# Predicting Cancer-Prevention Behavior: Disentangling the Effects of Utility Concavity and Risk Perceptions

Abstract

Studies such as Barsky et al. (1997), Anderson and Mellor (2008), and Dohmen et al. (2011) have concluded that risk aversion affects people's choices to take or avoid health risks. Another strand of this literature finds that these decisions are often rooted in misperceptions about health risks (Viscusi 1990, 1991; Lundborg and Lindgren 2002, 2004; Khwaja, Sloan and Chung 2007). We argue that if the health preference function follows a Rank-Dependent or Cumulative Prospect preferences, that both of these factors need to be considered. We use a survey of 474 men and women to investigate the influence of utility risk aversion, probability weighting represented by optimism, and cognitive ability on the choice to engage in behaviors that either increase or mitigate cancer risk. We measure optimism in two dimensions: baseline optimists are those who inaccurately believe their cancer risk to be below its expert-assessed level, while control optimists are those who believe they can reduce their risk of cancer (by changing their lifestyle choices) to a greater extent than is actually the case. Our results indicate that baseline optimism is significantly and positively correlated with subject's tendencies to engage in cancer risk-reducing behaviors, and negatively correlated with risky behaviors. Subjects' control misperceptions also appear to play a role in their tendency to engage in risky and prevention behaviors. When controlling for both of these types of risk misperception, utility risk aversion plays a much smaller role in determining health behaviors than found in past studies. This supports previous findings that health preferences are characterized by significant probability weighting.

JEL Codes: D80, D83, I12

Key Words: risk perception, optimism, risk aversion, cumulative prospect theory, probability weighting

# I. INTRODUCTION

Experimental and survey research spanning the last two decades has concluded that people who are more risk tolerant are more likely to engage in risky health activities such as smoking, heavy alcohol consumption, and are more likely to be obese (Barsky et al. 1997, Dave and Saffer 2007, Anderson and Mellor 2008, Dohmen et al. 2011). They base their conclusions on measures of curvature in the utility function derived from choices among financial gambles presented by the researchers. These studies either implicitly or explicitly assume that subjects are expected utility (EU) maximizers and that curvature in the risk-preference function arises solely from the utility function.

Increasingly, experiments have concluded that for most subjects, there is an additional source of curvature in the preference function arising from the tendency of subjects to overemphasize some probabilities in their decision making relative to EU theory. Cumulative prospect (CP) theory and rank dependent utility (RDU) theory allow probability weights to replace probabilities in the EU function and often better explain choices over risky financial prospects (Tversky and Kahneman 1992, Quiggin 1993, Rabin 1998, Starmer 2000, Bruhin, Fehr-Duda and Epper 2010, and Conte, Hey and Moffatt 2011). This research has been extended to the life-duration domain where researchers have similarly concluded that preferences are best modeled using CP or RDU, thereby allowing for an additional source of curvature in the health-preference function (Wakker and Deneffe 1996; Bleichrodt and Pinto 2000; Riddel and Kolstoe 2013; Conte, Moffatt, and Riddel 2015).

The bulk of experimental studies have documented a probability weighting function that takes on an inverse-s shape in the domain of financial risk. Overweighting of low probabilities means that subjects become more likely to engage in gambles over low-probability financial gains and more averse to low probability financial losses (Tversky and Kahnmean 1992, Tversky and Fox 1994, Wu and Gonzalez 1996, Prelec 1998). One interpretation of probability weighting is that it reflects optimism about having the best outcome occur and a corresponding pessimism about having the worst outcome realized (Abdellaoui, l'Haridon, and Zank 2010).

Studies such as Wakker and Deneffe (1996), Bleichrodt and Pinto (2000), and Riddel and Kolstoe (2013) offer support for the notion that probability weighting also affects choices in life-duration domain with a similar outcome: subjects tend to be overly attracted to low-probability health benefits and excessively averse to low probability disease outcomes. In the domain of cancer risk, this can be interpreted as optimism concerning a subject's risk of contracting cancer and/or optimism about the efficacy of prevention and risk-avoidance activities.[1]

While there are numerous studies that link risk perceptions with risky behavior (Viscusi 1990, 1991; Viscusi et al. 2000; Lundborg and Lingren 2002, 2004; Khwaja, Sloan, and Chung 2007), it is notable that none these control for utility risk aversion. Similarly, studies that control for utility risk aversion fail to control for probability weighting related to risk misperceptions.

The purpose of this paper is to examine the relationship between optimism and choices to engage in behaviors that either increase or decrease the risk of cancer, controlling for utility curvature and other factors such as cognitive ability and demographic variables that may affect health choices. The data is collected using a national survey of 474 men and women, aged 18 and over, in the US. We focus on four cancers: bladder and colon cancer in men and women as well as prostate cancer in men and breast cancer in women. On the risk-increasing side, we explore how subjective optimism concerning the risk of contracting one of these cancers

[1] A very interesting paper by Attema, Brouwer, and l'Haridon (2013) estimate CP preferences in the domain of life duration. They find significant deviations from EU preferences. However, they use only one probability, so it is difficult from their results to make any inference about probability weighting.

influences the decision to engage in activities that increase cancer risk such as smoking, eating excessive amounts of red meat, heavy alcohol consumption and being medically obese. On the cancer prevention side, we model regular physical activity, taking a daily aspirin, and taking a daily multivitamin. We elicit a coefficient of life-duration utility risk aversion from each subject using a sequential multiple price list (sMPL) auction (Holt and Laury 2002).

One of the challenges of estimating subjective optimism is having a solid basis for objective risks associated with different risk-increasing or risk-reducing behaviors. One of the strengths of the current study is that our measures of cancer-risk optimism rest on comparing subjective assessments of the relative risk of different health-related behaviors to objective estimates contained in the Harvard Cancer Risk Index (HCRI) (Colditz et al. 2000). The HCRI was established by a working group of "epidemiologists, clinical oncologists, and other Harvard faculty with quantitative expertise focused on cancer and risk assessment" at the Harvard Center for Cancer Prevention (Colditz et al. 2000). The HCRI establishes generally agreed on quantitative relative risk (RR) factors for demographic and behavioral characteristics that bear on the incidence of a given cancer.

Following Spinnewijn (2013), we construct our models based on two sources of optimism. Baseline optimists believe their cancer risk is lower than what an expert would assess their risk to be given their current and past behavior and health history. Control optimists believe that efforts to reduce cancer risk, including engaging in prevention measures and *avoiding risky activities*, are more effective than they what the current state of scientific knowledge would indicate.[2] We measure baseline and control optimism by comparing subjective

[2] It is important to note that our definition of control optimism differs from that used in the psychology literature. Control optimism in the current context arises from two sources. First, a subject is control optimistic if they believe that prevention efforts are more effective in reducing

assessments of cancer risk and efficacy of prevention with objective estimates of these same risks taken from the HCRI.

We find that optimism – in particular baseline optimism – plays an important role in people's decision to engage in cancer-prevention behavior and to avoid activities that increase cancer risks. In general, baseline optimists expend less effort to reduce their cancer risk than do their more pessimistic counterparts. Such persons are more likely to smoke, eat excessive amounts of red meat, and be medically obese. They are also less likely to engage in prevention behavior such as regular physical exercise and taking a daily aspirin or multivitamin.

The effects of control optimism on health behaviors appear to be less comprehensive. Control optimists engage in a generally higher level of prevention behavior and are less likely to smoke and drink excessively, but more likely to be obese. Control optimists are no more likely than their pessimist counterparts to engage in the specific prevention behaviors studied.

The model results also indicate that utility concavity is far less important in explaining risky behavior than baseline and control optimism. We show that while models that exclude optimism and cognitive ability frequently indicate a significant correlation between risky behavior and utility risk aversion, the relationship disappears when these other potentially important variables are controlled for. This result calls into question previous research linking utility risk aversion with risky health behaviors. We conclude that empirical models that fail to control for subjective beliefs about disease risk may erroneously conclude that utility risk aversion motivates health choices, when instead it is optimism that is actually driving the result.

---

cancer risk than what is reported in the HCRI. Second, a subject is control optimistic if they believe that *avoiding* risky behaviors such as smoking is *more effective* in reducing risk than what is stated in the HCRI. So for example, if a subject overestimates the risk of smoking they are classed as control optimistic with respect to smoking.

## II. LITERATURE REVIEW

Two strands of literature have developed that seek to explain why people either engage in risky health or prevention behaviors. Early work argued that people were frequently misinformed about health risk, so that risk perceptions, rather than objective assessments of risk, influenced behavior. The later strand of the literature correlates risk aversion, as measured by concavity in the utility function, with the choice to engage in risky and health behaviors. We discuss each strand of the literature in turn.

### *2.1 Risk Perceptions and Health Behaviors*

While the above articles examine utility risk aversion and behavior, numerous studies explore the relationships between subjective risk perceptions and behavior. Many of these examine the role risk perception plays in the decision to smoke. Viscusi (1990) uses a large survey of US subjects aged 16 and over to examine subjects' risk perceptions of the likelihood of contracting lung cancer and their smoking status. He finds that smokers and nonsmokers alike tend to overestimate the risk, with the overestimation somewhat higher in nonsmokers than smokers. Moreover, the higher the magnitude of overestimation i.e. the more control optimistic the subject is about lung-cancer risk, the less likely the subject is to smoke. Using the same dataset, Viscusi (1991) concludes that younger cohorts believe smoking risks are higher and are therefore less likely to smoke. Viscusi et al. (2000) comes to a similar conclusion about Spanish smokers. Like US smokers, they tend to overestimate lung cancer and other disease risks from smoking. They show that the number of cigarettes smoked per day is decreasing in the perceived loss in life expectancy due to smoking. They also find that the probability a subject is a current smoker is decreasing in the perceived lung cancer risk.

