

Subjective Expectations: Test for Bias and Implications for Choices
(Job Market Paper)

Yang Wang*

November 30, 2008

Abstract

I develop a new method to assess the rational expectations assumption using subjective expectations data. This approach allows the researcher to identify the distributions of the true underlying objective expectations and the expectation bias with respect to unobserved private information. I incorporate subjective longevity expectations into a structural dynamic discrete choice model to explain the smoking behavior of the elderly in the U.S. I find that subjective longevity expectations have important implications for smoking choices. Specifically, the elderly tend to overemphasize the importance of their genetic makeup but underestimate the effects on their longevity of their health-related choices, such as smoking. I also find that the elderly are more concerned with their health and are more forward-looking than under the rational expectations assumption. Policy experiments show that without the individuals' expectation bias the average smoking rate would decrease by 4%.

*I am indebted to my committee members, Frank A. Sloan, Peter Arcidiacono, Patrick Bayer, Hanming Fang, and Han Hong for their invaluable support and guidance. I am grateful to Jason Blevins, Ralph Boleslavsky, and participants at the 2008 Annual Meeting of American Society of Health Economists, the 6th Annual Meeting of Southeast Health Economics Study Group, and the Applied Microeconometrics Lunch Group at Duke University. All remaining errors are mine. Comments are welcome. Contact information: Department of Economics, Duke University, 213 Social Sciences Building, P.O. Box 90097, Durham, NC 27708-0097. Email: yang.wang@duke.edu.

1 Introduction

Expectations about future events play a central role in the decision-making process of individuals who consider dynamic implications of their current choices. How to model these expectations, therefore, is a fundamental issue for economic understanding of individual behavior. Indeed, because observed choice data alone could be consistent with various combinations of expectations and preferences, model predictions and policy implications depend crucially on the underlying assumptions about individual expectations (Manski (2004)).

The existing procedures to recover individuals' preferences from the observed choices are typically based on the rational expectations assumption.¹ This assumption states that agents use all relevant information in forming expectations about economic variables, and their expectations do not systematically differ from the realized outcomes. That is, although the future is not fully predictable, people do not make systematic mistakes when predicting the future, and deviations from actual results are due to random errors. In reality, however, individuals' subjective expectations about the future might be biased, and thus cannot be fully captured by the data on observed outcomes only.² Assuming rational expectations in these cases can lead to model misspecification and deceptive conclusions about individual preferences, which have important implications for economic analysis of the decision-making processes and public policy outcomes.

In this paper, I propose a new theoretical framework to assess the rational expectations assumption using elicited subjective expectations data. First, I establish the conditions for identification of the distributions of the true underlying objective expectations and the expectation bias. I then incorporate subjective expectations data directly into a structural model to infer individual preferences and explain individual behavior. As an application, I show that subjective longevity expectations have important implications for the smoking choices of the elderly in the U.S.

Using the subjective expectations elicited from the agents in economic models of individual behavior is intuitive and economically appealing. However, this approach faces several important challenges. First, subjective expectations data are not widely available for a range of important variables. Even when subjective expectations data are available, they are difficult to evaluate because of the possible measurement errors and the existence of private information. Further, using subjective expectations data to estimate a model of individual choices can be difficult, as it often requests good knowledge of the formation of subjective expectations; in addition, there might be a mismatch between the elicited subjective expectations and the subjective information required by the model.

Much effort has been taken to address these challenges. Since the 1990's, in addition

¹The rational expectations assumption was proposed by Muth (1961) and has been widely used in economics ever since. Recent studies assuming rational expectations include Rust and Phelan (1997), Keane and Wolpin (1997), Gilleskie (1998), and Arcidiacono, Sieg, and Sloan (2007).

²See Van den Steen (2004) for a brief review of different explanations for the existence of individual bias.

to some small-scale surveys, various large-scale surveys have started eliciting subjective expectations directly from respondents. Furthermore, economists have provided various explanations and solutions to account for the seemingly unreasonable responses.³ These developments have spurred a promising new literature on subjective expectations.

One strand of this new literature assesses the elicited subjective expectations, primarily by comparing subjective expectations with historical or actual realizations at the mean or individual level.⁴ The main findings have been quite positive about subjective expectations. Specifically, there are strong reasons to believe that individuals give meaningful answers when their subjective expectations are elicited. Furthermore, the changes in individuals' expectations over time respond in a qualitatively reasonable way to important life events.

The main underlying assumption in these studies is that there is no information in respondents' information sets which is not observed by the econometrician. The question of whether subjective expectations are biased is thus transformed into the question of whether there is any difference between subjective expectations and their "objective" counterparts constructed by the econometrician using only publicly observed information. The conclusions, therefore, depend crucially on the information available to, and used by, the econometrician to form these "objective" counterparts.⁵ Such methods, however, cannot deal with the possibility that individuals, when forming their subjective expectations, possess private information that is not observed by the econometrician but affects their subjective expectations in important ways.

In this paper, I show that subjective expectations data combined with observed realizations can be used to identify the key features of the distributions of the true underlying objective expectations and the expectation bias, which is defined as the difference between the subjective expectations and the true underlying objective expectations. In particular, under certain orthogonality assumptions, the whole distributions of the true underlying objective expectations and the expectation bias can be identified with respect to the unobserved private information.

I further show that after bringing subjective expectations data into a classic binary choice model, it is still possible to identify the model parameters and the distributions of the true underlying objective expectations and the expectation bias, under certain conditions. I then use the data on subjective longevity expectations and smoking decisions of the elderly in

³See, e.g., Hurd and McGarry (2002), Borsch-Supan (1998), Gan, Hurd, and McFadden (2003), Molinari and Manski (2008) and Beresteanu and Molinari (2008).

⁴For example, Hamermesh (1985) and Hurd and McGarry (1995) compare the means of elicited subjective longevity expectations with life tables. Dominitz and Manski (1997), Viscusi (1990, 1991), and Fischhoff, Parker, Bruine De Bruin, Downs, Palmgren, Dawes, and Manski (2000) compare adults or teens' mean expectations about certain life events with realizations of a different but comparable population. Dominitz (1998), Bernheim (1989), Hurd, McFadden, and Merrill (2001), Hurd and McGarry (2002), and Smith, Taylor, and Sloan (2001) compare individuals' expectations with their own realizations on weekly earnings, timing of retirement, onsets of health conditions and mortality.

⁵Cunha, Heckman, and Navarro (2005) develop an alternative approach to deal with individuals' unobserved information.

the U.S. from the Health and Retirement Study (HRS) in a structural and dynamic discrete choice model. In this model, individuals choose from a finite set of actions to maximize their lifetime utilities based on their expectations of future state transitions (Rust (1994)). To model the formation of individuals' subjective expectations about events multiple periods ahead, I set up an Expectation Tree, which solves the mismatch problem between the subjective expectations needed in the model and the ones available in the data.

The World Health Organization (World Health Report, 2002) estimates that more than four million people die from the consequences of smoking each year. The Centers for Disease Control and Prevention therefore describe smoking as the “most important preventable risk to human health” and the world’s leading cause of preventable premature deaths (See, also, Sloan, Ostermann, Picone, Conover, and Taylor (2004)). Due to its profound effects on people’s overall welfare, smoking behavior has been the focus of a large and growing literature in economics.⁶

Besides its importance in welfare, smoking behavior is also an excellent testing ground for the influence of subjective expectations on choices. First, the risks to health from smoking have been well documented and are relatively easy to identify and quantify. Second, the smoking decision is made under considerable uncertainty, with significant pecuniary and nonpecuniary consequences. Expectations about the future after the smoking choice is made, therefore, play a crucial role in the decision-making process.

The empirical results show that expectation biases do exist and play an important role in individuals' decision-making processes. Specifically in this application, individuals' genetic makeup, which is found to be statistically insignificant under the rational expectations assumption, is subjectively significant. Additionally, genetic makeup has an effect on individuals' longevities that is subjectively as significant as their health-related choices, such as smoking. I also find that individuals are subjectively more concerned with their health and more forward-looking than under the rational expectations assumption. The estimated utility parameter for health is about 50% higher, and the estimated discount factor increases from 0.61 to 0.70 when I directly incorporate subjective longevity expectations in the model.

The policy experiments conducted using the subjective expectations data and the inferred preferences further show that, if individuals had unbiased expectations about the marginal effects of smoking on their longevities, possibly following a more personalized information campaign, smoking rates would go down on average by 4% in this age group.

This paper stands at the intersection of economics literatures on subjective expectations, health behaviors, and dynamic discrete choice models, and makes the following main contributions. First and foremost, I propose a new method to assess the rational expectations

⁶See Grossman, Chaloupka, and Anderson (1998), Chaloupka, Tauras, and Grossman (2000), Liang, Chaloupka, Nichter, and Clayton (2003), or Chaloupka and Warner (2000) for reviews on smoking behavior in general. See Sloan and Wang (2008) for a review of literature on mature smokers' addictive smoking behavior. Preventive health choices in general are of great importance to individuals' welfare. For an overview of economic issues related to a variety of preventive health decisions, see Kenkel (2000).

assumption by identifying the distributions of the true underlying objective expectations and the expectation bias with respect to unobserved private information using subjective expectations data and observed realizations.

Second, this study belongs to a growing literature on analyzing individual behavior using subjective expectations data, which combines the revealed preferences approach with the stated preferences approach. Few studies have directly employed subjective expectations data to understand and predict choices. Delavande (2008b) analyzes how perceptions about the benefits and costs of different contraceptive methods affect women’s birth control choices; Lochner (2007) investigates the effects of perceptions of the justice system on youth criminal behavior.⁷ Both studies use a static structural model. To the best of my knowledge, this paper is among the first to use subjective expectations data in a structural and *dynamic* discrete choice model of (health-related) behaviors, and therefore also contributes to the large literatures on dynamic discrete choice models and health behaviors. The empirical results are important for the design of public policies aimed at further reducing smoking rates among adult smokers.

The rest of this paper proceeds as follows. Section 2 introduces the new theoretical framework to assess the rational expectations assumption and to identify the distributions of the expectation bias. Section 3 shows how to identify the preference parameters and the expectation bias when individual choices are introduced into the framework. I describe the data in Section 4, and then apply the theoretical framework to the data to identify the true underlying objective expectation and the expectation bias for survival in Section 5. Section 6 explains the estimation strategy and the empirical specifications. Section 7 presents estimation results and discusses the policy implications. Section 8 concludes.

2 Identification of Objective Expectations and Expectation Bias

In the absence of subjective expectations data, inferring individual preferences from observed choices typically requires strong assumptions about individuals’ expectations, such as the rational expectations assumption. This assumption does not allow for the possibility that respondents’ subjective expectations might be systematically different from the observed outcomes. The violation of this assumption has significant implications for economic understanding of the decision-making processes of the agents; hence, it is important to test for the rationality of individual expectations using available data.

⁷Other recent studies analyzing relationships between subjective expectations and individual behaviors include Nyarko and Schotter (2002) and Dominitz and Hung (2003), using experimental data; and Hurd, McFadden, and Gan (1998), Hurd, Smith, and Zissimopoulos (2004), Viscusi (1990, 1991), and Khwaja, Sloan, and Chung (2007), using survey data on various decisions such as life-cycle consumption patterns and smoking. der Klaauw and Wolpin (2002), on the other hand, treat subjective expectations data as auxiliary data and keep the rational expectations assumption.

The existing methods to assess the rational expectations assumption compare subjective expectations to their objective counterparts based on the public information. In this section, with the same available data, I go one step further and propose a new theoretical framework which can be used to identify the distributions of the true underlying objective expectations and the expectation bias with respect to the unobserved private information. In the next section, I further introduce individual choices to this framework and discuss the identification of the expectation bias and preference parameters.

2.1 Continuous Outcome

I start with the case in which the outcome of interest is a continuous random variable, such as income. The available data include the realizations of and individuals' subjective expectations about this random variable, alongside with other public information.

The key factors to explain the observed realizations and agents' expectations are the true underlying objective expectations, also known as "Nature's Law of Motion" (Manski (1993)); the expectation bias, which is defined as the deviation in individuals' subjective expectations from the true underlying objective expectations; and the realization shock, which is the difference between the true underlying objective expectation and the observed realization. These three components are generally unobserved to the econometrician due to individuals' private information, and have to be identified using the available data on observed realizations and subjective expectations. I show how to extract information on these three components from the two pieces of available information in a theoretical framework below.

Let S_t denote the state variables observed by the agent at time t . I assume a Markovian structure of these state variables, so that S_t fully describes the information set of the agent at time t . Part of S_t , denoted here as X_t , is observable to both the agent and the econometrician.

The expected values of the realizations of observable variables at time $t+1$ under the true objective underlying state transitions are denoted by $E(X_{t+1}|S_t)$. The agent has his/her own subjective expectations about the observable variables $\hat{E}(X_{t+1}|S_t)$, which might deviate from the true underlying expectations. The difference between these two expectations determines the agent's bias $\lambda_t = \lambda(S_t)$:

$$\lambda_t \equiv \hat{E}(X_{t+1}|S_t) - E(X_{t+1}|S_t), \quad (2.1)$$

while the difference between the realized states X_{t+1} at time $t+1$ and their expectations under the true underlying state transition defines the realization shock ξ_{t+1} :

$$\xi_{t+1} \equiv X_{t+1} - E(X_{t+1}|S_t). \quad (2.2)$$

With these notations, I can now establish sufficient conditions for the identification of the distributions of interest.

Proposition 1. *Assume the following conditions:*

1. $E(\lambda_t|X_t)$ is finite.
2. λ_t , $E(X_{t+1}|S_t)$, and ξ_{t+1} are mutually independent given X_t .
3. The random variables possess non-vanishing (a.e.) characteristic functions, conditional on X_t .

Then the conditional distributions of λ_t , $E(X_{t+1}|S_t)$, and ξ_{t+1} given X_t are identified.

Proof. From the definitions of the expectation bias and the realization shock, it immediately follows that

$$\begin{aligned} X_{t+1} &= E(X_{t+1}|S_t) + \xi_{t+1} \\ \hat{E}(X_{t+1}|S_t) &= E(X_{t+1}|S_t) + \lambda_t. \end{aligned} \tag{2.3}$$

where both X_{t+1} and $\hat{E}(X_{t+1}|S_t)$ are known by the econometrician. In addition, ξ_{t+1} is conditionally mean zero following its definition in Eq. (2.2), while the expected value of the expectation bias given the econometrician's information set, $E(\lambda_t|X_t)$, exists due to Assumption 1. Then, according to the theorem by Kotlarski (1967), given Assumptions 1 through 3, the distributions of the three random variables λ_t , $E(X_{t+1}|S_t)$, and ξ_{t+1} conditional on X_t are nonparametrically identified. \square

Assumptions 1 and 3 are technical conditions under which the existence of the conditional moments of the expectation bias and the true underlying state transitions, and the characteristic functions of their distributions given the econometrician's information set is guaranteed. Assumption 2 states that the expectation bias is independent of the true underlying expectation and the realization shock, and weakens the dependence between the underlying expectation and the realization shock to conditional independence. Under these assumptions, all the moments of the distribution of the true underlying state transitions can be obtained by exploring different moments of the joint distribution of the two left-hand side variables in Eq. (2.3), except for the first moment which will follow directly the fact that ξ_{t+1} has zero conditional mean. Once the distribution of the true underlying expectation is identified, the identification of the distributions of the other two unknowns is straightforward.

For an illustration, suppose we are interested in the true underlying objective expectation and the expectation bias for individuals' income levels in the next year, so X_{t+1} are next year's incomes which can be observed ex post in the data. We elicit from individuals their own subjective expectations about future income levels, so $\hat{E}(X_{t+1}|S_t)$ are also known. After

controlling for all the observed individual characteristics which affect individual incomes, such as gender, tenure, education, and occupation, there might still be private information in individuals' information set (S_t) about their future incomes, such as their ability. The complete information set determines individuals' true underlying objective expectations about future incomes ($E(X_{t+1}|S_t)$), which may differ from the realized income levels due to certain random shocks (ξ_{t+1}) such as unexpected absence from work for family reasons. Overly optimistic or pessimistic views about the economy may bias individuals' subjective expectations about future income (λ_t). According to Proposition 1, the econometrician can identify the distributions of the true underlying objective expectations of next year's income, the expectation bias, and the realization shock in the presence of the private information.

Intuitively, say, we want to know the variance of the distribution of the true underlying objective expectations of next year's income. The covariance of the realized and individuals' subjective expectations about next year's income, X_{t+1} and $\hat{E}(X_{t+1}|S_t)$, according to Eq. (2.3), can be written as the covariance of the two terms on the right hand side. Since the three components of these two terms are assumed to be mutually independent, the covariance of X_{t+1} and $\hat{E}(X_{t+1}|S_t)$ is exactly the variance of $E(X_{t+1}|S_t)$, the true underlying objective expectations of next year's income. Based on Kotlarski (1967) theorem, this argument can be extended to identify the whole distributions of the true underlying objective expectation and the expectation bias about future income levels, with respect to the unobserved private information, such as workers' ability.

The identification approach based on Kotlarski (1967) theorem has been applied in various fields including measurement errors (e.g., Li and Vuong (1998)), auctions (e.g., Krasnokutskaya (2003)) and education investment (e.g., Cunha, Heckman, and Navarro (2005)). The novel dimension of this paper is to adopt these methods into a framework to uncover unobserved distributions of underlying true expectations and expectation bias in the presence of private information.

