohort effects in the estimation of objective survival probabilities. The formula (3) is shown below and will be explained in more detail.

$$\text{Objective survival probability}_{it} = \text{Relative risk}_{it} * \text{AVG survival up to 75 (by age, year and gender)} \quad \textbf{(3)}$$

The first part of the formula (3) indicates an individual's relative risk (RR) of survival in a given year, relative to the average. Thus, an RR of 1.07, belonging to a 20-year-old male in 2018 who has never smoked, implies that this individual has a probability of reaching age 75 that is 1.07 times higher than the average probability of reaching age 75 for a 20-year-old male in 2018. The second part of the formula (3) indicates the average probability of surviving to age 75 in the Netherlands by age, year and gender. Moreover, gender-specific life tables published annually by the CBS have been transformed into tables showing the probability of surviving to

75 years by age (see Appendix B). This was done for all sample years, both genders and all ages between 17 and 74, resulting in 928 unique average probabilities.

The relative risks are computed on the basis of smoking-specific life tables, which were generated by Swanson et al. (2020). In these tables, the mortality rates by age, gender and smoking status are presented at five-year intervals (ages 20-90), with the mortality rate indicating the probability of dying before the next interval. These life tables have been transformed into tables showing the probability of surviving to age 75 for each age interval (see Appendix C). This allows the probability of reaching age 75 to be compared by smoking status.

However, these probabilities are unlikely to be representative of the Dutch population, as they correspond to a sample in California, United States. This can be evaluated by applying Swanson et al. (2020) their method to transform these lifetables to other populations. In order to apply this method, information on the proportion of never smokers, former smokers and current smokers in the population is required. The rationale is to weight the smoking-specific tables by their proportions in the Dutch population in a given year. This results in a table showing the weighted average probability of reaching age 75 by age interval. If the generated smoking-specific tables were representative of the Dutch population, the weighted average table should now display accurate average probabilities of surviving to age 75 in the Netherlands. Comparisons by age between the weighted average table and the actual average probability in the Netherlands show that the smoking-specific tables of Swanson et al. (2020) need to be adjusted to be representative of the Dutch population.

First of all, the smoking-specific tables are weighted by the Dutch proportions of smokers by gender in a given year. For example, among males in 2018, CBS found that 25.7% were current smokers, 34.6% former smokers and 39.7% never smokers. This results in a weighted average table consisting of 0.257*(survival probabilities current smoker) + 0.346*(survival probabilities former smoker) + 0.397*(survival probabilities never smoker). The next step in determining objective probabilities is to calculate the relative risks by dividing each group's survival probabilities by the weighted average. To illustrate, if a male never smoker aged 20 in 2018 has a 77% probability of surviving to age 75, while the calculated weighted average for a 20-year-old male is 72%. This would give in an RR of 1.07 (77/72). Because this is done for all age intervals, both genders and for all 8 years in the sample, this results in 704 unique relative risks ratios. To obtain objective probabilities that are representative of the Dutch population, we multiply the computed relative risks (relative to the weighted average) by the actual average in the Netherlands. So, if the actual average probability of surviving to age 75 in the Netherlands for a 20-year-old male in 2018 would be 76%. This would give this individual an objective probability of 81.32% (Objective probability= 1.07 RR * 76% AVG). While the actual average probabilities are available for all ages, the relative risks are provided for each age interval. Therefore, relative risks are assigned through the age categories presented in Table 1. Thus, an individual aged 28-32 receives the RR corresponding to a 30-year-old.

4.3 Structure of the analysis and methods used

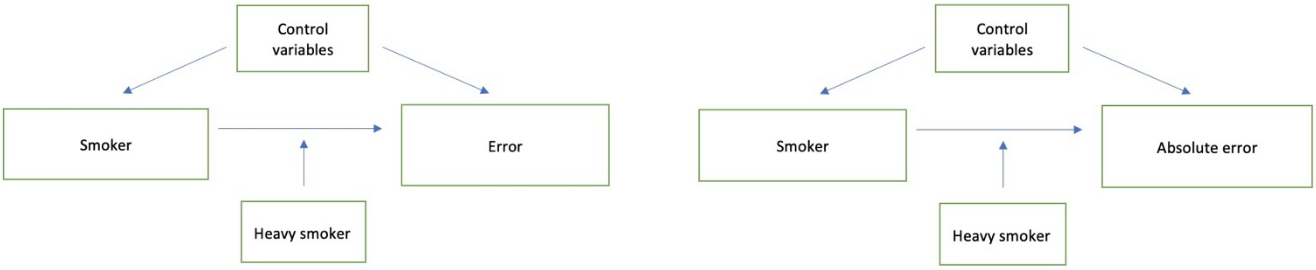
The results section of this study is divided into two parts. The first part examines associations between smoking status and survival probability beliefs using the full sample. This part includes both testing the hypothesis that smokers perceive the health risks of smoking and testing the hypothesis that smokers are relatively optimistic about their chances of survival compared to non-smokers. In addition to these main questions, the first part of the results section addresses three sub-questions. First, if smokers do update their subjective survival probability differently than non-smokers Second, if smokers are worse predictors of survival than non-smokers in general. Third, if the examined associations between smokers and survival probability beliefs differ for heavy smokers. The second part of the results section uses the adjusted sample and includes age, period and cohort categories to test the extent to which the associations in the first part differ for different APC groups.

To examine the extent to which smokers perceive the health risks, the association between smokers and their subjective probability of survival in comparison with non-smokers is evaluated. In addition, this study examines the sub-questions if smokers do update their subjective probability of survival differently than non-smokers and if there is a moderating effect for heavy smokers (Figure 2).

Figure 2.

The relative degree of optimism about survival is examined by evaluating the association between smokers and the error compared to never smokers. In addition, the association between smoking status and the absolute error is examined to assess accuracy of survival predictions in general. For both of these associations, the potential moderating effect of being a heavy smoker is analyzed (Figure 3).

Figure 3.



The second part of the results section focuses on how the examined associations between smokers and survival probability beliefs are affected by differences in age, cohort and period (Figure 4).

Figure 4.

The associations in Figure 2 and Figure 3 are examined by conducting panel data regressions on the full sample using statistical software STATA. In addition, regressions and plots using the adjusted sample and figures using the full sample are used to examine a moderating APC effect (Figure 4). For panel data, both the time-invariant and time-varying part of the error term should be uncorrelated with the variable of interest to produce unbiased causal estimators. However, zero correlation between the error term and the variable of interest over time is very unlikely when using longitudinal data. Therefore, this study examines the hypotheses using a random effects (RE) panel data regression as a non-causal estimation technique. In addition, standard errors are clustered at the individual level. The RE panel data regression allows to account for the panel structure of the data and eliminates part of the time-invariant and time-varying heterogeneity that could bias the association. Furthermore, the clustered standard errors allow to account for serial correlation within individuals.

To capture the effect of the included control variables on the analyses, two different specifications of the model are used. Specification (4) includes only smoking status as an independent variable. While specification (5) is an extension of the model with the inclusion of potential confounders.

$$Y_{it} = a + \beta_1 C_{it} + \beta_2 F_{it} + \beta_3 (C_{it} \times H_{it}) + \epsilon_{it} \quad (4)$$

$$Y_{it} = a + \beta_1 C_{it} + \beta_2 F_{it} + \beta_3 (C_{it} \times H_{it}) + \beta_x X_{it} + \epsilon_{it} \quad (5)$$

In these regression forms, Y_{it} indicates the outcome variable (subjective survival probability, change in subjective probability of survival, error and the absolute error), a is the constant, C_{it} is a dummy variable indicating whether a person is a current smoker, F_{it} is a dummy variable indicating whether a person is a former smoker, $C_{it} \times H_{it}$ is the interaction term between the dummies indicating current and heavy smoking, and ϵ_{it} indicates both the time-invariant and time-varying part of the error term. Furthermore, β is the estimated coefficient and the $_i$ and $_t$ stand for individual and time specific. In addition, specification (5) extends the model, where X_{it} is a vector of control variables (age, gender, survey year, birth cohort, level of education, origin, BMI and alcohol consumption).

In the estimation of the dependent variables, a positive β indicates a positive association between Y and the independent variable of interest. For categorical variables, this association is relative to the omitted category. Thus, a positive coefficient (β_1) for current smokers (C_{it}) means that current smokers are positively associated with the dependent variable (Y_{it}), relative to never smokers. In the estimation of the error, Y values can be either positive (subjective probability overpredicts the objective probability) or negative (subjective probability underpredicts the objective probability). So, if in this estimation the β for current smokers is positive, this means that current smokers are relatively optimistic about their probability of survival compared to never smokers. When examining the moderating APC effects, the control variables age, survey year and cohort in specification (5) are replaced by categorical variables. In this model, coefficients for age groups are interpreted relative to being in age group 11 (68-73), cohorts relative to belonging in 1956-1960 and periods relative to 2015. These are chosen considering sufficient observations.

5. Results

5.1 Examination of the association between smoking status and survival probability beliefs

The mean probability of surviving to age 75 and the corresponding errors by gender and smoking status are presented in Table 2. On average, male never smokers reported the highest subjective probability of survival (73.49%). While male heavy smokers reported the lowest subjective probability of survival (63.00%). The same pattern can be observed for females. As shown, female never smokers reported an average subjective probability of 72.88%. This was lower (67.48%) for current smokers and the lowest (63.73%) for heavy smokers (Table 2, column 1). The observed pattern of the mean subjective survival probabilities for both genders suggests that smokers recognize health risks of smoking when reporting longevity probabilities. However, the differences in means could also be due to differences between the groups in other factors that affect subjective survival probabilities such as age, cohort, alcohol consumption and BMI. Therefore, the regression analysis in Table 3 assesses the differences in means while holding constant other factors that are likely to affect the probability of survival.

