Are Mature Smokers Misinformed?*

Ahmed Khwaja, Fuqua School of Business, Duke University
Dan Silverman, Department of Economics, University of Michigan and NBER
Frank Sloan, Department of Economics, Duke University and NBER
Yang Wang, Department of Economics, Duke University

JEL Classification: D81, D84, I12, L15, L66
Keywords: Subjective beliefs, mortality, health, smoking

Abstract

While there are many reasons to continue to smoke in spite of its consequences for health, the concern that many smoke because they misperceive the risks of smoking remains a focus of public discussion and motivates tobacco control policies and litigation. In this paper we investigate the relative accuracy of mature smokers' risk perceptions about future survival, and a range of morbidities and disabilities. Using data from the Survey on Smoking (SOS) conducted for this research, we compare subjective beliefs elicited from the SOS with corresponding individual-specific objective probabilities estimated from the Health and Retirement Study. Overall, consumers in the age group studied, 50-70, are not overly optimistic in their perceptions of health risk. If anything, smokers tend to be relatively pessimistic about these risks. The finding that smokers are either well informed or pessimistic regarding a broad range of health risks suggests that these beliefs are not pivotal in the decision to continue smoking. Although statements by the tobacco companies may have been misleading and thus encouraged some to start smoking, we find no evidence that systematic misinformation about the health consequences of smoking inhibits quitting.

* This research has been supported by a grant from the Robert Wood Johnson Foundation's Substance Abuse Research Program 2004-07 "Why Mature Smokers Don't Quit." We thank Charlie Brown, Han Hong and two anonymous referees for their comments and suggestions that much improved the paper. All remaining errors are our own.

I. Introduction

While there are many reasons to continue smoking in spite of its well-documented consequences for health, the concern that many smoke because they misperceive its risks has been a focus of public discussion and has motivated tobacco control policies and litigation. Tobacco companies have been accused of manipulating beliefs about of the health risks of smoking (Glantz et al. 1996; Hilts 1996; Kluger 1996), in part by using marketing to affect perceptions (Hanson and Kysar 1999). These allegations have provided a basis for lawsuits filed by individuals and states, including litigation which resulted in the 1998 Master Settlement Agreement (MSA) between tobacco companies and state attorneys general, and which will bring millions of dollars to each of the states as compensation for the costs imposed by smoking. Public information campaigns and compulsory messages on cigarette packs about the harms of smoking also reflect an underlying assumption that smokers misperceive the risks of smoking.

This paper uses new and unique data to examine the relationship between subjective beliefs about longevity, diseases and disability, and the decision to smoke among older people. We do not attempt to answer the difficult question of whether being misinformed *causes* an older person to continue smoking. Instead, we look for evidence consistent with this hypothesis. Specifically, we investigate whether mature smokers are more prone to underestimate their health risks than their contemporaries who do not smoke. If, as the lawsuits allege and some tobacco control policies imply, misinformation causes people to underestimate their health risks and continue smoking, we would expect that mature smokers would make worse, and specifically more optimistic, predictions about their health risks.

Different from prior research, to investigate this hypothesis we do not compare the average beliefs of smokers and non-smokers; instead, we compare individual-specific subjective and objective probabilities about longevity, diseases and disabilities. Our analysis also accounts for two aspects of belief formation of beliefs that have not been examined before: we consider how

risk perceptions about future events are affected by (i) beliefs about future behaviors, and (ii) expectations about the planning horizon (or longevity).

Past research has investigated the relationship between subjective beliefs and smoking decisions with mixed results. For example, Viscusi (1990) finds that, in aggregate, smokers and non-smokers alike overestimate the risk of lung cancer from smoking, while Schoenbaum (1997) finds that in aggregate heavy smokers significantly underestimate their mortality risk. There is also a large literature on smoking in epidemiology and psychology that examines beliefs about risk (e.g., Weinstein 1986; Leventhal, Glynn, and Fleming 1987; McKenna, Warburton, and Winwood 1993; Ayanian and Cleary 1999). This literature generally elicits risk perceptions by asking respondents to compare their personal risk with those of others and the relative risk of various activities. Studies such as these do not examine the accuracy of subjective beliefs in an expected utility framework but rather base comparisons on judgments of higher versus lower risk.

To evaluate risk perceptions about health and longevity, we use responses to a new survey of smokers ages 50-70, the Survey on Smoking (SOS), conducted for our research. The SOS is unique in eliciting subjective beliefs about a wide range of health outcomes in a format and with a sample that allows direct comparisons with longitudinal data from the Health and Retirement Study (HRS). The SOS also addresses a gap in the HRS. While the HRS includes a large number of questions on onset of diseases and disabilities, there are few questions on subjective beliefs about future health. With the SOS data we can examine whether mature smokers appear to be unusually sanguine about a broad range of health risks.

Our focus on mature smokers has the advantage of allowing us to link subjective risk assessments with detailed objective health risk information from the HRS. It may be, however, that underestimates of health risks are more substantial and consequential for younger people who have yet to decide whether to start smoking (though Viscusi's (1991) findings suggest that this is not the case). Certainly, smoking initiation at younger ages is a key determinant of adult smoking levels; and since our study is limited to older persons, we cannot speak to questions about initiation.

Nevertheless, encouraging the millions of adult smokers to quit remains a major public health priority in the U.S;[1] and indeed the choice to quit for adult smokers is both common and consequential. At least 30 percent of smokers make at least one attempt to quit annually (Fiore et al. 1990; Hughes et al. 1992); and Sloan et al. (2003, pp. 129-36) report substantial quitting and relapsing in the first four waves of HRS. Taylor et al. (2002) estimate that even a person who quits at age 65 has an increased life expectancy of 2.0 years for males and 3.7 years for females. In addition, Sloan et al. (2003, pp. 79-88) described health benefits from quitting during middle age that go beyond improvements in longevity. Because adult quitting decisions are so common and consequential, we view the health beliefs and smoking decisions of older smokers to be relevant both for anti-smoking policy design and for evaluating a basis of litigation against tobacco companies.