2.2 Binary Outcome

In many cases of interest, such as individuals' survival, workers' job loss, and the arrival of an economic recession, the outcome is binary. In such situations, respondents report subjective probabilities of the occurrences of these events, while the econometrician observes only the binary outcome, e.g., alive or not. This binary case does not fall into the continuous case setup, since now Assumption 2 of Proposition 1 will no longer hold, as the realization shock will have a heteroskedastic variance and therefore cannot be independent of the true underlying objective expectation.⁸ In addition, binary outcomes provide far less information about the underlying distributions than continuous ones. Hence, the binary outcome case requires a special treatment relative to the continuous case outlined in the previous

⁸See, e.g., Mahajan (2006).

subsection.

Consider a binary realization X_{t+1} which takes two values 0 and 1. Call S_t the objective information observed by the agent, and ξ_{t+1} the realization shock. Then, the realization, X_{t+1} , can be written down in the following way:

$$X_{t+1} = I(S_t - \xi_{t+1} > 0). \quad (2.4)$$

I assume that ξ_{t+1} and S_t are independent, and the distribution of ξ_{t+1} , denoted as $\Phi(\xi_{t+1})$, is known to the agent and the econometrician.

Equation (2.4) implies that S_t determines the objective expectation of X_{t+1} given the agent's full information set. That is,

$$E(X_{t+1}|S_t) \equiv Pr(X_{t+1} = 1|S_t) = Pr(\xi_{t+1} < S_t|S_t) = \Phi(S_t). \quad (2.5)$$

The subjective expectation of X_{t+1} can differ from the objective one because of the agent's expectation bias. Denote by $\hat{E}(X_{t+1}|S_t)$ the subjective expectation of the agent given the full information set. Then, I can implicitly define the subjective expectation bias λ_t in the following way:

$$\hat{E}(X_{t+1}|S_t) = \Phi(S_t + \lambda_t). \quad (2.6)$$

Indeed, if the bias is zero ($\lambda_t = 0$), then the agent reports the true objective expectations of his/her survival:

$$\hat{E}(X_{t+1}|S_t, \lambda_t = 0) = \Phi(S_t) = E(X_{t+1}|S_t). \quad (2.7)$$

On the other hand, a positive bias $\lambda_t > 0$ implies that the agent overestimates the objective survival expectations, while a negative bias implies that s/he underestimates them. The following proposition establishes identification results for the distributions of S_t and λ_t .

Proposition 2. *Assume the following conditions,*

1. X_{t+1} and $\hat{E}(X_{t+1}|S_t)$ are observed, while S_t, λ_t and ξ_{t+1} are unobserved to the econometrician.
2. S_t, ξ_{t+1} , and λ_t are mutually independent.
3. ξ_{t+1} has a known distribution.

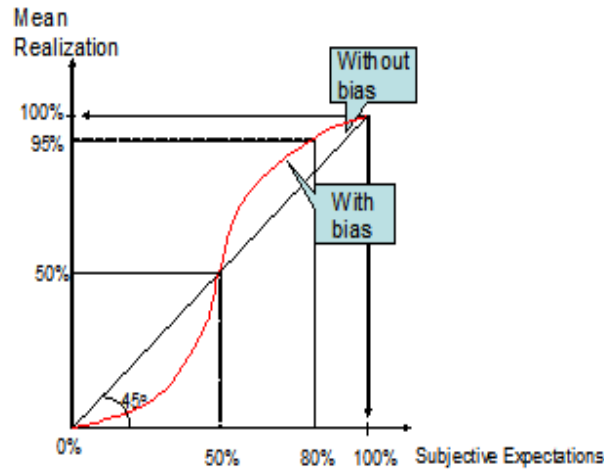
Then the distributions of S_t and λ_t can be identified.

Proof of this proposition and the corresponding estimation strategy is provided in Appendix A.⁹

⁹This proof draws insights from the literature on measurement errors. See, e.g., Li (2002) and Li and Vuong (1998). Chen, Hong, and Nekipelov (2008) provides a survey of nonlinear models of measurement errors. See, also, Heckman and Leamer (2007).

For example, let state X_{t+1} denote individuals' survival status, so that the expected value of this state equals the probability of being alive next period. Figure 1 provides a graphical illustration of the relationship between the expectation bias and subjective expectations in this binary outcome case. The vertical line shows the mean survival probabilities, while the horizontal line shows different levels of subjective survival expectations. Without bias, the plot of the objective versus elicited subjective survival expectations should follow a 45° line. Indeed, if individuals have rational expectations, those who report, say, 80% chance of survival should be alive next period with an 80% chance. With bias, however, this relationship may no longer follow this 45° line, but will lie to its right when agents are overly optimistic or to its left when agents are overly pessimistic about their survival. The red curve on this graph shows one possible relationship between subjective expectations and realizations in the presence of the expectation bias, where those with less than 50% chance of survival tend to be overly optimistic while those with more than 50% chance of survival are more pessimistic.

Figure 1: Realizations and Subjective Expectations

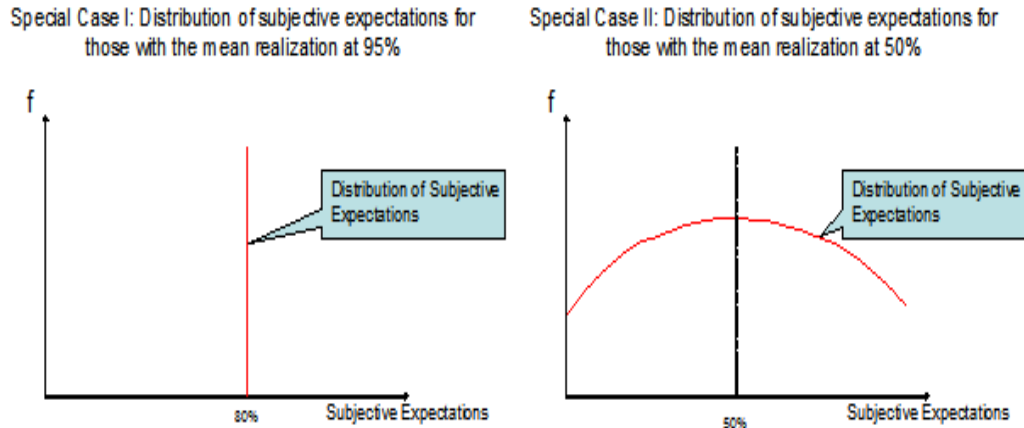


Graphical illustration of the relationship between expectation bias and subjective expectation for univariate binary outcome.

Two particular cases of interest for the distribution of the expectation bias shown graphically in Figure 2 are: 1) $E(\lambda_t|X_t) \neq 0, Var(\lambda_t|X_t) = 0$; and 2) $E(\lambda_t|X_t) = 0, Var(\lambda_t|X_t) > 0$. In the first case, everyone with the same observed characteristics and the same 95% objective chance of survival has the same subjective survival probability of 80%. So, everyone in this group is pessimistic. The econometrician can identify a non-zero systematic bias, which is the same for all agents in this group. In the second case, agents with the same observed states appear unbiased to the econometrician, as their mean objective and subjective survival probabilities are the same (50%). Without looking at the whole distribution of the

expectation bias, the conclusion would have been that individuals in this group have rational expectations. However, agents in this group actually have heterogeneous expectation biases about future states across the unobserved dimension. The proposed method can be used to identify the distribution of the expectation bias, even though it is not possible to pin down the expectation bias at the individual level.

Figure 2: Distributions of Expectation Biases



Graphical illustrations of the distributions of subjective expectation bias with respect to unobserved private information. The left panel shows the distribution of subjective bias with zero variance and non-zero mean; and the right panel shows the distribution of subjective bias with positive variance and zero mean.

3 Subjective Expectations and Individual Choices

In the previous section, I introduced a new theoretical framework to assess the rational expectations assumption using the objective realizations and subjective expectations data. In particular, I provided the conditions for the identification of the distributions of the true underlying objective expectation and the expectation bias with respect to the unobserved private information. The next step is to examine the implications of subjective expectations for individual choices without assuming rational expectations. In this section, I discuss the identification of preference parameters and the expectation bias using individual choices in addition to subjective expectations and observed realizations data. I start with a two-period binary choice model, and then consider a full-blown dynamic discrete choice model.

3.1 Subjective Expectations in a Binary Choice Model

In the current subsection, I consider individual behavior and extend the classic binary choice model, which allows me to identify preference parameters, the true underlying objective

expectations and expectation bias without relying on the rational expectations assumption.

I consider a two-period $(t, t + 1)$ binary choice $(A_t \in \{0, 1\})$ model with subjective expectations. I assume that choice 1 is associated with a higher instantaneous utility and higher probabilities of reaching less favorable states in the next period. Therefore, while the agent is faced with the tradeoff between short-term benefits and long-term costs when s/he makes a choice at time t , s/he will surely choose action 1 once reaching time $t + 1$. The econometrician observes the agent's optimal choice, although to simplify the exposition, I do not include the agent's choice in the vectors X_t or S_t .

I normalize the instantaneous utility associated with choice 0 to be 0, and assume that the instantaneous utility associated with choice 1 is linear in the observed variables X_t ,¹⁰

$$u_{1,t} = \theta' X_t. \quad (3.1)$$

Denote the exponential discount factor by β , and define the additively separable choice-specific utility shocks by ε_0 and ε_1 , respectively.¹¹ The difference in these shocks is given by ε ($= \varepsilon_1 - \varepsilon_0$), and for simplicity, the difference between the utility shocks in the second period is assumed to be 0. The difference in utility shocks (ε) exhibits the appropriate lack of dependence on observable information.

Given the structure of the model, the inter-temporal utilities associated with these two choices at time t satisfy:

$$U_1 = \theta' X_t + \varepsilon_1 + \beta\theta' \hat{E}(X_{t+1}|S_t, 1) \quad \text{and} \quad U_0 = \varepsilon_0 + \beta\theta' \hat{E}(X_{t+1}|S_t, 0), \quad (3.2)$$

where the expectations, $\hat{E}(X_{t+1}|S_t, A_t)$, are the agent's subjective beliefs about future states given his/her current information S_t and potential actions A_t , and β denotes the exogenous exponential discount factor between periods. The optimal choice of the agent, A_t^* , is determined by comparing the two utilities in Eq. (3.2):

$$A_t^* = I(\theta' X_t + \varepsilon + \beta\theta' \hat{E}(X_{t+1}|S_t, 1) - \beta\theta' \hat{E}(X_{t+1}|S_t, 0) > 0), \quad (3.3)$$

where the indicator function $I(\cdot)$ takes the value of 1 if the term inside the brackets, the difference in inter-temporal utilities between these two choices, is positive.

While the agent takes into account his/her subjective expectations for *both* possible actions when s/he makes the decision, the *factual* subjective expectations reported by the agent only correspond to the optimal choice A_t^* :

$$\hat{E}(X_{t+1}|S_t, A_t^*) = \hat{E}(X_{t+1}|S_t, 1) \times A_t^* + \hat{E}(X_{t+1}|S_t, 0) \times (1 - A_t^*). \quad (3.4)$$

¹⁰This linearity assumption is not critical and can be relaxed following arguments in Matzkin (1992, 1993, and 1994).

¹¹This additive separability assumption can also be relaxed following Matzkin (2003).

The relationship between subjective expectations, realization, true underlying expectations and expectation bias can be expressed in a way similar to that in the previous section, recognizing that the econometrician can now observe both X_{t+1} and A_t^* :

$$\begin{aligned} X_{t+1} &= E(X_{t+1}|S_t, A_t^*) + \xi_{t+1}, \\ \hat{E}(X_{t+1}|S_t, A_t^*) &= E(X_{t+1}|S_t, A_t^*) + \lambda_t. \end{aligned} \tag{3.5}$$

Using the same arguments as in Section 2, the econometrician can identify the distributions of ξ_{t+1} , $E(X_{t+1}|S_t, A_t^*)$, and λ_t , conditional on X_t and A_t^* . A more interesting question, however, is to characterize the distributions of these variables conditional on each of the two choices individually, not just upon the optimal one. For example, one of the goals of the model is to learn about individuals' expectations and biases for potential choice 1, which could depend on whether choice 1 is optimal. Another important concern is the identification of the utility parameters (θ) and the discount factor (β). I make the following alternative assumptions on the structure of the model to address these issues.

Assumption 3.1. (*Choice Independent Bias (CI)*) *Individuals' expectation biases are choice-independent: $\lambda_t(S_t, 1) = \lambda_t(S_t, 0)$.*

This first assumption, that individuals' expectation biases depend only on their states, not on their possible actions, is not as restrictive as it may seem. For example, if certain individuals are innately optimistic, then they will always have a rosy view about their probabilities of getting into a good state, no matter what choices they make.

Assumption 3.2. *About the true underlying objective expectations:*

1. (***Strong Irrelevance of Private Information (SIR)***) *The true underlying expectations depend only on observed public information and choices, not on unobserved private information: $E(X_{t+1}|S_t, A_t) = E(X_{t+1}|X_t, A_t)$.*
2. (***Weak Irrelevance of Private Information (WIR)***) *The dependence of the true underlying expectations on private information is the same for both possible choices individuals can make.*

The first statement in the second assumption that the true underlying state transitions do not depend on private information can potentially be strong. However, one nice feature of this assumption is that it is empirically testable. Since the distribution of the expectation biases for certain choices can be obtained using two different methods as explained in Appendix B, significantly different distributions of the expectation biases, resulting from these two methods, imply that the validity of Assumption SIR is in question. The second statement in Assumption 3.2 says that SIR can be relaxed to the extent that private information can affect the true underlying state transitions, but its effects are the same under

both choices. For example, an extremely noisy working environment might be an individual's private information, which will most likely affect this person's health transitions, but its impact can be the same regardless of his choices.

In an ideal but extremely rare case, the econometrician can observe individuals' subjective expectations under both choices. If it is true, then the following proposition establishes the identification results for model parameters and the conditional distributions of the objective expectations and expectation biases:

Proposition 3. *With both factual and counterfactual subjective expectations, and subject to appropriate lack of dependence of the difference in utility shock (ε) on the conditioning variables, conditional independence of expectation bias and underlying expectations given X_t , and existence of necessary moments, supports and characteristic functions,*

1. *Model parameters (θ, β) and the distribution of ε conditional on X_t and elicited subjective expectations are identified.*
2. *Given Assumption CI or SIR, the conditional marginal distributions of subjective expectations, true underlying expectations and expectation biases for both possible actions can also be identified.*

In a more realistic situation, only individuals' factual subjective expectations corresponding to the optimal choices are elicited, then,

Proposition 4. *Subject to appropriate lack of dependence of the difference in utility shock (ε) on the conditioning variables, conditional independence of expectation bias and underlying expectations given X_t , and existence of necessary moments, supports and characteristic functions,*

1. *Assumptions CI and SIR together identify conditional marginal distributions of subjective expectations, true underlying expectations and expectation biases for both possible actions.*
2. *Assumptions CI and SIR together identify model parameters (θ, β) and the distribution of ε conditional on X_t and elicited subjective expectations.*
3. *The conditions for identification of model parameters in 2 can be weakened, i.e., Assumption SIR can be replaced by Assumption WIR, at the expense of the identification of the difference in utility shocks.*

Proof of the above two propositions is shown in Appendices B and C, respectively.

Table 1 summarizes different assumptions for subjective expectations, expectation bias, and private information in different models discussed in this section. To analyze behaviors assuming rational expectations effectively assumes away any possible expectation bias and private information in individuals' subjective expectations. Once the rational expectations

Table 1: Identifying Assumptions for Subjective Expectations

| Components | Assumptions | Under Rational | With Subjective Expectations | | |
|-------------------------|-----------------------------|-------------------------|------------------------------|----------------|---|
| | | Expectations Assumption | Without Behaviors | With Behaviors | |
| Subjective Expectations | Factual | X | X | X | X |
| | Factual + Counterfactual | | | X | X |
| Expectation Bias | None | X | | | |
| | Choice Independent Flexible | | X | X | X |
| Private Information | Strong Irrelevance | X | | X | X |
| | Weak Irrelevance Flexible | | X | | X |

Necessary identifying assumptions on subjective expectations, expectations bias, and private information in models assuming rational expectations (Column 1), allowing for subjective expectations without behaviors (Column 2), and with behaviors (Column 3).

assumption is relaxed, the expectation bias can be dealt with conditional upon available data. With behaviors, certain assumptions about the expectation bias and the private information are required, but these assumptions are still weaker than the rational expectations assumption.

3.2 Dynamic Discrete Choice Model

This subsection extends the two-period binary choice model and incorporates subjective expectations data into a finite-horizon single-agent dynamic discrete choice model.¹² In this class of models, agents choose from a finite set of actions to maximize their expected lifetime utilities based on their expectations of future state transitions. This is also the model adopted in the following empirical analysis. Since the data used in this study, like most available data sets with subjective beliefs, contain only factual subjective expectations corresponding to the optimal choices, I follow Proposition 4 and assume that the expectation bias is choice independent (Assumption CI) and that the true underlying objective expectations do not depend upon the private information (Assumption SIR). In future work, I plan to further relax Assumption SIR by allowing the private information to affect the true underlying objective expectations.

This model is set up with the following components:

- A time index, $t \in \{0, 1, 2, \dots, T\}$.
- A state space, S , consisting of both observable and unobservable parts. Specifically, $S = (x, \varepsilon)$, where x is observable to all, while ε is observable to the agent but not to the econometrician.
- A choice space, A , with a finite number of discrete choices.
- Agents' subjective expectations about state transitions, $\hat{p}(s_{t+1}|s_t, a_t)$.

