

The Behavioral Dynamics of Youth Smoking

Donna B. Gilleskie and Koleman S. Strumpf

Abstract

Individual smoking behavior persists over time, but is this repeated behavior attributed to past use or individual heterogeneity? Using longitudinal data on teens from all 50 United States from 1988 to 1992, we find a significant causal role for endogenous past cigarette consumption even after controlling extensively for observed and unobserved heterogeneity. We also find measurable evidence of different sensitivities to cigarette price depending on past use. These two findings suggest that a cigarette price increase will have a larger aggregate effect in the long run than in the short run as more individuals accumulate in the price-sensitive non-smoking group.

Donna B. Gilleskie and Koleman S. Strumpf are associate professors of economics at the University of North Carolina at Chapel Hill. Dan Benjamin, David Blau, Bill Evans, Michael Grossman, Thomas Mroz, Rosalie Lliccardo Pacula, Jeanne Ringel, and seminar participants at the 2000 ASSA/AEA Annual Meetings, the Triangle Applied Microeconomics Conference, the UNC-CH Health Economics Workshop, Brown University, the Federal Trade Commission, and the Kenan-Flagler Business School provided valuable comments on the paper. We thank the Department of Education for access to the restricted use National Education Longitudinal Survey (NELS) data. Researchers interested in replicating this work should contact the authors for information about obtaining restricted use NELS data. Financial support from the National Institutes for Health (grant #1RO1HD42256-01) is gratefully acknowledged.

I. Introduction

Smoking is one of the primary public health concerns in the U.S. Many of the deleterious effects of smoking are associated with long-term use. This suggests that efforts to alter smoking behavior, whether to prevent initiation, encourage cessation, or reduce relapse, would benefit from an understanding of smoking dynamics -- the dependence of current smoking choices on past smoking behavior. It is well-known that individual smoking behavior is quite persistent. Individuals who smoked in the past are several times more likely to smoke today relative to previous non-smokers. For example, using data from the National Education Longitudinal Study, 75 percent of youth aged 14 to 18 who smoked in the past continue to smoke at the survey date; only 10 percent of non-smokers in this group begin smoking. Previous smoking intensity also seems to matter: 31 percent of light smokers quit smoking while less than 13 percent of heavy smokers quit. There is also movement toward higher levels of smoking in the future among those who were moderate or heavy smokers in the past.¹

In this paper we seek to measure the significance of previous smoking behavior as a causal contributor to smoking persistence. We refer to this possibility as *state dependence*. For example, previous smoking consumption might alter the current utility of different smoking alternatives, it might reflect different sensitivities to price, or it might impact health which influences the expected value of future smoking alternatives. What we call state dependence is often referred to as addiction.² However, it is also possible that smoking persistence stems largely from *individual heterogeneity*. Individual differences, both observed and unobserved, may lead some to be prone to smoke and others to abstain. For example, health-conscious individuals might never smoke while present-oriented individuals always smoke. These two explanations for

persistent behavior are laid out and formalized in Heckman (1981) and are now standard terminology in the economics literature. Our goal is to determine the importance of state dependence for smoking behavior after extensively controlling for individual heterogeneity.

Understanding the explanations for smoking persistence has important implications for public policy. Government taxes continue to be the main policy tool used to lower smoking rates, and the average state plus federal tax rate has doubled between 1998 and 2003. The effectiveness of these tax increases in reducing smoking rates over the long run depends crucially on the source of smoking persistence. Suppose one's smoking history does alter current choices, and that smokers are less price sensitive than non-smokers (a point we document later). Taxes are quite powerful in this case: by discouraging some individuals from smoking today, they will increase the proportion of relatively more price sensitive non-smokers in the future. This compositional change produces a multiplier effect as future taxes become more effective at reducing aggregate smoking rates. Alternatively, suppose smoking persistence is largely due to individual heterogeneity. In this case, taxes are not very powerful. Individual factors, rather than prices, are the primary determinant of current and long-run smoking decisions.

We disentangle the roles of previous smoking behavior and individual heterogeneity (that is, observed and unobserved individual characteristics) in smoking decisions using detailed panel data on U.S. teenagers. The National Education Longitudinal Survey (NELS) follows a representative sample of 8th graders in two year waves beginning in 1988 and is described in detail in section II. There are three reasons why we focus on teenagers and use the NELS data. First, teens are the focus of many policy discussions. NELS is rather unique because it tracks a wider population than other youth surveys. It includes teens from every state, which allows us to

utilize price variation in all 50 states, and follows school drop-outs. Second, a majority of smokers first try cigarettes when they are teenagers and youth smoking often persists into adulthood. While many other surveys have limited information about smoking initiation, which can occur years prior to the first interview and is often only known from retrospective self-reports, we accurately observe when smoking begins in our data. This allows us to avoid potentially serious sample selection problems or the need to estimate a simplified reduced-form initial condition equation representing previous smoking behavior. Third, studying teens provides insights into the state dependence of smoking decisions. We take advantage of the multiple observations per person to account for individual heterogeneity and to determine the extent to which previous smoking behavior causally influences current choices. Finding a significant role for past smoking behavior among teens would suggest that smoking, even the first short duration experience, can rapidly modify behavior.