By combining HRS and SOS data, our study makes four contributions. First, we examine the relative accuracy of smokers' perceptions about the risks of mortality and several chronic diseases and disabilities. By the comparing accuracy of smokers' beliefs with those of non-smokers we can evaluate the hypothesis that older smokers continue to engage in this risky behavior because they differentially misperceive the risks to their health. Second, we examine how beliefs about future behaviors, i.e., smoking cessation, affect risk perceptions. It might be that current smokers appear to be optimistic because they condition on their expectations of quitting in the future--information typically unavailable to the researcher. Third, we study how expected longevity affects risk perceptions about becoming afflicted with various chronic diseases and conditions. This last analysis allows us to draw inferences about how planning horizons enter into perceptions about future risk. Fourth, our evaluation is individual specific; we compare subjective and objective probabilities for each individual in the SOS sample, unlike previous research that has conducted aggregate comparisons using sample means.

[1] About a quarter of adults aged 45-64 and almost half this proportion among persons 65+ (U.S. Department of Commerce (2008, Table 192).

We find that individuals ages between 50 and 70 are, on average, not overly optimistic in their perceptions of health risk. Put differently, people in this age group appear to recognize their risk factors when forming beliefs about longevity and future health events. Depending on the specific risk under examination, the entire sample's subjective assessment is, on average, reasonably close to the objective risk. If anything, smokers tend to be relatively pessimistic about these risks, especially lung disease. They, like never-smokers, tend to exaggerate the risks of lung and heart disease, and stroke, but otherwise have quite accurate beliefs about the risks of various disabilities and about longevity. The finding that smokers are either well informed or pessimistic regarding a broad range of health risks suggests that these beliefs are not pivotal in the decision to *continue* smoking. In this way our results suggest that differences in tastes or the (induced) costs of quitting, rather than beliefs, are driving behavior among mature smokers.

The rest of the paper is organized as follows. Section II describes our data. We discuss methods in Section III. Results are presented and discussed in Section IV. Section V concludes.

II. Data

We use two primary data sources. The first is the HRS, a panel study that began with a nationally representative sample of those born between 1931 and 1941, and their spouses (www.hrsonline.isr.umich.edu). The HRS oversamples blacks, Hispanics and residents of Florida. Participants in the HRS have been interviewed every two years since 1992. While the primary respondent to the first wave of the HRS was between 50 to 61 years old; spouses received an identical interview and could be of any age.

Our second data set, the Survey on Smoking (SOS), was collected by the research firm Battelle from October 2004 to January 2005 at three sites where Battelle offices are located, Durham, North Carolina, St. Louis, Missouri, and Seattle, Washington. There were three interviews: a screener to determine age eligibility and smoking status, administered by telephone; a second longer interview also conducted by telephone; and an in-person computer-assisted

interview. All information used in this paper came from the screener and the longer telephone interview.

The SOS sample consisted of adults ages 50-70 at the interview date. Since much of the survey dealt with smoking status, current smokers were oversampled. The analysis sample from the longer telephone interviews consisted of 663 individuals, 252 current, 257 former, and 154 never smokers. The response rate for the longer telephone interviews was approximately 80%.

The SOS has several strengths for our study. Most important, it elicits subjective risk perceptions on a wide range of health conditions and longevity, as well as beliefs about quitting among persons who smoked at the time of the survey. The SOS asked respondents to evaluate their own risk of developing a variety of diseases and physical limitations. With this information we can compare these subjective assessments to objective risks estimated from the outcomes of respondents to the HRS. The HRS, in contrast to the SOS, only asks about subjective beliefs about survival to certain ages, not about beliefs about future onset of specific diseases and physical limitations. .

The SOS also has some weaknesses, and since the SOS was patterned on the HRS, these weaknesses tend to be common to both surveys. First, the SOS sample is limited to persons ages 50-70. While this age range matches the relevant HRS sample, it prevents us from studying young people. Second, Viscusi and Hakes (2003) have questioned the validity of the HRS survival expectations questions as representations of probabilities. Their empirical analysis indicates that the 0 to 10 scale used to elicit subjective probabilities in the first wave of the HRS does not satisfy all of the properties of well-behaved probabilities. Although Viscusi and Hakes only studied the survival expectations question, the same issues may arise in analysis of expectations about morbidity and disability onset which are included in the SOS.[2]

A third limitation pertains to the frequency of "focal" responses to expectations questions, i.e., the propensity to give subjective probabilities at extreme values or precisely in the middle (at

0.5).[3] Our analysis includes observations with these values since our goal is to compare subjective and objective probabilities for all persons in the sample. If, for example, those who give focal responses are especially misinformed, excluding them could produce misleading conclusions about the accuracy of subjective beliefs.[4] In analysis of the accuracy of subjective beliefs about survival using data from the HRS, Khwaja, Sloan, and Chung (2007) performed sensitivity analysis and found that their results were not importantly affected by whether the focal responses were included. A fourth limitation is that the SOS elicited responses in units of one between one and ten. Thus, converted to probabilities, the responses are coarser than we would like and require rescaling. Section III. B. provides more detail on the SOS measures of beliefs.

Table 1 shows summary statistics for the HRS at baseline in 1992 and for those who were both interviewed in 1992 and re-interviewed in 2004, and for the SOS in 2004-5. Compared to the HRS at baseline, higher proportions of SOS respondents were current or former smokers and, correspondingly, the proportion of never smokers was lower (Panel A). By 2004, the proportion of current smokers in the HRS was about half as large as at baseline. The HRS baseline and SOS samples are roughly comparable in terms of self-assessed health. SOS respondents were slightly older on average, and less likely to be of Hispanic, African American, or Asian origin, male, or currently married than were HRS respondents. They also had higher educational attainment on average. The health status of the HRS baseline sample who were re-interviewed in 2004 tended to be lower than either the baseline sample or the SOS. Panels B and C present probabilities of mortality, morbidity, and disability outcomes.

III Methods

To evaluate risk perceptions, we compare subjective measures of risk perceptions from the SOS with objectively estimated counterparts from the HRS for each *individual*. We think of the

[2] For a more detailed evaluation of the HRS survival probability questions, and by extension, the SOS questions on the same, see Sloan, Smith, and Taylor (2003, pp. 115-7).
[3] Börsch-Supan (1998, p. 306) argued that values of 0.5 could be round off errors.
[4] Gan, Hurd, and McFadden (2003) developed an approach for correcting focal responses to reduce the discrepancy between subjective and observed survival probabilities. Given our study goal (see the text), we do not perform this adjustment.

estimated measures of risk from the HRS as an objective benchmark against which subjective assessments can be compared. While any measure of objective risk is subject to specification error, our formulations are much more detailed than is typical of life tables which generally distinguish only between age, gender and race.

