LIVING RATIONALLY UNDER THE VOLCANO?
AN EMPIRICAL ANALYSIS OF HEAVY DRINKING AND SMOKING

Peter Arcidiacono
Holger Sieg
Frank Sloan

Living Rationally Under the Volcano?
An Empirical Analysis of Heavy Drinking and Smoking
Peter Arcidiacono, Holger Sieg and Frank Sloan
NBER Working Paper No. 8602
November 2001
JEL No. I1, C1, H4

ABSTRACT

This study investigates whether models of forward-looking behavior explain the observed patterns of heavy drinking and smoking of men in late middle age in the Health and Retirement Study better than myopic models. We develop and estimate a sequence of nested models which differ by their degree of forward-looking behavior. We also study models which allow for heterogeneity in discounting, and thus test whether certain types of individuals are more likely to show forward-looking behavior than other types. Our empirical findings suggest that forward-looking models with an annual discount factor of approximately 0.78 fit the data the best. These models also dominate other behavioral models based on out-of-sample predictions using data of men aged 70 and over. Myopic models predict rates of smoking and drinking for old individuals which are significantly larger than those found in the data on elderly men.

Peter Arcidiacono
Department of Economics
Duke University
Box 90097
Durham, NC 27708

Holger Sieg
GSIA
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213-3890
and NBER
hs25@econ.duke.edu

Frank Sloan
Department of Economics
Duke University
Box 90097
Durham, NC 27708
and NBER
fsloan@hpolicy.duke.edu

1 Introduction

Two competing theories have been proposed for explaining the consumption of potentially harmful and addictive goods. Early approaches typically attributed consumption of these goods to irrational or myopic behavior (Winston, 1980; Thaler and Sheffrin, 1981).[1] More recently, Stigler and Becker (1977) and Becker and Murphy (1988) have forcefully argued that addiction can be modeled as an outcome of rational behavior of forward-looking individuals with stable preferences.[2] These two theories primarily differ in their assumptions regarding the length of the planning horizon which is attributed to individuals. The myopic model assumes that the planning horizon is short and consists – in the limiting case – of only one time period. Individuals care about today, but ignore tomorrow and hence do not internalize the negative effects of smoking and drinking on health in the future. In contrast, rational addiction theory relies on the notion that individuals are forward-looking. Thus, individuals take into consideration the risks associated with heavy smoking or drinking.

Most prior empirical studies of the rational addiction theory follow Becker and Murphy (1988) and analyze first order conditions that prices and quantities need to satisfy, given individuals' (quadratic) utility functions.[3]. Chaloupka (1991) and Becker, Grossman, and Murphy (1991, 1994) apply this methodology and find that tobacco consumption typically

---

[1]One of most fascinating early analyses of alcohol abuse is due to Lowry (1947), which was published in its original version with a 17 year delay in the Prairie Schooner, XXXVII, 4, Winter 1963/64.

[2]An alternative to both approaches are models based on recent work by Laibson (1997), Harris and Laibson (2001) using hyperbolic discounting. In these models, individuals are forward-looking, but put less weight on future events than in standard forward-looking models which can give rise to time inconsistent consumption paths. See also Gruber and Koeszegi (2000).

[3]Exceptions are recent work by Gilleskie and Strumpf (2000) who use linear approximation of decision rules in a discrete choice model and Choo (2000) and Khwaja (2000) who estimate models using full-solution dynamic programming techniques.

responds to lagged, current and future price changes as predicted by rational addiction theory.[4]

This study differs from previous empirical studies in a number of important ways. Previous research has largely focused on young adults who are subject to experimentation, habit formation, and reinforcement (Orphanides and Zervos, 1995). However, an important characteristic of tobacco consumption, in particular consumption of cigarettes, is the long latency period between time of initiation and onset of adverse events.[5] Relatively few adverse health events occur in the first half of life. To illustrate, at age 35, the cumulative probability of survival is the same for males who have never smoked and smokers. At age 45 (65, 85), the corresponding ratio is 1.02 (1.18, 2.11) (Hodgson, 1992). Perfectly forward-looking individuals may therefore engage in heavy consumption of harmful substances at young ages because events in the distant future are heavily discounted. It is therefore difficult to distinguish between the two theories using data of young adults.

We focus our analysis on a sample of men in late middle age from the Health and Retirement Study (HRS). Persons over the age of 50 start to experience negative health shocks which are, at least partially, due to smoking and heavy drinking in the past. Much of the uncertainty about future health and the link between smoking or drinking and health outcomes is resolved during later years in life. We can thus study whether individuals rationally update their consumption behavior as they experience negative health shocks

---

[4]Chaloupka and Warner (2000) provide an overview of the existing empirical literature of the rational addiction model.

[5]The latency period for alcohol can be substantially less than for smoking, for example, due to accidents while being intoxicated.

and provide a new test of the rational addiction hypothesis.[6]

Earlier studies have typically assumed that individuals are either myopic or forward-looking. In contrast, we provide an empirical framework which allows us to control for varying degrees of forward-looking behavior of individuals. We adapt a dynamic discrete choice framework.[7] The myopic model is a special case of forward-looking models. Our framework thus nests the competing behavioral theories. Instead of *assuming* that individuals are rational, we compare a sequence of nested models which differ by their degree of forward-looking behavior. Moreover, we allow for heterogeneity in planning horizons, and *test* whether certain types of individuals are more likely to show forward-looking behavior than others.

Our empirical findings suggest that forward-looking models with annual discount factors less than 0.8 fit the data the best. To gain additional insights in the fit of the different model specifications, we predict the behavior for a sample of elderly individuals from the Asset and Health Dynamics Among the Oldest Old (AHEAD) data. We find that the out-of-sample predictions of our preferred forward-looking model clearly dominate those of the myopic model. We thus conclude that forward-looking models with an annual discount factor of approximately 0.78 provide better within sample and out-of-sample predictions than their myopic counterparts.

---

[6]Previous tests of the rational addiction model have been based on Euler equations, which rest on the implicit assumption that individuals must consume strictly positive amounts of alcohol or cigarettes. This is problematic if a large number of individuals do not consume tobacco or alcohol in each time period.

[7]Dynamic discrete choice estimation was first used by Wolpin (1984), Miller (1984), Pakes (1986) and Rust (1987). Recent applications include Hotz and Miller (1993), Keane and Wolpin (1994, 1997), Rust and Phelan (1997), Gilleskie (1998), Eckstein and Wolpin (1999) and Brien, Lillard, and Stern (2000).

Comparing myopic models to forward-looking models yields substantial differences in the consumption patterns of alcohol and tobacco over time. Larger discount factors imply larger declines in consumption with age. In forward-looking models, individuals take into account that the marginal adverse health effects of heavy drinking and smoking are higher later in life. The main drawback of the myopic model is that it predicts rates of smoking and heavy drinking for old individuals which are significantly larger than those found in the AHEAD. Myopic and forward-looking models also have very different behavioral interpretations. Estimates of the myopic model make alcohol and tobacco appear unattractive, particularly for the unhealthy. However, in forward-looking models, alcohol and tobacco are more attractive, but are not consumed because of losses in future utility caused by adverse health effects. This is particularly true for the unhealthy as the marginal adverse health effects of heavy drinking and smoking increase as the health of the individual deteriorates.

