Working Papers in Health Policy and Management

Vol. 2 / 09

Leonie Sundmacher, Andrew Jones and Nigel Rice

The Role of Health Shocks in Quitting Smoking

Evidence from the European Community Household Panel

Department Health Care Management / Reinhard Busse

Technische Universität Berlin

Sundmacher, Leonie

The Role of Health Shocks in Quitting Smoking

Working Papers in Health Policy and Management

Vol. 2/09

August 2009

HERAUSGEBER DER SCHRIFTENREIHE
FG Management im Gesundheitswesen
Prof. Dr. med. Reinhard Busse

REDAKTION
Miriam Blümel
Matthew Gaskins

VERLAG UND VERTRIEB
Universitätsverlag der Technischen Universität Berlin
Universitätsbibliothek
Fasanenstr. 88 (im VOLKSWAGEN-Haus), D-10623 Berlin
Tel.: (030)314-76131; Fax.: (030)314-76133
E-Mail: publikationen@ub.tu-berlin.de
http://www.univerlag.tu-berlin.de/

ISBN 978-3-7983-2182-3

ISSN 1869-6287

Inhaltsverzeichnis

Abstract

The European Union has stated interest in assessing the effectiveness and relevance of its messages about the adverse consequences of smoking in the context of its tobacco control policy. In the absence of disaggregated data on the direct relationship between health information and smoking decisions, we follow Clark et al. (2002) and investigate the impact of health shocks on the probability to quit daily smoking using eight waves of the European Union Community Household Panel (ECHP). Our intention is to assess whether individuals learn from changes in health i.e. successfully update new information about the consequences of tobacco consumption. As self assessed health is subjective and prone to reporting bias, we instrument self assessed health using “objective" health indicators and the socio-demographic variable age; the resulting variable is then used to model continuous and discrete changes in health, termed as health shocks. Estimating a discrete time hazard model with gamma distributed frailty, we find evidence that objective discrete health shocks increase the probability to quit daily smoking. Stratifying by gender reveals that in particular men above 55 quit following a negative health shock while the results for women are not statistically significant. Assuming that the increased hazard rate for men is associated with an increased perceived risk of coronary artery disease, we conclude that specific information about smoking related health shocks are the most effective health warnings.

Introduction

Tobacco is the single largest cause of avoidable death within the European Union accounting for over half a million deaths each year and over a million deaths in Europe as a whole. It is estimated that 25% of all cancer deaths and 15% of all deaths can be attributed to smoking.

In order to curb the epidemic, the European Union exhibits a tobacco control policy that supports Europe wide smoking prevention and cessation activities. Among other measures this involves raising the public awareness of the harmful effects of tobacco consumption in particular by means of information. In this context, the European Union has stated its interest to improve the effectiveness and relevance of the messages put across about the adverse consequences of smoking.

This dissertation seeks to contribute to the stated objective of the European Union, namely, to assess the effectiveness of health messages about the harmfulness of smoking. Ideally, we would like to investigate the relationship between health information policy and quitting behaviour. In the absence of good disaggregated data on such policies, Clark et al suggest to consider the impact of health changes on smoking decision in order to assess whether individuals successfully update new information about adverse health consequences.

If smokers adopt new information and accordingly change their smoking habits, we can conclude that smoker's link adverse health experience to their lifestyle, and may either be sensitized by health information or potentially responsive. Following the suggestions of Clark et al, we investigate the impact of health shocks on the decision to quit daily smoking using the first eight waves (1994-2001) of the European Union Community Household Panel (ECHP).

We specify four competing health shock measures; three of them are based on information on self assessed health (abbreviated SAH) provided by the ECHP. To tackle the problem of potential measurement errors in self reported health, we adopt a two step procedure first proposed by Bound (1991) and estimate SAH as function of (arguably) objective health measures; the resulting variable is then used to model continuous and discrete changes in health, termed as health shocks. Furthermore, we disentangle different types of health shocks and construct respective dummy variables in order to investigate their impact on the probability to quit daily smoking.

The competing model specifications are then estimated using a discrete time hazard model with gamma distributed unobserved heterogeneity. The model describes the probability of quitting as a function of covariates, health shocks and time spent in the initial state of daily smoking so that the time dynamics of smoking cessation are captured through duration dependence.

The dissertation is organized as follows. Section one presents a literature review of theoretical models that relate health shocks to smoking decision and previous econometric studies. Section two reflects on the appropriate measure to proxy health shocks. Section three presents our econometric specification for analysing the impact of health shocks in the decision to quit daily smoking. Section four discusses the data and section five follows with the presentation of results. The last section concludes with a discussion of the results.

1. Theory

The natural starting point of the literature review is Becker and Murphy's (1988) model of rational addiction. The model allows rational forward looking individuals who maximize their lifetime utility subject to a budget constraint to develop an addiction[1]. Utility depends at any time on a stock of past addictive consumption (defined as increasing utility of present consumption) and the individuals are rational about their addiction in the sense that they consider the future implications of their current consumption. An addicted smoker would quit if the adjustment costs are lower than the long term benefits of continued smoking. With respect to our question whether health shocks impact on smoking cessation, high perceived

[1] The model is based on Stigler and Becker (1977) model on addiction.

individual health costs should encourage abrupt quitting. On the other hand, positive health development increases overall utility and decreases the relative net benefit of the addictive consumption.

Recently, Clark et al developed the rational addiction model by adding variables on health changes. Rather than perceived costs of past health developments, the extension was intended to capture learning effects from health shocks. The underlying theoretical motivation has been provided by Etile (2000) who directly specifies a causal relationship between changes in health and smoking habits. His model differs from the rational addiction model such that it allows individuals to learn about the harmful consequences of tobacco consumption over a long time period while smoking. Adverse health developments provide smokers with new information about the (true) parameters of their health production function and enable the individual to revise their perceived smoking risk. Ceteris paribus, the updated knowledge about the harmful consequences of tobacco consumption changes on individual's belief and incentivises a reduction in cigarette consumption or to quit smoking in order to promote own health.

Etile's model specification is inspired by the Grossman model on the demand for health (Grossman, 2000) that offers fundamental insight into behavioural response to health shocks. In contrast to Etile's model, the Grossman model assumes that all relevant parameters of health production and the utility of health are known by the individual. Smoking enters the utility function as a consumption good and is an input factor into the health production function that reduces expected health at the end of period t. In general, the amount of healthy days produced in period t is the result of the initial health stock plus gross investment net or plus all changes in health in period t. Thus, a reduction in smoking in period t is expected to increase the health stock and could therefore compensate for adverse health developments. Ceteris paribus, the perceived loss in utility caused by a decrease in health needs to exceed the utility of cigarette consumption at the margin to incentivise the individual to reduce cigarette consumption or quit.

Furthermore, the Grossman model sheds light on the interaction between health information and education and health developments crossed with income or age; education increases the efficiency in the production of health at a given level of inputs. Better educated individuals are in a better position to process health information about adverse consequences of smoking and are therefore more receptive to changes in their smoking habits in order to compensate for a loss in health. Furthermore, an individual calculates the optimal health stock by equating the marginal product of health capital with the cost of gross investment. The marginal product depends on the benefit of additional healthy time when spent on both consumption and generating income. This implies that low income individuals face relatively less incentive to invest in health and compensate for a comparable absolute loss in health stock. Regarding the interaction of health changes and age, Clark et al. point out that the Grossman model predicts a decrease in daily smoking as the individual ages. This is because the initial health stock depreciates with age but may be augmented through investment such as quitting or cutting down tobacco consumption.

With respect to our initial interest in the relationship between health information and quitting decisions, we would like to know whether individuals successfully update new health information and accordingly adjust their smoking habits. Etile's learning model provides the theoretical basis for our empirical investigation. It predicts a positive correlation between negative health developments and the probability to quit smoking. However, the same holds true for the standard rational addiction model and the Grossman model; even without learning a rational addicted smoker would be (but less) prone to confine his consumption or quit after a negative shock (Clark et. al). Thus, a positive correlation of health shocks and smoking cessation is necessary to reflect a learning process but not sufficient to conclude one. However, the models slightly differ with respect to positive health developments: in the rational addiction models, perceived positive health changes may curb the addictive consumption. The effect in the Grossman model is ambiguous: positive objective health shocks increase the flow of healthy time available for consumption, while increased cigarette consumption decreases the health stock. Cigarette consumption may therefore stay stable.

Previous econometric studies

Clark et al apply their extended addiction model to seven waves of the British Household Panel Survey and find negative health shocks to be associated with reduced cigarette consumption and a higher probability to quit smoking. Whilst most variables have the expected sign, some misbehave. In the regression analysis with cigarette consumption as the dependent variable, young women and old women increase cigarette consumption as they suffer from decreases in health. Clark et al explain these counterintuitve results assuming the health variables do not reflect real developments in health. With respect to the probit quitting smoking equation, males between the ages of 15 and 25 and with a lung check up at time period t1 had a decreased probability of quitting smoking as did males between 55 and 65 years with heart problems at t1.