Lundborg and Lindgren (2004) use a cross-sectional survey data of Swedish adolescents to analyze risk perceptions and smoking behavior. Like Viscusi (1990) and Viscusi (1991), they find that subjects overestimate the risk of smoking and that the probability a subject smokes is decreasing in the perceive lung cancer risk. Khwaja, Sloan and Chung (2007) use data from the Health and Retirement study to compare subjective beliefs about expected longevity to objectively assessed probabilities. They conclude that, on average, people are fairly accurate about their beliefs, but smokers tend to be relatively optimistic about their likelihood of living another ten years.

Lundborg and Lindgren (2002) use a cross-sectional survey of young Swedes to study the link between risk perceptions of alcohol use and drinking behavior. They find that subjects tend to overestimate the risk of alcoholism, but the overestimation decreases with the age of the subject. The higher the perceived risk of alcohol consumption, the less likely the subject is to consume alcohol.

Others have taken a broader look at risk behavior and perceptions. Weber, Blais and Betz (2002) study the effects of both risk perceptions and risk attitudes on behavior. Roughly 500 US college students were asked to rate the likelihood they will engage in 101 different risky behaviors as well as their perception of the risks and benefits of the different behaviors.[3] The behaviors are categorized as falling within the following domains: financial, health/safety, recreational, ethics, and social. Subjects evaluated the likelihood that they would engage in behaviors within each domain, on a 5-point scale from 1 (extremely unlikely) to 5 (extremely likely). On the risk perception side, subjects were asked to grade the riskiness of the behaviors

[3] For example, in the health domain, subjects were queried about the perceived risk and likelihood of engaging in risky health behaviors such as drinking and driving.

on a 5-point scale from 1 (not at all risky) to 5 (extremely risky.) Subjects were also faced with real-stakes financial gambles, and their bets were recorded. Weber, Blais, and Betz (2002) find that risk perceptions and risky behaviors are negatively correlated, while perceived benefits are positively correlated with behavior. They conclude that risk preference measures must carefully control for perceptions of both probabilities and outcomes. If these are neglected, then risk preference estimates will likely be biased, as they depend not only on intrinsic risk preferences, but also on risk perception.

Carman and Kooreman (2014) study the link between subjective risk perceptions and engaging in prevention behaviors such as flu shots, mammograms, and taking aspirin to guard against heart disease. They compare subjective risks to individually predicted epidemiological risks. They find that subjects do not have very accurate risk beliefs, but beliefs have a significant influence on the choice to engage in prevention behaviors.

Winter and Wuppermann (2014) study subjective perceptions of obesity-related health risks for a sample of obese adults aged 50 – 62 years in the US. They calculate over and under-perception of risks for several obesity-related diseases by comparing each subject's assessment of their personal risk (on a 100-point scale) to an objective probability the subject will get the given disease. The objective probability measures are derived from a predictive model that is based on data from the Health and Retirement Study. The authors find that subjects tend to overestimate the risk of heart attack, stroke, and chronic lung disease but underestimate other health risks such as arthritis and high blood pressure. Comparing risk over- and under-estimation to those of a sample of normal-weight subjects, they conclude that all subjects tend to underestimate the risk of small probability diseases and overestimate higher-probability diseases. Moreover, this effect is more pronounced in obese subjects.

Two recent studies examine the relationship between risky behavior, cognitive ability, and optimism. Bijwaard, van Kippersluis, and Veenman (2015) investigate the influence of cognitive ability and education on survival of a Dutch cohort born between 1937 and 1941 with observed mortality between ages 55 to 75. They find that survival is increasing in both variables, with much of the survival differences across education cohorts explained by cognitive ability. While the current study looks at preventative and risky behaviors rather than survival, the findings of Bijwaard, van Kippersluis, and Veenman (2015) suggest that it may well be important to control for cognitive ability in models linking utility risk aversion, optimism, and behavior.

Riddel and Hales (2015) use data from a survey of adults in the US to elicit individual perceptions of cancer risk and demand for a hypothetical cancer-insurance policy. They find that demand for cancer insurance does not vary with utility risk aversion and cognitive ability, but is decreasing in baseline optimism. Moreover, they show that risk-reducing effort is decreasing in baseline optimism, increasing in control optimism, but unrelated to utility risk aversion or cognitive ability.

### *2.2 Risk Aversion and Health Behaviors*

Several studies correlate utility risk aversion with a greater propensity to engage in risky health behaviors. An early study by Barsky et al. (1997) uses data from the Health and Retirement Survey to calculate subject-specific coefficients of financial utility risk aversion based on choices over gambles of lifetime income. They measure risk tolerance using responses to a series of questions from the Panel Study of Income Dynamics of the form: would you take a job with 50-50 chance it will double your income and a 50-50 chance it will cut you income by x% where x = 50, 33, and 20. Given the responses, subjects can be categorized as one of four groups ranging from risk averse to risk loving. Assuming constant relative risk aversion, the choices define a range of the coefficient of utility risk aversion for each of the groups. They find that subjects who are more risk tolerant are more likely to smoke, drink alcohol, work in risky jobs, and have no health insurance. These results hold after controlling for demographic variables such as gender, ethnicity, and religious affiliation that may influence behavior.

Dave and Saffer (2007) study the effect of risk tolerance on the demand for alcohol using the same risk preference elicitation approach as Barsky et al (1997). They find that demand for alcohol is increasing in risk tolerance. Thus, they conclude that subjects who are more risk averse are less likely to consume alcohol than their risk-loving counterparts.

In a large experiment, Anderson and Mellor (2008) calculate subject-specific coefficients of financial utility risk aversion using Holt and Laury's (2002) sMPL elicitation scheme. Controlling for demographic characteristics, they find that the utility risk-averse subjects are less likely to engage in unhealthy behaviors such as smoking, heavy alcohol consumption, and being obese and more likely to wear seatbelts. They also show that utility risk aversion is decreasing in the number of risk behaviors engaged in.

Dohmen et al. (2011) analyze subjects' willingness to take risks in general, as well as over five domains including driving automobiles, financial matters, sports and leisure, career, and health based on a large representative survey (n=22,000) of the German population. They measure risk tolerance using two approaches. The first approach asks the subjects to grade their willingness to take risks in each of the domains and then generally on a 10-point scale. The second risk preference measure is determined by choices over a real financial-stakes lottery. They find that self-reported measures of willingness to take risks are highly correlated with utility risk aversion derived from the real-stakes lottery. They further show that smokers are likely to report a higher willingness to take health risks, and a higher willingness to take risks in general, than are nonsmokers.

Viscusi and Hersch (2001) compare compensating wage differentials of smokers and nonsmokers for risky jobs. They find that smokers are more likely to work in risky jobs, but also to accept lower wages than non-smokers. They conclude that smokers' relative preference for risk, together with a higher than average on-the-job injury rate, give rise to the observed wage differential.

At least one study fails to find a significant association between risk tolerance and risky behavior. Picone et al. (2004) finds that the probability of undergoing preventative medical tests such as breast self-exam, mammography, and pap smears are increasing in risk tolerance, but the associations are only marginally statistically significant.

Taken together, these studies appear to indicate that risk aversion influences health behaviors. However the risk aversion elicitation methods they use are only capable of determining risk aversion in the aggregate and cannot decompose that into that arising from the utility function and that arising from optimism or pessimism. Under CP and RDU models, both concavity in the utility function and probabilistic risk aversion will play a role in choice. In what follows, we discuss a survey designed to elicit utility risk aversion, optimism, current and past health behaviors, as well as demographic characteristics of subjects. We then model the choice to engage in risky and prevention behaviors as a function of these variables.

## III. THE SURVEY

We conducted an online survey of 474 men and women aged 18 and over living in the US using Amazon Mechanical Turk (AMT). AMT is a web-based tool for matching workers with employers who need small tasks performed which require human intelligence. Tasks range from completing marketing and academic surveys to interpreting photos for anomalies. Employers post descriptions of the tasks on the web platform along with compensation amounts. Employers are only required to pay for successfully completed tasks. However, employers who refuse to pay when tasks are adequately completed may well lose their ability to recruit subjects.

The ease of use and streamlined process for contacting subjects has made AMT increasingly popular over the past five years with social science and business researchers. Buhrmester, Kwang, and Gosling (2011) conclude that AMT is cost-effective source for high-quality data. They analyze the demographic variation in AMT survey subjects, determining that participants tend to be more diverse than a typical internet sample or a sample based on university students. They also found that the data quality is comparable to conventional internet or telephone surveys based on random sampling.

We described the task to potential respondents as a survey that aimed to elicit behavior and beliefs related to cancer risks. Subjects were given between $2.50 and $4.50 to complete the survey. The questionnaire has six components, described below.