The rest of the study is organized as follows. Section 2 discusses the data used in this paper which is based on the Health and Retirement Study. In Section 3, we develop a sequence of forward-looking models of decision making under uncertainty and discuss identification and estimation of these models. Sections 4 and 5 present the empirical results of this study. In Section 6, we analyze the fit of the models using within sample and out-of-sample predictions. Section 7 offers some concluding remarks.

2 Data

The data used in this study come from the Health and Retirement Study. The HRS is a national panel study of birth cohorts 1931 through 1941 and their spouses, if married. Participants in the HRS have been interviewed every two years since 1992. To date, four waves of the survey have been completed and released. Individuals in the first wave of the HRS range from 51 to 61 years of age with some spouses being younger or older than this. The average age in wave 1 of individuals in our sample is 58.2. We analyze smoking and heavy drinking behavior of males. We exclude women because they differ significantly from males in preferences for alcohol and tobacco, in their attitudes towards risk, and their willingness to engage in harmful consumption. Furthermore, women live longer and have different tolerance levels for alcohol and tobacco than males. Hence, transitions for health status and mortality differ significantly between males and females. Controlling for these differences using additive dummy variables is likely to be insufficient.

We use the panel structure of the data and include only persons who were not lost due to attrition and had complete information for the most relevant variables of this study. Our initial sample consists of 5,735 males. 1,281 men are lost due to attrition. We also remove those who abstained from any alcohol consumption throughout the period. This, and missing values, gives us 2,784 males on average in each wave, a total number of 8,352 male person-waves.

Drinking behavior questions in the HRS allow us to categorize wave 1 respondents as current drinkers or non-drinkers, and for drinkers the average number of drinks per day.

Table 1: Sample Means

		wave 1	wave 2	wave 3	wave 4
number of drinks:	= 0	.157	.226	.260	.306
	< 1	.558	.505	.505	.469
	1 − 2	.178	.178	.166	.158
	3 − 4	.080	.065	.050	.046
	> 4	.027	.026	.019	.020
smoking:	yes - no	.281	.256	.230	.203
health status:	excellent	.238	.200	.187	.128
	very good	.311	.311	.339	.308
	good	.287	.292	.290	.326
	fair	.114	.141	. 135	.172
	poor	.050	.056	.048	.066
total household income		65,593	58,917	56,232	52,997

Subsequent waves allow for tracking drinking status across time. For purposes of this study, we categorize individuals as drinkers and non-drinkers, and, among drinkers, by drinks per day: less than 1, 1-2, 3-4, and more than 4. We define heavy drinking as consuming three or more drinks on average per day.[8]

Table 1 shows that there is a steady decline in alcohol consumption, smoking, health status, and total income[9] during the four waves of the panel. For example, the percentage of individuals who do not drink increases from 15.7 percent to 30.6 percent during the observation period. Heavy drinking declines from 10.7 percent in wave 1 to 6.6 percent in

---

[8]See Graham (1985) for a discussion of alternative measures of alcohol consumption.

[9]The HRS includes detailed information on both labor and non-labor income. In this paper, we define total household income as a sum of the respondent's wage income, the spouse's wage income, household capital income and other household income. Income is measured in constant 1999 dollars.

wave 4.[10] Smoking, meanwhile, declines from 28.1 percent to 20.3 percent.

Table 2: Heavy Drinking and Smoking Transitions

		lagged consumption				
		a=0,c=0	a=0,c=1	a=1,c=0	a=1,c=1	Total
	a=0,c=0	5535	294	204	21	6054
		91.4%	4.9%	3.4%	0.3%	100.0%
		94.9%	16.6%	47.6%	6.6%	72.5%
	a=0,c=1	137	1384	6	133	1660
		8.3%	83.4%	.4%	8.0%	100.0%
		2.3%	78.0%	1.4%	42.0%	19.9%
current	a=1,c=0	151	8	211	17	387
consumption		39.0%	2.1%	54.5%	4.4%	100.0%
		2.6%	0.5%	49.2%	5.4%	4.6%
	a=1,c=1	9	88	8	146	251
		3.6%	35.1%	3.2%	58.2%	100.0%
		0.2%	5.0%	1.9%	46.1%	3.0%
	Total	5832	1774	429	317	8352
		69.8%	21.2%	5.1%	3.8%	100.0%
		100.0%	100.0%	100.0%	100.0%	100.0%

Additional information about alcohol and tobacco consumption is given in Table 2 which reports the transitions between heavy drinking (no, $a = 0$ and yes, $a = 1$) and smoking combinations. We find that the elements on the diagonal of the transition matrix are larger in magnitude than the elements on the off-diagonals, with the elements on the diagonal

---

[10]There is an issue of whether alcoholics give valid self-reports regarding their alcohol consumption. While there is some evidence in the literature that alcohol consumption is measured with error, there do not seem to be any systematic trends of over- or underreporting. See, for example, Watson, Tilleskjor, Hoodecheck-Schow, Purcel, and Jacobs (1984).

ranging from 54 and 91 percent. Thus, in spite of the general trend of reducing alcohol and tobacco consumption evident in the sample means, many men have relatively stable drinking and smoking patterns which is likely due to habit formation. However, there is still a fair amount of transitions between the different heavy drinking and smoking states, as reflected by the off-diagonal elements of the matrix.

The HRS also contains some qualitative information about past problem drinking and smoking behavior. This information is used in the empirical analysis to control for past habit formation. The first questionnaire of the HRS includes the CAGE instrument for clinical assessment of alcohol disorders. The acronym CAGE represents four questions that comprise the instrument: Have you ever felt you should *Cut down* on your drinking (32.7 percent)? Have people *Annoyed* you by criticizing your drinking (16.9 percent)? Have you ever felt bad or *Guilty* about drinking (21.1 percent)? Have you ever had a drink first thing in the morning (*Eye-opener*) to calm your nerves or to get over a hang-over (8.7 percent)? Item responses on the CAGE are scored 0 or 1, with a high score indicating the presence of an alcohol problem. We define "problem drinkers" as those persons with a score of 2 or more on the CAGE scale. This is a conservative measure since most of the difference is between 0 and 1. Edwards, Marshall, and Cook (1997) conclude that "these tests (describes another test–MAST as well as CAGE) generally tend to pick up the more extreme rather than the early cases ... they show remarkably good sensitivity and specificity (refers to type 1 and type 2 errors in doctor talk) for 'excessive drinking' as well as 'alcoholism' and may be superior to laboratory tests when used as screening instruments" (p. 197), 23.6 percent

of the males in our sample are classified as problem drinkers.[11]

With regard to past smoking behavior, the HRS asks when former smokers quit. To capture past smoking behavior, we construct a time invariant dummy variable which equals one if the individual quit less than 10 years ago or smoked at wave 1 and is zero otherwise. 34.5 percent of the individuals in the sample are classified as "problem smokers," people who entered the survey with smoking in their recent history.

High levels of alcohol use and smoking are highly correlated in our data and in other samples.[12] Thus, many of the same men who were at risk of poor health because of their alcohol consumption patterns were also at high risk because of their smoking. Although use of each substance carries its own risk when used independently of each other, some joint effects are appreciable.