The second column in Table 2 presents the mean computed objective probabilities by gender and smoking status. Regardless of their smoking status, females are more likely to survive to age 75 than males. Similar to the mean subjective probabilities, a pattern by smoking status can be observed. For example, female never smokers on average have an objective survival probability of 86,49%. This is lower (78.15%) for current smokers and the lowest (74.31%) for heavy smokers. However, in addition to smoking status and gender, these objective probabilities depend both on an individual's age and the year of survey. Therefore, relative differences in means by smoking status cannot be interpreted as effects on objective probabilities.

The third column in Table 2 demonstrates the mean computed error (difference between one's subjective and objective survival probability) by smoking status and gender. As shown, independent of smoking status and gender, all mean errors indicate pessimism regarding the probability of surviving to age 75. On the other hand, the smaller errors for smokers suggest that current smokers are relatively optimistic about survival (-5.82 and -10.66, for males and females respectively) compared to never smokers (-9.07 and -13.61, for males and females respectively). Furthermore, males seem relatively optimistic compared to females. However, both the mean subjective and objective probability of survival could have been affected by group differences, so the regression in Table 4 assesses these associations while holding constant factors that could affect the error.

Finally, the fourth column shows the overall accuracy in either direction of predicting the probability of survival by presenting the mean absolute differences. These results suggest that there are small differences between smoking groups in the ability to predict survival. Male smokers seem to make larger prediction errors (15.84) than those who have never smoked (13.36). Furthermore, males in general seem to be slightly more accurate (14.07) than females (16.04).

Table 2.

Mean probability of surviving to age 75 and errors by gender and smoking status

	Subjective prob.		Objective prob.		Error		Absolute error	
	Male	Female	Male	Female	Male	Female	Male	Female
Full sample	71.73 (17.52)	72.00 (16.34)	79.91 (4.95)	85.04 (3.64)	-8.18 (17.52)	-12.96 (16.30)	14.07 (13.26)	16.04 (13.28)
Never smoker	73.49 (16.15)	72.88 (15.74)	82.56 (2.25)	86.49 (1.89)	-9.07 (16.37)	-13.61 (15.88)	13.36 (13.11)	16.19 (13.24)
Former smoker	72.12 (17.81)	72.65 (16.5)	80.19 (3.54)	85.59 (2.47)	-8.07 (17.78)	-12.94 (16.44)	14.19 (13.41)	15.89 (13.62)
Current smoker	65.58 (19.23)	67.48 (17.22)	71.40 (4.38)	78.15 (3.74)	-5.82 (19.76)	-10.66 (17.17)	15.84 (13.17)	15.99 (12.35)
Heavy smoker	63.00 (18.66)	63.73 (19.29)	70.24 (4.94)	74.31 (3.51)	-7.24 (19.24)	-10.59 (18.91)	15.25 (13.76)	17.36 (12.93)

Note. Table provides means. The standard deviation is in parentheses.

Until now, associations between smoking status and the outcome variables of interest have been analyzed using group means. Table 3 and 4 show to what extent these associations are observable when controlling for factors that may have caused differences between groups. Table 3 presents the regression results on the subjective probability of survival and the annual change in this probability. Specification (4), which includes only the smoking status in the model, is presented in the first and third column of Table 3. In addition, the regression results of specification (5) are presented in the second and fourth columns. The constant in the first column shows that never smokers are estimated to report a subjective survival probability of 73.09%. Furthermore, being a current smoker (-3.46%) or former smoker (-1.58%) is significantly associated with a lower subjective survival probability compared to those who have never smoked, ceteris paribus. An even larger negative association (-4.83%) is found for heavy smokers, relative to never smokers. When potential confounders are added to the model in specification (2), the association remains significantly negative for current (-3.40%) and heavy smokers (-4.77%). These results, combined with the group means in Table 2, strongly suggest that smokers acknowledge, at least to some extent, the health risk of smoking on longevity. Furthermore, individuals with lower levels of education (-3.76%) and those with medium levels of education (-2.88%) are associated with lower subjective probabilities compared to those with higher levels of education, ceteris paribus. For non-Western immigrants, the subjective probability of survival is estimated to be 3.83 percent lower than for people with a Dutch background. Additionally, the subjective probability of survival is estimated to decrease 0.11% for every one-point increase in BMI (Table 3, col 2).

Columns three and four of Table 3 show that smokers (-0.09%) and heavy smokers (0.63%) do not update their subjective survival probability significantly different than those who have never smoked. However, former smokers (0.24% and 0.28%, without and with included factors respectively) are associated with a positive annual change in subjective survival probability, compared to never smokers. Thus, each year after quitting smoking is associated with a 0.28% increase in the subjective survival probability, compared to never smokers. Moreover, the constant in the third column shows that those who have never smoked, from year to year, also become more positive (0.15%) about the probability of surviving to age 75. Other factors excluded, this association should be present if reported rationally, as they get closer to 75 every year. Finally, non-Western immigrants are positively associated (0.67%) with updating their subjective probability of survival, compared to those with a Dutch background.

Table 3.

Linear regression results on dependent variables subjective survival probability and subjective survival probability change

	Subjective survival probability	Subjective probability change

	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)
Current smoker	-3.46*** (0.90)	-3.40*** (0.93)	-0.09 (0.28)	0.03 (0.29)
Former smoker	-1.58** (0.71)	-1.19 (0.75)	0.24* (0.13)	0.28* (0.15)
Heavy smoker	-4.83*** (1.39)	-4.77*** (1.40)	0.63 (0.58)	0.71 (0.59)
Age		-0.50 (0.43)		0.04 (0.18)
Male		-0.44 (0.78)		-0.11 (0.13)
Survey year		0.33 (0.45)		0.11 (0.19)
Cohort		-0.52 (0.43)		0.04 (0.18)
Lower educated		-3.76*** (1.07)		0.07 (0.18)
Medium educated		-2.88*** (0.88)		0.15 (0.14)
Western		-0.77 (1.58)		-0.14 (0.26)
Non-western		-3.83** (1.90)		0.67** (0.31)
BMI		-0.11* (0.06)		0.01 (0.02)
Alcohol		0.02 (0.16)		0.06 (0.05)
Constant	73.09*** (0.50)	458.36*** (135)	0.15* (0.09)	-316*** (99.37)
R^2 Within	0.00	0.00	0.00	0.00
R^2 Between	0.02	0.05	0.00	0.01
Observations	9912	9912	8673	8673

Note. Standard errors are in parentheses. Stars indicate the significance level; *** p<0.01, **p<0.05, *p<0.1. The number of observations for the subjective probability of change is lower (8673) due to the first change taking place in 2016. Alcohol coefficient reflects the categorical (1-8) amount of alcohol in the last 12 months, with 1 indicating almost every day, 2 indicates five or six days a week, 3 indicates three or four days a week, 4 indicates one or two days a week, 5 indicates once or twice a month, 6 indicates once every two months, 7 indicates once or twice a year and 8 indicates not at all.

The regression results on the error and the corresponding absolute error are presented in Table 4. The constant in the first column shows that never smokers are associated with a negative error (-11.6). This indicates that the subjective survival probability of those who have never smoked is estimated to be 11.6 percent lower than their objective survival probability. Regardless of smoking status, all groups appear to be pessimistic about their chances of surviving to age 75. However, the significant positive coefficients for current (5.32) and heavy smokers (7.92) indicate that smokers are relatively optimistic about their survival, compared to never smokers. This relative degree of optimism increases even more when the potential confounders are added in column 2. In the full model, being a current smoker is associated with a 5.75 increase in error, relative to never smokers. For heavy smokers, the survival error becomes even more optimistic, as it is estimated to increase by 8.08. In addition, this model shows a positive association (1.60) for former smokers relative to never smokers, ceteris paribus. Looking at the included confounders, males are associated with optimism (4.58) relative to females. For age, an additional year of age is negatively associated with the error (-0.83), ceteris paribus. Individuals with low (-3.78) and medium (-2.64) levels of education are relatively more pessimistic than those with higher levels of education. Lastly, non-Western immigrants are negatively associated (-3.76) with the error, relative to those with a Dutch background (Table 4, col 2).

For never smokers, the constant in column three shows that the absolute difference between the subjective and objective probability is estimated to be 14.85. Furthermore, columns three and four show that there are no significant differences between the smoking groups (-0.53, 0.85 and 0.77, for current, former and heavy smokers respectively) in the ability to accurately predict survival. On the other hand, columns one and two showed significant differences between these smoking groups in the regressions on error (Table 4, column 2). This suggests that these differences in error are due to (relative) optimism rather than smokers being better predictors of survival. The latter would have been an alternative explanation given that all smoking groups are found to be pessimistic about survival. In addition, the coefficients in column 4 show that an additional year of age is estimated to increase the absolute error by 0.85. Males are better able to predict longevity than females (-1.83). Individuals are associated with smaller absolute errors (-0.70) for each additional year surveyed, ceteris paribus. One-year increase in cohort is associated with a slightly larger absolute error (0.77). Low and medium educated individuals are estimated to predict longevity less accurately (3.57 and 2.06) than those with higher levels of education. Finally, non-Western immigrants are associated with larger absolute errors (4.02) compared to those with have a Dutch background.

Table 4.

