HEDG Working Paper 10/28

Smoking, Expectations, and Health: A Dynamic Stochastic Model of Lifetime Smoking Behavior

Michael Darden

October 2010

york.ac.uk/res/herc/hedgwp

Smoking, Expectations, and Health: A Dynamic Stochastic Model of Lifetime Smoking Behavior

Michael Darden*
The University of North Carolina at Chapel Hill‡
Job Market Paper

October 25, 2010

Abstract

This research discusses results obtained through formulation and estimation of a dynamic stochastic model that captures individual smoking decision making, health expectations, and longevity over the life cycle. The standard rational addiction model is augmented with a Bayesian learning process about the health marker transition technology to evaluate the importance of personalized health information in the decision to smoke cigarettes. Additionally, the model is well positioned to assess how smoking, and smoking cessation, impacts morbidity and mortality outcomes while taking into consideration the potential for dynamic selection of smoking behaviors. This research also provides a novel approach to the empirical construction of the theoretically common "smoking stock" that facilitates the estimation of investment and depreciation parameters. The structural parameters are estimated using rich longitudinal health and smoking data from the Framingham Heart Survey: Offspring Cohort. Results suggest that there exists heterogeneity across individuals in the pathways by which smoking effects health. Furthermore, upon smoking, the estimated parameters suggest a positive reinforcement effect and a negative withdrawal effect, both of which encourage future smoking. The paper also presents evidence of health selection in smoking behavior that, when not modeled, may cause an overstatement of the effect of smoking on morbidity and mortality. Finally, personalized health marker information is not found to significantly influence smoking behavior relative to chronic health shocks themselves.

*I would like to thank my dissertation adviser, Donna Gilleskie, as well as Vijay Krishna, Maarten Lindeboom, Brian McManus, Frank Sloan, Helen Tauchen, and participants in the UNC Applied Microeconomics Workshop, Second Annual Health Econometrics Workshop, and 7th iHEA World Congress for helpful comments. This paper was formerly circulated under the title "A Dynamic Learning Model of Smoking Behavior". Email:mdarden@email.unc.edu

†Department of Economics, 107 Gardner Hall, CB#3305, UNC-Chapel Hill, Chapel Hill, NC, 27599-3305.

‡The Framingham Offspring Study (FOS) is conducted and supported by the NHLBI in collaboration with the FOS Study Investigators. This manuscript was prepared using a limited access dataset obtained from the NHLBI and does not necessarily reflect the opinions or views of the FOS or the NHLBI.

Introduction

The decision to smoke has long interested social scientists and health policy researchers because of the seemingly irrational nature of such a choice. Why would an individual undertake an activity with such clear negative health consequences? A thorough review of this debate can be found in Sloan *et al.* (2003). Ultimately, those authors conclude that individuals make decisions within an environment that reflects individual preferences but one that is also subject to information acquisition costs. Gary Becker describes economic decision makers: "They (economic agents) are not expected to be perfect optimizers, as evaluated by the analyst, or dispassionate external observers; rather, people do the best they can, given their information and their cognitive abilities to understand it (qtd. in Sloan *et al.* (2003) pg. 25)." An important question addressed by the smoking literature has been: what determines and shapes "their information?" Furthermore, how does information influence smoking behavior? And, to what extent has information regarding the health effects of smoking been free from selection bias? These questions form the basis for the current paper.

The purpose of this work is to analyze the relationship between the consumption of cigarettes and health in a dynamic discrete choice framework that incorporates learning. In addition to assessing the role of personalized health information in the decision to smoke, this paper evaluates the effects of smoking cigarettes on morbidity and mortality outcomes while taking into consideration the potential for dynamic selection of smoking behaviors. I estimate the structural parameters of an individual's optimization problem with the following trade-off: current enjoyment of cigarette consumption versus the associated uncertain future utility and health consequences. I consider two dimensions of health: health markers and chronic health. Health markers are those factors (e.g., blood pressure, cholesterol, etc.) viewed by the medical literature to significantly predict the onset of chronic conditions (e.g., cardiovascular disease, cancer, etc.). Given a history of these health markers and smoking behaviors, an individual is able to more precisely evaluate the effect of smoking on her health markers levels which, in turn, helps to determine her chronic health probability. More generally, health markers offer information as to an individual's overall health condition. Endowed with this information, an individual makes the smoking choice that maximizes her present discounted expected utility. Smoking history is modeled as a capital stock and is measured in a novel way so as to facilitate the estimation of depreciation and investment coefficients while keeping the model computationally tractable. The structural parameters of the model are estimated with rich longitudinal data from the offspring of the original cohort of the Framingham Heart Survey.

The current paper fits into and extends the literature in four ways. First, the structural model embeds the standard rational addiction model of Becker and Murphy (1988). In

the current paper, forward-looking agents evaluate current smoking alternatives while taking into consideration the future health and utility consequences associated with past and current smoking behavior. A major contribution of the current paper is to model how health marker information may alter smoking behavior *prior* to major chronic health complications. I extend the literature on smoking responses to personalized health information that has only considered chronic health shocks (Smith *et al.*, 2001; Khwaja *et al.*, 2006; Arcidiacono *et al.*, 2007). This distinction is important if the potential gains from information from a chronic health shock are "too late". Furthermore, motivated by the Becker quote above, my model allows for the possibility of learning about the effects of smoking on health. A second contribution is the model's ability to assess the impact of smoking, and smoking cessation, on morbidity and mortality outcomes while accounting for dynamic selection of smoking behaviors. Using data from across the life cycle, I measure the role of health and mortality transition determinants by estimating these production technologies within the structural model of lifetime smoking decisions. I allow the unobserved errors that affect smoking, health, and mortality to be serially correlated through a common permanent unobserved component. My method improves upon recent papers that estimate health transitions outside the structural model (Adda and Lechene, 2001). Third, my model extends the empirical smoking literature with a novel construction of the theoretically common "smoking stock". Using factor analysis in a method similar to Sickles and Williams (2008), I create a continuous smoking stock index from several variables that reflect past smoking behavior. This state variable captures the unique smoking history that each individual brings into each decision making period. Measuring the smoking stock using this method also allows for the estimation of depreciation and investment parameters (Adda and Lechene, 2004). Finally, the model is solved using techniques common in current structural dynamic discrete choice modeling (Rust, 1987; Keane and Wolpin, 1994; Aguirregabiria and Mira, 2010). To incorporate learning, I combine several features of other recent structural papers that have explicitly modeled and estimated Bayesian learning processes (Ackerberg, 2003; Crawford and Shum, 2005; Chan and Hamilton, 2006; Mira, 2007; Chernew *et al.*, 2008). I incorporate permanent unobserved heterogeneity in a nonparametric fashion (Heckman and Singer, 1984; Mroz, 1999) following recent structural examples (Arcidiacono *et al.*, 2007; Blau and Gilleskie, 2008). This method amounts to a random effects specification of unobserved heterogeneity that is free from distributional assumptions. Conditional on the unobserved heterogeneity, I use the model to predict the initial conditions (Khwaja, 2010).

To evaluate the roles of learning and information, I use the model and the estimated structural parameters to simulate smoking behavior and health and mortality outcomes under different counterfactual scenarios. The results suggest that there exists heterogeneity across

individuals in the pathways by which smoking effects health. I find that the effect of an accumulated smoking stock on health markers varies widely across individuals relative to the mean effect. While the average variance in beliefs regarding this effect decreases by 20% after the first health exam, the estimated mean of the parameter distribution is small and thus, does not greatly impact smoking behavior.

Simulations of the structural model at the estimated parameter values suggest that health outcomes vary considerably by the intensity with which one smokes. Indeed, whereas some descriptive studies (Doll *et al.*, 1994, 2004) consider only whether one smokes cigarettes, the current paper finds that average life-expectancy is decreased by four and eight years for light ($\leq$ 1 pack/day) and heavy ($>$ 1 pack/day) smoking from age 18 relative to life-long nonsmokers, respectively. Furthermore, my results imply that quitting heavy smoking at ages 30, 40, 50, and 60 years of age increases life-expectancy by approximately 8, 7.75, 7, and 5.5 years, respectively. While these results suggest that there exist life expectancy gains from smoking cessation at any age, they are less severe in their overall assessment of the health effects of smoking than are the unconditional results presented in (Doll *et al.*, 1994, 2004). Indeed, I find that unobserved heterogeneity plays a major role in the dynamic relationship between smoking behavior and health outcomes. My results indicate that there exists a strong correlation between smoking tendencies and underlying factors that influence mortality outcomes.

Finally, consistent with the Becker and Murphy (1988) theory of rational addiction, smoking is found to be reinforcing; that is, the marginal utility of smoking is increasing in the smoking stock. I also find that the costs of withdrawal can prevent individuals from quitting smoking. The reinforcement and withdrawal effects both drive individuals to continue smoking.

This paper proceeds as follows. Section I provides background on the literature of the health effects of smoking, risk perception and subjective expectations as they relate to health, the econometric literature on structurally estimated Bayesian learning models, and preference structures in the context of smoking. Section II describes the structural model. Section III describes the Framingham Heart Survey and provides summary statistics. Section IV discusses identification and measurement issues as well as my econometric approach. Section V presents the main results of the paper and examines potential policy measures. Section VI offers a brief discussion and concludes.

I Background

Cigarette smoking is the single greatest preventable risk factor for mortality and morbidity. According to a 2004 Surgeon General report, cigarette smoking is causally linked to

cancers of the bladder, cervix, esophagus, kidney, larynx, lung, mouth, pancreas, and stomach. Furthermore, there exists a causal relationship between smoking and coronary heart disease, cerebrovascular disease, atherosclerosis, various respiratory diseases, and several reproductive maladies. 440,000 deaths are attributed to smoking in the United States each year. Illness from smoking is estimated to add $157 billion per year to national health expenditures. In short, a 2004 United States Surgeon General report on smoking concludes by stating: "Smoking harms nearly every organ of the body, causing many diseases and reducing the health of smokers in general."[1]

Behind much of the Surgeon General Report's results is the ongoing work of (Doll *et al.*, 1994, 2004). Those authors use survey data of British physicians over several decades to assess the impact of cigarette smoking on mortality. Their findings suggest that smoking cessation at ages 30, 40, 50, and 60 lead to improved life expectancies of 10, 9, 6, and 3 years respectively. Furthermore, life-long smokers face a roughly 25 percentage point increase in the probability of death during middle aged (35-69). However, Doll *et al.* do not control for the endogeneity of smoking with respect to health outcomes nor do they consider the possibility that doctors in different health states may select into smoking. If, independently of smoking, smokers are of a worse overall health status than non-smokers, standard statistical methods may overstate the effect of smoking on mortality. Indeed, only recently have papers in health economics begun to jointly model smoking and health outcomes. For example, Adda and Lechene (2001) show that potential life-span and smoking behavior are correlated along unobserved (to the econometrician) dimensions.

In the current paper, estimation of the primitive parameters of one's decision making optimization problem (e.g., preferences, constraints, and expectation parameters) allows me to assess the impact of smoking on morbidity and mortality outcomes while considering the potential for endogeneity of and dynamic selection into smoking behaviors. In addition, the introduction of serially correlated (permanent) unobserved heterogeneity that affects decision making over the life cycle (including observed initial conditions) allows for the recovery of parameters that measure the impact of smoking and health markers on morbidity and mortality that are free of selection bias. Unbiased estimation of these primitive parameters allows me to simulate the model and impose different patterns of smoking and quitting to examine the resulting changes in predicted health outcomes.

The other main goal of this paper is to assess the importance of health learning within the rational addiction framework of Becker and Murphy (1988). One branch of the rational addiction literature has studied the roles of information, risk perceptions, subjective expectations, and learning in the decision to smoke. For example, Viscusi (1990) models an individual's

[1] Centers for Disease Control and Prevention, 2004. http://www.cdc.gov/tobacco/basic_information/index.htm

beliefs regarding her health risk from cigarette smoking as a Bayesian function of three factors: a prior risk assessment, some measure of risk from experience (perhaps smoking history, age, etc.), and some new information regarding risk. An important question addressed by the literature has been: what exactly is this new risk information?

One type of new information can be categorized as any information that is directed toward a general audience. A widely publicized example was the landmark 1964 United States Surgeon General report that linked smoking to lung cancer and certain birth defects. Luther L. Terry, then Surgeon General, stated that the report "hit the country like a bombshell. It was front page news and a lead story on every radio and television station in the United States."[2] Did this information deter individuals from taking up smoking? Did smokers at the time respond to the report by quitting? On this question, the literature has been mixed. While much of the literature suggests that informational anti-smoking campaigns decrease cigarette demand for light to moderate smokers, Sloan *et al.* (2003) argue that heavy smokers "do not appear to update these perceptions (on the probability of illness/death due to smoking) in response to general information; they need the message to be *personalized*."[3]

Personalized health information may be an important motivator to quit if heavy smokers possess an "it won't happen to me" attitude. Khwaja *et al.* (2006), studying individuals from the Health and Retirement Survey (HRS), show that smokers only "learn" about the risks associated with smoking, as measured by a change in smoking behavior, from a shock to their own health. Those authors argue that if any health shock other than one's own would encourage smoking cessation, it should be that of a spouse. The authors however find no significant effect of spousal health shocks on smoking behavior. "The clear differences in the effects of smoking-related health shocks for current smokers suggest that personalized messages, relevant to their circumstances, are necessary to get their attention and induce changes in their beliefs (qtd. in Sloan *et al.* (2003) pg. 124)."

Nearly all previous work that has examined learning or expectation formation with respect to personalized health messages has studied behavioral changes after a *major* health shock to self or spouse (Smith *et al.*, 2001; Khwaja *et al.*, 2006; Arcidiacono *et al.*, 2007). Additionally, most papers focus on individuals above the age of 50, at which age we begin to observe the major health implications of smoking. I argue that personalized health information does not necessarily have to come in the form of a major health shock after age 50. Indeed, waiting for a major health shock to incite individuals to quit smoking may be too late in terms of life expectancy gains. The current paper examines the extent to which personalized health marker data at prime ages might inform. Using Framingham Heart Survey data,

[2] http://profiles.nlm.nih.gov/NN/Views/Exhibit/narrative/smoking.html

[3] Italics theirs.

Garrison *et al.* (1978) shows that cigarette smoking has a negative impact on high-density lipoprotein (HDL) cholesterol, or "good" cholesterol. Furthermore, due to the nicotine content in cigarettes, other studies have shown that smoking increases both heart rate and blood pressure (Bennett and Richardson, 1984; Omvik, 1996). These health markers, among others, have been directly tied to the risk of cardiovascular disease. To my knowledge, no study has examined the impact of personalized health marker information on the decision to smoke. The extent to which information on these health markers may affect the decision to smoke are central questions of this research. My model builds on the Bayesian learning structural models of other papers in pharmaceutical demand (Crawford and Shum, 2005; Chan and Hamilton, 2006), fertility and infant mortality (Mira, 2007), marketing (Ackerberg, 2003), and health plan report cards (Chernew *et al.*, 2008).