Self-reported health status is measured on a 5-point Likert scale ranging from excellent (1) to poor (5). Table 1 shows the sample means for each of the five categories as they evolve over time. Self-reported health status declines throughout most of the observational period. However, the sample means do not decline monotonically as one may have expected, especially for adjacent response categories, which may be due to measurement error in the variable. We therefore investigate whether we can aggregate the information in self-reported health status to a simple bivariate health indicator. We find that a health status variable which is equal to one if the individual is in good health (self-reported health status less

---

[11]The CAGE testing procedure is also discussed in King (1986) and Bernadt, Taylor, Smith, and Murray (1982).

[12]See, for example, Rosengren, Wilhelmsen, and Wedel (1988), Schlecht, Franco, J. Pintos, and et. al. (1999) and Palfai, Ostafin, Monti, and Hutchinson (2000).

than 4) and zero otherwise captures the main differences between the individuals in good health and bad health.[13]

We also control for a number of demographic characteristics. Educational attainment is measured by an indicator variable equal to one if the individual has a college education and is zero otherwise. 25.9 percent of men in our sample have a college degree. We also create a variable for the person's expected longevity. We experimented with different family background variables and chose an indicator based on the life span of the mother. 23.8 percent of the sample report that their mothers died by age 70.

We supplement the HRS data with price information for cigarettes, liquor, beer and wine collected by American Chamber of Commerce Researchers' Association (ACCRA) for a large number different localities. A price index for alcohol is constructed by weighting the prices for liquor, beer and wine using the expenditure shares in the Consumer Price Index. We assign alcohol and cigarette prices to individuals in the HRS sample using prices of the nearest locality for which we have ACCRA data.[14]

3   The Framework

Previous analysis of the rational addiction hypothesis has been typically based on versions of the Becker-Murphy (1988) model which rests on the assumption that individuals make continuous choices. As discussed above, the HRS only contains discrete measures of smoking

---

[13]For a discussion of self-reported versus objective measures of health see Bound (1991).

[14]For computational reasons, we need to discretized prices for alcohol and tobacco. We split the sample into 16 subsamples and use the mean prices in the subsamples as points in the grid.

and heavy drinking. We, therefore, adapt an alternative modelling framework based on the dynamic discrete choice literature.

3.1 A Model of Forward-Looking Behavior

Consider the following model of individual decision making. Let J denote the number of alternatives in the choice set (D_t) available to an individual in a given period. In the our model, individuals can engage in heavy drinking or not. They also choose whether or not to smoke.[15] Hence there are four mutually exclusive elements in the choice set. Let $d_{jt} \in \{0, 1\}$ denote an indicator variable which equals one if the individual chooses alternative j at time t and is zero otherwise. Let the vector $d_t = (d_{1t}, ..., d_{Jt})$ characterize choices of an individual at t. Since the alternatives are mutually exclusive,

$$\sum_{j=1}^{J} d_{jt} = 1 \tag{3.1}$$

Besides the control variables, d_t, there is a vector of state variables, s_t. State variables can be decomposed into an unobserved component, ϵ_t, and an observed component, x_t. Observed state variables in our model are the following:

- y: income in 25 states ranging from \$4,000 to \$100,000.

- p_a: price of alcohol in 4 states.

---

[15] As noted in the previous section, heavy drinking and smoking is strongly correlated. Failing to control for either heavy alcohol use or smoking is likely to bias the findings, both regarding the effects of alcohol and tobacco consumption on health, earnings and other outcomes.

- p_c: price of tobacco in 4 states.

- t: age in 20 states ranging from 50-90 years of age.

- h: health status good, bad or dead.

- la: lagged decision to consume 3 or more drinks of alcohol daily.

- lc: lagged decision to consume tobacco.

- as: indicator for problem drinking.

- cs: indicator for problem smoking.

- l: indicator for expected longevity.

- ed: educational status: college education or not.

The construction of the state variables thus mimics the definition of the main variables in the data set discussed in the previous section.

Individuals have beliefs about uncertain future states of the world. These beliefs are captured by a Markov transition density of the state variables $f_t(s_{t+1}|s_t, d_t)$. Let $u_t(d_t, s_t)$ denote the single period utility function which depends on the state and control variables in t. We assume that individuals have preferences which can be represented by a time separable (expected) utility function with constant discount factor β. Individuals are rational and forward-looking and behave according to an optimal decision rule $\delta_t(s_t) = d_t$ which solves

the following maximization problem:

$$\max_{\delta=(\delta_1,\ldots,\delta_T)} E_\delta \left\{ \sum_{t=0}^{T} \beta^t \, u_t(d_t, s_t) \,\Big|\, s_0 = s \right\} \tag{3.2}$$

where E_δ denotes the expectation with respect to the controlled stochastic process $\{s_t, d_t\}$ induced by the decision rule, δ.

3.2 A Parameterization of the Model

Let $u_t(d_t, x_t) = \sum_{j=1}^{J} d_{jt} u_j(x_t)$ denote the single period utility function which depends on the state and control variables in t. We use the following specification:[16]

$$u_j(x_t) = \theta_{0j} + \theta_{1j} \ln(y_t - e(a_{jt}, c_{jt})) + \theta_{2j} \, h_t + \theta_{3j} \, as_t + \theta_{4j} \, cs_t$$

$$+ \theta_5 \, 1\{\text{quit drinking}\} + \theta_6 \, 1\{\text{quit smoking}\} \tag{3.3}$$

where $1\{\cdot\}$ denotes an indicator function which is equal to one if the event inside the brackets happens and is zero otherwise. This specification implies that each choice has a fixed effect denoted θ_{0j}. The utility derived from income, θ_{1j}, and health, θ_{2j}, varies across choices. Utility depends on net income. Unfortunately, the HRS does not contain data on expenditures on alcohol and tobacco; we therefore impute expenditures using the following expression $e(a_{jt}, c_{jt}) = \min\{a_{jt} p_t^a + c_{jt} p_t^c, \alpha y_t\}$.[17] Using the minimum of mean expenditures and a fraction of income, αy_t, guarantees that imputed expenditures on alcohol and tobacco

---

[16]For a discussion of utility functions that depend on health status, see Viscusi and Evans (1990).

[17]a_{jt} (c_{jt}) are the mean quantity of alcohol (tobacco) consumption in the sample for individuals who consume alcohol given choice j. It is zero if choice j is abstaining.

are not excessively high for low income people who are likely to consume cheaper products than the average individual in our sample. Utility also varies systematically among observed types of individuals. Problem drinkers and smokers may have different preferences than other individuals. These differences in utility are captured by θ_{3j} and θ_{4j}.