Smith et al. (2001) use panel data from the health and retirement study (HRS) and evaluate how participants between age 51 to 61 change their longevity expectations in response to an exogenous health shock. They find that smokers, non-smokers and ex smokers have different rules to revise their assessment of longevity such that smokers update their expectations more dramatically downward after a shock. Furthermore, smokers revise their risk following a severe smoking-related health shock (heart attacks, congestive failure, stroke, smoking-related cancer, severe lung diseases) considerably while there is almost no effect after a general health shock. This implies that specific information about smoking related health events is most likely to update smoker's belief.

Hsieh (1998) explores quitting rates among elderly Taiwanese smokers and finds that hazard rates increase with increased perceived health risk. The results are reported to be robust to changes in three different health measure specifications (self reported health, number of chronic diseases, number of physical limitations) provided by the two period national survey.

2. Construction of health shock measures

Another theoretical issue concerns the choice of appropriate health measures used to construct health shocks. In Etile's model, individuals revise their perceived smoking risk as they suffer from non-anticipated objective health shocks. Thus, the smoking decision depends (among other things) on objective health shocks and the initial smoking risk parameter; if the model holds true, smokers learn from a shock, revise their risk and the probability to quit daily smoking increases. The leading question of this section is which health measure best captures changes in objective health in order to explain the dynamics of smoking cessation as proposed by the model.

Our first intuition is that instructional objective health changes (this involves revision of the belief about the risk of smoking) should be reflected in a (temporary) change in reported SAH; learning demands the processing of information and enforces revision of latent self assessed health. Given the change in latent health is large enough, as we would expect from an instructional health shock[2], it further leads to a change in reported SAH. From this perspective, SAH would be a good proxy.

However, an instructional objective health shock might necessarily be reflected in a change in perceived health but changes in self assessed health are not sufficient to conclude comparable objective shocks across individuals. There are two main explanations to this statement.

First, the subjective assessment of a given "true" latent health stock might differ by subgroups. Lindeboom et al (2006) find - inter alia - evidence for changes in cut points set to map latent health to one of the categories of SAH for age and gender. With respect to age, elderly people are more likely to report their health positive given the same objective health stock; thus, reported objectives health changes at young age might not matter at older age and the resulting reporting heterogeneity makes comparison of SAH across age groups difficult.

Second, unobserved heterogeneity between individuals might generate a bias if respondents who quit smoking differ systematically from deathless smokers with respect to their self perception of health. It is intuitive that a fragile self perception of health may be related to both: volatile self reported health and a disposition to quit smoking and therefore lead us to overestimate the impact of health shocks on cessation decisions.

A strategy to approach these problems and explore the "true" link between health and smoking decisions is to use "objective" indicators of ill health in the empirical specification. Bound et al. (1999) argues that these proxy variables suffer from similar problems like self assessed health and are prone to measurement error. Accordingly, we argue that the "objective" indicators in the ECHP only provide a glimpse on the full spectrum of individual health and are not adequate to explore the impact of health variations i.e. health shocks. The number of hospital nights and GP visits in the last year, observations on mental illness, and the variable chronically ill are characterised by excess zero observations leading to left censored distributions and do not provide non-zero observations on changes in health for a vast majority of individuals in the sample. Thus, it would be sensible to use the information provided by the five category self assessed health variable ranging over very good, good,

[2] We could argue that only changes on the latent self assessed health scale that cross a cut point can be termed health shocks.

fair, bad and very bad health on the true variations in health that are intrinsic to this "subjective measure".

To tackle the problem, we adopt an approach first suggested by Bound (1991) and later used by Bound et al (1999), Disney et al. (2006) and Jones et al. (Chapter seven, forthcoming). We construct a latent health stock as a function of personal characteristics and health indicators using "subjective" reported health. The constructed variable is then used to explore the relationship between variations in health over time (termed as health shocks) and quitting daily smoking. Referring to Disney et al., we argue that constructing the latent health stock standardizes self assessed health on arguably objective health indicators while preserving the ordered categorisation of self assessed health and its reflection on true health.

Adopting notation and approach of Jones et al. we assume that health considerations that lead an individual to quit daily smoking, h_{it}^{Q}, are a function of objective health indicators and age, denoted by Z_{it}^{Q}:

(1) $$h_{it}^{Q} = Z_{it}^{Q} + \varepsilon_{it} \qquad \text{i=1,2,...,n; t=1,2,...} \; T_i$$

where ε_{it} is a time varying error term exogenous to Z_{it}. The latent health status h_{it}^{Q} is not observed. Instead, individuals introspect and map their assessed latent health h_{it}^{Q} to one of the five categories of self reported health, providing a measure of SAH, h_{it}^{S}. As discussed above, h_{it}^{S} is subject to reporting heterogeneity and therefore potential measurement errors so that in fact, it is observed as a function of its latent counterpart denoted by h_{it}^{*}:

(2) $$h_{it}^{*} = h_{it}^{Q} + \eta_{it}$$

where η_{it} represents measurement errors. η_{it} may be correlated to the propensity to quit smoking and represent person-specific errors brought about by demographic different use of the self assessed health scale. Substituting (1) in (2) yields

(3) $h_{it}^{*} = Z_{it}`\beta + \varepsilon + \eta = Z_{it}`\beta + \nu_{it}$

Latent self assessed health is a function of age, health indicators and a composed error term. We assume a normal distribution for ν_{it} and use an ordered probit regression to estimate the betas with the intention to predict $\hat{h}_{it}^{*}$ and sweep the measurement errors out of the SAH variable. Predicted $\hat{h}_{it}^{*}$ is then used to estimate the impact of health shocks on the probability to quit smoking.

We include the socio-demographic variable age for two related reasons in the probit equation. First, it should predict “true” health that is intrinsic to the self assessed health measure and second, serve as predictor for latent health if all indicator variables take a zero value for the respondent. Thus, if an individual has not been to the hospital in the year of the interview, is neither chronically nor mentally ill, age serves as predictor for its latent health status. We are conscious that including the age variable may perpetuate the associated person-specific reporting; with respect to health shocks, we would then expect less reported shocks at old age compared to young age given a comparable objective change in health. However, cut point shifts at increased age reflect anticipated worsening of health. In Etile’s model smokers learn from non-anticipated health shocks. Therefore, we hope not to underestimate the influence of non-anticipated health shocks at higher age.

Exploration in health shocks

By conditioning on the initial health stock, we may interpret variations in predicted latent health as departures from its initial health stock, termed as health shocks. A drawback of the latent health variable concerns its continuous nature. It does not allow us to specify various types of health shock variables in order to identify their respective impact on the probability to quit daily smoking.

As a starting point, we define a health shock as a departure of any size and magnitude from the health stock in the preceding time period. Using the five categories of self assessed health as reference scale of discrete health developments, we can summarize the possible manifestations in a table.

Type of health shocks	Presence	Direction	Size	Direction & Size	Preceding level	Direction, size and level (in square brackets)
Value or/and sign	0/1	+ve	1	- 4	1	- 4 [5]
		- ve	2	- 3	2	- 3 [5,4]
			3	- 2	3	- 2 [5,4,3]
			4	- 1	4	- 1 [5,4,3,2]
				0	5	0
				1		1 [4,3,2,1]
				2		2 [3,2,1]
				3		3 [2,1]
				4		4 [1]

Table 1

A health shock either occurs or it does not. If it does, it is positive and improving in health or negative and associated with a worsening of health. Its magnitude ranges from one to four in either direction; as magnitude and direction interact, we can already identify eight different types of health shocks. For instance, -2 in column “Direction and Size” denotes a worsening in health over two categories of the discrete SAH scale[3]. Considering the level from which the health stock departed in the preceding period, we obtain twenty different manifestations. As an example, - 4 [5] in column “Direction, size and level” describes a decrease in health (-ve) from very good health [5] over four categories of the SAH scale (4) to consequently very bad health. Thus, health shocks differ by sign, size and the level of health in the preceding time period.

A priori theoretical expectation regarding the relationship between different types of changes in health and smoking cessation is vital to guide our empirical analysis. With respect to the theory discussed in section one, the conventional rational addiction model predicts that smokers quit as the disutility from smoking exceeds the utility of addictive consumption. This may lead us to conclude that the probability to quit smoking increases with the magnitude of a negative health shock. We might also expect that there is a break point in latent self assessed health, for example, a cut off marking the change from reported good to fair health, at which the disutility from worsened health exceeds the utility of addictive consumption.