### *3.1 Risk Perception*

In this section, we elicit subjective beliefs about the likelihood of contracting three cancers: breast, bladder, and colon for women and prostate, bladder and colon for men. Subjects were asked to gauge their risk for each of the three cancers relative to someone their age and gender. Allowed responses were assigned both a numerical and textual description. There were eight categories of allowed responses that ranged from 0 (there is no risk of me getting this cancer) to very much above average (risk is 5 times or more than the average) with six other categories in between.

Following that, subjects were asked to rate the increase (decrease) in their risk arising from engaging in risky (preventative) activities. The risk factors were taken directly from Colditz' et al (2000) HCRI and varied with the cancer. For example, risk factors for bladder cancer included smoking and exposure to chemicals at work, whereas preventative activities related to colon cancer included taking an aspirin daily and regular physical activity. The relative risks for each characteristic or behavior were presented as both ranges of relative risk and a qualitative descriptor as follows no risk increase, small risk increase, risk is higher but less than double the average risk, moderate risk increase where the risk is 2 to 4 times the average risk, large risk increase of 4 to 8 times the average risk, and very large risk increase where risk is more than 8 times the average risk.

*3.2 Risk Preference.*

We used a sequential multiple price list auction to elicit a utility risk aversion coefficient based on the Constant Relative Risk Aversion utility function defined over life-duration risk. Details on the exact choice list can be found in Riddel and Kolstoe (2013). Briefly, the following text was presented to the subjects:

*Hypothetical Health Risk: Assume you have been diagnosed with a disease that will certainly be fatal in a year without treatment. There are two treatments, but neither is effective 100% of the time. Assume the costs of the treatment are the same, and neither treatment has side effects.*

The subjects were subsequently presented with a sequence of paired lotteries and asked to choose the one they preferred. For example, the first gamble presented was:

*Treatment A means a 30% chance of 8 more years of life and a 70% chance of 2 more years. Treatment B gives a 90% chance of 1 more year (the treatment fails) and a 10% chance of 13.5 more years.*

In the subsequent gambles, the outcome in treatment B was altered so that E[A]-E[B] declines, eventually becoming negative. The researcher notes where the subject switches from choosing lottery A to lottery B. The switch defines an inequality over the prospects, with later switch points indicating relatively higher levels of utility risk aversion.

Subjects were not compensated for their choices because it is clearly impossible to play out life-duration gambles. We are optimistic that the lack of incentives did not bias the estimates. In a review of 74 studies that analyzed the influence of compensation in experiments, Camerer and Hogarth (1999) concluded that “the effects of incentives are mixed and complicated”. They generally concluded that when financial gambles were considered, the mean response was unaffected, but the variance tended to be higher when incentives were not offered.

### *3.3 Health History and Objective Relative Risks from HCRI*

The survey elicited a detailed health history related to the three cancers in question. Questions included their family history of the cancers, risky or preventative behaviors such as smoking, exercise, vitamin use, dietary habits, chemical exposure, and alcohol use. In the current study, we are interested in a subset of the behaviors including the following preventative behaviors 1) taking a daily multivitamin 2) taking a daily aspirin and 3) regular exercise as defined at 30 minutes most days or 3 or more hours of exercise per week. Of the risky behaviors, we evaluate 1) being a current smoker 2) being an ever smoker described as having smoked at least 100 cigarettes in the subject’s lifetime 3) eating more than three servings of red meat per week, 4) having a BMI of greater than 30, and 5) drinking excessive amounts of alcohol defined as more than two drinks daily.

The HCRI is constructed using a series of multiplicative risk factors, such that when multiplied together one can objectively estimate one’s overall risk of a given cancer, relative to those of the same age and gender. For example, a relative risk (RR) factor of 1.3 for a given risky activity (or demographic factor) implies that those who regularly engage in this activity (or belong to the demographic) are 30 percent more likely to get the particular cancer, ceteris paribus, than the US population average for those of the same gender and age. Similarly, an RR factor of 0.8 for a healthy behavior (or demographic factor) implies that those who regularly engage in this activity (or belong to the demographic) have a cancer risk that is 20 percent lower than other Americans of their same gender and age. By multiplying all of the RR risk behavioral factors (as well as certain demographic factors that the HCRI assesses also affect cancer risk) that apply to each subject, one arrives at a comprehensive relative risk factor for the individual in question, again compared with the average US person of the same age and gender. This is described in more detail in Section 4.1.

### *3.4 Cognitive Ability*

Subjects were given a series of 7 questions commonly asked in the Wonderlic cognitive ability test. Subjects scored a point for each correct answer. Their total cognitive ability score was the sum of the individual scores.

### *3.5 Demographic Variables*

Subjects were queried about their age, income, insurance status, education level, and marital status.

Table 1 gives descriptive statistics for cognitive ability (mean score out of possible 7 correct), risk aversion (mean CRRA coefficient) and the demographic variables. The mean risk aversion coefficient is 0.685, indicating that the average subject is modestly life-duration risk averse. This value is similar to that found in other studies (Riddel and Kolstoe 2013). For the most part, the demographics compare reasonably well to the US population. Approximately 45 percent of the subjects are male compared to the US population percentage of 49 percent. The median income of the sample of $45,839 is somewhat less than the national median of $51,939. The survey subjects are relatively well educated, with over 43 percent having a Bachelor's degree or higher compared to the US percent of just under 29 percent. Sample subjects are less likely to be married: 36 percent married for the sample compared to 51 percent for the US population. They are slightly less likely to have health insurance than the general US population.

Table 1. Summary Statistics for Demographic, Risk Aversion, and Cognitive Ability

| | Sample | | US Census (2013) |
|---|---|---|---|
| Variable | mean or median | std. dev. | mean or median |
| Risk Averse (mean CRRA coefficent) | 0.685 | 0.320 | ---- |
| Cognitive Ability (mean out of 7) | 4.814 | 1.227 | ---- |
| Male (percent) | 45.600 | 49.850 | 49.200 |
| Age (median) | 32.000 | 10.903 | 36.800 |
| Income (median $thous) | 45.839 | 31.019 | 51.939 |
| 4 year College Degree (percent) | 43.700 | 49.700 | 28.800 |
| Married (percent) | 36.400 | 48.200 | 51.200 |
| Uninsured (percent) | 11.300 | 31.700 | 13.300 |

## IV. INDEXES FOR BASELINE AND CONTROL OPTIMISM[4]

Based on subjects' perceptions of their personal cancer risk and their beliefs about the risk and efficacy of risky behaviors and prevention efforts, we seek to calculate measures of baseline and control optimism, respectively. Although there is no generally agreed on method for calculating optimism in these two dimensions, we maintain that a key criterion for such measures is that they clearly represent the magnitude of the deviation of actual risk from perceived risk. Consistent with this thinking, we developed the measures described below.

### *4.1 Measures of Baseline Optimism*

Subjects are defined as baseline optimistic if they underestimate their true cancer risk relative to a person in the US of their same age and gender. Based on this, we calculate the variable *Baseline Optimism* for each subject by comparing his/her stated, subjective population-relative

[4] Note that in this section we apply indexes for baseline and control optimism that we developed and described in earlier work (Riddel and Hales, 2015).

risk estimate of cancer incidence with the subject's "actual" population-relative risk factor ("ARR"). We calculate the ARR by applying each subject's responses to demographic, family history, and lifestyle questions in the survey to the risk estimates recorded for those behaviors in the HCRI. As such, we calculate the ARR of subject *i*'s risk of incidence of cancer *j* as follows:

$$ARR_{ij} = \frac{1}{PD\left(\underline{x}_i\right)} \prod_{k=1}^{k_o(j)} RR_{jk}\left(\underline{x}_i\right) \qquad (1)$$

where $k_0\left(j\right)$ is the number of relative risk factors for cancer *j* identified in the HCRI, $\underline{x}_i$ is a vector of subject *i*'s demographic characteristics, family history, and lifestyle choices, $RR_{jk}\left(\underline{x}_i\right)$ is the HCRI-relative risk measure for subject *i* for factor *k* of cancer *j*, and $PD\left(\underline{x}_i\right)$ is a population denominator derived from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program.[5] The resulting population-relative risk factor gives expert-opinion derived estimates of a given subject's risk of incidence of cancer *j*, relative to the US population of persons the same age and gender (Colditz et al. 2000). An $ARR_{i,j}$ value of 1 implies that subject *i* has an average risk of cancer *j* incidence equal to the average of persons in the U.S. of the same age and gender; a value of 2.0 suggests cancer risk that is twice the average, etc.

As reported above, the survey asked respondent to report their perceived risk of contracting each of three cancers, again relative to people of their same age and gender. Following the methodology outlined in Colditz et al. (2000), we constructed allowed responses to range from "Very much below average risk," corresponding to a relative risk value of 0.2, to

[5] Note that as we did not have access to the SEER population denominator for prostate cancer, we used an estimate of 1.107372, based on the average (non-normalized) relative risk factors of our sample of 218 men.