¹²For detailed reviews of this literature on dynamic discrete choice models, see, for example, Eckstein and Wolpin (1989), Rust (1994), and Akerberg, Benkard, Berry, and Pakes (2005).

- A long-run exponential discount factor, $\beta \in [0, 1]$.¹³
- An instantaneous period utility function, $u(s_t, a_t)$.

Agents have the following inter-temporal utility function at time t :

$$U_t = u(s_t, a_t) + \sum_{\tau=t+1}^T \beta^{\tau-t} u(s_\tau, a_\tau) \quad (3.6)$$

where $u(s_t, a_t)$ is the instantaneous utility at time t , while the second term on the right-hand side is the summation of discounted future utilities from the time $(t + 1)$ forward.

The agent's value function at the time of choice can, in turn, be expressed as the maximum value of the inter-temporal utility function, with the maximization taken over the action space:

$$V(s_t) = \max_{a_t \in A} [u(s_t, a_t) + \beta \int V(s_{t+1}) \hat{p}(s_{t+1} | s_t, a_t) ds_{t+1}].$$

Intuitively, in a dynamic discrete choice model, agents are faced with a set of discrete choices, each of which is associated with a total amount of utility, as shown by Eq. (3.6). As future states are uncertain, the agents need to calculate these utilities based on their expectations about the future given their current choices and states. So, to recover individuals' preferences (u) from observed choices (a), we should use individuals' *own* subjective expectations (\hat{p}) about the future states following their current choices.

Following Rust (1987), I make the following assumptions on the unobservable part in the preferences:

A1. Additive Separability: $u(s_t, a_t) = u(x_t, a_t) + \varepsilon_t(a_t)$.

A2. Conditional Independence:

$$\begin{aligned} \hat{p}(x_{t+1}, \varepsilon_{t+1}(a_{t+1}) | x_t, \varepsilon_t(a_t), a_t) &= q(\varepsilon_{t+1}(a_{t+1}) | x_{t+1}) \hat{p}(x_{t+1} | x_t, a_t) \\ q(\varepsilon_{t+1}(a_{t+1}) | x_{t+1}) &= q(\varepsilon). \end{aligned}$$

A3. Extreme Value Error Distribution: ε is i.i.d with extreme value distribution.

Assumption A1 shows that the period utility function consists of two parts: $u(x_t, a_t)$, which depends only on the observed components – state variable x and choice a at time t , and the unobserved state variable ε .

Assumption A2 places simplifying restrictions on the transition probabilities by assuming that the transition of the unobserved state variable is independent of the observed state variables and the agent's choice, and that the transition probabilities of the whole state space

¹³An alternative approach models individuals as hyperbolic discounters who put less weight on future than in the model considered here and therefore make time inconsistent choices. See Laibson (1997), Harris and Laibson (2001), Fang and Silverman (2007), Gruber and Koszegi (2001), and Fang and Wang (2008).

are multiplicatively separable in the observed and unobserved state variables, conditional upon the agent's lagged choice and observed state variables. Furthermore, lagged unobserved state variables have no implications for the evolution of future state variables.

The distribution of the unobserved state variable is difficult to identify without making strong parametric assumptions about its functional form. This variable is therefore assumed to be known, with little loss of generality. The particular extreme value distribution in Assumption A3 guarantees a closed-form solution to the ex ante value function, also known in the literature as the social surplus function (Rust (1994)), and a convenient logistic functional form for the conditional choice probabilities.

The expected maximum value of inter-temporal utilities is expressed as the ex ante value function through the following relationship:

$$V(x_t) = E_\varepsilon \max_{a_t \in A} [u(x_t, a_t) + \varepsilon_t(a_t) + \beta \int V(x_{t+1}) \hat{p}(x_{t+1} | x_t, a_t) dx_{t+1}]. \quad (3.7)$$

The maximization above is again taken over the choice space, and the expectation is taken over the whole distribution of the unobserved state variable.

Note that the only unobserved part inside the big bracket on the right-hand side of Eq. (3.7) is ε . The remaining observed part is therefore collected and named the *choice-specific value function*:

$$V(x_t, a_t) = u(x_t, a_t) + \beta \int V(x_{t+1}) \hat{p}(x_{t+1} | x_t, a_t) dx_{t+1}, \quad (3.8)$$

which is a function of only the observed state variables (x) and the observed choice (a).

Under Assumptions A1 - A3, Eqs. (3.7) and (3.8) together imply that the ex ante value function is related to the choice-specific value functions through the following expression:

$$\begin{aligned} V(x_t) &= E_\varepsilon \max_{a_t} \{V(x_t, a_t) + \varepsilon\} \\ &= G(V(x_t, a_t), a_t = 0, \dots, \#A) = \ln \left\{ \sum_{a \in A} \exp[V(x_t, a_t)] \right\}, \end{aligned} \quad (3.9)$$

where the first equality in Eq. (3.9) is obtained by replacing the first and third terms inside the big bracket on the right-hand side of Eq. (3.7) using the definition given in Eq. (3.8); while the second equality directly follows the extreme value error distribution of Assumption A3.¹⁴

Without loss of generality, consider a two-choice case where $A = \{0, 1\}$.¹⁵ The literature has made it clear that what can be identified from the data is the difference in the choice-specific value functions (Hotz and Miller (1993)). In particular, under Assumption A3, the difference in the logarithms of the conditional choice probabilities is equal to the difference

¹⁴For a more detailed derivation of Eq. (3.9), see Rust (1987).

¹⁵Generalization to a case with multiple choices is straightforward.

in the choice-specific value functions, denoted here as $d(x)$. That is,

$$d(x_t) \equiv V(x_t, 1) - V(x_t, 0) = \ln P(1|x_t) - \ln P(0|x_t). \quad (3.10)$$

If we further assume that the instantaneous utility functions take on the following linear form,

$$u(x_t, 1) = x'_{1t}\theta_1 \quad \text{and} \quad u(x_t, 0) = x'_{0t}\theta_0,$$

then the utility parameters can be identified through the following relationship, using the difference in the choice-specific value functions, $d(x)$, and the state transition probabilities, $\hat{p}(x_{t+1}|x_t, a_t)$, both of which can be obtained from the data:

$$\begin{aligned} & d(x_t) - \sum_{s=t+1}^T \beta^{s-t} \hat{E} [\log(1 + e^{-d(x_s)})|x_t = x, a = 1] \\ & \quad + \sum_{s=t+1}^T \beta^{s-t} \hat{E} [\log(1 + e^{d(x_s)})|x_t = x, a = 0] \\ & = \sum_{s=t}^T \beta^{s-t} \hat{E} [x_{1s}|x_t = x, a = 1]' \theta_1 - \sum_{s=t}^T \beta^{s-t} \hat{E} [x_{0s}|x_t = x, a = 0]' \theta_0. \end{aligned} \quad (3.11)$$

The derivation of Eq. (3.11) is provided in Appendix D.

We can identify the utility parameters, as described above, for a given discount factor.¹⁶ Magnac and Thesmar (2002) discuss the nonparametric identification of dynamic discrete choice models using a method based on the insights from Hotz and Miller (1993), who view Bellman equations as moment conditions. One of the conclusions Magnac and Thesmar have reached is that the discount factor can be identified if there exist certain exclusive restrictions that shift the expected future utilities (through, say, the transition of the state variables), but do not enter the agents' instantaneous utility functions.

This somewhat abstract argument has an intuitive appeal. Suppose two individuals are identical in all ways except for one. This difference affects their future state transitions but not their instantaneous utilities. That is, this one variable is exogenous to the utility function but still relevant to the state transitions, which makes this variable the exclusive restriction by definition. If these two individuals are completely myopic with a zero discount factor, they should make exactly the same choices because their choices depend only on their identical instantaneous utilities. Therefore, systematic differences in their behaviors should not be expected.

If these two individuals are not completely myopic, then we should expect some differences in their behaviors because their state transitions and therefore the total utilities associated with each choice are different due to the exclusive restriction. The more forward-looking they are, the greater the magnitude of this difference in their behavior. This relationship

¹⁶This study, as most studies in the literature, treats time preference as exogenous. Becker and Mulligan (1997), on the other hand, show a model where time preference can change endogenously as a result of individuals' investment. This possibility is not considered here.

helps pin down the discount factor.

4 Data

This study uses the data from the Health and Retirement Study. The HRS is a nationally representative biennial panel study, in which the baseline interviews were conducted in 1992 (wave 1) with birth cohorts 1931 through 1941 and their spouses, if married. New birth cohorts have been added to the initial sample of 12,652 persons in 7,702 households, and the most recent available data are for year 2006 (wave 8).¹⁷

A. Subjective Longevity Expectations and Smoking Behaviors

To measure the subjective longevity expectations, which are the main focus of this paper, I use the survey answers to the following questions: 1) “*What is the percent chance that you will live to be 75 or more?*” and 2) “*What is the chance that you will live to be 85 or more?*”. These questions are asked only of those under age 65 and 75 at the time of the interview, respectively.¹⁸ For consistency, responses across all waves were re-scaled to fall within $[0, 1]$.

The validity of the elicited subjective longevity expectations might be in question in the following two cases: 1) the response is 0% or 100%, or 2) the elicited probabilities of living to age 75 do not exceed those of living to age 85. These two cases are not mutually exclusive.

The first case might exist because people round their answers to the closest integers.¹⁹ Following the literature (e.g., Khwaja, Sloan, and Chung (2007)), I add 10% (1%) to the response if 0% is the answer to the first (second) question, and subtract 1% (10%) from the response if 100% is the answer to the first (second) question.

The second case usually implies certain mistakes by the respondents or their misunderstanding of the questions. It might be due to rounding again if the responses to the two questions are the same, which I deal with by subtracting 1% from the elicited subjective percent chance of living to age 75 and then using the resulting number as the subjective percent chance of living to age 85, so the probability of living to age 75 is always at least 1% higher than that of living to age 85.²⁰

¹⁷The survey history and design are described in more details in Juster and Suzman (1995). Data flow and other information are also available at <http://hrsonline.isr.umich.edu>.

¹⁸The questions in 1992 were slightly different from those in the following waves, “Using any number from 0 to 10 where 0 equals *absolutely no chance* and 10 equals *absolutely certain*, what do you think are the chances that you will live to be 75(85) or more?”. In addition, from the 5th wave in 2000, the target age in the second longevity question has been based on respondents’ ages at the time of interview.

¹⁹See Molinari and Manski (2008) and Beresteanu and Molinari (2008) for an analysis of the possible rounding problem with subjective probabilities.

²⁰As part of the sensitivity test not shown here, I tried adding 1% to the elicited subjective percent chance of living to age 85 and then using the resulting number as the subjective percent chance of living to age 75. I also tried estimating the whole model excluding those observations with any of these two cases of special responses. Results are robust. Appendix Table 8 shows the percentages of observations in the final analysis sample with any of the special responses. Appendix Figure 11 shows the distributions of both the original and the transformed (‘corrected’) subjective probabilities of living to age 75 and 85.

A response of 50% may also suggest the incapability or carelessness of the respondents to answer those questions, or it may be an expression of uncertainty rather than a quantitative probability (see, e.g., Fischhoff and Bruine de Bruin (1999), and Bruine de Bruin, Fischhoff, Halpern-Felsher, and Millstein (2000)). However, other studies argue that respondents who reported a value of 50% were drawing from nearby probability points (e.g., Hurd and McGarry (2002)), and that 50% could be merely round-off errors instead of cognitive dissonance (Borsch-Supan (1998), p.306). Furthermore, the responses generally are not excessively more bunched at 50% than at adjacent round values (see, e.g., Manski (2004)). I, therefore, do not consider any further adjustment to those responses of 50%.

From the very first survey in 1992, the HRS has asked the respondents about their smoking behaviors using the questions: “*Do you smoke cigarettes now?*” and “*Have you ever smoked before?*” By smoking, the HRS refers to the consumption of more than 100 cigarettes in a respondent’s lifetime, excluding pipes and cigars. Using the answers to these two questions, I categorize the respondents into current smokers, former smokers, and never smokers. Because never smokers might abstain from smoking for various reasons, such as religious beliefs and dislike of the flavor, while the HRS does not provide any information on the exact causes of abstinence, I exclude never smokers from the study.

B. Additional Explanatory Variables

In addition to subjective longevity expectations and smoking choices, I also collect information on individuals’ health, genetic makeup, and other demographic characteristics.

Self-reported health status is measured on a 5-point Likert scale. Respondents were asked, “*What do you think is your current health status: 1. excellent; 2. very good; 3. good; 4. fair, 5. poor?*” To avoid possible measurement errors and to simplify the estimation procedure, I summarize the information on self-reported health status by a binary health indicator that is set to 1 if the respondent is in bad health (fair or poor) and 0 otherwise (excellent, very good, or good).²¹

I also control for a number of demographic characteristics, including the respondents’ age, gender (female or not), race (non-Hispanic White or not), real household income in “1992 dollars” (calculated using the Consumer Price Index), and the longevity of the respondent’s same-gender parent which is summarized by a binary variable set to 1 if the parent is still alive or died at an age greater than 70, and 0 otherwise.²² This last variable is used to capture the respondents’ private information regarding their expected longevity.

The effect of a parent’s death on self-assessed survival probabilities is potentially both biological and psychological. For example, if the parent died of a type of disease known to have a genetic link, the child might reassess his/her own longevity expectancy accordingly. In

²¹See Bound (1991) for a discussion of self-reported and objective measures of health.

²²I also tried using only mothers’ or fathers’ longevity information for all the observations and mothers’ longevity information for males and fathers’ longevity information for females. Results are robust.

addition, a parent’s death may also affect the respondent’s subjective longevity expectations because it reminds the respondent of his or her own mortality. It would admittedly be helpful to know the cause of the same-gender parent’s death, e.g., due to accident, some choice-related sources, or genetic reasons. Unfortunately, the HRS does not provide detailed information on this matter.

Previous research has focused largely on smoking behaviors of young adults (see, e.g., Liang, Chaloupka, Nichter, and Clayton (2003)), yet one important characteristic of cigarette consumption is the long latency period between the time of initiation and the onset of adverse health shocks.²³ And, even for smokers who quit at age 65, Taylor, Hasselblad, Henley, Thun, and Sloan (2002) show that men gained on average 1.4 to 2.0 years of life, and women 2.7 to 3.7 years. Furthermore, Mendez and Warner (2000, 2004) show that the goal of cutting the smoking prevalence among U.S. adults to 12% by 2010 (*Healthy People 2010*, U.S. Department of Health and Human Services, 2000) cannot be achieved unless the rate of smoking cessation among adults increases. I therefore focus on the smoking behaviors of respondents aged 51 to 61.

Our final analysis sample excludes observations with missing values for any of the variables mentioned above.²⁴

Table 2: Summary Statistics

| Variables | Mean | Std. Dev. |
|--|-------|-----------|
| Age | 56.7 | 2.98 |
| Female | 0.53 | 0.50 |
| White | 0.82 | 0.38 |
| Current smoker at period 1 | 0.38 | 0.49 |
| Former smoker at period 1 | 0.62 | 0.49 |
| Same gender parent alive or died at age > 70 | 0.67 | 0.47 |
| Self-rated health at period 1 | 2.65 | 1.16 |
| Bad health at period 1 | 0.23 | 0.42 |
| Household income at period 1 (‘\$k) | 53.4 | 65.6 |
| Self-rated health at period 2 | 2.63 | 1.25 |
| Bad health at period 2 | 0.24 | 0.43 |
| Household income at period 2 (‘\$k) | 51.6 | 68.4 |
| Subjective probability of living to 75 | 0.63 | 0.29 |
| Subjective probability of living to 85 | 0.42 | 0.30 |
| Observed deaths in two years | 0.021 | 0.14 |

Summary statistics of variables for the final analysis sample. HRS panel data from 1992 to 2006.

Table 2 provides the summary statistics for our final analysis sample. Approximately 38% of the observations in the sample are current smokers, and the other 62% have a history of cigarette smoking. About 53% of our sample observations are female, with an average age of 56.7. Around 82% of our observations are Non-Hispanic Whites. 67% of our respondents’

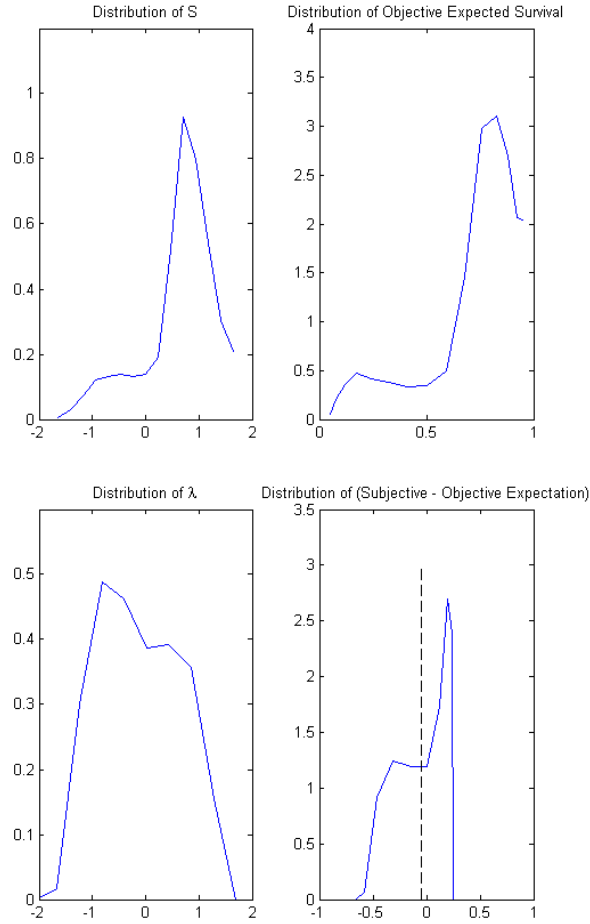
²³At age 35, the cumulative probability of survival is the same for males who have never smoked and smokers. At age 45 (65, 85), the corresponding ratio is 1.02 (1.18, 2.11) (Hodgson (1992)).