[3] The previous level is not considered yet. Therefore it may for instance describe a drop in health from good to bad or very good to fair health.

However, Etile's model refines this prediction in the sense that adverse health developments initiate a learning process that changes the output parameter of the health production function, ceteris paribus, leading to higher cut backs in tobacco consumption and probability to quit daily smoking. Thus, we would expect smoker's to learn as well from smaller changes in health.

Following the discussion of types of health shock and related theory, we argue that a discrete measure of latent health would be useful to explore the impact of various shock types on the probability to quit smoking. We construct this discrete version of latent health by relating the predicted values back to the original SAH scale using the estimated cut-offs from the ordered probit model.

(4) $\hat{h}_{it}^{*D} = 1 \text{ if } \hat{h}_{it}^{*} < \hat{\mu}_1$

$\hat{h}_{it}^{*D} = 2 \text{ if } \hat{\mu}_1 \leq \hat{h}_{it}^{*} < \hat{\mu}_2$

$\hat{h}_{it}^{*D} = 3 \text{ if } \hat{\mu}_2 \leq \hat{h}_{it}^{*} < \hat{\mu}_3$

$\hat{h}_{it}^{*D} = 4 \text{ if } \hat{\mu}_3 \leq \hat{h}_{it}^{*} < \hat{\mu}_4$

$\hat{h}_{it}^{*D} = 5 \text{ if } \hat{\mu}_4 \leq \hat{h}_{it}^{*}$

where $\hat{\mu}_{ii}$ denotes the estimated cut off i.

There are two further advantages of this approach. First, by defining differences in discrete categories all minimal changes in latent self assessed health within the categories are not defined as health shocks. Second, individuals are required to map their latent health status to the five category scale of self assessed health. This involves a decision concerning the cut offs at which the reported categories change. We argue that a change in latent self assessed health that crosses the estimated cut off might come closest to the idea of a health shock in the sense of a conscious (as prerequisite for learning) change in "objective" health.

3. Econometric Specification

We use a discrete time hazard model to describe the probability of quitting as a function of covariates and time spent in the initial state of daily smoking. Thus, in the spirit of duration models the time dynamics of smoking cessation are captured through duration dependence.

In particular, we adopt a stock sampling approach originally proposed by Jenkins (1995) and our comments follow closely his explanations. The approach relies on defining the duration model stock sample such that only individuals at risk of the event "quitting daily smoking" are included in the analysis. Consequently our risk set consists of N individuals i = 1, ... , N , who each smoke daily at time t=0.

The starting point for the econometric specification is the hazard rate function for person i at time t >0. It describes the instantaneous probability of transit into a state other than daily smoking, conditional on the survival time t under the assumption of proportional hazards.

(5) $$\lambda_{it} = \lambda_0(t) \cdot \exp\left(X_{it}{}' \beta \right)$$

where $\lambda_0(t)$ is the baseline hazard function describing time dependency, χ_{it} is a vector of covariates summarizing the observed differences between individuals at t and β is the vector of parameters to be estimated.

Considering the discrete nature of the data, the underlying (assumed) continuous time duration t is only observed in disjoint intervals $[0 = a_0, a_1), [a_1, a_2), [a_2, a_3), \ldots, [a_k, a_{k+1,} a_k = \infty)$ .The associated probability of failure in the j-th time interval is therefore given by the probability that duration lasts up to time interval j net the probability of survival until interval j-1. Using the concept of the survival function, we can write

(6) $$prob\{T \in [a_{j-1}, a_j)\} = S(a_j; X_{it}) - S(a_{j-1}; X_{it}).$$

where T denotes the end of the spell for person i (quit daily smoking) and S(.) the survival function. It should be noted that the covariates are allowed to vary between the time intervals but have a fixed effect on the probability of quitting within them.

Under the assumption of proportional hazard, the probability of survival up to the j-th interval is just given as a function of the covariates and duration dependence

(7) $$S(a_j; X_{it}) = \exp\left[-\exp(X_{it}'\beta + \delta_j)\right]$$

where $\delta_j = \log(\int_0^t \lambda_0(\tau)d\tau$ is the log of the integrated baseline hazard at t for all j=1,...k

The hazard of failure in the j th interval is consequently given by

(8) $$h_{it}(X_{it}) = 1 - \exp\left[-\exp(X_{it}'\beta + \delta_j)\right]$$

In the analysis, we record durations for each person i corresponding to the interval $[t_{i-1}, t_i)$ where t is in units of years. The person either quits daily smoking and leaves the interval or stays. Following Jenkins, we define the dependent variable as indicator variable y_{jt}=1 if individual i smokes daily during the interval $[t_{i-1}, t_i)$ and y_{jt}=0 otherwise. Using (4), we base the log likelihood on the sample hazard function specification

(9) $$h_{it} = 1 - \exp\left[-\exp(X_{it}'\beta + \theta_j)\right]$$

where $\theta(j)$ is the duration dependence. We may then write the sample log likelihood in binary response form such that

(10) $$\log L = \sum_{i=1}^{n_i} \sum_{j=1}^{t_i} \left\{y_{ij} \log h_j(X_{ij}) + (1 - y_{ij}) \log\left[1 - h_j(X_{ij})\right]\right\}$$

The log likelihood may be specified to allow for a non-parametric baseline hazard implying more flexible time effects. However, a non-parametric baseline does not suit the small variation among failure rates between the three time durations in our short panel (as pointed out in section four). This is why we decided to use a log Weibull specification for the baseline hazard assuming continuous time dependency. Thus $\theta(j)$ is specified as ln(t).

Furthermore, the log likelihood written in binary response form[4] allows us to estimate the model as discrete complementary log log or a random effects model with normally distributed unobserved heterogeneity. However, in our preferred specification[5], a gamma distributed random variable uncorrelated with the explanatory variables is added to describe unobserved heterogeneity. The instantaneous hazard rate of the incidental mixed proportional hazard model is then specified as:

(11) $$\lambda_{it} = \lambda_0(t) \cdot \varepsilon_i \cdot \exp\left(X_{it}{}' \beta \right) = \lambda_0(t) \cdot \exp\left[X_{it}{}' \beta + \log(\varepsilon_i) \right]$$

where ε_i is a gamma distributed random variable with unit mean and variance $\sigma^2 = \nu$ giving the corresponding discrete-time hazard function

(12) $$h_j(X_{ij}) = 1 - \exp\{-\exp\left[X_{ij}{}' \beta + \gamma_j + \log(\varepsilon_i) \right]\}$$

Again following Jenkins, the log likelihood of the model with unobserved gamma heterogeneity is then specified as

(13) $$\log L = \sum_{i=1}^{N} \log\{(1 - c_i) \cdot A_t + c_i \cdot B_i\}$$

---

4 And in praxis, a data set organised according to the time intervals each person is at risk of the event.

5 It is shown in section five that the gamma frailty model with log Weibull baseline hazard performs superior in terms of the likelihood and duration dependence fit. As a result we just present the full econometric specification of our preferred model.

Where $A_i = \left[1 + v\sum_{i=1}^{N} \exp\left[X_{ij}{}'\beta + \theta(j)\right]\right]^{-(1/v)}$

And $B_i = \left[1 + v\sum_{i=1}^{N} \exp\left[X_{ij}{}'\beta + \theta(j)\right]\right]^{-(1/v)} - A_i$, if ti>1or just

$= 1 - A_i$, if ti=1.

As mentioned above the functional form of $\theta(j)$ is specified as ln(t) assuming a continuous log Weibull baseline hazard. STATA sets the starting value of the gamma variance by default equal to .37. The limiting case of the log likelihood function is given when the gamma variance approaches zero[6].

4. Data

The European Community Household Panel Users Database (abbreviated ECHP-UDB) is an annual panel with approximately 130,000 individuals of 16 years and older in 60,000 households conducted in the European Union Member States. It provides eight waves (1994 – 2001) with microdata on demographics, income, social transfers, individual health, housing, education and employment collected with a standardized questionnaire. We used the last four waves (1997-2001) to analyse the impact of health shocks on the probability to quit daily smoking and the first eight waves to estimate latent self assessed health (1994 – 2001).

The first wave covers the 15 EU Member States in 1994. Austria and Finland joined when they become EU members in 1995 and 1996, respectively. In the first waves the ECHP was replaced by existing national surveys in Germany (SOEP), UK (BHPS) and Luxembourg (PSELL). From the fourth wave onward, the ECHP were substituted by adjusted data from the national surveys. Sweden did not take part in the ECHP but provided data on living conditions from its national database (Jones et al. 2005).