"Very much above average risk," corresponding to a relative risk value of 5.0.[6] We define subject *i*'s stated estimate of relative risk of cancer *j* as $SRR_{ij}$, and create a measure of subjects' baseline optimism as follows:

$$Baseline\ Optimism_{ij} = log_2\left(\frac{ARR_{ij}}{SRR_{ij}}\right). \tag{2}$$

As defined, a *Baseline Optimism* of 0.0 indicates that a subject's own estimate of cancer risk for cancer j is identical to the HCRI estimate. A *Baseline Optimism* value of 1.0 indicates that the subject's estimates of cancer incidence risk are half of the expert-derived value (making her risk perceptions relatively optimistic), and a *Baseline Optimism* value of -1.0 indicates the subject's cancer risk estimates are twice that of the expert value (making her risk perception relatively pessimistic). The logarithmic nature of our index means that each increase (decrease) of one point in our measure thus has the effect of doubling the amount by which expert risk assessments exceed (are exceeded by) the subjects' own-risk estimates.[7]

Next, we calculate an overall estimate of each subject's tendency to exhibit baseline optimism by averaging the separate measures for each of the three cancers considered in our study:

$$Baseline\ Optimism_i = \left(\tfrac{1}{3}\right)\sum_{j=1}^{3} Baseline\ Optimism_{ij} = \left(\tfrac{1}{3}\right)\sum_{j=1}^{3} log_2\left(\frac{ARR_{ij}}{SRR_{ij}}\right). \tag{3}$$

Figure 1 gives the distribution of *Baseline Optimism* for males and females. Roughly three-fourths of subjects of each gender are baseline optimistic. That is, they believe their risk of contracting one of the cancers is lower than that predicted by the HCRI. The mean and variance

[6] In addition to the seven levels of relative risk suggested in Colditz *et al.* 2000, we also allowed survey respondents to select "No risk," which we code as a relative risk factor equal to 0.01.
[7] Note that we chose to use base-2 logarithms as the resulting measures have a more straightforward interpretation than would measures using a natural or base-10 logarithm.

are somewhat higher for females (mean=0.72, std. dev. =1.55) than males (mean = 0.67, std. dev.=1.24).

### *4.2 Measures of Control Optimism*

A subject is "Prevention Control Optimistic" if she believes that engaging in beneficial activities is more effective in reducing cancer risks than it actually is. Similarly, a subject is "Risk Control Optimistic" if she believes that engaging in a particular risky activity is more likely to lead to cancer than it actually is; we therefore infer that she overestimates her ability to *reduce* cancer risks by *avoiding or curtailing* the risky activity in question.

Our survey contained a set of questions for each cancer that elicited subjects' perceptions of the relative impact different activities have on cancer risk. For a given cancer *j* and beneficial activity *k*, subjects were asked to estimate risk-reducing factors between no risk reduction effect (RR=1.0) and a risk reduction of ten-fold (RR=0.1). Table 1 gives the mean subjective relative risk for each of the cancers for each activity. Comparing these estimates with "actual" expert estimates for each cancer and preventative measure associated with each cancer, subject *i*'s level of prevention control optimism is then:

$$Prevention\ Control\ Optimism_i = \frac{1}{\sum_{j=1}^{3} k_{prev}(j)} \cdot \sum_{j=1}^{3} \sum_{k=1}^{k_{prev}(j)} log_2 \left( \frac{APRR_{jk}}{SPRR_{jk}} \right) \tag{4}$$

where $k_{prev}(j)$ is the number of preventative measures identified in the HCRI for cancer *j*, $APRR_{jk}$ is the HCRI-assessed "actual" post-preventative behavior *k* relative risk of cancer *j*, and $SPRR_{jk}$ is the subject's estimates of relative risk of incidence of cancer *j*, assuming behavior *k* (with possible responses coded with RR values ranging from 0.1 to 1.0).

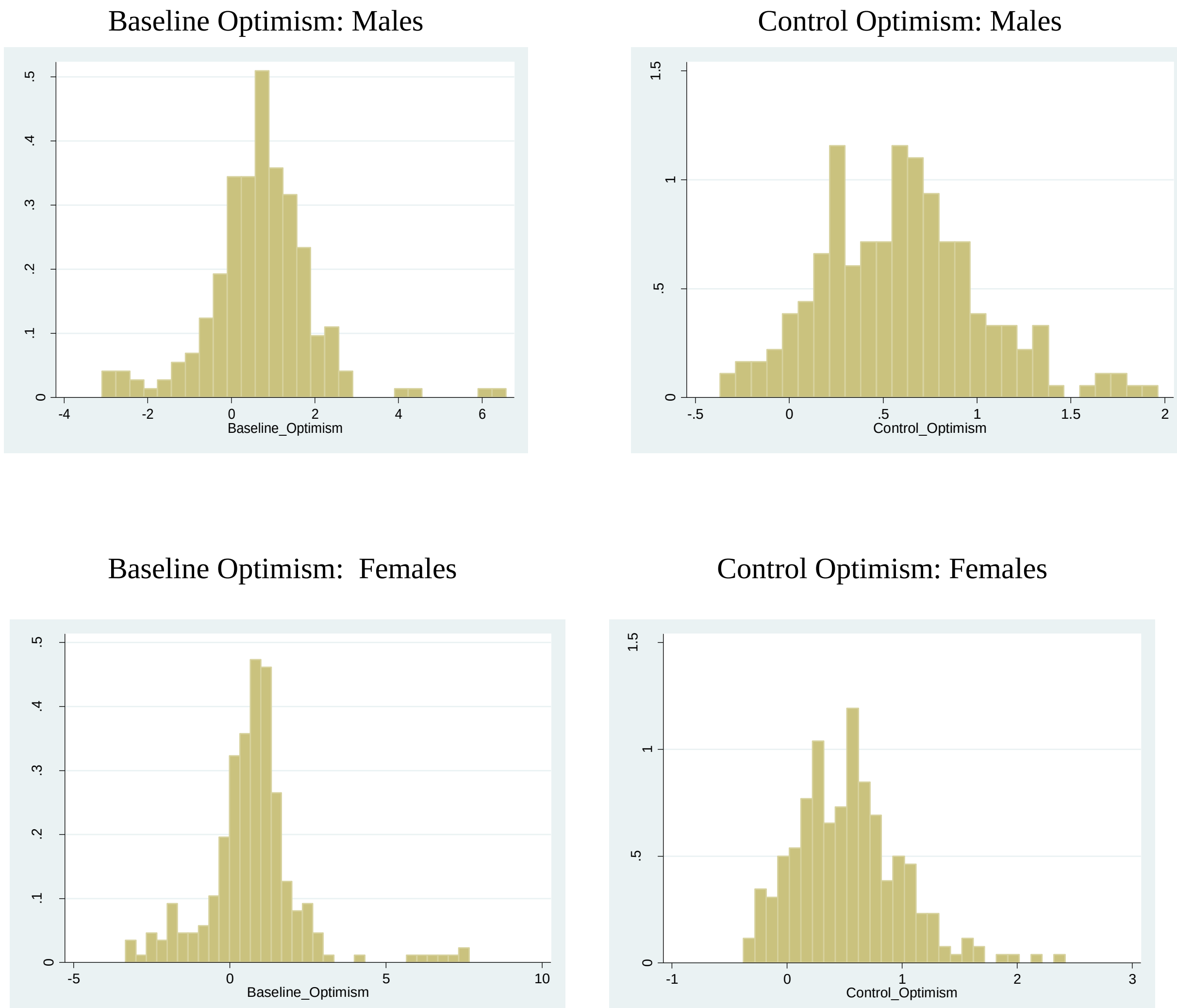
By taking the base-2 logarithm of this ratio, and averaging over the total number of preventative measures identified for each of the three cancers in question, we arrive at a measure of optimism exhibited by subject *i* for a typical preventative measure. A *Prevention Control Optimism* value of 0.0 suggests the subject's estimates of prevention effectiveness are, on average, equal to the "actual" expert estimates. A measure of 1.0 implies that on average, the subject believes preventative measures are twice as effective as they actually are in reducing cancer risk; a measure of -1.0 implies that on average, the subject believes preventative measures are half as effective as they actually are.

Similarly, but with one crucial difference, we estimate risk control optimism as follows:

$$Risk\ Control\ Optimism_i = \frac{1}{\sum_{j=1}^{3} k_{risk}(j)} \cdot \sum_{j=1}^{3} \sum_{k=1}^{k_{risk}(j)} log_2\left(\frac{SRRR_{jk}}{ARRR_{jk}}\right) \tag{5}$$

where $k_{risk}(j)$ is the number of risky activities identified in the HCRI for cancer *j*, $ARRR_{jk}$ is the HCRI-assessed "actual" post-preventative behavior *k* relative risk of cancer *j*, and $SRRR_{jk}$ is the subject's estimates of relative risk of incidence of cancer *j*, assuming behavior *k* (with responses coded with RR values ranging from 1.0 to 5.0). Note that to produce a consistent meaning, the ratio between stated and "actual" risk factors is inverted relative to preventative activities. When *Risk Control Optimism* = 1.0, the subject believes that avoiding risky activities is twice as effective in reducing cancer risk than is actually the case; a measure of -1.0 implies that the subject believes exerting such effort is half as effective in reducing cancer risk than it actually is.

Figure 1. Histograms of Baseline and Control Optimism for Males and Females




Tables 2 and 3 give the mean stated relative risk, the actual relative risk, and the measure of control optimism based on (4) and (5) for each of the prevention and risky behaviors, respectively. Care must be taken in interpreting these measures. For example, the true relative risk for regular physical activity with respect to colon cancer is 0.6, indicating that a subject who exercises regularly has a risk of colon cancer that is 60% of an individual of their age and gender who doesn't engage in exercise. The average subject in the sample believes the relative risk reduction is slightly more at 52.9%. Thus, the average subject overestimates the risk reduction from physical exercise and is therefore optimistic about the effectiveness of exercise in reducing colon-cancer risk.