Finally, given the nature of the current paper's learning model, the issue of how to define preferences under varying degrees of uncertainty requires further discussion. The von Neumann-Morgenstern axioms that define the expected utility framework dominate the empirical smoking literature. Under this standard framework, agents are indifferent as to the timing of the resolution of uncertainty. Furthermore, a common criticism when empirically testing hypotheses derived from the expected utility framework has been the inverse relationship between the coefficient of risk aversion and the intertemporal elasticity of substitution. Kreps and Porteus (1978) deviate from the expected utility framework by incorporating preferences over the timing of the resolution of uncertainty. A strand of theoretical and empirical literature has flowed from Kreps and Porteus (1978). (See Kreps and Porteus (1978), Kreps and Porteus (1979), and Epstein and Zin (1991)) In the current paper, uncertainty over the individual specific match value is never resolved, rather just lessened (i.e., reduced posterior variance). Implementing nonexpected utility preferences is currently beyond the scope of this paper. Interesting future work might consider several preference specifications (e.g., expected utility, Kreps Porteus preferences, hyperbolic preferences á la Gruber and Koszegi (2001), etc.) in the context of smoking and addiction.

II Theoretical Model

I specify a dynamic stochastic model of smoking behavior that incorporates learning. This section outlines the basic theoretical model. Given the limitations of the data, changes to the model in the empirical implementation are discussed in section IV. Furthermore, the appendices provide derivations and details of my solution method.

Consider a mixed discrete/continuous-state, discrete-time model of smoking behavior.

The model has a finite horizon in the sense that, while an individual may die prior to period T, the probability of death equals one in period T. A period is indexed by subscript t and is assumed to be one year in length. Each period, a forward-looking individual makes a smoking decision to maximize her lifetime discounted expected utility. Let the decision for individual i be given by $d_{it} = d$, where smoking alternative d is:

$$d = \left\{ \begin{array}{ll} 0 & \text{Do not smoke} \\ 1 & \text{Smoke} \leq 1 \text{ Pack/day} \\ 2 & \text{Smoke} > 1 \text{ Pack/day} \end{array} \right\}$$

The set of factors that influence individual $i's$ smoking decision in period t are given by the state space S_{it}. Define S_{it} as follows:

$$S_{it} = \{A_{it}, R_{it}, \tau_{it}, \psi_{it}, H_{it}, X_{it}\}$$

where A_{it} is individual $i's$ smoking stock entering period t; R_{it} is her health marker index; τ_{it} and ψ_{it} are her mean and variance respectively of her posterior belief distribution; H_{it} is her chronic health status; and X_{it} is her set of demographic characteristics. Additionally influencing behavior, but not listed here, are a preference error ϵ_{it} and a permanent heterogeneity term μ that are both assumed to be known to the individual but *unobserved to the econometrician*. Assumptions about these error terms that aid estimation are discussed in section IV.

At the beginning of period t, an individual undergoes her period t health exam and realizes her period t chronic health state H_{it}, health marker index R_{it}, and smoking stock A_{it}. Using the information from the period t health exam, an individual then updates her beliefs regarding the evolution of future state variables. The smoking decision is then made and utility (to be defined below), as a function of the decision and period t state variables, is realized.

In the subsections below, I expand upon each of the observed state variables and preferences.

II.1 Smoking Stock

Following the rational addiction literature, A_{it} represents the accumulated smoking "stock". The concept of a smoking stock is not immediately intuitive. Broadly speaking, the rational addiction literature treats the stock as a measure of past smoking. Medically, however, we might consider the stock as some accumulation of tar in the lungs that influences health. Alternatively, we might think of the stock as a measure of dependence on nicotine. For the purposes of this paper, the stock may be interpreted as a continuous summary of an individual's smoking history. Here, the extent to which A_{it} influences health is an empirical question to be discussed

in section IV. Formally, the stock is defined as:

$$A_{it} = \left\{ \begin{array}{ll} \exp\left\{\delta_1 \ln(A_{it-1}) + \delta_2 \mathbf{1}[d_{it-1}=1] + \delta_3 \mathbf{1}[d_{it-1}=2] + \rho^A \mu + \eta_{it}\right\} & \text{if } \sum_{n=1}^{t-1} d_{in} > 0 \\ 0 & \text{otherwise} \end{array} \right\} \quad (1)$$

Equation 1 says that individual i's time t smoking stock is normalized to 0 if she has not smoked in *any* previous period. Conditional on any past smoking, the stock is specified as a function of the previous period stock and the previous period decision. δ_1 can be interpreted as one minus the depreciation rate of the stock in percentage terms. The nonlinear investment of light and heavy smoking into the smoking stock are captured by δ_2 and δ_3, respectively. Unobserved permanent heterogeneity is captured by the μ term and its factor loading ρ^A.[4] Also influencing the stock is an i.i.d. white noise term, η_{it}, which is distributed $\mathcal{N}(0, \sigma_\eta)$.[5] Consistent with the interpretation of the stock as a summary of an individual's smoking history, the stock is assumed to be known by the individual in each period.

II.2 Health Marker Index

Define R_{it} as a continuous scalar summary of a variety of health markers (e.g., blood pressure, cholesterol, etc.) that is realized by the individual in each period. Similar to the smoking stock (A_{it}), R_{it} is a scalar representation of numerous health factors. I assume that R_{it} evolves as follows:

$$R_{it} = \zeta R_{it-1} + X_{it}\phi + \kappa_{it} + \rho^R \mu. \quad (2)$$

Here, X_{it} is a vector of sociodemographic characteristics of individual i. I assume that the technology associated with these characteristics (i.e., ϕ) is known by the individual. ζ captures the dynamic aspect of the health markers and is also assumed to be known to the individual. Time invariant and unobserved (to the econometrician) heterogeneity is captured by the μ term and its factor loading ρ^R. Let κ_{it} represent the input from the smoking stock plus an idiosyncratic, i.i.d. error term that is defined as:

$$\kappa_{it} = \theta_i A_{it} + \nu_{it}. \quad (3)$$

Because the individual observes or knows R_{it}, ζ, $X_{it}\phi$, and $\rho^R\mu$, κ_{it} is also observed by the individual. The medical literature suggests that there exists heterogeneity across individuals in the health effects of smoking. I theorize that each individual is endowed with a time invariant, unknown (to both the individual and econometrician) match value θ_i that captures the

[4] See section IV for a discussion of estimation and interpretation issues regarding the permanent unobserved heterogeneity.

[5] Given the exponential stock evolution equation, η is a log normal shock.

idiosyncratic effect of the smoking stock, A_{it} on the health marker index, R_{it}. I assume that θ_i is drawn from a known population distribution given by:

$$\theta_i \sim \mathcal{N}(\overline{\theta}, \sigma_\theta^2).$$

κ_{it} therefore serves as an information signal. Over time, by having health exams and thus observing a sequence of signals, θ_i is learned in a Bayesian fashion. Learning is, however, confounded by the i.i.d. noise term, ν_{it}. Indeed, without ν_{it}, an individual would perfectly learn their match value θ_i at the first health exam (i.e., the first realization of κ_{it}). While ν_{it} is unknown, its distribution is known and given by:

$$\nu_{it} \sim \mathcal{N}(0, \sigma_\nu^2).$$

Because θ_i is time invariant, and because the distributions of θ_i and ν_{it}, as well as the stock A_{it}, are known, over time, an individual can learn their idiosyncratic value of θ_i.[6]

II.3 Learning

Let an individual's period t posterior beliefs, those with which she forecasts future health markers, be given by τ_{it}, her posterior mean, and ψ_{it}, her posterior variance. I assume rational expectations such that an individual's initial belief, prior to any health exams, regarding her true θ_i (the marginal effect of one's smoking history, A_{it}, on health markers, R_{it}) is the population distribution.[7] Initial beliefs ($t = 0$) are:

$$\tau_{i0} = E_0(\theta_i) = \overline{\theta}$$

$$\psi_{i0} = V_0(\theta_i) = \sigma_\theta^2.$$

Expectations about future health marker transitions evolve in the current model with the receipt of personalized health information. In deriving posterior beliefs, consider an individual in period t with smoking stock A_{it}. This individual has two fundamental sources of information: her prior beliefs, $(\tau_{it-1}, \psi_{it-1})$, and the observed results from her period t health exam, κ_{it}. Appealing to the assumption of conjugate prior and signal distributions, the period t beliefs have closed form solutions that are given via Bayes' Rule. The posterior mean and variance are :[8]

$$\tau_{it} = E(\theta_i | \kappa_{it}, A_{it}, \tau_{it-1}, \psi_{it-1}) = \frac{A_{it}^2 \psi_{it}}{\sigma_\nu^2} \hat{\theta}_{it} + \frac{\psi_{it}}{\psi_{it-1}} \tau_{it-1} \tag{4}$$

[6]The assumption that an individual knows the technology of the health production function is ubiquitous in health economics. That is, typically $\theta_i = \overline{\theta}\ \forall\ i$ and $\sigma_\theta = 0$. $\overline{\theta}$ is then estimated and assumed to be the marginal product that all individuals use to solve optimization problems.

[7]The rational expectations assumption is what is typically made in most models of health transitions.

[8]Derivations of these equations can be found in Appendix B.

$$\psi_{it} = Var(\theta_i | A_{it}, \psi_{it-1}, \sigma_\nu) = \frac{\psi_{it-1}\sigma_\nu^2}{A_{it}^2\psi_{it-1} + \sigma_\nu^2}. \tag{5}$$

Here, $\hat{\theta}_{it}$ is the least squares estimate of κ_{it} on A_{it} from the within individual variation of the t^{th} health exam. Note that these beliefs have the following appealing properties. First, the posterior mean is a weighted average of $\hat{\theta}_{it}$ and the original prior mean τ_{it-1}. Second, the weight placed on the period t signal (i.e., $\hat{\theta}_{it}$) is increasing in the smoking stock. Finally, the posterior moments of an individual for whom the stock equals zero (i.e., $A_{it} = 0$) collapse to the prior moments.

II.4 Chronic Health

Let H_{it} represent an individual's overall health state. An individual's overall health state is determined by the presence of any chronic conditions. Let $H_{it} = h$, where outcome h is as follows:

$$h = \left\{ \begin{array}{lll} 1 & \text{if} & \text{Chronic Condition} \\ 0 & \text{if} & \text{No Chronic Condition} \end{array} \right\}$$

What differentiates H_{it} and R_{it} is "reversibility". While R_{it} changes each period, I assume that upon diagnosis of a chronic condition, an individual has the condition forever.[9] Let the probabilities of transiting to different chronic health states in period $t+1$ be:

$$\pi_{it+1}^{h=0} = \left\{ \begin{array}{ll} [1 - P(H_{it+1} = 1 | S_{it}, d_{it}, \mu)] & \text{if } H_{it} = 0 \\ 0 & \text{if } H_{it} = 1 \end{array} \right\}$$

$$\pi_{it+1}^{h=1} = \left\{ \begin{array}{ll} P(H_{it+1} = 1 | S_{it}, d_{it}, \mu) & \text{if } H_{it} = 0 \\ 1 & \text{if } H_{it} = 1 \end{array} \right\}.$$

Define the relevant probability, $P(H_{it+1} = 1 | S_{it}, d_{it}, \mu)$, with the following binary logit equation:

$$\frac{exp(\lambda_0 + \lambda_1 R_{it} + \lambda_2 R_{it}^2 + \lambda_3 \mathbf{1}[1980s] * R_{it} + \lambda_4 \mathbf{1}[1990s] * R_{it} + [\lambda_5 + \lambda_6 R_{it}] * d_{it} + \lambda_7 X_{it} + \rho^H \mu)}{1 + exp(\lambda_0 + \lambda_1 R_{it} + \lambda_2 R_{it}^2 + \lambda_3 \mathbf{1}[1980s] * R_{it} + \lambda_4 \mathbf{1}[1990s] * R_{it} + [\lambda_5 + \lambda_6 R_{it}] * d_{it} + \lambda_7 X_{it} + \rho^H \mu)} \tag{6}$$

Here, R_{it} is the health marker index defined above, X_{it} is a vector of exogenous individual characteristics, d_{it} is the smoking choice and μ is an individual, time invariant unobserved heterogeneity term. The factor loading superscript H simply differentiates it from other factor loadings in the model. λ_6 and λ_7 capture changes over time in how health markers affect the probability of chronic disease incidence (perhaps due to advances in medical technology, pharmaceuticals, etc.).

In forecasting future chronic health transitions, I follow the literature and assume that an individual has rational expectations and that she understands the technology associated

[9] This assumption captures the fact that upon having an heart attack, for example, an individual is in a fundamentally different health state even if they don't have repeated heart attacks (Khwaja *et al.*, 2006).

with the chronic health transition probability. A natural question becomes, why do individuals in the model learn about how smoking affects health markers but not chronic conditions? By modeling learning about the effect of A_{it} on R_{it}, however, individuals are indirectly updating their expectations about future chronic health transitions because the health marker index enters the chronic health transition probability. Furthermore, the purpose of this paper is to explore the importance of health information prior to major health shocks. Imposing that individuals understand the technology (i.e., the λs) associated with covariates in the chronic health transition equation is the standard approach. While future work may incorporate learning about health transition probabilities, such learning is currently beyond the scope of this paper.

II.5 Mortality

While an individual may die prior to period T, death is assumed to occur with probability one in period T. Define an indicator for death at the end of period t, M_{it+1}=1, and let its corresponding probability, $\varsigma_{it+1} = P(M_{it+1} = 1|S_{it}, d_{it}, \mu)$, be given by:

$$\frac{exp(\omega_0+\omega_1 R_{it}+\omega_2 R_{it}^2+\omega_3 H_{it+1}+[\omega_4+\omega_5 R_{it}+\omega_6 H_{it+1}]*d_{it}+\omega_7 \mathbf{1}[1980s]*H_{it+1}+\omega_8 \mathbf{1}[1990s]*H_{it+1}+\omega_9 X_{it}+\rho^M \mu)}{1+exp(\omega_0+\omega_1 R_{it}+\omega_2 R_{it}^2+\omega_3 H_{it+1}+[\omega_4+\omega_5 R_{it}+\omega_6 H_{it+1}]*d_{it}+\omega_7 \mathbf{1}[1980s]*H_{it+1}+\omega_8 \mathbf{1}[1990s]*H_{it+1}+\omega_9 X_{it}+\rho^M \mu)} \quad (7)$$

Here, H_{it+1}, is individual $i's$ chronic health state at the end of period t.[10] Again, the superscript on the factor loading simply differentiates it from other factor loadings. The technology for the death transition equation is assumed to be known by the individual. ω_8 and ω_9 capture the fact that, conditional on having some chronic illness, the probability of death from that illness may have changed over time due to medical advances. Furthering the discussion above, because the health marker index enters the death transition equation directly (and indirectly through the chronic health term H_{it}), individuals are indirectly updating their expectations about death transitions conditional on their smoking choice through the learning process. Assuming that the ωs are known by the individuals is the standard approach and one that can be relaxed in future work.