The last two parameters, θ_5 and θ_6, are designed to capture habit persistence of smoking and heavy drinking. We expect that θ_5 (θ_6) is positive indicating that quitting heavy drinking (smoking) is costly to individuals because of withdrawal effects. Since it is hard to assign quitting costs to the choices, we prefer to think of these parameters of the utility function as measures of habit persistence and estimate them as part of the parameters of the utility function.[18]

Furthermore we assume that the transition probabilities are given by:

$$
\begin{aligned}
f_t(s_{t+1}|s_t, d_t) &= f_{1t}(y_{t+1}\,|y_t, h_t, as_t, cs_t, ed_t, d_t) \times f_{2t}(h_{t+1}\,|h_t, as_t, cs_t, l_t, d_t) \\
&\times f_{3t}(p^a_{t+1}, p^c_{t+1}|p^a_t, p^c_t)
\end{aligned}
\tag{3.4}
$$

The transitions for income are implicitly given by a log-normal regression model. The transition probabilities for health and mortality are logit models. These specifications are implicitly based on a number of plausible, but admittedly *ad hoc* exclusion restrictions. We assume that income does not affect health transitions after controlling for lagged health. Expected longevity affects the health status transitions, but is excluded from the utility and

---

[18]Since discrete choices only depend on relative utility levels, we can normalize a set of coefficients. We set $\theta_{11} = 1$ and $\theta_{i1} = 0$ for $i \neq 1$.

earnings functions. Education only affects earnings. These exclusions give rise to variation in the choice probabilities which helps in estimating some of the structural parameters. Identification is also aided by the inherent nonlinearities of the model. To simplify the analysis, we assume that prices evolve deterministically and that individuals have correct point expectations about current and future prices.[19]

3.3 Estimation

The parameters of our model consist of the parameters of the utility function and the transition probabilities. Let θ denote the parameter vector to be estimated. Since a Markov decision process yields deterministic decision rules, we need to rely on unobserved state variables to generate a properly defined econometric model. Rust (1987) shows that if the unobserved state variables satisfy the assumptions of additive separability (AS) and conditional independence (CI), conditional choice probabilities are well defined. If we additionally assume that the unobservables follow a Type I extreme value distribution, we obtain Rust's multinomial dynamic logit specification:

$$P_t(d_{jt} = 1 \mid x_t) \quad = \quad \frac{\exp(v_{jt}(x_t, \theta))}{\sum_{i=1}^{J} \exp(v_{it}(x_t, \theta))} \tag{3.5}$$

where the conditional value function $v_{jt}(x, \theta)$ is recursively defined as:

$$v_{jT}(x_T, \theta) \quad = \quad u_{jT}(x_T, \theta) \tag{3.6}$$

$$v_{jt}(x_t, \theta) \;=\; u_{jt}(x_t, \theta) \;+\; \beta \int \left(\log \left[\sum_{j=1}^{J} \exp\{v_{jt+1}(x_{t+1}, \theta)\} \right] \right) \; p(dx_{t+1}|x_t, d_t, \theta)$$

Substituting these value functions into equation (3.5) yields the conditional choice probabilities of the dynamic logit model. We observe a panel of N individuals over T periods. The model can be estimated using a Maximum Likelihood Estimator (MLE):

$$\hat{\theta} \;=\; \operatorname*{argmax} \prod_{n=1}^{N} \prod_{t=1}^{T} P_t(d_{nt} \mid x_{nt}, \theta) \; f_t(x_{nt} \mid x_{n,t-1}, d_{nt-1}, \theta) \tag{3.7}$$

To compute the conditional choice probabilities, we need to know the optimal decision rules and hence must solve the decision problem for each individual in the sample repeatedly. To implement the estimator, we can use a two-step procedure: (1) estimate the transition densities for the state variables; and (2) estimate the remaining structural parameters using the conditional choice probabilities.

3.4   Myopic Behavior

As discussed above, heavy drinking and smoking can also stem from myopic behavior which implies that individuals ignore future risks of harmful consumption goods. The myopic model is nested in the forward-looking model described above. In the limiting case of the myopic model, individuals ignore the future.[20] We can obtain this version of the myopic model by setting the discount factor β equal to zero. The choice probabilities are then given

---

[20]We condition our analysis on behavior prior to age 50. Hence, our analysis does not allow us to make statements about behavior of younger adults who may be more likely to engage in myopic behavior. See Gruber and Koeszegi (2000).

by the simple (static) logit model:

$$P_t(d_{jt} = 1 \mid x_t) \;\; = \;\; \frac{\exp(u_{jt}(x_t, \theta))}{\sum_{i=1}^{J} \exp(u_{it}(x_t, \theta))} \tag{3.8}$$

More generally, we can thus distinguish between myopic and forward-looking models based on the value of β. Low estimated values of β suggest myopia; large values of the two-period discount factor β in the magnitude of 0.9 are an indication of "fully rational" behavior. Alternatively, we could vary the planning horizon T and obtain different versions of our model. If T exceeds t by a few time periods, the planning horizon is short. If T is set equal to the maximum life expectancy, we obtain the fully rational model which is typically considered in the rational addiction theory.[21]

3.5   Heterogeneity in Discount Factors

Thus far, we have assumed that individuals are either myopic or forward-looking. A more general framework should control for the fact that individuals differ in their degree of forward-looking behavior. These differences may be due to observable characteristics.

We can extend our framework to allow for observed differences in the planning horizon of individuals. The basic idea is to make the discount factor a function of observables, ph_n, with good predictive power regarding the length of the planning horizon. The HRS contains questions about attitudes toward risk and the length of the planning horizon in

---

[21]The utility function in the simple logit model can also be interpreted as a linear approximation of the conditional value function. Following that line of reasoning, we are then comparing dynamic models based on full solution algorithms with those based on linear approximations of value functions. It is hard to distinguish between the two alternative interpretations of the simple logit model.

wave 1. We use these variables to construct a discrete time-invariant random variable, ph_n, which classifies our sample into two subsamples. It is be reasonable to expect that individuals who differ in their answers to these types of questions, will also differ in their degree of forward-looking behavior. More generally, we assume that the discount factor is a function of observable characteristics, ph_n, and that:

$$\beta_n = \frac{\exp(\gamma_1 + \gamma_2 \ ph_n)}{1 + \exp(\gamma_1 + \gamma_2 \ ph_n)} \qquad n = 1, ...N \qquad (3.9)$$

The logit specification for the discount factor constrains the discount factor to be between 0 and 1. Testing for significance of γ_2 allows us to analyze whether individuals have different planning horizons.

Alternatively, differences in planning horizons may be due to unobserved factors. It is common practice in econometrics to account for unobserved state variables using a non-parametric approach which allows for a finite mixture of types, each comprising a fixed proportion of the population.[22] Following this approach, we assume that there are M discrete types of individuals that differ in their discount factor:

$$\beta_i = \frac{\exp(\gamma_i)}{1 + \exp(\gamma_i)} \qquad i = 1, ...M \qquad (3.10)$$

We then need to solve the model conditional on a discount factor and compute the expected

---

[22]This approach was introduced into the econometric literature by Heckman and Singer (1984). Keane and Wolpin (1997) discuss how to use this approach in a dynamic discrete choice framework. See also Arcidiacono and Jones (2000).

likelihood function:

$$\hat{\theta} \;\; = \;\; \text{argmax} \prod_{n=1}^{N} \sum_{m=1}^{M} \pi_m \prod_{t=1}^{T} P_t^m(d_{nt} \mid x_{nt}, \theta) \tag{3.11}$$

where π_m denotes the probability of type m and superscript m in $P_t^m(\cdot)$ indicates that choices depend on the unobserved type m.[23]

4 Estimation Results: Transition Probabilities

In the previous section, we saw that the behavior predicted by the forward-looking model depends on beliefs that individuals hold about the evolution of the main state variables. In our model, individuals have beliefs about transitions of their health status, mortality, and income. Estimating these beliefs for reasonably parametrized dynamic programming models is a daunting task since beliefs are likely to depend on a number of variables which may only be partially observed by the econometrician. They may also vary among individuals. We must therefore rely on a number of simplifying assumptions to implement our approach.