Linear regression results on dependent variables error and absolute error

	Error		Absolute error	
	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)
Current smoker	5.32*** (0.94)	5.75*** (0.94)	-0.53 (0.76)	-0.82 (0.77)

Former smoker	0.15 (0.73)	1.60** (0.75)	0.85 (0.59)	0.15 (0.61)
Heavy smoker	7.92*** (1.42)	8.08*** (1.41)	-0.77 (1.09)	-1.07 (1.11)
Age		-0.84* (0.43)		0.85** (0.36)
Male		4.58*** (0.78)		-1.83*** (0.59)
Survey year		0.40 (0.45)		-0.70* (0.37)
Cohort		-0.67 (0.43)		0.77** (0.35)
Lower educated		-3.78*** (1.07)		3.57*** (0.82)
Medium educated		-2.64*** (0.88)		2.06*** (0.67)
Western		-0.89 (1.54)		0.59 (1.28)
Non-western		-3.76** (1.91)		4.02*** (1.38)
BMI		-0.08 (0.06)		0.06 (0.05)
Alcohol		0.00 (0.16)		0.04 (0.13)
Constant	-11.6*** (0.52)	529.85*** (135)	14.85***(0.42)	-132.41 (112)
R^2 Within	0.01	0.02	0.00	0.00
R^2 Between	0.00	0.07	0.00	0.05
Observations	9912	9912	9912	9912

Note. Standard errors are in parentheses. Stars indicate the significance level; *** p<0.01, **p<0.05, *p<0.1. Alcohol coefficient reflects the categorical (1-8) amount of alcohol in the last 12 months, with 1 indicating almost every day, 2 indicates five or six days a week, 3 indicates three or four days a week, 4 indicates one or two days a week, 5 indicates once or twice a month, 6 indicates once every two months, 7 indicates once or twice a year and 8 indicates not at all.

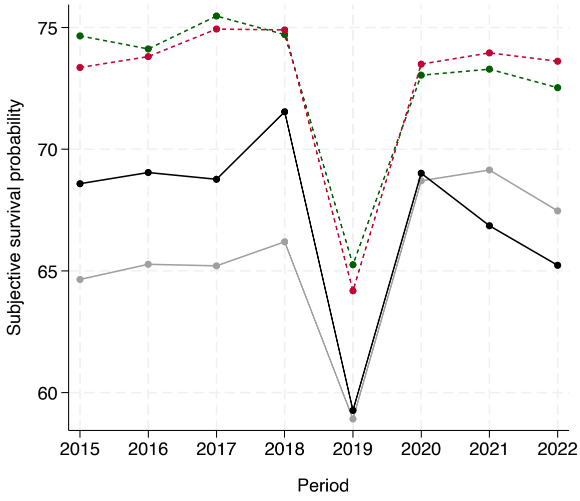
5.2 Examination of moderating age, period and cohort effects

Part two of the results section examines the extent to which the associations observed in part one differ for differences in age, period and cohort. As shown in Table 3, smoking status is associated with subjective beliefs about survival. In addition, the results in Table 4 demonstrated an association between smoking status and the relative degree of optimism about survival. The moderating effect for the APC groups is examined first by looking at group means figures for the full sample, and second by running a regression with corresponding marginal plots using the adjusted sample.

Figure 5.

Mean subjective survival probability and error over time, by smoking status and gender

5a.



5b.

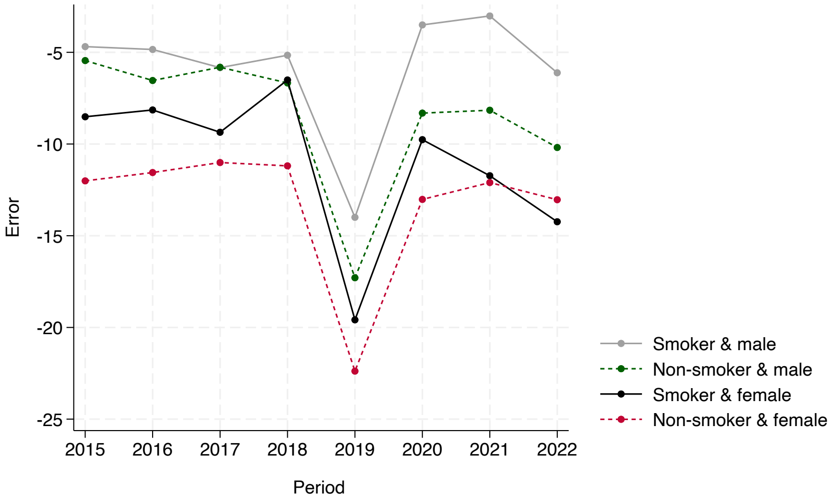


Figure 5 shows the differences over time between the group means of subjective survival probability and error. As indicated in the legend, the groups distinguish between smokers, non-smokers, males and females. During these years, various anti-smoking policies have been implemented in The Netherlands which could have affected smokers their subjective probability of survival such as, the ban on the use of color in cigarette packaging in 2020, the prohibition for retailers to place tobacco products in plain view of the consumer in 2021 and the ban on all indoor smoking areas in 2022. Figure 5a shows that the mean subjective probability of survival is higher for non-smokers than for smokers in all years for both males and females. For all groups, there appears to be a negative association between 2019 and one's subjective probability of survival. This is also illustrated in Appendix A, which shows the distribution of subjective survival probabilities by year. In addition, compared to 2015-2017, male smokers in 2020-2022 seem to have increased their subjective probability of survival. On the other hand, compared to all years prior to 2019, female smokers in 2021-2022 seem to be associated with a slightly lower subjective probability of survival. Figure 5b shows that compared to males, females are more pessimistic about their survival probability in all years. Except for females in 2022, smokers are more optimistic about survival than non-smokers of the same sex. The observable negative association between 2019 and the errors can be attributed to the change in subjective survival probability (see Figure 5a), as there were no significant differences in objective survival probabilities during this period.

The differences between the groups means of subjective survival probability and error by age category are presented in Figure 6. The patterns in Figure 6a show that differences between smokers and non-smokers in survival beliefs primarily emerge from age group 4 (33-37 years) onwards. This could be because people in this age group begin to experience adverse effects of smoking or realize that they are addicted. The subjective probability of survival appears to increase significantly from age group 9 (58-62 years) onwards, except for female smokers. The patterns in Figure 6b show that smokers of both genders are relatively optimistic about survival compared to non-smokers. However, this association seems to be reversed for people in age groups 10 and 11 (63-73 years). As in these age groups, the errors made by both male and female smokers are more pessimistic than those who don't smoke. Finally, although we observe a decreasing trend in the error across the age groups for both smokers and non-smokers, the trend is steeper for smokers compared to those not smoking (Figure 6b).

Figure 6.

Mean subjective survival probability and error by age category, smoking status and gender

6a. *6b.*

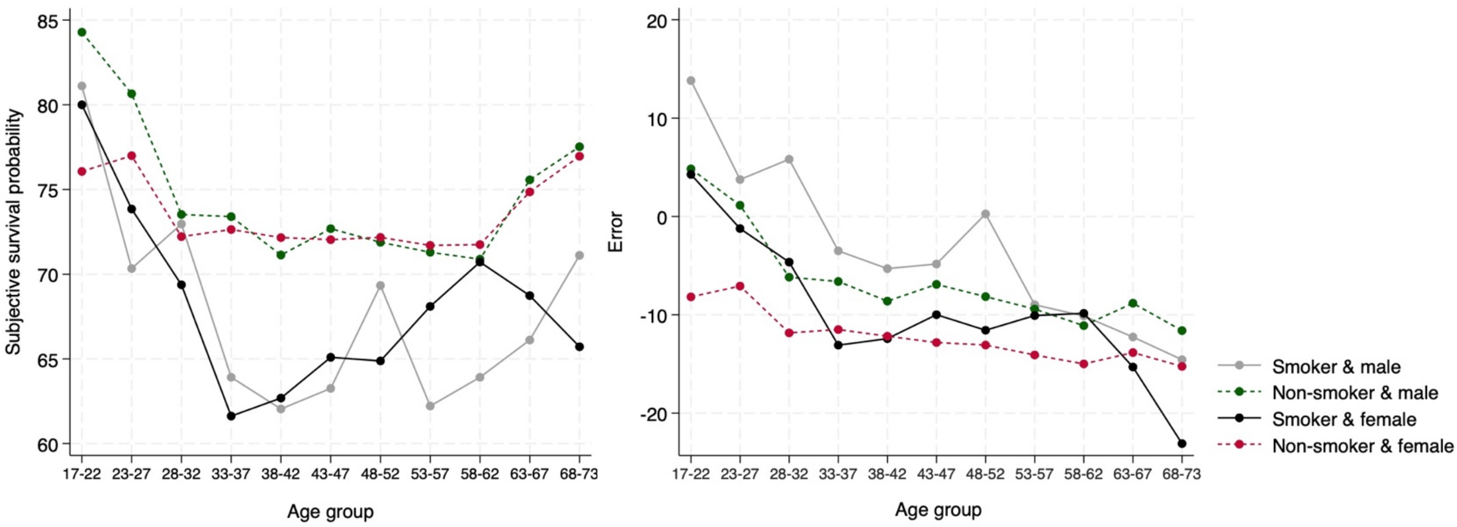
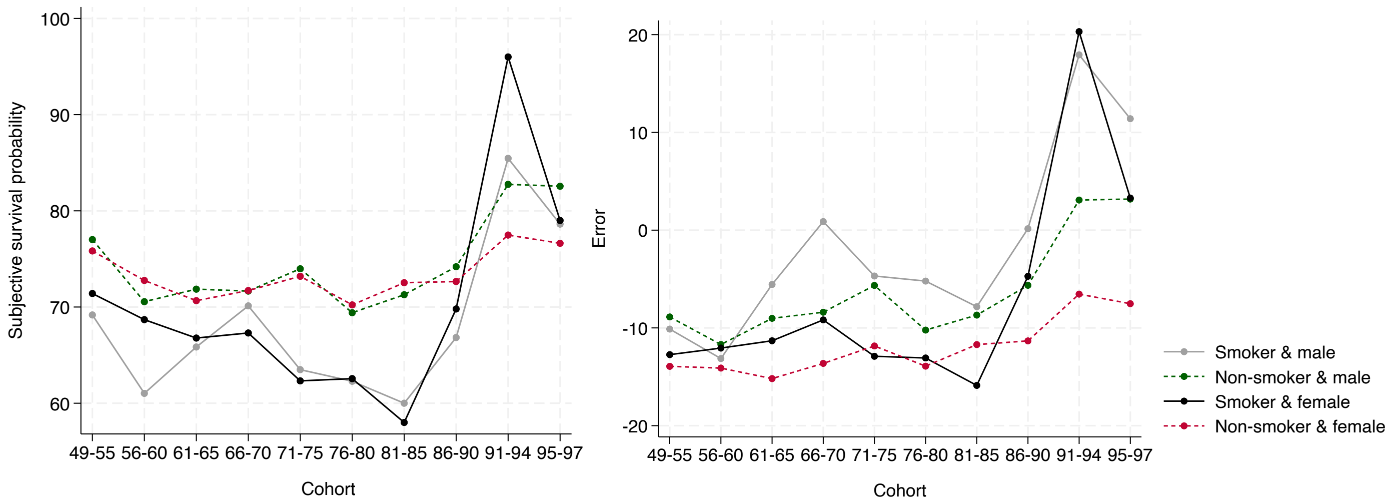


Figure 7 presents the differences in the group means of subjective probability of survival and error by cohort group. Figure 7a clearly demonstrates the difference in subjective survival between smokers and non-smokers across all cohorts, while there is almost no difference between males and females. As shown earlier in Figure 6a, younger people and therefore (in this analysis) people from recent cohorts tend to be associated with a higher subjective expectation of survival (Figure 7a). This is particularly the case for smokers who were born after 1986. Similarly, Figure 7b shows that smokers born after 1986 are relatively optimistic about their chances of survival, compared to non-smokers.