III. A. Objective Probabilities

Following Khwaja, Sloan, and Chung (2007), we use the HRS to estimate conditional hazard rates for mortality, and the onset of morbidities and disabilities, with data from seven waves of the HRS (1992-2004). These estimated hazards, which condition on characteristics of the individual (described below and including smoking status), are then used to compute the objective probabilities of each person being alive, developing specific diseases by 75, and acquiring specific disabilities within six years.

Specifically we estimate the relationships between observables and mortality, morbidities and disabilities using a hazard function, which allows for unobserved heterogeneity, and assumes a Weibull distribution.[5] The hazard function at time *t* for individual i in the HRS sample with vector of observables X_i is given by,

$$\lambda(t; X_i, \theta, \mu \mid \eta_i) = \lambda_0(t)*\exp(X'_i\,\theta)*\eta_i, \qquad (1),$$

where $\lambda_0(t) = \mu * t^{\mu - 1}$ is the baseline hazard for the Weibull distribution with the shape parameter μ, and $\exp(X'_i\ \theta)$ is the proportional hazard with parameters θ. We account for time invariant individual-specific unobserved heterogeneity by including a multiplicative term η_i (>0), distributed as gamma with mean 1 and variance σ (see e.g., Lancaster 1979).[6] Time to failure is defined as years elapsed from the interview to death or the morbidity outcome. All survivors to the 2004 interview, as well as individuals who were lost from the sample but who could not be verified as dead (N=1,673), are treated as censored.[7] We estimate the parameters using maximum likelihood.

[5] Lancaster (1990) provides an excellent review of the estimation of duration models.
[6] Three assumptions required to account for unobserved heterogeneity are: (1) the heterogeneity term is independent of the covariates; (2) the heterogeneity term has a distribution known up to some parameters; and (3) the heterogeneity term enters the hazard function multiplicatively. The mean also needs to be normalized to one for identification.
[7] The individuals who were lost from the sample are treated as censored one year after the last interview.

For each individual i in the SOS sample, we use the parameter estimates and the individual's characteristics to compute the objective probability of surviving to age 75, or not having a specific health outcome by age 75, $\hat{O}_i(t; X_i, \hat{e}, \hat{i})$. This is done using the following relationships:

$$\hat{O}_i(t; X_i, \hat{e}, \hat{i} \mid \hat{\varsigma}) = \exp\{-[\hat{\varsigma}_i * \exp(X_i' \hat{e})] * t^{\hat{\mu}}\} \qquad (2a),$$

and,

$$\hat{O}_i(t; X_i, \hat{e}, \hat{i}) = E_\eta[\hat{O}_i(t; X_i, \hat{e}, \hat{i} \mid \hat{\varsigma})] = [1 + \hat{\delta} * (\exp(X_i' \hat{e})) * t^{\hat{\mu}})]^{-1/\hat{\delta}} \qquad (2b),$$

where $\hat{e}$, $\hat{\mu}$ and $\hat{\delta}$ are estimated from (1), and t is the number of years to age 75 from the current age of the individual.

From the HRS data, we construct the health and disability variables as follows. For mortality, we define a binary variable for death by selecting individuals who were between 50 and 62 at HRS wave 1 (1992), observing their survival through all seven waves (1992-2004). For the morbidities, we define binary variables for lung disease, stroke, and heart problems. We exclude from the analysis those who had the disease at baseline. For disability and onset of physical difficulties, we define binary variables indicating difficulty walking several blocks, climbing stairs, or lifting modest weights. A binary variable is created if an individual developed the disability in the six years following the baseline HRS interview. A period of six years is used because the HRS (and hence the SOS) questions about subjective beliefs for these outcomes had a six year time frame (see section III. B). For the following physical conditions, feet swelling, shortness of breath, fatigue, and wheezing, the HRS did not collect information until wave 3 (1996). For these variables we use wave 3 as the baseline and followed the respondents for the next eight years constructing binary variables for onset of conditions as described above.

In estimating the hazard we use information from wave 1 (1992) on the following characteristics (X_i): smoking status (current smoker, former smoker, and the omitted category, never smoker), self-reported health status (excellent--the omitted category, very good, good, fair,

and poor), and demographic characteristics (age, non-Hispanic white, male, married, educational attainment).[8]

III. B. Subjective Probabilities

The SOS elicited subjective beliefs about survival and three chronic diseases (heart and lung disease, and stroke) by age 75. The questions were worded similarly for survival and the three chronic diseases. For survival, the question was, "On a scale from 1 to 10, with 1 being 'not at all likely' and 10 being 'very likely,' what is the chance you will be alive at age 75?"

For the following disabilities, difficulty walking several blocks, difficulty climbing several flights of stairs, difficulty lifting or carrying weights over 10 pounds, the question was worded as "On a scale from 1 to 10, with 1 being 'not at all likely' and 10 being 'very likely,' what is the chance you will experience (name of disability) within the next six years?" For four physical conditions, feet swelling, shortness of breath, fatigue, and wheezing, the questions were worded as "What is the percentage chance you will experience (name of disability) within the next six years?" This question used the term percentage to keep it consistent with the corresponding question in the HRS.

To convert responses on the 1-10 scale into probabilities, we rescale these to be in the range 0-1.[9] There was some clustering of the raw responses at 1 or 10. For the 11 health outcomes of interest, the clustering at the low end of the scale ranged between 4.2% (for survival to age 75) to 39.2% (difficulty walking) of the sample. At the upper end, the range was 2.4% (stroke) and 42.5% (survival to age 75) of the sample.

An ideal investigation of the role played by beliefs in smoking decisions would obtain individual assessments of the joint distribution of adverse health events, conditional on smoking behavior. For many reasons, it is not practical to gather such complex information from survey respondents. Indeed many social scientists have voiced skepticism about the ability or willingness

[8] We do not describe the regressors in the hazards in detail as these are self-explanatory.
[9] For a value x between 1-10, the rescaled value is calculated as (x-1)/9.

of individuals to convey even unconditional beliefs in numerical form.[10] Moreover, even among those who interpret the responses to numerical risk questions as measuring expectations, there remains controversy regarding the appropriate methods for eliciting such responses and about their precise interpretation. Alternative methods include third-person frequency questions (as in "out of 100, how many will get lung cancer") or second-person chance questions (as in "what is the chance you will get lung cancer"); and for each of these types of questions, there are issues of scaling.