Table 2 indicates that there is a mix of optimism and pessimism for the effects of the different prevention behaviors on the four cancers in question, but optimism appears to dominate. Of the 26 combinations of behaviors and cancers, they are pessimistic about 11 and optimistic about the remaining 15. Subjects are more pessimistic about reducing colon-cancer risk than bladder cancer, where they are universally optimistic.

Table 3 gives the individual measures of optimism for the *risky* behaviors for each cancer. Again, care must be taken when reviewing the results. A subject is *Risk Control Optimistic* if they over-estimate the riskiness of the activity, thereby indicating that avoiding the activity has more risk-reducing power than what science might predict. The results indicate that the average subject is almost universally risk control optimistic. Of the 41 behavior and cancer combinations, subjects are only pessimistic about two i.e. having a vasectomy to prevent prostate cancer and the risk of prostate cancer given African American descent. For the remaining 39 behaviors, they believe these activities raise cancer risk more than they do. For example, the value of 1.5 for *Risk Control Optimism* for the effect of smoking 1 – 15 cigarettes per day on the risk of colon cancer indicates that the average subject believes the risk is more than double the actual risk. For smoking 15 – 25 cigarettes per day, they believe the risk is slightly more than double the actual risk (Risk Control Optimism equal to 1.06).

Table 2. Perceived Risk and Risk Control Optimism

| Variable | mean stated relative risk | actual relative risk | Prevention Control Optimism= $\log_2$ (actual risk/stated risk) |
|---|---|---|---|
| Colon | | | |
| 30 min/day of Physical Activity | 0.529 | 0.600 | 0.18 |
| Take multivitamin daily | 0.684 | 0.500 | -0.45 |
| Take birth control pills 5+ years | 0.923 | 0.700 | -0.40 |
| Take hormones 5+ years | 0.921 | 0.800 | -0.20 |
| Take daily aspirin | 0.835 | 0.700 | -0.25 |
| Recommneded Vitamin D dose | 0.722 | 0.600 | -0.27 |
| Recent colon screening test | 0.427 | 0.670 | 0.65 |
| Maintain healthy weight | 0.410 | 1.000 | 1.29 |
| Bladder | | | |
| 30 min/day of Physical Activity | 0.566 | 1.000 | 0.82 |
| Take multivitamin daily | 0.681 | 1.000 | 0.55 |
| Take birth control pills 5+ years | 0.948 | 1.000 | 0.08 |
| Take hormones 5+ years | 0.943 | 1.000 | 0.08 |
| Take daily aspirin | 0.862 | 1.000 | 0.21 |
| Maintain healthy weight | 0.475 | 1.000 | 1.07 |
| Breast | | | |
| 30 min/day of Physical Activity | 0.752 | 0.800 | 0.09 |
| Take multivitamin daily | 0.824 | 1.000 | 0.28 |
| Take daily aspirin | 0.937 | 1.000 | 0.09 |
| Breastfeed 1+years | 0.807 | 0.800 | -0.01 |
| Two or more births | 0.847 | 0.850 | 0.01 |
| Take tamoxifen 5+years | 0.881 | 0.600 | -0.55 |
| First period 15+ years old | 0.881 | 0.800 | -0.14 |
| prostate | | | |
| 30 min/day of Physical Activity | 0.806 | 0.800 | -0.01 |
| Take multivitamin daily | 0.867 | 1.000 | 0.21 |
| Take daily aspirin | 0.935 | 1.000 | 0.10 |
| Asian descent | 0.920 | 0.400 | -1.20 |
| Eat tomatoes 3+ times per week | 0.860 | 0.700 | -0.30 |

Table 3. Perceived Risks and Prevention Control Optimism

| Variable | mean stated relative risk | actual relative risk | Risk Control Optimism= $\log_2$ (stated risk/actual risk) |
|---|---|---|---|
| Colon | | | |
| 3+ drinks/day | 3.17 | 1.40 | 1.18 |
| Excess redmeat | 3.64 | 1.20 | 1.60 |
| BMI>30 | 5.21 | 1.50 | 1.80 |
| Sibling with this cancer | 6.01 | 1.80 | 1.74 |
| Taller than 5'7 (females) 5'10 (males) | 1.99 | 1.30 | 0.62 |
| Irritable bowel syndrome | 5.02 | 1.50 | 1.74 |
| Less than 3 servings dairy/day | 1.82 | 1.30 | 0.49 |
| Smoke 1 - 15 cigaettes/day | 2.11 | 1.00 | 1.08 |
| Smoke 15 - 25 cigarettes per day | 2.71 | 1.00 | 1.44 |
| Smoke 25 or more cigarettes per day | 3.18 | 1.50 | 1.08 |
| Bladder | | | |
| Smoke 1 - 15 cigaettes/day | 3.68 | 1.30 | 1.50 |
| Smoke 15 - 25 cigarettes per day | 4.81 | 2.30 | 1.06 |
| Smoke 25 or more cigarettes per day | 5.53 | 3.00 | 0.88 |
| Work in chemical industry 5 - 20 years | 5.93 | 2.50 | 1.25 |
| Work in chemical industry 20+ years | 6.42 | 5.00 | 0.36 |
| Exposed to arsenic 20+ years | 5.87 | 2.00 | 1.55 |
| Family history of colon cancer | 6.46 | 1.50 | 2.11 |
| Breast | | | |
| 3+ drinks/day | 2.80 | 1.30 | 1.11 |
| Currently taking birth control pills | 2.52 | 1.40 | 0.85 |
| Sister with breast cancer | 5.61 | 1.80 | 1.64 |
| Mother with breast cancer | 6.50 | 1.80 | 1.85 |
| Jewish heritage | 2.21 | 1.20 | 0.88 |
| Taller then 5'7" | 1.64 | 1.30 | 0.34 |
| Gain 22 - 44 lbs since age 19 | 2.33 | 1.61 | 0.53 |
| Gain 45+ lbs since age 18 | 3.42 | 1.99 | 0.78 |
| Taking estrogen alone 5+ years | 3.23 | 1.30 | 1.31 |
| Taking estrogen +progesterone < 5 years | 3.02 | 1.30 | 1.21 |
| Taking estrogen +progesterone 5+ years | 3.44 | 2.30 | 0.58 |
| First birth over 35 years old | 2.74 | 1.50 | 0.87 |
| Begnign breast disease | 4.37 | 1.50 | 1.54 |
| Birthweight>8.5 lbs | 1.63 | 1.50 | 0.12 |
| Prostate | | | |
| 3+ drinks/day | 1.874 | 1.000 | 0.906 |
| Family history of prostate cancer | 3.235 | 1.800 | 0.846 |
| African American descent | 1.697 | 1.800 | -0.085 |
| Taller than 5'10" | 1.451 | 1.300 | 0.158 |
| Gain 22 - 44 lbs since age 19 | 1.520 | 1.000 | 0.604 |
| Gain 45+ lbs since age 18 | 2.128 | 1.000 | 1.089 |
| Smoke 1 - 15 cigaettes/day | 2.892 | 1.000 | 1.532 |
| Smoke 15 - 25 cigarettes per day | 2.340 | 1.000 | 1.226 |
| Smoke 25 or more cigarettes per day | 1.911 | 1.000 | 0.934 |
| Vasectomy | 1.233 | 1.700 | -0.463 |

In our modeling effort, it is useful to have an aggregate measure of optimism that combines prevention and control optimism. Thus, we average the values of the two variables for each subject, to arrive at a characteristic level of control optimism for each subject:[8]

$$Control\ Optimism_i = \tfrac{1}{2}\left(Prevention\ Control\ Optimism_i + Risk\ Control\ Optimism_i\right) \qquad (6)$$

An overall *Control Optimism* measure of 0.0 indicates that the subject accurately assesses the efficacy of prevention efforts and danger of risky activities, in aggregate. A value of 1.0 indicates they believe control efforts are double what science would predict while a value of -1.0 indicates efforts are half as effective as the HCRI predicts.

The distribution of *Control Optimism* for males and females is given on the right-hand side of Figure 1. Roughly 93% of males and 90% of females are control optimistic, overestimating factors that increase cancer risk and overestimating the value of prevention efforts in the aggregate. Both the male and female *Control Optimism* distributions have significant right skew, with the male distribution being markedly platykurtic, and the female distribution somewhat less so.

*4.3 Scope Tests for Model Variables*

Due to the stated preference nature of the survey, it is important to gauge whether subjects' responses are consistent with rational choices. A number of rationality tests have been proposed for stated preference surveys, but the scope test is likely the most popular (Viscusi 2013). The current survey contained a single-bounded contingent valuation (CV) question addressing willingness to pay for insurance that covered all monetary costs associated with a cancer diagnosis. A companion paper (Riddel and Hales 2015) uses the results of the contingent

[8] We elected to weight values for prevention and risk control optimism equally in this estimate, rather than weighting by the total number of preventative and risk-related attributes for each cancer. We did this to avoid overweighting the influence of risk-related attributes, of which more are identified in the HCRI than are preventative-related attributes.

valuation question to investigate adverse and advantageous selection in the market for cancer insurance. However, the CV data also provides for a test of scope. Economic theory indicates demand for insurance, hence willingness to pay for the full coverage should be 1) increasing in income if cancer insurance is a normal good, and 2) increasing in the insured risk. Moreover, the coefficient of the offered premium should be negative, indicating that the higher the cost of the insurance, the less likely the subject will purchase the coverage.