We describe our empirical methods in section III. We jointly estimate a set of equations that capture smoking participation and conditional demand. The empirical model allows previous smokers and non-smokers to have different tastes for smoking and different price sensitivities. Tastes are distinguished by observed characteristics-- which include time-varying information on individual demographics, parents, schools, and the community-- as well as by unobserved (to the econometrician) individual heterogeneity that is treated in a general semi-parametric fashion. Thus the model allows changes or persistence in behavior to be driven by individual heterogeneity as well as by previous smoking choices. We use the discrete factor random effects methodology introduced by Heckman and Singer (1984) and extended by Mroz (1999). This method allows correlation in unobservables across multiple equations without imposing a distributional form and has been used in a wide variety of empirical applications including health

(Goldman 1995; Cutler 1995), welfare participation (Hoynes 1996), child care (Blau and Hagy 1998; Hu 1999), disability insurance (Kreider and Riphahn, 2000), and program evaluation (Angeles et al. 1998).

Our estimates are presented in section IV. The results depend crucially on two modeling features: whether individual smoking history is treated as exogenous or endogenous and whether unobserved individual heterogeneity is omitted or included. We find that previous smoking behavior has a causal effect, even after allowing for extensive individual differences. For example, after controlling for both observed and unobserved heterogeneity, smoking lightly in 8th grade (as opposed to not smoking) doubles the probability of choosing to smoke two years later, and smoking participation rates are one-fifth higher four years later. The strong role of previous use is impressive given that the teens in our sample have been smoking for no more than a few years. Alternatively, if we do not control for the unobserved heterogeneity associated with persistent behavior, this effect is biased upward and overstates the role of previous smoking behavior. For example, smoking lightly in 8th grade quadruples participation rates two years later and more than doubles them four years later. We also find that previous non-smokers are more price sensitive than previous smokers with participation price elasticities of -0.48 and -0.16, respectively. Our preferred results suggest that prices serve both to reduce consumption of smokers and also to reduce the number of new smokers. Alternatively when previous behavior is presumed exogenous, prices only affect consumption among smokers. Our modeling extensions allow us to simulate the long-run effects of policy changes as the population composition of smokers and non-smokers is altered. In total, a dollar increase in cigarette prices reduces age 18 smoking participation predictions by a third more when unobserved individual heterogeneity and behavior modification are taken into account than when they are ignored.

Our paper contributes to the large economics literature on cigarette consumption that is reviewed in Chaloupka and Warner (2001). To place our methods and results in context, we briefly mention here some of the work upon which our paper builds. We focus on papers that use individual-level panel data, as opposed to aggregate time series or repeated cross sections, since multiple observations on individuals allow a researcher to disentangle state dependence from individual heterogeneity. We first discuss efforts in the relevant literature to incorporate previous smoking behavior and then the addition of future consumption and prices.

To account for the endogeneity of past smoking behavior, Chaloupka (1991) instruments lagged consumption using several lags of prices and taxes in a two-stage approach. The recently developed estimation strategies that we employ lead to significant efficiency improvements over the instrumental variable approach and its potential weak instrument critique.³ Other research that explicitly considers previous smoking behavior focuses only on the decisions to start or stop smoking.⁴ DeCicca et al. (2002) also use NELS data to determine the immediate effects of prices, non-tax policies, and schooling on the single transition from non-smoking to smoking among teens. Tauras and Chaloupka (2001) estimate smoking cessation hazards, while Douglas (1998) separately estimates smoking initiation and quitting decisions. We unify the approaches in these papers by jointly modeling all decisions (that is, initiation, continuation, cessation, and relapse) and by considering the general impact of previous consumption, both participation and intensity, on current smoking decisions. This allows us to more broadly understand the role of previous smoking behavior, since we study participation of both former non-smokers and former smokers. Continuing to follow all individuals, even those changing their behavior, permits us to simulate dynamic responses to policy changes. Labeaga (1999) has a focus similar to our work. He separately estimates dynamic participation and conditional intensity equations and accounts

for unobserved heterogeneity. We disaggregate his results which rely on household, rather than individual, data and explain expenditures on all tobacco-related products, rather than consumption of cigarettes.