²⁴Table 7 compares the final analysis sample with the sample with missing values on subjective longevity expectations and shows that these two samples are not significantly different in important variables such as smoking behaviors and two-year mortality rates.

same-gender parents are still alive or died after age 70. At period one, the average self-rated health level is 2.65, which is between *very good* and *good* health. The average household income is about \$53.4K. At period two, the average self-rated health level drops to 2.63, with the average household income decreasing to \$51.6K. The average subjective probabilities of living to ages 75 and 85 are 63% and 42%, respectively. About 2.1% of the observations do not survive to period 2.

5 Survival Expectations and Bias

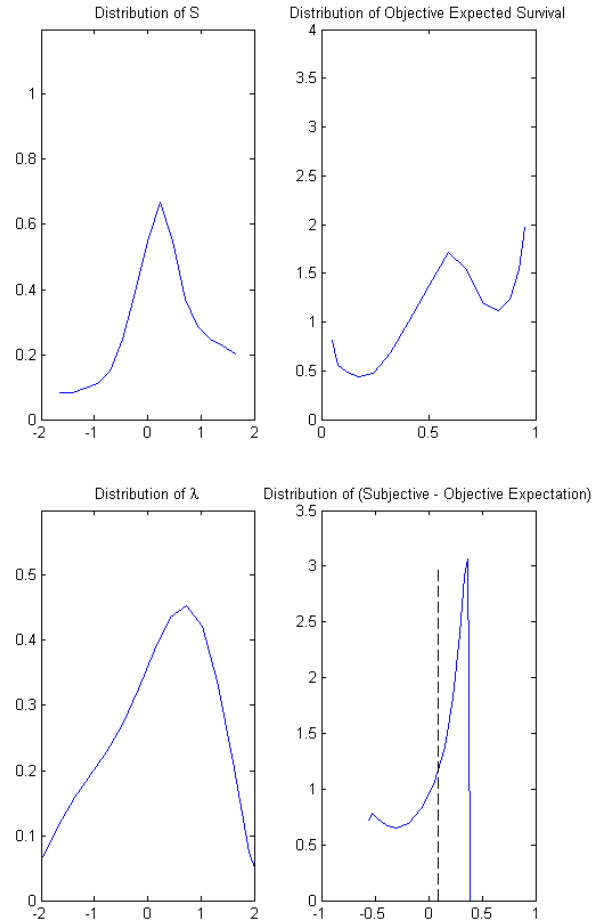
Figure 3: **Objective Expected Survival and Bias, All**



Top row shows distributions of S_t (left panel) and the expected objective survival $E(X_{t+1}|S_t)$ (right panel). Bottom row shows Distribution of λ_t (left panel) and the distribution of the difference between subjective and objective expected survival (right panel) at average S_t , for the whole sample. The dashed line indicates the average difference. Whole sample.

Recall in the binary outcome specification in Section 2, an individual information set S_t determines objective expectations, $E(X_{t+1}|S_t) = \Phi(S_t)$; while the expectation bias λ_t is defined in $\hat{E}(X_{t+1}|S_t) = \Phi(S_t + \lambda_t)$. In this section, I consider the binary outcome of survival status at age 75 for those who were 61 to 65 years old and reported their subjective

Figure 4: **Objective Expected Survival and Bias, Smokers**

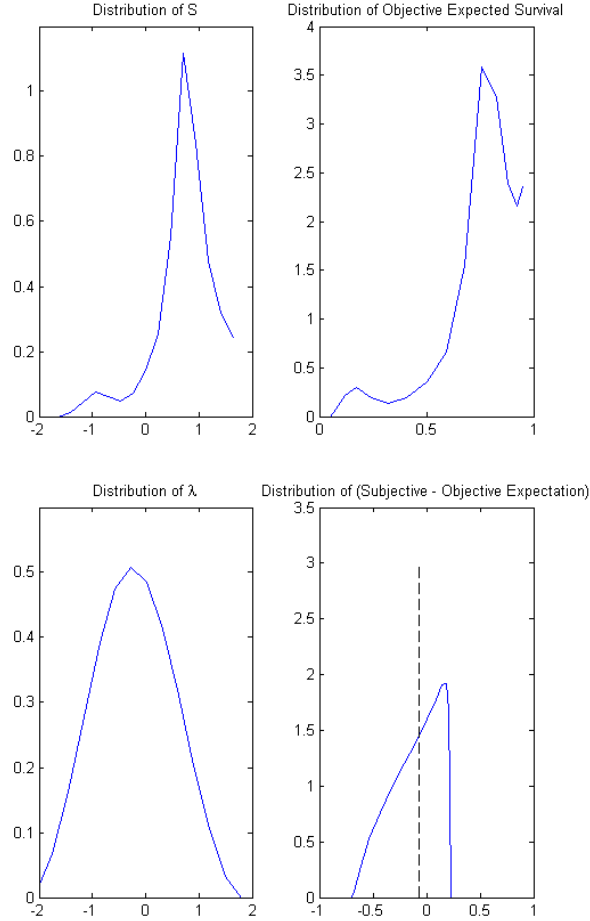


Top row shows distributions of S_t (left panel) and the expected objective survival $E(X_{t+1}|S_t)$ (right panel). Bottom row shows Distribution of λ_t (left panel) and the distribution of the difference between subjective and objective expected survival (right panel) at average S_t , for the sample. The dashed line indicates the average difference. Smokers only.

expectations of surviving to age 75 in 1992. This particular analysis sample is used in this subsection, due to the availability of information on both the observations' survival outcomes and their subjective survival probabilities by age 75. Younger age groups have not reached the target age, so their survival outcomes cannot be observed yet; while older age groups' subjective survival probabilities to age 75 were not elicited. For the estimation of the structural dynamic discrete choice model explained in the following sections, the analysis sample is the one described in Subsection 4.

Figures 3, 4 and 5 show the distributions of the true underlying objective expectations and the expectation bias for the whole sample, smokers, and nonsmokers. Specifically, on the left panels, I plot the distributions of S_t (top panel) and λ_t (bottom panel), respectively. For interpretation purposes, I also show the implied distributions of the objective expectations ($\Phi(S_t)$, top right) as well as the difference between the subjective and objective

Figure 5: Objective Expected Survival and Bias, Nonsmokers



Top row shows distributions of S_t (left panel) and the expected objective survival $E(X_{t+1}|S_t)$ (right panel). Bottom row shows Distribution of λ_t (left panel) and the distribution of the difference between subjective and objective expected survival (right panel) at average S_t , for the sample. The dashed line indicates the average difference. Nonsmokers only.

expectations for average S_t (bottom right) in each sample. The shapes of the distributions are qualitatively similar for all samples. However, overall, smokers are objectively less likely to survive than nonsmokers, as evident from heavier left tails of the distributions of S_t and $\Phi(S_t)$ (top two panels). On the other hand, as shown in the bottom two panels, the expectation bias for smokers is positively skewed with a large mass above zero, implying that the majority of smokers are optimistic about their survival; while the opposite is true for nonsmokers.

I quantify the differences in the means and variances of the true underlying objective expectations and expectation bias in Table 3. As shown in this table, for the whole sample, the mean objective survival expectations is 0.72, with a small negative average expectation bias implying that this group slightly underestimates its survival probabilities by 0.04. Smokers, on the other hand, have a lower average objective survival probability, but tend to

Table 3: Distribution of Objective Expectation and Expectation Bias

| | $E(X_{t+1} S_t)$ | $E(S_t)$ | $Std(S_t)$ | $E(\lambda_t)$ | $Std(\lambda_t)$ | $\Phi(E(S_t) + E(\lambda_t)) - \Phi(E(S_t))$ |
|-----------|------------------|----------|------------|----------------|------------------|--|
| All | 0.72 | 0.63 | 0.67 | -0.13 | 0.78 | -0.04 |
| Smoker | 0.58 | 0.26 | 0.75 | 0.24 | 0.99 | 0.09 |
| Nonsmoker | 0.76 | 0.72 | 0.56 | -0.22 | 0.73 | -0.07 |

$E(X_{t+1})$ is average survival rate in the sample, $E(\cdot)$ and $Std(\cdot)$ are the estimated mean and variance of S_t and λ_t , respectively, while the last column $\Phi(E(S_t) + E(\lambda_t)) - \Phi(E(S_t))$ calculates the average difference between subjective and objective survival expectation. Data are from the 1992 HRS.

overestimate their survival probabilities. On average, the difference between subjective and objective survival probabilities are 0.09 for smokers, relative to -0.07 for nonsmokers who underestimate their survival probabilities. Smokers are also found to have more dispersed distribution of the expectation bias, which has a standard deviation of 0.90 compared to 0.73 for nonsmokers and 0.78 for the whole sample.

Similar analysis can be conducted based on other individual characteristics, e.g., self-reported health status. Respondents who report having bad health on average have lower objective expectations of survival (0.51) than their healthier counterparts (0.79). Individuals with bad health are also found to be subjectively more pessimistic about their survival, with substantially more dispersed expectation bias (standard deviation of 0.88 versus 0.35 for those with good health). Distributions of objective expectations and expectation bias are qualitatively similar in shape to those depicted in Figures 3-5, and are omitted in the interest of space.

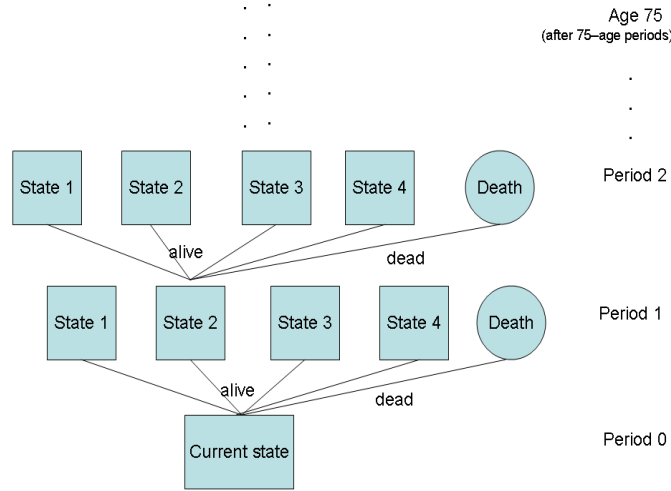
6 Empirical Estimation

6.1 Estimation of Survival Probabilities

The dynamic discrete choice model requires knowledge of one-period ahead subjective expectations of state transitions. The available subjective longevity expectations, however, are about survival probabilities at a certain target age which is at least 14 years in the future for the observations in the sample. To solve this mismatch problem, I propose a new method, an Expectation Tree, to analyze the formation of individuals' subjective expectations of certain events multiple periods ahead. The main characteristic of this new method is that it puts dynamics into individuals' formation of such subjective expectations by explicitly taking into consideration forward-looking agents' uncertainty about future state transitions.

Specifically, according to the Expectation Tree illustrated in Figure 6, when asked at the time of interview (period 0) about their subjective longevity expectations to certain target age, say 75, respondents first think about the probabilities that they will die before the next period (period 1); then, if alive at period 1, the probabilities that they will be in any of the

Figure 6: Expectation Tree: Formation of Longevity Expectations



Graphical illustration of Expectation Tree for survival outcome.

possible (alive) states at period 1. Now, given that they actually reach one of those possible states at period 1, say state 2, again, they think about the probabilities that they will not survive to period 2; and if still alive, the probabilities that they will reach all the different possible states. They keep thinking forward in this way until reaching the target age.

Therefore, the subjective expectations elicited from the respondents, 20 percent or 90 percent, are the weighted summations of probabilities of getting into all the possible states at the target age, taking into consideration the probabilities of dying along the way (from age at period 0 to the target age), conditional upon their states at the time of interview. That is, the elicited subjective expectations are results of forward-looking individuals' careful calculations with all possible situations in the relevant future taken into consideration.

This new dynamic method has the following features. First, this Expectation Tree provides a new way of understanding and analyzing individuals' formation of subjective expectations, which is consistent with the decision-making process of agents in a dynamic discrete choice model. Specifically, this Expectation Tree can also be viewed as the decision tree forward-looking individuals use to make discrete choices by taking the future into consideration. That is, when faced with options A and B, individuals first consider the probabilities of getting into each and every one of all the possible states following either choice. Once in one of the possible states after their choices, they again need to think about the probabilities of getting into all the possible states following options A and B in the next period. They keep thinking forward in this manner and finally make the optimal decisions after comparing the net benefits of the two options.

Note that, in a dynamic discrete choice model, there are utilities attached to all of the possible states, which individuals use to make the optimal choices which have the highest

expected total utilities. When individuals form their subjective expectations, however, they only need to do the forward-looking calculations, no utility maximization needed. This new method provides a connection between individuals' subjective expectations and their discrete choices in a dynamic context where both outcomes – expectations and choices – are generated in the same manner and therefore offer consistent information for the recovery of individual preferences.

Second, this method provides an alternative to the two main methods currently used in the literature in dealing with subjective expectations data: subjective hazard model and Bayesian-style information updating mechanism.²⁵ Compared with the subjective hazard model, this new method does not require any assumption on the functional form of the underlying subjective hazard, but can still be used to infer subjective (longevity) expectations between any two points in time. Unlike the Bayesian-style information updating mechanism, this new method does not require multiple measures of subjective expectations to infer their formation. Notwithstanding these differences, all three methods can complement each other in dealing with subjective expectations data in different contexts.

To estimate the Expectation Tree, first denote by $F(xa, T)$ the probability of surviving another T periods to the target age under current status xa . Recursively, it follows that

$$F(xa, T) = F(xa, 1) * \sum_{xa'} p(xa'|xa)F(xa', T - 1), \quad (6.1)$$

where $F(xa, 1)$ and $F(xa', T - 1)$, by definition, refer to the probabilities of surviving another 1 and $T - 1$ periods, given the current status xa or xa' , respectively. $P(xa'|xa)$ denotes the probability of reaching status xa' , given current status, xa .

Eq. (6.1) shows that the probability of surviving T periods is the product of the probability of surviving to the next period ($F(xa, 1)$) and the weighted summation of the probabilities of surviving the remaining $T - 1$ periods, where the weights are the probabilities of reaching different statuses in the next period given the current status. Note that *status* here includes both the agents' states (x) excluding their vital status (alive or not), and their choices a . That is,

$$p(xa'|xa) \equiv p_a(a'|x') * p_x(x'|x, a, \text{alive}') \quad (6.2)$$

where p_a and p_x are the conditional choice and transition probabilities, respectively. Variables with ' are those in the next period.

²⁵For subjective hazard model, see, e.g., Hamermesh (1985), Khwaja, Sloan, and Chung (2007), Khwaja, Sloan, and Salm (2006), Picone, Sloan, and Taylor (2004), Gan, Gong, Hurd, and McFadden (2004), and Gan, Hurd, and McFadden (2003); for Bayesian-style information updating mechanism, see, e.g., Viscusi (1985), Viscusi and O'Conner (1984), Viscusi and Magat (1987), Smith, Taylor, Sloan, Johnson, and Desvousges (2001), and Lochner (2007). Delavande (2008a) proposes a new metric to measure revisions to subjective expectations.

When $T = 2$, we can write

$$F(xa, 2) = F(xa, 1) * \sum_{xa'} p(xa'|xa)F(xa', 1). \quad (6.3)$$

Once we assume a certain functional form for $F(xa, 1)$ with a set of parameters, θ , we can write $F(xa, 2)$ according to Eq. (6.3), which can in turn be used to write $F(xa, 3)$ in a similar way by replacing the second $F(xa, 1)$ on the right-hand side of Eq. (6.3) with $F(xa, 2)$, and eventually lead us to $F(xa, 75 - \text{age})$. We can then obtain parameter estimates ($\hat{\theta}$) in $F(xa, 1)$ using optimization methods by matching the elicited subjective expectations to the predicted ones.

Eqs. (6.1) and (6.3) show that the elicited longevity expectations are results of subjective expectations of future state transitions and choices. An ideal investigation of subjective expectations would obtain individual assessments of all possible events in different contexts. For many reasons, it is currently impractical to gather such complex information from survey respondents. Therefore, with only the elicited subjective longevity expectations, we need to make certain compromises. That is, to estimate the subjective probability of surviving to the next period ($F(xa, 1)$), which is an essential part of the dynamic discrete choice model of individuals' smoking behavior, we need to estimate the transition probabilities ($p(xa'|xa)$), including the conditional choice probabilities and the conditional state transitions, using only objectively observed data.