The results of probit models relating demand for full coverage as a function of the offered premium, perceived cancer risks, and income by gender are reported in Table 4. The dependent variable is equal to one if the subject agreed to pay the offered premium and zero otherwise. Note that the ratio of the coefficient of the perceived risk and income variables to the coefficient of the premium amount is equal to the marginal willingness to pay for that variable. Due to high correlation between the individual perceived cancer risks, the equations are estimated separately.

The offered premium is negative and statistically significant in all ten models, indicating that the higher the premium offered to the subject, the less likely they are to agree to purchase the hypothetical cancer insurance coverage. The coefficients of the perceived cancer risk variables are all positive and all but one are statistically significant for a p-value of 0.05 or better. Thus, willingness to pay for full coverage is increasing in the perceived risk of contracting the different cancers. Finally, willingness to pay is also increasing in income, supporting the hypothesis that cancer insurance is a normal good.

Taken together, the scope tests indicate that people respond consistently to the survey. In the next section, we present our behavioral models and results.

Table 4. Scope Tests for the Premium, Individual Perceived Cancer Risks, and Income. Probit model where dependent variable equals one if the subject agreed to pay the offered premium. P-values in parentheses.

| Variable | Females | | | | |
|---|---|---|---|---|---|
| Premium | -0.004 (0.000) | -0.011 (0.000) | -0.008 (0.000) | -0.007 (0.000) | -0.009 (0.000) |
| Perceived Bladder Risk | | 0.633 (0.000) | | | |
| Perceived Breast Risk | | | 0.162 (0.001) | | |
| Perceived Colon Risk | | | | 0.212 (0.001) | |
| Income | | | | | 0.009 (0.000) |
| | | | | | |
| Variable | Males | | | | |
| Premium | -0.006*(0.000) | -0.010 (0.000) | -0.009 (0.000) | -0.008 (0.000) | -0.010 (0.000) |
| Perceived Bladder Risk | | 0.360 (0.014) | | | |
| Perceived Prostate Risk | | | 0.246 (0.007) | | |
| Perceived Colon Risk | | | | 0.090 (0.237) | |
| Income | | | | | 0.008 (0.005) |

## V. BEHAVIOR MODEL RESULTS

The goal of this paper is to understand how utility risk aversion, cognitive ability, and the different types of optimism influence risky and prevention behavior, controlling for demographic variables that may affect behavior. Toward that end, we estimate probit models first for the five risky behaviors in question followed by the three prevention behaviors. The model results are reported in the following subsections.

Table 5 gives the results of ten probit models that evaluate the influence of optimism, utility risk aversion, cognitive ability, and demographic control variables on the probability of engaging in the risky activities. The risky behaviors evaluated are being a current smoker, being an ever smoker, excessively consuming red meat (eating red meat more than three times per week), being obese [having a Body Mass Index (BMI) of 30 or higher], and heavy alcohol consumption (more than two drinks per day). For each behavior, we give a full model including all of the variables of interest and a reduced model that excludes cognitive ability and the optimism variables. The table gives the marginal effects of a change in the independent variable on the probability the subject engages in the risky activity. P-values in parentheses are calculated using robust standard errors.

The results indicate that older, lower income subjects are significantly more likely to be current smokers than their younger, wealthier counterparts. College graduates are less likely to smoke currently than subjects with less education. Gender, marital status, and insured status do not have a statistically significant influence on the decision to smoke currently.

The first two columns address the probability the subject is a current smoker. Subjects who are relatively baseline optimistic about their risk of contracting one or more of the three cancers are significantly more likely to smoke (p-value=0.036). Subjects who are control optimistic i.e. those who overestimate the efficacy of prevention efforts and avoiding risky activities, in aggregate, are less likely to smoke (p-value=0.053). Cognitive ability is also statistically significant: subjects with higher cognitive abilities are less likely to smoke. Interestingly, the marginal effect of utility risk aversion in the reduced model is half that of the full model that controls for optimism and cognitive ability. Moreover, the coefficient of utility risk aversion is significant in the reduced model but becomes insignificant in the full model. These results suggest that controlling for optimism significantly reduces the influence of utility risk aversion on the decision to smoke.

Table 5. Probit Models of Risky Behavior: Depdendent Variable Equals 1 if the Subject Engages in that Behavior and Zero Otherwise

| | Current Smoker | | Ever Smoker | | Red Meat | | High BMI | | Heavy Drinker | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variable | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) |
| Risk Aversion | -0.007 (0.346) | -0.014 (0.05) | -0.015 (0.111) | -0.021 (0.020) | 0.002 (0.818) | 0.001 (0.87) | -0.009 (0.348) | -0.010 (0.276) | -0.037 (0.206) | -0.054 (0.053) |
| Cognitive Ability | -0.043 (0.000) | ----- | -0.032 (0.036) | ----- | -0.006 (0.707) | ----- | -0.007 (0.680) | ----- | -0.084 (0.077) | ----- |
| Baseline Optimism | 0.027(0.036) | ----- | 0.015 (0.364) | ----- | 0.029 (0.076) | ----- | 0.034 (0.024) | ----- | -0.002 (0.976) | ----- |
| Control Optimism | -0.080 (0.053) | ----- | -0.063 (0.226) | ----- | -0.049 (0.332) | ----- | 0.100 (0.054) | ----- | -0.402 (0.021) | ----- |
| Male | 0.042 (0.288) | -0.026 (0.493) | 0.073 (0.137) | 0.022 (0.634) | 0.085 (0.071) | 0.078 (0.079) | 0.249 (0.000) | 0.248 (0.000) | 0.089 (0.555) | -0.112 (0.419) |
| Age | 0.003 (0.020) | 3.00E-05 (0.975) | 0.009 (0.000) | 0.006 (0.000) | -3.99E-04 (0.826) | -0.001 (0.449) | 0.002 (0.243) | 0.002 (0.328) | -0.009 (0.114) | -0.018 (0.000) |
| Income ($000) | -0.002 (0.008) | -0.003 (0.000) | -0.002 (0.011) | -0.002 (0.002) | -0.001 (0.356) | -0.001 (0.216) | -0.001 (0.130) | -0.001 (0.163) | -0.004 (0.123) | -0.006 (0.018) |
| College | -0.103(0.007) | -0.125 (0.001) | -0.174 (0.000) | -0.187 (0.000) | -0.052 (0.276) | -0.061 (0.189) | -0.059 (0.201) | -0.079 (0.076) | 0.044 (0.779) | 0.031 (0.832) |
| Married | 0.035 (0.414) | 0.024 (0.583) | -0.007 (0.893) | -0.008 (0.874) | 0.009 (0.860) | 0.017 (0.732) | 0.085 (0.092) | 0.093 (0.063) | 0.199 (0.227) | 0.118 (0.453) |
| Uninsured | 0.002 (0.971) | -0.035 (0.549) | -0.046 (0.537) | -0.071 (0.327) | -0.075 (0.294) | -0.068 (0.335) | -0.089 (0.188) | -0.010 (0.129) | 0.015 (0.950) | -0.099 (0.674) |
| Log Likelihood | -228.540 | -245.760 | -300.840 | -308.830 | -317.770 | -322.970 | -277.630 | -298.83 | -181.984 | -194.523 |

As with current smokers, ever smokers tend to be older, lower income, and less likely to have a college degree. Marital status, insurance status, and gender are not statistically significant. This is true of both the full and reduced model. When subjects who have quit smoking are included in the smoker category, neither optimism variable is statistically significant. As in the current smoker model, subjects with higher cognitive ability are less likely to have ever smoked in their lives. Risk averse subjects are less likely to have ever smoked, but the effect is only marginally significant (p-value=0.11). However, when optimism and cognitive ability are excluded, utility risk aversion becomes strongly significant (p-value=0.02). Again, when optimism and cognitive ability are considered, the effect of utility risk aversion on optimism is muted, with the marginal effect falling from -0.021 in the reduced model to -0.015 in the full model.

Excessive red-meat consumption is increasing in baseline optimism so that those who underestimate their personal risk of the three cancers are more likely to eat more than three servings of red meat per week. Utility risk aversion, cognitive ability, and control optimism are not statistically significant determinants of red-meat consumption in the full model. Utility risk aversion remains insignificant in the reduced model. Of the demographic variables, only gender is significant.

The model results suggest that both types of optimism increase the likelihood that a subject has a BMI of 30 or higher. On one hand, subjects who are baseline optimistic and therefore underestimate their cancer risk are more likely to be obese than subjects with accurate baseline risk perceptions. On the other hand, subjects who believe that prevention activities and avoiding risky activities are a more effective strategy for reducing cancer risk than what experts would believe are also more likely to be obese. This latter correlation is counterintuitive, since we expect that control optimists would tend to over-invest in risk-reducing effort. Males and married subjects in our sample are more likely to be obese, while income, education, and insurance status are not statistically significant predictors of obesity. Cognitive ability also does not influence the probability of obesity when controlling for optimism and demographic variables. Similarly, utility risk aversion is not statistically significant in either the full model or the reduced model.