In addition to dependence on previous smoking decisions, theoretical models of smoking behavior suggest that cigarette consumption today also depends on future consumption and future cigarette prices. The rational addiction model (Becker and Murphy, 1988) implies that individuals respond to future events. Empirical tests of such rational behavior have found that leads of consumption and leads of price are statistically significant predictors of current smoking behavior.⁵ We do not include these future variables in our specifications. The empirical rational addiction model assumes these variables are known with perfect foresight. We contend that current behavior is a function of the entire distribution of future values for all consumption options and prices, not just those that are observed. These distributions are a function of currently available information, potentially including all lags of behavior and prices. Our empirical strategy controls for all of these lags. However, we cannot test whether future variables matter in a non-structural framework. Our empirical model does not explicitly estimate the underlying preference parameters or how individual expectations are formed, and so it is mute about whether there is forward-looking behavior.⁶ Hence, one may interpret our model as myopic in the accepted rational addiction framework. We believe that a true test of both the addictive nature of smoking behavior and the rational aspect of behavior requires solution of the individual's forward-looking optimization problem and estimation of the primitives of the model. We do not intend to estimate these structural parameters here, and reserve that for future work.

Our work contributes to the smoking demand literature in two ways. First, we provide empirical evidence of state dependence -- the statistically significant effect of endogenous previous smoking consumption on current smoking decisions -- within a unified framework that accounts for unobserved heterogeneity and allows for all possible transitions in behavior. Second, our work refines long-run simulations of smoking behavior because our predictions reflect the changing composition of smokers and non-smokers, and the subsequent differences in price sensitivities portrayed by these different groups. This approach improves upon long-run predictions that ignore the role of previous behavior, presume it is exogenous, or loosely proxy for it using instruments. Failing to model previous smoking choices precludes measuring the compositional changes in the population over time and hence significantly under-estimates the long-run price response.⁷

II. Description of the Data

Despite a great interest in the U.S. in understanding youth smoking behavior, few nationally representative data sets are available that chronicle the behavior of the same children over multiple periods of time. The National Education Longitudinal Study of 1988 (NELS) is one exception. NELS, a continuing study sponsored by the U.S. Department of Education's National Center for Education Statistics, began in 1988 with the specific purpose of collecting information on educational, vocational, and personal development of a nationally representative sample of 8th graders as they transition from middle school into high school, through high school, and into postsecondary institutions and the work force. Approximately 24,500 8th graders in more than 1,000 public and private schools in all 50 states participated in the first wave of the study. In addition to the student questionnaires, supplementary questionnaires were administered to the

students' parents, teachers, and school principals and provide a wealth of information on the early social and academic environment of the students. Through special agreement with the U.S. Department of Education, we obtained access to restricted-use NELS data that include geographic information. These data, supplemented with state-level price and tax data (Tobacco Institute 1997) and measures of inflation, allow us to determine the appropriate real cigarette price each individual faces in every year.

The first follow-up, administered in the spring of 1990, includes responses from approximately 17,500 of the students from the 1988 base year interview, while the second follow-up, administered in the spring of 1992, includes approximately 16,500 students from the original cohort.⁸ One of the many unique features of the NELS data is that youth who leave high school prior to graduation continue to be interviewed throughout the longitudinal study and are asked the same questions pertaining to smoking behavior. It is therefore possible to examine the smoking behavior of all youth, including those not represented in other national school-based surveys such as Monitoring the Future. The NELS data contain information on the student's background, upbringing, early family environment, early school environment, and other behaviors. It provides many variables that have been found to be significant risk factors for smoking such as school performance, religious affiliation, family structure and living arrangement, and parental education. Since parents are surveyed in the base year and second follow-up, it is possible to obtain time-varying information on family background and socioeconomic characteristics that the student would not be as informed about. In the first and second follow-up, school principals and teachers continue to be surveyed, making it possible to control for important school environmental characteristics as well.

We model the behavior of youths who are observed in each year (1988, 1990, and 1992) of the survey; we do not model attrition from the full sample. We keep only those youth who were on grade during the sample period or who were permanent dropouts (12,954 youth). We are forced to drop 2237 kids for whom smoking behavior is unobserved. Because prices differ by state, another 270 are dropped if we cannot identify the state in which they live or go to school, and 196 are dropped if they do not reside in the same state in all three waves.⁹ After deleting 18 individuals for whom other important variables are missing, our sample consists of three years of observations on 4755 males and 5478 females (30,699 person-year observations).

Information on smoking behavior is collected in each wave of the survey. In each year, youths are asked, “How many cigarettes do you currently smoke in a day?” Responses are limited to the following categories: do not smoke, smoke less than one cigarette a day, smoke one to five cigarettes, smoke about a half pack (6-10), smoke more than half a pack but less than two packs (11-39), and smoke two packs or more (40+). Table 1 reveals that 4.6 percent of youths reported smoking in the 8th grade (1988), while 22.7 percent reported smoking four years later. The dramatic increase in smoking rates is not surprising given that smoking initiation typically occurs during the late teens. We also form indicators of the quantity smoked conditional on being a smoker.¹⁰ Table 1 indicates that smokers appear to be smoking larger quantities over time.

