BLS WORKING PAPERS

**Lighting the Fires: Explaining Youth Smoking Initiation
and Experimentation in the Context of a Rational
Addiction Model with Learning**

Brett Matsumoto, U.S. Bureau of Labor Statistics

Lighting the Fires: Explaining Youth Smoking Initiation and Experimentation in the Context of a Rational Addiction Model with Learning

Brett Matsumoto[†‡]

Bureau of Labor Statistics

Sept 2016

Abstract

This paper examines the dynamics of youth smoking behavior using a model of rational addiction with learning. Individuals in the model face uncertainty regarding the parameters that determine their utility from smoking. Through experimentation, individuals learn about how much they enjoy smoking cigarettes as well as the effects of reinforcement, tolerance, and withdrawal. The addition of learning to the dynamic optimization problem of adolescents provides an explanation for the experimentation of the non-smoker. I estimate the parameters of the model using data from the National Longitudinal Survey of Youth 1997 and compare the overall fit of the model to the model without learning. The estimated model is also used to analyze the effect of cigarette taxes and anti-smoking policies. I find that the model with learning is better able to fit the observed data and that cigarette taxes are not only effective in reducing the level of youth smoking, but can even increase welfare for some individuals.

JEL Classification: I12, D83, D12, C61, C63

Keywords: Dynamic Discrete Choice, Uncertainty and Learning, Youth Smoking, CCP Estimation

1 Introduction

Despite its historically low level in the U.S., cigarette smoking remains a major public health concern. The Surgeon General estimates that tobacco use causes approximately 480,000 deaths per year in the United States and is estimated to cause between \$289-332.5 billion in economic costs (USDHHS, 2013).[1] Tobacco use is the leading preventable cause of death, yet people continue to smoke despite the high level of public awareness of its adverse health effects. Because cigarettes are addictive, it may be easier to discourage smoking initiation than to encourage smoking cessation. Also, cigarette manufacturers have historically targeted their advertisements to young people in the hopes of cultivating lifelong customers. Among adults who become daily smokers, approximately 90 percent smoke for the first time before age 18 (USDHHS, 2012). For these reasons, policy interventions aimed at reducing the level of smoking in the population often target young people.

The decision to engage in a harmful addictive behavior, such as smoking, seemingly presents a problem for standard economic models. For a forward-looking utility-maximizing agent, consuming a harmful addictive substance would be an irrational act. The Rational Addiction (RA) model of Becker and Murphy (1988) shows that consumption of an addictive substance can be explained using the standard economic framework. Their explanation of addictive behavior centers around the concept that past utilization of addictive goods impacts current utility from consumption of these goods. A major criticism of the Becker and Murphy model is the implication that individuals are always acting optimally, so addicts do not regret their decision to consume the addictive good. In their model addiction is not a problem or even an undesirable outcome, so there is no place for policy intervention to treat or prevent addiction. Empirical evidence suggests that many individuals regret their decision to smoke. Approximately 70% of adult smokers wish to quit smoking entirely and over half have attempted to quit smoking in the past year (NHIS, 2010).

Another limitation of the RA model as a model of youth smoking behavior is that it treats smoking initiation as exogenous. In this paper, I extend the RA model so that it is better able to explain the individual's smoking initiation decision. Specifically, I relax the assumption of perfect information in the RA model by incorporating learning about one's preferences. The parameters that determine the utility one receives from smoking are initially unknown, but the individual has beliefs about their true value. As an individual experiments with smoking, he receives utility signals and updates his beliefs. The addition of uncertainty and learning to the optimization problem of adolescents provides an explanation

---

[1] Economic costs include direct medical costs in addition to the lost productivity attributable to smoking related illnesses. Estimates are for the years 2009-2012.

for the experimentation of the non-smoker and allows for the possibility that an individual who starts to smoke may later regret that decision. Therefore, policies that prevent an individual from experimenting with cigarettes may be welfare improving as the individual would be prevented from making a decision he may later regret.

The main purpose of this paper is to quantify the effectiveness of anti-smoking policies and to evaluate the resulting impact on individual welfare. In order to do this, I recover the policy-invariant utility function parameters of a rational addiction model with learning by fitting a dynamic discrete choice model of optimal smoking decision making to the observed data. As the first attempt to estimate the structural parameters of a rational addiction model with learning about preferences, this research allows for empirical testing of the perfect information assumption in the RA model (i.e., the assumption that individuals know their utility function parameters). I estimate the model parameters using the National Longitudinal Survey of Youth 1997 (NLSY97).

Estimation of the parameters of a dynamic discrete choice model is generally computationally intensive as each iteration over the parameter space requires re-solving the dynamic optimization problem. The inclusion of uncertainty and learning over multiple parameters further complicates estimation of the model. To circumvent these computational issues, I use the Expectation Maximization (EM) algorithm in conjunction with Conditional Choice Probability (CCP) estimation and Monte Carlo simulation to estimate the model parameters. The estimation procedure provides a significant computational advantage, which allows for the estimation of a more complex model than is feasible using full-solution techniques.

Preliminary estimation results demonstrate that allowing for uncertainty and learning in a dynamic model of youth smoking significantly improves the overall fit of the model. Results from counterfactual policy simulations suggest that policies that impact individuals' initial beliefs about their utility function parameters are effective in reducing youth smoking. Taxes are also shown to be effective in reducing the level of smoking. The estimated model predicts that a doubling of the price of cigarettes would reduce the prevalence of youth smoking by 41.6% and adult smoking by 35.1%. An increase in the legal purchasing age from 18 to 19 years old would decrease youth smoking by 21.6%. However, there would be no effect on adult smokers as the higher legal purchasing age would only cause a delay in smoking initiation. The results of the welfare analysis show that increasing cigarette taxes would only lead to a relatively small loss in total welfare as the welfare gains to keeping those who would later regret the decision to smoke from starting to smoke offset the loss of welfare from smokers having to pay a higher price for cigarettes.

The remainder of the paper proceeds as follows: Section 2 reviews the related literature. Section 3 presents the model. Section 4 discusses the data. Section 5 develops the estimation

routine. The estimation results are presented in section 6, and section 7 concludes.

2   Related Literature

Becker and Murphy (1988) developed the RA model to show that seemingly irrational behavior could be explained using a standard economic framework of a forward-looking utility-maximizing agent. The model's welfare implications have caused many to abandon the general framework of the RA model and to develop "irrational" models to explain the time inconsistency of addictive behavior. These alternative theoretical models generally feature dual-states of the world or individuals with dual-selves.[2] Addiction results when an individual is in an addictive state of the world or if the behavior of the individual is being controlled by the self that is more prone to addiction.

Other models of the consumption of addictive goods generate time-inconsistent behavior by deviating from the standard assumptions regarding how future utility is discounted. The simplest deviation is the myopic model. A myopic individual completely discounts future utility and only considers the current period's utility when making decisions. Other deviations from the standard assumptions regarding time preferences include an endogenous discount factor (Orphanides and Zervos, 1998) or hyperbolic discounting (Gruber and Koszegi, 2001). Finally, Orphanides and Zervos (1995) argue that the problem with the RA model is not the assumption of a rational, forward-looking agent but the assumption of perfect information. An individual in their model can be one of two types (addict or not an addict). The individual learns which type he is if he consumes the addictive good. The model estimated in this paper is an extension of the theoretical model proposed by Orphanides and Zervos (1995).

The RA model assumes that individuals are forward-looking, and there have been many studies that attempt to test the validity of this assumption empirically in the context of consumer demand for an addictive good. The evidence is generally consistent with forward looking behavior (Becker et al., 1994; Chaloupka, 1991).[3] One of the limitations of the empirical addiction literature is that papers primarily attempt to compare the rational addiction model to the myopic model. No work (of which the author is aware) has been done to estimate alternative models or to empirically test the other assumptions of the RA model. Most of the literature involves reduced form estimation, but a few papers have estimated the structural parameters of an addiction model (Arcidiacono et al., 2007; Choo, 2000; Gordon and Sun, 2009; Darden, 2011).[4]

---

[2]Papers that use the dual-state approach include Winston (1980) and Bernheim and Rangel (2004). Papers that use the dual-self approach include Thaler and Shefrin (1981) and Benabou and Tirole (2004).

[3]See Chaloupka and Warner (2000) for a thorough summary of the empirical literature.

[4]There is a learning component to the life-cycle model of Darden (2011), but the learning is over the health

Much of the analysis in the economics literature of policy interventions on youth smoking has focused on cigarette taxes. The rational addiction framework implies that individuals who are not currently consuming the addictive good should be more responsive to changes in the price of that good than current users. Many studies have found a significant effect of taxes on smoking initiation. Some studies, however, have found that cigarette taxes have little to no significant effect on youth smoking initiation (DeCicca et al., 2002, 2008; Emery et al., 2001). Importantly, some of the studies in this literature find that nonsmokers are more price sensitive than smokers while also controlling for unobserved heterogeneity (Fletcher et al., 2009; Gilleskie and Strumpf, 2005). Finally, some studies have found that taxes merely delay smoking initiation rather than prevent people from becoming smokers (Glied, 2002). There have been fewer papers that examine the effect of other anti-smoking policies on youth smoking and the results have been mixed (Tworek et al., 2010).[5]

One of the main applications of learning models in economics is in the area of consumer learning from experience goods (Erdem and Keane, 1996; Ackerberg, 2003).[6] These models estimate the learning process involved when consumers purchase unfamiliar goods. The consumer learns about the utility he receives from consuming these goods and updates his beliefs each time the good is consumed. This paper fits into the structural learning literature because the utility that the individual receives from consuming an addictive good is initially unknown and is learned over time if the individual consumes the addictive good. This paper extends the standard models used by incorporating the unique features of consuming an addictive good.

3   Model

This section sets up the individual's decision problem regarding optimal smoking behavior. An individual receives utility from consuming cigarettes as well as the consumption of other goods. In order to incorporate the features of consuming an addictive good, the individual's utility in the current period also depends on past levels of smoking in a manner consistent with the scientific literature on addiction (Laviolette and van der Kooy, 2004; Nestler and Aghajanain, 1997). Past consumption of the addictive good affects current utility through *reinforcement*, which occurs when the marginal utility of smoking is increasing in the level of past smoking. As the body becomes accustomed to consuming an addictive substance, larger quantities of the substance must be consumed to achieve a similar effect. This physical

---

effects of smoking. Individuals are assumed to know their preferences (i.e., utility function parameters).

    [5]For an overview of the effectiveness of anti-smoking legislation in general, see Goel and Nelson (2006).

    [6]See Ching et al. (2011) for an overview of the empirical economic applications of learning models.

transition is referred to as developing *tolerance*. Habitual use of an addictive good also generates physical dependence. As a result, the individual experiences adverse effects from attempting to lower the level of consumption of the addictive good. This transition may result in a *withdrawal* effect. Withdrawal is modeled as an asymmetric adjustment cost, i.e. a cost associated with decreasing the amount consumed.[7] These effects are parameterized in the model (ρ, τ, and ω for reinforcement, tolerance, and withdrawal respectively), the magnitude of these effects depends on the level of past smoking, and these parameters vary across individuals. For certain combinations of these individual specific parameter values, the combined effect of reinforcement, tolerance, and withdrawal generates *adjacent complementarity* in the consumption of cigarettes. Adjacent complementarity, which Becker and Murphy (1988) use as the defining characteristic of addiction, occurs when current consumption of a good is increasing in past consumption.

3.1 Utility

Each year, individual n makes an annual smoking decision and chooses his level of smoking from a discrete set of alternatives, $a_j \in \{a_1, a_2, \ldots, a_J\}$, which reflect the average daily cigarette consumption during the year. The decision not to smoke is represented by the level of smoking a_1. The price of a single cigarette in period t is denoted p_t. The addictive stock is denoted as $S_{n,t}$ and is defined as the level of smoking in the prior year.[8] The contemporaneous utility associated with alternative $j > 1$ for individual n at time t if the individual did not smoke in the previous period ($S_{n,t} = 0$) is:

$$u_{n,t}^j = \left(\alpha_n + \xi_j X_{n,t}\right) z(a_j) - \gamma_n p_t a_j + \epsilon_{n,t}^j \tag{1}$$

where ϵ is a vector of independent and identically distributed alternative-specific preference shocks that follow a Generalized Extreme Value (GEV) distribution. The utility from smoking depends upon the individual-specific match parameter α_n, demographic variables ($X_{n,t}$), the level of smoking through the function $z(a)$ (explained below), and the expenditure on smoking (which depends on both the price of cigarettes and level of smoking). The parameter γ_n measures the individual's sensitivity to the price of cigarettes and is a function of age, work status, and income.[9] Additionally, the individual's demographic variables affect

---

[7]This approach of explicitly modeling withdrawal effects as asymmetric adjustment costs to achieve adjacent complementarity in a rational addiction model was developed by Suranovic et al. (1999).