The main assumption is that individuals subjective beliefs correspond to probability measures that can be estimated based on observed data. Furthermore, we impose additional structure on the functional forms of the transition probabilities to reduce the "curse of dimensionality" encountered in the analysis. In this section, we discuss the main findings of the estimation of the transition probabilities.

---

[23]We can use the same approach to control for unobserved heterogeneity in the transition probabilities. While there are many other ways proposed in the econometric literature to control for unobserved heterogeneity, the procedure adapted in this study is powerful for the types of models estimated here.

Table 3: Transition Probabilities for Health Status, Survival and Income

	Health Status	Survival	Total Household Income
Intercept	5.201	11.695	6.742
	(0.385)	(0.862)	(0.174)
Lagged Income			0.503
			(0.011)
Lagged Bad Health	-2.846	-1.646	-0.242
	(0.071)	(0.151)	(0.025)
Age	-0.045	-0.110	-0.027
	(0.006)	(0.013)	(0.002)
Mother Dead by 70	-0.307	-0.113	
	(0.073)	(0.161)	
Problem Drinker	-0.428	-0.343	-0.036
	(0.073)	(0.156)	(0.021)
Lagged Heavy Drinking	-0.032	0.047	-0.043
	(0.111)	(0.238)	(0.035)
Problem Smoker	-0.310	-0.664	-0.017
	(0.102)	(0.213)	(0.029)
Lagged Smoking	-0.223	-0.382	-0.014
	(0.109)	(0.211)	(0.033)
College			0.288
			(0.024)
Drinking $\times$ College			-0.007
			(0.080)
Smoking $\times$ College			-0.118
			(0.055)
Estimated standard errors are in parenthesis.			

As discussed earlier, we assume that health status and mortality evolve according to logit processes. Transitions on total household income are assumed to follow an AR(1), with normally distributed errors. Estimates of the transition parameters are given in Table 3.[24] Our estimates suggest that lagged health, age and expected longevity (as measured by mother's age of death) are significantly different from zero in all models and have the expected signs. Heavy drinking in the past and smoking have more important negative effects on health than does current heavy drinking.

Previous medical studies have documented the potential negative health effects of heavy drinking among the elderly. Equal doses of alcohol produce higher blood alcohol levels among the elderly compared to the young. Heavy alcohol use increase the risks of injuries and illnesses, including falls, depression, and cognitive impairment, and mortality (Andreasson, 1998). The elderly are also more susceptible to the ill effects of alcoholism because of medication and drug-alcohol interactions (Adams, Gary, Rhyne, Hunt, and Goodwin, 1990). We may not pick up these effects as strongly as do studies based on clinical tests. Our findings are also consistent with the large medical literature documenting adverse effects of smoking on health.

We also estimate logit regressions with survival as the dependent variable. The dependent variable is equal to zero if the individual died in the two years between adjacent waves of the panel and one otherwise. Estimating these transitions is important because smoking and heavy drinking are likely to reduce the life span of an individual, and thus

---

[24]We have estimated a large number of transition model sand only report the findings for those models that we use to estimate preferences.

reduce lifetime utility. Out-of-sample predictions of survival rates are needed to solve the computational dynamic programming model. Accurate estimation of these transitions is therefore an important component of the overall modelling strategy.[25]

Our main findings are quite plausible. Past health and age are significant and have the expected sign. Heavy current drinking has, at best, small effects on the probability of death in the next period. In contrast, past heavy drinking indicated by the coefficient of the problem drinking variable as well as smoking, both currently and in the past, significantly increase the probability of dying in the subsequent period.

Estimating beliefs that individuals hold about income transitions is also complex. Current income of individuals obviously is largely affected by current and past labor market participation, retirement and saving decisions. Modelling these decisions within a well-defined dynamic programming model is exceedingly complicated as documented, for example, by Rust and Phelan (1997). We do not seek to improve upon these efforts. Instead, we analyze interactions between consumption of various harmful goods and their health effects. We therefore adapt a reduced form approach for modelling the income process. The main approach for modelling income transitions is based on log-normal income regressions.

In general, our findings are quite reasonable. The parameters have the expected sign in almost all cases and are often estimated precisely. Not surprisingly, income follows a strong autoregressive process.[26] Past health and age are significant in all three models

---

[25]The main problem encountered in estimating conditional survival rates is that there are only a small number of individuals in the HRS that die during the length of the panel. Brown and Sieg (2001) explore methods based on Imbens and Lancaster (1994) to estimate mortality models more efficiently using a combination of micro data from the HRS and aggregate data form life-tables.

[26]Income is measured as total household income and therefore one would expect a lower coefficient than

and have the expected negative sign. College education is positive and significant. Both past and current drinking, as well as past and current smoking, negatively affect future income. Smoking interacted with college had a particularly strong negative effect on future household income.

In summary, we have shown how to estimate beliefs of individuals based on observed outcomes using a variety of different techniques. Despite the inherent problems of estimating beliefs, our results are quite reasonable. Most of the parameters have the correct sign and are often significantly different from zero. Current and past smoking and heavy drinking affect health status, mortality rates and income even after controlling for lagged dependent variables in the analysis. Given these findings, we assess different behavioral models of high levels of alcohol consumption and smoking.

5   Estimation Results: Preferences

5.1   Myopic and Forward-looking Models

We estimate the preference parameters of the utility function specified in equation (3.3) using myopic and forward-looking model specifications. The myopic model is obtained by setting the discount factor equal to 0.0. The fully rational model uses a two-period discount rate of 0.9. We also estimate an unconstrained dynamic model by searching over a grid for β. Our models control for habit formation in two different ways. We use time invariant measures of problem drinking and smoking. We also control for switching cost by including

estimates typically found for labor income.

lagged choices.

Most of the coefficients in the utility functions have the expected sign and are statistically significant. Our results suggest that, on average, lower income individuals are more likely to engage in smoking and heavy drinking. Including variables for habit formation improves the fit of the model considerably. There are substantial costs for quitting smoking or heavy drinking. Quitting smoking, however, is much harder than quitting heavy drinking. The effect of health on the decisions to drink or smoke point to problems in the fully rational model. Heavy drinking and smoking is more attractive for those in bad health which is counterintuitive.

The values of the likelihood functions indicate that the myopic model ($\beta = 0$) fits the data better than the fully rational model ($\beta = 0.9$). The unconstrained estimate of the model yields a point estimate for the two-period discount factor β of approximately 0.62, which translates into an annual discount factor of approximately 0.78.[27] We also estimate a number of models excluding any measure of habit formation. These models are clearly rejected by the data. Even the fully rational model fits the data much better than these specifications.

To further illustrate the findings, we trace out the likelihood function as a function of β. The results are shown in Figure 1. The likelihood function is relatively flat for a large number of β's, but drastically decreases once the two-year discount factor exceeds 0.65. Thus it appears that a class of models with annual discount factors less than 0.8

---

[27]Discount rates both higher and lower have been reported in the literature. These studies use different approaches for estimating discount rates than we do. See Hausman (1979), Dreyfus and Viscusi (1995), Moore and Viscusi (1988, 1990), and Warner and Pleeter (2001).