Table 1 also illustrates the pattern of smoking behavior over time by conditioning current smoking choices on past smoking consumption.¹¹ The probability of quitting if one was a moderate (6-10 cigarettes/day) or heavy (11+ cigarettes/day) smoker decreases dramatically as these youth age. Smokers who continue to smoke also tend to migrate to higher levels of

smoking over time. For example, one-half of previous heavy smokers who continue to smoke continue smoking heavily in 1990 while almost three-fourths of them continue to smoke heavily in 1992. Additionally, the proportion of non-smoking youth who remain non-smokers is fairly constant after 1988, and the proportion of light smokers (1-5 cigarettes/day) who quit smoking is not declining.

The rich set of variables used to explain variations in individual smoking behavior are summarized in Table 2. These include both stationary and time-varying variables describing characteristics of the student, his family, and his school. In general (although not shown in the table), adolescent smokers are older white youth with lower test scores and socioeconomic status than non-smokers. They are more likely to have older siblings, to have siblings who dropped out of school, to have one parent absent from the home, and to report no religion. Additionally, our data contain significant variation in youth smoking rates and cigarette prices across states and within states over time. Smoking rates by state vary by over 10 percentage points in 1988 and by as much as 17 percentage points by 1992. There is about a one dollar range in real cigarette prices per pack over the 1988-1992 period, with prices averaging \$1.41.¹²

III. Empirical Model

The theoretical model of Becker and Murphy (1988) characterizes the framework on which most empirical analyses of smoking are based. Recognizing the potential dynamics of smoking behavior, they relax the additive separability assumption common in demand analyses. In their rational addiction model they suggest that current period utility of smoking alternatives is influenced by previous smoking behavior implying that the first order demand equation contains the lagged smoking decision. Individuals are also forward looking and consider the influence of

today's smoking choice on future smoking behavior. Hence, the current period demand equation, derived from maximization of discounted lifetime utility subject to per-period budget constraints, contains expected future smoking consumption. Let C_t represent smoking consumption in period t . Then (presuming a quadratic utility function), the linear demand equation for individual i at time t is

$$(1) \quad C_{it} = \alpha_0 + \alpha_1 P_{it} + \alpha_2 C_{it-1} + \alpha_3 C_{it+1}^* + \alpha_4 P_{it-1} + \alpha_5 P_{it+1}^* + \alpha_6 X_{it} + \varepsilon_{it}$$

where P_t is the period t price of cigarettes, X_t denotes contemporaneous exogenous characteristics, the α -terms are non-linear combinations of underlying structural preference parameters, and “*” indicates an expected future value.

The empirical formulation above suggests that lags and leads of consumption influence current behavior. Observations on the same individual over time would allow estimation of the impact of past smoking decisions on current smoking behavior. Because longitudinal data is often unavailable, in the empirical literature lagged and lead prices serve as proxies for lagged and lead behavior in equation (1)¹³ In this context, a statistically significant effect of lagged prices on current consumption is often taken as evidence of state dependence.¹⁴ Alternatively, we investigate the dependence of current smoking decisions on endogenous past smoking behavior by using observations on the same individuals over time and allowing for correlation through permanent individual unobservables.

While future consumption influences the value of current smoking alternatives in equation (1), it is the *optimal expected* value based on currently available information rather than the actual *ex post* value that matters. Optimal future consumption depends on the distribution of

future unobserved preferences. Similarly, while expectations of future cigarette prices are important to today's smoking decision, future prices are not known with perfect foresight, (especially in our biannual data). Rather, they are forecasted as a function of current information. Gilleskie and Strumpf (2000) incorporate these concerns about expectations and provide an approximation to the individual's dynamic optimal demand to which we now turn. While the empirical specification is based on forward-looking decision making, it resembles a myopic model since the expectations process is not estimated.

Our specification of current period cigarette consumption is

$$(2) \quad C_{it} = \gamma_0 + \gamma_1 P_{st} + \gamma_2 C_{it-1} + \gamma_3 C_{it-1} P_{st} + \gamma_4 X_{it} + \sigma_s + \sigma_t + \varepsilon_{it}$$

where we include an interaction of previous smoking status with current prices (in individual i 's state of residence s) to allow for different price elasticities among former smokers and non-smokers, and we exclude future consumption and future prices. The specification also allows for state and time fixed-effects (σ_s and σ_t , respectively). We further decompose the error term ε_{it} into two components to reflect permanent individual unobservables and serially-uncorrelated errors. That is,

$$(3) \quad \varepsilon_{it} = \rho\mu + u_{it}$$

where μ (an unobserved factor) is the permanent unobservable individual heterogeneity and ρ (the factor loading) measures its effect. Analogously, one could consider an individual fixed effect, $\sigma_i = \rho\mu$ where σ_i varies by individual. Instead, as discussed in more detail below, we consider μ to be a random effect estimated by a discrete distribution of mass points. The error u_{it} is an idiosyncratic, i.i.d. component. The model is dynamic because of the testable assumption

that smoking behavior and price sensitivity today depend on the smoking state occupied in the previous period. To recap, the model has two important implications that are the focus of our empirical work: previous smoking history influences current smoking decisions and cigarette price sensitivity may be different for individuals with different histories of smoking.