Variables & stock sample

---

[6] The model can be estimated using the STATA pgmhaz command. It is programmed to set the gamma variance equal to zero, i.e. run the log variance towards a very high number. This might cause problems in the estimation. Changing the start value of the gamma variance (it has been changed to .8 in our model specification) might help. The corresponding command is lnvar0(.). The trace option can be used to investigate whether the gamma variance causes problems in the estimation (Jenkins).

With respect to data on smoking, individuals were asked whether they smoke daily, smoke occasionally, used to smoke daily, used to smoke occasionally or have never smoked. We defined a binary dependent variable taking a value of one if the respondent is a daily smoker and zero if she smokes occasionally, used to smoke occasionally and used to smoke daily. Thus, in the duration analysis, hazard refers to the transit to the state of used to smoke daily, smoke occasionally and used to smoke occasionally[7].

The discrete time duration analysis requires us to organize the stock sample so that there is an observation at each time interval that a subject is at risk of failure. Thus, only individuals who are daily smokers in wave five, provided a complete sequence of responses until attrition or hazard and are observed from wave one on entry the analysis. Latent self assessed health is estimated on the complete stock sample from wave one to at least wave five in order to make use of all available information on health developments. Thus, we imposed the last restriction with the intention to estimate latent self assessed health using wave one up to wave eight on the same sample of individuals that enter the duration analysis[8]. Due to the late inclusion of the smoking variable in the survey from wave five on, we then drop wave one to four and conduct the analysis on a reduced sample with a maximum of four waves and three durations per individual .

Following the restrictions set by the definition of the stock sample, we drop Finland, Austria and Sweden due to missing waves, Germany and UK because of missing observations on objective health information on mental problems and inpatient stays that compare to information provided by other country surveys in wave one to four, and France and the Netherlands since consecutive observations on smoking are missing beyond wave five. Thus, our pooled data set consists of observations on Denmark, Belgium, Ireland, Italy, Greece, Portugal and Finland.

The estimation procedure on the resulting sample then follows three steps. First, four models with competing health shock variables are estimated. Three of the four health shock proxies

---

[7] Allowing for transition to the state used to smoke occasionally allows for slightly incoherent respond. This is because all persons included in the analysis claim to have smoked daily at some point in their former life so that - strictly speaking – the only feasible transition would be to the state of “used to smoke daily” or “smoke occasionally”. However, individuals who claim that they have never smoked after having reported to smoke daily in an earlier time period, are excluded from the analysis. This is because this “degree” of inconsistent response in the data led us to doubt the general quality of the data provided by the respective individual.

[8] As a result, we have information on lagged health shocks for wave five.

are binary variables just indicating the presence of a health development. Second, the preferred discrete shock measure from step one is used to explore the impact of different direction, size and preceding levels of a discrete change in health on the probability to quit daily smoking. Four models compete. Third, the preferred models from step one and two are stratified by gender. Beyond, the interaction of health shocks with education, income and age is considered in step one and three.

Construction of health shock measure

According to step one, we define four competing variables that proxy health shocks. The first is a binary variable taking a value of one if the person reports a change in her disability status and zero otherwise. It is derived from a question in the ECHP that asks all persons whether they are hampered in their daily activities by any physical or mental health problems, illness and disability[9]. The second health shock proxy utilizes self assessed health; the respective binary variable takes a value of one if self assessed health differs from the category reported in the preceding time period and zero otherwise. We argue that modelling variations in health, i.e. health shocks, eliminates the influence of person specific characteristics on shifts of the thresholds values used to map latent health to one of the categories of self assessed health. The third and fourth health shock measures are based on latent assessed health. The indicators in the latent health equation (3) presented in section two, as provided by the ECHP, refer to mental health problems, inpatients status, duration of hospital stay and chronic disease status. The exact questions are described in the appendix. The set of variables and estimated coefficients used to obtain predicted latent self assessed health are shown in the table two below:

[9] Possible responses are “severely”, “to some extent” and “not hampered”. The responds “Severely and to some extent” have been summarized to indicate the presence of a self assessed disability status.

Ordered probit regression

	Coef	Std. Error
Illness	-.5193107***	.0143603
Mentalprob	-.642007***	.0258848
Inpat	-.342503***	.0156508
Hospnight	-.0058723***	.0006751
Chronsev	-1.859303***	.0204274
Chronsome	-1.204351***	.0135996
Age	-.0211191***	.0002447

***significant at .001 level

Log likelihood = -106824.69 Number of observations 100307 Pseudo R^2 0.1492

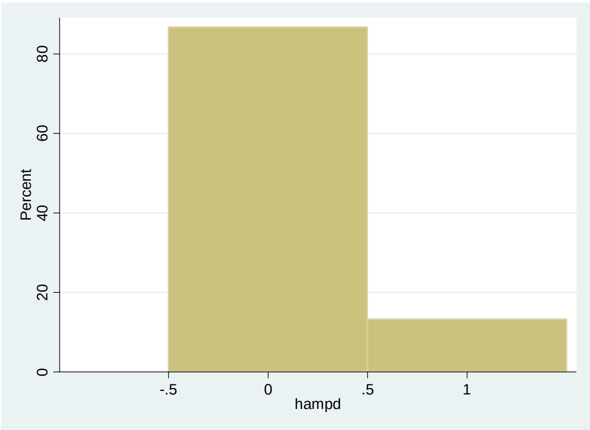
Table 2

Conditioning on the initial health stock in the first period enables us to interpret all variations in constructed latent health as a sensible, third, measure for health shocks. The fourth proxy measure, the discrete version of latent self assessed health is comparably inert. The sketch below shows how the discrete version translates the distribution of latent self assessed health into predicted SAH categories using the estimated cut-offs (as described in the previous section).

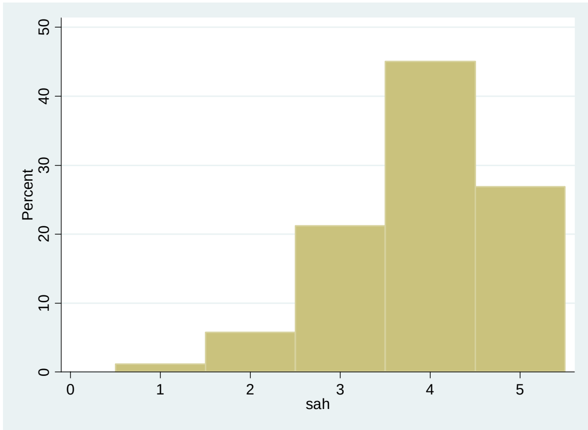
```
<------------------------------------------------------------------------------------------------------------------------→
                    -4.464188*                 -3.141247*              -1.816494*            -.420256*           0
SAHLAT

<---------------------[  cut  1  ------------------------[  cut2  --------------------[  cut3  --------------[  cut4  -------------------------→
                                                                                                        0              SAHLAT

<  --------------------[  cut  1   ------------------------[  cut  2  --------------------[  cut3---------------[  cut4  -------------------------→
                                                                                                        0       PRSAH
```

As specified for the self assessed health shock variable, we define any departure of discrete latent health from its preceding health stock to demonstrate a health shock.

The empirical distributions of the four health measures used to construct the respective four shock variables for the defined stock sample are shown below.

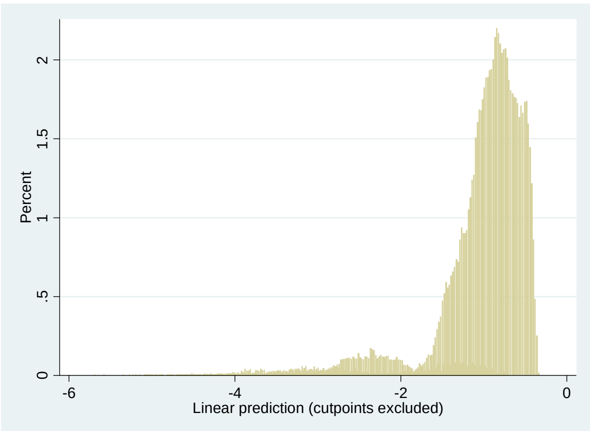


Graph 1

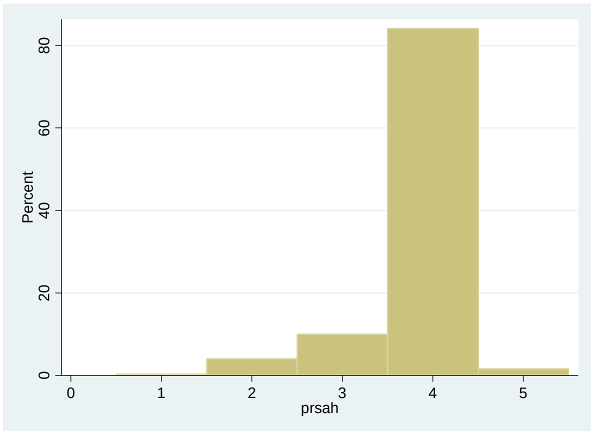


Graph 2

Graph one and two refer to the distribution of the disability status and self assessed health respectively. In approximately 15% of the sample observations, individuals indicate that they are (severely or mildly) hampered in their daily activities. The distribution of self assessed health is right centred. Most people report that they are in good health, followed by observations on very good and fair health. Only approximately 5% and 2% report bad and very bad health, respectively.