Figure 7.

Mean subjective survival probability and error by cohort, smoking status and gender

7a. *7b.*



Until now, the moderating APC effect has been analyzed with graphically presented group means. Table 5 and the corresponding marginal plots demonstrate to what extent these associations are observable when controlling for factors that may have caused differences between groups. The regressions conducted in Table 5 examine moderating age, period and cohort effects on the presented associations in part one of the results section between smokers and the outcome variables of interest. To examine this, the APC categories are added to specification (5) using the adjusted sample, which only includes individuals who reported being current smokers in the baseline year 2015. Columns one and two in Table 5 show that becoming a former smoker is not associated with a different subjective probability of survival, relative to current smokers. As previously suggested by Figure 7a, smokers in the 1991-1994 cohort are significantly associated with a higher (19.84%) subjective probability of survival relative to smokers in the 1956-1960 cohort. In addition, compared to smokers in 2015, smokers in 2019 are associated with lower (-6.63%) subjective survival probabilities. By contrast, 2020 is positively associated (3.04%) with subjective survival probabilities compared to 2015. For lower (-4.14%) and medium (-3.91%) educated smokers, the subjective probability of survival is estimated to be lower than for smokers with a higher level of education, ceteris paribus.

As previously illustrated in Figure 6a, smokers in age group 4 are estimated change their subjective survival probability by -11.5%, relative to smokers in age group 11. (Table 5, column 2). Although the figures and the coefficients suggest a similar association for smokers in the age groups 1 to 3, these are not found to be significant. In addition, smokers in the cohorts 1981-1985 (11.19%) and 1986-1990 (11.28%) are associated with a positive annual subjective survival probability change, relative to smokers in the cohort 1956-60. Finally, column 2 shows that relative to smokers in 2016, smokers in 2019 are associated with a -10.57% change in subjective survival probability. While smokers in 2020 are associated with positively (7.95%) updating their subjective probability of survival, compared to smokers in 2016.

The third column in Table 5 shows the regression on the error. Smokers who become former smokers are associated with pessimism (-5.88) about their chances of survival, relative to current smokers. One possible explanation could be that these former smokers quitted because they experienced adverse health effects. This has led them to not or negatively change their beliefs about their own subjective survival, while the computed objective survival probability increased. However, as motivated in Section 4.1, former smokers are of less interest. The third column in Table 5 shows that smokers in the age groups 5 to 10 (38-67 years) are significantly optimistic about their chances of survival, compared to smokers in age group 11 (68-73 years). This observed relative pessimism for smokers in later life is also supported by the pattern in Figure 6b. Consistent with the trend presented in Figure 7b, smokers in cohorts 1991-1994 (26.47) and 1995-1997 (17.11) are associated with optimistic chances of survival, relative to smokers in the 1956-1960 cohort. The year 2019 is associated with pessimism (-8.84) about smokers' chances of survival relative to 2015. In addition, males who smoke are relatively optimistic (3.44) about their survival probability compared to females who smoke, ceteris paribus.

Table 5.

Linear regression results on dependent variables subjective survival probability, subjective probability change and error

	Subjective survival probability	Subjective probability change	Error
	Coefficient (*SE*)	Coefficient (*SE*)	Coefficient (*SE*)

Former smoker	-0.07 (0.99)	0.32 (0.61)	-5.88*** (0.99)
Heavy smoker	-1.85 (1.32)	-0.46 (0.64)	1.90 (1.31)
Male	-2.82 (1.95)	0.40 (0.33)	3.44* (1.95)
Lower educated	-4.13* (2.42)	0.47 (0.50)	-3.70 (2.44)
Medium educated	-3.91* (2.32)	0.32 (0.39)	-3.58 (2.34)
Western	-1.36 (3.05)	0.07 (0.68)	-1.38 (3.01)
Non-western	-1.38 (4.59)	1.32 (0.91)	-1.58 (4.66)
BMI	0.13 (0.12)	0.09* (0.05)	0.15 (0.12)
Alcohol	0.10 (0.31)	0.00 (0.10)	0.12 (0.31)
Age groups			
Age G1 (17-22)	6.31 (9.63)	-9.86 (8.00)	9.89 (9.60)
Age G2 (23–27)	0.86 (8.47)	-9.62 (7.54)	5.31 (8.46)
Age G3 (28–32)	0.40 (7.51)	-10.57 (7.03)	5.21 (7.48)
Age G4 (33–37)	-0.88 (6.82)	-11.55**(5.52)	4.80 (6.81)
Age G5 (38-42)	3.03 (5.25)	-3.19 (3.92)	9.18* (5.25)
Age G6 (43-47)	4.21 (4.48)	-2.02 (3.39)	10.32** (4.46)
Age G7 (48-52)	3.04 (3.79)	-1.56 (2.84)	9.10** (3.77)
Age G8 (53-57)	1.02 (2.95)	0.71 (2.31)	6.66** (2.91)
Age G9 (58-62)	0.99 (2.38)	1.14 (1.61)	5.68** (2.34)
Age G10 (63-67)	1.72 (1.36)	-0.41 (0.81)	4.64*** (1.31)
Cohort			
1949-1955	4.30 (3.26)	0.10 (1.16)	2.15 (3.34)
1961-1965	1.90 (3.14)	-1.57 (1.11)	3.32 (3.13)
1966-1970	0.14 (3.54)	0.66 (1.83)	2.58 (3.54)
1971-1975	-2.82 (4.61)	2.36 (2.30)	0.63 (4.61)
1976-1980	-1.35 (5.03)	4.65 (3.00)	2.36 (5.05)
1981-1985	-7.75 (6.92)	11.19** (4.95)	-3.28 (6.86)
1986-1990	2.44 (7.01)	11.28* (6.28)	7.94 (7.03)
1991-1994	19.84** (9.72)	9.04 (6.73)	26.47*** (9.96)
1995-1997	9.91 (8.83)	9.03 (7.18)	17.11* (8.82)
Period			

2016	1.00 (1.04)	-	0.50 (1.04)
2017	1.07 (1.20)	-1.10 (1.79)	-0.17 (1.20)
2018	1.83 (1.29)	-0.66 (1.46)	0.46 (1.30)
2019	-6.63*** (1.55)	-10.07*** (1.46)	-8.84*** (1.55)
2020	3.04* (1.67)	7.95*** (1.52)	1.45 (1.68)
2021	2.76 (1.95)	-2.16 (1.40)	1.43 (1.95)
2022	2.00 (1.88)	-2.60 (1.60)	-0.19 (1.89)
Constant	64.96*** (4.55)	-1.66 (2.15)	-20.46*** (4.60)
R^2 Within	0.08	0.10	0.13
R^2 Between	0.07	0.09	0.09
Observations	1816	1589	1816

Note. Standard errors are in parentheses. Stars indicate the significance level; *** p<0.01, **p<0.05, *p<0.1. Alcohol coefficient reflects the categorical (1-8) amount of alcohol in the last 12 months, with 1 indicating almost every day, 2 indicates five or six days a week, 3 indicates three or four days a week, 4 indicates one or two days a week, 5 indicates once or twice a month, 6 indicates once every two months, 7 indicates once or twice a year and 8 indicates not at all.

In addition, to visualize and facilitate the interpretation of the moderating age, period and cohort effects on the relationship between smokers and the survival probability error. Marginal effects of the regression in column three of Table 5 are plotted in Figures 8, 9 and 10. These marginal effects show how each category separately affects the error. Figure 8 shows that older smokers (-17.47 for smokers aged 68-73) are relatively pessimistic about their probability of survival compared to younger smokers (-6.14 for smokers aged 43-47). In addition, Figure 9 shows that smokers from recent cohorts (14.08 for smokers from cohort 91-94) are relatively optimistic compared to smokers from earlier cohorts (-15.66 for smokers from cohort 81-85). Except for 2019, which affected one's subjective survival probability regardless of smoking status, there is no evidence that periods and anti-smoking policies implemented during this period did affect smokers their survival probability errors.

Figure 8.