[10]The timing convention here is due to data aggregation. Clearly, any chronic health event occurring in period t must occur at or before the time of death, if death also occurs in t. Therefore, to accommodate the frequent observation in the data of an individual dying from a chronic health event, the appropriate chronic health data point in this equation is H_{it+1}.

II.6 Preferences

Following the standard expected utility framework, the deterministic portion of per period utility associated with health state h, $(h = 0, 1)$, and smoking alternative $d_{it} = d$ is:

$$\begin{aligned}\overline{U}_{it}^{h}(A_{it}, d_{it} = d, R_{it}, X_{it}, \mu) = \quad & \alpha_{0h} + (\alpha_{1h} + \alpha_{2h}A_{it} + \alpha_{3h}R_{it} + \alpha_{4h}Age_{it}) * [d_{it} = 1] \\ & + (\alpha_{5h} + \alpha_{6h}A_{it} + \alpha_{7h}R_{it} + \alpha_{8h}Age_{it}) * [d_{it} = 2] \\ & + \alpha_{9h} * [d_{it-1} \neq 0] * [d_{it} = 0] + \alpha_{10h}A_{it} + \alpha_{11h}A_{it}^{2} + \rho^{Uhd}\mu \end{aligned} \tag{8}$$

The specification accommodates any nonlinearity in the effects of light and heavy smoking on utility. While $\alpha_{1.}$ ($\alpha_{5.}$) is the direct marginal utility of light (heavy) smoking, $\alpha_{2.}$ ($\alpha_{6.}$) captures the extent to which past consumption reinforces current consumption. $\alpha_{2.}$ ($\alpha_{6.}$) captures a part of the intertemporal trade-off in utilities. The extent to which the health marker index affects the marginal utility of smoking is captured by $\alpha_{3.}$ ($\alpha_{7.}$). Note that higher values of R_{it} and A_{it} imply worse health and a higher smoking stock respectively. The sign and magnitudes of $\alpha_{2.}$ $\alpha_{3.}$, $\alpha_{6.}$, and $\alpha_{7.}$ are empirical questions. $\alpha_{4.}$ ($\alpha_{8.}$) captures changes in the marginal utility of smoking across the lifespan. Specific withdrawal costs from quitting, which also capture part of the intertemporal utility trade-off, are captured by $\alpha_{9.}$. Finally, $\alpha_{10.}$ and $\alpha_{11.}$ capture tolerance in smoking. That is, the extent to which a given level of stock affects utility is captured here regardless of smoking behavior.

Relative preferences over smoking alternatives hinge on two main factors. First, preferences vary by the chronic health state ($H_{it} = h$). The extent to which the marginal utility of smoking varies across chronic health states remains an open question. Generally, the marginal utility of consumption of any normal good is thought to be lower in worse health states (Viscusi and Evans, 1990; Gilleskie, 1998). If however smoking provides relaxation and comfort when stricken with a chronic illness, the overall marginal utility of smoking may be larger in worse health states. Estimation of the structural parameters will therefore empirically test for the sign of the marginal utility of smoking across health states. Second, as seen in equations 1, 2, and 3, current period smoking affects the size of the next period smoking stock, which in turn affects the next period health marker index and next period utility. Given the dynamic nature of the model, individuals evaluate smoking alternatives while considering the future marginal utility of smoking as well as the future consequences of a higher A_{it}.

Following Rust (1987), let the total current period utility be the sum of the deterministic utility from equation 8 and an additive i.i.d. preference shock that is alternative and health-state specific:

$$U_{it}^{h}(A_{it}, d_{it} = d, R_{it}, X_{it}, \mu, \epsilon_{it}) = \overline{U}_{it}^{h}(A_{it}, d_{it} = d, R_{it}, X_{it}, \mu) + \epsilon_{it}^{dh}.$$

In the empirical implementation below, ϵ_{it}^{dh} is simply an additive econometric error; however, in the theoretical model, ϵ_{it}^{dh} is given a structural interpretation as an unobserved state variable (Aguirregabiria and Mira, 2010). The alternative specific lifetime value function in health state h, conditional on unobserved heterogeneity μ, is:

$$V_d^h(S_{it}, \epsilon_{it}^{dh}|\mu) = \overline{U}_{it}^h(A_{it}, d_{it} = d, R_{it}, X_{it}, \mu) + \epsilon_{it}^{dh} + \beta\big[(1-\varsigma_{it+1})\sum_{a=0}^{1}\pi_{it+1}^a E_{it}[V^a(S_{it+1}|\mu)|d_{it} = d]\big].$$

Here, $V^a(S_{t+1}|\mu)$ is the maximal expected lifetime utility of being in health state a in period $t+1$. The value function is conditional on the unobserved heterogeneity component μ. The expectation operator is taken over the time t posterior distribution of θ, as well as other shocks that determine future state variables and preference shocks. Given the unitary dimension of the posterior distribution, as well as the i.i.d. nature of other shocks to the model, I use a Monte Carlo method to evaluate the expectation within solution to the model .[11] Let $\overline{V}_d^h(\cdot) = V_d^h(\cdot)$ - ϵ_{it}^{dh}. If we assume that ϵ_t^d has an Extreme Value Type I distribution, then the maximal (EMAX function) expected lifetime utility has the following closed form solution:

$$V^h(S_{it+1}|\mu) = EC + ln\Big(\sum_{d=1}^{D} exp\big(\overline{V}_d^h(S_{it+1}|\mu)\big)\Big) \quad \forall t, \quad \forall h. \tag{9}$$

Here, EC is Euler's constant. Furthermore, because the error term ϵ_t^d is additively separable, the conditional choice probabilities take the following dynamic multinomial logit form:

$$p(d_{it} = d|S_{it}, \mu) = \frac{exp\big(\overline{V}_d^h(S_{it}|\mu)\big)}{\sum_{d=0}^{2} exp\big(\overline{V}_d^h(S_{it}|\mu)\big)} \ \forall t, \forall h \tag{10}$$

To preview the empirical implementation, the conditional smoking choice probability in equation 11 enters the likelihood function. The parameters that dictate the choice probability are structural in the sense that they are follow from the above maximization problem. Also to enter the likelihood function are the health and death transition probabilities, as well as health marker index and smoking stock transition equations.

III Data: The Framingham Heart Survey

The Framingham Heart Survey is one of the longest running panel studies in the world. With the stated goal to "identify the common factors that contribute to cardiovascular disease", the survey contains repeated observations of individuals over a 50 year period.[12] Beginning in 1948, the Framingham Heart Survey began collecting biennial health data from 5107 individuals living in Framingham, Massachusetts. These individuals formed what became known

[11] See Appendix A for additional details.

[12] The Framingham Heart Survey: http://www.framinghamheartstudy.org/index.html

as the Original Cohort. In 1971, the Framingham Heart Survey began following the offspring of the Original Cohort to form the Offspring Cohort. Each cohort represents a different panel study that has continued into the 21st century. The main drawback of these data is that all participants in the survey are from Framingham. Therefore, there is no geographic, and limited demographic, variation. Another drawback of these data is the lack of income measures. However, the data contain a wealth of health and smoking information that are ideal for analyzing the trade-off between smoking and the potential for future health shocks.

The structural model above is estimated with data from the Framingham Offspring Cohort.[13] The decision to focus on data from the Offspring Cohort stems from the consistency with which the health exams were administered. Smoking and health questions changed over time in the Original Cohort; thus, constructing uniform measures of smoking history, per-period behavior, and health variables (especially health markers) proved to be difficult. In constructing the sample used in estimation, I drop all individuals with a missing exam and all those lost to attrition.[14] Table 1 explains my process of sample construction.[15] The final sample consists of 19,461 person/year observations.

Table 1: Sample Construction

N	Description
4989	Framingham Heart Survey Offspring Cohort Participants - Restricted Sample
3730	Sample after dropping those individuals that skipped one or more of the health exams
3008	Sample after dropping all person/year observations of individuals who attrit

3008 unique individuals yields 19461 person/year observations.

Source: The Framingham Heart Survey, Offspring Cohort.

III.1 Sample Statistics

The sample statistics given in this section are by Framingham Heart Survey exam. Individual range in age from 13 to 62 at the first exam (between 1971 and 1975). Offspring Cohort health exams have been administered at roughly four year intervals from 1971 to the present.

[13] In another study, I am examining the intergenerational transfer of smoking preferences between Original and Offspring Cohort participants.

[14] 671 individuals are lost to attrition (i.e., some reason other than death) at some point during the seven exams. This constitutes approximately 18% of my sample. The decision to drop these individuals is based on the computational tractability of modeling attrition. Simple t-tests for difference of means suggest that those that attrit are slightly more likely to be women, have a three point lower level of systolic blood pressure on average, and have a statistically insignificant difference in coronary heart disease incidence than their nonattriting counterparts. Those that attrit are on average slightly more likely to smoke.

[15] The full sample contains 5124 individuals. For this work, I only have access to data for those individuals from whom consent for distribution was granted.

While variation in the timing of the health exams may seem detrimental to implementing the structural model above, as I discuss in section IV, in estimation I exploit this variation to help identify key parameters of the model. Indeed, section IV provides the majority of the sample statistics for data used in estimation. In the empirical implementation of the model, I expand these data to reflect the yearly decision making model presented above. I therefore postpone the presentation of smoking and health transition statistics by age until the empirical implementation description in section IV.

I have data for each individual in the Framingham Heart Survey Offspring Cohort for up to seven health exams. Initial health exams were conducted between 1971 and 1975. For each participant, subsequent exams occurred at varying time intervals. Table 2 provides information on the average timing of each exam across individuals, in addition to demographic information. Because attrition has been eliminated, the number of individuals at each exam

Table 2: Sample Characteristics by Exam

Exam	Mean Year	Mean Age	St. Dev. Age	% Female	% Married	# Individuals
1	1973	37.0	(10.28)	50.0	80.5	3008
2	1981	44.3	(10.05)	50.1	82.9	2921
3	1985	48.3	(9.99)	51.1	83.0	2849
4	1988	51.5	(9.99)	51.5	80.6	2796
5	1992	55.0	(9.83)	52.1	79.9	2709
6	1996	58.6	(9.69)	52.7	77.2	2613
7	1999	61.5	(9.58)	53.1	74.7	2565

Ages in the sample range from 13 in exam 1 to 88 in exam 7.

reflects only those that have survived. Over the health exams, the sample becomes slightly more weighted toward female and non-married individuals. For confidentiality reasons, all survey participants are white. Table 2 also shows the great variability in ages across the sample. At the first exam, there are individuals who are as old as the average age at the final exam. Indeed, over the entire sample, ages range from 13 to 88. Table 3 gives sample percentages of the maximum number of years of education by category. The sample reflects a rather well educated cohort for the time period. Nearly 89% of the sample has a high school degree or better.

Table 4 breaks down the sample by smoking prevalence over exams. Over the seven exams, the sample smoking prevalence drops from roughly 41% to 11%. Interestingly, at the first exam, smoking prevelence in the sample is roughly consistent with that of the United States average prevelence (37% of Americans smoked in 1973). However, by the final exam, the sample percentage of smokers has decreased to roughly 11% whereas the national average fell to only 23.3%. The sample is also clearly older than the general population by the seventh

Table 3: Education

Education Years	% of Sample
0-4	3.2%
5-8	1.0
9-11	6.1
12	32.8
13-16	43.2
17+	13.8

$N = 3008$. Percentages reflect highest attained level of education.

exam.[16]

Table 5 shows the percentage of the sample living with a chronic condition at each health exams. I define a chronic condition to include a wide variety of cardiovascular diseases (e.g.,

Table 4: Smoking Behavior by Exam

Exam	Nonsmokers	Light Smokers $\leq$ 1 Pack/Day	Heavy Smokers $>$ 1 Pack/Day
1	59.0%	26.7%	14.3%
2	61.3	24.4	14.2
3	77.2	14.3	8.5
4	81.2	12.8	6.0
5	85.2	11.0	3.9
6	87.9	9.5	2.6
7	88.9	8.7	2.5

coronary heart disease, myocardial infraction, cerebrovascular accidents, congestive heart failures, etc.) and cancers (lung, larynx, tongue,esophagus, etc.). The decision to aggregate the data to this level stems from the computation burdens of estimating additional parameters in the structural model. As in the theoretical model, I assume that upon transiting to a chronic health state, an individual remains in that state for life. The incidence of new chronic conditions is in column 3 of table 5.

IV Empirical Implementation

In implementing the model described in section II, there are four main hurdles. First, state variables A_{it} and R_{it} must be constructed from the Framingham data in such a way as to capture an individual's smoking history and health markers respectively. For each of these variables, I employ principal component analysis in a method similar to that of Sickles and

[16] Centers for Disease Control and Prevention: http://www.cdc.gov/tobacco/basic_information/index.htm

Table 5: Chronic Health by Exam

Exam	Chronic Condition at Exam	Newly Chronically Ill at Exam
1	0.2%	0.0%
2	4.0	3.8
3	7.0	3.4
4	9.6	3.1
5	12.2	3.9
6	16.0	5.0
7	20.3	4.8

Williams (2008). The second hurdle lies in the timing of the health exams. While the data contain only (at most) seven exams over a 40-year period, the theoretical model is based on a yearly decision making process over a finite time horizon. As explained below, I exploit retrospective questions in the data to construct a dataset that mirrors the timing of the model. I use predictions from solution to the model to integrate over "off years" and to explain the initial condition for each individual. The third hurdle involves modeling permanent unobserved heterogeneity and resolving any initial conditions problems. A final hurdle is identification of the key parameters of the model.

IV.1 Continuous State Variable Construction

For the model to both remain computationally tractable and be consistent with the assumption of conjugate distributions, I need continuous, scalar representations of both an individual's smoking history and her health markers. My solution is to employ principal component analysis (PCA) in the construction of each variable. PCA is a nonparametric technique that summarizes the total variation in a set of variables into an ordered set of continuous, scalar principal components. The first principal component is constructed as a linear combination of data and of factor loadings from the highest eigenvalue eigenvector from an eigenvector decomposition of the variables' correlation matrix.[17] The trade-off with PCA is accuracy and completeness. Only considering the first principal component implies that any remaining variation in the data (i.e. the second, third, fourth, etc. principal components) is lost. In the context of most structural models, however, reducing the dimension of the data is clearly adventageous.