Table 4: Parameter Estimates of the Utility Function

	Myopic		Dynamic		Dynamic	
	coefficient	std. error	coefficient	std. error	coefficient	std. error
α	0.1541	0.0062	0.1528	0.0391	0.1533	0.0015
β (two-year)	0.0	——	0.9	——	0.6157	0.0110
c=1, a=1						
intercept	-5.029	0.9545	1.7709	0.7051	-1.3937	0.1503
income	0.8222	0.0883	0.7430	0.0647	0.8287	0.0425
health	-0.2220	0.2031	2.0558	0.1530	0.3074	0.2217
problem d	0.2925	0.1845	-0.1089	0.1033	0.1023	0.1362
problem s	1.6784	0.3443	1.0529	0.2529	1.0742	0.2639 –
c=0, a=1						
intercept	-1.8118	0.6134	1.0371	0.4118	-0.1107	0.4712
income	0.8385	0.0575	0.8246	0.0393	0.8439	0.0405
health	-0.1632	0.1615	0.3705	0.1268	-0.0514	0.1409
problem d	0.2614	0.1290	0.1293	0.0825	0.1905	0.1052
problem s	0.0578	0.1408	0.0982	0.0928	0.0515	0.1169
c=1, a=0						
intercept	-2.4074	0.4282	1.8800	0.3891	-0.1880	0.4368
income	0.8433	0.0418	0.7775	0.0360	0.8477	0.0386
health	-0.1958	0.1159	1.2444	0.0844	0.1744	0.1267
problem d	0.2729	0.1098	-0.0094	0.0497	0.1222	0.0752
problem s	1.4781	0.1443	0.8570	0.0903	0.9403	0.1187
Quitting Costs						
a=1 − > a=0	-3.2993	0.0956	-3.3172	0.0986	-3.3002	0.1006
c=1 − > c=0	-4.2073	0.1336	-4.2120	0.1358	-4.1888	0.1240
likelihood	-12,712		-12,760		-12,705	

explain the main features of the data reasonably well with adequate controls for past habit formation.

FIGURE 1

5.2 Unobserved Heterogeneity and Heterogeneity in Discount Factors

The above models did not account for unobserved heterogeneity among individuals. Failure to account for unobserved heterogeneity may seriously bias the estimates of the parameters of the model. Adding mixture distributions allows for the errors in the transitions to be correlated and also correlated with the errors in the preferences. Further, the mixture distributions allow the errors in the preferences to be correlated over time.

We estimated a mixture distribution with two types and allowed the types to be correlated across all stages of the model. From a variety of starting values, the model consistently converged to a place where over 99% of the observations were of one type. We surmise that this occurs because of the combination of using lagged variables and a short (6-year) panel. Lagged income and health, with the other explanatory variables, account for much of the variation in the transitions. Lagged smoking and excess alcohol use, with variables characterizing problem smoking or drinking, absorbs most the variation in preferences. With a short panel, there is apparently little scope for unobserved heterogeneity. As seen below, even without controlling for unobserved heterogeneity, the model with an estimated discount factor fits the data very well.

We also examined models with observed and unobserved heterogeneity in the discount

factor. Estimating the model with two types where types only differed in the discount factor yielded no change in the estimates. We also tested whether individuals with different cognitive skills, financial planning horizons, and education had different discount factors. In each case, we split the sample into groups: one with high values for the variable (cognitive skills, etc.) and one with low values. Coefficients across the groups were constrained to be equal except for the discount factor. The differences in the discount factors were at most a quarter of a percentage point and not statistically different from one another. Large differences in the discount factor would produce large differences in the drinking and smoking profiles which is inconsistent with the data.

6 Goodness of Fit and Out-of-Sample Predictions

In this section, we investigate the predictive power of the different models that we estimated in the previous sections. Table 5 reports statistics which are commonly used to evaluate the within sample predictions of models.

Table 5 suggests that there are only small differences in the models' predictions. All of the models match the overall trend in the data well. However, heavy drinking and smoking behavior start at higher rates and drop off more quickly, the higher the discount factor. However, we conclude that all four models fit the proportion of individuals heavy drinking and smoking reasonably well.

A more interesting comparison of the different models examines out-of sample predictions. In this study, we focus on predicted heavy drinking and smoking profiles as a function

Table 5: Model Fit

		wave 2	wave 3	wave 4
Proportion Drinking	Myopic	.086	.075	.066
	Fully Rational	.090	.074	.060
	Estimated β	.085	.076	.065
	Data	.091	.069	.066
Proportion Smoking	Myopic	.255	.221	.204
	Fully Rational	.266	.218	.183
	Estimated β	.259	.222	.199
	Data	.256	.230	.203

of age. We run the following experiment. Take the set of individuals in our data and forecast out the probabilities of heavy drinking and smoking at each age. We condition on being alive by dividing the unconditional probabilities of smoking and heavy drinking by the mean probability of being alive at a particular age. That is, individuals with a high probability of dying are weighted less in later years than those who have a low probability of dying. The main results are shown in Figures 2 and 3.

On these same figures, we also plot the profiles for HRS data as well as data from the Asset and Health Dynamics Among the Oldest Old (AHEAD). AHEAD is a companion data set to HRS which contains individuals that were aged 70 or over in 1993, and their spouses who could be any age. In particular, the HRS data are used for ages 52 to 72, and then the wave 1 AHEAD data are attached for ages 72 to 90. One caution is necessary when examining the AHEAD data: the model was estimated with permanent abstainers removed; this was not possible for the AHEAD data.

FIGURES 2 and 3

Vast differences across the models are apparent when examining the age-smoking pro-
files. The myopic model predicts very flat profiles and even slightly increasing smoking rates
between the ages of 55 and 76. As individuals age, they become more unhealthy which makes
smoking less attractive. However, they also become poorer which makes smoking more at-
tractive. This latter effect appears to be slightly stronger until age 76 in the myopic model.
The drop off after 76 occurs because those people who find smoking most attractive are
dying off rapidly.

In contrast, the fully rational model predicts sharp declines in smoking rates. Here,
individuals smoke more than in the myopic model up until the age of 57. The decline
in smoking rates occurs because the marginal adverse health effects of smoking increase
as individuals age. Individuals in the fully rational model know when, and when not, to
smoke. At about the same point as the myopic model shows decreases in smoking rates,
the fully rational model shows *increases* in smoking rates. This occurs because individuals
know they are going to die soon so they might as well enjoy the little time they have left.

The model in which the discount factor is estimated yields profiles in between the fully
rational and myopic model until age 83. After age 83, the model predicts lower smoking
rates than either of the myopic or the fully rational model. This model clearly matches
the trends in the data best, both in and out of sample. That some sort of forward-looking
behavior is present is clear from the upward trend in the smoking behavior at the end of

the life cycle.[28] This holds despite the fact that those who were most attracted to smoking had most likely died at this point. Indeed, an upward trend is not necessary for a model with forward-looking behavior to be correct. Those who have a propensity to smoke die at a higher rate than those who do not.

Although the model with an estimated discount factor predicts smoking behavior well both in and out of sample, this is not true for heavy drinking. In fact, both the myopic and the estimated discount factor model predict very flat age-drinking profiles. The fully rational has individuals engaging in heavy drinking at a much higher rate late in the life cycle which is the exact opposite of what occurs in the data. The data show sharp declines in alcohol consumption later in the life-cycle. However, a deficiency in the construction of the data is important: AHEAD did not have the information needed to remove permanent abstainers. This would obviously affect the profiles for heavy drinking much more than for smoking.