[8]This definition of the addictive stock implies full depreciation which is justified by the frequency of the smoking decision. Future versions of this paper will test whether the parameter estimates of the model change if this assumption is relaxed.

[9]The utility from consuming one's entire income in other goods is normalized to zero.

utility by imposing additional costs or benefits on different levels of smoking. If a variable only affects the utility of smoking versus not smoking and does not affect the decision of how much to smoke conditional on smoking, then the coefficient ξ_j will be constant for $j > 1$. Variables in this category include the individual's race or religion. These variables may affect the social acceptance of smoking within the individual's culture. Variables that potentially affect utility differently for different levels of smoking could include whether the individual is under 18 years of age or whether the individual has older siblings. These variables were shown to be significant in the smoking decision of young people in Gilleskie and Strumpf (2005).

If the individual has a positive level of smoking stock (i.e., $S_{n,t} > 0$) then the utility for alternative $j > 1$ is:

$$u_{n,t}^j = \left(\alpha_n + \rho_n g(S_{n,t}) + \xi_j X_{n,t}\right) z(a_j) - \tau_n S_{n,t} - \omega_n q(a_j, S_{n,t}) \mathbf{1}[a_j < S_{n,t}] - \gamma_n p_t a_j + \epsilon_{n,t}^j \quad (2)$$

The addictive stock affects the marginal utility of smoking through the reinforcement, tolerance, and withdrawal terms. The reinforcement effect, $\rho g(S)$, increases the marginal utility of smoking for every positive level of smoking. The tolerance effect, τS, enters current period utility for positive levels of past and current consumption and decreases the utility associated with each positive level of smoking. The adjustment cost or withdrawal cost, $\omega q(a, S)$, only enters the current period's utility when the individual reduces his consumption from one period to the next. The utility of not smoking ($j = 1$) is normalized to only include the withdrawal term (if $S_{n,t} > 0$) and the preference shock. The functions $z, g,$ and q have the following properties:

1. $z'(a) > 0$, $z''(a) < 0$, $\lim_{a \to 0} z'(a) < \infty$

2. $g(0) = 0$, $g'(S_{n,t}) > 0$, $g''(S_{n,t}) < 0$

3. $q(a_j, S_{n,t}) \geq 0$ for all $a_j \leq S_{n,t}$ and $q(a_j, S_{n,t}) = 0$ if $a_j = S_{n,t}$

The assumptions on the function z allow for a corner solution since the marginal utility from smoking is finite when the individual chooses not to smoke. The function q, which is a component of the withdrawal effect, is also assumed to be increasing in the size of the decrease in smoking from one period to the next. The functions g and q allow the reinforcement and withdrawal effects to be nonlinear.[10] The Estimation section discusses the specific functional forms used. The individual's smoking preference parameters are $\theta_n = (\alpha_n \ \rho_n \ \tau_n \ \omega_n)'$. The parameter α_n determines the individual's match quality for smoking. The parameters $\rho_n, \tau_n,$

---

[10]The tolerance term could also be allowed to be nonlinear in S.

and ω_n correspond to the effects of reinforcement, tolerance and withdrawal, respectively. The parameters in θ_n vary across individuals and are jointly normally distributed in the population: $\theta_n \sim N(\bar{\theta}, \Sigma)$.

3.2 Timing

The individual does not initially know the value of his smoking preference parameters (θ_n). He makes an annual smoking decision based on his beliefs about the parameters. At the start of the period, the individual observes prices, government tobacco policies, demographic variables (X), and the alternative specific preference shock. Then, the individual chooses a level of smoking and receives a utility signal. The individual uses this signal to update his beliefs at the end of the period.

An individual who has never smoked before the current period faces a sequential smoking decision within the period, where he first decides whether to experiment with smoking before making a smoking consumption decision for the year. The consumer learning literature generally finds that learning about match quality occurs relatively quickly. Since it would not take a full year to learn the match quality parameter α, an individual who has never smoked must first decide whether to experiment with smoking. Let a^E denote the level of consumption associated with experimentation. If he chooses to experiment, he learns his true value of α and proceeds to make a smoking decision for the rest of the period. If he chooses not to experiment, his smoking consumption for the period is zero and he will face the experimentation decision again in the next period. In periods after the individual experiments, the only decision is about annual smoking consumption. The utility of experimenting is:

$$u_{n,t}^E = (\alpha_n + \xi^E X_{n,t}) z(a^E) - \gamma p_t a^E + \epsilon_{n,t}^E \tag{3}$$

The utility shock for experimenting is assumed to be from a Type I Extreme Value distribution. For the sequential decision, individuals observe the preference shock for experimenting at the start of the period but do not observe the preference shock for the smoking decision until after they experiment.

3.3 Beliefs and Learning

3.3.1 Learning over the Utility Function Parameters

The individual's initial prior beliefs are denoted as $\theta_{n,0} \sim N(m_{n,0}, \Sigma_{n,0})$. Assuming Rational Expectations, the mean and variance of the individual's initial prior beliefs equal the

population mean and variance of θ.[11]

The individual updates his beliefs according to a Bayesian learning process based on the signals received. After experimenting, the individual learns his true value of α. Without loss of generality, assume that the individual first experiments with the addictive good in period 0. The initial prior for the period 0 consumption decision is the initial prior distribution conditional on the realized value of α. Let $m_{n,0|\alpha_n}$ and $\Sigma_{n,0|\alpha_n}$ denote the mean and covariance matrix of the initial prior distribution conditional on $\alpha = \alpha_n$. This conditional distribution becomes the initial prior distribution for the subsequent learning over the parameters ρ, τ, and ω.

In every period that an individual chooses to smoke, he receives utility signals about the value of the reinforcement and tolerance parameters. If the individual reduces his level of smoking in period t from the level in period $t-1$, he receives a signal for the withdrawal parameter. For the level of smoking a_j and past smoking $\{S_{n,l}\}_{l=0}^{t}$, the signals are as follows:

$$\delta_{n,t} = \begin{cases} (\rho_n + \lambda_{n,t})\mathbf{1}[a_j > 0] & \lambda_{n,t} \sim i.i.d.\ N(0, \frac{\sigma_\lambda^2}{a_j(1+g(S_{n,t}))}) \\ (\tau_n + \psi_{n,t})\mathbf{1}[a_j > 0] & \psi_{n,t} \sim i.i.d.\ N(0, \frac{\sigma_\psi^2}{1+S_{n,t}}) \\ (\omega_n + \eta_{n,t})\mathbf{1}[a_j < S_{n,t}] & \eta_{n,t} \sim i.i.d.\ N(0, \frac{\sigma_\eta^2}{S_{n,t}-a_j}) \end{cases} \tag{4}$$

The variation in the observed signal around its true value is assumed to be uncorrelated with the other parameters. The accuracy of the reinforcement signal is proportional to the quantity consumed as well as the level of past consumption, which implies that individuals face a trade-off between the speed of learning and the risk of becoming addicted. The accuracy of the tolerance signal is greater for higher levels of past consumption, and the accuracy of the withdrawal signal increases with larger decreases in consumption. The individual uses this utility signal to update his beliefs about his true parameters. I assume that the individual is able to distinguish between the signals if multiple signals are received in a given period and that the signal noises are uncorrelated (conditional on a_j and $S_{n,t}$).

The individual's posterior beliefs at the end of period t after choosing a level of smoking equal to a_j (i.e., the individual's beliefs after receiving the signals associated with the smoking decision) are:

$$\theta_{n,t+1|\alpha} \sim N(m_{n,t+1|\alpha}, \Sigma_{n,t+1|\alpha}) \tag{5}$$

where

$$m_{n,t+1|\alpha} = \Sigma_{n,t+1|\alpha}^{-1}(\Sigma_{n,t+1|\alpha}^{-1}m_{n,t|\alpha} + \Phi_{n,t}^{-1}\delta_{n,t}) \tag{6}$$

---

[11]Some restriction on the initial prior beliefs is required for identification. It may be possible to introduce heterogeneity in the initial priors by allowing the parameters of the initial prior beliefs to vary by observable characteristics.

$$\Sigma_{n,t+1|\alpha} = (\Sigma_{n,t|\alpha}^{-1} + \Phi_{n,t}^{-1}B_{n,t})^{-1} \tag{7}$$

$$\Phi_{n,t}^{-1} = \begin{pmatrix} \frac{a_j(1+g(S_{n,t}))}{\sigma_\lambda^2} & 0 & 0 \\ 0 & \frac{1+S_{n,t}}{\sigma_\psi^2} & 0 \\ 0 & 0 & \frac{S_{n,t}-a_j}{\sigma_\eta^2} \end{pmatrix} \tag{8}$$

$$B_{n,t} = \begin{pmatrix} \mathbf{1}[a_j > 0] & 0 & 0 \\ 0 & \mathbf{1}[a_j > 0] & 0 \\ 0 & 0 & \mathbf{1}[a_j < S_{n,t}] \end{pmatrix} \tag{9}$$

Equations (6) and (7) are the updating equations for the mean and variance of the individual's beliefs. The updated mean is a weighted average of the prior mean and the signal, where the weights are the precision (inverse of the variance) of the prior and the signal. Φ is a diagonal matrix of the signal precision, and B is a diagonal matrix with indicators for a given signal being received. As the individual receives more signals, the precision of his beliefs increases. Since the signals are unbiased, the individual's beliefs converge to the true parameter values.

Note that even though the signal noises are uncorrelated, the learning process for each parameter is not independent of the learning process for the other parameters. Since the parameters are correlated in the population and the population covariance matrix is the variance of the individual's initial prior beliefs, there is correlation in the learning process among the parameters. Even if the individual never receives a withdrawal signal, his beliefs about the value of his withdrawal parameter will change as he receives more information about the value of his other parameters.

3.3.2   Expectation of Future Prices, Policies, and State Variables

There are two components of the retail price of cigarettes: the manufacturer's price of the product and state and federal excise taxes. Determinants of the price of the product include the price of tobacco, production technology, labor costs, and other costs of production and distribution. Since surveyed individuals are not typically asked about their subjective expectations for future prices, some assumption must be made for how individuals forecast prices. One possible specification is to assume that the base component of the price follows a simple stochastic process (e.g., time trend with an AR(1) error). The justification for this specification is that individuals likely have some idea as to any time trend in the price as well as some realization that price shocks are persistent over time.

The other component of price, the excise tax, is much more difficult for the individual to forecast because it is determined by the political system. Specifying how individuals form

expectations over other future tobacco policies presents a similar challenge. Estimates of the model presented in this work will impose the likely unrealistic assumption of perfect foresight.[12]

The endogenous state variables include the individual's beliefs and the addictive stock. The addictive stock is defined as the prior period's level of smoking, so the addictive stock evolves deterministically conditional on a particular smoking choice. The individual uses his current beliefs about smoking preferences to evaluate the different smoking alternatives, while taking into account the potential information that he will receive from each possible choice. The individual also has perfect foresight regarding the observed exogenous state variables in X.[13]

3.4   The Individual's Problem

Each period, the individual chooses a level of smoking that maximizes his expected discounted lifetime utility given his beliefs and the value of the other state variables. The individual evaluates his expected discounted lifetime utility using backwards recursion. Let T denote the final period the individual is observed in the data, and let $d_{n,t}^j$ be an indicator variable that equals one if the individual selects alternative j in period t. Then the value function in period T is:

$$V_{n,T}(S_{n,T}, m_{n,T}, \Sigma_{n,T}, X_{n,T}) = \mathbf{E}\Big[\max_j \; d_{n,T}^j \Big(u^j(\theta_{n,T}, S_{n,T}, X_{n,T}) \tag{10}$$

$$+\beta \mathbf{E}[V_{n,T+1}(S_{n,T+1}, m_{n,T+1}, \Sigma_{n,T+1}, X_{n,T+1}, H_{n,T+1}) \mid S_{n,T}, m_{n,T}, \Sigma_{n,T}, X_{n,T}, d_{n,T}^j = 1])\Big)\Big]$$

The continuation value function V_{T+1} contains an additional state variable H that contains the individual's cumulative smoking history (i.e., total number of years smoked at each level of smoking).[14] The cumulative smoking history affects the individual's utility later in life through potential adverse health effects of smoking. The expectation over the discounted future value term is taken with respect to the future state variables. Current period utility is the expected utility given the current period's prior beliefs. Since the parameters in θ enter the utility function linearly, the expected utility for the current period is just the utility evaluated using the mean of the individual's current prior. The value function for earlier

---

[12]Other possibilities include assuming that the individual expects current tobacco taxes and policies to continue indefinitely or that individuals form expectations regarding the frequency and magnitude of excise tax changes based upon recent experience (i.e., a form of adaptive expectations).