One additional adaptation is needed given the smoking outcome available in our data. Our data contains a polychotomous outcome categorizing individuals as non-smokers, light smokers, moderate smokers, or heavy smokers in each period. For reasons we discuss below, current smoking behavior is modeled in two parts: (i) no smoking vs. smoking, and (ii) the quantity of cigarettes smoked conditional on smoking. Undoubtedly there is correlation between the error terms, which reflect unobserved individual differences, associated with these smoking behaviors. For example, individuals who heavily discount the future may be more likely to smoke and, conditional on smoking, to smoke more. We allow such permanent individual heterogeneity (μ) to influence both equations. Assuming that the idiosyncratic components of the error terms (u_i) are independently Extreme Value distributed, the smoking and conditional smoking intensity probabilities (logit and multinomial logit, respectively) are:

$$(4) \quad p(C_{it} = 0 \mid \mu) = \frac{\exp(\delta_0 + \delta_1 P_{st} + \delta_2 C_{it-1} + \delta_3 C_{it-1} P_{st-1} + \delta_4 X_{it} + \sigma_{s1} + \sigma_{t1} + \rho_1 \mu)}{1 + \exp(\delta_0 + \delta_1 P_{st} + \delta_2 C_{it-1} + \delta_3 C_{it-1} P_{st-1} + \delta_4 X_{it} + \sigma_{s1} + \sigma_{t1} + \rho_1 \mu)}$$

$$(5) \quad p(C_{it} = j \mid C_{it} > 0, \mu) = \frac{\exp(\eta_{0j} + \eta_{1j} P_{st} + \eta_{2j} C_{it-1} + \eta_{3j} C_{it-1} P_{st-1} + \eta_{4j} X_{it} + \sigma_{s2} + \sigma_{t2} + \rho_{2j} \mu)}{\sum_{j'=1}^J \exp(\eta_{0j'} + \eta_{1j'} P_{st} + \eta_{2j'} C_{it-1} + \eta_{3j'} C_{it-1} P_{st-1} + \eta_{4j'} X_{it} + \sigma_{s2} + \sigma_{t2} + \rho_{2j'} \mu)}$$

where $j = 1, \dots, J$ indicates increasing levels of smoking intensity. This two-equation specification is preferred over a single ordered logit model because it allows for different

sensitivities to prices (and other variables) by smoking outcome. More formally, we can reject the ordered logit assumption of common coefficients across outcomes using a likelihood ratio test. A discussion of alternative model specifications is presented in subsection IV.A.¹⁵

The dynamic equations of our empirical model are estimated jointly and are linked by dependence on the common individual unobservables (μ). We follow Heckman and Singer (1984) and Mroz (1999) and treat the unobservables as random effects that are integrated out of the model. That is, we assume a discrete distribution of μ and estimate the points of support of the distribution (μ_1, \dots, μ_M where $\mu_m \in [0,1]$), the probability weights on each point of support ($\theta_1, \dots, \theta_M$ where $\sum_{m=1}^M \theta_m = 1$), and the factor loadings in each equation ($\rho_1, \rho_{21}, \dots, \rho_{2J}$), with the appropriate normalizations. This procedure addresses the joint endogeneity of outcomes arising from common unobserved factors, but imposes no distributional assumption such as joint normality on the unobserved factors. Mroz shows there are strong econometric advantages to using this approach rather than parametric random effects techniques. Also, estimation of the random effects as described above improves efficiency over a model of individual fixed effects which would introduce thousands of additional parameters to be estimated. Moreover, individual fixed effects models yield biased parameter estimates unless there are several observations per individual unit (Hsiao 1986; Nickell 1981). The system of equations is estimated using full information maximum likelihood techniques.

Several features of our model and data provide statistical identification in our jointly estimated system of smoking equations, both participation and conditional intensity. Because the system allows for dependence on previous smoking behavior and is estimated using a panel of individual observations, lags of the endogenous variables enter as explanatory variables. This

feature provides a major source of identification. Since we are modeling (estimating) this endogenous previous behavior, previous exogenous covariates such as cigarette prices serve as implicit instrumental variables for the lagged endogenous variables. In fact, all exogenous variables from the previous periods help identify the system since they directly influence past, but not current, behavior. The key is to have sufficient sources of variation in each period. While we use several exogenous time-varying variables, prices alone achieve identification. Put differently, the behavior of two identical individuals (in terms of both their time-invariant and time-varying individual characteristics) who live in different states is identified by the entire observed *history* of cigarette prices/taxes within each state.¹⁶

The behaviors are further identified from the rich set of covariates, many of which vary over time. We exclude specific individual and school variables in particular equations to more precisely estimate the separate behaviors. Notice that these are over-identifying restrictions, since the earlier dynamic argument is sufficient for identification. In the conditional smoking intensity equation, age within the cohort and population density around the school are omitted. Using likelihood ratio tests to compare the empirical specification of our model to ones that include these variables, we cannot reject the null hypotheses that the coefficients on the variables are zero. Cigarette availability is the basis of our argument for excluding these variables. That is, older teens within a cohort are more likely to have access to cigarettes through illegal purchases; they are less likely to be prohibited from buying because they look older. But, conditional on purchase, age is unlikely to influence the number of cigarettes smoked. Similarly, population density is likely related to the number of nearby stores selling cigarettes and hence to cigarette access, but should have little effect on smoking intensity. The parameter estimates below are consistent with these interpretations.