In what follows, when we compare the SOS subjective risk assessments to objective risk levels, we will interpret the former as beliefs. While evidence discussed below indicates that average responses are remarkably stable with respect to different methods of questioning, we are nevertheless cautious in this interpretation. We acknowledge the difficulties in obtaining subjective risk measures, but view the responses as far from uninformative. Instead we attempt to stake out a middle ground and focus our analysis on comparisons between smokers and nonsmokers. We ask whether, relative to their objective risk of morbidity and mortality, smokers are more or less optimistic in their beliefs than non-smokers. If the mapping between the relevant belief and the response to the survey question is the same for smokers and nonsmokers, the biases induced by the questions may be alleviated by these comparisons.

III. C. Analysis of Beliefs by Smoking Status

To evaluate beliefs by smoking status, we estimate regressions of the following form.

$$(S_i - \tilde{O}_i) = \alpha + \delta^* d_i + \beta^* p_i + \varphi^* q_i + \varepsilon_i \qquad (3),$$

where S_i is the subjective probability of a morbidity or disability outcome by age 75 or within six years from the SOS, $\tilde{O}_i = (1-\hat{O}_i)$ (see eq. 2b) is the objective probability counterpart to S_i that is predicted using the parameters estimated from the HRS, d_i is an indicator of smoking status (current, former, and never, the omitted group), p_i is the subjective probability of living to 75, q_i is the self-assessed probability by current smokers of quitting smoking in two years and is zero for others, and

[10] See Manski (2004) for a discussion of this skepticism.

ε_i is an error term assumed to be uncorrelated with the regressors. As eq. 3 is specified, the dependent variable can either be positive if the subjective probability overpredicts that an outcome will occur or conversely negative for an underprediction. If the estimate of δ for current smokers is positive, this means that smokers are relatively pessimistic about an adverse health outcome occurring and conversely for a negative estimate of δ. When the outcome being analyzed is surviving to a particular age, the interpretation is reversed. Since subjective beliefs about living to age 75 may be correlated with the dependent variable and also may be endogenous to the dependent variable in eq. 3, we show results with and without this explanatory variable.

The rationale for this specification that compares subjective and objective probabilities in levels is that in a standard economic model of decisions under uncertainty, choices depend on the levels of subjective beliefs. We interpret the smoking status parameter estimates as the differences in risk perceptions by smoking status. An alternative specification would put the measure of objective probabilities on the right hand side. We do not do this for two reasons. First, the current specification (eq. 3) allows the simple interpretation of the coefficient on the smoking status dummies as differences in risk perceptions by smoking status. Second, $\hat{O}_i$ is a predicted value and measured with error. Putting it on the right hand side would lead to biased estimates in the absence of an instrumental variable approach; and finding a variable that affects subjective probabilities only through its effect of objective probabilities and can be excluded from the main equation is a virtual impossibility.[11]

The specification also includes the subjective probability of living to 75 as a regressor. This is to account for the possibility that those believe they will live longer may also think that they have a higher probability of future onset of particular diseases and disabilities before age 75 since they are more likely to be alive. Alternatively, people with a higher expected longevity may believe that they will be healthier at an older age as well. It is possible that individuals may appear optimistic or pessimistic in the absence of accounting for these expectations.

[11] Instead we, in effect, assume that the coefficient on $\hat{O}_i$ in the alternative specification equals one.

Subjective beliefs of current smokers about quitting in the future are included as a regressor since people who think they will quit may have different beliefs about their future health relative to persons who believe otherwise. For example, smokers who condition on their expectation that will have quit in the near future may appear to be optimistic whereas in reality they may be reporting beliefs that are consistent with those of former smokers, which is what they expect to be. The specification is parsimonious because the objective probability variable, $\tilde{O}_i$, on the left hand side is constructed from a hazard that is estimated using many individual characteristics, e.g., demographics.

We also estimate a regression for the difference in the subjective and objective probabilities of living to age 75 with the following specification

$$(p_i - \hat{O}_i) = \beta + \gamma^* d_i + \xi^* q_i + \upsilon_i, \qquad (4),$$

where all of the variables are as defined above. Equation (4) allows us to examine the differences in beliefs about mortality by smoking status and whether smokers account for their quitting intentions in forming beliefs about future survival. If we find that beliefs about quitting have a significant impact on beliefs about survival, then this could explain, for example, why certain smokers have beliefs about survival more similar to former smokers.

IV. A. Results and Discussion

Table 2 presents Weibull hazard estimates for survival and onset of various diseases by age 75. These estimates are used to predict objective probabilities of these outcomes by age 75. Current smokers are more likely to experience each of the adverse health outcomes than are never smokers. In three out of the four cases, former smokers have a significantly higher probability of an adverse outcome. Even holding smoking status constant, worse self-reported health at the HRS baseline is associated with a lower likelihood of survival and higher likelihood of disease onset;[12] the mortality and the disease onset hazards are all monotonically increasing as self-reported health

[12] In results not reported, we include interaction terms between smoking status and self-reported health. The results are almost identical to those reported.

worsens. The other covariates, age,[13] gender, race-ethnicity, marital status, and educational attainment, matter in some of the regressions.

Table 3 shows Weibull hazard estimates for onset of various types of disabilities and physical conditions, again using data from the HRS. With one exception, feet swelling, being a current smoker is positively associated with the probability of onset of disabilities and physical conditions. The only statistically significant difference between former and never smokers is for wheezing. As before, hazards are monotonically increasing as self-reported health worsens. The patterns in the estimated hazards uniformly suggest that smoking is bad for a person's health and, judging from the parameter estimates on former smokers, that quitting the habit is health-improving relative to continuing to smoke.

Table 4 reports mean values of subjective and predicted (based on the estimated hazards reported above) objective probabilities of survival, onset of morbidity, disability, and physical conditions. The objective probabilities range from 0.046 (chronic lung disease by age 75 for never smokers) to 0.908 (survival to age 75 for never smokers). The largest differences between current and never smokers' objective probabilities are for survival, and onset of chronic lung disease, difficulty climbing, carrying weight, and walking several blocks.