[13]For some variables, such as the individual's age, this assumption is not unrealistic.

[14]The state variable H is suppressed in the value functions of earlier periods to simplify notation. Although the cumulative history does not affect utility in earlier periods, it still impacts the individual's behavior by changing the discounted expected future lifetime utility in period T.

periods can be defined recursively starting from the terminal period value function:

$$V_{n,t}(S_{n,t}, m_{n,t}, \Sigma_{n,t}, X_{n,t}) = \mathbf{E}\bigg[\max_j \; d^j_{n,t}\Big(u^j(\theta_{n,t}, S_{n,t}, X_{n,t})$$
$$+ \beta \mathbf{E}[V_{n,t+1}(S_{n,t+1}, m_{n,t+1}, \Sigma_{n,t+1}, X_{n,t+1}) \mid S_{n,t}, m_{n,t}, \Sigma_{n,t}, X_{n,t}, d^j_{n,t} = 1]\Big)\bigg] \quad (11)$$

If the individual has never smoked prior to period t, the value function for the experimentation decision is defined as:

$$V^E_{n,t}(m_{n,0}, \Sigma_{n,0}, X_{n,t}) =$$
$$\max \left\{ u^E_{n,t} + \mathbf{E}_\alpha[V_{n,t}(m_{n,0|\alpha}, \Sigma_{n,0|\alpha}, 0, X_{n,t})] \; , \; \beta \mathbf{E}[V^E_{n,t+1}(m_{n,0}, \Sigma_{n,0}, X_{n,t+1})] \right\} \quad (12)$$

The first term inside the max operator is the value from experimenting in the current period. This term includes the utility from experimenting plus the value of the consumption decision for the current period. The value of the consumption decision depends upon a particular realization of α, which is unknown at the time of the experimentation decision, so the expected value of the consumption decision is calculated by integrating over potential realizations of α. The second term inside the max operator is the value associated with not experimenting, which is the discounted expected future value of the next period's experimentation decision.

The individual's problem is to choose the optimal sequence of experimentation and consumption in order to maximize his discounted lifetime expected utility. In the first period, the individual's beliefs are the initial prior beliefs and the individual has no experience with smoking.

4 Data

The data used to estimate the structural parameters of the model are from the NLSY97. The first wave of the survey was conducted in 1997 and included 8,984 individuals who were born between 1980 and 1984 (age at first interview ranged from 12 to 18). Subsequent waves have been conducted annually and are ongoing. This paper uses the first 13 waves of the data (through the 2009 wave). There are several advantages of using this data set for the study of youth smoking initiation. First, the individuals in the data set are surveyed at a young age during which the decision to begin smoking is made. Second, the survey is conducted annually, which is generally the shortest interval between observations in large nationally-representative panel data sets. The learning process is better identified with

annual observations as opposed to less frequent observations.[15] Finally, the questions related to smoking are asked every wave. I supplement the geocoded restricted use version of the NLSY97 data set with tobacco policy data by matching individuals with the tobacco policies in their state. Relevant policies for this study include the cigarette excise tax, restrictions on tobacco advertising, spending on anti-smoking policies, and indoor smoking bans.

4.1 Sample Selection and Attrition

In a dynamic structural model, missing choice data add additional complexity in estimation. If an individual is in the sample, leaves, and later re-enters the sample, then the estimation routine has to integrate over all possible sequences of choices in the missing periods to calculate the value of the state variable when the individual re-enters the sample. One alternative is to only estimate the model on individuals who are observed in each time period. Restricting the sample to individuals observed in every time period avoids the difficulties in estimation, but the resulting sample may no longer be representative of the population if attrition is non-random. Table 1 reports the proportion of individuals with a given number of missing waves. Only about 60% of the original sample (5,385 of the original 8,984 individuals) is observed in every wave. Approximately 11% of this sample has one missing observation, and an additional 10% have either two or three missing observations. The preliminary estimation sample only includes the individuals who are observed in every wave. An additional 598 individuals are excluded due to missing smoking, demographic, or geographic data. The preliminary estimation sample contains the 4,787 individuals observed in every wave with nonmissing data for the key variables.

Table 1 : Individual Level Survey Participation

Total years missing	0	1	2	3	4	5	6	7+	Total
Frequency	5,385	1,011	582	378	330	254	219	825	8,984
Percent	59.94	11.25	6.48	4.21	3.67	2.83	2.44	9.18	100

---

[15]If the individuals are only observed infrequently, then it is likely that much of the uncertainty would be resolved after a relatively small number of observations. It would be difficult to identify the dynamic learning process if the econometrician only had a few observations per individual where uncertainty and learning mattered.

4.2 Data Summary and Construction of Key Variables

In the NLSY97, individuals are asked whether they have smoked since the previous interview. If the answer is yes, the individuals are asked about their smoking behavior over the month prior to the interview. Specifically, the question asks, "during the past 30 days, on how many days did you smoke a cigarette?" If the answer is greater than zero, the next question asks, "when you smoked a cigarette during the past 30 days, how many cigarettes did you usually smoke each day?" I construct a categorical smoking variable from the answers to these two questions. The total number of cigarettes smoked in the past month is simply the product of the answer to these two questions and is divided by 30 to give the average number of cigarettes smoked per day. The range of possible values for the average number of cigarettes smoked per day is divided into four intervals to create the discrete choice variable a_j. These intervals correspond to not smoking, light smoking (0-5 cigarettes per day), moderate smoking (5-15 cigarettes per day), and heavy smoking (more than 15 cigarettes per day).

Table 2 : Categorical Smoking Statistics

Smoking Level	Range (cigarettes per day)	Frequency (in person years)	Percent	$\mathbf{E}[a\vert a_j]$
None	$a_1 = 0$	$44,186$	69.65	0
Light	$0 < a_2 \leq 5$	$9,714$	16.50	1.63
Moderate	$5 < a_3 \leq 15$	$5,469$	9.21	10.50
Heavy	$15 < a_4$	$2,862$	4.64	23.51

Table 2 reports the range of each of the intervals, the number of observations (in person years) in each interval, and the mean of average cigarettes smoked per day conditional on being in the range of the interval. The distribution of the average cigarettes smoked per day is skewed to the right with the majority of the observations concentrated at the mass point of zero.

Table 3 reports the transition probabilities for the smoking categories. The transition probabilities illustrate several key features of the data. First, individuals increase their level of smoking gradually. Individuals are more likely to increase to the next highest level than they are to jump several levels. Also, for any given level of smoking, there is a high probability that individuals will transition to a lower level of smoking. For light and moderate levels of smoking, the probability that individuals decrease the amount they smoke is approximately 30%. For the heaviest smokers, this probability is almost 40%. The amount of decreases in the level of smoking observed in the data is difficult to reconcile with the standard RA model,

Table 3 : Cumulative Smoking Transition Probabilities

Smoking level at $t-1$	Smoking level at t			
	None	Light	Moderate	Heavy
None	0.900	0.081	0.014	0.005
	(36,748)	(3,321)	(562)	(193)
Light	0.297	0.537	0.141	0.026
	(2,679)	(4,848)	(1,269)	(233)
Moderate	0.097	0.175	0.575	0.154
	(482)	(871)	(2,862)	(767)
Heavy	0.065	0.059	0.249	0.627
	(169)	(155)	(650)	(1,635)

Note: frequencies in parentheses

Table 4 : Under 18 Smoking Transition Probabilities

Smoking level at $t-1$	Smoking level at t			
	None	Light	Moderate	Heavy
None	0.882	0.099	0.015	0.004
Light	0.360	0.448	0.151	0.040
Moderate	0.116	0.146	0.517	0.221
Heavy	0.076	0.093	0.271	0.559

but is consistent with the model of behavior that incorporates uncertainty and learning. Table 4 reports the transition probabilities for individuals under 18 years old. Relative to the full sample there is less persistence in smoking choices, with more movement (both upward and downward) between smoking categories. This is consistent with the learning model since it will take some experience before individuals are able to determine what level of smoking is optimal for their specific utility function parameters.

Table 5 presents summary statistics for smoking behavior and demographic variables in three of the early waves.[16] Over these waves, the proportion of individuals who currently smoke increases, however, it does fall in later waves. The other variables included in the

Table 5 : Summary Statistics of Smoking and Demographic Variables in Select Years

| | Year | | | | | |
| | 1997 | | 1999 | | 2001 | |
Time-Varying Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.363	0.481	0.517	0.500	0.598	0.490
Current smoker	0.156	0.363	0.265	0.441	0.315	0.464
Number of cigarettes per day	0.541	2.559	1.618	5.064	2.368	6.135
Age	14.23	1.474	16.82	1.432	18.88	1.430
Employed	0.447	0.497	0.530	0.499	0.708	0.455
Real annual income*	247.5	762.6	1,208	3,178	3,889	6,295
Income ¿ $20,000	0.000	0.014	0.010	0.098	0.028	0.164
Married	0.000	0.014	0.013	0.115	0.053	0.224
Any children in household	0.007	0.085	0.047	0.212	0.107	0.310
High School student	0.982	0.133	0.691	0.462	0.292	0.455
College student	0.000	0.020	0.128	0.334	0.303	0.460
High School graduate	0.001	0.035	0.240	0.427	0.605	0.489
Time-Invariant Variables						
Female	0.536	0.499				
Black	0.255	0.436				
Father's educ (years)	10.32	5.752				
Mother's educ (years)	11.79	4.242				

* In year 2000 dollars.

table enter the individual's decision to smoke, either directly through the utility received from smoking or through the cost of smoking. The NLSY does not ask about parent's smoking

---

[16]See the Data Appendix for summary statistics for all waves.

behavior. Parental smoking behavior potentially enters the individual's smoking decision through the individual's beliefs as well as through the cost of smoking. Parental education and other parental characteristics could serve as a proxy for parent smoking behavior.

Figure 1 presents the proportion of individuals in each smoking category by age. The proportion of individuals choosing to smoke increases steadily during the teenage years, reaches a peak for individuals in their early 20s, and declines slightly as individuals progress through their 20s. The decline in smoking rates for individuals in their 20s is primarily due to a lower proportion of light smokers. The proportion of moderate and heavy smokers remains relatively constant after reaching a peak around the age of 20. Figure 2 presents the proportion of current smokers by gender and race. Blacks have a substantially lower rate of smoking compared to other ethnic groups, and females have a lower smoking rate than males.

4.3   Cigarette Prices and State Excise Tax Data

The cigarette tax and price data used in this paper are from Orzechowski and Walker's *Tax Burden on Tobacco*. The price used is a sales weighted average of the premium brand cigarettes sold in a given year. Cigarettes are taxed at the federal and state level. In some instances they are also taxed at the county and municipal level. The federal cigarette tax in 2011 was $1.01 per pack. The tax rates vary considerably across states. In 2011, state cigarette taxes ranged from a low of $0.17 per pack in Missouri to a high of $4.24 in New York. At the start of the sample period in 1997, state cigarette taxes ranged from a low of $0.025 in Virginia to a high of $0.825 in Washington. Historically, the states with the lowest tax rates on tobacco are the tobacco-producing states of the southeast. From 1997-2011, only two states have had a constant tax rate, and most states have had multiple tax increases over the period. The variation in tax rates is largely responsible for the variation in the retail price of cigarettes across states. In 2011, the average retail price of cigarettes per pack ranged from $4.70 in Missouri to $10.29 in New York.

Table 6 present the summary statistics across the 50 states and the District of Columbia of the real price of cigarettes as well as the real total tax. The average real price approximately doubles over the sample period, and the amount of the average real tax increases by about three times. Over the time frame, the variability in both the prices and taxes across states increases. Most years the average real price increases due to increases in taxes. In years when there are no tax changes in a state, the real price of cigarettes falls as the nominal price increases less than inflation.