[17] For example, given a set of k variables, employing PCA will yield k principal components. If, however, the first two principal components account for 70% of the total variation in the k variables, and $k > 2$, the researcher may find it adventagous to only use the first two principal components as data.

In constructing R_{it}, the health marker index, I use PCA with the following (standardized) variables: systolic blood pressure, diastolic blood pressure, total cholesterol, high-density lipoprotein (HDL) cholesterol, and a diabetes dummy.[18] These health markers are identified by the Framingham Heart Survey as significant predictors for coronary heart disease (Wilson *et al.*, 1998).[19] Table 6 provides summary statistics of these variables at each health exam.

Table 6: Health Markers by Exam

Exam	Systolic Blood Pressure	Diastolic Blood Pressure	High-Density Lipoprotein (HDL)	Total Cholesterol	Diabetes
1	122.3	79.0	50.8	197.1	1.8%
	(16.1)	(10.7)	(15.0)	(39.4)	
2	122.3	78.2	48.6	203.9	2.6
	(16.5)	(9.8)	(13.6)	(39.1)	
3	123.7	79.0	51.1	212.2	3.5
	(16.7)	(9.6)	(14.8)	(41.2)	
4	126.8	79.1	49.9	207.5	4.8
	(18.8)	(10.0)	(14.8)	(38.5)	
5	126.4	74.5	49.9	205.6	6.9
	(18.8)	(10.1)	(15.2)	(36.5)	
6	128.2	75.3	51.0	206.0	9.6
	(18.4)	(9.5)	(16.0)	(37.7)	
7	127.2	73.8	53.3	200.6	11.1
	(18.7)	(9.7)	(16.8)	(36.6)	

Standard deviations are in parenthesis.

The first principal component of these variables explains approximately 33% of the total variation. Admittedly, this is not high. I do however now have a continuous index of health markers. I see two main justifications for using the first principal component as my measure for the health markers. First, the theory places no restriction on the amount of information that R_{it} must convey, only that it conveys some information. Any computationally tractable definition of R_{it} will have to be an approximation. That I can explain a third of the variation in the variables that the medical literature view as signficant will at least inform to some degree. Second, most papers that use PCA use first principal components that explain between $20\% - 40\%$ of the total variation.[20]

To provide intuition as to the weights used to create the health index, Table 7 presents

[18] PCA is most effective when there exists significant correlation between the variables. As one might expect, the correlation between these health markers is high.

[19] The Framingham Heart Survey: Risk Score Profiles:http://www.framinghamheartstudy.org/risk/index.html

[20] In the context of socioeconomic indices, see Vyas and Kumaranayake (2006) for a good overview of PCA.

results from an OLS regression of the health index on the above health markers.[21] In

Table 7: Health Index Regression

Variable	Coefficient		(Std. Err.)
Systolic Blood Pressure	0.657	***	(0.000)
Diastolic Blood Pressure	0.637	***	(0.000)
Total Cholesterol	0.306	***	(0.000)
High-Density Lipoprotein (HDL)	-0.177	***	(0.000)
Diabetes	0.193	***	(0.000)

Significance Levels: ***1% Level, **5% Level, *10% Level

this context, the continuous health index can be interpreted as a measure of bad health (i.e. higher values of the index imply worse overall health). Note in Table 7 that only HDL, or "good" cholesterol, negatively affects the health index.

As discussed above, the smoking stock summarizes all past smoking decisions prior to period t. Again using PCA, I define the index A_{it} as the first principal component of the following four standardized variables: total number of years smoking at time t (experience), number of years smoking at time t since last year not smoking (tenure), number of years at time t not smoking since last year smoking (cessation), and the intensity of smoking in the previous period, $t-1$.[22] I term these variables experience, duration, cessation, and intensity respectively. Table 8 gives sample averages by exam of the number of years of duration, tenure and cessation. The first principal component explains nearly 52% of the total variation in these four variables.

To aid in interpretation of both the resulting smoking stock and the associated parameters to be estimated, I normalize the smoking stock as follows. First, I run PCA on just those with some smoking history. That is, individuals with any observed or reported past smoking in each period are included in the PCA. For example, if an individual takes her first exam at age 18 and begins smoking at age 22, all observations from this individual after age 22 are included in the PCA, whereas observations prior to 22 are not included. Second, I shift the distribution of the resulting index such that the person with the lowest value has a stock approximately equal to zero. Finally, for individuals with no smoking history, I assign a stock value of zero. Table 9 reports the results of a regression of the smoking stock index on the four variables of interest (excluding a constant).[23] Notice that while experience, duration,

[21] The regression is run without a constant. Point estimates correspond to the eigenvector values from the first principal component for the corresponding variables.

[22] Intensity is measured as the average number of cigarettes per day. Each of these smoking variables is measured as the value entering the examination.

[23] Point estimates correspond to the eigenvector values from the first principal component for the corresponding variables.

and smoking intensity of an individual all increase the stock index, cessation from smoking decreases the stock. I therefore interpret higher values of the index as more accumulated smoking stock capital.

Table 8: **Smoking History by Exam**

Exam	Experience	Duration	Cessation
1	10.6	7.3	1.9
2	13.0	7.9	4.1
3	14.1	7.6	4.8
4	14.7	6.2	6.1
5	15.1	5.3	7.9
6	15.1	4.4	9.8
7	15.3	3.8	11.3

Table 9: Smoking Stock Regression

Variable	Coefficient	(Std. Err.)
Experience	0.309***	(0.000)
Duration	0.589***	(0.000)
Cessation	-0.517***	(0.000)
Intensity	0.540***	(0.000)

Significance Levels: ***1% Level, **5% Level, *10% Level

IV.2 Solution and Timing

While I only observe individuals at seven health exams over a 40-year period, the theoretical model is based on a yearly decision making process. To reconcile this difference, I proceed in the following steps. First, in solution to the model, I specify the final period, T, to be at age 100. That is, the probability of death at the end of period T equals one. The yearly model is then solved recursively back to age 7, at which point I assume that all individuals have a smoking stock of zero (i.e. $A_{i7} = 0$, $\forall i$).[24] Second, the data from section III are expanded based upon retrospective questions. With the exception of the health marker information needed to construct the health marker index, R_{it}, data are available to construct a yearly dataset from age seven until an individual either dies or completes their seventh exam.[25] Figure 1 shows the sample probabilities for each smoking choice by age. Furthermore, Table 10 shows general

[24]The solution appendix gives greater detail on my solution method

[25]Data in years prior to an individual's first exam were constructed based on questions at the first and second exams that asked, if applicable, the first age at which one started smoking and the age at which one stopped smoking. For later years in between health exam years, smoking data were imputed based on history and adjacent health exam data. Specific dates are available in the data for any chronic health and mortality events.

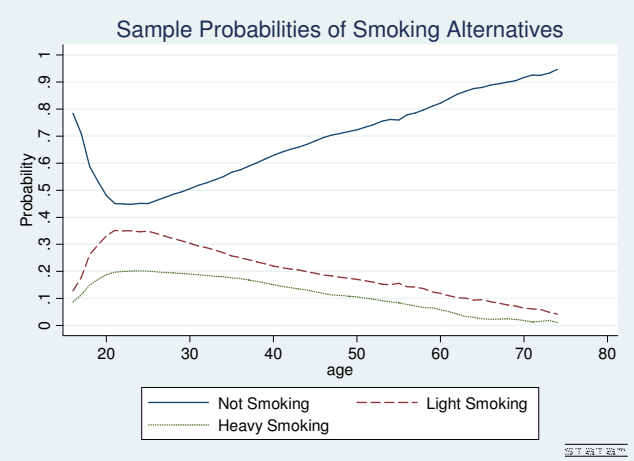


Figure 1: Sample Choice Probabilities by Age

smoking summary statistics from the expanded data. Table 11 reports smoking behavior transitions around three events: health exams, chronic health shocks, and transiting to a health marker index at or above the 75th percentile. Conditional upon an individual's smoking behavior one period prior to each event, the table reports percentages in each smoking option one and three years after the event. For example, of those individuals smoking heavily one

Table 10: Smoking Summary Statistics

	Mean (Median)	**S.D.**	**Min**	**Max**
First Age Smoking (Conditional on Ever Smoking)	19.57 (18)	7.45	7	67
Total Years Smoking (Conditional on Ever Smoking)	24.78 (24)	14.07	1	68
Tenure Smoking (Years) (Conditional on Ever Smoking)	21.13 (19)	14.85	1	68
Last Age Smoking (Conditional on Ever Smoking and Quitting)	44.76 (45)	12.84	13	76

period prior to a chronic health shock, 70.24% continued to smoke heavily one period after the shock and 47.06% were still smoking heavily three periods after the shock. The table provides at least antidotal evidence that each event alters smoking behavior. I compare these descriptive statistics with predictions from the model in section V.

I conduct the smoking stock PCA from the preceding section on yearly experience, duration, cessation, and intensity questions and thus, generate a yearly stock variable. Conditional on parameters and data from years with a health exam, I integrate over the health marker index distribution in solution to the model in "off years" (i.e. those years with no health exam and thus no health marker information). Consistent with the learning process in the model, I only allow updating of the posterior distribution in periods in which an individual had a health exam. Therefore, from the individual's perspective, when forecasting future health, an individual uses the updated posterior distribution from the most recent health exam.

Table 11: Observed Transitions

	Behavior One Period Post			Behavior Three Periods Post		
Behavior One Period Prior	Not Smoking	Light Smoking	Heavy Smoking	Not Smoking	Light Smoking	Heavy Smoking
Transitions Around Health Exams						
Not Smoking	96.38%	3.20%	0.42%	97.00%	2.55%	0.45%
Light Smoking	21.40	70.29	8.30	29.45	63.50	7.06
Heavy Smoking	14.50	16.47	69.03	20.62	16.48	62.90
Transitions Around Chronic Health Shocks						
Not Smoking	98.94	0.89	0.18	97.18	2.59	0.24
Light Smoking	24.78	73.45	1.77	39.36	54.26	6.38
Heavy Smoking	21.43	8.33	70.24	39.71	13.24	47.06
Transitions Around Health Marker Index Shocks: 75^{th} Percentile						
Not Smoking	97.49	2.33	0.18	98.02	1.79	0.20
Light Smoking	24.52	62.50	12.98	33.33	54.87	11.79
Heavy Smoking	17.12	23.29	59.59	25.18	23.74	51.08

IV.3 Permanent Unobserved Heterogeneity and Initial Conditions

Permanent unobserved heterogeneity enters the model in a linear fashion through the μ term and the associated factor loadings. The factor loadings allow for a different effect of the unobserved μ term everywhere it enters. Rather than placing a distributional assumption on the underlying unobserved heterogeneity, I approximate its distribution with a step function and estimate the factor loadings and mass point probabilities with other parameters in the model (Heckman and Singer, 1984). This discrete factor method has been shown to approximate both Gaussian and non-Gaussian distributions well (Mroz, 1999).

I first observe individuals at various points in their life cycle (i.e., different ages at the first health exam) and with a variety of health histories. Failing to properly model these histories would lead to an initial conditions problem. Furthermore, the initial conditions problem may lead to an issue of dynamic selection into smoking behaviors. That is, individuals in some permanently lower (unobserved) health state may select into smoking. However, solution to the model generates individual probabilities of choice behavior and health/death transitions for all ages beginning at age seven. Recall that data exist for all smoking, chronic health, and death events from age seven until either death or the final health exam (exam 7 in the data). At age seven, I assume that each individual has a smoking stock of zero and has no chronic health problems. The only remaining initial condition is the initial health marker index upon entering the sample. Using the model, I can simulate a health marker index for each period from age seven until the first observed health exam. Hence, I use the model to generate probabilities of an individual's health history when they are first observed in the sample. (Khwaja,

2010).[26] Furthermore, individual variation in the data at the first exam (the initial condition) helps to identify parameters of the model.

IV.4 Identification

The following sets of parameters are estimated.

$$
\begin{array}{ll}
\text{Utility Parameters:} & \Theta_U = \left\{\alpha_{0h}, \ldots, \alpha_{11h}\right\}_{h=0}^{1} \\
\text{Health Transition Parameters:} & \Theta_H = \left\{\lambda_0, \ldots, \lambda_{10}\right\} \\
\text{Death Transition Parameters:} & \Theta_M = \left\{\omega_0, \ldots, \omega_{12}\right\} \\
\text{Smoking Stock Parameters:} & \Theta_A = \left\{\delta_1, \delta_2, \delta_3, \sigma_\eta\right\} \\
\text{Learning and Risk Parameters:} & \Theta_R = \left\{\overline{\theta}, \sigma_\theta, \sigma_\nu, \phi, \zeta\right\} \\
\text{Factor Loadings:} & \Theta_\rho = \left\{\left\{\left\{\rho^{Uhd}\right\}_{h=0}^{1}\right\}_{d=0}^{2}, \rho^H, \rho^M, \rho^R, \rho^A\right\}
\end{array}
$$

Additionally, I estimate the probability weights of the mass points for the discretized distribution of the permanent unobserved heterogeneity, μ. Let $\Theta = \left\{\Theta_U, \Theta_H, \Theta_M, \Theta_A, \Theta_R, \Theta_\rho\right\}$. In order to identify the preference parameters, I normalize the utility of death to be zero. Relative to this normalization, identification of the preference parameters comes mainly from variation in smoking behavior and health and death transitions over time. For example, different smoking choices across in smoking stock, health marker index, and age levels identifies the interaction preference parameters. Furthermore, while the withdrawal parameter, $\alpha_{9\cdot}$, is identified off of variation in the choices of individuals after a period of smoking, the direct impacts of the stock on utility, $\alpha_{10\cdot}$ and $\alpha_{11\cdot}$, reflect tolerance in smoking and are identified by individuals that progress from light to heavy smoking.

In the absence of subjective expectation data, the structure of the model is needed to identify the presence of learning. Mira (2007) notes that learning can no more be identified than can rational behavior. If, however, the prior distribution of beliefs is proved to be degenerate (i.e., if the null hypothesis that $\sigma_\theta = 0$ is not rejected), then the results would suggest an absence of learning. The identification strategy of the specific learning parameters is therefore quite subtle. While identification of $\overline{\theta}$ comes from variation in the smoking stock and health marker index, variation in smoking by individual over time identifies σ_θ (Crawford and Shum, 2005). If, indeed, individuals are learning over time, choices at the end of the time frame relative to the beginning should better reflect an individual's true match value, θ_i. An additional source of variation that helps to identify the learning parameters is the variation

[26] Individuals enter the sample aged between 13 and 62 years. At age seven, I assume that the lagged value of the health marker index is in the 90th percentile (e.g., good health) of each health marker that is used to construct the index. I then use the weights from the principle component analysis to construct the lagged value. Recall that the simulated health marker index is scaled by demographic characteristics, X_{it}, as well as the unobserved heterogeneity, μ, term and its factor loading.

across individuals in the timing of health exams. There exists considerable variation in the number of years between exams across individuals; thus, two similar individuals that receive health information at different frequencies may develop different smoking patterns. Finally, the identification of chronic health and death transition parameters comes from variation in the state variables and the observed incidence of chronic health and death.