6.2 Estimation of Utility Parameters and Discount Factors

For each value of the discount factor β within the reasonable range, we can obtain the parameter estimates, $\hat{\theta}$, from Eq. (3.11) using certain optimization methods. One of the simplest ways is to regress the left-hand side dependent variable which contains only information directly obtained from the data:

$$d(x_t) - \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{-d(x_s)}) | x_t = x, a = 1] \\ + \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{d(x_s)}) | x_t = x, a = 0],$$

on the right-hand side generated regressors:

$$\sum_{s=t}^T \beta^{s-t} E [x_{1s} | x_t = x, a = 1]' \theta_1 \quad \text{and} \quad \sum_{s=t}^T \beta^{s-t} E [x_{0s} | x_t = x, a = 0]' \theta_0,$$

where the only unknowns are θ .²⁶

Given the utility parameter estimates at each possible value of the discount factor, β , we can then use a linear line search to locate the discount factor with the maximum likelihood,

²⁶ Appendix E shows the steps to obtain $\sum_{s=t}^T \beta^{s-t} E [x_a^s | x_t = x, a]$.

where the likelihood is evaluated at both the discount factor and the corresponding utility parameter estimates.

Conditional choice probabilities are flexibly estimated using the Logit model, with up to fourth order polynomials and interaction terms of the state variables. Conditional transition probabilities of the state variables (excluding the vital status) are estimated non-parametrically.²⁷ For comparison purposes, I also estimate the model by assuming rational expectations. The only difference in this objective estimation (since there are no subjective expectations data) is that it uses a Logit model with the observed mortality as the dependent variable for the estimation of survival probabilities.

The estimation steps are:

- Step 1 Flexibly estimate the dependent variable and the regressors in Eq. (3.11). This estimation requires estimations of conditional choice probabilities, conditional health and income transitions, and survival probabilities.
- Step 2 For each discount factor β inside the reasonable range $[0, 1]$, recover $\hat{\theta}$ using Eq. (3.11).
- Step 3 Construct the sample likelihood at β and $\hat{\theta}(\beta)$.
- Step 4 Repeat Steps 2 and 3 for each discount factor β along a line search to find the β with the maximum likelihood.

The estimation procedure described above assumes there are no persistent unobserved differences in the transition of the state variables and in the preferences for smoking and health. It is also possible to account for unobserved heterogeneity following the insights of Heckman and Singer (1984), Keane and Wolpin (1997), Kasahara and Shimotsu (2008), and Hu and Shum (2008), and the methods of empirical implementation suggested by Arcidiacono, Sieg, and Sloan (2007) and Arcidiacono and Miller (2008).

6.3 Empirical Specifications of States and Utilities

The empirical specification follows the framework of the demand for health introduced by Grossman (1972).²⁸ In each period, individuals choose whether to smoke ($a = 1$) or not ($a = 0$), after they observe all the state variables, including the utility shocks, which are unobserved to the econometrician. The complete set of state variables includes whether the individuals are alive or not, whether they are in bad health or not, (the logarithm of) their real household income,²⁹ age, gender, race (Non-Hispanic White or not), and same-gender

²⁷For estimation purposes, I discretize the continuous variable – the logarithm of household income – into a finite number of discrete values. Furthermore, due to the large number of discrete states and the relatively sparse data for each value of the discrete state variables, I kernel smooth over the discretized income levels when I nonparametrically estimate the conditional transition probabilities of the state variables. For details on kernel smoothing, see Appendix F.

²⁸See, also, Becker and Murphy (1988), Becker, Grossman, and Murphy (1991), and Becker, Grossman, and Murphy (1994).

²⁹Household income per capita is also used here as a robustness check.

parents' longevity. This last variable, as mentioned in Subsection 4, is created to control for the differences in the agents' expected longevity due to (unobserved) familial and genetic reasons. This variable also serves as the exclusive restriction for the identification of the discount factor (see Section 3.2). Specifically, I assume that same-gender parents' longevity affect the agents' own expectations about future survival and health, without affecting the agents' instantaneous utilities associated with their smoking choices in a different way.

The instantaneous period utility when the agents choose not to smoke ($a = 0$), depends upon whether they have bad health and (the logarithm of) their household income which is used here to measure the composite good. If the agents decide to smoke ($a = 1$), the period utility depends only on the (logarithm of) their household income. The utility of a deceased individual is normalized to be zero. That is,

$$u_0 = \alpha_0 + \alpha_1 * \text{bad health} + \alpha_2 * \log(\text{household income})$$

$$u_1 = \log(\text{household income}).$$

Agents' uncertainty about future comes from uncertainty about future survival status, whether they will be alive or not, and if alive, future health and income states. The survival probabilities and the conditional probabilities of having bad health and a certain amount of household income in the next period depend upon individuals' choices (whether to smoke or not), the state variables that reflect individuals' natural and biological initial conditions (age, race, gender, same-gender parents' longevity), and the state variables that are linked to the decisions that have been made up to the current period (alive or not, current health status, and household income).³⁰

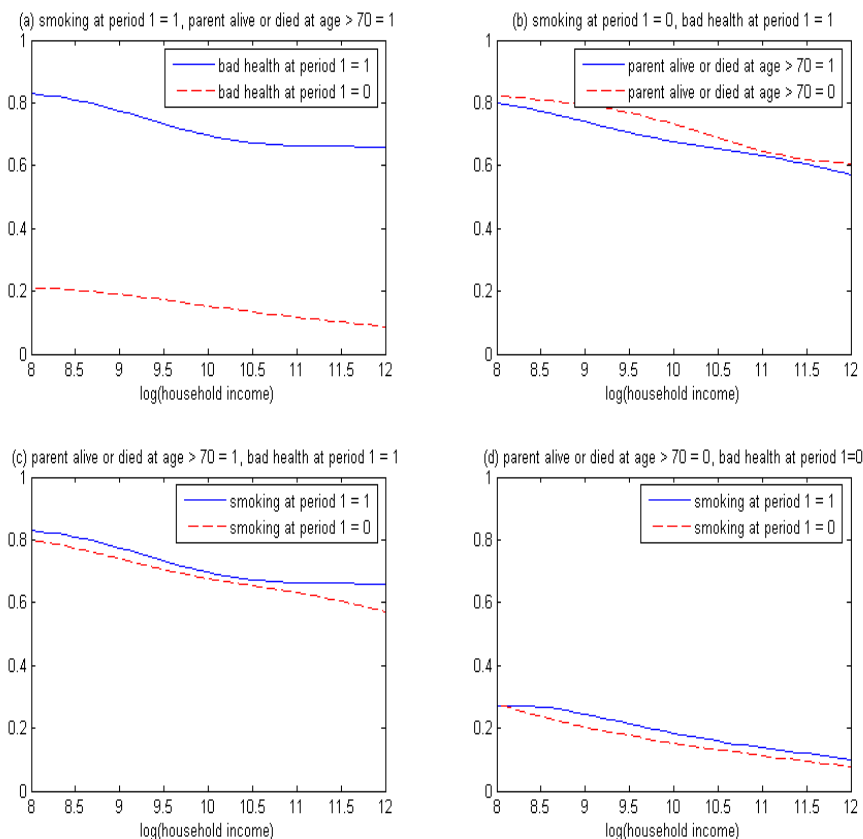
7 Estimation Results

7.1 Conditional Transition Probabilities

Figure 7 shows the estimation results for conditional health transitions. As expected, individuals who are in bad health at period one or do not have long-lived same-gender parents are much more likely to have bad health in the next period (Figure 7, (a) and (b)). Smoking today increases one's chance of having bad health tomorrow, regardless of the current health state (Figure 7, (c) and (d)). Higher household income, as is clear from every panel in Figure 7, predicts a lower probability of having bad health in the next period.

³⁰See Ehrlich (2000) and Ehrlich and Yin (2005) for economic justification for the variables included in this specification.

Figure 7: Health Transition Conditional Upon Survival



Probabilities of having bad health in two years as a function of household income by current health status, same-gender parents' longevity, and smoking status.

7.2 Survival Probabilities

Table 4 reports the estimation results for the probabilities of dying before the next period (in 2 years in this case). The first two columns are the results of the objective estimation based on the rational expectations assumption, with observed deaths before the next period as the dependent variable. The results are as expected: currently smoking, having had health, and aging increase one's probability of dying in 2 years, while White females with higher household incomes and long-lived same-gender parents are less likely to die in 2 years. All estimates are statistically significant at the 1% level, except for the variable measuring same-gender parents' longevity, which, although with the expected sign, is neither economically nor statistically as significant as other factors.

The last two columns of Table 4 report the estimation results of the probabilities of dying in 2 years using subjective longevity expectations data. First, note that each parameter estimate has the same sign as its objective counterpart, except for "being White". That is, people understand correctly the general effects of different determinants on their survival

Table 4: Estimation Results for Probability of Dying in Two Years

| Variable | Obj. Coef. | Std. Err. | Subj. Coef. | Std. Err. |
|--------------------------------|------------|-----------|-------------|-----------|
| Currently smoke | 0.42** | 0.11 | 0.39** | 0.19 |
| Same-gender parent's longevity | -0.10 | 0.11 | -0.33** | 0.04 |
| Bad health | 1.65** | 0.13 | 3.19** | 0.66 |
| Log(household income) | -0.25** | 0.05 | -0.13** | 0.08 |
| White | -0.25 | 0.13 | 0.34** | 0.04 |
| Female | -0.68** | 0.12 | -0.25** | 0.03 |
| Age | 0.08** | 0.02 | 0.07** | 0.02 |
| Constant | -6.13** | 1.30 | -8.01** | 1.53 |

Probabilities are parameterized using Logistic specification. Objective estimation uses a logit model with observed mortality as the dependent variable. Subjective estimation relies on elicited multiple periods ahead subjective longevity expectations using the Expectation Tree method.

*Statistically significant at 5% level; **statistically significant at 1% level.

Table 5: Marginal Effects on Probabilities of Dying in Two Years

| Variable | Obj. Coef. | Std. Err. | Subj. Coef. | Std. Err. |
|--------------------------------|------------|-----------|-------------|-----------|
| Currently smoke | 0.006* | 0.002 | 0.004** | 0.001 |
| Same-gender parent's longevity | -0.001 | 0.002 | -0.004** | 0.001 |
| Bad health | 0.040** | 0.004 | 0.034** | 0.002 |
| Log(household income) | -0.004** | 0.001 | -0.001** | 0.0004 |
| white | -0.004 | 0.002 | 0.004** | 0.001 |
| female | -0.010** | 0.002 | -0.003** | 0.002 |
| Age | 0.001** | 0.0002 | 0.001** | 0.0003 |

Objective estimation uses a logit model with observed mortality as the dependent variable. Subjective estimation relies on elicited multiple periods ahead subjective longevity expectations using the Expectation Tree method. Marginal effects are calculated holding other variables constant at mean levels. Discrete changes for binary variables from 0 to 1 are reported.

*Statistically significant at 5% level; **statistically significant at 1% level.

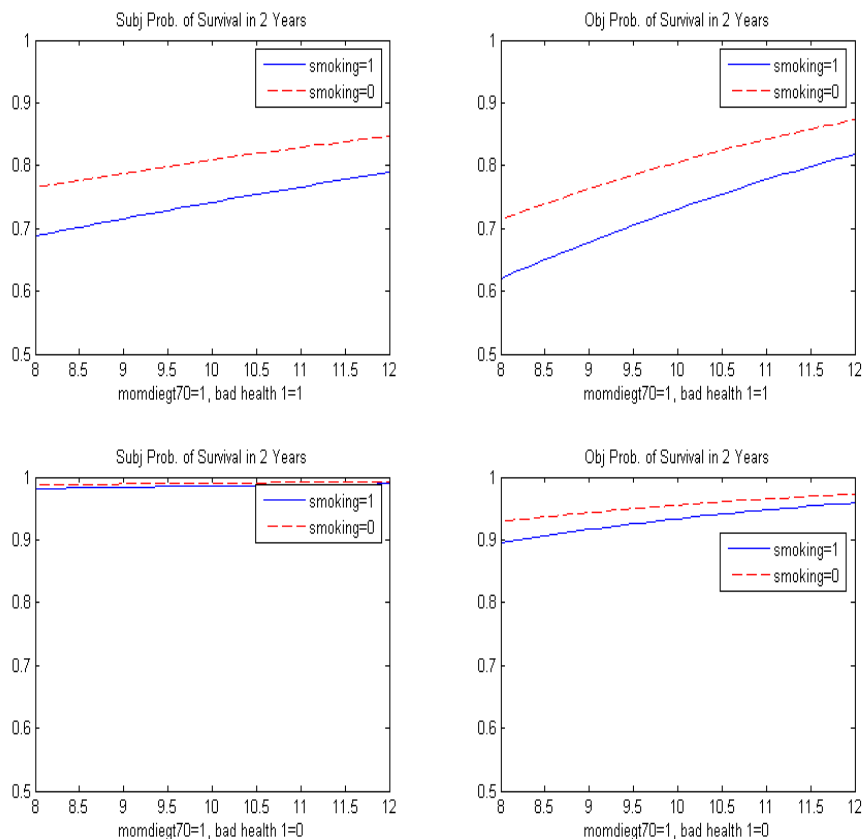
probabilities. For example, having bad health right now will increase their chances of dying in 2 years, and having a long-lived same-gender parent is positively associated with survival. However, Whites in our sample, as mentioned in the previous paragraph, are objectively estimated to be less likely to die, which is consistent with the life tables,³¹ but subjectively, our respondents think that being White lowers ones' chances of survival. The result that individuals have the right idea about the directions of the effects of different determinants is also found by others using different methods for different data sets (see, e.g., Smith, Taylor, and Sloan (2001), Hurd, McFadden, and Merrill (2001), and Hurd and McGarry (2002)).

Because both sets of survival parameters are estimated using nonlinear methods, the marginal effects are also reported (Table 5). The magnitudes of the estimates show that currently being in bad health has the greatest effect on the probability of dying in 2 years, objectively or subjectively. Being a smoker is almost equivalent to adding 4 to 6 years to one's age in terms of its effects on mortality.

One main difference in the relative magnitudes of objective and subjective parameter estimates is noteworthy. Genetic information, summarized by the same-gender parents'

³¹See, for example, <http://www.cdc.gov/nchs/products/pubs/pubd/lftbls/decenn/1991-89.htm> for the life table constructed by CDC.

Figure 8: Two-Year Survival Probabilities for Non-White Males



Probabilities of surviving another two years as a function of household income by smoking status for Non-White males. Left panels: subjective estimation based on the Expectation Tree. Right panels: objective estimations using observed mortality data.

longevity, is given a disproportionately heavy weight subjectively, making it almost as important as current smoking status. One interpretation is that, when forming their longevity expectations, people think that having “good” genes is as important as the health-related choices they make. Alternatively put, people believe that as long as they have good genes, the negative effects of their bad health behaviors such as smoking can be canceled out.

That people tend to overestimate the importance of their parents’ longevity for their own life expectancy was also found by Hamermesh (1985) using a smaller sample of younger age groups. One explanation to this overestimation might be what Tversky and Kahneman (1974) call the “availability heuristic,” an over-reliance on apparently relevant information. However, the fact that individuals form expectations differently than described by the rational expectations assumptions, regardless whether they are justified to think this way, is the main message here.

As a graphical example, Figure 8 shows the 2-year survival probabilities for Non-White males with long-lived parents. The right two figures are recovered from subjective longevity

expectations, while the left two figures are based on objective estimation assuming rational expectations. We can make several observations from this figure. First, subjective survival probabilities are almost everywhere above their objective counterparts, showing that this group is optimistic about their longevity. Second, regardless of their health status and income levels in period one, smokers are estimated to be less likely to survive to the next period, subjectively and objectively. And finally, those who are in bad health in period one (top 2 figures) have a lower survival probabilities than their healthier counterparts.

7.3 Utility and Time Preferences

Table 6 reports the objective and subjective estimates of utility and time preferences. The first similarity between these two sets of estimation results is the negative sign associated with having bad health, implying that bad health hurts nonsmokers more than smokers. Alternatively put, current smokers are less concerned with having bad health than nonsmokers, which could potentially explain their smoking choices in the first place. This result is consistent with what is found in the literature. For example, Khwaja, Sloan, and Wang (2008), using the stated preferences approach, also find that current smokers have substantially lower willingness-to-pay than nonsmokers to be rid of the highly costly and mostly smoking-related disease: Chronic Obstructive Pulmonary Disease (COPD).

Table 6: Utility Parameters and Discount Factor

| Variables | Obj.Estimates | Subj.Estimates |
|-----------------------|---------------|----------------|
| Bad health | -0.46 | -0.77 |
| Log(household income) | 1.35 | 1.33 |
| Constant | -3.48 | -3.42 |
| β | 0.61 | 0.70 |

Objective estimation is based on the rational expectations assumption (Column 1). Subjective estimation uses subjective longevity expectations (Column 2).

The second similarity lies in the sign and magnitude of parameter estimates for the logarithm of real household income: both are positive and greater than 1. Since the utility of current smokers is normalized to be the logarithm of their household income, these estimates show that: 1) marginal utility of income is higher for nonsmokers than for current smokers, and 2) the difference in individuals' instantaneous utilities by smoking status increases with (logarithm of) household income, so individuals with higher household incomes will experience more utility loss if they choose to smoke than their lower-income counterparts. These findings concur with previous studies' results that smoking tends to be more prevalent among low-income consumers,³² and are consistent with the idea of state-dependent utilities introduced by Viscusi and Evans (1990).

³²See, e.g., Wasserman, Manning, Newhouse, and Winkler (1991), Townsend, Roderick, and Cooper (1994), and Chaloupka and Warner (2000).

Two differences in the estimated utility parameters and discount factors between objective and subjective estimations are economically and statistically significant. First, although they share the same sign, the subjective estimation of the parameter for having bad health is more than 60% greater than the objective estimation. This result implies that the difference in how much they care about having bad health between current smokers and nonsmokers is much more significant than we can estimate under the rational expectations assumption. Without subjective expectations data, we would have underestimated this difference.