Figure 3 shows how real retail cigarette prices and taxes have changed over time in
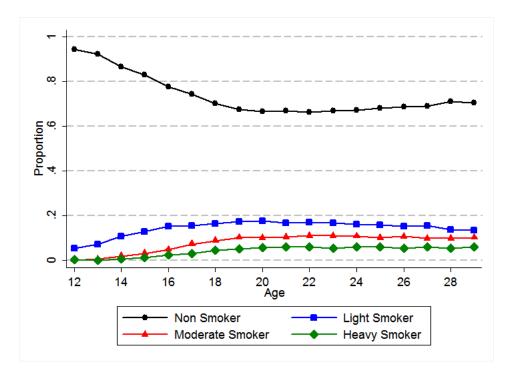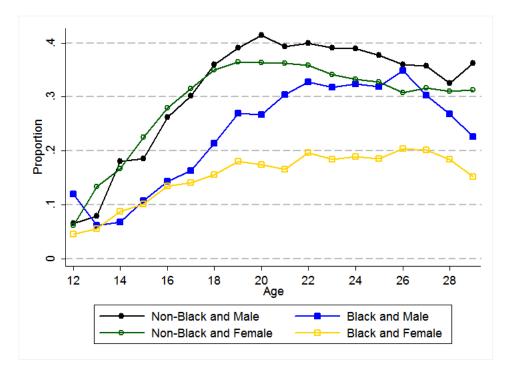
Figure 1: Smoking Choice Probabilities by Age

Figure 2: Gender and Racial Differences in Smoking Rates by Age

18

Table 6 : Summary Statistics of State Tobacco Price and Taxes

Year	Real Price				Real Tax (State + Federal)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
1997	2.265	0.327	1.796	3.305	0.633	0.217	0.284	1.143
1998	2.477	0.353	2.013	3.576	0.661	0.256	0.280	1.310
1999	3.200	0.361	2.698	4.304	0.670	0.271	0.274	1.282
2000	3.318	0.393	2.777	4.512	0.760	0.279	0.365	1.450
2001	3.500	0.370	3.035	4.458	0.752	0.291	0.355	1.410
2002	3.787	0.550	3.107	5.671	0.927	0.432	0.397	1.819
2003	3.843	0.567	3.157	5.452	1.041	0.453	0.388	2.284
2004	3.815	0.615	3.088	5.343	1.064	0.515	0.378	2.598
2005	3.847	0.643	3.095	5.292	1.157	0.528	0.406	2.513
2006	3.772	0.673	2.899	5.365	1.135	0.527	0.393	2.533
2007	3.883	0.652	2.906	5.520	1.191	0.518	0.382	2.462
2008	3.896	0.737	2.893	5.687	1.244	0.579	0.368	2.511
2009	4.711	0.846	3.406	6.458	1.856	0.628	0.867	3.588

New York and North Carolina. Much of the price difference between these two states can be attributed to the difference in their cigarette taxes. Also, the increase in the price of cigarettes over time is driven by the increase in the tax rates. Other factors behind the increase in cigarette prices over this time period are the Tobacco Master Settlement Agreement in 1998, and the increase in the federal cigarette tax rate in 2009.[17] Figure 4 shows the distribution of state cigarette tax rates over time. At the beginning of the sample period, state cigarette taxes were relatively low. Over time, both the mean and variance of the state cigarette tax distribution increased.

4.4 State Level Tobacco Policy Data

In addition to tobacco excise taxes, there are many other policies that states can pursue to influence the level of youth smoking. Some of these policies enter into the individual's problem through the budget constraint by imposing non-monetary costs on obtaining tobacco. Some examples of policies that enter the individual's problem in this way are restrictions

---

[17]In 1998, 46 states came to an agreement with the four largest cigarette manufacturers. The states agreed to drop their lawsuits against the tobacco companies, which sought compensation for the treatment of tobacco-related illnesses in the Medicaid system. In exchange, the tobacco companies agreed to a monetary settlement, restrictions on the marketing of tobacco products to young people, and the funding of a national anti-smoking organization. The tobacco companies raised the price of cigarettes by 45 cents per pack in response to the settlement to cover the payments to the states.
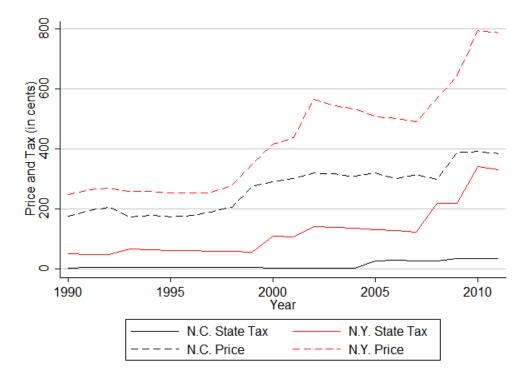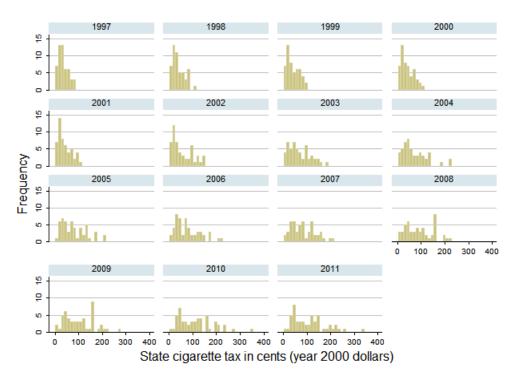
Figure 3: Real Cigarette Taxes and Prices in NY and NC (in year 2000 dollars)

Figure 4: Distribution of Real State Cigarette Taxes by Year

on the sale of tobacco to minors, bans on the sale of tobacco in vending machines, and restrictions on free samples of tobacco products. Another way for tobacco policies to influence behavior is through restrictions on tobacco consumption. The overall utility one receives from smoking will be less if there are restrictions on where and when one can smoke. Examples of restrictions on tobacco consumption are indoor smoking bans and smoke-free schools. Finally, some tobacco policies influence the individual's beliefs and expectations. In the context of this paper, these policies influence the individual's initial prior beliefs. Examples include restrictions on cigarette advertisements, funding of tobacco prevention and education programs, and requiring tobacco education in schools. The data on state tobacco policies are from the Centers for Disease Control (CDC), the National Cancer Institute (NCI), and the Substance Abuse and Mental Health Services Administration (SAMHSA).

5 Estimation

5.1 Likelihood Function

Define the conditional value function for alternative j as the deterministic portion of flow utility from that alternative (i.e., utility minus the preference shock) plus the discounted expected future value of lifetime utility conditional on alternative j being chosen. Then, the conditional value function associated with alternative j in period t is given by:

$$
v_{n,t}^j(S_{n,t}, \Gamma_{n,t}, X_{n,t}) = \big(\alpha_n + \mathbf{E}_t[\rho_n|\Gamma_{n,t}]g(S_{n,t}) + \xi_j X_{n,t}\big)z(a_j) - \mathbf{E}_t[\tau_n|\Gamma_{n,t}]S_{n,t}\mathbf{1}[a_j > 0]
$$
$$
- \mathbf{E}_t[\omega_n|\Gamma_{n,t}]q(a_j, S_{n,t})\mathbf{1}[a_j < S_{n,t}] - \gamma_n p_t a_j + \beta\mathbf{E}_t[V_{n,t+1}(S_{n,t+1}, \Gamma_{n,t+1}, X_{n,t+1})|d_{n,t}^j = 1]
$$
(13)

where

$$
V_{n,t+1}(S_{n,t+1}, \Gamma_{n,t+1}, X_{n,t+1}) = \mathbf{E}[\max_j v_{n,t+1}^j(S_{n,t+1}, \Gamma_{n,t+1}, X_{n,t+1}) + \epsilon_{n,t+1}^j] \qquad (14)
$$

The expectation over the future value term is taken with respect to the distribution of future beliefs, future demographic state variables, and future prices. The evaluation of current period utility depends upon the mean of the prior beliefs only. The variance of the prior does affect the expectation over future beliefs. The utility from not smoking is normalized to include the cost of withdrawal only, so $\xi_1 = 0$. The state variables are the level of smoking stock (i.e., last period's smoking decision) and the individual's beliefs, denoted by Γ, which

include beliefs about parameter values and future prices.[18] I assume an i.i.d. type I extreme value (EV) preference shock.[19] The choice probabilities after experimentation are given by:

$$P_{n,t}^{j} = \frac{e^{v_{n,t}^{j}}}{\sum_{k=1}^{J} e^{v_{n,t}^{k}}} \quad \text{for } j = 1, \ldots, J \tag{15}$$

For individuals who have never smoked, they first choose whether or not to experiment, and then, conditional on experimenting, they decide the level of smoking. Let $d_{n,t}^{E}$ be a dummy variable that equals one if the individual experiments in period t. The conditional value of experimenting is:

$$v_{n,t}^{E} = (\mathbf{E}[\alpha_n | \Gamma_{n,t}] + \xi^{E} X_{n,t}) z(a^{E}) - \gamma p_t a^{E} + \mathbf{E}_t[V_{n,t} | d_{n,t}^{E} = 1] \tag{16}$$

The conditional value function of not experimenting is simply the discounted expected maximum of the next period's value function conditional on not experimenting and not consuming any of the addictive good. The probability for experimenting, $P_{n,t}^{E}$, is given by the Logistic cumulative distribution function. For an individual who has never smoked prior to period t, the behavior in period t is captured by the joint probability of experimenting and level of smoking $(P_{n,t}^{E} P_{n,t}^{j})$. The decision to experiment is made based on the individual's belief about his level of α, so $P_{n,t}^{E}$ is calculated based on an individual's beliefs. If he decides to experiment, he learns his true level of α, so $P_{n,t}^{j}$ is calculated using the individual's true value of α.

There are a total of N individuals, and each individual is observed for a total of $T + 1$ periods. The likelihood of individual n making the sequence of choices $\{\cup_j \{d_{n,t}^{j}\}, d_{n,t}^{E}\}_{t=1}^{T}$ is:

$$L_n\left(\gamma, \xi \mid \theta_n, \Gamma_{n,0}, \Lambda_n\right) = \prod_{t=0}^{T} \left((\prod_{j=1}^{J} P_{n,t}^{j\, d_{n,t}^{j}})^{A_{n,t}} * \left[(1 - P_{n,t}^{E})^{1-d_{n,t}^{E}} \left(P_{n,t}^{E} \prod_{j=1}^{J} P_{n,t}^{j\, d_{n,t}^{j}} \right)^{d_{n,t}^{E}} \right]^{1-A_{n,t}} \right) \tag{17}$$

where $A_{n,t}$ is an indicator for the individual having ever smoked prior to period t. If the individual has smoked prior to period t (i.e., $A_{n,t} = 1$), the individual makes a consumption decision. If the individual has not smoked prior to period t (i.e., $A_{n,t} = 0$), then the

---

[18]The price process has yet to be formally incorporated into the model, so the following estimation routine assumes perfect knowledge of future prices. The proposed estimation routine can be extended to estimate the parameters of a random price process.

[19]One of the major limitations of the multinomial logit model is the assumption that the shocks are uncorrelated over alternatives (i.e., the Independence of Irrelevant Alternatives (IIA) assumption). The use of random parameter, or mixed, logit can overcome the limitations of this assumption. In fact, mixed multinomial logit can approximate any discrete choice model derived from a random utility model to within any arbitrary degree of precision (McFadden and Train, 2000).

individual makes a sequential experimentation and consumption decision. This individual likelihood is conditional on the individual's true addictive parameters (θ_n), the distribution of individual's initial prior beliefs $(\Gamma_{n,0})$, and a given sequence of signal noise draws $(\Lambda_n = \{\psi_{n,t}, \lambda_{n,t}, \eta_{n,t}\}_{t=0}^{T})$. This formulation is equivalent to conditioning on the individual's beliefs at time t since the beliefs in time t are completely determined by the individual's initial prior, the sequence of signal noise, and the sequence of choices. Since the individual's true parameters and signal noise sequences are not observed by the researcher, the unconditional likelihood is calculated by integrating the conditional likelihood over the distribution of these unobserved variables:

$$L_n(\gamma, \xi, \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma) =$$
$$\int_\theta \int_\Lambda L_n\left(\gamma, \xi \mid \theta_n, m_{n,0}, \Sigma_{n,0}, \Lambda_n\right) dF(\Lambda | \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2) \, dF(\theta | \bar{\theta}, \Sigma) \quad (18)$$

and the full log-likelihood function is given by:

$$\mathscr{L}(\gamma, \xi, \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma) = \sum_n log\left(L_n(\gamma, \xi, \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma)\right) \quad (19)$$

The total dimensions of unobserved variables is $3 * T + 4$.[20] The integrals do not have a closed form solution, so they must be approximated numerically. The parameters to be estimated include the utility function parameters (γ, ξ), the mean and covariance matrix of the population distribution of the rational addiction parameters $(\bar{\theta}, \Sigma)$, and the variances of the signal noise distributions $(\sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2)$.