As expected, heavy drinking is decreasing in control optimism: subjects who have optimistic views about the health benefits of avoiding risky activities and engaging in prevention measures are less likely to drink heavily. Subjects with higher cognitive ability are also less likely to consume detrimental amounts of alcohol. The probability of heavy drinking is decreasing in both age and income, but the coefficients are only marginally significant in the full model (p-values of 0.11 and 0.12, respectively). None of the remaining demographic variables are significant. As with the model for current smoking, the coefficient of utility risk aversion is much smaller in the full model than in the reduced model. Also note that the coefficient of utility risk aversion is not statistically significant in the full model that controls for optimism and cognitive ability, but becomes significant when these variables are excluded from the regression. Again, it appears that controlling for optimism and cognitive ability moderates the effect of utility risk aversion on risky behavior.

Table 6 gives the marginal effects for a change in the probability of engaging in a cancer risk-reducing behavior. Three healthy behaviors are studied: physical activity (engaging in 30 minutes or more of physical activity most days or at least three hours per week), taking a daily aspirin, and taking a daily multivitamin. We hypothesize that engaging in health behaviors will be increasing in utility risk aversion as more risk-averse subjects are more likely to take steps to reduce health risks. We expect that baseline optimists will be less likely to engage in these risk-reducing behaviors since they tend to underestimate their cancer risk and therefore underinvest in risk reduction. Finally, we expect the probability of engaging in the healthy behaviors to be increasing in control optimism since these subjects over-estimate the efficacy of risk-reducing effort and would therefore engage in more of these activities than their more pessimistic counterparts.

Table 6. Probit Models of Healthy Behaviors: Dependent Variable Equals One if the Subject Engages in the Activity

| | Physically Active | | Aspirin Daily | | Multivitamin | |
|---|---|---|---|---|---|---|
| Variable | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) | dy/dx (p-value) |
| Risk Aversion | -0.004 (0.657) | -0.005 (0.572) | -0.003 (0.178) | -0.0003 (0.828) | 0.016 (0.070) | 0.013 (0.137) |
| IQ | -0.011 (0.541) | ----- | -0.013 (0.007) | ----- | 0.001 (0.967) | ----- |
| Baseline Optimism | -0.077 (0.000) | ----- | -0.010 (0.033) | ----- | -0.010 (0.000) | ----- |
| Control Optimism | 0.029 (0.572) | ----- | -0.006 (0.612) | ----- | 0.047 (0.364) | ----- |
| Male | 0.118 (0.008) | 0.113 (0.009) | -0.005 (0.659) | -0.001 (0.853) | -0.057 (0.207) | -0.045 (0.310) |
| Age | -0.003 (0.094) | -0.003 (0.131) | 0.001 (0.137) | 0.001 (0.006) | 0.004 (0.046) | 0.003 (0.114) |
| Income ($000) | 0.001 (0.461) | 0.001 (0.504) | -0.001 (0.052) | -0.001 (0.023) | 0.0006 (0.384) | 0.0004 (0.611) |
| College | 0.118 (0.008) | 0.123 (0.005) | -0.008 (0.502) | -0.004 (0.591) | 0.010 (0.034) | 0.110 (0.014) |
| Married | -0.033 (0.510) | -0.049 (0.314) | 0.016 (0.248) | 0.016 (0.228) | -0.018 (0.713) | -0.028 (0.562) |
| Uninsured | 0.058 (0.392) | 0.035 (0.606) | (omitted) | (omitted) | 0.119 (0.110) | 0.093 (0.191) |
| Log Likelihood | -280.54 | -294.47 | -44.02 | -50.52 | -278.22 | -298.83 |

The model results indicate that baseline optimists are significantly less likely to meet the exercise minimums, but utility risk aversion does not influence physical activity in either the full or reduced model. Younger and college educated subjects are more likely than their older, less educated counterparts to achieve the minimum exercise requirement. Males are more likely to engage in regular physical activity than female subjects. Cognitive ability, control optimism, income, marital status, and insurance status are not statistically significant determinants of exercise frequency.

Daily aspirin consumption is decreasing in baseline optimism cognitive ability, and income. None of the other variables are statistically significant predictors of the probability of taking a daily aspirin.

More utility risk-averse subjects are more likely to take multivitamins. This is true in the full model, but the significance level of the variable declines in the reduced model. As expected, the probability of daily multivitamin use is decreasing in baseline optimism. Older subjects and college-educated subjects are also more likely to take multivitamins. Cognitive ability, control optimism, income and marital and insurance status do not affect the choice to take a daily multivitamin.

## VI. DISCUSSION

The models indicate that optimism significantly influences people's decisions about engaging in risky activities. On one hand, baseline optimists – those who underestimate their risk of contracting one of the three cancers – are more likely to be current smokers, to eat excessive amounts of red meat, and to be medically obese. On the other hand, they are less likely to engage in the prevention behaviors such as regular physical activity, taking a daily aspirin and multivitamin. Accordingly, baseline optimistic subjects engage in these risky behaviors and fail to undertake preventative health measures because they believe that they are healthier, at least in terms of cancer risk, than they actually are. This result suggests that one strategy a physician or other medical consultant could use to encourage people to mitigate their cancer risk would be to provide patients with accurate information about their cancer risk given current behaviors.

The relationship between control optimism and risky and prevention behavior is rather more complex. While control optimism appears to deter smoking and heavy drinking, it actually leads to an increase in the probability a subjects is overweight. None of the other risky or prevention behaviors are influenced by control optimism. These findings support those of Viscusi's (1990) that subjects who tend to overestimate the specific risk of contracting lung cancer from smoking are less likely to smoke. This is what one would predict from standard models of preferences: subjects with higher perceptions of the risk and therefore expected costs of engaging in an activity will be more likely, ceteris paribus, to avoid the costly activity.

The finding that obesity is increasing in control optimism is difficult to explain. We would expect that control optimists would engage in more prevention than those with accurate perceptions since they believe that the marginal return to effort is higher than it actually is. Presumably, one of the outcomes of increased effort would be a lower BMI and a decline in the likelihood that the subject is obese. Of course body weight is an accumulation of past decisions about eating and exercise. It could well be that many subjects with a high BMI live currently healthy lifestyles i.e. engage in a significant amount of prevention activities, but continue to suffer from obesity. A more sophisticated analysis of current and past behavior may be able to give a deeper insight into control optimism, investing in prevention effort, and obesity.

One of the key goals of this paper is to examine the complex interaction between utility risk aversion, baseline and control optimism, and cognitive ability and their joint influence on engaging in risky and prevention behaviors. While utility risk aversion is significant in the reduced models for current smoking, ever smoking, and heavy drinking, the variable becomes insignificant when controlling for optimism and cognitive ability and the size of the effect, as measured by the marginal probability, is reduced by half or more. Only the positive correlation between utility risk aversion and multivitamin use remains when controlling for optimism and cognitive ability. Risk aversion is not correlated with BMI or physical activity in either the full or reduced model.

It is possible that a high correlation between optimism, utility risk aversion, and cognitive ability are inflating the standard errors of the respective coefficients and making the model estimates unstable. This is less likely to be of concern in a probit model because the inherent nonlinearity between the dependent and independent variables help identify the coefficients. Nevertheless, to test for multicollinearity, we calculated variance inflation factors (VIFs) for the utility risk aversion, cognitive ability and the optimism variables for each of the models. The average VIF is 1.08 with a maximum VIF of 1.16. A frequently used rule of thumb suggests that multicollinearity is a problem if a variable or set of variables have a VIF of 10 or higher. Since the VIFs for the current set of variables are all well below 2, it is highly unlikely that multicollinearity is responsible for the high standard errors for the utility risk aversion variable in the full regressions. Rather, we conclude that for the overwhelming majority of our models, it is optimism, rather than utility risk aversion, that is determining the amount of effort invested in risky and prevention activities. This indicates that health-preference functions are best represented by models such as Cumulative Prospect or Rank-Dependent preferences that allow probability weighting to influence health choices.

We are not the first authors to conclude that health preferences are best modeled using Cumulative Prospect or Rank Dependent preferences. Wakker and Deneffe (1996), Bleichrodt and Pinto (2000) and Riddel and Kolstoe (2013) have all found evidence of significant probability weighting in the health domain. Moreover, numerous studies have found that over-estimation of health risks can lead people to avoid risky behaviors. Our main contribution is to disentangle the influence of probability weighting and utility risk aversion on behavior.

Our findings are consistent with the previous work such as Barsky et al. (1995) and Anderson and Mellor (2008) that more “risk averse” subjects are less likely to smoke, less likely to drink excessively, and less likely to be obese. Risk aversion is a general property of preference functions: when the function is concave, then the subject is deemed risk averse. If EU preferences are assumed, then the only source of risk aversion is the utility function. Our findings suggest that it is probabilistic risk aversion, represented by baseline and control optimism, rather than curvature in the utility function that motivates many subjects to avoid risky behaviors and engage in prevention.