The second main difference lies in the estimated discount factors. The (two-year) discount factors are estimated to be 0.61 and 0.70 for objective and subjective estimations, respectively. Both are within the reasonable range suggested by the literature, implying that agents in this analysis sample are rather forward-looking.³³ However, the subjective discount factor is much greater than the objective one, showing that individuals are much more forward-looking or patient with time than we would have concluded under the rational expectations assumption using only objective data. Without subjective expectations data, we would have underestimated individuals' patience.

Combining the estimation results on survival probabilities with the estimated utility parameters and discount factors, we can reach two important conclusions. First, individuals care greatly about their health, much more than we would have concluded. Second, individuals are also more patient and pay more attention to their future than under the rational expectations assumption. Both conclusions have important implications as we evaluate the effectiveness of various public policies. One possible explanation of the fact that we can still find so many smokers is that they attach disproportionately large weight to their genetic makeup when they form their longevity expectations. They even think that "good" genes can protect them against the detrimental effects of harmful health behaviors, such as smoking.

This finding that there is still gap between individuals' subjective beliefs in the relative effects of their smoking choices and other characteristics on their survival and health and those objectively identified calls for further effort in policy intervention such as information campaign. Specifically, more personalized anti-smoking messages which highlight both the absolute and the relative importance of quitting smoking to mortality and morbidity may further shrink the gap between subjective and objective expectations and lead to lower smoking rates.³⁴

7.4 Goodness of Fit

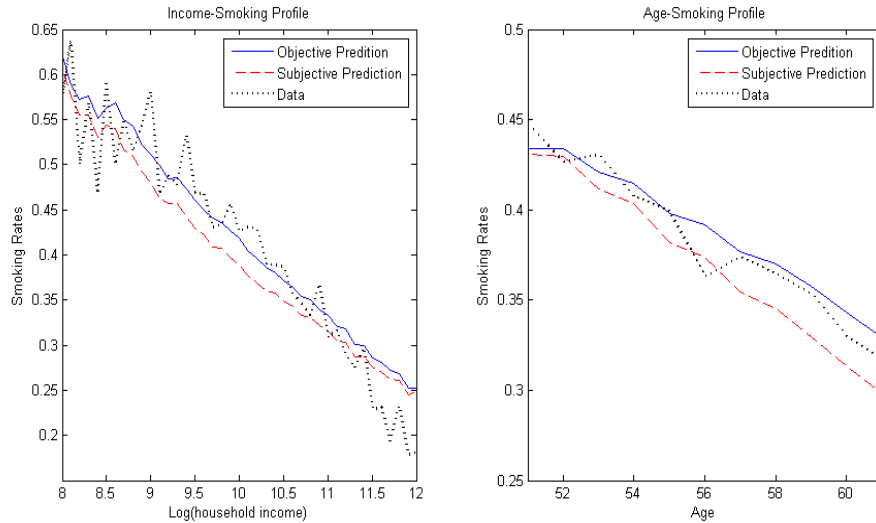
Figure 9 shows two measures of within sample goodness of fit of the estimated parameters. Specifically, I compare the predicted smoking rates using both subjectively and objectively

³³See Appendix Table 9 for examples of discount factors estimated in the literature.

³⁴Estimation results with unobserved heterogeneity are work in progress.

estimated parameters with those from the data at different ages and (logarithm of) household income levels. These two graphs show that both estimation methods have done a great job in matching smoking rates with the data: the predicted smoking rates are right around the true smoking rates from the data, and both methods capture the decreasing trend in smoking rates over age and income levels.

Figure 9: Within-Sample Goodness of Fit



Model implied smoking probabilities based on the rational expectations assumption, subjective expectations data, and from the data. Left panel depicts the smoking rates with respect to household income; right panel depicts the smoking rates with respect to age.

The finding that estimation using subjective longevity expectations data alone, without any objectively observed information on mortality, can still perform very well in fitting the data is consistent with the literature. Gan, Gong, Hurd, and McFadden (2004), using the data from Asset and Health Dynamics Among the Oldest Old, also find that parameter estimates using subjective mortality risk perform better in predicting out-of-sample wealth levels than estimates using life table mortality risks. Similarly, Lochner (2007) shows that the perceived crime rate among youth can better predict their criminal behavior than the official neighborhood crime rate.

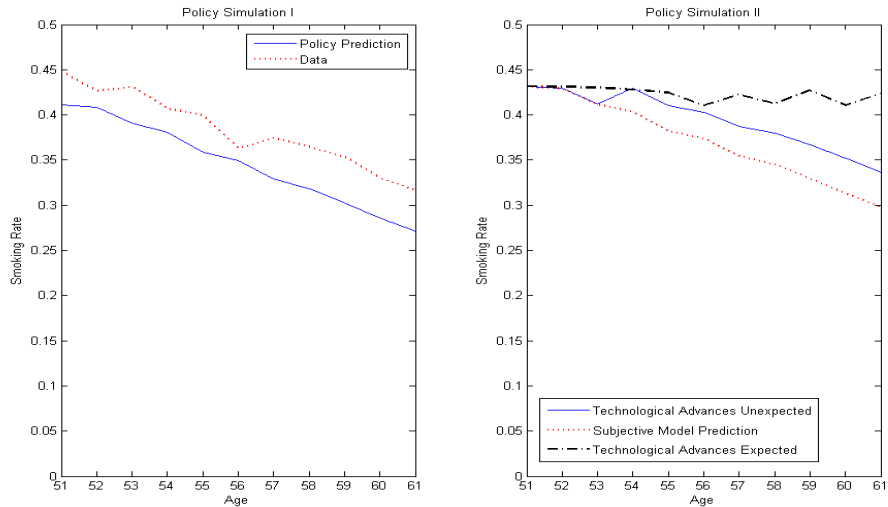
7.5 Policy Implications

Subjective expectations data and the estimated utility and time preferences can be used to assess by simulation precisely what the impact would be of different policy interventions on smoking choices. As an illustration, I consider below two specific public policies. The first one is an information campaign, which has been proven to be effective in informing individuals of the harmful effects of smoking and consequently reducing the smoking rate. The case we study here is when the information campaign is so effective that subjective marginal effects of smoking on survival are exactly the same as those objectively estimated

under the rational expectations assumption. This policy experiment aims at showing the effects of (at least partially) reducing the expectation bias.

The second policy we consider is the effects of advances in medical technology. We focus on changes in the age-smoking profile given anticipated or unanticipated advances in medical technology, which are assumed to make smoking less detrimental to health by lowering the harmful effects of smoking on survival by 50 percent. We focus on people aged 51 and assume that the advances in medical technology occur when they are at age 54.

Figure 10: Policy Experiments



Left panel: smoking rates from the data and those predicted by policy experiment which sets the subjective marginal effects of smoking at the objectively estimated levels. Right panel: smoking rates from the data and those predicted by policy experiments with expected and unexpected technological advances which cut down the negative effects of smoking by half.

The left panel of Figure 10 shows that, if individuals' subjective expectations about the marginal effects of smoking on their survival are exactly like the objective ones, the smoking rate for the age group studied in this paper will decrease. This drop in smoking rate ranges from 2% to 6% with an average of 4%.

The right panel of Figure 10 shows the results of the second policy experiment. As expected, if individuals anticipate the advances in medical technology, the smoking rate will go up even before the advances actually occur. This is because individuals expect the health costs of smoking to be smaller in the future, which makes smoking a more attractive option now. If individuals do not anticipate the advances in medical technology, then the smoking rate will not jump before the advances occur. Anticipated or not, the medical advances will increase the smoking rate, but the smoking rate will still decrease with age.

8 Conclusion and Discussion

In this paper, I present a new approach to assess the rational expectations assumption using subjective expectations data. I also show how to relax the rational expectations assumption by bringing subjective expectations data into discrete choice models to improve our inference on choice behaviors. As one of the few studies using subjective expectations data in a dynamic structure model, this paper shows that these data are critical in conducting analysis under weaker assumptions than what is usually imposed in the literature.

My central empirical finding is that individuals care more about their health and are more forward-looking than we would have concluded under the rational expectations assumption. To recover these true preferences, I analyze individuals' own longevity expectations and find that expectation bias does exist. Specifically, individuals put disproportionately heavier weights on genetic information, yet pay less attention to the effects of their health related behaviors. Policy experiments further show that if individuals do not have biased expectations about the marginal effects of smoking on their mortality, the average smoking rates will be around 4% lower than their current values.

Individuals in our model are assumed to have possibly biased yet mature opinions on the effects of different determinants on certain outcomes. This assumption can be justified in the current sample of the elderly in the U.S. who one can argue have passed the initial information-gathering stage. An interesting extension of this model is to allow for learning or belief revision which could play an important role in explaining adolescent behaviors.³⁵

The empirical results are specific to the empirical question asked and the analysis sample used in this paper. The method proposed and the model used here, however, can be generalized to other decision-making processes such as consumption and saving patterns, decisions on labor supply and education investment choices.

Even though this paper is able to relax certain strong assumptions about individuals' expectations, it still requires other assumptions in its estimation mainly due to the available subjective expectations data. For example, the availability of counterfactual subjective expectations data would allow for more flexible forms of expectation bias or a greater role of private information in ones' decision-making processes.

³⁵See, e.g., Viscusi and O'Conner (1984), Viscusi (1985), Viscusi and Magat (1987), Bernheim (1990), and Delavande (2008a) for analysis of belief revision.

Appendices

A Proof of Proposition 2

Proof. For simplicity, I will drop time subscripts t and $t + 1$ in this proof. Denote $\hat{z} = \Phi^{-1}(\hat{E})$. Then, as $\Phi(\cdot)$ is monotonic and known, \hat{z} is effectively observable, so we can rewrite the equation (2.6) in the following way.

$$\hat{z} = S + \lambda. \quad (\text{A.1})$$

Due to monotonicity of $\Phi(\cdot)$, the information in \hat{E} is the same as in \hat{z} . It is, however, easier to work with the transformed variable due to linearity.

Then, from equation (2.4),

$$\begin{aligned} P(X|\hat{z}) &= P(S - \xi > 0|\hat{z}) = P(\varepsilon < S|\hat{z}) = \int P(\xi < S|S, \hat{z})f(S|\hat{z})dS \\ &= \int \Phi(S)f(S|\hat{z})dS = \int \Phi(S)\frac{f(S, \hat{z})}{f(\hat{z})}dS = \int \Phi(S)f(S)f_\lambda(\hat{z} - S)dS\frac{1}{f(\hat{z})} \end{aligned}$$

So,

$$P(X|\hat{z})f(\hat{z}) = \int \Phi(S)f(S)f_\lambda(\hat{z} - S)dS$$

Take Fourier Transform of both sides. Since the left-hand side is observed, $\int P(X|\hat{z})f(\hat{z})e^{it\hat{z}}d\hat{z}$ is observed and is equal to the convolution between $\Phi(S)f(S)$ and $f(\lambda)$:

$$\int P(X|\hat{z})f(\hat{z})e^{it\hat{z}}d\hat{z} = \int \Phi(S)f(S)e^{itS}dS \int f_\lambda(\lambda)e^{it\lambda}d\lambda \quad (\text{A.2})$$

where $\int \Phi(S)f(S)e^{itS}dS$ is defined as $Ee^{itS}\Phi(S)$, and $\int f_\lambda(\lambda)e^{it\lambda}$ is defined as $Ee^{it\lambda}$. That is, Fourier Transform of convolution is equal to the product of Fourier transforms.

Now, from Eq. (A.1), $Ee^{it\hat{z}} = Ee^{itS}Ee^{it\lambda}$. So, define

$$\Phi_{\hat{z}}(t) = \frac{\int P(X|\hat{z})f(\hat{z})e^{it\hat{z}}d\hat{z}}{Ee^{it\hat{z}}}, \quad (\text{A.3})$$

where $\Phi_{\hat{z}}(t)$ is observable from the data. From (A.2) it follows that

$$\Phi_{\hat{z}}(t) = \frac{\int \Phi(S)f(S)e^{itS}dS}{\int f(S)e^{itS}dS}, \quad \forall t,$$

which leads to

$$\int (\Phi(S) - \Phi_{\hat{z}}(t))f(S)e^{itS}dS = 0, \quad \forall t. \quad (\text{A.4})$$

Eq. (A.4) can be considered a “linear functional relation” of $f(S)$, which must be satisfied for all t , subject to the integrability constraint on $f(S)$. Once $f(S)$ is identified, identification of the distribution of λ is straightforward. \square

Equation (A.4) provides a basis for empirical identification of the unknown distribution $f(S)$ in the data. In the empirical analysis, I take $\Phi(\xi)$ to be standard Normal distribution. I discretize the support of $f(S)$ into N equidistant points $\{S_1, S_2, \dots, S_N\}$ with increment ΔS , and consider M possible values for t $\{t_1, t_2, \dots, t_M\}$.

Denote the vector of values of $f(S)$ at the selected points $f(\mathbf{S})$:

$$f(\mathbf{S}) = \left[f(S_1) \quad f(S_2) \quad \dots \quad f(S_N) \right]'. \quad (\text{A.5})$$

Let us construct a complex-valued $M \times N$ matrix H , whose (j, k) element is

$$H(j, k) = (\Phi(S_k) - \Phi_z(t_j)) e^{it_j S_k} \Delta S. \quad (\text{A.6})$$

This matrix can be constructed from the data; further, given equation (A.4), for $N \rightarrow \infty$ it should be true that

$$Hf(\mathbf{S}) \rightarrow 0. \quad (\text{A.7})$$

For a fixed N , I minimize the sum of squared products of real and imaginary parts of rows of H and the unknown function $f(\mathbf{S})$, subject to the integrability constraint on $f(\mathbf{S})$.

Denote $W = \begin{bmatrix} \text{Real}(H) \\ \text{Imag}(H) \end{bmatrix}$, and let W_k denote the k th row of W . I then select $f(\mathbf{S})$ according to the following objective function

$$\min_{f(\mathbf{S})} \sum_k (W_k f(\mathbf{S}))^2, \quad (\text{A.8})$$

subject to the integrability constraint $\sum_n f(S_n) \Delta S = 1$. Taking first-order conditions, the optimal solution to $f(\mathbf{S})$ is then given by

$$f(\mathbf{S}) \propto \left(\sum_k W_k' W_k \right)^{-1} \times \mathbf{1}, \quad (\text{A.9})$$

where the proportionality coefficient is chosen to ensure integrability of $f(\mathbf{S})$ to 1.

In principle, this approach permits the estimation of the unknown distribution function of S , $f(\mathbf{S})$, for an arbitrarily large grid N . In practice, I find that this approach works quite well for N around 10 – 30, and for a 30 – 50 discrete values of t ranging from 0 to about 10 – 20 (negative values of t are redundant in W).

After I identify the distribution of S , it is straightforward to obtain the distribution of the expectation bias (λ) using inverse Fourier methods. Indeed, I can estimate the characteristic

function of λ using

$$Ee^{it\lambda} = Ee^{it\hat{z}} / Ee^{itS}, \quad (\text{A.10})$$

and then recover the probability distribution function $f_\lambda(\lambda)$ using inverse Fourier transform:

$$f_\lambda(\lambda) = \frac{1}{2\pi} \int e^{-it\lambda} Ee^{it\lambda} dt \quad (\text{A.11})$$

B Proof of Proposition 3

Proof of Statement 1: In the ideal case that econometrician observes both factual and counterfactual subjective expectations, the only unobserved variable in the choice equation (3.3) is ε , and the binary choice problem falls into a standard framework analyzed by Manski (1988), Matzkin (1992) and Cunha, Heckman, and Navarro (2005), among others. The conditions provided in these works (e.g., independence of ε from observed state variables) can be used for identification of model parameters θ and β , as well as the conditional distribution of ε in Eq. (3.3) given X_t , $\hat{E}(x_{t+1}|S_t, 1)$ and $\hat{E}(x_{t+1}|S_t, 0)$.³⁶

Proof of Statement 2: When both factual and counterfactual subjective expectations are elicited, the conditional distributions of the true underlying state transition, the expectation bias, and the realization shock given X_t and A_t are identified for $A_t = (0, 1)$, following the arguments in Section 2.

Let us first add Assumption 3.1 on choice-independent expectation bias. Then,

$$\begin{aligned} \hat{E}(X_{t+1}|S_t, 0) &= E(X_{t+1}|S_t, 0) + \lambda_t \\ \hat{E}(X_{t+1}|S_t, 1) &= E(X_{t+1}|S_t, 1) + \lambda_t. \end{aligned} \quad (\text{B.1})$$

The conditional mean of λ_t given X_t can be identified using the optimal choice and expectations data in the following way,

$$E(\lambda_t|X_t) = E(\lambda_t|X_t, A_t^* = 1) \times Pr(A_t^* = 1) + E(\lambda_t|X_t, A_t^* = 0) \times Pr(A_t^* = 0).$$

As the conditional distribution of λ_t given X_t and optimal choice A_t^* is identified (see Section 3.1) and the optimal actions are observable, $E(\lambda_t|X_t)$ is identified as well. Hence, the means of the conditional distributions of $E(X_{t+1}|S_t, 0)$, $E(X_{t+1}|S_t, 1)$ and λ_t in Eq. (B.1) are identified. Applying the Kotlarski theorem, the marginal distributions of these variables can be identified given that $E(X_{t+1}|S_t, 0)$, $E(X_{t+1}|S_t, 1)$ and λ_t are mutually independent (conditional on X_t) and characteristic functions exist.