5.2   Identification

The model parameters are identified through the observed sequences of smoking decisions. The parameters ξ and γ are identified through differences in smoking decisions between individuals with different observable characteristics. The price sensitivity parameter γ is identified by both cross-sectional variation and variation over time in the price of cigarettes. The utility from not smoking when the smoking stock is zero is normalized to zero. The parameter α affects the utility for each level of smoking regardless of past smoking. The reinforcement parameter captures the effect of the interaction between the current level of smoking and the smoking stock. The tolerance parameter only depends on the smoking

---

[20]The dimension of the unobserved signals is likely to be less than 3*T since some of the signals are observed by the researcher. Based upon the sequence of actions, the researcher knows whether or not a signal is received in a given period.

stock, so, for a given level of smoking stock, a change in the tolerance parameter only affects the probability of smoking versus not smoking. The reinforcement parameter affects the probability of smoking versus not smoking, but it also affects the probability of each level of smoking. The withdrawal parameter only affects the utility of a reduction in the level of smoking from one period to the next, so this parameter is identified by smokers who reduce their level of smoking or quit smoking entirely. The match, tolerance, reinforcement, withdrawal, and price sensitivity parameters do not vary across alternatives. Differences in utility for the different levels of smoking for these parameters are ultimately a result of the functional form assumptions.

The individual-specific parameters are not point identified for each individual. There is no way to estimate a specific value of these parameters for each individual. Also, since these parameters are continuous, a distributional assumption is required for the population distribution of parameters. Then, given that the conditional value function is defined over the support of the distribution of the unobserved continuous variables, the parameters of the population distribution (mean and covariance) are identified. The identification behind the learning process is driven by the fact that the valuation an individual attributes to each alternative depends upon the individual's current beliefs only and not the individual's true parameters. The individual's beliefs converge to the true parameters as the individual receives additional signals. Therefore, individuals with a lot of experience will behave according to their true parameter values. Also, if an individual knows his true parameter values, he can use the model to calculate an optimal consumption sequence. Differences between the optimal consumption sequence if the individual knows his true parameter values and the decisions of the individual when he is inexperienced are driven by the difference between the individual's beliefs and his true parameter values. The speed at which the individual's consumption sequence converges to the optimal consumption sequence with full knowledge identifies the speed of learning (i.e., the variance of the signals). Additional restrictions on the learning process are necessary for identification. These include restrictions on the initial prior beliefs (Rational Expectations), distributional assumptions for the beliefs and signals (both Normal), and Bayesian updating.

5.3 Estimation Procedure

There are several computational requirements that make estimation of the parameters of the model by Full Information Maximum Likelihood difficult. The main issue is that the evaluation of the log likelihood function requires integrating over the continuous distribution of population parameters and over all possible sequences of signal noise. Simulated maximum

likelihood is one method that is used to overcome this problem. The unconditional likelihood function is approximated numerically by taking random draws from the distribution of the unobserved variable, evaluating the conditional likelihood, and taking the average of the conditional likelihoods over the draws. Evaluating the conditional likelihood, however, for a single draw still involves significant computation. The solution to the individual's problem requires integrating over future beliefs, which are multidimensional continuous variables. One way to reduce the computational burden of evaluating the value function is to use the Conditional Choice Probability (CCP) method of Hotz and Miller (1993).

Hotz and Miller (1993) show that when the preference shock has a GEV distribution, the future value term in the conditional value function can be expressed as a function of future flow utilities and conditional choice probabilities (CCPs). For certain classes of problems (e.g., optimal stopping problems), taking the difference in conditional value functions leads to the future value term only containing one period ahead flow utilities and CCPs. In other problems, the future value term associated with the difference in conditional value functions contains flow utilities and CCPs for a finite number of future periods. This property is called finite dependence, and it is a feature of the problem in this paper.[21] Standard CCP estimation involves estimating the CCPs in a first stage using the data and using the estimated CCPs to calculate the individual's value function. One limitation of the standard method is that it do not allow for unobserved heterogeneity. Arcidiacono and Miller (2011) develop a method of CCP estimation that allows for a finite distribution of unobserved heterogeneity by using the Expectation Maximization (EM) algorithm. The unobserved heterogeneity in this paper are the individual's beliefs and the individual's true parameter values, which are both continuous. Matsumoto (2015) extend the work of Arcidiacono and Miller (2011) to allow for a continuous distribution of unobserved heterogeneity.

It can be shown that the values of the parameters that maximize the likelihood function (19) also maximize the following transformed likelihood function:[22]

$$\mathscr{L}(\gamma, \xi, \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma) = \sum_n \int_\theta \int_\Lambda \pi_n(\theta_n, \Lambda_n) \Bigg(\sum_t$$

$$(1 - A_{n,t}) \big[(1 - d_{n,t}^E) log(1 - P_{n,t}^E) + d_{n,t}^E log(P_{n,t}^E) \big] + \sum_j d_{n,t}^j log(P_{n,t}^j) \Bigg) d\Lambda \, d\theta \quad (20)$$

where π is the conditional probability that the parameter values are θ, θ_0, and Λ given the

---

[21] See the Estimation Appendix for the derivation of the CCP representation of the future value term.

[22] This is the expected conditional (on the unobserved variables) likelihood, where the expectation is taken with respect to the distribution of the unobserved variables conditional on the observed variables and the choices.

observed choices. This conditional probability is given by:

$$\pi_n(\theta_n, \Lambda_n) = \frac{f(\theta_n|\bar{\theta}, \Sigma) f(\Lambda_n|\sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2) \prod_t L_{n,t}(\theta_n, \Lambda_n)}{\int_\theta \int_\Lambda \prod_t L_{n,t}(\theta_n, \Lambda_n) f(\Lambda_n|\sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2) f(\theta_n|\bar{\theta}, \Sigma) \, d\Lambda \, d\theta} \tag{21}$$

The estimation routine in this paper used the likelihood function in equation 20. The procedure starts by taking M draws from the distribution of the unobserved variables for each individual as well as initial guesses for the values of the parameters and the CCPs. The estimation proceeds by using the EM algorithm, specifically a simulated EM algorithm (SEM). The EM algorithm is an iterative procedure that alternates between an expectation step (or E-step) and a maximization step (or M-step). The E-step updates the CCPs and π using the prior iteration values of the parameters and CCPs. The M-step updates the value of the parameters by maximizing the likelihood function using the updated CCPs and π. The estimation continues to iterate over these two steps until the parameter estimates converge. The use of the EM algorithm to incorporate unobserved heterogeneity has several advantages.[23] The most significant advantage is that the EM algorithm, or the SEM algorithm in the current context, reintroduces additive separability of the likelihood function. This property allows for sequential estimation of the likelihood function. In the current context, additive separability of the likelihood function allows for the parameters of the experimentation and consumption decisions to be estimated separately. The estimation procedure is presented in greater detail in the Estimation Appendix.

5.4   Initial Conditions

The first period that individuals are observed in the NLSY97 is not the same as the initial period of the individual's optimization problem. That is, individuals may enter the estimation sample having already smoked. The values of the state variables in the initial wave of data depend on prior decisions and state variables that are not observed by the researcher. Some individuals have never smoked by the first wave. Others have smoked at some point prior to the first wave but are not observed to smoke in the first wave. Finally, some individuals are regular smokers at the first wave. The latter two groups present an initial conditions problem both in that the prior year's smoking is not observed in the first period and it is not observed how much they have learned. Individual's initial prior beliefs also present an initial conditions problem. I assume that individual initial priors are identical to the population distribution of the parameters (i.e., Rational Expectations).[24]

---

[23]See Arcidiacono and Jones (2003) for a full discussion.

[24]In future work, I will attempt to parameterize the initial priors by allowing the mean of the initial priors (and perhaps the variance as well) to be functions of individual characteristics and state tobacco policies.

For individuals who have smoked prior to the first wave, the amount smoked in the period prior to the first wave is treated as discrete unobserved heterogeneity. The individual likelihood is calculated for each possible alternative in period $t = 0$. The probability that the individual selected alternative j in period $t = 0$ is:

$$P_{n,0}^j = \frac{1}{1 + exp(\xi_{IC}^j X_{n,0}^{IC})} \tag{22}$$

The individual likelihood is calculated by multiplying the likelihood conditional on selecting alternative j in period $t = 0$ by the probability $P_{n,0}^j$ and summing over the alternatives.

5.5 Functional Forms

The utility for the smoking level associated with alternative j, contains several modifying functions. The purpose of these functions is to allow for utility to be nonlinear in both the level of smoking and the level of past smoking. In order to estimate the parameters of the model, these generic functions must be replaced with specific functional forms. The function $z(a)$ incorporates the standard utility function assumptions except that the marginal utility of smoking is positive for a level of smoking equal to zero. Also, the utility from not smoking is normalized to zero. The function $z(a)$ is assumed to take the following form: $z(a) = log(1 + a)$. The function that modifies the effect of reinforcement takes the following form: $g(S_t) = \sqrt{S_t}$. Finally, the function that modifies the withdrawal effect has the following form:

$$q(a_j, S_{n,t}) = S_{n,t} * \left(1 - exp\left(c * (a_j - S_{n,t})\right)\right) \tag{23}$$

When the individual smokes the same amount as the prior period, $q = 0$. If the individual smokes less than the prior period, the withdrawal cost is positive. For a given level of last period smoking, the withdrawal cost decreases as the individual smokes more in the current period. This decrease occurs at an increasing rate. The parameter c affects the curvature of the function q as well as the maximum possible withdrawal cost. This parameter is initially fixed at a value of 0.15. Finally, the discount parameter β is set to 0.95.

6 Results

6.1 Parameter Estimates

This section presents the parameter estimates for the model. The estimation sample includes white males who are observed in every time period. The version of the model that is estimated

Table 7 : Estimation Results, Population Distribution Parameter Estimates

Parameter	Description	Model with Learning	Model without Learning
$\bar{\alpha}$	Mean of match parameter	-0.925	-1.351 (0.042)
$\bar{\rho}$	Mean of reinforcement parameter	0.361	0.363 (0.011)
$\bar{\omega}$	Mean of withdrawal parameter	0.466	0.150 (0.014)
$Var(\alpha)$	Variance of match parameter	0.124	4.723 (0.221)
$Var(\rho)$	Variance of reinforcement parameter	0.204	0.167 (0.009)
$Var(\omega)$	Variance of withdrawal parameter	0.472	0.038 (0.004)
$Cov(\alpha,\rho)$	Covariance of match and reinforcement	0.063	0.867 (0.031)
$Cov(\rho,\omega)$	Covariance of reinforcement and withdrawal	-0.104	0.061 (0.006)
$Cov(\alpha,\omega)$	Covariance of match and withdrawal	-0.116	0.274 (0.032)
σ_λ	Standard deviation of reinforcement signal	0.994	-
σ_η	Standard deviation of withdrawal signal	1.086	-

Note: Standard Errors in parentheses

differs from the model presented earlier in that the tolerance parameter τ is not estimated and set to zero. Table 7 presents the parameter estimates for the model with learning as well as the model without learning. The match parameter is negative for a large majority of the population. Even individuals with a negative match parameter could receive positive utility from smoking due to the effect of reinforcement. Individuals below the age of 18 experience a utility cost from smoking, which is likely due to their inability to purchase cigarettes legally. This cost is increasing in the level of smoking. The variance of the signals is significantly different from zero, which suggests that the learning component of the model is significant.

In order to test the importance of learning, I estimate a version of the model without learning. In the model without learning individuals are assumed to know the value of their parameters, but the parameters vary across individuals.[25] Estimates of the parameters from the model without learning differ in important ways from the parameter estimates from the model with learning. The mean of the population distribution of the match parameter is larger in magnitude in the model without learning and has much higher variability in the population. The mean value of the reinforcement parameter is nearly the same in both models, and the mean of the withdrawal parameter is smaller in the model without learning. The population variance of the reinforcement and withdrawal parameters is smaller in the model without learning.