Marginal effects of age group on the error, with 95% confidence intervals

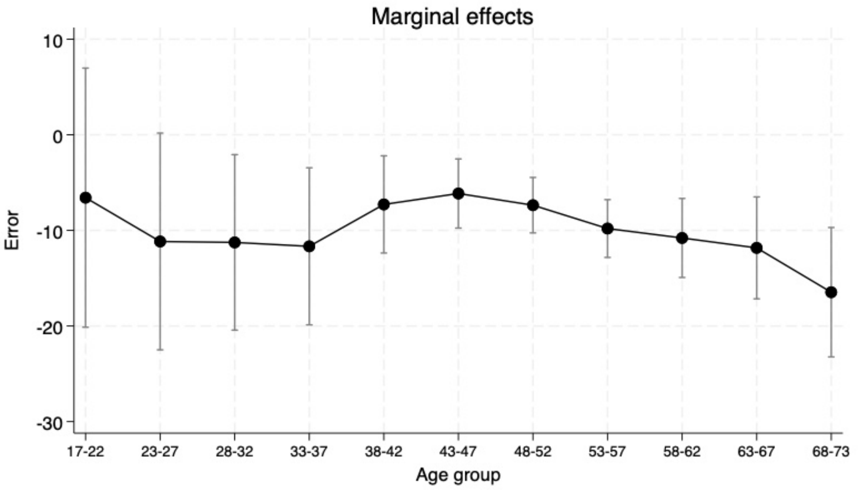


Figure 9.

Marginal effects of cohort on the error, with 95% confidence intervals

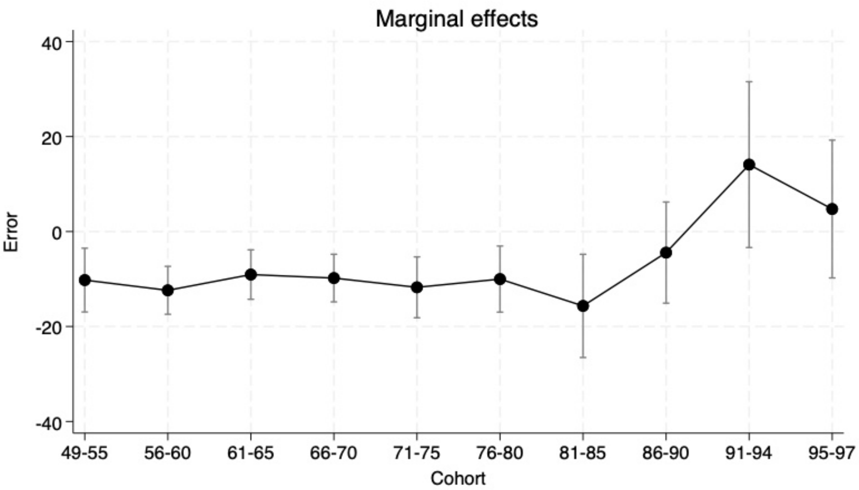
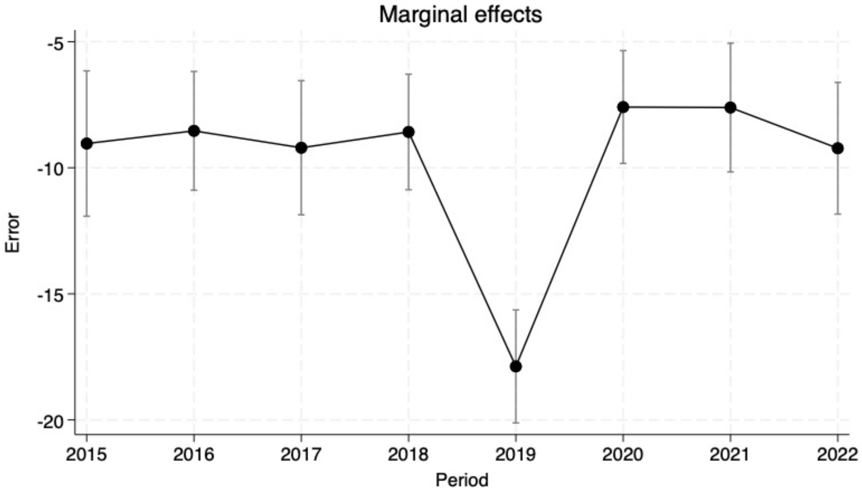


Figure 10.

Marginal effects of period on the error, with 95% confidence intervals



6. Discussion

While there may be alternative reasons for smoking behavior, such as addiction, time inconsistency, self-control problems or short time horizons. This study looked specifically at risk (mis)perception to explain smoking behavior. In the first part of the results section, several associations between smoking and longevity perceptions were presented. First of all, the magnitude of the negatively associated subjective probability of survival strongly suggests that smokers recognize health risks of smoking on longevity. This negative association is even stronger for heavy smokers. However, both smokers and heavy smokers do not seem to update their subjective probability of survival differently each year, compared to those who have never smoked. Since smoking-specific life tables were used to predict objective probabilities, the relatively optimistic survival probability errors, compared with never smokers, suggest that smokers do not take into account the full harm of smoking on longevity. This applies even more to those who are heavy smokers. The ability to accurately predict survival did not differ between the smoking groups. The latter suggests that the differences in error are due to relative optimism rather than smokers being better predictors of survival.

The second part of the results section demonstrated that the association between smoking and longevity perceptions differs for different age, period and cohort groups. In summary, smokers seem to be relatively optimistic about their chances of survival at particularly early ages, whereas this relationship is reversed for smokers in age group 11 (68-73 years). Most likely because smokers in this age group experience adverse effects that affect their subjective beliefs about longevity. For cohorts, the association between smokers and their survival probability errors remained largely unchanged. However, compared to the earlier cohorts, cohort groups 8, 9 and 10 (1986-1997) seem to be relatively optimistic about their probability of surviving to age 75. Whereas we would expect smokers from recent cohorts to be relatively pessimistic given the increase in information available and the anti-smoking policies implemented. However, potential explanations for these observed associations could be the not yet occurred adverse effects or the believe that they will quit smoking in the future, as outcomes of people from recent cohorts cannot be interpreted independently of those from people with a young age in the study period of only 8 years. In addition, regardless of smoking status, 2019 is associated with relatively pessimistic survival probability errors, compared to the other survey years. This is due to the lower subjective probabilities of survival being reported by all individuals (see Appendix A)

Consistent with previous literature, this study finds that smokers and particularly heavy smokers do perceive health risks of smoking on longevity (Viscusi and Kip,1990; Schoenbaum,1997; Kwaja et al.,2009). In contrast to Sloan et al. (2018), smokers do not seem to update their subjective probability of survival differently than those who never smoked. This may be due to the fact that Sloan et al. (2018) used a sample of young people (aged 12-18), for whom the initiation of smoking is likely to have a greater impact on subjective survival than for mature individuals. Conflicting with Kwaja et al. and Schoenbaum, who found accurate predictions on average, individuals are found to be highly pessimistic about their probability of survival to age 75 regardless of smoking status. This may be explained by the fact that these studies used a mature sample (aged 50-70), for whom it is easier to predict their probability of survival to age 75. However, in line previous research, smokers are found to be relatively optimistic about their chances of survival compared to those who have never smoked. This relative optimism is even stronger for those who are classified as heavy smokers (Schoenbaum, 1997; Kwaja et al., 2009).

Despite the consistency of this study's findings across the analyses performed, there are several things to bear in mind. The subjective survival probabilities are obtained from the LISS panel using a scale ranging from 0 to 10. As a result, the outcomes are coarse when converted to probabilities and could therefore not represent the true perceptions. In addition, the design of the question could cause a high frequency of focal (extreme values or the value that is in the middle) responses. The distribution of subjective survival probabilities in Figure 1 shows that focal responses are not over-represented. In addition, prior research using the similar scale to elicit subjective survival probabilities and future health beliefs concluded that these responses were able to reflect information about health events and behaviors that individuals have about themselves (Smith, Taylor, & Sloan, 2001; Hurd & McGarry, 2002). Moreover, Kwaja et al. (2009) concluded that individuals were quite accurate in reporting (0-10) longevity beliefs, which supports the LISS survey question being an accurate reflection of the perception of survival. Even if the design of this question meant that responses did not reflect true perceptions, relative differences between smoking groups could still be interpreted to answer the research questions in this study. Another thing to bear in mind regarding the subjective survival probabilities is that changes in subjective survival are likely to be underweighted because they are perceived from a starting point between approximately 0.5 and 0.9 (Tversky & Fox, 2000)

Although the objective survival probabilities in this study are much more detailed than used in prior research, all computed objective risk measures are subject to specification error. The use of individual and time-specific objective probabilities eliminated the bias of different period, cohort and age effects on the error of survival. Smoking-specific relative risks were multiplied by the averages in the Netherlands so that the objective survival probabilities would be an accurate counterpart to the subjective survival probabilities. Nevertheless, it is possible to imagine some unobserved scenarios in which this would not be the case, for example if a person has an underlying disease or decides to quit smoking because of occurred adverse effects. The latter motivates the focus on the comparison between smokers and those who have never smoked. In both scenarios, the objective survival probability would be overestimated. However, controlling for reported underlying diseases in the sensitivity analysis did not significantly affect the results and the answers to the research questions.

Finally, the relative optimistic survival probability errors should not necessarily be interpreted as smokers misperceiving the risks of smoking. Although this is a logical explanation given the large independent effect of smoking on survival and the magnitude of the observed association, the relative optimism may also be driven by smokers' optimism about the risk of death from all causes. Because, additionally to optimism about the risks of death from smoking, this optimism may also be driven by rosy thoughts about the future such as, medical technological improvements or the believe of quitting. The latter has been examined by Kwaja et al. (2009), who concluded that quitting expectations did not significantly affect survival expectations. Nevertheless, the alternative explanation that smokers are associated with rosy thoughts about premature mortality independent of their smoking behavior, awaits further research. Thus, smokers' relatively optimistic perceptions about survival compared to those who have never smoked suggest misperceptions about smoking risks. However, this should not be interpreted as evidence, as this study could not control for all potential confounders. It should be noted that only people who were surveyed in all the years were included in the sample. Therefore, this study does not examine people that died during this period. So, if optimistic survival errors cause smokers to continue smoking (which does increase the probability of dying during the years surveyed), this may have biased the relative optimism of smokers downwards.