IV.5 Likelihood Function

Consider first the contribution of individual i to the likelihood function. Given that $\eta_{it} \sim \mathcal{N}(0, \sigma_\eta^2)$ and $\nu_{it} \sim \mathcal{N}(0, \sigma_\nu^2)$, we can express the probability density functions of A_{it} and R_{it} respectively as:

$$\Lambda_{it} = f(\eta_{it}|A_{it-1}, d_{it-1}, \mu, \rho_A, \Theta_A) = \frac{1}{\sigma_\eta}\phi\left([\log A_{it} - \delta_1 \log A_{it-1} - \delta_2 d_{t-1} - \rho_A \mu]/\sigma_\eta\right) \quad (11)$$

and

$$\Omega_{it} = g(\nu_{it}|X_{it}, \kappa_{it}, \mu, \rho_R, \Theta_R) = \frac{1}{\sigma_\nu}\phi\left([R_{it} - \zeta R_{it-1} - X_{it}\phi - \kappa_{it} - \rho^R \mu]/\sigma_\nu\right) \quad (12)$$

where $\phi(\cdot)$ is the standard normal distribution. Recall, however, that the health marker index, R_{it}, and only the health marker index, is unobserved in periods in which a health exam was not taken.[27] I must, therefore, integrate over R_{it} in all periods with no health exam. For ease of exposition, define the dummy y as follows:

$$y_{it} = \left\{ \begin{array}{ll} 1 & \text{if An exam was taken in year t} \\ 0 & \text{if No exam was taken in t} \end{array} \right\}$$

Define $Z_{it}^{y=1}|\mu$ as individual $i's$ likelihood contribution in period t when $y_{it} = 1$ and conditional on unobserved heterogeneity term μ:

$$Z_{it}^{y=1}|\mu = \prod_{d=0}^{2}\left(p(d_{it} = d|s_{it}, \mu) * \Lambda_{it} * \Omega_{it} * \prod_{h=0}^{1}\left(\pi_{it+1}^{dh}|\mu\right)^{\mathbf{1}[H_{it+1}=h]} * \prod_{m=0}^{1}\left(\varsigma_{it+1}^{dm}|\mu\right)^{\mathbf{1}[M_{it+1}=m]}\right)^{\mathbf{1}[d_{it}=d]}. \quad (13)$$

Here, π_{it+1}^h represents the probability of transiting to health state h in period $t+1$ and ς_{it+1}^m is the probability of transiting to death state m in period $t+1$. Unless a health exam was taken in the period directly before t, the lagged value of the health marker index in equation 13 is unobserved. In practice, I use the expected health marker index given the model parameters

[27] All right-hand side terms in the health marker equation *are* observed in these "off" years due to retrospective questions and/or imputation with the exception of the lagged value of the health marker index when the previous period did not contain a health exam.

as the lagged value. In periods in which $y_{it} = 0$, define the expected health marker index, conditional on the model parameters as:

$$\widetilde{R}_{it} = \mathrm{E}_{\nu}(R_{it}|\Theta_R, S_{it}) \tag{14}$$

Here, the expectation operator is taken over the i.i.d. noise term, ν. Other probabilities in the model are conditional on $\widetilde{R}_{it}$ for years in which $y_{it} = 0$.[28] In the period directly after a health exam, the lagged value of the health marker index (i.e., from the exam and not the simulated term) is used in the construction of $\widetilde{R}_{it}$. Therefore, define $Z_{it}^{y=0}|\mu$ as individual $i's$ likelihood contribution in period t when $y_{it} = 0$:

$$\prod_{d=0}^{2}\left(p(d_{it}=d|s_{it},\widetilde{R}_{it},\mu)*\Lambda_{it|\widetilde{R}_{it}}*\prod_{h=0}^{1}\left(\pi_{it+1}^{dh}|\widetilde{R}_{it},\mu\right)^{\mathbf{1}[H_{it+1}=h]}*\prod_{m=0}^{1}\left(\varsigma_{it+1}^{dm}|\widetilde{R}_{it},\mu\right)^{\mathbf{1}[M_{it+1}=m]}\right)^{\mathbf{1}[d_{it}=d]}. \tag{15}$$

The total conditional (on μ) likelihood contribution from individual i for all time periods $7,\ldots,T_i$, where T_i is either the period of an individual's death or their final exam, is:

$$L_i(\Theta|\mu) = \prod_{t=7}^{T_i}\Big[\prod_{y=0}^{1}\left(Z_{it}^{y}|\mu\right)^{\mathbf{1}[Y_{it}=y]}\Big]. \tag{16}$$

Because of the discretized distribution of the unobserved heterogeneity, each individual's *unconditional* contribution will be a finite mixture of likelihoods. Given K points of support in the estimated distribution of μ, the unconditional likelihood function contribution for individual i is:

$$L_i(\Theta) = \sum_{k=1}^{K}\xi_k L_i(\Theta|\mu_k). \tag{17}$$

Where ξ_k is the estimated probability weight placed on mass point k. The full sample log-likelihood function is:

$$L(\Theta) = \Big[\sum_{i=1}^{N} log L_i(\Theta)\Big]. \tag{18}$$

V Results

This section begins with a description of the parameter estimates. It then discusses model fit and examines the predictions of lifetime smoking behavior under alternate smoking, learning, and health scenarios.

[28] In practice, I numerically integrate over ν_{it}. For each draw of ν_{it}, all other probabilities in the model are constructed. The resulting probabilities are then averaged over the draws. See the solution appendix for further details.

V.1 Parameter Estimates

The parameter estimates in Θ are estimated via a nested solution method (Rust, 1987). The inner algorithm solves the dynamic model for each individual conditional on a given set of parameters and for all mass points of the unobserved heterogeneity distribution. Using the resulting probabilities, the outer algorithm calculates the unconditional likelihood function, $L(\Theta)$, and attempts to improve the likelihood value via a BHHH gradient method. The BHHH method is standard in estimating dynamic structural models because, as opposed to traditional gradient methods such as Newton-Raphson that explicitly construct the Hessian matrix of the likelihood function. BHHH approximates the Hessian by exploiting the fact that the likelihood function ($L(\Theta)$) is the sum of individual log-likelihood contributions. Calculating the second derivatives of the likelihood function would be computationally infeasible for nearly all dynamic structural models. At the parameters that maximize the log-likelihood function, however, the average outer-product over individuals is the covariance matrix of the scores of the sample. Furthermore, at the true parameters, the covariance matrix of the scores is equal to minus the expected Hessian matrix (Train, 2009).

Table 12 reports the main parameter results. The estimated utility constants, α_{00} and α_{01}, for the absence of a chronic health condition and a chronic health condition respectively, are quite intuitive given that the utility of death has been normalized to zero. The total marginal utility of current period light and heavy smoking is a function of $\alpha_{1\cdot} \ldots \alpha_{8\cdot}$. A key component of rational addiction theory, indeed the defining feature of an addictive good under rational addiction, is that past consumption reinforces current consumption. That is, the marginal utility of smoking is increasing in the smoking stock. My results are consistent with this adjacent complementarity defined in Becker and Murphy (1988). In the absence of a chronic illness, both light and heavy smoking are found to be reinforcing (i.e., α_{20}, α_{60}>0). Indeed, I find that heavy smoking is much more "reinforcing" than light smoking. My results also suggest that the marginal utility of light smoking in the absence of a chronic condition is invariant to the health marker index but increasing in age. Interestingly, the marginal utility of heavy smoking is decreasing in the health marker index and invariant to age when free of a chronic condition; however, when chronically ill, the marginal utility of heavy smoking is increasing in the health markers ($\alpha_{71} = 0.001$) and decreasing in age. Withdrawal from smoking, (i.e., smoking in period $t-1$ and not smoking in period t) is negative for all health states and larger in magnitude when free of a chronic illness. The withdrawal effect, in addition to the strong reinforcement effect, both drive smokers to continue smoking. Finally, the tolerance effect ($\alpha_{10\cdot}$) flips sign across health states. In the absence of a chronic condition, smoking is found to have a tolerating effect (i.e., lower utility from a larger smoking history).

Several interesting trends emerge from these results. First, note that baseline marginal

utility of both light and heavy consumption is negative with the exception of heavy smoking *with* a chronic condition. As suggested by the rational addiction literature, the model cannot explain why individuals start smoking. Consider that over 90% of smokers in the data start smoking before age 25 and *no* individuals in the data under the age of 25 have a chronic condition. The estimated preference parameters in the absence of a chronic illness suggest that, for a never smoker under the age of 25, there is no incentive to begin smoking because the marginal utility of smoking is negative. Furthermore, the dynamic considerations of the model suggest that smoking will increase the probability of future chronic illness and death through the smoking stock and the health marker index. However, upon commencing smoking, the resulting positive smoking stock drives the dynamics forward. Competing effects for a new smoker include the reinforcement and withdrawal effects, which both encourage more smoking, and the increased probability of chronic disease and death, which encourage cessation.

The second main trend from the estimated preference parameters is the reversal in sign of several preference parameters upon succumbing to a chronic illness. The baseline marginal utility of heavy smoking when in the chronic health state (α_{51}) flips to positive. Along with the positive reinforcement ($\alpha_{21}, \alpha_{61} > 0$) and the flip in the sign of the effect of the stock on utility ($\alpha_{101} > 0$), individuals now face a positive marginal utility from heavy smoking.

The model finds evidence of a small degree of individual variation in the effect of the smoking stock (A_{it}) on the health marker index (R_{it}) as the estimated standard deviation of θ, σ_θ, is nonzero. Recall further that the null hypothesis of σ_θ equaling zero is my explicit test for the presence of learning. While the results do suggest the presence of learning, the signals received at each health exam are quite noisy. The estimated standard deviation of the random error term (σ_ν) is large relative to $\overline{\theta}$ and σ_θ.

Table 12: Main Parameter Estimates

Description	Chronic Condition	Parameter	Estimate	ASE
Utility Parameters				
Constants				
	No	α_{00}	25.947	1.808
	Yes	α_{01}	1.364	0.272
Consumption - Light Smoking				
Constant	No	α_{10}	-6.128	0.117
Consumption*Smoking Stock	No	α_{20}	0.001	0.000
Consumption*Health Marker Index	No	α_{30}	0.000	0.000
Consumption*Age	No	α_{40}	0.070	0.001
Consumption	Yes	α_{11}	-7.479	0.194
Consumption*Smoking Stock	Yes	α_{21}	2.479	0.018
Consumption*Health Marker Index	Yes	α_{31}	-0.005	0.001
Consumption*Age	Yes	α_{41}	0.002	0.001
Consumption - Heavy Smoking				
Constant	No	α_{50}	-18.753	0.043
Consumption*Smoking Stock	No	α_{60}	1.704	0.010
Consumption*Health Marker Index	No	α_{70}	-0.001	0.000
Consumption*Age	No	α_{80}	0.000	0.000
Consumption	Yes	α_{51}	0.015	0.004
Consumption*Smoking Stock	Yes	α_{61}	2.483	0.018
Consumption*Health Marker Index	Yes	α_{71}	0.001	0.000
Consumption*Age	Yes	α_{81}	-0.068	0.004
Withdrawal				
	No	α_{90}	-6.927	0.046
	Yes	α_{91}	-1.539	0.133
Smoking Stock				
	No	α_{100}	-0.025	0.007
	Yes	α_{101}	2.636	0.051
Smoking Stock Squared				
	No	α_{110}	-0.002	0.001
	Yes	α_{111}	-0.596	0.005
Learning Parameters				
Mean Effect		$\overline{\theta}$	0.003	0.000
Standard Deviation of θ_i		σ_θ	0.098	0.004
Standard Deviation of ν		σ_ν	1.024	0.004
Additional Health Marker Index Parameters				
Lagged Health Marker Index		ζ	0.807	0.001
Age in Years		ϕ_1	0.005	0.000
Female		ϕ_2	-0.122	0.003
Education in Years		ϕ_3	-0.011	0.001
Married		ϕ_4	0.000	0.000
Constant		ϕ_5	1.039	0.013

Table 13 provides estimates of all other estimated model parameters. These estimates are not marginal effects and therefore are difficult to interpret because each outcome (health marker index, chronic health, death, etc.) is a complex function of entering period states and per-period decisions. In the simulation subsection below, I describe the results of simulations that isolate the effects of each variable on the system. However, a casual interpretation of

the results in Table 13 does yield some interesting insights. The parameter estimates of the smoking stock evolution equation indicate that an individual's stock of smoking depreciates faster than suggested by the medical literature. δ_1 suggests that, given cessation from smoking over the cycle of one year, the smoking stock is reduced by approximately 57%.[29] In the context of the model, 57% depreciation implies that after about six years of smoking cessation, an individual may have roughly the same health marker index and chronic health and death tranistion probabilities as a lifelong nonsmoker, all else equal. Additionally, the estimated magnitude of investment return in the smoking stock is greater for heavy compared to light smoking ($\delta_2 < \delta_3$).

As noted above, the estimated mean effect of the smoking stock on the health markers is positive ($\overline{\theta} = 0.003$). A greater smoking history therefore implies a higher, and thus worse health marker index. According to Table 13, a higher health marker index implies a higher probability of chronic illness (through the positive sign on λ_1), albeit at a decreasing rate (λ_2<0), and death (through the positive signs on $\omega_1, \omega_2, \omega_3, \omega_5$, and ω_6. Furthermore, given a chronic illness, the probability of death is lower during the 1980s (ω_7<0) and 1990s (ω_8<0) both relative to before 1980 to capture exogenous advances in medical technology over time.

[29]Note that while this suggests a large amount of depreciation, the factor loading on unobserved heterogeneity for the stock equation slows that depreciation.