Except for feet swelling, there are statistically significant differences in subjective probabilities between current and never smokers on survival to age 75 and in onset of morbidity and physical limitations. Differences in subjective probabilities between former and never smokers are much less likely to be statistically significant than are those between current and never smokers. The differences in subjective probabilities between current and never smokers strongly suggest that smokers are aware of their compromised health and of their increased chances of a number of adverse health outcomes. Although statistical significance is often lacking for the differences in subjective probabilities between former and never smokers, the patterns of subjective probabilities

[13] These hazards also increase with age. For at least two reasons, one would expect the hazard to decrease with age. Not only is the person closer to age 75 but is also a "survivor." However, the hazard of dying increases substantially as people get older. Furthermore, age in our specification also proxies for health effects not captured by the other health regressors.

are in the same directions as their objective counterparts. We analyze the differences between subjective and objective probabilities next.

Tables 5 and 6 present our results on the accuracy of risk perceptions. In Specification I, we only include smoking status; Specification II adds longevity expectations (p_i), and Specification III adds quitting expectations (q_i) for current smokers; q_i is set to zero for both former and never smokers.

In Specification I, never smokers slightly underpredict their objective probability of survival to age 75 by just 0.039 (see intercept, Table 5). The mean objective probability of survival to age 75 for never smokers is 0.908 (see Table 4). In this sense, never smokers' assessments appear quite accurate on average. The mean age of sample persons in the SOS is 59.6, implying a 15-year projection period on average. Given the length of the time horizon, this would seem to be relatively difficult prediction to make and thus mature individuals are, on average, remarkably accurate.

Current smokers slightly overpredict the probability of living to 75. The subjective exceeds the objective probability by 0.015 (i.e., 0.054-0.039). The mean objective probability for current smokers is 0.698, implying a two percent overprediction. By this metric, they are even more accurate than never smokers in predicting their survival to age 75. Former smokers are not statistically different from never smokers in their beliefs about survival. The estimate of ξ which gives the effect of q_i on survival in Specification III is not statistically different from zero, suggesting that quitting expectations do not affect survival expectations.[14]

In contrast to survival, individuals appear to overpredict the onset of serious chronic diseases by age 75, irrespective of smoking status. For never smokers the levels of overprediction range from 0.097 for chronic lung disease to 0.203 for stroke in Specification I. For never smokers, the mean objective probability of contracting lung disease by age 75 is 0.046; for stroke, the mean objective probability is 0.056 and for heart problems it is 0.188, implying that overestimates are

[14] Specification II is not estimated because the subjective probability of survival to age 75 is used to construct the dependent variable.

quite large. Current smokers are *more* pessimistic about onset of chronic lung disease and stroke than are never smokers, especially for chronic lung disease. Former smokers are not statistically different from never smokers for all three diseases.

In Specification II, persons who expect to live longer have a smaller difference between the subjective and objective probabilities, perhaps because they attach lower probabilities to onset of stroke and heart problems. This implies that they are less pessimistic about acquiring these diseases, presumably because they think they are healthier in the current period. However, longevity expectations do not affect risk perceptions for chronic lung disease.

In Specification III, which adds quitting expectations, the results are unchanged. Quitting intentions raise the subjective probability of onset of heart problems. It may be that smokers who expect their health to worsen are more likely to quit.

Compared to assessments of the risk of chronic lung disease, stroke, and heart problems, people tend to be more accurate in assessing the probabilities of onset of difficulties in walking several blocks, climbing stairs, and carrying weights within 6 years of the survey (Table 6). One reason may be that the time horizon considered for the 3 diseases is 15 years on average while the time horizon for these physical activities in Table 6 is 6 years.

In Specification I, none of the intercepts is statistically significant. For carrying weights, current smokers tend to underpredict the probabilities of onset of difficulty. There are no statistical differences between former and never smokers.

Adding the subjective probability of living to 75 in Specification II does not change these conclusions. In contrast to onset of stroke and heart problems, longevity expectations do not affect risk perceptions for the onset of these physical difficulties. In Specification III, including expectations about quitting smoking also have no effects on perceptions of these risks and leave the conclusions from Specification I unchanged.

The last four columns of Table 6 contain results for onset of physical conditions, feet swelling, shortness of breath, fatigue, and wheezing. Current, former, and never smokers are

equally pessimistic about onset of these physical conditions. In Specification I, never smokers are pessimistic for shortness of breath, fatigue, and wheezing, with intercepts of 0.048, 0.031, and 0.039, respectively. This compares to mean objective probabilities of onset of 0.069 for shortness of breath, 0.095 for fatigue, and wheezing of 0.049. There is no statistical difference between current and never and former and never smokers.

For feet swelling, shortness of breath and fatigue, people who expect to live longer have smaller differences between the subjective and objective probabilities (Specification II). As for stroke and heart problems, people who expect to live longer, may be healthier in ways not captured by HRS health questions and hence less pessimistic about acquiring these physical conditions. Beliefs about quitting in two years from the SOS interviews (Specification III) have no statistically significant effect on the differences between the subjective and objective probabilities except for feet swelling.

Even though smokers and others are clearly not systematically optimistic about their future health and longevity, they may be prone to large errors in forecasting future events in either direction. To learn whether or not current smokers are unusually misinformed, we compute proportions of persons with absolute differences between subjective and objective probabilities greater than 0.95, 0.85, and 0.75, respectively. Current smokers do not differ significantly from never smokers in being unusually misinformed about survival. However, current smokers are more likely to be unusually misinformed about onset of the chronic conditions and the disabilities than are never smokers.

IV. B. Alternative Methods of Belief Elicitation

Social scientists have pointed to several concerns about the difficulty of measuring subjective probabilities. Some have questioned whether individuals are even be able to form probabilities (see, e.g., Hanson and Kysar 2001), while others point to sensitivity of responses to the framing of the questions (e.g., Tversky and Kahneman 1981). There is evidence that people overpredict very low and underpredict very high probability events (e.g., Lichtenstein, et al. 1987 and Slovic 2000). In addition, there is evidence that people overweight changes in probabilities

when they change from values near zero or one. Thus a 0.1 change in the probability of an outcome is perceived to be greater when the starting point is near zero or one than when the starting point is in the mid-range of probabilities (see e.g., Tversky and Fox (2000, p. 95)).