IV. Estimation Results

A. Comparative Results

Before discussing the results from our preferred model that explicitly captures the dynamic influences of one's smoking history, we discuss estimates obtained from typical specifications found in the literature. We then make several additions to the simple model, some of which have been included in the relevant literature, which help motivate our full model. In each case we estimate a logit model of smoking participation and a multinomial logit model of conditional smoking intensity. We use the estimates from each specification to generate behavioral responses to price changes. In the interest of brevity we omit the parameter estimates and summarize the resulting behavioral changes in Table 3 when cigarette prices are increased by \$0.25 in all years of the sample.¹⁷ We follow the literature and present price elasticities and indicate their statistical significance based on bootstrapped standard errors.

The top left panel of Table 3 displays results from a typical model estimated in the literature that includes few covariates, no observations on dropouts, and no state fixed effects (which we refer to as Model 1). The first column shows there are large reductions in smoking participation following the price increase with an overall price elasticity of -0.87. The remaining three columns of the panel show that, conditional on smoking, the price increase noticeably increases light smoking and reduces moderate and heavy smoking intensities. Also, the participation price elasticity falls over time. The reduced price sensitivity of NELS youth as they grow older could stem from behavioral differences related to the aging process, some national trend in smoking behavior, or differences in the composition of smokers.

We take advantage of the details in NELS and add to Model 1 additional individual, parental, and school characteristics and include dropouts in the sample (Model 2). At this point, however, we do not include variables measuring one's smoking history. These additions slightly reduce the participation response (the elasticity is -0.78) and increase the shift to light consumption among smokers. These results indicate that using a richer set of observable exogenous variables provides better estimates of the price effect by purging it of some omitted variable bias. This issue is a particularly important problem for the older literature that was often based on sparse sets of covariates, aggregated individual observations, or retrospective data that may not include time-varying information on exogenous characteristics.

Adding state fixed effects to the specification dramatically changes the results (Model 3). Prices now have a much smaller and statistically insignificant influence on smoking participation (the elasticity is -0.04) while there is a larger but generally insignificant response on conditional smoking behavior. This suggests that there are state-specific factors that influence both individual smoking behavior and cigarette price variation. This is consistent, for example, with an anti-smoking sentiment that boosts prices by increasing state cigarette taxes and independently reduces smoking rates. Failure to model such state specific differences incorrectly attributes the negative effect of anti-smoking sentiment to prices. The more negative price elasticity in Model 2 relative to Model 3 is evidence of a spurious negative correlation between prices and smoking propensity.

Finally, we investigate the role of previous smoking behavior by including exogenous indicators of previous use and previous intensity (Model 4).¹⁸ While the overall behavioral change remains roughly the same as Model 3, Model 4 allows us to explore how previous

smoking status influences price sensitivity. This model suggests that a price increase has a slightly negative, but statistically insignificant, influence on previous non-smokers' participation, but significantly *increases* previous smokers' participation. It must be emphasized, however, that these price sensitivities may be biased since the endogeneity of previous use is not modeled. The estimates could reflect an overestimate of smoking persistence, since there is no control for permanent unobserved heterogeneity. For example, the price sensitivity of previous non-smokers will be underestimated if the empirical model overstates their likelihood of continuing to abstain.

Our preliminary results are roughly consistent with previous studies of teen smoking behavior.¹⁹ DeCicca et al. (2002), who also use NELS data, find that prices and taxes have a statistically insignificant and sometimes positive effect on smoking initiation during high school. They also find that the price elasticity of smoking participation becomes less negative as the cohort ages. In contrast, Gruber (2001) finds prices have a statistically significant negative effect on smoking participation and that this effect is more marked for older teens. (Gilleskie and Strumpf (2000) help reconcile these differences.)

We examined the sensitivity of our results to a variety of changes in specification and data construction (and these results are available upon request). First, we used state taxes in place of cigarette prices to control for the simultaneity of prices and aggregate demand for cigarettes. The simulated behavioral changes are virtually identical for smoking participation and conditional smoking intensity, though there is a slightly greater reduction in heavy smoking under taxes than prices. Second, we checked whether the results are sensitive to the exact timing of prices and taxes since NELS individuals are surveyed on different dates. We considered several monthly dates for taxes (which are available on a day-by-day basis) and several interpolated monthly

dates and averages for prices (which are surveyed once a year). Shifting these dates did not alter the results significantly. Third, we estimated the smoking participation and intensity equations separately by year and by previous smoking status. The price sensitivities were similar, and thus we report results from the model with fewer estimated parameters. Fourth, we investigated whether price has a non-linear effect on smoking decisions. It is not possible to reject a null hypothesis that price squared, and its interactions with previous behavior indicators, should be excluded.