---

[25]The model without learning corresponds to a restricted version of the model with learning. Specifically, the model without learning is equivalent to the model with learning where the mean of the initial prior is set to the individual's true parameter value and the variance of the initial prior is set to zero.

Table 8 : Estimation Results, Coefficients on Observable Variables

Preference Shifting Variables		
Variable	Model with Learning	Model without Learning
Years until age 18, light smoking	-0.238	-0.435 (0.052)
Years until age 18, moderate smoking	-0.316	-0.903 (0.176)
Years until age 18, heavy smoking	-0.074	-1.016 (0.337)
Years until age 18 squared, light smoking	-0.034	0.005 (0.012)
Years until age 18 squared, moderate smoking	-0.184	0.001 (0.039)
Years until age 18 squared, heavy smoking	-0.311	-0.013 (0.074)
Married, light smoking	-0.212	-0.052 (0.039)
Married, moderate smoking	-0.215	-0.127 (0.051)
Married, heavy smoking	-0.511	-0.041 (0.048)
Has children in household, light smoking	0.114	0.167 (0.039)
Has children in household, moderate smoking	-0.113	0.107 (0.051)
Has children in household, heavy smoking	0.026	-0.032 (0.049)
Price Sensitivity Variables		
$\bar{\gamma}$, Mean Price sensitivity	0.870	0.442 (0.025)
Under age 18	0.036	-0.053 (0.112)
Employed	0.004	-0.009 (0.018)
Income greater than $20k	-0.070	-0.024 (0.009)
Unobserved Prior Consumption Variables		
Constant, light smoking	1.228	5.913 (4.978)
Constant, moderate smoking	-0.354	1.047 (1.189)
Constant, heavy smoking	-0.379	-0.543 (1.623)
Years since first smoked, light smoking	2.107	-3.581 (5.137)
Years since first smoked, moderate smoking	-0.271	-0.135 (0.213)
Years since first smoked, heavy smoking	-0.558	-0.153 (0.330)
Experimentation Variables		
Years until age 18	-4.697	-
Years until age 18 squared	0.186	-
Married	2.20	-
Has children in household	3.047	-
a^E	3.565	-

Table 8 presents the estimates of the coefficients on the observable variables. The first panel includes the estimates for the variables that enter the utility function as preference shifters. The next panel includes the variables that affect price sensitivity, and is followed by the parameters that affect the probability of different levels of prior unobserved consumption. The last panel includes the variables that enter the utility of experimentation. Note that there is no experimentation decision in the model without learning since individuals already know they value of the match parameter. Other than the age variables, the coefficients on observable characteristics tend to be relatively small in magnitude

6.2 Model Fit

Table 9 presents the observed transition probabilities from the data as well as the transition probabilities from simulated outcomes generated using the model with learning and the estimated parameters. The model is able to fit the observed transition probabilities well. For smoking transitions for individuals under 18, the simulated data tends to overstate the persistence in smoking behavior, particularly for remaining a nonsmoker and a heavy smoker. The model is better able to fit the transition probabilities for individuals over 18 years old.

Table 9 : Transition Probabilities, Observed and Simulated Data

Under 18 years old								
	Smoking level at t							
	Observed Data				Simulated Data			
Smoking level at $t-1$	None	Light	Moderate	Heavy	None	Light	Moderate	Heavy
None	86.81	10.47	1.83	0.88	92.39	5.23	1.39	0.99
Light	35.21	44.92	14.25	5.62	29.04	49.89	14.98	6.09
Moderate	15.70	10.74	49.59	23.97	11.98	16.71	44.91	26.40
Heavy	9.84	8.20	22.95	59.02	10.40	6.67	15.82	67.10
Over 18 years old								
	Smoking level at t							
	Observed Data				Simulated Data			
Smoking level at $t-1$	None	Light	Moderate	Heavy	None	Light	Moderate	Heavy
None	89.60	8.09	1.52	0.79	86.59	10.78	1.88	0.75
Light	27.12	57.88	12.24	2.76	28.33	56.14	12.44	3.09
Moderate	9.90	16.30	55.82	17.99	11.92	14.75	55.17	18.15
Heavy	6.30	5.09	21.26	67.35	6.98	4.06	17.77	71.18

Figure 5 shows the proportion of individuals in each smoking category by age for both the observed and simulated data using the estimated model with learning. The simulated data closely match the observed age profile of smoking behavior. Figure 6 compares the proportion of individuals in each smoking category by age for the observed data and for simulated data

using the model without learning. The model without learning does a relatively poor job in matching the observed data.

6.3 Policy Simulations

In this section, I use the parameter estimates from the model with learning to conduct policy counterfactual experiments. I consider policies that alter cigarette prices, beliefs about withdrawal, and the legal smoking age.[26]

6.3.1 Prices

The tobacco excise tax is a popular policy tool among policymakers and anti-smoking advocates to reduce the level of smoking. The specific policy experiment is doubling the price of cigarettes. Under the counterfactual policy, individuals are faced with a price of cigarettes that is two times what is observed in the data. This counterfactual measures the long run impact of a change in the price of tobacco. Figure 7 depicts the smoking rates by age for the baseline simulation and the simulated data under the counterfactual prices.

Doubling prices has a dramatic effect on the proportion of smokers. For individuals over 18 years old, the proportion of nonsmokers increases by over 30 percentage points as a result of the higher prices. The proportion of light smokers decreases by around 90 percent, and the proportion of moderate and heavy smokers decreases by around 75 percent.[27]

6.3.2 Beliefs

Prior to experimentation with cigarettes, a young individual has beliefs about his smoking preference parameters. The model imposes rational expectations for the initial beliefs. However, the individual's initial beliefs are likely influenced by a number of factors and could potentially be influenced by anti-smoking policies. For example, advertisements that highlight the addictive nature of cigarettes and the difficulty of quitting smoking may affect the individual's beliefs about the value of the parameters that govern the effects of reinforcement, tolerance, and withdrawal.

In this counterfactual experiment, the mean initial prior for the withdrawal parameter is increased by one standard deviation of the population distribution. Only the initial belief about the withdrawal parameter changes. The actual distribution of the withdrawal parameter in the population is the same.

---

[26]The legal smoking age is the minimum age at which an individual can legally purchase tobacco products.

[27]One concern with this counterfactual is that the CCPs are estimated using the data, so they are only identified for values of the state variables that are observed in the data.
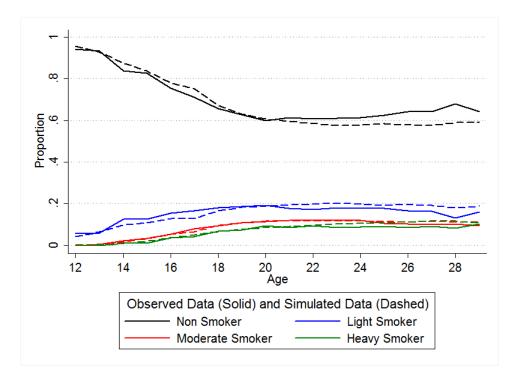
Figure 5: Smoking Rates by Age, Observed and Simulated Data from the Model with Learning
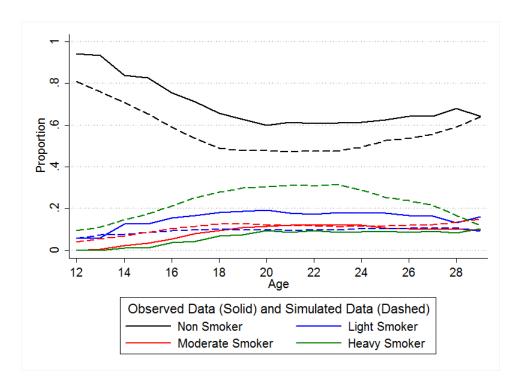

Figure 6: Smoking Rates by Age, Observed and Simulated Data from Model without Learning

Increasing the mean value of the withdrawal parameter in one's initial prior beliefs causes a reduction in the overall level of smoking. Individuals are now less likely to experiment with smoking given the higher anticipated cost of quitting. The higher expected withdrawal cost results in a large reduction in the proportion of moderate and heavy smokers. The proportion of light smokers is about the same relative to the baseline simulation. Some individuals who were light smokers under the baseline simulation decide not to smoke under the counterfactual. These individuals are offset by those who were moderate and heavy smokers under the baseline who remain light smokers for a longer period before transitioning to higher levels of smoking under the counterfactual. The increase in the expected withdrawal cost has the effect of extending the experimentation period.

6.3.3   Legal Smoking Age

Another possible policy tool that targets youth smoking is the minimum legal age to purchase tobacco. In this counterfactual experiment, the effect of increasing the minimum legal purchase age to 19 years old. Relative to the baseline simulation, increasing the purchase age is effective in reducing smoking among teenagers. However, increasing the legal purchase age only delays smoking rather than preventing it. The smoking rates converge to the baseline simulation for all smoking categories once individuals are able to legally purchase tobacco.

These counterfactual simulations confirm that increasing the price of cigarettes is an effective policy tool to reduce the prevalence of smoking. Changing the legal smoking age would have the effect of reducing youth smoking, but would likely have only a minimal impact in reducing smoking in the broader population. Policies that target an individual's initial prior beliefs about the utility of smoking could be very effective in reducing smoking. Specifically, increasing an individual's beliefs about the withdrawal cost would lead to a relatively large reduction in the probability that individual would become a heavy smoker.

6.4   Welfare Analysis

Thus far, the policy analysis has evaluated the effectiveness of alternative policies in reducing the level of smoking without taking into account the effect of the policies on an individual's welfare. One of the key advantages of the model with learning is the ability of the model to explain regret. In the standard RA model, policies that increase the cost of smoking will lower the welfare of every individual. In the model with learning, a policy that increases the cost of smoking may lower the welfare of some individuals, but it may increase the welfare of others. If an individual who would later regret the decision to become a smoker decided not to smoke because of the policy, that policy would increase his welfare (assuming the policy
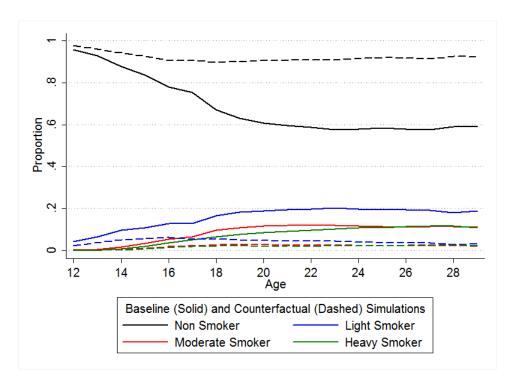
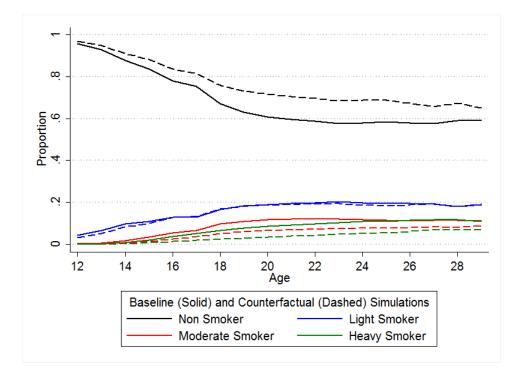Figure 7: Smoking Rates by Age, Baseline Simulation and Price Counterfactual Data

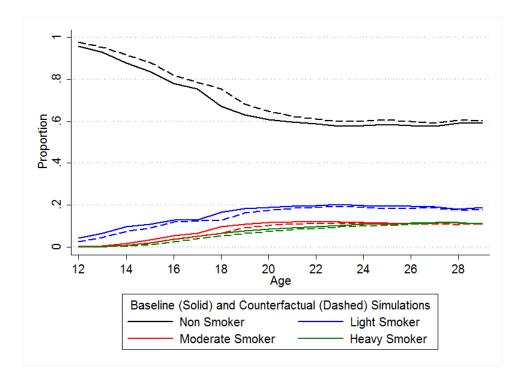Figure 8: Smoking Rates by Age, Baseline Simulation and Beliefs Counterfactual Data

Figure 9: Smoking Rates by Age, Baseline Simulation and Smoking Age Counterfactual Data

did not affect the utility from not smoking).