Table 13: Other Parameter Estimates

Description	Parameter	Estimate	ASE
Smoking Stock Parameters			
Depreciation Rate	δ_1	0.430	0.002
Investment, Light Smoking	δ_2	0.335	0.001
Investment, Heavy Smoking	δ_2	0.411	0.002
Standard Deviation of η	σ_η	0.134	0.000
Chronic Health Parameters			
Constant	λ_0	-12.040	0.057
Health Marker Index	λ_1	0.207	0.016
Health Marker Index Squared	λ_2	-0.010	0.001
1980s*Health Marker Index	λ_3	0.000	0.000
1990s*Health Marker Index	λ_4	0.000	0.000
Choice	λ_5	0.336	0.024
Choice*Health Marker Index	λ_6	-0.008	0.002
Age	λ_7	0.119	0.001
Education	λ_8	0.007	0.001
Gender	λ_9	0.019	0.005
Married	λ_{10}	-0.070	0.005
Mortality Parameters			
Constant	ω_0	-8.805	0.104
Health Marker Index	ω_1	0.001	0.000
Health Marker Index Squared	ω_2	0.001	0.000
Chronic Health State	ω_3	4.868	0.094
Choice	ω_4	0.000	0.000
Choice*Health Marker Index	ω_5	0.013	0.002
Choice*Chronic Health State	ω_6	0.503	0.021
1980s* Chronic Health State	ω_7	-0.086	0.017
1990s* Chronic Health State	ω_8	-0.214	0.030
Age	ω_9	0.041	0.002
Gender	ω_{10}	-0.061	0.014
Education	ω_{11}	-0.135	0.005
Married	ω_{12}	-0.204	0.028
Heterogeneity Parameters			
Utility: No Chronic Condition			
Not Smoking	ρ_{u00}	0.066	0.018
Light Smoking	ρ_{u01}	2.664	0.161
Heavy Smoking	ρ_{u02}	8.619	0.103
Utility: Chronic Condition			
Not Smoking	ρ_{u10}	0.964	0.148
Light Smoking	ρ_{u11}	-0.081	0.021
Heavy Smoking	ρ_{u12}	0.132	0.033
Stock	ρ_A	0.647	0.002
Health Marker Index	ρ_R	0.000	0.000
Chronic Health	ρ_H	0.001	0.000
Mortality	ρ_M	1.027	0.122
Mass Points and Probabilities			
Mass Point 1	μ_1	0.000	-
Mass Point 2	μ_2	1.270	0.122
Mass Point 3	μ_2	1.000	-
Coef. Weight on Mass Point 1	θ_1	-2.622	0.688
Coef. Weight on Mass Point 2	θ_2	-1.231	0.204
Miscellaneous Parameters			
Discount Factor	β	0.950	-
Log-Likelihood Value	$L(\Theta)$	-30481.266	

Mass points 1 and 3 are fixed at 0 and 1 respectively. Mass point 2 is estimated and its location is $\frac{exp(1.270)}{1+exp(1.270)} = 0.781$. The corresponding probabilities of mass points 1 through 3 are 0.053, 0.214, and 0.733.

The model is estimated with three points of support for the discretized unobserved heterogeneity distribution. Heterogeneity located to the right of the distribution is associated with a greater likelihood of experiencing both chronic health and mortality shocks. Furthermore,

the health and smoking alternative specific factor loadings in the utility function suggest that, when free of a chronic illness, the marginal utility of smoking is shifted upward for individuals with higher values of the unobserved heterogeneity. These effects are exacerbated by the fact that this type of heterogeneity also implies a smoking stock that depreciates less rapidly and never fully depreciates. Even worse, this heterogeneity characterizes continued future smoking (through the reinforcement and withdrawal effects) which also effects chronic health and death shocks (through the positive signs on λ_5 and ω_4-ω_6).

V.2 Model Fit

Figure 2 summarizes the relationship between the model's predicted probabilities and the observed data by age. Each pane of the figure represents one specific smoking option. For each individual, I compare observed smoking decisions and predicted smoking probabilities for periods up to either her final exam (exam seven) or death. I then average the results across individuals at each age.[30] The model predictions generated from the solution routine fit the data well even at ages for which there are not many observations.

Table 14 reports sample and predicted smoking probabilities by health exam and health state. I do not include a table on model fit by exam conditional on being in the chronic health state because less than one, three, and seven percent of individuals have a chronic condition in exams one, two, and three respectively. Note however that the average predicted choice probability across all health exams conditional on being in the chronic health state mirrors the observed probabilities in the data fairly well. Table 14 suggests that the model does a good job of predicting whether or not an individual smokes at all. The model slightly under predicts light smoking and slightly over predicts heavy smoking.

[30] Despite the fact that the model is solved from age 7 to 100, the figure only presents results for ages 20 to 75. Outside of the 20 to 75 age range, there are insufficient data for an informative comparison.

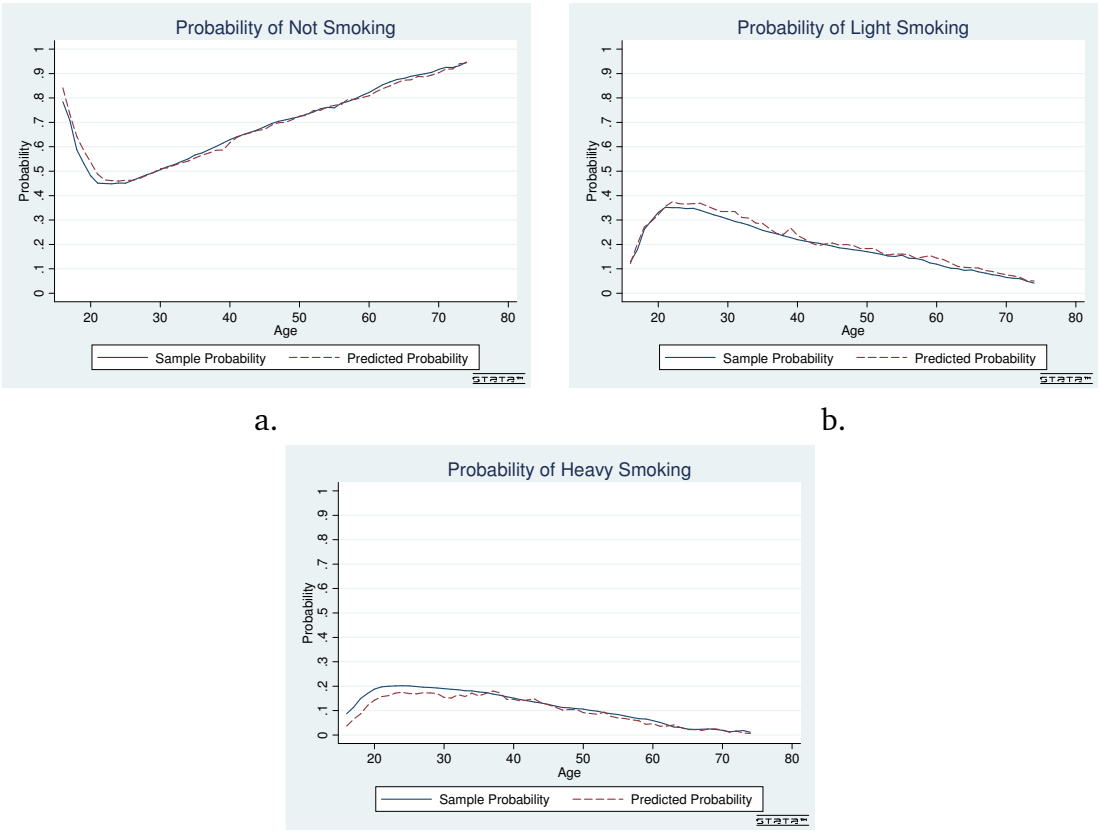


Figure 2: Smoking Behavior by Age: Predicted and Sample Probabilities

Figure 3 compares the observed sample probabilities of chronic health with the predicted health probabilities, as generated by the model at the estimated parameter values. As in Figure 2, Figure 3 averages predicted and sample probabilities across individuals by age only for those individuals with an observation at that age. Figure 3 reflects both transitions to and surviving members of the chronic health state. This is because solution to the model yields a predicted probability of transiting to a chronic health state of one for individuals already in that state. Note that for most ages, the model slightly over predicts the probability of being in a chronic health state.

Table 14: **Model Fit: Choice Probabilities**

	Not Smoking		Light Smoking		Heavy Smoking	
Exam	**Predicted**	**Observed**	**Predicted**	**Observed**	**Predicted**	**Observed**
Unconditional on Chronic Health State. # Person/Year Obs.=19,461						
1	60.68	59.01	25.68	26.70	13.64	14.30
2	68.74	61.35	19.97	24.44	11.29	14.21
3	71.32	77.22	18.64	14.25	10.03	8.53
4	78.35	81.22	14.50	12.77	7.15	6.01
5	82.53	85.16	12.00	10.96	5.47	3.88
6	86.72	87.87	9.30	9.53	3.98	2.60
7	90.21	88.89	6.82	8.65	2.97	2.46
Mean	76.94	77.25	15.27	15.33	7.79	7.43
Conditional on No Chronic Condition, ($H_{it} = 0$). # Person/Year Obs.=17,601						
1	60.68	58.99	25.71	26.72	13.61	14.29
2	68.91	61.43	20.07	24.24	11.02	14.33
3	71.64	77.31	18.76	14.23	9.60	8.46
4	78.38	80.93	14.77	13.02	6.85	6.05
5	82.47	84.91	12.45	11.10	5.08	3.99
6	86.58	87.42	9.80	9.85	3.61	2.73
7	90.01	88.22	7.25	9.14	2.75	2.64
Mean	76.95	77.03	15.54	15.47	7.50	7.50
Conditional on Chronic Health, ($H_{it} = 1$). # Person/Year Obs.=1,860						
Mean	75.69	79.27	11.03	13.64	13.28	7.09

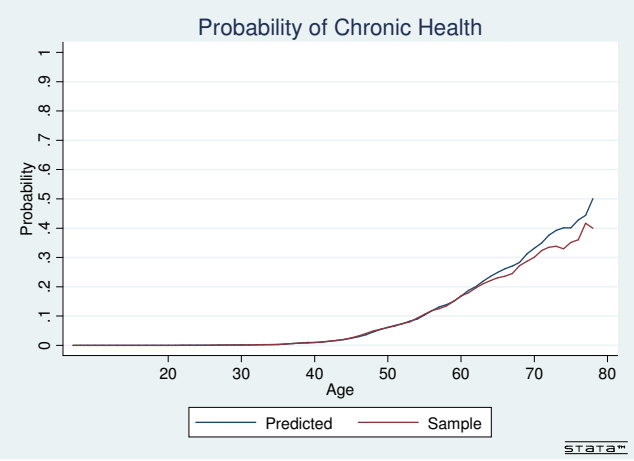


Figure 3: Chronic Health State by Age: Predicted and Sample Probabilities

V.3 Model Simulation

In this section, I simulate smoking behavior and health outcomes using the structural model and the estimated parameters. I address how smoking affects morbidity and mortality outcomes as well as how learning from personalized information may impact these behaviors and outcomes. My simulations proceed as follows. First, I construct a simulated sample of 1000 individuals that mirrors the joint distribution of observable demographic characteristics (education, gender, marriage, and initial age upon entering the Framingham survey) of the Framingham sample. Next, for each simulated individual i, I construct 50 sets of match value, unobserved heterogeneity, and error draws over the estimated time frame.

$$\left\{\theta_{ik}, \mu_{ik}, \left\{\nu_{ikt}, \eta_{ikt}, \left\{\epsilon_{iktd}\right\}_{d=0}^{2}\right\}_{t=7}^{100}\right\}_{k=1}^{50}.$$

Smoking behavior and health outcomes are then simulated for each of the 50,000 observations from age seven until death.

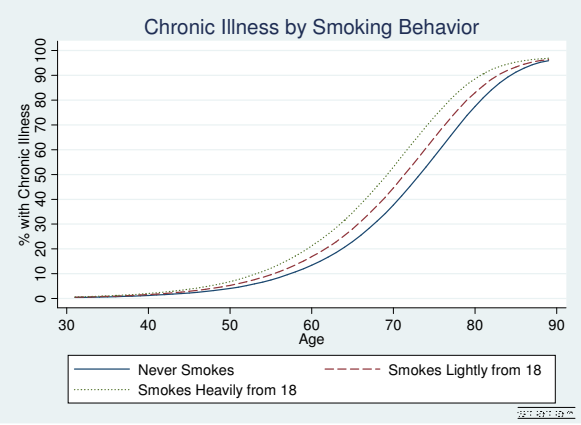
First, I reconstruct Table 11 using the simulated smoking behavior to evaluate the model's performance in capturing smoking transitions around significant events. These results are reported in Table 15.[31] For those simulated to be not smoking prior to a health exam, a chronic health shock, or transiting to at least the 75th percentile of the health marker index distribution, the simulated smoking probabilities one and three periods after these events mirror those from the data. The model does less well in simulating behavior conditional on lagged light or heavy smoking. While the simulated probabilities of not changing behavior after one of the three events reflect those from the data, the model tends to under predict the probability of quitting and over predict the probability of switching to a different smoking intensity. However, the model does capture the general trend that more individuals have quit three years after an event when compared to one year after.

[31] Transitions around health exams and transiting to the 75th percentile of the health marker index distribution are unconditional on chronic illness.

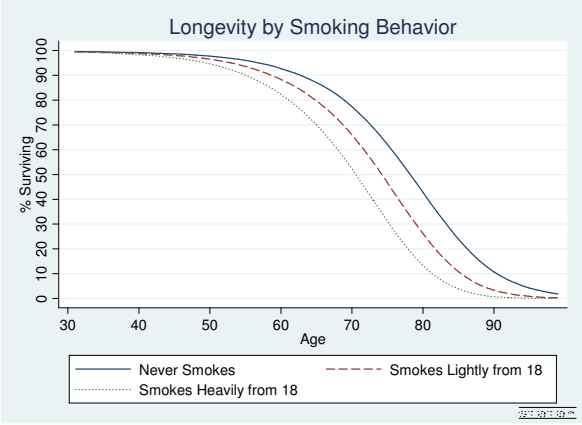
Table 15: Predicted Transitions

Behavior One Period Prior	Behavior One Period Post: Not Smoking	Light Smoking	Heavy Smoking	Behavior Three Periods Post: Not Smoking	Light Smoking	Heavy Smoking
Transitions Around Health Exams						
Not Smoking	98.73%	1.22%	0.05%	97.75%	1.98%	0.27%
Light Smoking	11.76	63.81	24.43	21.75	54.01	24.24
Heavy Smoking	4.66	37.90	57.44	9.91	42.49	47.60
Transitions Around Chronic Health Shocks						
Not Smoking	99.16	0.50	0.34	98.92	0.65	0.43
Light Smoking	32.78	44.09	23.13	55.86	27.55	16.59
Heavy Smoking	11.49	34.65	53.86	24.22	27.97	47.81
Transitions Around Health Marker Index Shocks: 75^{th} Percentile						
Not Smoking	99.04	0.93	0.03	98.18	1.65	0.17
Light Smoking	11.66	62.62	26.22	21.33	54.94	26.73
Heavy Smoking	4.22	33.12	62.66	8.87	37.91	53.22

Next, I use the simulated model to address how smoking impacts the age of chronic health onset and death. Figure 4a reports, by age, the percentage of the simulated sample with a chronic condition while forcing individuals to 1.) never smoke, 2.) smoke lightly from age 18, and 3.) smoke heavily from age 18.[32] Under these same forced behaviors, Figure 4b shows, by age, the percentage of the simulated sample that remains alive. The results in Figure 4 confirm the findings in Sloan *et al.* (2003) that the detrimental effects of smoking occur largely after the age of 50. Indeed, the gap in the percentage of the sample in the chronic



a.



b.