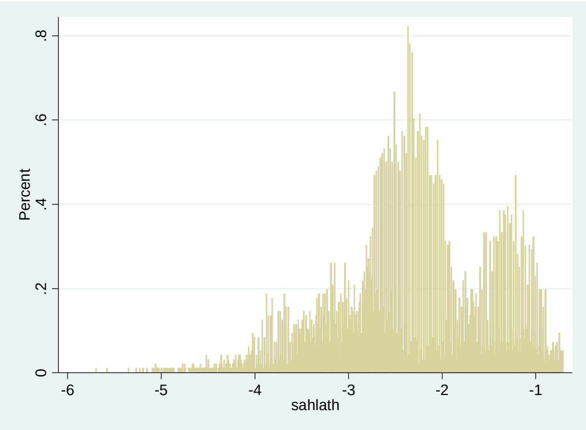


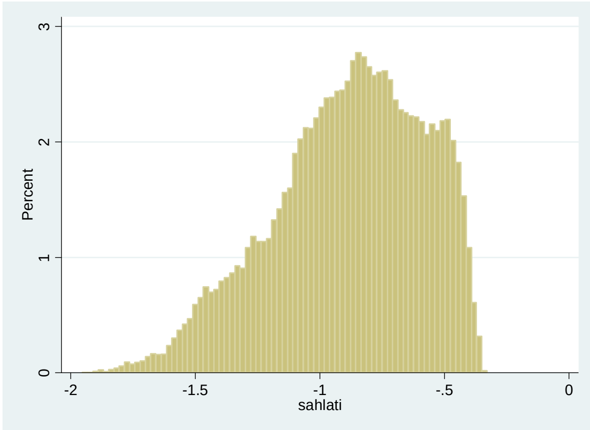
Graph 3



Graph 4

Graph three and four show the distribution of latent assessed health (Graph 3) and its discrete version. Latent health is right centred with a global peak at minus one (corresponding to reported good discrete latent health) and small local peak at approximately -2.5 (corresponding to reported fair discrete latent health). With respect to the effect of predicting SAH as function of „objective“ health indicators, it flattens the outer edges of reported SAH and concentrates the predicted values in the discrete category good health.





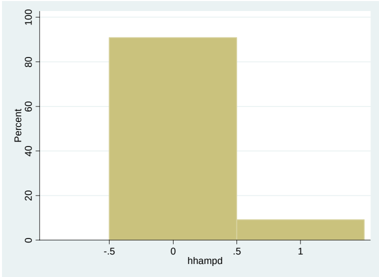
Graph 5

Graph 6

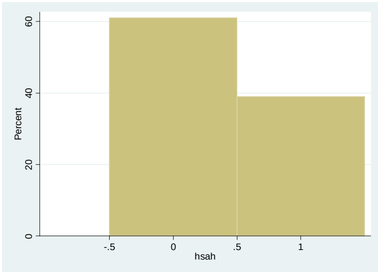
The two-peak distribution of continuous latent health (Graph 3) can be disentangled in individuals with at least one non-zero observation in the health indicators (Graph 5) and persons who report no objective health impairment (Graph 6).

In each of the four model specification in step one, we condition on the initial level of the health measure that we use to construct the respective health shocks variable. In doing so, we intend to control for the different probabilities to quit daily smoking as the start level varies.

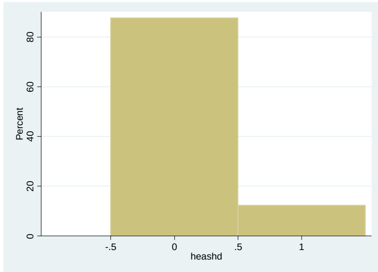
The three graphs below compare the resulting discrete health shock proxies. As one would expect, there are few observations on people who change the hampered status (Graph 7) while SAH is volatile (Graph 8). Forty percent of observations on self assessed health differ from the previously reported category. The discrete predicted self assessed health then settles the high amplitude in health shocks to ten percent (Graph 9).



Graph 7



Graph 8



Graph 9

Furthermore, to avoid the chicken or egg problem i.e. to disentangle whether individuals suffer from a health shock as a consequence of quitting smoking (presumably an improvement) or quit smoking as a consequence of a change in health, all shock measures are lagged. This implies a delayed change in smoking habits in response to health developments. The idea is illustrated below.

Wave 5	Wave 6	Wave 7	Wave 8
Predicted good health	Predicted fair health	Predicted good health	Preditced good health
No health shock	Health shock	Health shock	No health shock
No lagged health shock	No lagged health shock	Lagged health shock	Lagged health shock
Daily smoker	Daily smoker	Used to smoke daily	Daily smoker
Enters risk set	Exists	Dies	No obervation

Table 3

The bold black frame describes the respondent’s participation in the duration analysis. If the individual quits smoking in time period seven, its behaviour is explained by a lagged health shock (a change in health in wave six). Furthermore, since the example individual departed from the daily smoking state, it leaves the analysis (it is absorbed in the state of being an ex-smoker). Thus, his resumed smoking in wave eight is not considered in the analysis.

Exploring health shocks

With respect to step two, four competing models are specified to explore health shocks using the best performing of our three discrete health shock variables (in terms of the log likelihood). The first includes a dummy for negative shocks, the second expands the first with a dummy on positive health development, the third considers the interaction with magnitude and the fourth adds variables including the preceding health level. Referring to section two, the table below shows the four models.

Features of health shocks	Presence	Direction	Size	Direction & Size	Precending level	Direction, size and level (in square brackets)
Value or/and sign	0/1	+ve - ve	1 2 [3] [4]	(-4)* (-3)* - 2 - 1 (0) 1 2 (3)* (4)*	1 2 3 4 5	- 4 [5] - 3 [5,4] - 2 [5,4,3] - 1 [5,4,3,2] 0 1 [4,3,2,1] 2 [3,2,1] 3 [2,1] 4 [1]
Model I		Dummy for –ve health shock				
Model II		Dummies for –ve and +ve health shock				
Model III				Dummies for interaction of –ve and +ve health shocks with one and two changes in SAH		
Model IV						AB and AABB Dummies incorporating health level before shock

*restricted to a cell size of at least 30

Table 4

Model IV incorporates an idea implied by the conventional rational addiction model. It predicts that individuals quit smoking once the derived utility from addictive consumption is not sufficient to outweigh its harm to health. By intuition, we assume the point where disutility from a loss in health intersects with utility of smoking is at the cut off from good to fair health. Therefore a variable that takes value one if the predicted health status changes from very good or good health to fair, bad or very bad health and zero otherwise is constructed.

Resulting State

Initial State	Very good	Good	fair	Bad	Very bad
Very good	0 A	1	2	3 B	4
Good	-1	0	1	2	3
Fair	-2	-1	0	1	2
Bad	-3 BB	-2	-1	0 AA	1
Very bad	-4	-3	-2	-1	0

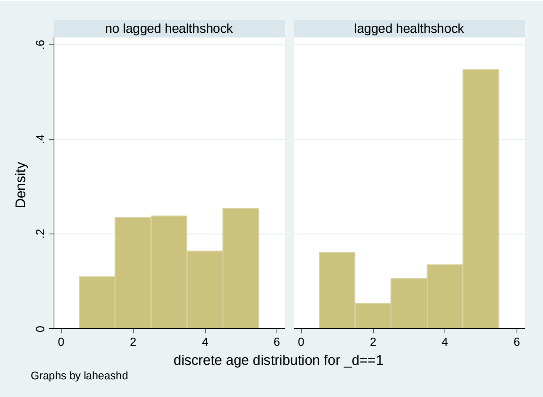
Table 5

With respect to table 5, the constructed variable takes a value of one if the respondent changes from her initial state A (blue frame) to the resulting state B (red frame). We also consider incisive positive health developments and define an AABB variable indicating moves from fair, bad and very bad health to very good or good health.