To evaluate the effect of the alternative policies on welfare, I calculate the expected life-time utility (ELU) for each individual's simulated sequence of choices. This measure includes the deterministic portion of utility evaluated using the individual's true parameters.[28] The objective in constructing this measure of welfare is to determine the effect of the alternative policies on the individual's ex-post welfare. The welfare measure is normalized with respect to the sequence of never smoking. The expected flow utility portion of the welfare measure is already normalized from the normalization of the utility function.

Table 10 shows the summary statistics for the ELU measure of individual welfare for the baseline simulation as well as under the different counterfactuals. The ELU of an individual is zero for an individual who never smokes. The mean ELU for the baseline simulation is 1.773 and the median is -1.900. A vast majority of individuals have an ELU that is less than zero, indicating that they are ex-post worse off than if they had never smoked. In the price counterfactual, the proportion of individuals who never smoke increases. By increasing the proportion of individuals who never smoke, these policies decrease the proportion of individuals with a negative ELU, but the policies also decrease the proportion of individuals

---

[28]Note that expected lifetime utility refers to the ex-post deterministic flow utility up to the end of the sample period.

with a positive ELU. Individuals who enjoy smoking are negatively affected by the increase in the price of cigarettes. Increasing initial beliefs about the difficulty of quitting smoking decreases welfare on average as it decreases the proportion of individuals with a positive ELU and increases the proportion of individuals with a negative ELU.

Table 10 : Welfare Analysis

	Mean ELU	Std ELU	Med ELU	Pr(ELU<0)	Pr(ELU=0)	Pr(ELU>0)
Baseline	1.773	27.172	-1.900	0.726	0.049	0.225
Price Cf	-0.460	13.381	-0.172	0.719	0.095	0.186
Belief Cf	-0.085	20.073	-1.049	0.739	0.070	0.191
Age Cf	1.126	24.886	-1.658	0.730	0.052	0.218

7 Conclusion

This research develops a model of rational addiction with learning in order to explain the smoking initiation decision of young people. Estimation of the structural parameters of the model requires significant computational resources, and is not computationally feasible using a full solution estimation routine unless significant restrictions are placed on the model. Therefore, this paper proposes the use of an alternative estimation routine. This estimation routine uses the EM algorithm and CCP estimation, which reduces the computational burden of estimating the structural parameters of the model.

Overall, the model is able to fit the data well. In particular, the model with learning fits the data significantly better than the model without learning. The estimated parameters of the model are used to conduct counterfactual policy experiments. Since an individual's decision to smoke depends upon his beliefs about his smoking preference parameters, policies that affect one's beliefs can have a significant impact on smoking behavior. Increasing individuals' beliefs about the difficulty of quitting smoking is effective at reducing the number of heavy smokers. Increasing cigarette prices is shown to be an effective policy tool to reduce youth smoking, although the model without learning overstates the importance of the price of cigarettes. An increase in the legal age to purchase cigarettes would lead to a decrease in the number of youth smokers, but it would only delay smoking initiation so adult smoking behavior would not be affected. The analysis of individual welfare supports the use of taxes as an anti-smoking policy tool. An increase in the price of cigarettes improves the ex-post level of utility for some individuals by discouraging those who would later regret the decision to smoke from ever experimenting with cigarettes, but hurts those who do enjoy smoking.

The results of this paper suggest several potential avenues of future research. First, the analysis performed considers the demand side of the market. Although the analysis in this paper demonstrates the importance of learning in explaining cigarette demand, the model would need to be extended to incorporate optimal firm behavior in order to better capture the general equilibrium effects of policy changes. An individual's initial beliefs are an important determinant of early smoking behavior, as changing the initial beliefs was shown to have a large effect on behavior in the counterfactual simulation. Additional exploration of how these initial beliefs are formed would be useful. In particular, to what degree are the individual's initial beliefs influenced by the smoking behavior of others (e.g., parents, siblings, peers) or by advertising (either pro- or anti-smoking). Finally, the importance of learning in explaining youth smoking behavior begs the question of how learning about cigarette smoking preferences may impact learning about preferences for consuming other addictive goods such as alcohol or illegal drugs. There are potential knowledge spillovers about the dynamic effects of consuming an addictive good (i.e., tolerance, reinforcement, and withdrawal), which may be correlated across different addictive goods for an individual.

References

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

Arcidiacono, P. and J. Jones (2003). Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm. *Econometrica 71*(3), 933–946.

Arcidiacono, P. and R. Miller (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica 7*(6), 1823–1868.

Arcidiacono, P., H. Sieg, and F. 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 (1994). An Empirical Analysis of Cigarette Addiction. *American Economic Review 84*(3), 396–418.

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

Benabou, R. and J. Tirole (2004). Willpower and Personal Rules. *Journal of Political Economy 112*(4), 848–886.

Bernheim, B. and A. Rangel (2004). Addiction and Cue-Triggered Decision Processes. *American Economic Review 94*(5), 1558–1590.

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

Chaloupka, F. and K. Warner (2000). The Economics of Smoking. In J. Newhouse and A. Cuyler (Eds.), *Handbook of Health Economics*, pp. 1539–1627. New York: North Holland.

Ching, A., T. Erdem, and M. Keane (2011). Learning Models: An Assessment of Progress, Challenges, and New Developments. Working Paper.

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

Darden, M. (2011). *A Dynamic Stochastic Model of Lifetime Smoking Behavior*. Ph. D. thesis, University of North Carolina at Chapel Hill.

DeCicca, P., D. Kenkel, and A. Mathios (2002). Putting out the Fires: Will Higher Taxes Reduce the Onset of Youth Smoking? *Journal of Political Economy 110*(1), 144–169.

DeCicca, P., D. Kenkel, A. Mathios, Y. Shin, and J. Lim (2008). Youth Smoking, Cigarette Prices, and Anti-smoking Sentiment. *Health Economics 17*, 733–749.

Emery, S., M. White, and J. Pierce (2001). Does Cigarette Price Influence Adolescent Experimentation? *Journal of Health Economics 20*, 261–270.

Erdem, T. and M. Keane (1996). Decision-Making under Uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science 15*, 1–20.

Fletcher, J., P. Deb, and J. Sindelar (2009). Tobacco Use, Taxation and Self Control in Adolescence. NBER Working Paper Series, No. 15130.

Gilleskie, D. and K. Strumpf (2005). The Behavioral Dynamics of Youth Smoking. *Journal of Human Resources 40*(4), 822–866.

Glied, S. (2002). Youth Tobacco Control: Reconciling Theory and Empirical Evidence. *Journal of Health Economics 21*, 117–135.

Goel, R. and M. Nelson (2006). The Effectiveness of Anti-smoking Legislation: A Review. *Journal of Economic Surveys 20*(3), 325–355.

Gordon, B. and B. Sun (2009). A Dynamic Structural Model of Addiction, Promotions, and Permanent Price Cuts. Working Paper.

Gruber, J. and B. Koszegi (2001). Is Addiction "Rational"? Theory and Evidence. *Quarterly Journal of Economics 116*(4), 1261–1303.

Hotz, V. and R. Miller (1993). Conditional Choice Probabilities and the Estimation of Dynamic Models. *Review of Economic Studies 60*(3), 497–529.

Laviolette, S. and D. van der Kooy (2004). The Neurobiology of Nicotine Addiction: Bridging the Gap from Molecules to Behavior. *Nature Reviews: Neuroscience 5*, 55–65.

Matsumoto, B. (2015). Estimating Dynamic Discrete Choice Models with Continuous Unobserved Heterogeneity using Conditional Choice Probabilities with an Application to Learning Models. Working Paper.

McFadden, D. and K. Train (2000). Mixed Multinomial Logit Models for Discrete Response. *Journal of Applied Econometrics 15*, 447–470.

Nestler, E. and G. Aghajanain (1997). Molecular and Cellular Basis of Addiction. *Science 278*(5335), 58–63.

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

Orphanides, A. and D. Zervos (1998). Myopia and Addictive Behavior. *The Economic Journal 108*(446), 75–91.

Suranovic, S., R. Goldfarb, and T. Leonard (1999). An Economic Theory of Cigarette Addiction. *Journal of Health Economics 18*, 1–29.

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

Train, K. (2007). A Recursive Estimator for Random Coefficient Models. Working Paper.

Tworek, C., R. Yamaguchi, D. Dloska, S. Emery, D. Barker, G. Giovino, P. O'Malley, and F. Chaloupka (2010). State-level Tobacco Control Policies and Youth Smoking Cessation Measures. *Health Policy 97*, 136–144.

USDHHS (2012). *Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General.* Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

USDHHS (2013). *The Health Consequences of Smoking - 50 years of progress.* Washington D.C.: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

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

A   Data Appendix

Table A1 presents the summary statistics by year for the sample of individuals who are observed in every wave of the survey. The proportion of individuals who smoke increases over the first few waves. The proportion of smokers peaks at around 36% in 2002 and remains in the low 30's for the rest of the sample period.

B   Estimation Appendix

B.1   CCP Representation and Finite Dependence

When the preference shock is GEV, the future value term in the conditional value function has a closed form solution. With type I EV errors, the future value term can be expressed as the one period ahead CCP and conditional value function of any alternative. The closed form expression of the future value term is:

$$\mathbf{E}[\max_j v_{t+1}^j] = log(\sum_{k=1}^{J} e^{v_{t+1}^k}) + e.c. \tag{24}$$

where $e.c.$ is Euler's constant.[29] To express the future value term in terms of the conditional value function and CCP of alternative 1, consider the probability of choosing alternative 1 in period $t+1$:

$$P_{t+1}^1 = \frac{e^{v_{t+1}^1}}{\sum_{k=1}^{J} e^{v_{t+1}^k}} \tag{25}$$

---

[29]Individual subscripts are suppressed for simplicity.

Table A1: Summary statistics by year for individuals observed every period
($N = 5,385$)

	Year					
	1997		1998		1999	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.373	0.484	0.473	0.499	0.532	0.499
Current smoker	0.164	0.370	0.239	0.427	0.279	0.448
# of cigs/day	0.560	2.572	1.274	4.566	1.674	5.137
Age	14.23	1.472	15.87	1.430	16.84	1.439
Employed	0.441	0.497	0.501	0.500	0.522	0.500
Income	243.5	749.5	626.3	1,720	1,189	3,168
Married	0.000	0.019	0.004	0.064	0.013	0.112
	2000		2001		2002	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.579	0.494	0.618	0.486	0.648	0.478
Current smoker	0.316	0.465	0.335	0.472	0.361	0.480
# of cigs/day	2.122	5.616	2.451	6.168	2.703	6.535
Age	17.91	1.435	18.90	1.427	19.90	1.402
Employed	0.611	0.488	0.697	0.459	0.750	0.433
Income	2,191	4,565	3,785	6,281	4,877	7,702
Married	0.026	0.159	0.052	0.222	0.074	0.262
	2003		2004		2005	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.697	0.460	0.668	0.471	0.684	0.465
Current smoker	0.350	0.477	0.361	0.480	0.357	0.479
# of cigs/day	2.936	7.019	2.772	6.349	2.880	6.631
Age	22.85	1.420	20.85	1.424	21.88	1.418
Employed	0.830	0.376	0.785	0.411	0.804	0.397
Income	10,832	13,255	6,327	8,995	8,485	11,807
Married	0.108	0.310	0.143	0.350	0.182	0.386
	2006		2007		2008	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Ever smoked	0.709	0.454	0.717	0.450	0.722	0.448
Current smoker	0.354	0.478	0.343	0.475	0.328	0.470
# of cigs/day	2.818	6.462	2.740	6.488	2.760	6.827
Age	23.81	1.420	24.77	1.432	25.78	1.426
Employed	0.854	0.353	0.870	0.336	0.867	0.340
Income	14,024	15,555	17,402	18,905	20,711	20,735
Married	0.217	0.412	0.248	0.432	0.277	0.447

Now, take the log of both sides:

$$log(P^1_{t+1}) = v^1_{t+1} - log(\sum_{k=1}^{J} e^{v^k_{t+1}}) \tag{26}$$

Substituting the log-sum term into the future value term gives the CCP representation of the future value term:

$$\mathbf{E}_t[\max_j v^j_{t+1}] = u^1_{t+1} + \mathbf{E}V_{t+2} - log(P^1_{t+1}) + e.c. \tag{27}$$