Of course, our results are not directly comparable to Barsky et al. (1995) and Anderson and Mellor (2008) because their measure of risk aversion is constructed over financial choices rather than life-duration choices used in the current study. Arguably, the most notable influence on engaging in risky behaviors that increase cancer risk would be an aversion to life-duration gambles, however a cancer diagnosis can also have significant financial costs as well. It could be that if we considered financial lotteries that the probability of engaging in these risky behaviors would be decreasing in financial utility risk aversion. That said, previous research has found that financial and life-duration utility risk aversion measures are highly correlated (Einav et al 2012, Dohmen et al. 2011). Still, our results indicate that at least in the domain of cancer-related health factors, optimism – particularly baseline optimism – is a more important factor than utility risk aversion in predicting individual health-risk decisions.

## VII. CONCLUSIONS

In this paper, we report the results of a survey of 474 men and women that analyzes variables that influence cancer-prevention activities. In particular, we seek to sort out the effects of utility risk aversion, probability weighting revealed through baseline and control optimism, and cognitive ability on the choice to engage in risky or prevention activities.

Our results indicate that utility risk aversion is far less influential than previously believed. In contrast, we find that choice patterns that were previously attributed to utility risk aversion more likely arise from optimism about one's cancer risk paired with optimism about the efficacy of prevention and risk avoidance behaviors.

This finding is potentially important to public health experts. If choices to engage in risky behavior and prevention effort are primarily motivated by diminishing marginal utility of health, then there is not much room for information to change their behavior since the shape of the utility function is rooted in individual preferences. However, if these decisions rest largely on misperceptions of the risk of contracting cancer, then interventions in the form of accurate information about scientifically assessed estimates of baseline risk as well as the efficacy of prevention may well guide subjects to engage in more prevention effort.

This is the first empirical study we know of that investigates the role of risk misperception in the domain of cancer-related health behaviors while simultaneously controlling for cognitive ability and utility risk aversion. However, there are some limitations to our analysis. For one, we ask people about their *perceptions* of the risk and benefits of different behavior. Risk perceptions are difficult to elicit and problems with numeracy may well induce measurement error in these variables (Reyna et al. 2009). We control for cognitive ability, which should offset bias introduced by numeracy issues. However, numeracy issues could still be a problem in the estimates. Also, we use self-reported measures of risky and prevention behavior. Some researchers question the validity of self-reported smoking and other risky behaviors, believing that subjects may understate how much risky behavior they engage in either because they underestimate the amount or intentionally provide a lower number because they don't want to be associated with an undesirable behavior. Patrick et al. (1994) found that under-reporting was less likely to happen in self-administered surveys such as the current survey. Nevertheless, significant under reporting of risky behavior or over-reporting of prevention behavior could potentially cloud our results.

REFERENCES

Abdellaoiu, A., O. L'Harridon, and H. Zank. (2010). Separating curvature and elevation: A parametric probability weighting function, *Journal of Risk and Uncertainty,* Vol. 41:39–65.

Anderson, L.R. and J.M. Mellor. (2008). Predicting health behaviors with an experimental measure of risk preference, *Journal of Health Economics,* Vol. 27(5): 1260-1274.

Attema, A.,W.B.F. Brouwer and O. l'Haridon. (2013). Prospect theory in the health domain: A quantitative assessment, *Journal of Health Economics,* Vol. 32: 1057-1065.

Barsky, R., Kimball, M., Juster, F.T., and Shapiro, M. (1995). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey, *National Bureau of Economic Research Working Paper 5213*.

Bijwaard, G.E, H. van Kippersluis, and J. Veenman. (2015). Education and health: The role of cognitive ability, *Journal of Health Economics,* Vol. 42: 29-43.

Bleichrodt, H. & J. L. Pinto. (2000). A parameter-free elicitation of the probability-weighting function in medical decision analysis. *Management Science,* 46: 1485-96.

Bruhin, A., H. Fehr-Duda, and T. Epple. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion, *Econometrica*, Vol. 78(4): 1375-1412.

Buhrmester,M. T. Kwang and S. D. Gosling. (2011). Amazon's Mechanical Turk
A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* January 2011**,** vol. 6(1): 3-5**.**

Camerer, C. and R. M. Hogarth. (1999). The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty,* Vol. 19:7-42.

Carman, K. and P. Kooreman. (2014). Probability perceptions and preventative care. *Journal of Risk and Uncertainty*, Vol. 49(1): 43-71.

Colditz, G.A., K.A. Atwood, K. Emmons, R.R. Monson, W.C. Willett, D. Trichopoulos & D.J. Hunter. (2000). "Harvard Report on Cancer Prevention Volume 4: Harvard Cancer Risk Index," *Cancer Causes and Control* Vol. 11: pp. 477-488.

Conte, A., J. D. Hey and P. G. Moffatt. (2011). Mixture models of choice under risk, *Journal of Econometrics*. Vol. 162: 79-88.

Conte, A., P.G. Moffatt and M. Riddel. (2015). Heterogeneity in risk attitudes across domains: A bivariate random preference approach. Working paper. University of East Anglia, Centre for Behavioural and Experimental Social Science, Discussion Paper 15-10.

Dave, D., H. Saffer. (2007). Risk tolerance and alcohol demand among adults and older adults. National Bureau of Economic Research Working Paper No.13482

Dohmen, T.; A. Falk, D. Huffman, U. Sunde, J. Schupp, and G.G. Wagner. (2011). Individual Risk Attitudes: Measurement, Determinants and Behavioral Consequences, *Journal of the European Economic Association*, Vol. 9(3): 522-550.

Einav, L., A. Finkelstein, J. Pascu, and M.R. Cullen. (2012). How General are Risk Preferences? Choices under uncertainty in different domains, *American Economic Review* Vol. 102(6): 2606-2638.

Holt, C. and S. Laury. (2002). Risk aversion and incentive effects. *American Economic Review*, Vol. 95, 1644-1655.

Khwaja, K. F. Sloan and S. Chung. (2007). The relationship between individual expectations and behaviors: Mortality expectations and smoking decisions. *Journal of Risk and Uncertainty*, Vol. 35(2): 179-201.

Lundborg, P. and B. Lindgren. (2002). Do risk perceptions and alcohol consumption. *Journal of Risk and Uncertainty*, Vol. 25(2): 165-183.

Lundborg, P. and B. Lindgren. (2004). Do they know what they are doing? Risk perceptions and smoking behavior among Swedish teenagers. *Journal of Risk and Uncertainty*, Vol. 28(3): 261-286.

Quiggin, J. (1993). Generalized Expected Utility Theory. The Rank-Dependent Model. Kluwer. Boston, Dordrecht/London.

Patrick, D.L, A. Cheadle, D.C. Thompson, P. Diehr, T. Koepsell, and S. Kinne. (1994). Self-Reported Smoking. *American Journal of Public Health*. Vol. 84: 1086-1093.

Picone, G., Frank, S., Taylor Jr., D., (2004). Effects of risk and time preference and expected longevity on demand for medical tests. *The Journal of Risk and Uncertainty* Vol. 28(1): 39–53.

Prelec, D. (1998). The probability weighting function, *Econometrica* Vol. 66(3):497-527.

Rabin, M. (1998). Psychology and economics, *Journal of Economic Literature*, Vol. 36(1): 11-46.

Reyna, V.F, W.L. Nelson, P.K. Han, and N.F. Dieckman. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*. Vol. 135(6):943-973.

Riddel, M. and D. Hales. (2015). Risk misperceptions and selection in insurance markets: An application to demand for cancer insurance. *Working Paper*, University of Nevada, Las Vegas.

Riddel, M. and S. Kolstoe. (2013). Heterogeneity in life-duration preferences: Are risky recreationists really more risk loving? *Journal of Risk and Uncertainty*, Vol. 46(2): 191-213.

Spinnewijn, J. (2013). "Insurance and perceptions: How to screen optimists and pessimists," *The Economic Journal*, June 2013, Vol. 123(569), pp. 606-633.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, *Journal of Economic Literature*, Vol. 38(2):332-382.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, Vol. 5: 297–323.

Tversky, A. and C.R Fox (1994). Weighting risk and uncertainty, *Psychological Review*, Vol. 102: 269-283.

Viscusi, W.K. and J. Hersch. (2001). Cigarette smokers as job risk takers, *The Review of Economics and Statistics*, Vol. 83(2): 260-280.

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

Viscusi, W.K. (1991). Age variations in risk perceptions and smoking decisions. *The Review of Economics and Statistics*, Vol. 73(4): 577-588.

Viscusi, W.K, I. Carvalho, F. Antonanzas, J. Rovira, F.J. Brana and F. Portillo. (2000). Smoking risks in Spain III-Determinants of behavior, *Journal of Risk and Uncertainty,* Vol. 21(2): 213-234.

Viscusi, W.K. (2013). The Value of Individual and Societal Risks to Life and Health. In Handbook in Economics: Economics of Risk and Uncertainty, Vol. I. eds. M. Machina and W.K. Viscusi. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: North Holland, Elsevier.

Wakker, P. and D. Deneffe. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, Vol. 42, 1131-50.

Weber, E. U., A.R. Blais, and N. E. Betz. (2002). A Domain-specific risk-attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making*, Vol. 15: 263–290.

Winter, J. and A. Wupperman. (2014). Do they know what is at risk: Health risk perception among the obese, *Health Economics*, Vol. 23: 564-585.

Wu, G. and R. Gonzalez. (1996). "Curvature of the probability weighting function. *Management Science.* Vol. 42: 1676-1690.