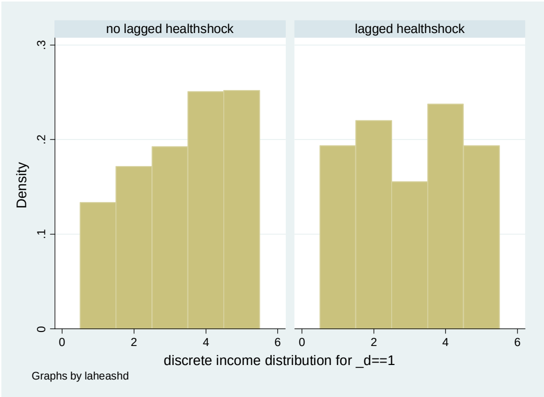
Interaction with socioeconomic characteristics

To control for socioeconomic characteristics that influence the probability to quit daily smoking, we include the logarithm for household income, dummies for stage three and below stage two education, a variable for age, marital status (dummies for separated, divorced, never been married), employment status (dummies for unemployed, self employed, retired, inactive, housework) and construct a proxy dummy variable for pregnancy. Country dummies control for national differences. Thus, the benchmark individual is Portuguese, employed, married not pregnant and holds a stage two education.

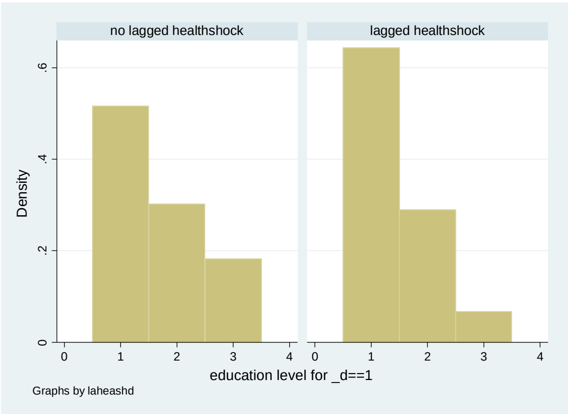
Furthermore, following the predictions of the Grossman model for demand of health, discrete negative health shocks are crossed with income, education and age. A glance at the descriptive statistics gives a first impression if the predicted causal relations will hold. The graphs below show the distributions of age, income, education and sex for the selected sample of failures (persons who quit daily smoking) divided again into those who experienced a health shock and those in stable health.



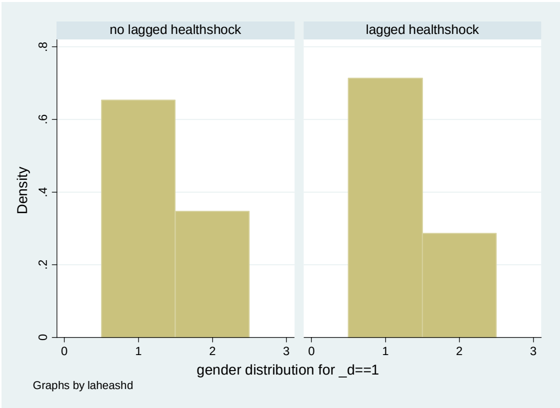
Graph 10



Graph 11



Graph 12



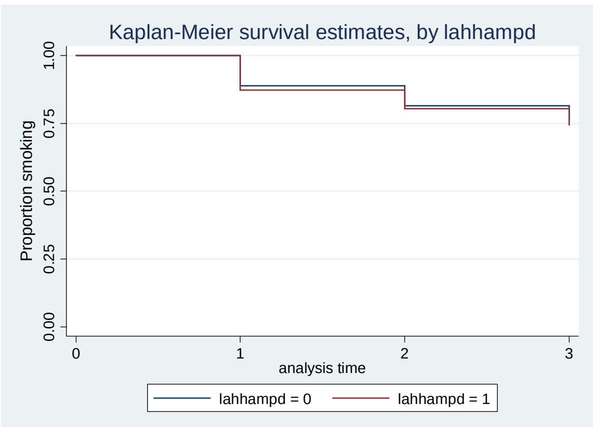
Graph 13

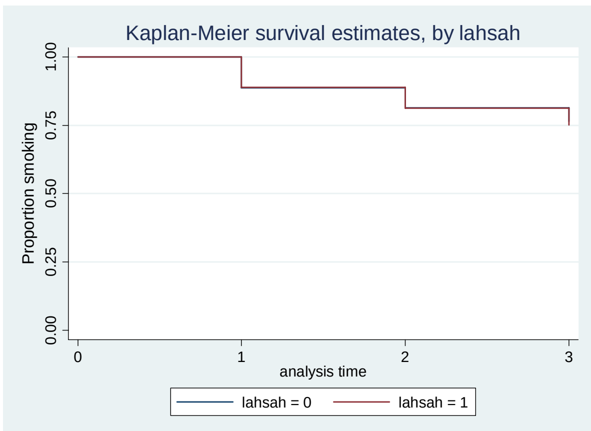
The proportion of age above 55 is considerably higher among the set of quitters with a lagged health shock compared to their counterparts in stable health (Graph 10). Quitters with observation on lagged health changes are evenly distributed across income quintiles (Graph 11). With respect to graph 12, the observations one, two and three refer respectively to less than second stage, second stage and a third stage education. Thus, we obtain a counterintuitive result; low education is relatively more observed among quitters with health changes. Finally, female quitters are more often associated with non-zero observations on health shocks compared to men (Graph 14).

To explore the causal relationship between age, income and education, the respective dummies are each crossed with lagged health shocks. The interaction with age is represented by two dummies, the first taking a value of one if the person is aged between 16 and 39 years and has a lagged health shock and zero otherwise, and the second indicating a health shock at age above 55 years. Both compare to having a health shock when between 40 and 54 years[10].

Descriptive Statistics

Lifetables provide us with estimates, known as Kaplan Meier survival estimates, of failure (other options are survival and hazard) as the underlying survival time is assumed to be continuous but has been observed in grouped form. It requires us to make an assumption about the underlying continuous hazard rate; following STATA's default we assume that failures occurs at a uniform rate within the intervals so that the estimates reflect the midpoint of the intervals (idea of acturial adjustment); Plotting the respective failure functions (one minus the survivor function) stratified by subgroups according to lagged health shock versus no lagged health shock for the three competing binary health shock measures yields the following graphs.

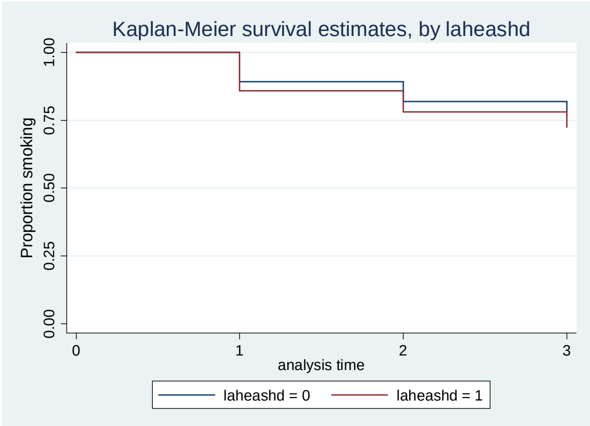




Graph 14 Graph 15

[10] The interaction dummies for income and education are described in the appendix.

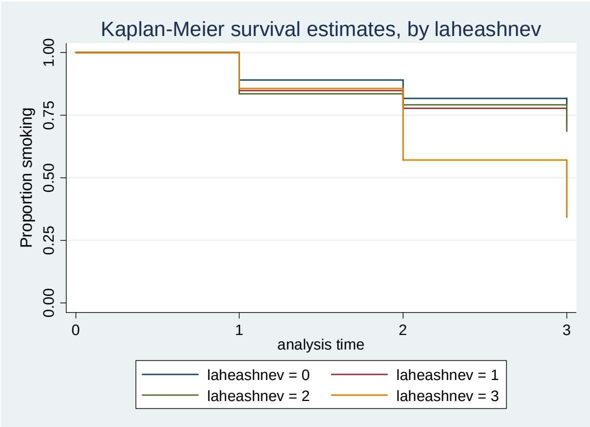
Respondents who experience a change in the disability status appear to have a slightly higher hazard rate (Graph 14). There is reason to doubt the significance of the small eyeballed difference, and the log rank test and likelihood ratio confirm that we fail to reject the null hypothesis of no subgroup difference in the failure function for all conventional significance levels. The empirical Kaplan Meier estimates and respective log rank and log likelihood tests can be found in the appendix. Furthermore, there is no distinguishable difference in failure for person in stable health and those who report lagged self assessed health shocks either (Graph 15).