Figure 4: Percentage of simulated sample a.) in the chronic health state and b.) remaining, by age and quit status

[32]Recall from the structural model that I assume that, upon transiting to a chronic health state, an individual remains in that state for life.

health state between never smokers and heavy smokers widens from less than 10% at age 50 to more than 17% at age 70. Similarly, while the difference in those surviving to age 50 between heavy and never smokers is five percentage points, that gap widens to 30 percentage points at age 70. These results are roughly inline with those of Doll *et al.* (2004). Those authors find a difference of approximately 28 percentage points at age 70 when considering never smokers and smokers. The first half of Table 16 reports the mean age of onset for various health outcomes. Individuals who are forced to smoke lightly and smoke heavily from age 18 onwards face a mean age of chronic health onset that are approximately two and four years earlier than those forced to never smoke. While Doll *et al.* (2004) report that smoking shortens the lifespan by ten years, my results suggest the reduction is approximately four and eight.[33]

While Doll *et al.* (2004) only condition their results on decade of birth and gender, I report results that are conditional on both observed and unobserved factors. Here, I highlight the importance of incorporating unobserved heterogeneity. Figure 5 plots the same two graphs as in Figure 4 but now conditions each result by unobserved "type". Panels a. and b. report health outcomes under the baseline rational choice simulation whereas panels c. and d. report health outcomes assuming that all simulated individuals never smoke. Note that while unobserved heterogeneity does not play a significant role in chronic health transitions, the model predicts that type three individuals face lower expected longevity in both the baseline and nonsmoking simulations. Recall that the alternative specific factor loadings in the utility function greatly increase the marginal utility of smoking for individuals of a higher type. Indeed, the model predicts that only individuals with the largest mass point, type three, will ever smoke. Therefore, Figure 5 demonstrates that, independent of smoking, individuals of a higher type face lower expected longevity.

[33] Doll *et al.* (2004) do not take into account intensity of smoking in these calculations. My results indicate that, conditional on smoking, the intensity with which one smokes is an important factor explaining health outcomes.

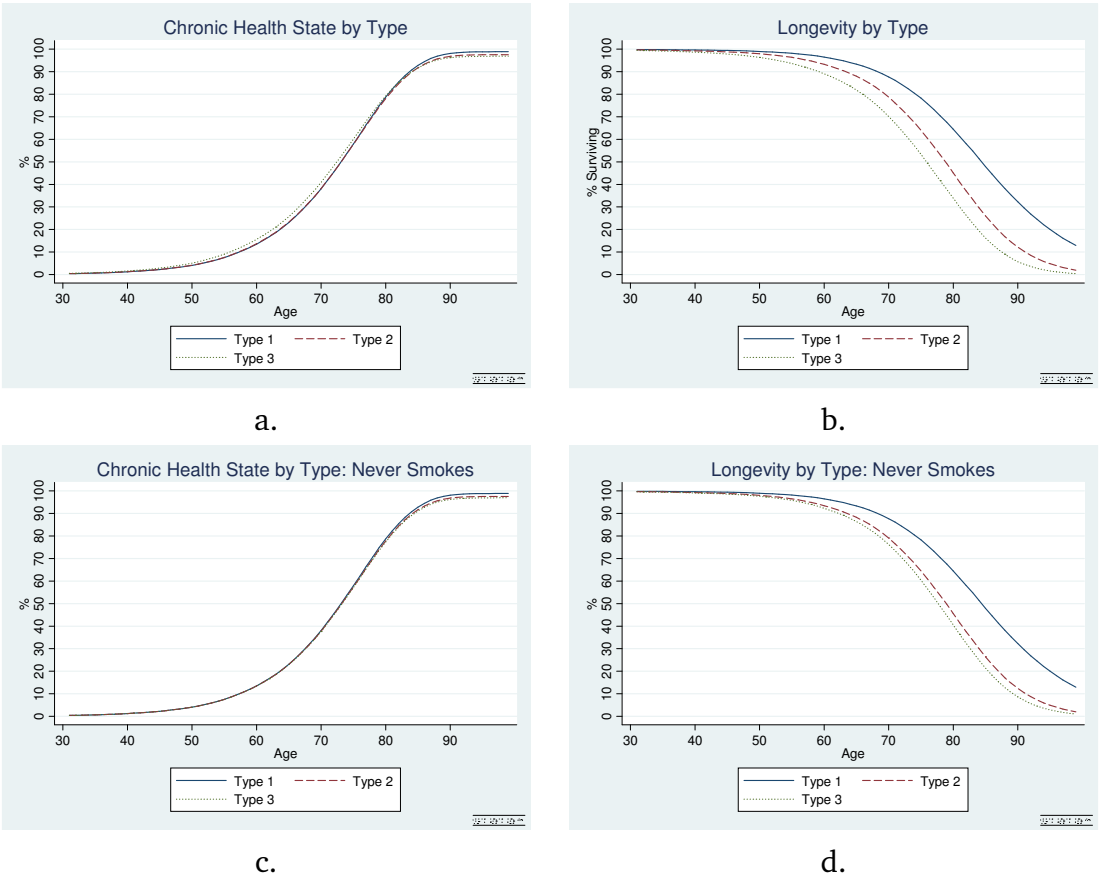


Figure 5: a.) Simulated chronic health state by age at baseline. b.) Simulated longevity by age at baseline. c.) Simulated chronic health state by age assuming no smoking. d.) Simulated longevity by age assuming no smoking.

Next, I use the model to simulate chronic health and death outcomes under different lifetime smoking paths to assess the impact of smoking cessation on these outcomes. I simulate health outcomes assuming that an individual smokes heavily from age 18 and quits forever at ages 30, 40, 50, and 60. The results, reported in Table 16, imply that quitting smoking at ages 30, 40, 50, and 60 years of age increases life-expectancy by approximately 8, 7.75, 7, and 5.5 years, respectively. These results suggest clear life expectancy gains from quitting at *all stages of the life cycle*.

Table 16: Age of Chronic Health Onset and Death

Variable	**Mean Age of Chronic Health Shock**	**Mean Age of Death**
Never Smokes	70.75	77.60
	(10.72)	(11.60)
Smokes $\leq$ 1 Pack/day from Age 18	68.91	73.32
	(10.89)	(11.19)
Smokes $>$ 1 Pack/day from Age 18	66.79	69.58
	(10.87)	(10.94)
Smokes $>$ 1 Pack/day from Age 18 *and*		
quits at Age 30	70.77	77.58
	(11.00)	(11.85)
quits at Age 40	70.54	77.32
	(11.33)	(12.28)
quits at Age 50	69.98	76.60
	(11.83)	(11.02)
quits at Age 60	66.55	74.99
	(10.91)	(13.85)

Figure 6 shows the survival percentages by age for the different smoking patterns. Note that for individuals that quit at age 30, their expected longevity is roughly identical to never smokers. Similarly, quitting by age 40 has minimal effects on mortality probabilities. Individuals that smoke into their fifties and sixties, however, have a much more likely chance of dying prematurely.

Given the other main focus of this paper on the value of personalized health information, I next evaluate policies that alter either the timing or the frequency with which information is received. First, to demonstrate the speed at which individuals learn, Table 17 reports the change in the average posterior variance after each health exam of the baseline simulation. Note that after the first exam (i.e., the first signal of information) the posterior variance decreases by nearly 20%. By the seventh exam, the mean posterior variance has been decreased by 40%. In spite of the "honing in" on individuals' true match values, smoking behavior appears to only slightly be influenced by learning. As a natural benchmark, I compare the predictions of the baseline model to results from specifications with no learning (i.e., $\sigma_\theta = 0$), complete information (i.e., $\tau_{it} = \theta_i \ \forall t$), and a situation where an individual undergoes yearly health exams as opposed to every four years. Figure 7 presents the mean percentage difference of simulated individuals choosing each smoking option for each information scenario relative to the baseline prediction. Somewhat counter intuitively, the

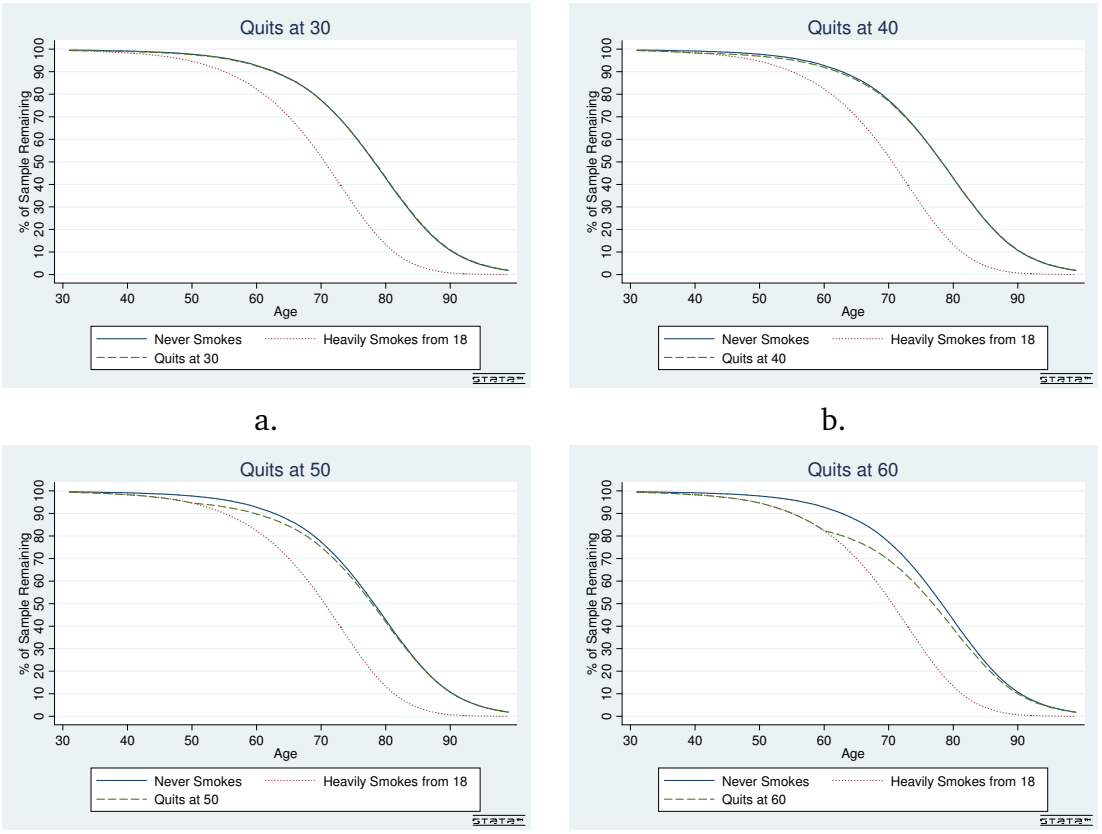


Figure 6: Percentage of simulated sample remaining, by age and quit status

Table 17: Posterior Variance by Exam: The Speed of Learning

Exam	Mean Posterior Variance	% Decrease	Cummulative % Decrease
Initial Prior	0.0095	-	-
1	0.0076	19.9%	19.9%
2	0.0069	8.7%	26.9%
3	0.0066	4.7%	30.3%
4	0.0064	3.6%	32.8%
5	0.0061	4.0%	35.5%
6	0.0059	3.8%	37.9%
7	0.0057	2.9%	39.7%

simulations suggest that the effect of more information, that is, yearly exams, is only to encourage individuals to smoke lightly in later life. In the extreme, with complete information, individuals are more likely to smoke lightly at all ages. In both cases, there is no apparent change in heavy smoking. One possible explanation for this finding is that, because the effect

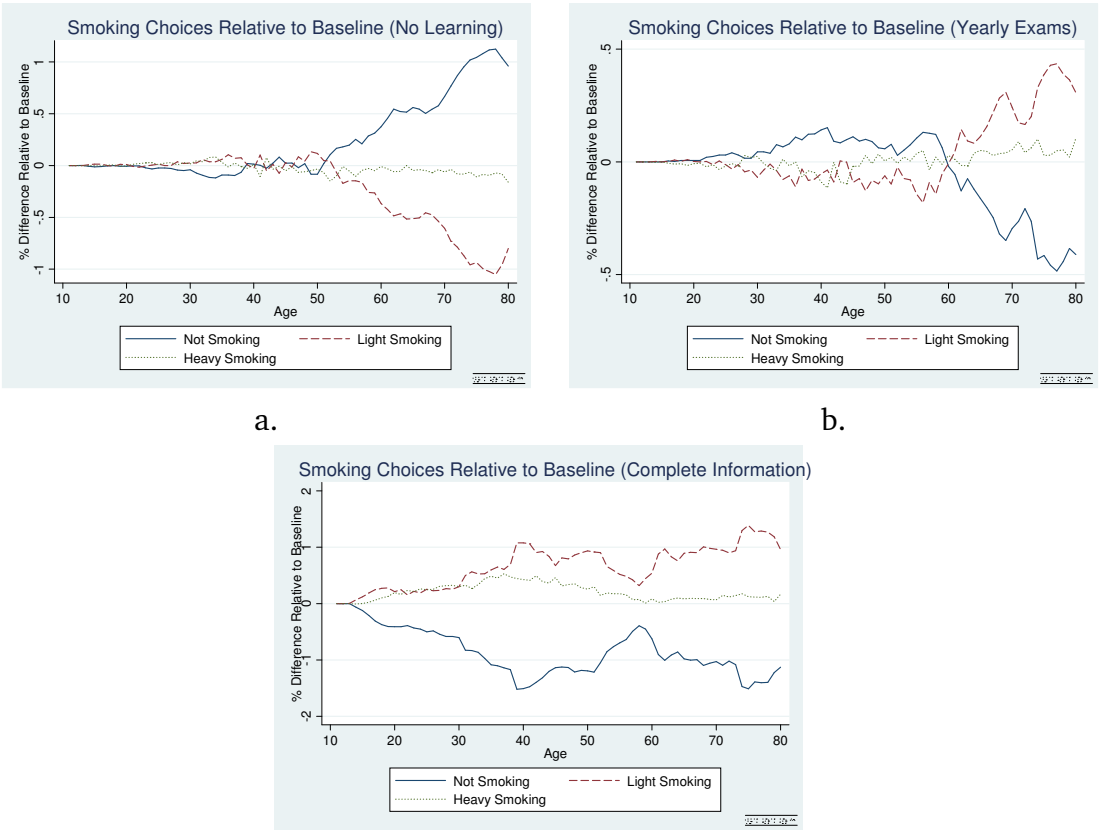


Figure 7: Average difference in smoking probabilities, relative to baseline choices, by age and across different policy scenarios

of the smoking stock on the health marker index is small ($\overline{\theta} = 0.003$) and because the estimated standard deviation of the effect is large relative to the mean, upon learning their true match value, individuals feel that the health effects of smoking are manageable.[34] Ultimately, the effects of different information regimes are quit small. Even with yearly exams, by age 70, the difference in average smoking rates relative to the baseline model predictions are only approximately 0.06% higher.