7   Conclusions

This study has analyzed smoking and heavy drinking in a sample of elderly individuals drawn from the HRS. The main objective of the analysis has been to investigate whether models of forward-looking behavior explain the main regularities found in the data better than myopic models. Earlier studies have typically assumed that individuals are either myopic or forward-looking. In contrast, we provide an empirical framework which allows

---

[28]However, only twenty observations were available for the age 90 category.

us to control for varying degrees of forward-looking behavior of individuals and, therefore, nest the competing theories. Estimating the discount factor is clearly desirable and helps us to distinguish between the competing hypothesis which have been put forward to explain the consumption of harmful and addictive goods. Our analysis suggests that models with moderate discount factors above the myopic but below the level commonly used in modelling forward-looking behavior and which control for observed habit formation fit the data the best. They also have the best predictive power for a sample of mean above age 70. Our analysis thus provides strong evidence for the hypothesis that older individuals are forward-looking and take future risks associated with smoking and heavy drinking into consideration when determining their choices. However, events outside a planning horizon of five years are heavily discounted and receive little weight in current decision making.

Our analysis relies on a number of simplifying assumptions. Most of the exclusion restrictions reflect computational considerations. Models based on hyperbolic discounting present an alternative to the two types of models considered here. Our framework could also be used to estimate a behavioral model with hyperbolic discount factors. Future research, with longer panels, should investigate whether these hybrid models provide better fits of the data than models we analyzed. Although estimating hyperbolic-discounting models is feasible, in principle, additional problems arise because of the need to estimate two separate discount factors. Given the availability of only short panels, it may prove to be quite challenging to precisely estimate these types of models. Nevertheless, our results are quite promising for further work which combines formal decision-theoretic analysis and estimation to address questions regarding the degree of forward-looking behavior imposed

in modelling the consumption of potentially harmful and addictive goods.

References

Adams, W., Gary, P., Rhyne, R., Hunt, W., and Goodwin, J. (1990). Alcohol Intake in the Healthy Elderly. *JAGS*, *38*, 211–216.

Andreasson, S. (1998). Alcohol and J-Shaped Curves. *Alcoholism: Clinical and Experimental Research*, *22*, 359–364.

Arcidiacono, P. and Jones, J. (2000). Finite Mixture Distributions, Sequential Likelihood and the EM Algorith. Working Paper.

Becker, G., Grossman, M., and Murphy, K. (1991). Rational Addiction and the Effect of Price on Consumption. *American Economic Review*, *81 (2)*, 237–242.

Becker, G., Grossman, M., and Murphy, K. (1994). An Empirical Analysis of Cigarette Addiction. *American Economic Review*, *84 (3)*, 396–418.

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

Bernadt, M., Taylor, J. M. C., Smith, B., and Murray, R. (1982). Comparison of Questionaire and Laboratory Tests in the Detection of Excessive Drinking and Alcoholism. *Lancet*, 325–328.

Bound, J. (1991). Self-reported and Objective Measures of Health in Retirement Models. *Journal of Human Resources*, *26*, 106–138.

Brien, M., Lillard, L., and Stern, S. (2000). Cohabitation, Marriage, Divorce in a Model of Match Quality. University of Virginia Working Paper.

Brown, D. and Sieg, H. (2001). Estimating the Impact of Smoking on Mortality Using A Combination of Panel and Aggregate Data Sets. Working Paper.

Chaloupka, F. (1991). Rational Addictive Behavior and Cigarette Smoking. *Journal of Political Economy*, *99 (4)*, 722–742.

Chaloupka, F. and Warner, K. (2000). The Economics of Smoking. In *Handbook of Health Economics*, pp. 1151–1627. North Holland.

Choo, E. (2000). Rational Addiction and Rational Cessation: A Dynamic Structural Model of Cigarette Consumption. Working Paper.

Coppejans, M. and Sieg, H. (2001). Price Uncertainty and Rational Addiction. Working Paper.

Dreyfus, M. and Viscusi, K. (1995). Preferences and Consumer Valuations of Automobile Safety and Fuel Efficiency. *Journal of Law and Economics*, *38(1)*, 79–105.

Eckstein, Z. and Wolpin, K. (1999). Why Youth Drop out of High School: The Impact of Preferences, Opportunities and Abilities. *Econometrica*, *67*, 1295–1340.

Edwards, G., Marshall, E., and Cook, C. (1997). *The Treatment of Drinking Problems: A Guide for the Helping Professions*. Cambridge University Press. Cambridge.

Gilleskie, D. (1998). A Dynamic Stochastic Model of Medical Care Use and Work Absence. *Econometrica*, *66*, 1–46.

Gilleskie, D. and Strumpf, K. (2000). The Behavioral Dynamics of Youth Smoking. NBER Working Paper.

Graham, K. (1985). Identifying and Measuring Alcohol Abuse among the Elderly: Serious Problems with Exisitng Instrumentation. *Journal of the Studies of Alcohol*, *47*, 322–326.

Gruber, J. and Koeszegi, B. (2000). Is Addiction Rational? Theory and Evidence. NBER Working Paper 7507.

Harris, C. and Laibson, D. (2001). Dynamic Choice of Hyperbolic Consumers. *Econometrica*, *69 (4)*, 935–958.

Hausman, J. (1979). Individual Discount Rates and the Purchase and Utilization of Engery-Using Durables. *Bell Journal of Economics*, *10(1)*, 33–54.

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

Hodgson, T. (1992). Cigarette Smoking and Lifetime Medical Expenditures. *The Milbank Quarterly*, *70(1)*, 85–125.

Hotz, J. and Miller, R. (1993). Conditional Choice Probabilities and Estimation of Dynamic Models. *Review of Economic Studies*, *60*, 497–529.

Imbens, G. and Lancaster, T. (1994). Combining Micro and Macro Data in Microeconometric Models. *Review of Economic Studies*, *61*, 655–680.

Keane, M. and Wolpin, K. (1994). The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence. *Review of Economics and Statistics*, *76*, 648–672.

Keane, M. and Wolpin, K. (1997). The Carrer Decisions of Young Men. *Journal of Political Economy*, *105(3)*, 473–523.

Khwaja, A. (2000). Health Insurance, Habits and Health Outcomes: A Dynamic Stochastic Model of Investment in Health. Working Paper.

King, M. (1986). At Risk Drinking Among General Practice Attenders: Validation of the CAGE Questionaire. *Psychological Medicine*, *16*, 213–217.

Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly-Journal-of-Economics*, *112(2)*, 443–77.

Lowry, M. (1947). *Under the Volcano*. Reynal and Hitchcock. New York.

Miller, R. (1984). Job Matching and Occupational Choice. *Journal of Political Economy*, *92 (6)*, 1086–1120.

Moore, M. and Viscusi, K. (1988). The Quality-Adjusted Value of Life. *Economic Inquiry*, *26(3)*, 369–388.

Moore, M. and Viscusi, K. (1990). Models for Estimating Discount Rates in Long-Term Health Risks Using Labor Market Data. *Journal of Risk and Uncertainty*, *31(4)*, 381–401.

Orphanides, A. and Zervos, D. (1995). Rational Addiction with Learning and Regret. *Journal of Political Economy*, *103 (4)*, 739–758.