Despite these limitations, this study is likely to have improved on previous research examining the associations between smoking and longevity beliefs. Firstly, by extending the scope of the existing studies with the examination of moderating age, period and cohort effects. Secondly, by improving the computation of objective survival probabilities as a counterpart to the reported subjective probabilities. Because the computed objective survival probabilities in this study are unique for each age, gender, survey year and smoking status, while previous studies have ignored period and cohort effects in the estimation of objective survival probabilities. Thus, this study includes and separately estimates period and cohort effects, which have been a source of bias in previous research.

7. Conclusion

Taken together, this study has looked at the association between (mis)perceptions of survival and smoking and the extent to which this association differs for different age period and cohort groups. Both male and female smokers seem to perceive adverse effects of smoking when reporting their subjective probability of survival to age 75. This is particularly the case for smokers over the age of 33. However, smokers and heavy smokers do not update their subjective probability of survival differently each year than those who have never smoked. Irrespective of smoking status, people are found to be highly pessimistic about their chances of survival to age 75. On the contrary, smokers and particularly heavy smokers are found to be relatively optimistic about their probability of survival to age 75, compared to those who have never smoked. The ability to accurately predict survival did not differ between the smoking groups. Examination of the APC effects showed that, compared to younger smokers, smokers aged 68-73 are found to be relatively pessimistic about their chances of survival to age 75. For cohorts, smokers from recent cohorts are associated with relatively optimistic survival probability errors, compared to smokers from earlier cohorts. In future studies, the time span of the study could be extended to be better able to distinguish between age and cohort effects. Over time, the survival probability errors among smokers shows a constant trend, except for 2019, which is associated with relatively pessimistic survival beliefs. However, the latter is the case regardless of smoking status.

Anti-smoking policies considering the APC-specific variations could focus on improving the risk perception of smoking for particularly younger smokers, which would discourage smoking behavior and reduce associated externalities. Whether the relative optimism about survival for particularly younger smokers is driven by misperceptions about the full harm or by rosy thoughts about the future awaits further research. In addition, to optimize the anti-smoking policies, research could examine the extent to which smoking behavior is determined by misperceptions of risks or by other internalities such as, time-inconsistent decision-making, self-control problems and short time horizons. If the smoking decision is made without considering the risks to longevity, smokers would not quit even when accurately perceiving the risks of smoking. In this case, anti-smoking campaigns should focus on other factors than the increased health risks, such as the addictiveness of smoking. Nevertheless, the magnitude of the association between smokers and optimistic (mis)perceptions of survival suggest that the latter is an important factor in the decision to smoke, which varies over a lifetime and over time.

Reference list

Breton, M. O., Gillespie, D., Pryce, R., Bogdanovica, I., Angus, C., Alava, M. H., Brennan, A., & Britton, J. (2021). Understanding long-term trends in smoking in England, 1972–2019: an age–period–cohort approach. *Addiction*, *117*(5), 1392–1403.

Carbone, J. C., Kverndokk, S., & Røgeberg, O. J. (2005). Smoking, health, risk, and perception. *Journal of Health Economics*, *24*(4), 631–653.

Centraal Bureau voor de Statistiek. (2022, October 18) *Meer hoogopgeleiden en beroepsniveau steeg mee.* Geraadpleegd van https://www.cbs.nl/nl-nl/nieuws/2022/42/meer-hoogopgeleiden-en-beroepsniveau-steeg-mee

Ehrlich, I., & Becker, G. S. (1972). Market Insurance, Self-Insurance, and Self-Protection. *Journal of Political Economy*, *80*(4), 623–648.

Fox, C. R., & Tversky, A. (2000). Ambiguity aversion and comparative ignorance. In *Cambridge University Press eBooks* (pp. 528–542). https://doi.org/10.1017/cbo9780511803475.031

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

Hsieh, C. (1998). Health risk and the decision to quit smoking. *Applied Economics*, *30*(6), 795–804.

Hurd, M. D., & McGarry, K. (2002). The predictive validity of subjective probabilities of survival. □*the* □*Economic Journal/Economic Journal*, *112*(482), 966–985.

Kahn, B. E., & Sarin, R. K. (1988). Modeling ambiguity in decisions under uncertainty. □*the* □*Journal of Consumer Research/Journal of Consumer Research*, *15*(2), 265.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263.

Khwaja, A., Sloan, F., & Wang, Y. (2009). Do smokers value their health and longevity less? □*the* □*Journal of Law & Economics/*□*the* □*Journal of Law & Economics*, *52*(1), 171–196.

Leening, M. J. G., Siregar, S., Vaartjes, I., Bots, M. L., Versteegh, M. I. M., Van Geuns, R. M., Koolen, J. J., & Deckers, J. W. (2013b). Heart disease in the Netherlands: a quantitative update. *Netherlands Heart Journal*, *22*(1), 3–10.

Liu, J.-., & Hsieh, C.-. (1995). Risk perception and smoking behavior: Empirical evidence from Taiwan. *Journal of Risk and Uncertainty*, *11*(2), 139–157.

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*(3), 261–286.

Marcon, A., Pesce, G., Calciano, L., Bellisario, V., Dharmage, S. C., Garcia-Aymerich, J., Gislasson, T., Heinrich, J., Holm, M., Janson, C., Jarvis, D., Leynaert, B., Matheson,

M. C., Pirina, P., Svanes, C., Villani, S., Zuberbier, T., Minelli, C., & Accordini, S. (2018). Trends in smoking initiation in Europe over 40 years: A retrospective cohort study. *PloS One*, *13*(8), e0201881.

Meza, R., Cao, P., Jeon, J., Fleischer, N. L., Holford, T. R., Levy, D. T., & Tam, J. (2023). Patterns of Birth Cohort–Specific Smoking Histories by Race and Ethnicity in the U.S. *American Journal of Preventive Medicine*, *64*(4), S11–S21.

Schoenbaum, M. (1997). Do smokers understand the mortality effects of smoking? Evidence from the Health and Retirement Survey. *American Journal of Public Health*, *87*(5), 755–759.

Sloan, F., & Platt, A. (2011). Information, risk perceptions, and smoking choices of youth. *Journal of Risk and Uncertainty*, *42*(2), 161–193.

Smith, V. K., Taylor, D. H., Sloan, F. A., Johnson, F. R., & Desvousges, W. H. (2001). Do smokers respond to health shocks? *the Review of Economics and Statistics*, *83*(4), 675–687.

Snow, A. (2010). Ambiguity aversion and the propensities for self-insurance and self-protection. *Journal of Risk and Uncertainty*, *42*(1), 27–43.

Swanson, D. A., Chow, S., & Bryan, T. (2020). Constructing Life Tables from the Kaiser Permanente Smoking Study and Applying the Results to the Population of the United States. In *Applied demography series* (pp. 115–151).

Jeon, J., Cao, P., Fleischer, N. L., Levy, D. T., Holford, T. R., Meza, R., & Tam, J. (2023). Birth Cohort–Specific smoking Patterns by family income in the U.S. *American Journal of Preventive Medicine*, *64*(4), S32–S41.

Viscusi, W. K. (1991). Age variations in risk perceptions and smoking decisions. *the Review of Economics and Statistics*, *73*(4), 577.

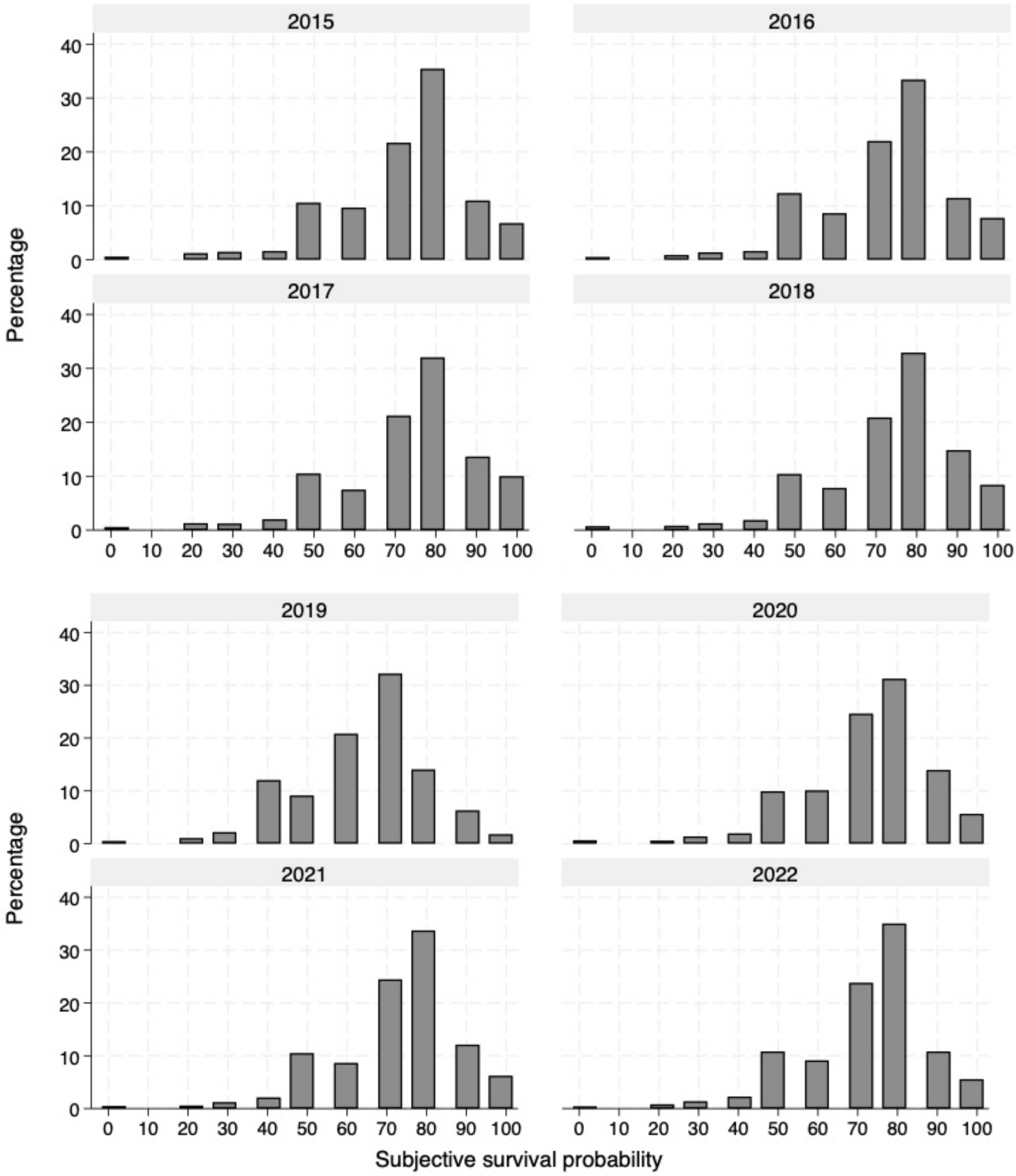
Viscusi, W. K. (1990). Do smokers underestimate risks? *Journal of Political Economy*, *98*(6), 1253–1269.