Now, let us use Assumption SIR instead. Then, the underlying expectation $E(X_{t+1}|S_t, A_t)$ is equal to $E(X_{t+1}|X_t, A_t)$, and can be directly identified from the objective data by look-

³⁶See Appendix A in Heckman and Navarro (2006) for a complete list of Matzkin's conditions.

ing at the optimal choices of the agents. The difference between objective and subjective expectations for the two choices allows to identify the distribution of the expectation bias.

C Proof of Proposition 4

Proof of Statement 1: Without counterfactual subjective expectations, Assumptions CI and SIR can be combined to identify the underlying true expectations $E(X_{t+1}|S_t, A_t)$ ($= E(X_{t+1}|X_t, A_t)$) from the objective data and agents' optimal choices (Assumption SIR). Matching factual subjective and objective expectations, we can further identify the expectation bias, which by Assumption CI is independent of the choice.

Proof of Statement 2: Under Assumptions CI and SIR, the counterfactual subjective expectations can be obtained using the choice-independent bias and the identifiable underlying true expectations. Once the counterfactual subjective expectations are known, we can bring them back to Eq. (3.3) and the identification of the model parameters and the distribution of ε proceed as discussed above.

Proof of Statement 3: Without Assumptions CI and SIR, non-parametric identification of the distributions of the expectation bias, the realization shock, and the true underlying expectations conditional upon *optimal* choices (A_t^*) are still possible based on the arguments in Section 2. However, identification of model parameters (θ, β) and the distribution of ε in Eq. (3.3) are in general only possible to a certain extent.

One alternative is to keep Assumption CI and to relax Assumption SIR in the following way. Write the true underlying objective expectation as:

$$E(X_{t+1}|S_t, A_t) = E(X_{t+1}|X_t, A_t) + \nu(S_t, A_t),$$

where $\nu(\cdot)$ denotes the only part in the true underlying objective expectation that depends on the unobserved private information in S_t . Then, the difference in the true underlying expectations by choices, $\Delta_{1-0}E(X_{t+1}|S_t)$, is:

$$\Delta_{1-0}E(X_{t+1}|S_t) = \Delta_{1-0}E(X_{t+1}|X_t) + \Delta_{1-0}\nu(S_t, A_t). \quad (\text{C.1})$$

Given Assumption CI, we can identify the model parameters (θ, β), as well as the distribution of the term ($\varepsilon + \beta\theta'\Delta_{1-0}\nu(S_t, A_t)$), assuming $\varepsilon + \beta\theta'\Delta_{1-0}\nu(S_t, A_t)$ has zero median conditional upon X_t .

Indeed, since Assumption CI still hold, using Eq. (C.1) in which the first term on the

right-hand side can be obtained from the data, we have:

$$\Delta_{1-0}\hat{E}(X_{t+1}|S_t) = \Delta_{1-0}E(X_{t+1}|S_t) = \Delta_{1-0}E(X_{t+1}|X_t) + \Delta_{1-0}\nu(S_t),$$

which can be used in Eq. (3.3) to give us:

$$A_t^* = I(\theta'X_t + \beta\theta'\Delta_{1-0}E(X_{t+1}|X_t) + \varepsilon + \beta\theta'\Delta_{1-0}\nu(S_t)) > 0).$$

According to Manski (1988), Matzkin (1992) and Cunha, Heckman, and Navarro (2005), if we know that $\varepsilon + \beta\theta'\Delta_{1-0}\nu(S_t)$ has zero median conditional upon X_t , we can identify θ , β , and the distribution of $\varepsilon + \beta\theta'\Delta_{1-0}\nu(S_t)$.

D Derivation of Eq. (3.11)

Eq. (3.9) implies that the ex ante value function can be expressed as a function of choice specific value functions, which can be further written as a function of the differences in choice specific value functions ($d(x) \equiv V(x, 1) - V(x, 0)$) and one baseline choice specific value function.

With two choices, we can write Eq. (3.9) in the following way

$$\begin{aligned} V(x_t) &= \log \left\{ \exp(V(x_t, 1)) + \exp(V(x_t, 0)) \right\} \\ &= \log \left\{ \exp(V(x_t, 1) - V(x_t, 0)) + 1 \right\} + V(x_t, 0) \\ &= \log \left\{ \exp(V(x_t, 0) - V(x_t, 1)) + 1 \right\} + V(x_t, 1) \end{aligned} \quad (\text{D.1})$$

where the first equality is from the two-choice assumption, and the second and third equalities are obtained by adding and subtracting $\ln(\exp(V(x_t, 0)))$ and $\ln(\exp(V(x_t, 1)))$, respectively.

By inserting the second line of Eq. (D.1) into Eq. (3.8) and rearranging the terms, we can have the following expression with $d(x_t)$

$$\begin{aligned} &V(x_t, 0) - \beta \int V(x_{t+1}, 0)\hat{p}(x_{t+1}|x_t, 0)dx_{t+1} \\ &= u(x_t, 0) + \beta \int \log(1 + e^{d(x_{t+1})})\hat{p}(x_{t+1}|x_t, 0)dx_{t+1} \end{aligned} \quad (\text{D.2})$$

where the second term on the right hand side can be obtained directly from the data. Taking this term as given, the left hand side subsequently defines a backward induction relation that recursively expresses $V(x_t, 0)$ for all t as functions of $u(x_t, 0)$ only.

Specifically, in the last period, the continuation value is equal to the instantaneous utility, for all x and a

$$V(x_T, a_T) = u(x_T, a_T) \quad (\text{D.3})$$

For any period $t < T$, we can rewrite Eq. (D.2) and apply it recursively by moving the time index forward

$$\begin{aligned}
V(x_t, 0) &= u(x_t, 0) + \beta \int \log(1 + e^{d(x_{t+1})}) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&\quad + \beta \int V(x_{t+1}, 0) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&= u(x_t, 0) + \beta \int \log(1 + e^{d(x_{t+1})}) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&\quad + \beta \int u(x_{t+1}, 0) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&\quad + \beta^2 \int \int \log(1 + e^{d(x_{t+2})}) \hat{p}_{t+1, t+2}(x_{t+2}|x_{t+1}, 0) \hat{p}_{t, t+1}(x_{t+1}|x_t, 0) dx_{t+2} dx_{t+1} \\
&\quad + \beta^2 \int \int V(x_{t+2}, 0) \hat{p}_{t+1, t+2}(x_{t+2}|x_{t+1}, 0) \hat{p}_{t, t+1}(x_{t+1}|x_t, 0) dx_{t+2} dx_{t+1} \\
&= \dots
\end{aligned}$$

which can be (relatively more succinctly) written as:

$$\begin{aligned}
V(x_t, 0) &= u(x_t, 0) + \beta \int \log(1 + e^{d(x_{t+1})}) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&\quad + \beta \int V(x_{t+1}, 0) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&= u(x_t, 0) + \beta \int \log(1 + e^{d(x_{t+1})}) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} \\
&\quad + \beta \int u(x_{t+1}, 0) \hat{p}(x_{t+1}|x_t, 0) dx_{t+1} + \beta^2 \int \log(1 + e^{d(x_{t+2})}) \hat{p}_{t, t+2}(x_{t+2}|x_t, 0) dx_{t+2} \\
&\quad + \beta^2 \int V(x_{t+2}, 0) \hat{p}_{t, t+2}(x_{t+2}|x_t, 0) dx_{t+2} \\
&= \dots
\end{aligned}$$

which leads to the following backward induction relation

$$\begin{aligned}
V(x_t, 0) &= \sum_{s=t}^T \beta^{s-t} E[u(x_s, 0)|x_t = x, a = 0] \\
&\quad + \sum_{s=t+1}^T \beta^{s-t} E[\log(1 + e^{d(x_s)})|x_t = x, a = 0] \quad (D.4)
\end{aligned}$$

In the expectation here, the distribution of x_s given $x_t = x, a = 0$ is induced by the transition of the state variables x from t to s as if action $a = 0$ is always taken between periods t to $s - 1$.

Analogously, if we instead insert the third line of Eq. (D.1) into Eq. 3.8, then we have

$$\begin{aligned}
&V(x_t, 1) - \beta \int V(x_{t+1}, 1) \hat{p}(x_{t+1}|x_t, 1) dx_{t+1} \\
&= u(x_t, 1) + \beta \int \log(1 + e^{-d(x_{t+1})}) \hat{p}(x_{t+1}|x_t, 1) dx_{t+1} \quad (D.5)
\end{aligned}$$

And for any period $t < T$, Eq. (D.5) translates into the following backward induction relation

$$\begin{aligned}
V(x_t, 1) &= \sum_{s=t}^T \beta^{s-t} E[u(x_s, 1)|x_t = x, a = 1] \\
&\quad + \sum_{s=t+1}^T \beta^{s-t} E[\log(1 + e^{-d(x_s)})|x_t = x, a = 1] \quad (D.6)
\end{aligned}$$

In the above conditional expectations, the distribution of x_s given $x_t = x, a = 1$ is induced by the transition of the state variables x from t to s as if action $a = 1$ is always taken between periods t to $s - 1$.

Given that the data identify $d(x_t) (\equiv V(x_t, 1) - V(x_t, 0))$, taking the difference between Eqs. (D.4) and (D.6), and rearranging terms shows that the difference

$$\sum_{s=t}^T \beta^{s-t} E [u(x_s, 1) | x_t = x, a = 1] - \sum_{s=t}^T \beta^{s-t} E [u(x_s, 0) | x_t = x, a = 0]$$

is identified through the data as

$$\begin{aligned} d(x_t) - \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{-d(x_s)}) | x_t = x, a = 1] \\ + \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{d(x_s)}) | x_t = x, a = 0]. \end{aligned}$$

Given the linear form of the instantaneous period utility functions, difference between Eq. (D.5) and Eq. (D.2) can be written as

$$\begin{aligned} d(x_t) - \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{-d(x_s)}) | x_t = x, a = 1] \\ + \sum_{s=t+1}^T \beta^{s-t} E [\log(1 + e^{d(x_s)}) | x_t = x, a = 0] \\ = \sum_{s=t}^T \beta^{s-t} E [x_{1s} | x_t = x, a = 1]' \theta_1 - \sum_{s=t}^T \beta^{s-t} E [x_{0s} | x_t = x, a = 0]' \theta_0 \end{aligned} \quad (D.7)$$

which is exactly Eq. (3.11).

E Derivation of $\sum_{s=t}^T \beta^{s-t} E [x_a^s | x_t = x, a]'$

With discrete state variables, we can write the “regression” relation in Eq. (D.7) as

$$\begin{aligned} d(x_t) - \sum_{s=t+1}^T \beta^{s-t} P_1^{t,s} \log(1 + e^{-d(x_s)}) + \sum_{s=t+1}^T \beta^{s-t} P_0^{t,s} \log(1 + e^{d(x_s)}) \\ = \left[\sum_{s=t}^T \beta^{s-t} P_1^{t,s} x_1 \right]' \theta_1 - \left[\sum_{s=t}^T \beta^{s-t} P_0^{t,s} x_0 \right]' \theta_0 \end{aligned} \quad (E.1)$$

where $d(x)$ denotes the vector of differences in the choice specific value functions with length equal to the number of discrete states, and $P_1^{t,s}$ and $P_0^{t,s}$ are defined as

$$\begin{aligned} P_1^{t,s} &= P_1^{t,t+1} P_1^{t+1,t+2} \dots P_1^{s-1,s} \\ P_0^{t,s} &= P_0^{t,t+1} P_0^{t+1,t+2} \dots P_0^{s-1,s} \end{aligned}$$

with

$$\begin{aligned} P_1^{t,t} &= I \\ P_0^{t,t} &= I \end{aligned}$$

More precisely, we can write

$$\begin{aligned} \sum_{s=t}^T \beta^{s-t} P_1^{t,s} &= I + \beta P^{t,t+1} + \beta^2 P^{t,t+2} + \dots + \beta^{T-t} P^{t,T} \\ &= I + \beta P^{t,t+1} + \beta P^{t,t+1} \beta P^{t+1,t+2} + \dots + \beta P^{t,t+1} \times \dots \times \beta P^{T-1,T} \end{aligned}$$

It is worth noting that when death is allowed and the states in P_1 and P_0 do not include death, P_1 and P_0 only have to be “sub” stochastic matrices, in the sense that their rows sum to < 1 instead of 1.

If we define

$$Q_t = \sum_{s=t}^T \beta^{s-t} P_1^{t,s}$$

then we can compute Q_t through recursive relations.

First of all,

$$Q_T = I$$

Secondly, for all $t < T$

$$Q_t = \beta P^{t,t+1} Q_{t+1} + I$$

It is also worth noting that there is a certain relation between the matrices on the left hand side and the right hand side of Eq. (E.1). In particular, the left hand side matrices are $\beta P_1^{t,t+1}$ and $\beta P_0^{t,t+1}$ multiplied by the left hand side matrices shifted forward by one period

$$\sum_{s=t+1}^T \beta^{s-t} P_1^{t,s} = \beta P_1^{t,t+1} \left(\sum_{s=(t+1)}^T \beta^{s-(t+1)} P_1^{(t+1),s} \right),$$

and

$$\sum_{s=t+1}^T \beta^{s-t} P_0^{t,s} = \beta P_0^{t,t+1} \left(\sum_{s=(t+1)}^T \beta^{s-(t+1)} P_0^{(t+1),s} \right).$$

F Kernel Smoothing

For estimation purposes, I discretize the continuous variable – the logarithm of household income – into a finite number of discrete values. Furthermore, due to the large number of discrete states and the relatively sparse data for each value of the discrete state variables, I kernel smooth over the discretized income levels when I nonparametrically estimate the conditional transition probabilities of the state variables. Kernel weights are calculated as

follows

$$w(b_2, l_2 | b_1, l_1, s_1, m) = \frac{p(b_2, l_2 | b_1, l_1, s_1, m)}{\sum_{i=1}^n p(b_{2i}, l_{2i} | b_{1i}, l_{1i}, s_{1i}, m_i)}$$

where

$$p(b_2, l_2 | b_1, l_1, s_1, m) = \frac{\frac{1}{nh^2} \sum_{i=1}^n k\left(\frac{l_{2i}-l_2}{h}\right) k\left(\frac{l_{1i}-l_1}{h}\right) I(b_{2i} = b_2) I(b_{1i} = b_1) I(s_{1i} = s_1) I(m_i = m)}{\frac{1}{nh} \sum_{i=1}^n k\left(\frac{l_{1i}-l_1}{h}\right) I(b_{1i} = b_1) I(s_{1i} = s_1) I(m_i = m)}$$

is the kernel estimate of the conditional joint density of the log(household income) and health status in period 2 conditional on log(household income), health status, and smoking status in period 1 as well as the indicator of same-gender parent's longevity. b denotes bad health, l denotes log(household income), s denotes smoking status, m denotes the indicator of whether the same-gender parent is still alive or died at an age greater than 70. $k(\cdot)$ is the kernel function and h is the bandwidth used in the nonparametric probability estimation and adapted from "Silverman's Rule of Thumb" (Silverman (1986)).³⁷

³⁷See Hardle (1990) for a discussion of bandwidth choice.

G Additional Tables and Figures

Table 7: Comparing Final Sample and the Sample with Missing Subjective Probabilities.

| Variable | Final Sample | | Missing Subj. Prob. | |
|--|--------------|-----------|---------------------|-----------|
| | Mean | Std. Dev. | Mean | Std. Dev. |
| Age | 56.7* | 2.98 | 56.7 | 2.97 |
| Female | 0.53** | 0.50 | 0.51 | 0.50 |
| White | 0.82** | 0.38 | 0.81 | 0.39 |
| Current smoker at period 1 | 0.38 | 0.49 | 0.38 | 0.49 |
| Former smoker at period 1 | 0.62 | 0.49 | 0.62 | 0.49 |
| Same gender parent alive or died at age > 70 | 0.67 | 0.47 | 0.67 | 0.47 |
| Self-rated health at period 1 | 2.65** | 1.16 | 2.69 | 1.17 |
| Bad health at period 1 | 0.23** | 0.42 | 0.24 | 0.43 |
| Household income at period 1 ('\$k) | 53.4 | 65.6 | 53.0 | 78.3 |
| Self-rated health at period 2 | 2.63* | 1.25 | 2.65 | 1.27 |
| Bad health at period 2 | 0.24** | 0.43 | 0.29 | 0.50 |
| Household income at period 2 ('\$k) | 51.6 | 68.4 | 50.8 | 68.2 |
| Observed deaths in two years | 0.021 | 0.14 | 0.023 | 0.15 |

Differences in means between the final analysis sample and the sample with missing subjective probabilities.
 *Statistically significant at 5% level; **statistically significant at 1% level.