Our finding that peoples' beliefs about longevity in our sample of 50-70 year olds are accurate on average (Table 5) is reassuring in that it suggests that SOS subjects could form risk perceptions about distant future events, and that the question and scaling reasonably reflect those perceptions. Furthermore, other research using responses to questions about subjective probabilities of survival and future health outcomes asked by the HRS (and used for additional health outcomes in SOS) has found that these responses do reflect information about health events and behaviors individuals have about themselves (e.g., Smith, Taylor, and Sloan 2001, Hurd and McGarry 2002).[15]

Some previous research has taken a different approach to eliciting beliefs. The analysis in Viscusi (1990) was based on responses to the following question: "Among 100 cigarette smokers, how many of them do you think will get lung cancer because they smoke?" Using data on the identical question included in the SOS, we replicate Viscusi's finding. The mean response to this question from SOS, 43, is identical to that reported by Viscusi, 43.[16] Conditional on being a current smoker, the mean response from SOS to this question is 40 while Viscusi reported 37.

The surveys were conducted two decades apart, and even though adverse health effects of smoking were known and well publicized when Viscusi's data were collected, there has been considerable additional publicity about the threat of smoking to health since then. For this reason, the mean response from SOS might have been higher. However, at the same time, individuals in the 50-70 age group have probably been exposed to the health experiences of persons who smoke

[15] Smith, Taylor, and Sloan (2001) using the HRS found that the HRS question on subjective beliefs about survival to age 75 predicts future mortality among the same individuals. Using several waves of the HRS, the subjective probability of surviving to age 75 declined as survey respondents approached their own deaths. Also using HRS, Hurd and McGarry (2002) found that subjective survival probabilities predict survival. Persons who survived in the panel reported about a 50 percent higher probability at baseline of surviving to age 75 than did those who subsequently died. Results from both studies indicate that responses to the HRS longevity expectations questions provide reasonably good predictions of future mortality.

[16] Based on data from reports by the U.S. Surgeon General, Viscusi took the objective risk of lung cancer among smokers to be between 0.05 and 0.10 in the early 1980s, which led him to conclude that individuals overestimate the risk of lung cancer.

for more years than a general population has. This factor would tend to improve the accuracy of risk perceptions of onset of smoking-related diseases that occur in late adulthood. Since people tend to overestimate the probability of smokers getting smoking-related diseases, greater accuracy means a reduction in the beliefs about the onset of such diseases relative to other adults.

Previous research has found that, under identical exposure to risk, people often perceive themselves to be at a lower risk than others (Weinstein 1989), i.e., are more optimistic about their chances than those of others. This would imply that posing questions about risk in the second person should yield different risk assessments than obtained from responses to questions posed in the third person. Our results appear fairly robust to variation in the method of elicitation; comparing average assessments of the risk of getting lung cancer among smokers, the responses are remarkably similar when the question is asked in the third person (as in Viscusi 1990) and when the question takes the usual SOS format phrased in the second person on a 1-10 scale. The mean probability of getting lung cancer from the question used by Viscusi (1990) repeated with identical wording in the SOS is 0.40. The probability derived from the 1-10 question asked in the second person of current smokers is 0.41. Thus similar to Viscusi we find that relative to estimates of incidence of lung cancer in the epidemiological literature, beliefs about smokers getting lung cancer are substantially overestimated by smokers and non-smokers. We cannot compare subjective with objective probabilities of getting lung cancer in our empirical analysis since the HRS does not ask about specific types of cancer, except in its baseline interviews, so a hazard cannot be estimated.

IV. C. Beliefs about Quitting

The SOS elicited information from current smokers about how many cigarettes they thought they would be smoking two years from the survey date. The question was "Roughly how many cigarettes do you expect to smoke per day two years from now?" Those smokers who said that would be smoking zero cigarettes were classified as self-assessed quitters. The self-assessed probability of quitting was compared with actual quit rates from the HRS. The mean subjective

probability from SOS of having quit in two years was 0.45. In the HRS by comparison, the corresponding, objective two-year quit rate was 0.16.[17] However, our regression results indicate that smokers do not condition on quitting in forming their beliefs about survival and onset of diseases and disabilities. In this sense, smokers are misinformed. One interpretation of the lack of effect of the self-assessed chances of quitting as implying that people tend to say that they will quit, but they do not really believe it themselves.[18]

V. Conclusions

Taken together, our results indicate that mature individuals irrespective of their smoking status are quite accurate on average about the probability of surviving to age 75 and the onset of physical difficulties (i.e., walking several blocks, climbing stairs, and carrying a 10-pound weight). In assessing the probability of onset of three chronic diseases (chronic lung disease, stroke, heart problems), individuals in the 50 to 70 age group are highly pessimistic regardless of smoking status. These findings are robust to accounting for expectations about future quitting and longevity. For the condition which is most closely linked to smoking, chronic lung disease, current smokers are even more pessimistic than others about the probability of onset of these diseases. Since the estimated hazards used to predict the objective probabilities account for smoking status, this extra degree of pessimism may reflect the many messages smokers receive about specific health risks of smoking. Similar pessimism, irrespective of smoking status, is found for onset of physical conditions (i.e., feet swelling, shortness of breath, fatigue, and wheezing).

Viewed another way, people in their fifties and sixties appear to take account of objective risk factors when forming beliefs about longevity and future health events. If anything, such persons tend to be too pessimistic about their future health, contradicting the notion that people underestimate their risks. This is true for mature smokers as well. The lack of association between

[17] We also estimated probits which included age, marital status, gender, educational attainment, race, and self-rated health status as regressors to correct for differences in sample characteristics between the SOS and HRS. The results were almost identical to those reported.

[18] A referee pointed out that this result could also mean that smokers reason that quitting in the sixth or seventh decade of life would not improve one's health. We acknowledge that this is a possibility. However, that quitting at 65 yields an appreciable increase in life expectancy on average (Taylor, Hasselblad, Hensley et al. 2002) may not have been known by many experts either prior to recent publication of study findings on this issue.

smoking and optimism in risk perceptions, casts doubt on the idea that continued smoking can be attributed to a rosy view of future risks.