When forming the choice probabilities in the likelihood function, all that matters is the difference in conditional value functions. Finite dependence occurs when two sequences of choices lead to the same future state in expectation. Then when taking the difference in conditional value functions, the remaining future value terms in the CCP representation cancel. The state variables are the individuals beliefs and the prior period's decision. The expectation in the current period of future mean priors is simply the mean of the current period priors for any future sequence of signals (i.e., $\mathbf{E}_t[m_{t+k}] = m_t$, $\forall k$ and $\cup_j \{d^j_{t+1}, \ldots, d^j_{t+k-1}\}$). The variance of the priors only depends on the number and intensity of the signals; the timing of the signals does not matter. So the expected distribution of a future period's beliefs will be the same along any two sequences that generate the same number and intensity of the signals. The other state variable is the prior period's decision, which will be the same as long as the two sequences end with the same alternative. The following table gives the sequences that generate finite dependence.

period	$t-1$	t	$t+1$	$t+2$	$t+3$
sequence 1	0	a_j	0	0	
sequence 2	0	0	a_j	0	
		For any $a_j > 0$			
sequence 1	$a_{j'}$	a_j	1	$a_{j'}$	0
sequence 2	$a_{j'}$	0	$a_{j'}$	a_j	0
		For any $a_j, a_{j'} > 0$			

Consider the simpler case, which is when the individual did not smoke in the prior period. The conditional value function in period t for any alternative $j > 1$ and $j = 1$ are:

$$
\begin{aligned}
v^j(d^1_{t-1} = 1, \Gamma_t) &= u^j(m_t) + \beta * \mathbf{E}[V(\Gamma_{t+1}|d^j_t = 1)] \\
v^1(d^1_{t-1} = 1, \Gamma_t) &= \beta * \mathbf{E}[V(\Gamma_{t+1})|d^1_t = 1]
\end{aligned}
$$

$$\tag{28}$$

The CCP representation of the future value term in the conditional value function for alternative $j > 1$ is:

$$
\begin{aligned}
\mathbf{E}[V(\Gamma_{t+1}|d_t^j = 1)] &= u^1(\mathbf{E}[m_{t+1}|d_t^j = 1]) - \mathbf{E}_m[log(P^1(m_{t+1}|d_t^j = 1))] \\
&+ \beta u^1(\mathbf{E}[m_{t+2}|d_t^j = 1, d_{t+1}^1 = 1]) - \beta \mathbf{E}_m[log(P^1(m_{t+2}|d_t^j = 1, d_{t+1}^1 = 1))] \\
&+ \beta^2 \mathbf{E}[V(\Gamma_{t+3}|d_t^j = 1, d_{t+1}^1 = 1, d_{t+2}^1 = 1)] \quad (29)
\end{aligned}
$$

and the CCP representation of the future value term in the conditional value function for alternative $j = 1$ is:

$$
\begin{aligned}
\mathbf{E}[V(\Gamma_{t+1})|d_t^1 = 1] &= u^j(\mathbf{E}[m_{t+1}|d_t^1 = 1]) - \mathbf{E}_m[log(P^j(m_{t+1}|d_t^1 = 1))] \\
&+ \beta u^1(\mathbf{E}[m_{t+2}|d_t^1 = 1, d_{t+1}^j = 1]) - \beta \mathbf{E}_m[log(P^1(m_{t+2}|d_t^1 = 1, d_{t+1}^j = 1))] \\
&+ \beta^2 \mathbf{E}[V(\Gamma_{t+3}|d_t^1 = 1, d_{t+1}^j = 1, d_{t+2}^1 = 1)] \quad (30)
\end{aligned}
$$

When calculating the choice probability in the likelihood function, all that matters is the difference between these conditional value functions. The $t+3$ expected future value term is the same for the alternative $j > 1$ and $j = 1$ conditional value functions, so it will cancel out in the difference term. All that remains are the flow utilities for periods t, $t + 1$, and $t + 2$ as well as CCPs for periods $t + 1$ and $t + 2$. Note that the CCPs are functions of the mean prior beliefs (m), which depend on the realized value of the signal. If no signal is received, then the CCP can be evaluated using the current period beliefs. If, however, a signal is received, then calculating the expectation requires integrating over possible realizations of the signal. Approximating these integrals numerically adds to the computational burden of the estimation procedure, but the computational requirements are much less than would be needed to fully solve the dynamic learning problem.

An additional advantage of the CCP representation of the value function with finite dependence is that it is not necessary to estimate a separate closing function for the value function in the final period that the individuals are observed, which is the case when solving for the value function using backwards recursion.

The experimentation decision is an optimal stopping problem, which is one of the original class of problems where the CCP representation was applied. The conditional value function of experimenting is:

$$
v_{n,t}^E = u_{n,t}^E + \mathbf{E}_t[V_{n,t}|d_{n,t}^E = 1] \quad (31)
$$

In periods after the individual experiments, he no longer faces an experimentation decision. So, the conditional value function for experimenting does not include a future value term

for future experimentation decisions. The expectation over the value of the consumption decision is over the iid preference shock as well as potential realizations of the value of α. The conditional value function for not experimenting is the discounted expected value of the next period's experimentation decision. The CCP representation of this future value term is:

$$
\begin{aligned}
v_{n,t}^{NE} &= \beta \mathbf{E}_t[V_{n,t+1}^E] \\
&= \beta \mathbf{E}_t[v_{n,t+1}^E - log(P_{n,t+1}^E) + ec] \\
&= \beta \mathbf{E}_t[u_{n,t+1}^E + \mathbf{E}_{t+1}[V_{n,t+1}|d_{n,t+1}^E = 1] - log(P_{n,t+1}^E) + ec]
\end{aligned}
\tag{32}
$$

Note that the CCP representation does not contain a future value term for the experimentation decision for period $t+2$. Now both conditional value functions contain an expected value of the consumption decision in the period that the individual experiments. The expected value of the consumption decision in the conditional value function for experimenting can be expressed as a function of the CCP of not smoking and the conditional value function of not smoking:

$$
\mathbf{E}_t[V_{n,t}|d_{n,t}^E = 1] = \mathbf{E}_t[u_{n,t}^1 + \beta \mathbf{E}_t[V_{n,t+1}|d_{n,t}^1 = 1] - log(P_{n,t}^1) + ec]
\tag{33}
$$

The future value term will cancel with the future expected value of the consumption decision in the conditional value function of not experimenting. Then the difference in the conditional value functions of not experimenting and experimenting is:

$$
v_{n,t}^{NE} - v_{n,t}^E = -u_{n,t}^E + \mathbf{E}[log(P_{n,t}^1)] - ec + \beta(u_{n,t}^E - \mathbf{E}[log(P_{n,t}^E)] + ec)
\tag{34}
$$

The probability that an individual experiments is:

$$
P_{n,t}^E = \frac{1}{1 + exp(v_{n,t}^{NE} - v_{n,t}^E)}
\tag{35}
$$

B.2   Estimation Procedure

This section describes the details of the estimation procedure. The estimation procedure uses the EM algorithm to estimate the parameters that maximize the likelihood function (equation 20). The integrals in the likelihood function are approximated numerically, so the likelihood function becomes a simulated likelihood function in estimation.[30] The procedure

---

[30]When the EM algorithm is used to maximize a simulated expectation (the likelihood being maximized is the expected conditional likelihood), it is called a simulated EM (SEM) algorithm.

begins with initial guesses for the parameters and the CCPs as well as M vectors of draws from the standard normal distribution for each individual, $\{z_n^m\}_{m=1}^M$. These draws are used to form a sample of $N * M$ simulated individuals. Each iteration proceeds according to the following steps:

1. Calculate the value of the unobserved state variables for each individual using the current estimates of the population distribution parameters and the M draws using the following equations and the corresponding elements of the vector z:

$$\theta_n^m = \bar{\theta} + Ch' * z_n^m \tag{36}$$

$$\{\lambda_{n,t}^m\}_{t=1}^T = \sigma_\lambda * z_n^m, \ \{\psi_{n,t}^m\}_{t=1}^T = \sigma_\psi * z_n^m, \text{and } \{\eta_{n,t}\}_{t=1}^T = \sigma_\eta * z_n^m \tag{37}$$

where Ch is the Cholesky decomposition of Σ. These values of the addictive parameters and the noisy component of the signals are used to calculate the individual's prior beliefs for each period.

2. E step, part 1: Use the prior iteration parameter values and CCPs, denoted $\hat{P}$, to update π:

$$\pi(\theta_n^m, \Lambda_n^m) = \frac{\prod_t L_{n,t}(\theta_n^m, \Lambda_n^m, \hat{P})}{\sum_m \prod_t L_{n,t}(\theta_n^m, \Lambda_n^m, \hat{P})} \tag{38}$$

3. E step, part 2: Use the updated values of π to update the CCPs. There are several methods for updating the CCPs. The method used in this paper is to estimate a weighted multinomial logit model of the observed choices on a flexible polynomial of the state variables (both observed and unobserved), where the values of π are the weights. The coefficients from this multinomial logit are used in order to approximate the CCPs at the relevant combinations of state variables in the solution to the individual's problem. This method for updating the CCPs is analogous to least squares value function interpolation. The only heterogeneity in utility from experimentation is in observable characteristics. Therefore, the CCPs for the experimentation decision can be estimated outside of the main estimation routine. Similarly, state transition probabilities that do not depend on the unobserved heterogeneity, or that only depend on the unobserved heterogeneity through the smoking choice, can also be estimated in a first stage.

4. M step: Using the updated CCPs and π, the parameters are updated by maximizing

the simulated log-likelihood function:

$$\tilde{\mathscr{L}}(\gamma, \xi, \sigma_\psi^2, \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma) =$$

$$\sum_n \frac{1}{M} \sum_m \pi(\theta_n^m, \Lambda_n^m) \mathscr{L}(\theta_n^m, \Lambda_n^m, \hat{P}, \gamma, \xi | \sigma_\psi^2 \sigma_\lambda^2, \sigma_\eta^2, \bar{\theta}, \Sigma) \quad (39)$$

The parameters of the population distribution of unobserved heterogeneity can be estimated separately and have a closed form solution (Train, 2007). The updated parameters are simply the weighted mean (for $\bar{\theta}$) and variance (for Σ, σ_λ^2, and σ_η^2) of the values of θ_n^m and Λ_n^m, where the weights are the values of π. The remaining parameters are estimated using simulated maximum likelihood.

These steps are repeated until the parameters converge. The criteria for convergence can either be based on changes in the parameter values or changes in the likelihood function. In practice, there are a wide range of criteria used to determine the convergence of the SEM algorithm. Also, the algorithm may not converge to the global maximum, so to confirm any potential maximum, the algorithm must be rerun using different starting values. The convergence criteria used for preliminary estimation results are that the parameters change by less than one half of one percent, which is the criteria suggested by Train (2007). A feature of the EM algorithm is that the likelihood function weakly increases from one iteration to the next. Performing the full maximization in the M-step yields the largest possible increase in the likelihood but may be computationally intensive. The computational burden is particularly great if an the derivative and Hessian must be approximated using finite differences. In order to reduce the computational burden, I use an alternative version of the EM algorithm. This alternative version of the EM algorithm replaces the full optimization of the M-step, which gives the greatest possible likelihood improvement, with a procedure that is simply guaranteed to improve the likelihood function. This version of the EM algorithm is called a Generalized EM algorithm (GEM) and is commonly implemented by replacing the maximization in the M-step with a single Newton-Raphson iteration. GEM algorithms share similar convergence properties as the EM algorithm, although they converge at a slower rate. Even though GEM algorithms require more iterations to converge, each iteration requires much fewer evaluations of the likelihood function.

This estimation procedure is still computationally demanding, although standard procedures would likely be infeasible.[31] Evaluating the likelihood for a single simulated individual only takes a fraction of a second, but with $N * M$ simulated individuals, a single evaluation

---

[31]Using Simulated Maximum Likelihood to estimate the version of the model without learning took around 12 hours for a modest number of draws.

of the likelihood function can take hours.[32]

---

[32]There are two factors that influence the number of calculations needed to evaluate the likelihood function for a single simulated individual. The most significant determinant of the number of necessary calculations is the number of draws used to approximate the future value terms. Increasing the number of draws by a given factor increases estimation time by nearly the same factor. The second determinant of estimation time is the number of terms used in the interpolation of the CCPs. Increasing the number of terms by using a higher order polynomial approximation increases the calculations needed to evaluate the likelihood function. However, the most significant effect of increasing the number of interpolation terms comes in the increase in the time it takes to update the CCPs.