Before proceeding with our preferred method of modeling the endogeneity of previous smoking behavior, we explored the importance of this endogeneity. A standard approach is to instrument for the endogenous regressor and to perform a Hausman test. We focus on the smoking participation logits of Model 4 and use a two-stage procedure to form predicted previous smoking behavior.²⁰ Using the specification in Model 4 with a fitted value for previous smoking participation, the null of smoking exogeneity can be rejected at the 99 percent confidence level. These results suggest that smoking endogeneity is important; after instrumenting, previous behavior plays a smaller role in smoking decisions. This approach, however, results in a serious efficiency loss: standard errors on the smoking history covariates increase by a factor of five. Intuitively, the two-stage approach ignores the substantial variation in previous smoking behavior that is not captured by the instruments. Similarly, we re-estimated Model 4 with individual fixed effects. While the estimates are comparable, the standard errors increase by a factor of two.

Because the standard approaches give imprecise estimates, we now turn to results from our preferred empirical approach that allows for state dependence and formally models the unobserved heterogeneity that influences smoking decisions over time.

B. Results from the Preferred Model

Having motivated the importance of including both observed and unobserved individual heterogeneity, allowing for unobserved state differences, and modeling previous behavior, we now examine results from a model that incorporates all of these concerns (Model 5). We begin by discussing the effect of previous behavior on smoking decisions when permanent individual unobserved heterogeneity is modeled. Because subgroups have different price sensitivities, we continue to interact price with the smoking history variables and year indicators. We then discuss differences in the immediate and long-run effects of previous smoking behavior using this preferred model.

If the smoking history variables reflect part of the permanent unobservable differences among youth, then the estimated coefficients on these behavioral variables and their price interactions are biased in Model 4 where previous behavior is assumed to be exogenous. Model 5 controls for these unobserved differences, and the parameter estimates for all variables are listed in Table 4.²¹ There are two main differences relative to the preliminary models. First, the effect of previous behavior on current smoking decisions is mitigated. For example, Model 4 suggests a 38 percentage point (4.75 times) increase in the predicted probability of smoking if an individual smoked 6-10 cigarettes in the previous period relative to not smoking, while Model 5 has only an eight percentage point (0.87 times) increase. These substantial differences can be attributed to upwardly biased estimates in Model 4. While unobserved differences may drive certain

individuals to smoke or not smoke in every period, Model 4 incorrectly incorporates this effect into the coefficient on previous behavior. When this endogeneity of previous smoking behavior is modeled, smoking history has a smaller, but still statistically significant, effect on current behavior. Hence, previous behavior matters and persistence cannot be fully explained by unobserved heterogeneity.

The second major change under endogenous previous behavior is an increase in the estimated price sensitivity. The top panel of Table 5 indicates a statistically significant participation price elasticity of -0.41 compared to essentially no sensitivity measured in Model 4. Elasticities by year also become more negative (the price elasticity in 1992 is not statistically different from zero).²² The conditional smoking intensity price elasticities also change, with noticeably more negative values for moderate and heavy smoking relative to the model with no heterogeneity. Modeling the heterogeneity allows us to control for intense unobserved tastes and therefore better measure price sensitivity. That is, some individuals will smoke or not smoke regardless of the price. Model 4 does not control for such heterogeneity, and so it estimates that all individuals are relatively price insensitive. This difference is particularly noticeable in the participation price elasticity of previous non-smokers, which is -0.47 (and statistically significant) in Model 5 and only -0.05 (and statistically insignificant) in Model 4.

We can now predict the behavior of individuals over time. Two routes are possible: we can use the observed history of individuals each period to simulate behavior or we can update behavior each period using simulated outcomes from the predictions. The former provides an estimate of the *immediate* response to price changes (as in Table 3 and the top panel of Table 5), since previous behavior is not updated to reflect changes in smoking decisions under the new

prices. The latter provides the *long-term* response and appropriately accounts for changes in past behavior resulting from price changes over time. This calculation requires simulating behavior of the sample in 1988 and updating their smoking histories using the simulated behavior. Subsequently, behavior in 1990 is simulated using simulated 1988 behavior and histories are updated. Finally, we simulate behavior for 1992 using the updated histories.²³ This updating procedure leads to different predictions because previous behavior influences current decisions as well as current price sensitivity. These changes now reflect *long-run* responses as opposed to *immediate* responses to price changes.