Weinstein, N. D. (1989). Effects of personal experience on self-protective behavior. *Psychological Bulletin*, *105*(1), 31–50.

Appendices

Appendix A

Distribution of the subjective probability of survival to age 75, by year.



Note: x-axis is the subjective probability of surviving to age 75 (percentage).

Appendix B

Example (Male, 2020) of how the CBS life tables by year and gender have been transformed to tables showing the probability of survival to age 75.

17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
0,99984																				
0,99958	0,99974																			
0,99926	0,99942	0,99968																		
0,9989	0,99906	0,99932	0,99964																	
0,998551	0,998711	0,99897	0,99929	0,99965																
0,998221	0,998381	0,998641	0,99896	0,99932	0,99967															
0,997902	0,998062	0,998321	0,998641	0,999	0,99935	0,99968														
0,997473	0,997632	0,997892	0,998211	0,998571	0,99892	0,99925	0,99957													
0,997134	0,997293	0,997553	0,997872	0,998231	0,998581	0,99891	0,99923	0,99966												
0,996825	0,996984	0,997243	0,997563	0,997922	0,998271	0,998601	0,99892	0,99935	0,99969											
0,996356	0,996515	0,996775	0,997094	0,997453	0,997802	0,998131	0,998451	0,99888	0,99922	0,99953										
0,995928	0,996087	0,996346	0,996665	0,997024	0,997373	0,997702	0,998022	0,998451	0,99879	0,9991	0,99957									
0,995519	0,995679	0,995938	0,996256	0,996615	0,996964	0,997293	0,997612	0,998042	0,998381	0,998691	0,99916	0,99959								
0,994902	0,995061	0,99532	0,995639	0,995997	0,996346	0,996675	0,996994	0,997423	0,997762	0,998071	0,998541	0,99897	0,99938							
0,994464	0,994623	0,994882	0,995201	0,995559	0,995908	0,996236	0,996555	0,996984	0,997323	0,997632	0,998101	0,998531	0,99894	0,99956						
0,994027	0,994186	0,994444	0,994763	0,995121	0,995469	0,995798	0,996117	0,996545	0,996884	0,997193	0,997662	0,998091	0,998501	0,99912	0,99956					
0,99343	0,993589	0,993848	0,994166	0,994524	0,994872	0,9952	0,995519	0,995947	0,996286	0,996595	0,997064	0,997492	0,997902	0,998521	0,99896	0,9994				
0,992874	0,993033	0,993291	0,993609	0,993967	0,994315	0,994643	0,994962	0,99539	0,995728	0,996037	0,996505	0,996934	0,997343	0,997962	0,998401	0,99884	0,99944			
0,992139	0,992298	0,992556	0,992874	0,993231	0,993579	0,993907	0,994225	0,994653	0,994991	0,9953	0,995768	0,996196	0,996605	0,997223	0,997662	0,998101	0,9987	0,99926		
0,991375	0,991534	0,991792	0,992109	0,992467	0,992814	0,993142	0,99346	0,993887	0,994225	0,994533	0,995001	0,995429	0,995837	0,996455	0,996894	0,997333	0,997931	0,998491	0,99923	
0,990751	0,990909	0,991167	0,991484	0,991841	0,992189	0,992516	0,992834	0,993261	0,993599	0,993907	0,994374	0,994802	0,99521	0,995827	0,996266	0,996704	0,997303	0,997862	0,9986	0,99937
0,98979	0,989948	0,990206	0,990523	0,990879	0,991226	0,991553	0,991871	0,992297	0,992635	0,992943	0,99341	0,993837	0,994245	0,994861	0,995299	0,995738	0,996335	0,996894	0,997632	0,998401
0,988889	0,989047	0,989304	0,989621	0,989978	0,990324	0,990651	0,990968	0,991394	0,991732	0,992039	0,992506	0,992933	0,99334	0,993956	0,994394	0,994831	0,995429	0,995986	0,996724	0,997492
0,988068	0,988226	0,988483	0,9888	0,989156	0,989502	0,989829	0,990146	0,990572	0,990909	0,991216	0,991682	0,992108	0,992515	0,993131	0,993568	0,994006	0,994602	0,99516	0,995897	0,996664
0,987051	0,987208	0,987465	0,987781	0,988137	0,988483	0,988809	0,989126	0,989551	0,989888	0,990195	0,99066	0,991087	0,991493	0,992108	0,992545	0,992982	0,993578	0,994135	0,994871	0,995638
0,985797	0,985955	0,986211	0,986527	0,986882	0,987228	0,987554	0,98787	0,988295	0,988631	0,988937	0,989402	0,989828	0,990234	0,990848	0,991284	0,991721	0,992316	0,992872	0,993607	0,994373
0,984653	0,984811	0,985067	0,985382	0,985737	0,986082	0,986408	0,986724	0,987148	0,987484	0,98779	0,988255	0,98868	0,989085	0,989699	0,990135	0,99057	0,991165	0,99172	0,992455	0,99322
0,983482	0,983639	0,983895	0,98421	0,984564	0,984909	0,985234	0,985549	0,985973	0,986309	0,986615	0,987079	0,987503	0,987908	0,988521	0,988956	0,989392	0,989986	0,99054	0,991274	0,992038
0,982183	0,982341	0,982596	0,982911	0,983265	0,983609	0,983934	0,984249	0,984672	0,985007	0,985312	0,985776	0,9862	0,986604	0,987216	0,987651	0,988086	0,988679	0,989233	0,989965	0,990728
0,980425	0,980582	0,980837	0,981151	0,981505	0,981848	0,982172	0,982487	0,982909	0,983244	0,983549	0,984011	0,984434	0,984838	0,985449	0,985883	0,986317	0,986909	0,987462	0,988193	0,988955
0,978759	0,978915	0,97917	0,979483	0,979836	0,980179	0,980503	0,980817	0,981238	0,981572	0,981877	0,982338	0,982761	0,983164	0,983774	0,984207	0,98464	0,985231	0,985783	0,986513	0,987274
0,976713	0,976869	0,977123	0,977436	0,977788	0,978131	0,978453	0,978767	0,979188	0,979521	0,979824	0,980285	0,980707	0,981109	0,981718	0,98215	0,982582	0,983172	0,983723	0,984452	0,98521
0,974545	0,974701	0,974954	0,975266	0,975618	0,975959	0,976281	0,976594	0,977014	0,977346	0,977649	0,978109	0,97853	0,978931	0,979538	0,97997	0,980401	0,98099	0,981539	0,982266	0,983023
0,972303	0,972459	0,972712	0,973023	0,973374	0,973714	0,974036	0,974348	0,974767	0,975098	0,975401	0,975859	0,976279	0,97668	0,977285	0,977716	0,978146	0,978733	0,979282	0,980007	0,980762
0,969513	0,969668	0,96992	0,970231	0,97058	0,97092	0,97124	0,971551	0,971969	0,9723	0,972601	0,973059	0,973477	0,973876	0,974481	0,97491	0,975339	0,975924	0,976471	0,977194	0,977947
0,966401	0,966555	0,966807	0,967116	0,967464	0,967803	0,968123	0,968433	0,968849	0,969179	0,969479	0,969935	0,970352	0,97075	0,971353	0,97178	0,972208	0,972792	0,973337	0,974057	0,974808
0,963028	0,963182	0,963433	0,963741	0,964088	0,964426	0,964744	0,965053	0,965468	0,965796	0,966096	0,96655	0,966966	0,967362	0,967963	0,968389	0,968815	0,969397	0,96994	0,970658	0,971406
0,959686	0,95984	0,960089	0,960397	0,960743	0,961079	0,961396	0,961704	0,962118	0,962445	0,962743	0,963196	0,96361	0,964006	0,964604	0,965028	0,965453	0,966033	0,966574	0,96729	0,968035
0,956203	0,956356	0,956604	0,95691	0,957255	0,95759	0,957906	0,958213	0,958625	0,958951	0,959249	0,9597	0,960113	0,960506	0,961102	0,961525	0,961949	0,962526	0,963065	0,963779	0,964521
0,952617	0,952769	0,953017	0,953322	0,953665	0,953999	0,954314	0,95462	0,95503	0,955355	0,955651	0,956101	0,956512	0,956904	0,957498	0,95792	0,958341	0,958917	0,959454	0,960164	0,960904
0,948492	0,948644	0,94889	0,949194	0,949536	0,949868	0,950182	0,950486	0,950895	0,951218	0,951513	0,951961	0,95237	0,952761	0,953352	0,953772	0,954192	0,954764	0,955299	0,956007	0,956744
0,943873	0,944024	0,944269	0,944572	0,944912	0,945243	0,945555	0,945857	0,946264	0,946586	0,94688	0,947325	0,947732	0,948121	0,948709	0,949127	0,949545	0,950115	0,950647	0,951351	0,952084
0,93854	0,93869	0,938934	0,939235	0,939573	0,939902	0,940212	0,940513	0,940918	0,941238	0,94153	0,941972	0,942378	0,942764	0,943349	0,943764	0,94418	0,944747	0,945276	0,945976	0,946705
0,93227	0,93242	0,932662	0,932961	0,933297	0,933623	0,933932	0,934231	0,934633	0,93495	0,93524	0,93568	0,936083	0,936467	0,937048	0,93746	0,937873	0,938436	0,938962	0,939657	0,940381
0,925428	0,925576	0,925816	0,926113	0,926446	0,926771	0,927077	0,927373	0,927772	0,928088	0,928376	0,928812	0,929212	0,929593	0,93017	0,930579	0,930989	0,931548	0,93207	0,93276	0,933479
0,918237	0,918384	0,918623	0,918917	0,919248	0,91957	0,919873	0,920168	0,920564	0,920877	0,921162	0,921595	0,921992	0,92237	0,922942	0,923348	0,923755	0,924309	0,924827	0,925512	0,926225
0,910359	0,910504	0,910741	0,911033	0,911361	0,91168	0,911981	0,912273	0,912665	0,912975	0,913259	0,913688	0,914081	0,914456	0,915023	0,915426	0,915829	0,916379	0,916892	0,917571	0,918278
0,900791	0,900935	0,901169	0,901458	0,901782	0,902098	0,902396	0,902685	0,903073	0,90338	0,90366	0,904085	0,904474	0,904845	0,905406	0,905805	0,906204	0,906748	0,907256	0,907928	0,908627
0,891026	0,891169	0,8914	0,891686	0,892007	0,892319	0,892614	0,8929	0,893284	0,893587	0,893865	0,894285	0,89467	0,895037	0,895592	0,895986	0,89638	0,896919	0,897421	0,898086	0,898778
0,88004	0,880181	0,88041	0,880691	0,881008	0,881317	0,881608	0,88189	0,882269	0,88257	0,882843	0,883258	0,883638	0,884001	0,884549	0,884939	0,885328	0,88586	0,886356	0,887012	0,887696
0,868212	0,868351	0,868577	0,868855	0,869168	0,869472	0,869759	0,870037	0,870412	0,870708	0,870978	0,871387	0,871762	0,87212	0,872661	0,873045	0,873429	0,873954	0,874443	0,875091	0,875765
0,855423	0,85556	0,855783	0,856057	0,856365	0,856665	0,856948	0,857222	0,857591	0,857882	0,858148	0,858552	0,858921	0,859273	0,859807	0,860185	0,860564	0,86108	0,861563	0,862201	0,862865
0,84148	0,841615	0,841833	0,842103	0,842406	0,842701	0,842979	0,843249	0,843612	0,843899	0,84416	0,844557	0,844921	0,845267	0,845792	0,846164	0,846536	0,847045	0,847519	0,848147	0,848801
0,826813	0,826945	0,82716	0,827425	0,827723	0,828013	0,828286	0,828551	0,828908	0,82919	0,829447	0,829837	0,830194	0,830534	0,83105	0,831415	0,831781	0,832281	0,832747	0,833364	0,834006
0,809987	0,810117	0,810328	0,810587	0,810879	0,811163	0,811431	0,81169	0,812039	0,812316	0,812568	0,81295	0,813299	0,813633	0,814138	0,814496	0,814855	0,815344	0,815801	0,816405	0,817034
0,792095	0,792221	0,792427	0,792681	0,792967	0,793244	0,793506	0,79376	0,794101	0,794372	0,794618	0,794992	0,795334	0,79566	0,796153	0,796504	0,796854	0,797333	0,79778	0,79837	0,798986
0,772672	0,772796	0,772997	0,773245	0,773523	0,773794	0,774049	0,774297	0,77463	0,774894	0,775134	0,775498	0,775832	0,77615	0,776632	0,776974	0,777316	0,777782	0,778218	0,778794	0,779395
0,752359	**0,752479**	**0,752675**	**0,752916**	**0,753187**	**0,753451**	**0,7537**	**0,753941**	**0,754265**	**0,754522**	**0,754756**	**0,755111**	**0,755435**	**0,755745**	**0,756214**	**0,756547**	**0,75688**	**0,757334**	**0,757759**	**0,75832**	**0,758904**