SOS respondents who smoked had done so for many decades, and they presumably had a greater opportunity to observe that specific adverse outcomes are more likely to occur among persons who smoke regularly than younger persons do. We do not have information on these individuals' views about adverse health effects of smoking when they were teenagers and deciding whether or not to become smokers, and in any case, it is possible that even young persons in recent years have become more knowledgeable about the adverse health effects than persons who were born between 1934 and 1955.

Our results that people tend to be pessimistic about onset of some diseases and disabilities, and realistic about longevity is also consistent with the notion that individuals are optimistic about future technological improvements in medical care that will allow them to survive health threats (e.g., Cutler 2004). The result that individuals are pessimistic about onset of bad health contrasts with previous research finding people tend to be too optimistic about risks that can attributed to personal choices (Weinstein 1989). In contrast to our study, this body of research asks respondents to compare their personal risk to the risks of others. Our analysis is not based on such comparative risk judgments which might account for the different conclusion, and may have implications for survey methods.

Even if people are fully informed about the health consequences conditional on smoking, they may be misinformed about the addictiveness of tobacco, especially at the time the smoking initiation decision is made. Even mature smokers appear to be overly optimistic about their probability of quitting.[19] Thus, risk perceptions appear to be unbiased in some aspects of the health domain, but our finding may not generalize to other consequences of personal decisions.

Finally, although statements by the tobacco companies may have been misleading, we find no evidence that mature consumers are overly optimistic about the health effects of smoking.

[19] Hanson and Kysar (1999) discuss these two forms of potential biases separately.

Given the age of our sample and the year in which the survey on subjective beliefs was conducted, we cannot say anything about initiation of smoking. Our results do not necessarily imply that consumers may never have been misled by product claims of tobacco companies or that young persons are currently misled as much as they were in the past. Extending this analysis to younger persons would be a very useful next step.

References

Ayanian, J. Z. and Cleary, P. D., "Perceived Risks of Heart Disease and Cancer Among Cigarette Smokers," *Journal of the American Medical Association*, 281(11), 1999, pp. 1019-1021.

Börsch-Supan, A., "Comment on 'Subjective Survival Curves and Life Cycle Behavior'," in Wise, D. A., *Inquiries in the Economics of Aging*. Chicago: University of Chicago Press, 1998, pp. 306-309.

Cutler, D.M., Gruber, J., Hartman, R.S., Landrum, M.B., Newhouse, J.P., and Rosenthal, M.B., "The Economic Impacts of the Tobacco Settlement," *Journal of Policy Analysis and Management*, 21. 2002, pp. 1-19.

Cutler, D., *Your Money or Your Life: Strong Medicine for America's Health Care System*. New York: Oxford University Press, 2004

Fiore, M., Novotny, T., Pierce, J., Giovino, G., Hatziandreu, E., Newcomb, P. Surawicz, T., and Davis, R. "Methods Used to Quit Smoking in the United States: Do Cessation Programs Help?" *Journal of American Medical Association* 263 (20), 1990, pp. 2760 – 2765.

Gan, L., Hurd, MD., and McFadden, D., "Individual Subjective Survival Curves," *NBER Working Paper 9480*, 2003.

Glantz, S.A., Slade, J., Biro, L.A., Hanauer, P., and Barnes, D.E., *The Cigarette Papers*. Berkeley: University of California Press, 1996.

Hamermesh, D., "Expectations, Life Expectancy, and Economic Behavior," *Quarterly Journal of Economics*, 95, 1985, pp. 537-557.

Hanson, J.D. and Kysar, D.A., "Taking Behavioralism Seriously: Some Evidence on Market Manipulation," *Harvard Law Review*, 112(7), 1999, pp. 1420-572.

Hanson, J.D. and Kysar, D.A., "The Joint Failure of Economic Theory and Legal Regulation," in Slovic, P, ed., *Smoking: Risk, Perception, & Policy*. Thousand Oaks, CA: Sage Publications, 2001, pp. 229-276.

Hanson, J.D. and Logue, K.D., "The Cost of Cigarettes: The Economic Case for Ex Post Incentive-Based Regulation," *Yale Law Journal* 107(5), 1998, pp. 1163-1361.

Hilts, P.J., *Smoke Screen: The Truth Behind the Tobacco Industry Cover-Up*. Reading, MA: Addison-Wesley, 1996.

Hughes, J.R., Gulliver S.B., Fenwick J.W., Valliere, W.A., Cruser, K., Pepper S., Shea P., Solomon L.J., and Flynn B. S. "Smoking Cessation among Self-Quitters," *Health Psychology* 11, 1992, pp. 1331 – 1334.

Hurd, M. D., and McGarry, K., "The Predictive Validity of Subjective Probabilities of Survival," *The Economic Journal,* 112, October 2002, pp. 966-985.

Kessler, D., *A Question of Intent: A Great American Battle with a Deadly Industry*. New York: PublicAffairs, 2001.

Khwaja, A., Sloan, F.A., and Chung S. "The Relationship Between Individual Expectations and Behaviors: Evidence on Mortality Expectations and Smoking Decisions," *Journal of Risk and Uncertainty*, 35(2), 2007, pp. 179-201.

Kluger, R., *Ashes to Ashes: American's Hundred-Year Cigarette War, the Public Health, and the Unabashed Triumph of Philip Morris*. New York: Alfred A. Knopf, 1996.

Kedzi, G. and Willis, R.J., "Who Becomes a Stockholder? Expectations, Subjective Uncertainty, and Asset Allocation," paper prepared for the Fifth Annual Conference on Retirement Research Consortium 'Securing Retirement Income for Tomorrow's Retirees,' May 15-16, 2003, Washington, D.C.

Lancaster, T., "Econometric Methods for the Duration of Unemployment," *Econometrica,* 47(4), July 1979, pp. 939-956.

Lancaster, T., *The Econometric Analysis of Transition Data*. New York: Cambridge University Press, 1990.

Leventhal, H., Glynn, K., and Fleming, R., "Is the Smoking Decision an 'Informed Choice'? Effect of Smoking Risk Factors on Smoking Beliefs," *Journal of the American Medical Association*, 257(24), 1987, pp. 3373-3376.

Lichtenstein, S., Slovic, P., Fischoff, B, Layman, M., and Combs, B., "Judged Frequency of Lethal Events," *Journal of Experimental Psychology: Human Learning and Memory*, 4, 1987, pp. 55-78.