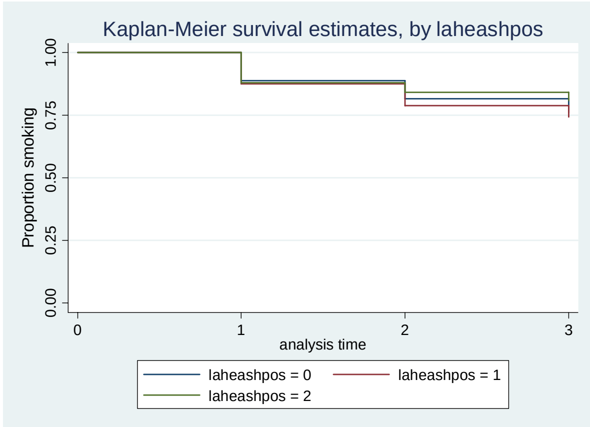


Graph 16

The discrete self assessed health shock variable shown above performs best in terms of separating the subgroups by health shocks and we find that predicted lagged health shocks are associated with statistically higher failure rates among respondents who quit daily smoking (Graph 16).

To disentangle the impact of different types on health shocks on the decision to quit smoking, we further stratify subgroups according to direction (positive and negative) and magnitude of health shocks.





Graph 17

Graph 18

The first graph shows negative shocks varying by magnitude (Graph 17). The blue line at the top displays the benchmark of no health shock. All kinds of negative health shocks are associated with a higher probability of failure and there is some evidence that the failure rate increases with the magnitude of the health shock in the last duration. The chi-square value associated with the log rank test is sufficiently large (25.67) to us to reject the null hypothesis of no subgroup differences. The likelihood that the observed difference occurred by chance is less than .001.

Lagged positive health shocks only show a differentiated picture in the third duration (Graph 18): the blue benchmark line is framed by a two category shock from above and one category shock from below. However, there is no statistical evidence for a subgroup difference.

Results

Results from step one: Estimating the competing models with the four specified health shock measures using a discrete time duration model with gamma frailty, we find the discrete predicted health shock measure and a log Weibull baseline hazard to perform superior in terms of the log likelihood. Strong evidence for unobserved heterogeneity confirms the correct model choice[11] and this result holds as we compare to the log likelihood of a random effects model with normally distributed unobserved heterogeneity (Table 7). All outputs can be found in the Appendix.

[11] The log likelihood test can be found in the Appendix. The null hypothesis of no unobserved heterogeneity (rho is equal to zero) can clearly be rejected at the 0.001 significance level.

Health shock specification (statistical signifcance)	Log likelihood	
	Gamma frailty model	Random Effects model
Baseline Hazard Specification	Weibull lnt	Weibull lnt
Lagged change in hampered status	-5453.85	-5502.72
Lagged change in reported self assessed health variable	-5521.37	-5547.50
Lagged change in variation in latent self assessed health	-5335.02	-5371.63
Lagged change in predicted discrete latent health **	-5168.70	-5240.16

** significant at 0.5 level

Table 7

The preferred discrete predicted health shock measure is statistically significant at the 0.05 level. There is no statistical evidence that the other health shock measures help to explain the probability to quit daily smoking. Neither are any of the initial health stock variables statistically significant. The null hypothesis of no unobserved heterogeneity can be rejected at the 0.001 level. The coefficients of the preferred specification Model A are presented below.

Model A

Discrete time duration model with unobserved gamma frailty and log Weibull baseline hazard		
	Coefficients	Standard error
Log time dependence	-.5591***	(.1281)
Dummy for lagged change in discrete latent self assessed health	.1626**	(.0817)
Level of initial discrete latent self assessed health stock	.0162	(.0498)
Logarithm of yearly equivalised household income	-.1201***	(.0311)
Age	-.0627***	(.0103)

Age squared	.0007***	(.0001)
Third level education	.3370***	(.0828)
Less than second level education	-.2128***	(.0627)
Pregnant	.9730***	(.2155)
Female	-.0754	(.0642)
Seperated	-.5858**	(.2281)
Divorced	-.2868*	(.1628)
Never been married	-.1529**	(.0757)
Self employed	-.0313	(.0781)
Unemployed	-.1145	(.1112)
Retired	.1208	(.1051)
Housework	.0448	(.1033)
Inactive	.2166	(.1566)
Dummy for danmark	-.0096	(.1293)
Dummy for belgium	-.0036	(.1363)
Dummy for ireland	.5625***	(.1326)
Dummy for italy	.3583***	(.1061)
Dummy for Greece	.1133	(.1075)
Dummy for Spain	.4332***	(.1057)

***significant at .1, ** significant at .05, * significant at .001

Table 8

Surprisingly, the probability to quit smoking decreases in log household income. To investigate whether this result depends on regional differences, we crossed log household income with the country dummies and added the interaction variable to Model A specification. In Spain, the probability to quit daily smoking increases with income. All other interaction dummies are not significant (but the group of dummies may be). However, the statistically non-significant interaction dummies are collinear with the country dummies causing potential problems in the implementation of the STATA pgmhaz command. Therefore, we decided to proceed with the former model. The output can be found in the appendix.

Furthermore, the probability to quit daily smoking is decreasing in age. Cube age, included to pick up partial effects that vary with the level of age, sets off some of the decreasing effect more than proportionately as age increases. There is a clear education gradient in the sense that the probability to quit daily smoking increases with the stage of education. The proxy for pregnancy has a comparably high positive impact. The coefficients reflecting marital status are all statistically significant. Being separated, divorced or never been married adversely affects smoking cessation compared to being married. The same holds for unemployment and self employment in comparison to usual work arrangements. The coefficients in retirement and housework are positive but not statistically significant. One could argue that the negative significant characteristics of marital status and employment proxy increase distress in relation to their comparators. If this assumption holds true, the rational addiction model explains the decreased smoking behaviour by a general lower baseline utility resulting in relatively higher utility of addictive consumption. Furthermore, being Irish, Italian and Spanish significantly increases the probability to quit in comparison to being Portuguese.

Results from step two: We use the preferred discrete health shock measure, that is predicted latent health shock (from Model A), to construct different manifestations of health shocks. Four competing models - as presented in section four – are specified and estimated using the preferred discrete hazard model with gamma frailty. When disentangling the direction of health shocks only negative health shocks are found to be statistically significant. Crossed with the size of the change, only negative one category changes are significant. There could be two reasons for this finding: The number of high magnitude health shocks is not sufficient to pick up statistically significant effects or the probability to quit does indeed not increase with the magnitude of a health shock. However, smokers respond to small negative health shocks; this finding is consistent with Etile's idea of a dynamic learning process. Furthermore, changes from good or very good health to bad, fair or very bad health have significant predictive power. In terms of the likelihood, the specification with a negative lagged health shock performs best. However, the log likelihood of the benchmark model A with a binary variable for a lagged discrete health shock indicates an overall better model fit.

Discrete time duration model with unobserved gamma frailty and log Weibull baseline hazard					
	Benchmark Model A	Model 1 Direction	Model 2 Direction	Model 3 Direction & Size	Model 4 Direction, Size & level
Predicted initial latent health stock	.0162 (.0498)		.0245 (.0520)	.0296 (.0543)	.0186 (.0532)
Lagged binary health measure	.1626** (.0817)				
Lagged –ve health shock		.2452** (.0974)	.2445** (.0989)		
Lagged +ve health shock			-.0048 (.1155)		
Lagged –ve health shock with one category in SAH				.2730** (.1086)	
Lagged –ve health shock with two category change in SAH				.1310 (.2648)	
Lagged +ve health shock with one category change in SAH				.0755 (.1228)	
Lagged +ve health shock with tw category change in SAH				-.6358 (.4287)	
Lagged change from very good or good health to fair, bad or very bad health (AB variable)					.2451** (.1109)
Lagged change from fair, bad, very bad to good or very good health					-.0003 (.1169)
Log Likelihood	-5168.7	-5264.05	-5264.05	-5403.76	-5406.10

Table 9

As income, education and age interacting with negative health shocks (following theoretical guidance offered by the Grossman model) are added to the preferred Model A, we find no statistical evidence that the interaction with income and education add explanatory power to the equation. Consequently, the variables are dropped. In contrast, there is statistical evidence for an age gradient: the probability to quit smoking following a health shock is increasing in age. The result is given in the Appendix. The log likelihood of the model is as well improved.

Results from step three: Stratifying, both Model A and its version with age interaction dummies (named Model B), by gender shows interesting differences. In the model for women, none of the health related variables are significant. The key quitting smoking driver is pregnancy. Once we drop the proxy for pregnancy and estimate the model new, the log likelihood decreases significantly:

Model A Log likelihood without pregnancy	Model A Log likelihood with pregnancy
-2481.18	-1783.52

Table 10

In the model for men, the initial health stock and age above 55 crossed with health shocks are significant. The probability of quitting is decreasing the more health stock the respondent holds. Furthermore, compared to respondents with health shocks between 40 and 54 years, males above 55 with negative shocks are more likely to quit. The stratified models are presented below.