Appendix C

Example of how the smoking-specific life tables by Swanson et al. (2020) are transformed into tables showing the probability of survival to age 75.

Male	Never	Former (2-1(	Former (11-2	Former 20+	Former aver.	Current<20	Current 20-3	Current 40+	Current ave	Heavy	Weighted Average
20y	0,7701644	0,703677	0,7258601	0,7519797	0,7271723	0,7384103	0,6282764	0,5686202	0,6451023	0,6102477	0,7263529
25y	0,7703801	0,7039938	0,7260489	0,7522505	0,7274311	0,7391051	0,6287228	0,5690242	0,6456173	0,6109258	0,7266525
30y	0,7709814	0,7048396	0,726572	0,7529206	0,7281107	0,7406456	0,6298691	0,5700617	0,6468588	0,6125306	0,7274285
35y	0,7723562	0,7067266	0,7277947	0,7543539	0,7296251	0,7436201	0,6323543	0,5723109	0,6494284	0,6157881	0,7291269
40y	0,7751312	0,7104422	0,7303436	0,7571402	0,732642	0,7488093	0,6371328	0,5766357	0,6541926	0,6217507	0,7324441
45y	0,7802496	0,7171693	0,7352181	0,7621858	0,7381911	0,757268	0,6456227	0,5843195	0,6624034	0,6318862	0,7384245
50y	0,7890874	0,7286084	0,7439747	0,7708814	0,7478215	0,7703799	0,6598558	0,5972011	0,6758123	0,6482546	0,7485911
55y	0,8036004	0,7471987	0,7590029	0,7853155	0,763839	0,7899712	0,6827837	0,6262333	0,6996627	0,6737843	0,7657783
60y	0,8266219	0,776504	0,7839886	0,8086699	0,7897208	0,8185042	0,7188105	0,6703992	0,7359046	0,712773	0,7928415
65y	0,862268	0,8218713	0,8247737	0,8459155	0,8308535	0,8593852	0,7747806	0,7373425	0,7905028	0,7718003	0,8347669
70y	0,91677	0,89171	0,89105	0,9051	0,8959533	0,91748	0,86194	0,83975	0,8730567	0,86149	0,8994301

Female	Never	Former (2-1(	Former (11-2	Former 20+	Former aver.	Current<20	Current 20-3	Current 40+	Current ave	Heavy	Weighted Average
20y	0,8576591	0,8343004	0,8382803	0,842457	0,8383459	0,836432	0,7690201	0,6950405	0,7668309	0,7125164	0,8351298
25y	0,8577535	0,8343087	0,8383306	0,8425834	0,8384076	0,8365073	0,7693201	0,6953117	0,7670464	0,7127872	0,8352355
30y	0,858028	0,8343337	0,8385151	0,8429458	0,8385982	0,8367416	0,7700902	0,6960077	0,7676132	0,7135222	0,8355358
35y	0,8587064	0,8344422	0,8390185	0,8437812	0,8390806	0,8373613	0,771765	0,6975213	0,7688825	0,7151957	0,8362579
40y	0,8601687	0,8347928	0,8402032	0,845506	0,8401673	0,8387536	0,7750123	0,7004562	0,7714074	0,7185586	0,8377912
45y	0,863008	0,8357707	0,8426891	0,8487567	0,8424055	0,8415729	0,780798	0,7056854	0,7760187	0,7247552	0,8407571
50y	0,8681212	0,8381594	0,8474689	0,854499	0,8467091	0,8468317	0,7904654	0,7144228	0,7839066	0,7354561	0,8461107
55y	0,8768193	0,8434903	0,8560983	0,8641428	0,8545771	0,8560773	0,8059065	0,7363817	0,7994552	0,7530781	0,8557763
60y	0,8909769	0,8546174	0,870948	0,8797317	0,8684324	0,8716005	0,8297622	0,7694688	0,8236105	0,781169	0,8716367
65y	0,9132885	0,8766835	0,895604	0,9042365	0,8921747	0,8968098	0,8658238	0,8186361	0,8604232	0,8250972	0,8970163
70y	0,94767	0,9191	0,93564	0,94207	0,93227	0,93689	0,91969	0,89144	0,9160067	0,89332	0,9370426

Example of how the relative risk of surviving to age 75 is computed by dividing the smoking-specific probabilities by the weighted average.

RR never	RR former	RR current	RR heavy
1,060317117	1,0011281	0,888139	0,8401532
1,060176743	1,0010715	0,8884816	0,84074
1,059872477	1,0009379	0,8892404	0,8420493
1,059289141	1,0006832	0,8906932	0,8445555
1,058280362	1,0002702	0,8931639	0,8488713
1,056641051	0,9996839	0,8970496	0,8557221
1,054096779	0,998972	0,902779	0,8659663
1,049390502	0,9974676	0,9136623	0,8798687
1,042606661	0,9960639	0,9281863	0,8990107
1,03294464	0,9953119	0,9469742	0,9245698
1,019278763	0,9961345	0,9706776	0,9578176

RR never	RR former	RR current	RR heavy
1,026977015	1,003851	0,9182176	0,8531804
1,026960012	1,0037978	0,9183594	0,8533967
1,026919505	1,0036652	0,9187077	0,8539696
1,026844051	1,0033755	0,9194323	0,8552335
1,026710081	1,0028362	0,9207632	0,8576821
1,026465262	1,0019606	0,9229999	0,8620269
1,026013753	1,0007072	0,9264823	0,8692197
1,024589394	0,9985988	0,9341871	0,8799942
1,022188375	0,9963238	0,9449011	0,8962094
1,018140355	0,9946025	0,9592058	0,919824
1,011341427	0,9949067	0,9775507	0,9533398