Pakes, A. (1986). Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, *54*, 755–785.

Palfai, T., Ostafin, B., Monti, P., and Hutchinson, K. (2000). Effects of Nicotine Deprivation on Alcohol-Related Information Processing and Drinking Behavior. *Journal of Abnormal Psychology*, *109*, 96–105.

Rosengren, A., Wilhelmsen, L., and Wedel, H. (1988). Separate and Combined Effects of Smoking and Alcohol Abuse in Middle-aged Men. *Acta Med Scand*, *223*, 111–118.

Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurchner. *Econometrica, 55 (5)*, 999–1033.

Rust, J. and Phelan, C. (1997). How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets. *Econometrica, 65*, 781–832.

Schlecht, N., Franco, E., J. Pintos, J., and et. al. (1999). Interaction between Tobacco and Alcohol Consumption and the Risk of Cancers of the Upper Aero-digestive Tract in Brazil. *American Journal of Epidemiology, 150*, 1129–37.

Stigler, G. and Becker, G. (1977). De Gustibus Non Est Disputandum. *American Economic Review, 67*, 76–90.

Thaler, R. and Sheffrin, H. (1981). An Economic Theory of Self-controll. *Journal of Political Economy, 89(2)*, 392–406.

Viscusi, K. and Evans, W. (1990). Utility Functions that Depend on Health Status: Estimates and Economic Implications. *American Economic Review, 80*, 353–375.

Warner, J. and Pleeter, S. (2001). The Personal Discount Rate: Evidence from Military Downsizing Programs. *American Economic Review, 91(1)*, 33–53.

Watson, C., Tilleskjor, C., Hoodecheck-Schow, E., Purcel, J., and Jacobs, L. (1984). Do Alcoholics Give Valid Self-Reports?. *Journal of Studies on Alcohol, 45*, 344–348.

Winston, G. (1980). Addiction and Backsliding: The Theory of Compulsive Consumption. *Journal of Economic Behavior and Organization, 1*, 295–324.

Wolpin, K. (1984). An Estimable Dynamic Stochastic Model of Fertility and Child Mortality. *Journal of Political Economy, 92*, 852–874.

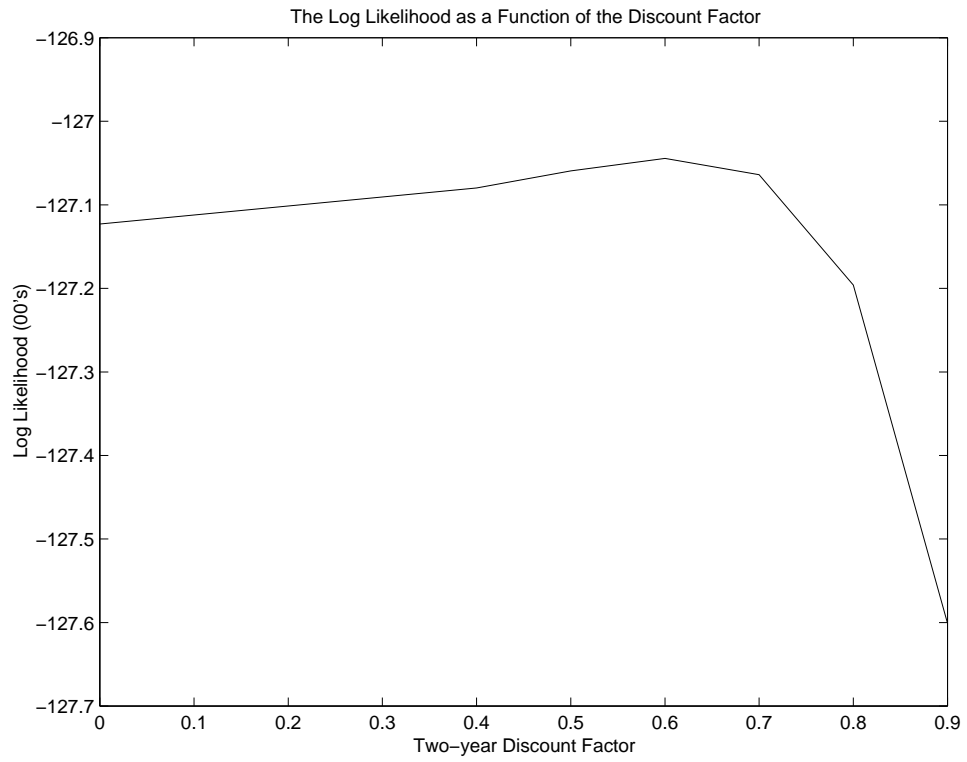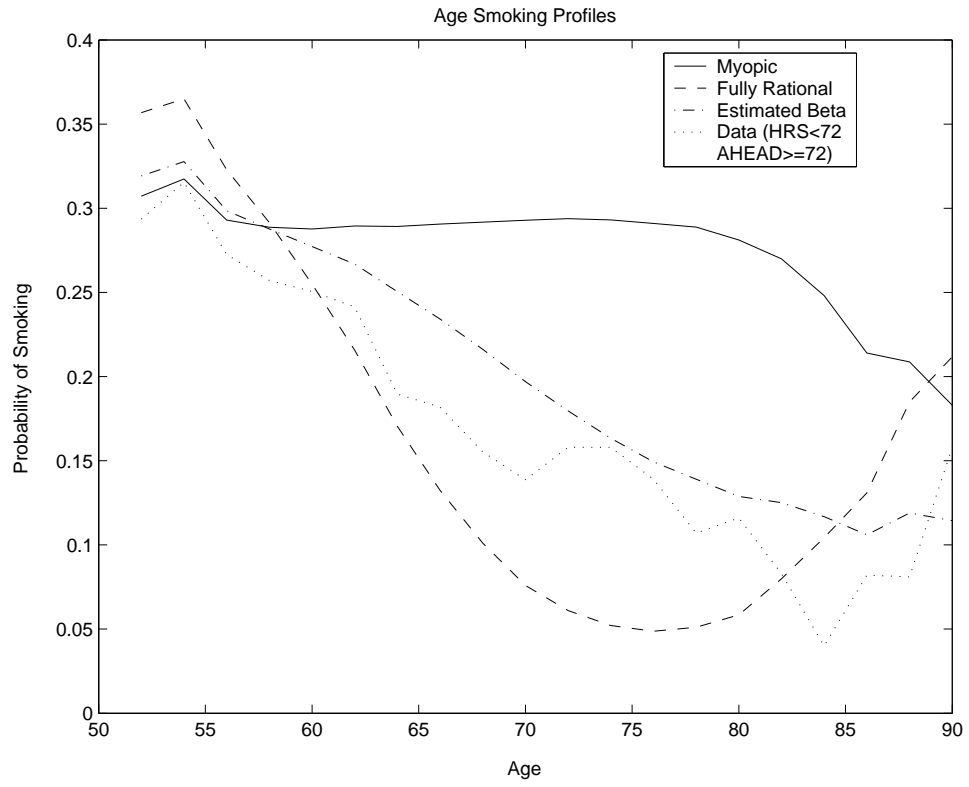Figure 1: The Log Likelihood Function

Figure 2: Age-Smoking Profiles

Age Smoking Profiles

Figure 3: Age-Heavy Drinking Profiles

Age Heavy Drinking Profiles