Discrete time duration model with unobserved gamma frailty and log Weibull baseline hazard				
	Female		Male	
	Model A	Model B	Model A	Model B
Log time dependence	-.4355*** (.1275)	-.4645*** (.1311)	-.2444* (.1327)	-.3596*** (.1174)
Level of initial discrete latent self assessed health stock	.1366 (.086)	.0897 (.0885)	-.0810 (.0620)	-.1197** (.0591)
Dummy for lagged change in discrete latent self assessed health	-.0782 (.1559)	-.1640 (.1980)	.2106** (.0937)	.0336 (.1135)
Lagged health shock crossed with observation age 16 to 39	-	.1428 (.4272)	-	-.3013 (.4077)

Lagged health shock crossed with observation above age 55	-	.1556 (.3344)	-	.5087*** .1582
Logarithm of yearly equivalised household income	-.1892*** (.0504)	-.1893*** (.0513)	-.0757* (.0413)	-.0873** (.0387)
Age	-.0445** (.0158)	-.0357** (.0180)	-.0716*** (.0142)	-.0597*** (.0129)
Age squared	.0004*** (.0001)	.0004** (.0001)	.0008*** (.0001)	.0006*** (.0001)
Third level education	.3931*** (.1177)	.3859*** (.1211)	.3677*** (.1079)	.3677*** (.1030)
Less than second level education	-.3131*** (.1051)	-.3135** (.1075)	-.1572* (.0780)	-.1340*** (.0750)
Pregnant	.9600*** (.2079)	.9710*** (.2204)	-	-
Seperated	-.4075 (.2973)	-.4027 (.2977)	-.6318* (.2854)	-.5858** (.2742)
Divorced	-.4601 (.2255)	-.4632** (.2259)	-.2067 (.2340)	-.1914 (.2221)
Never been married	-.1326 (.1235)	-.1384 (.1284)	-.2085** (.0953)	-.2335** (.0913)
Self employed	-.1309 (.1870)	-.1298 (.1902)	-.0065 (.0866)	-.0002 (.0824)
Unemployed	-.0838 .1786829	-.0098 .1837938	-.0537 .1351028	-.0266 .1313926
Retired	.2381 (.2008)	.2422 (.2015)	.1135 (.1285)	.0994 (.1205)
Housework	.1399 (.1189)	.1620 (.1208)	-	-
Inactive	.2388 (.3295)	.2499 (.3303)	.1719 (.1777)	.1812 (.1686)
Dummy for danmark	-.1774 (.2240)	-.2196 (.2272)	.1584 (.1636)	.1798 (.1564)
Dummy for belgium	-.3739	-.4271*	.2489*	.2782*

	(.2456)	(.2502)	(.1609)	(.1534)
Dummy for ireland	.2136 (.2218)	.1721 (.2257)	.8966*** (.1710)	.8597*** (.1601)
Dummy for italy	.0459 (.2091)	.0023 (.2127)	.5640*** (.1252)	.5579*** (.1197)
Dummy for Greece	-.1715 (.2169)	-.1780 (.2195)	.1879 (.1247)	.1847 (.1191)
Dummy for Spain	.1214 (.1982)	.0731 (.2018)	.6166*** (.1245)	.5910** (.1177)
Log likelihood	1783.52	-1718.03	-4792.69	-4678.37
Lnvarg	-14.26	-12.19		

The dummy housework has been dropped since in the model for men since the cell size is below 30

Table 11

5. Discussion

We find that a discrete change in objective health incentivises smoking cessation. The discrete latent shock variable significantly increases the probability to quit smoking while there is no evidence that subjective self assessed health or self assessed disability are relevant. The result is consistent with Etile's learning model and the demand for health model; in contrast, the conventional rational addiction model stresses the importance of perceived health in smoking decisions. Furthermore, continuous latent health, the sensitive measure for health shocks, is not statistically significant either. This supports the idea that only discrete – incisive – health developments may explain a change in smoking habits.

Further differentiation between types of health shocks shows that only objective negative changes significantly increase the probability to quit daily smoking. With respect to Etile's model, individuals learn from negative shocks, update their smoking risk and revise the parameters of their health production function. Apparently, small – one category - changes in health would be sufficient to initiate a learning process. In addition, we find health changes that cross the cut off from good to fair latent health to be significant. Thus, not only direction and size of a shock but also its preceding level of health matters. Positive changes do not significantly increase the probability to quit smoking as it is predicted by the rational addiction models.

Stratifying by gender reveals a new pattern. As yet, education, marital and employment status have the same sign and (presumably) about same significant impact. Clearly pregnancy is the key quitting driver in the model for women while health shocks do not seem to have an effect. In contrast, men are responsive to negative shocks and a decrease in health is even more important in the decision to quit smoking as age increases. In accordance with the Grossman model, we can argue that men intend to hold a certain health stock and use smoking cessation as compensatory tool for a loss in health. This is also consistent with our finding that especially shocks to low levels of health, i.e. fair, bad or very bad health are important. However, the naturally arising question is why do men but not women respond to health shocks.

We argue that the increased smoking hazard for men above 55 years is likely to be associated with a perceived increased risk of coronary heart disease. Health warnings concerning the relationship between smoking and the risk to develop coronary heart disease for men are widespread. The responsiveness to negative shocks is presumably a reaction to these warnings. With respect to health warnings, stratifying by men and women could then be viewed as natural experiment between two groups that bear a smoking related risk of coronary artery disease but only one group, here men, that is sufficiently informed about its risk. We argue that the natural experiment arises since the risk of coronary artery disease for women increased in the past years but this development has not infilitrated the public consciousness and health messages.

If this scenario comes close to reality, then men only learn from health shocks and quit as they have prior (superior) information about the link of their objective health developments to lifestyle and therefore the effectiveness of smoking cessation in the production of health. Women are not provided with information about their risk; they do not link an objective negative health shock to the harm of smoking and, thus, do not recognize cessation as appropriate tool to control the health stock i.e. produce health.

Thus, smokers update information about adverse consequences of tobacco consumption and are more likely to quit following an objective significant change in health. However, they apparently only do so as they have specific prior information about the influence of smoking on their adverse health developments that enables them to assess the effectiveness of smoking cessation in the production of owns health. Thus, a learning process that leads to smoking cessation might follow two steps: first, knowledge about potential adverse effects of smoking is accumulated and then, secondly, the information is linked to own health developments. Following this thoughts, we conclude that effective tobacco control policy provides specific information about smoking related adverse health developments that enables the individual to link his own health experience to the health warning.

References

Becker and Murphy (1988): “A theory of rational addiction” *Journal of Political Economy* 96, 675-700.

Bound, J (1991): “Self reported versus objective measures of health in retirement models” *Journal of Human Resources*, 26(1), 106-138

Bound, J. Schoenbaum, M, Stinebrickner, TR., Waidmann, T (1999) “The dynamic effect of health on the labor force transition of older workers” *Labour Economics*, 6 (2), 179-202

Clark A, Etile F (2002): “Do Health Changes affect smoking? Evidence from the British panel data” *Health Economics*, 21, 533-562

Disney R, Emmerson C, and Wakefield M.(2006): “Ill-health and retirement in Britain: a panel data-based analysis”. *Journal of Health Economics*

Etile (2000): Usage de drogues et dependance: une analyse economique. PhD Thesis. Univerity of Paris-1. Paris.

Grossman M (2000): “The Human Capital Model”, in The Handbook for Health Economics (Culyer, A. et Newhouse, P. eds.), Paris, vol. 1A, chap.7, 347-408

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

Jenkins, Stephens P. “Easy Estimation Methods for Discrete-Time Duration Models.” Oxford Bulleting of Economics and Statistics., 1995, 57 (1), 129-138

Jones, A.M. (1994): “Health, addiction, social interaction and the decision to quit smoking”. *Journal of Health Ecoonomics* 13, 93-110

Jones, van Doerslaer, Bago d’Uva, Balia, Gambin, Hernandez Quevedo, Koolman and Rice (2005): “Health and Wealth: empirical findings and political consequences”; in *Perspektiven der Wirtschaftspolitik,* 7 (Special Issue): 93-112

Jones A M, Rice N, Bago d’Uva T, Balia S: Applied Health Economics (forthcoming)

Lindeboom M, Doorslaer E.V., Bago d’Uva T, O’Donnel O, Chatterji S (2006): “Does reporting heterogeneity bias the measurement of health disparities?” *Tinbergen Institute Discussion Paper No. 2006-033/3*,

Smith K, Taylor D J, Sloan F A, Johnson F, Desvogues W H (2000): “Do smokers respond to health shocks”. Working paper 8. Duke University. Department of Economics.

Stigler, G.J., Becker, G.S. (1977). “De Gustibus Non Est Disputandum” *American Economic Review* 67. 76-90.