Lillard, L.A., and Willis, R.J., "Cognition and Wealth: The Importance of Probabilistic Thinking," paper prepared for the Third Annual Conference on Retirement Research Consortium 'Making Hard Choices About Retirement,' May 17-18, 2001, Washington, D.C.

Manski, C. F., "Measuring Expectations," *Econometrica*, 72(5), September 2004, pp. 1329-1376.

McKenna, F.P., Warburton, D.M., and Winwood, M., "Exploring the Limits of Optimism: The Case of Smokers' Decision Making," *British Journal of Psychology*, 84, 1993, pp. 389-394.

Schoenbaum, M. "Do Smokers Understand the Mortality Effects of Smoking?" Evidence from the Health and Retirement Survey," *American Journal of Public Health*, 87(5), May 1997, pp. 755-758.

Sloan, F.A., Carlisle, E.S., Ratliff, J.R., and Trogdon, J., "Determinants of States' Allocations of the Master Settlement Agreement Payments," *Journal of Health Politics, Policy, and Law*, 30(4), August 2005, pp. 643-86.

Sloan, F.A., Smith, V.K., and Taylor, D.H., Jr., *The Smoking Puzzle: Information, Risk Perception, and Choice*. Cambridge, MA: Harvard University Press, 2003.

Sloan, F.A., Ostermann, J., Picone G., Conover C., and Taylor, D.H., *The Price of Smoking*. Cambridge MA: MIT Press, 2004.

Slovic, P., "Perception of Risk," in P Slovic, *The Perception of Risk.* London. Earthscan Publications, 2000, pp. 220-31.

Smith, V. K., Taylor, D. H., and Sloan, F. A., "Longevity Expectations and Death: Can People Predict Their Own Demise?" *American Economic Review,* 91(4), September 2001, pp. 1126-1134.

Smith, V. K., Taylor, D. H., Sloan, F. A., Johnson, F. R., and Desvouges, W. H., "Do Smokers Respond to Health Shocks?" *Review of Economics and Statistics,* 83(4), 2001, pp. 675-687.

Summers, L.H., "Some Simple Economics of Mandated Benefits," *American Economic Review,* 79(2), May 1989, pp.177-83.

Taylor, D.H., Jr., Hasselblad, V, Hensley, S.J., Thun, M.J., Sloan, F.A., "Benefits of Smoking Cessation for Longevity," *American Journal of Public Health*, 92, 2002, pp. 990-996.

Tversky, A., and Fox, C.R., "Weighing Risk and Uncertainty," in D. Kahneman and A. Tversky, Choices, Values, and Frames. Cambridge University Press, 2000, pp. 93 – 117.

Tversky, A. and Kahneman, D., "The Framing of Decisions and the Psychology of Choice," *Science*, 211, 1981, pp. 453-458.

Van der Klaauw, W. and Wolpin, K., "Social Security and the Retirement and Savings Behavior of Low Income Households," mimeo, University of North Carolina-Chapel Hill and University of Pennsylvania, 2005.

Viscusi, W. K. "Do Smokers Underestimate Risks?" *Journal of Political Economy,* 98(6), December 1990, pp. 1253-1269.

Viscusi, W.K., :Age Variations in Risk Perceptions and Smoking Decisions," *The Review of Economics and Statistics*, 73(4), November 1991, pp. 577-588.

Viscusi, W.K., *Smoking: Making the Risky Decision*. New York: Oxford University Press, 1992.

Viscusi, W.K., *Smoke-Filled Rooms: A Postmortem on the Tobacco Deal.* Chicago: University of Chicago Press, 2002.

Viscusi, W.K. and Hakes, J.K., "Risk Ratings That Do Not Measure Probabilities," *Journal of Risk Research*, 6(1), 2003, pp. 23-43.

Weinstein, N.D., "Unrealistic Optimism about Susceptibility to Health Problems: Conclusions from a Community-wide Sample," *Journal of Behavioral Medicine*, 10(5), 1986, pp. 481-500.

Weinstein, N.D., "Optimistic Biases about Personal Risks," *Science*, 246, 1989, pp. 1232-1233.

Table 1. Summary Statistics

	HRS (1992)		HRS (2004)		SOS	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
A. Individual characteristics†						
Subjective probability of living to age 75	0.642	0.298			0.768	0.283
Smoking status						
Current smoker	0.273	0.445	0.146	0.353	0.380	0.486
Former smoker	0.361	0.480	0.464	0.499	0.388	0.488
Never smoker	0.366	0.482	0.390	0.488	0.232	0.423
Self-rated Health						
Excellent	0.217	0.412	0.106	0.308	0.172	0.378
Very good	0.278	0.448	0.283	0.451	0.299	0.458
Good	0.280	0.449	0.320	0.466	0.286	0.453
Fair	0.143	0.351	0.204	0.403	0.149	0.357
Poor	0.082	0.274	0.087	0.282	0.090	0.287
Quitting smoking in 2 years					0.447	0.498
Demographic characteristics						
Age	55.774	3.391	67.591	3.360	59.620	5.823
Non-Hispanic white	0.793	0.405	0.807	0.395	0.863	0.344
Male	0.467	0.499	0.444	0.497	0.357	0.479
Married	0.748	0.434	0.672	0.469	0.587	0.493
Years of education completed	12.035	3.229	12.194	3.199	14.334	2.583
Number of observations	10,439		7,282		663	
B. Incidence of mortality and morbidity outcomes††						
Deaths	0.142	0.349				
Lung disease	0.072	0.258			0.212	0.409
Heart problems	0.153	0.360			0.192	0.394
Stroke	0.053	0.224			0.058	0.233
Difficulty walking	0.294	0.455			0.321	0.467
Difficulty climbing stairs	0.398	0.490			0.422	0.494
Difficulty carrying weights	0.266	0.442			0.263	0.441
C. Incidence of physical conditions†††						
Feet swelling	0.163	0.370			0.135	0.342
Shortness of breath	0.146	0.353			0.245	0.431
Feeling fatigued	0.163	0.369			0.213	0.410
Wheezing	0.130	0.336			0.203	0.402

† Baseline characteristics for HRS data from wave 1, 1992.
†† HRS data is from Waves 1 – 7, 1992 – 2004, conditional on not having the condition in 1992.
††† HRS data is from Waves 3 – 7, 1996 – 2004, conditional on not having the condition in 1996.