[34]For match values that are negative, there may be an incentive to continue to smoke because an increased smoking stock will decrease the health marker index, which in turn, will lower chronic health and death probabilities. Other experiments in which health signals where positively amplified, that is, while the health marker index evolved according to the estimated structural parameters, individuals *received signals* that suggested "scary" results, induced individuals to quit significantly more rapidly than the baseline results.

VI Discussion

This study formulates and estimates a dynamic stochastic model of smoking behavior. The model extends the classic rational addiction model to allow for health learning. Estimates of the structural parameters suggest that there exists heterogeneity in the effect of the accumulated smoking stock on an index of health markers. Significant reinforcement and withdrawal effects both drive smoking dynamics by altering the future marginal utilities of smoking. The theoretically well-known smoking stock, empirically constructed using principal components analysis, is found to depreciate at approximately 57% per year. Simulations of the structural model suggest that, when controlling for unobserved heterogeneity, the effects of smoking on chronic health and mortality outcomes may be slighly less extreme than previously estimated. I find that smoking heavily from age 18 can reduce life expectancy by eight years relative to life-long non-smokers and by four years relative to those smoking only lightly ($\leq$ 1 pack/day) from age 18. Furthermore, quitting smoking by age 30 implies relatively few chronic health or mortality differences, on average, from life-long non-smokers; however, waiting to quit until age 60 implies that the health consequences may be severe. Indeed, as suggested by the literature, the major effects of smoking on health are realized after age 50. Finally, health markers, at least in this setting, do not appear to significantly inform smokers about the long-term health consequences of smoking. In fact, learning about how smoking effects health markers may actually increase moderate smoking in older individuals.

Appendices

A Solution

Given the long time frame of the model and the mixed discrete/continuous nature of the state space, I employ a variant of the Keane and Wolpin (1994) value function interpolation method for approximating the value function. The iterative solution method proceeds in two main steps. First, I solve the model generally for each possible age at which an individual may have taken her first health exam.[35] Starting in the final period T, the method is as follows:

- Beginning in period T, I draw n state vectors and sequences of past smoking behavior $\mathbf{D}_{iT-1} = \{d_{i1}, \ldots, d_{iT-1}\}$.[36] For each of the n draws, I construct the main equations of the model.

- Next, I posit a relationship between the n calculated value functions and a set of regressors. The regressors include the drawn state variables in addition to interaction and higher-order terms. Using the calculated regression coefficients, I can calculate predicted values of the lifetime value function for any value of the state vector.[37]

- I then repeat the above steps for period $T-1$. When calculating the value function in period $T-1$, I use the regression coefficients from period T to approximate the expected future value function.

- I repeat the above process for all periods back to $t = 7$. That is, I solve the model for all ages between 7 and 100.

The above first-stage process yields value function regression coefficients at every time period and for every possible initial age upon entering the sample. These coefficients are conditional on the parameters used to solve the model. The second main step, conditional on the same parameters and using the above regression coefficients, involves solving the model for each individual. The resulting probabilities enter the likelihood function. The total process continues, updating parameters at each iteration, until the likelihood function is maximized.[38]

[35] In the data, the ages range from 13 to 62.

[36] In practice I set $n = 100$.

[37] Note that because the probability of death in period T equals one, each of the choice specific value functions in period T simply equals the current period utility from the choice.

[38] To maximize the likelihood function, I use the BHHH numerical maximization technique. I assume convergence at the maximum of the likelihood function when the percentage change of the likelihood value over an iteration is at or below 0.000001. The model is solved and estimated using parallel processing techniques for Fortran 90 code. In practice, I use 51 64-bit processors.

B Bayesian Updating

Here, I derive the posterior beliefs discussed in the main text (Equations 4 and 5). I assume rational expectations such that an individual's initial belief upon entering the sample regarding their true θ_i is the population distribution:

$$E_0(\theta_i) = \tau_{i0} = \overline{\theta}$$

$$V_0(\theta_i) = \psi_{i0} = \sigma_\theta^2.$$

Consider an individual in period t with smoking stock A_{it}. For ease of exposition, assume that an individual takes a health exam each period. When deriving the posterior beliefs in period t, an individual considers only her prior beliefs $(\tau_{it-1}, \psi_{it-1})$ and her signal of information k_{it}. According to Bayes' Rule, the posterior distribution, f_t, of θ_i is given as:

$$f_t(\theta_i|\kappa_n, \tau_{it-1}, \psi_{it-1}) \propto f_{t-1}(\theta_i) g(\kappa_{it}|A_{it}, \theta_i, \sigma_\nu). \tag{B.1}$$

Note that while $g(\kappa_{it}|A_{it}, \theta_i, \sigma_\nu)$ conveys information about κ^{it}, an individual knows A_{it} and, because θ_i is time invariant, can therefore infer information about θ_i over time. This will become more clear in the interpretation of the posterior mean and variance. First consider $g(\kappa_{it}|A_{it}, \theta_i, \sigma_\nu)$:

$$g(\kappa_{it}|A_{it}, \theta_i, \sigma_\nu) = \frac{1}{(2\pi\sigma_\nu^2)^{\frac{1}{2}}} \exp\Big(\frac{1}{2\sigma_\nu^2}(\kappa_{it} - \theta_i A_{it})^2\Big). \tag{B.2}$$

Note that because we are concerned with the distribution of θ_i, any term that does not include θ_i can be treated as part of the normalizing constant. We can ignore the first term within the parenthesis:

$$\propto \exp\Big(\frac{-1}{2\sigma_\nu^2}\big(-2\theta_i\kappa_{it}A_{it} + \theta_i^2 A_{it}^2\big)\Big).$$

Simplifying and completing the square yields:

$$\propto \exp\Big(-\frac{A_{it}^2}{2\sigma_\nu^2}\big(\theta_i - \frac{\kappa_{it}A_{it}}{A_{it}^2}\big)^2\Big).$$

Notice that the term subtracted from θ_i is the within (individual i) variation ordinary least squares estimate of θ_i from the n^{th} signal of information. Define $\hat{\theta_{it}} = \frac{\kappa_{it}A_{it}}{A_{it}^2}$. Substituting for $\hat{\theta_{it}}$, we have that:

$$g(\kappa_{it}|A_{it}, \theta_i, \sigma_\nu) \propto \exp\Big(-\frac{A_{it}^2}{2\sigma_\nu^2}\big(\theta_i - \hat{\theta_{it}}\big)^2\Big). \tag{B.3}$$

Now consider the prior probability distribution of θ_i:

$$f_{t-1}(\theta_i) = \frac{1}{(2\pi\psi_{it-1})^{\frac{1}{2}}} \exp\Big(\frac{1}{2\psi_{it-1}}(\theta_i - \tau_{it-1})^2\Big). \tag{B.4}$$

The nice aspect of the conjugate distribution assumption is that we can characterize the posterior distribution sufficiently with closed form expressions for the posterior mean and variance. Therefore, we only have to characterize that part of the posterior density that captures the mean and variance. In that light, consider the product of the exponential portions of Equations B.3 and B.4 after rearranging terms and absorbing those without θ_i into the normalizing constant:

$$f_t(\theta_i) \propto \left(-\frac{1}{2\psi_{it-1}\sigma_\nu^2}\left(\theta_i^2\left(A_{it}^2\psi_{it-1}+\sigma_\nu^2\right) - 2\theta_i\left(A_{it}^2\psi_{it-1}\hat{\theta}_{it}+\sigma_\nu^2\tau_{it-1}\right)\right)\right). \tag{B.5}$$

After rearranging and completing the square, we have the kernel of a normal distribution representing the posterior distribution:

$$f_t(\theta_i) \propto \left(-\frac{A_{it}^2\psi_{it-1}+\sigma_\nu^2}{2\psi_{it-1}\sigma_\nu^2}\left(\theta_i - \left(\frac{A_{it}^2\psi_{it-1}\hat{\theta}_{it}+\sigma_\nu^2\tau_{it-1}}{A_{it}^2\psi_{it-1}+\sigma_\nu^2}\right)^2\right)\right).$$

The posterior mean and variance is:

$$\tau_{it} = E(\theta_i|\kappa_t, \tau_{it-1}, \psi_{it-1}) = \left(\frac{A_{it}^2\psi_{it-1}}{A_{it}^2\psi_{it-1}+\sigma_\nu^2}\right)\hat{\theta}_{it} + \left(\frac{\sigma_\nu^2}{A_{it}^2\psi_{it-1}+\sigma_\nu^2}\right)\tau_{it-1} \tag{B.6}$$

$$\psi_t = Var(\theta_i|\psi_{it-1}, \sigma_\nu) = \frac{\psi_{it-1}\sigma_\nu^2}{A_{it}^2\psi_{it-1}+\sigma_\nu^2}. \tag{B.7}$$

Rearranging these equations yields the posterior mean and variance equations above.

References

Ackerberg, D. (2003). Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination. *International Economic Review*, **44** (3), 1007–1040.

Adda, J. and Lechene, V. (2001). Smoking and Endogenous Mortality: Does Heterogeneity in Life Expectancy Explain Differences in Smoking Behavior, discussion Paper Series, Department of Economics, University of Oxford, ISSN 1471-0498.

— and — (2004). On the identification of the effect of smoking on mortality, discussion Paper Series, Department of Economics, University of Oxford, Working Paper Number 184.

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey. *Journal of Econometrics*, **156** (1), 38–67.

ARCIDIACONO, P., SIEG, H. and SLOAN, F. A. (2007). Living Rationally Under the Volcano? An Empirical Analysis of Heavy Drinking and Smoking. *International Economic Review*, **48** (1), 37–65.

BECKER, G. S. and MURPHY, K. (1988). A Theory of Rational Addiction. *Journal of Political Economy*, **96** (4), 675–700.

BENNETT, C. H. and RICHARDSON, D. R. (1984). Effects of Chronic Tobacco Smoke Exposure on Arterial Blood Pressure Regulation. *American Journal of Physiology: Heart and Circulatory Physiology*, **247**.

BLAU, D. M. and GILLESKIE, D. B. (2008). The Role of Reitree Health Insurance in the Employment Behavior of Older Men. *International Economic Review*, **49** (2), 475–514.

CHAN, T. Y. and HAMILTON, B. (2006). Learning, Private Information, and the Economic Evaluation of Randomized Experiments. *Journal of Political Economy*, **114** (6), 997–1040.

CHERNEW, M., GOWRISANKARAN, G. and SCANLON, D. P. (2008). Learning and the Value of Information: Evidence from Health Plan Report Cards. *Journal of Econometrics*, **144**, 156–174.

CRAWFORD, G. S. and SHUM, M. (2005). Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, **73** (4), 1137–1173.

DOLL, R., PETO, R., BOREHAM, J., GRAY, R. and SUTHERLAND, I. (2004). Mortality in Relation to Smoking: 50 Years' Observations on Male British Doctors. *British Medical Journal*, **328**, 1519–1528.

—, —, WHEATLEY, K., GRAY, R. and SUTHERLAND, I. (1994). Mortality in Relation to Smoking: 40 Years' Observations on Male British Doctors. *British Medical Journal*, **309**, 901–911.

EPSTEIN, L. G. and ZIN, S. E. (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. *The Journal of Political Economy*, **99** (2).

GARRISON, R., KANNEL, W., FEINLEIB, M., CASTELLI, W., MCNAMARA, P. and PADGETT, S. (1978). Cigarette Smoking and HDL Cholesterol. *Atherosclerosis*, **30**, 17–25.

GILLESKIE, D. B. (1998). A Dynamic Stochastic Model of Medical Care Use and Work Absence. *Econometrica*, **66** (1), 1–45.

GRUBER, J. and KOSZEGI, B. (2001). Is Addiction "Rational"? Theory and Evidence. *The Quarterly Journal of Economics*, **116** (4).

HECKMAN, J. J. and SINGER, B. (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, **52** (2), 271–320.

KEANE, M. P. and WOLPIN, K. I. (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *The Review of Economics and Statistics*, **76** (4).

KHWAJA, A. (2010). A Life Cycle Analysis of the Effects of Medicare on Individual Health Incentives and Health Outcomes. *Journal of Econometrics*, **156**, 130–147.

—, SLOAN, F. A. and CHUNG, S. (2006). Learning About Individual Risk and the Decision to Smoke. *International Journal of Industrial Organization*, **24** (), 683–699.

KREPS, D. M. and PORTEUS, E. L. (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica*, **46** (1).

— and — (1979). Dynamic Choice Theory and Dynamic Programming. *Econometrica*, **47** (1).

MIRA, P. (2007). Uncertain Infant Mortality, Learning, and Life-Cycle Fertility. *International Economic Review*, **48** (3), 809–846.

MROZ, T. (1999). Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome. *Journal of Econometrics*, **92** (2), 233–274.

OMVIK, P. (1996). How Smoking Affects Blood Pressure. *Blood Pressure*, **5**, 71–77.

RUST, J. (1987). Optimal Replacement of Gmc Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, **55** (5), 999–1033.

SICKLES, R. C. and WILLIAMS, J. (2008). Turning from Crime: A Dynamic Perspective. *Journal of Econometrics*, **145**, 158–173.

SLOAN, F. A., SMITH, V. K. and TAYLOR, D. H. (2003). *The Smoking Puzzle: Information, Risk Perception, and Choice*. Harvard University Press.

SMITH, V. K., TAYLOR, D. H., SLOAN, F. A., JOHNSON, F. R. and DESVOUSGES, W. H. (2001). Do Smokers Respond to Health Shocks? *The Review of Economics and Statistics*, **83** (4), 675–687.

TRAIN, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edn.

VISCUSI, W. K. (1990). Do Smokers Underestimate Risks? *Journal of Political Economy*, **98** (6), 1253–1269.

— and EVANS, W. N. (1990). Utility Functions that Depend on Health Status: Estimates and Economics Implications. *American Economic Review*, **80** (3), 353–374.

VYAS, S. and KUMARANAYAKE, L. (2006). Constructing Socio-Economic Status Indices: How to Use Principal Component Analysis. *Oxford University Press in association with the London School of Hygiene and Tropical Medicine*.

WILSON, P. W., D'AGOSTINO, R. B., LEVY, D., BELANGER, A. M., SILBERSHATZ, H. and KANNEL, W. B. (1998). Prediction of Coronary Heart Disease Using Risk Factor Categories. *Circulation*, **97**, 1837–1847.