Table 8: Special Cases of Subjective Probabilities of living to 75 and 85

| Cases | % |
|--|-------|
| Live to 75 = 0 | 0.064 |
| Live to 85 = 0 | 0.170 |
| Live to 75 = 1 | 0.200 |
| Live to 85 = 1 | 0.087 |
| P75 > P85 | 0.019 |
| Both probabilities = 0 | 0.064 |
| Both probabilities = 1 | 0.085 |
| P75 >= P85 | 0.290 |
| Any of the special case, with P75 > P85 | 0.380 |
| Any of the special case, with P75 >= P85 | 0.500 |

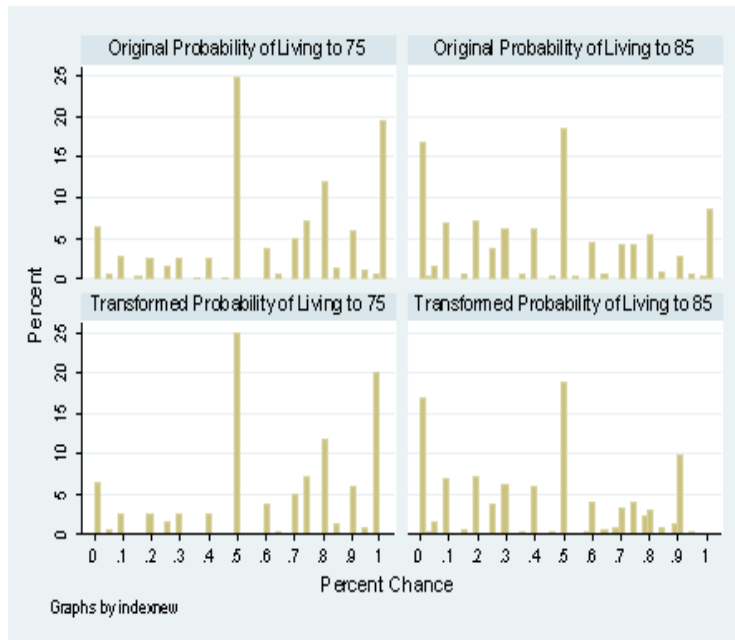
Frequencies of special elicited subjective expectations in the data.

Table 9: Discount Factors in the Literature

| Sources | Estimated Discount Factor |
|-------------------------------------|---------------------------|
| Moore and Viscusi (1990) | 0.88 to 0.99 |
| Moore and Viscusi (1988) | 0.89 to 0.91 |
| Arcidiacono, Sieg, and Sloan (2007) | 0.91 |
| Dreyfus and Viscusi (1995) | 0.85, 0.88, 0.90 |
| Hausman(1979) | 0.83 |
| Warner and Pleeter (2001) | 0.77 to 1 |

Examples of discount factors estimated in the literature.

Figure 11: Original and Transformed Subjective Probability of Living to Ages 75 and 85



Distributions of original elicited subjective responses and transformed answers.

References

- ACKERBERG, D., L. BENKARD, S. BERRY, AND A. PAKES (2005): "Econometric Tools for Analyzing Market Outcomes," in *Handbook of Econometrics*, ed. by J. J. Heckman, and E. E. Leamer, vol. 6, Amsterdam. North Holland.
- ARCIDIACONO, P., AND R. MILLER (2008): "CCP Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity," Revise and Resubmit.
- ARCIDIACONO, P., H. SIEG, AND F. A. SLOAN (2007): "Living Rationally Under the Volcano? an Empirical Analysis of Heavy Drinking and Smoking," *International Economic Review*, 48(1), 37 – 65.
- BECKER, G., M. GROSSMAN, AND K. MURPHY (1991): "Rational Addiction and the Effect of Price on Consumption," *American Economic Review*, 71(3), 933 – 946.
- (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review*, 84(3), 396 – 418.
- BECKER, G., AND C. B. MULLIGAN (1997): "The Endogenous Determination of Time Preference," *The Quarterly Journal of Economics*, pp. 729 – 758.
- BECKER, G., AND K. MURPHY (1988): "A Theory of Rational Addiction," *Journal of Political Economy*, 96(4), 675 – 701.
- BERESTEANU, A., AND F. MOLINARI (2008): "Asymptotic Properties for a Class of Partially Identified Models," *Econometrica*, 76(4), 763 – 814.
- BERNHEIM, B. D. (1989): "The Timing of Retirement: A Comparison of Expectations and Realizations," in *The Economics of Aging*, ed. by D. Wise, pp. 335 – 355. Chicago: The University of Chicago Press.
- (1990): "How do the Elderly Form Expectations: An Analysis of Responses to New Information," in *Issues in the Economics of Aging*, ed. by D. Wise, pp. 259 – 283. Chicago: The University of Chicago Press.
- BORSCH-SUPAN, A. (1998): "Comment on 'Subjective Survival Curves and Life Cycle Behavior'," in *Inquiries in the Economics of Aging*, ed. by D. A. Wise, pp. 306 – 309. Chicago: University of Chicago Press.
- BOUND, J. (1991): "Self-reported and Objective Measures of Health in Retirement Models," *Journal of Human Resources*, 24, 106–138.
- BRUINE DE BRUIN, W., B. FISCHHOFF, B. HALPERN-FELSHER, AND S. MILLSTEIN (2000): "Expressing Epistemic Uncertainty: It's a Fifty-Fifty Chance," *Organizational Behavior and Human Decision Processes*, 81, 115 – 131.
- CHALOUKKA, F., J. A. TAURAS, AND M. GROSSMAN (2000): *Tobacco Control in Developing Countries*chap. The Economics of Addiction, pp. 107 – 1290. Oxford: Oxford University Press.
- CHALOUKKA, F., AND K. WARNER (2000): *Handbook of Health Economics*chap. The Economics of Smoking, pp. 1539 – 1627. New York: North Holland.
- CHEN, X., H. HONG, AND D. NEKIPELOV (2008): "Nonlinear Models of Measurement Errors," Under Review.
- CUNHA, F., J. J. HECKMAN, AND S. NAVARRO (2005): "Separating Uncertainty from Heterogeneity in Life Cycle Earnings," *Oxford Economic Papers*, 57(2), 191 – 261, Oxford University Press.
- DELANVANDE, A. (2008a): "Measuring Revisions to Subjective Expectations," *Journal of Risk and Uncertainty*, 36(1).
- (2008b): "Pill, Patch or Shot? Subjective Expectations and Birth Control Choice," *International Economic Review*, 49(3).
- DER KLAUW, W. V., AND K. WOLPIN (2002): "Social Security, Pensions, and the Savings and Retirement Behavior of Households," Department of Economics, University of North Carolina at Chapel Hill.
- DOMINITZ, J. (1998): "Earnings Expectations, Revisions, and Realizations," *Review of Economics and Statistics*, 80, 374 – 388.
- DOMINITZ, J., AND A. HUNG (2003): "Homogeneous Actions and Heterogeneous Beliefs: Experimental Evidence on the Formation of Information Cascades," working paper, Heinz School of Policy and Management, Carnegie Mellon University.
- DOMINITZ, J., AND C. MANSKI (1997): "Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations," *Public Opinion Quarterly*, 61, 261 – 287.
- ECKSTEIN, Z., AND K. I. WOLPIN (1989): "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey," *The Journal of Human Resources*, 24, 562 – 598.

- EHRlich, I. (2000): "Uncertain Lifetime, Life Protection, and the Value of Life Saving," *Journal of Health Economics*, 19, 341 – 367.
- EHRlich, I., AND Y. YIN (2005): "Explaining Diversities in Age-Specific Life Expectations and Values of Life Saving: A Numerical Analysis," *Journal of Risk and Uncertainty*, 31(2), 129 – 162.
- FANG, H., AND D. SILVERMAN (2007): "Time-inconsistency and Welfare Program Participation: Evidence from the NLSY," *International Economic Review*, forthcoming.
- FANG, H., AND Y. WANG (2008): "Estimating Dynamic Discrete Choice Models with Hyperbolic Discounting, with an Application to Mammography Decisions," Working Paper.
- FISCHHOFF, B., AND W. BRUINE DE BRUIN (1999): "Fifty-Fifty = 50?," *Journal of Behavioral Decision Making*, 12, 149 – 163.
- FISCHHOFF, B., A. M. PARKER, W. BRUINE DE BRUIN, J. DOWNS, C. PALMGREN, R. DAWES, AND C. F. MANSKI (2000): "Teen Expectations for Significant Life Events," *Public Opinion Quarterly*, 64, 189 – 205.
- GAN, L., G. GONG, M. HURD, AND D. MCFADDEN (2004): "Subjective Mortality Risk and Bequests," working paper.
- GAN, L., M. HURD, AND D. MCFADDEN (2003): "Individual Subjective Survival Curves," NBER Working Paper 9480.
- GILLESKIE, D. (1998): "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica*, 66, 1 – 46.
- GROSSMAN, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80(2), 223 – 255.
- GROSSMAN, M., F. CHALOUKKA, AND R. ANDERSON (1998): "A Survey of Economic Models of Addictive Behavior," *Journal of Drug Issues*, 28, 631 – 643.
- GRUBER, J., AND B. KOSZEGI (2001): "Is Addiction "Rational"? Theory and Evidence," *Quarterly Journal of Economics*, 116(4), 935 – 958.
- HAMERMESH, D. (1985): "Expectations, Life Expectancy, and Economic Behavior," *Quarterly Journal of Economics*, 95, 537 – 557.
- HARDLE, W. (1990): *Applied Nonparametric Regression*. Cambridge University Press.
- HARRIS, C., AND D. LAIBSON (2001): "Dynamic Choice of Hyperbolic Consumers," *Econometrica*, 69(4), 935 – 958.
- HECKMAN, J. J., AND E. E. LEAMER (2007): *Handbook of Econometrics*, vol. 6B. Elsevier.
- HECKMAN, J. J., AND S. NAVARRO (2006): "Dynamic Discrete Choice and Dynamic Treatment Effects," *Journal of Econometrics*, 136, 341 – 396.
- HECKMAN, J. J., AND B. SINGER (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for 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 R. MILLER (1993): "Conditional Choice Probabilities and Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497 – 529.
- HU, Y., AND M. SHUM (2008): "Nonparametric Identification of Dynamic Models with Unobserved State Variables," Working Paper.
- HURD, M. D., D. MCFADDEN, AND L. GAN (1998): *Inquiries in the Economics of Aging*chap. Subjective Survival Curves and Life Cycle Behavior, pp. 353 – 387. University of Chicago Press.
- HURD, M. D., D. MCFADDEN, AND A. MERRILL (2001): *Themes in the Economics of Aging*chap. Predictors of Mortality Among the Elderly, pp. 171 – 197. University of Chicago Press.
- HURD, M. D., AND K. MCGARRY (1995): "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study," *Journal of Human Resources*, 30, S268 – S292.
- (2002): "The Predictive Validity of Subjective Probabilities of Survival," *The Economic Journal*, 112, 996 – 985.
- HURD, M. D., J. P. SMITH, AND J. M. ZISSIMOPOULOS (2004): "The Effects of Subjective Survival on Retirement and Social Security Claiming," *Journal of Applied Econometrics*, 19, 761 – 775.

- JUSTER, F. T., AND R. SUZMAN (1995): "An Overview of the Health and Retirement Study," *Journal of Human Resources*, 40, 934 – 940.
- KASAHARA, H., AND K. SHIMOTSU (2008): "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete choices," *Econometrica forthcoming*.
- KEANE, M. P., AND K. I. WOLPIN (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473 – 522.
- KENKEL, D. S. (2000): "Prevention," in *Handbook of Health Economics*, ed. by A. Culyer, and J. Newhouse, vol. 1, pp. 1675 – 1720. North Holland.
- KHWAJA, A., F. A. SLOAN, AND S. CHUNG (2007): "The Relationship Between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions," *Journal of Risk and Uncertainty*, 35(2), 179 – 201.
- KHWAJA, A., F. A. SLOAN, AND M. SALM (2006): "Evidence on Preferences and Subjective Beliefs of Risk Takers: The Case of Smokers," *International Journal of Industrial Organization*, 24, 667 – 682.
- KHWAJA, A., F. A. SLOAN, AND Y. WANG (2008): "Do Smokers Value Their Health and Longevity Less?," forthcoming, *Journal of Law and Economics*.
- KOTLARSKI, I. (1967): "On Characterizing the Gamma and Normal Distribution," *Pacific Journal of Mathematics*, 20, 729 – 738.
- KRASNOKUTSKAYA, E. (2003): "Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity," Revise and resubmit in *Review of Economic Studies*.
- LAIBSON, D. (1997): "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112(2), 443 – 477.
- LI, T. (2002): "Robust and Consistent Estimation of Nonlinear Errors-in-Variables Models," *Journal of Econometrics*, 110, 1 – 26.
- LI, T., AND Q. VUONG (1998): "Nonparametric Estimation of the Measurement Error Model Using Multiple Indicators," *Journal of Multivariate Analysis*, 65, 139 – 165.
- LIANG, L., F. CHALOUPEK, M. NICHTER, AND R. CLAYTON (2003): "Prices, Policies and Youth Smoking, May 2001," *Addiction*, 98, 105 – 122.
- LOCHNER, L. (2007): "Individual Perceptions of the Criminal Justice System," *American Economic Review*, 97, 444 – 460.
- MAGNAC, T., AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 20(2), 801 – 816.
- MAHAJAN, A. (2006): "Identification and Estimation of Regression Models with Misclassification," *Econometrica*, 74(3), 631 – 665.
- MANSKI, C. F. (1988): "Identification of Binary Response Models," *Journal of the American Statistical Association*, 83(403), 729 – 728.
- (1993): "Dynamic Choice in Social Settings," *Journal of Econometrics*, 58, 121 – 136.
- (2004): "Measuring Expectation," *Econometrica*, 72(5), 1329 – 1376.
- MATZKIN, R. L. (1992): "Nonparametric and Distribution-Free Estimation of the Binary Threshold Crossing and the Binary Choice Models," *Econometrica*, 60, 239 – 270.
- (1993): "Nonparametric Identification and Estimation of Polychotomous Choice Models," *Journal of Econometrics*, 58, 137 – 168.
- (1994): "Restrictions of Economic Theory in Nonparametric Methods," in *Handbook of Econometrics*, ed. by R. Engle, and D. McFadden, pp. 2523 – 2558. North-Holland, New York.
- (2003): "Nonparametric Estimation of Nonadditive Random Functions," *Econometrica*, 71(5), 1339 – 1375.
- MENDEZ, D., AND K. E. WARNER (2000): "Smoking Prevalence in 2010: Why the Healthy People Goal is Unattainable," *American Journal of Public Health*, 90, 401 – 403.
- (2004): "Adult Cigarette Smoking Prevalence: Declining as Expected (Not as Desired)," *American Journal of Public Health*, 94, 251 – 252.
- MOLINARI, F., AND C. MANSKI (2008): "Rounding Probabilistic Expectations in Surveys," Working Paper.
- MUTH, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315 – 335.

- NYARKO, Y., AND A. SCHOTTER (2002): "An Experimental Study of Belief Learning Using Elicited Beliefs," *Econometrica*, 70, 971 – 1005.
- PICONE, G., F. A. SLOAN, AND D. TAYLOR (2004): "Effects of Risk and Time Preference and Expected Longevity on Demand for Medical Tests," 28, 39 – 53.
- RUST, J. (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics*, ed. by R. Engle, and D. McFadden, vol. 4, pp. 3082 – 3139. North Holland.
- RUST, J., AND C. PHELAN (1997): "How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets," *Econometrica*, 65(4), 781 – 831.
- SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- SLOAN, F. A., J. OSTERMANN, G. PICONE, C. CONOVER, AND D. TAYLOR (2004): *The Price of Smoking*. MIT Press.
- SLOAN, F. A., AND Y. WANG (2008): "Economic Theory and Evidence on Smoking Behavior of Adults," *Addiction*, 103, 1777 – 1785.
- SMITH, V. K., D. H. TAYLOR, AND F. A. SLOAN (2001): "Longevity Expectations and Death: Can People Predict Their Own Demise?," *American Economic Review*, 91(4), 1126 – 1134.
- SMITH, V. K., D. H. TAYLOR, F. A. SLOAN, F. R. JOHNSON, AND W. H. DESVOUSGES (2001): "Do Smokers Respond to Health Shocks," *Review of Economics and Statistics*, 83(4), 675 – 687.
- TAYLOR, D. H., V. HASSELBLAD, S. J. HENLEY, M. J. THUN, AND F. A. SLOAN (2002): "Benefits of Smoking Cessation for Longevity," *American Journal of Public Health*, 92(6), 990 – 996.
- TOWNSEND, J., P. RODERICK, AND J. COOPER (1994): "Cigarette Smoking by Socioeconomic Group, Sex, and Age: Effects of Price, Income, and Health Publicity," *British Medical Journal*, 309, 923 – 927.
- TVERSKY, A., AND D. KAHNEMAN (1974): "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185(4157), 1124 – 1131.
- VAN DEN STEEN, E. (2004): "Rational Overoptimism (and Other Biases)," *The American Economic Review*, 94(4), 1141 – 1151.
- VISCUSI, W., AND W. EVANS (1990): "Utility Functions that Depend on Health Status: Estimation and Economic Implications," *American Economic Review*, 80, 353 – 375.
- VISCUSI, W. K. (1985): "A Bayesian Perspective on Biases in Risk Perception," *Economics Letters*, 17, 59 – 62.
- (1990): "Do Smokers Underestimate Risks?," *Journal of Political Economy*, 98(6), 1253 – 1269.
- (1991): "Age Variations in Risk Perceptions and Smoking Decisions," *The Review of Economics and Statistics*, 73(4), 577 – 588.
- VISCUSI, W. K., AND W. A. MAGAT (1987): *Learning About Risk: Consumer and Worker Responses to Hazard Information*. Cambridge, MA: Harvard University Press.
- VISCUSI, W. K., AND C. J. O'CONNOR (1984): "Adaptive Responses to chemical Labeling: Are Workers Bayesian Decision Makers?," *American Economic Review*, 74, 942 – 956.
- WASSERMAN, J., W. G. MANNING, J. P. NEWHOUSE, AND J. D. WINKLER (1991): "The Effects of Excise Taxes and Regulations on Cigarette Smoking," *Journal of Health Economics*, 10(1), 43 – 64.