The bottom panel of Table 5 describes the changes in behavior and elasticities when we use the updating procedure. Relative to not updating (top panel of Table 5), the participation elasticity is slightly more negative (-0.46 vs. -0.41) as is this price elasticity in 1990 and 1992.²⁴ The greatest changes are among previous smokers who now have a much smaller and negative, but still insignificant, price sensitivity (-0.16 vs. 0.42). This difference has an intuitive explanation. The updating calculation allows endogenous previous smoking behavior to reflect past price changes, while the earlier simulation simply used the observed previous smoking history. A price increase in the initial periods, when price sensitivity is higher, is particularly effective at deterring smoking. Some individuals who first took up smoking at this time will no longer do so. This in turn makes them less likely to smoke in later periods. Thus, the effect of the updating calculation is reflected most notably in the behavior of previous smokers, individuals who smoked in earlier periods in the data. Previous non-smokers continue not to smoke under higher prices. We also observe larger negative price sensitivities in the conditional level of smoking. In particular, using the updating procedure we observe more movement away from

heavy and moderate smoking levels when prices increase. Overall, the total price elasticity in our preferred model is -1.11 compared to -0.51 in Model 4 (see footnote 19).

Despite the role of unobserved heterogeneity in explaining smoking patterns, these simulations demonstrate that price changes can modify behavior and influence subsequent smoking decisions of youth. Prices have a dynamic effect as they prevent smoking today which in turn makes teens more price sensitive and less likely to smoke in future periods. We suspect that long run price sensitivities will be even larger when horizons of longer than four years are considered.

C. Policy Simulations

It is important to see whether our refined estimates imply economically meaningful effects, both in terms of responses to policy changes as well as the role of state dependence. We show the importance of modeling unobserved individual heterogeneity and accounting for the behavior modification associated with previous price changes by simulating the smoking participation predictions of Models 3, 4, and 5. While we do not estimate the policy-invariant parameters of the individual's optimization problem, the policies that we simulate exist within the available data. That is, we simulate price increases that fall within the price variation in the data and we simulate forced previous behavior where all smoking choices are observed in the data.

Table 6 displays the proportion of individuals who are predicted to smoke at the unchanged original prices in the data and under various permanent price increases. The overall smoking participation rates are larger in Models 3 and 4 than in Model 5 for each price change simulation. Table 6 also demonstrates the degree of prediction inaccuracy of population

subgroups when previous behavior is not modeled. There are marked differences across the three models for previous smokers and non-smokers. Model 3 simply relies on observed covariates other than previous behavior to explain differences between these groups. Model 4, on the other hand, overstates previous smokers' smoking propensities since the coefficient on previous behavior (assumed exogenous in this model) is strongly correlated with unobserved heterogeneity and is therefore biased. Separated by year, we see that the probability of smoking is overestimated in Models 3 and 4 by as much as seven percentage points following a \$1/pack increase.²⁵

We next illustrate the magnitude and importance of state dependence in cigarette consumption. Consider the effect of a one-time cigarette tax (or price) increase. This is a convenient, if unlikely, policy to consider since any change in smoking behavior in future years is the result of modified behavior in the year of the tax increase.²⁶ Using our preferred specification, we simulated a \$1 increase in real prices in 1988, with prices held at their original 1990 and 1992 real levels. Relative to behavior at original prices, there is a 2.46 percentage point drop in smoking participation rates in 1988. This in turn results in a 0.67 percentage point (four percent of base) decrease in smoking rates in 1990 and a 0.25 percentage point (one percent of base) decrease in 1992. The decrease in future participation rates is observed among both non-smokers and smokers. Subsequent movement away from moderate and heavy levels of smoking among smokers also continues after the one-time 1988 tax increase. Despite low smoking rates among 8th graders (and hence, few individuals for whom the simulated price increase will alter smoking behavior), the participation changes are large relative to their baseline levels. These results confirm that price changes have important dynamic implications.

Table 7 shows our final experiment that measures the pure effect of previous behavior (state dependence) on smoking decisions. As we have argued above, the parameters on the smoking history variables in Model 5 more precisely measure the influence of past smoking behavior because there are controls for observed and unobserved heterogeneity that influence the smoking decisions over time. The estimates from Model 4 can only provide the combined effect of both past behavior and unobserved preferences. To measure these effects, we perform identical simulations and use the updating procedure to generate smoking histories for both Models 4 and 5. The simulations force particular types of smoking behavior in order to demonstrate the role of state dependence. We first force all individuals in the sample to be non-smokers in 1988. We then update their history and simulate behavior in 1990 and 1992. Similarly, we allow no one to smoke in 1988 and 1990 and predict behavior in 1992. These experiments are akin to a *perfectly enforced* minimum age purchase restriction (that is, age 16 and age 18 respectively). Compared to observed participation, the top panel of Table 7 shows that this enforcement could reduce smoking rates in 1992 by five percentage points (17.76 compared to an observed rate of 22.69).

This procedure is repeated forcing all individuals to be either light, moderate, or heavy smokers in 1988 and simulating their behavior in 1990 and 1992. From Model 5, Table 7 shows the pure effect of smoking lightly in 1988 as opposed to not smoking in 1988 is to increase participation rates by almost 15 percentage points (29.50 - 14.63) in 1990 and by 4.6 percentage points in 1992 (26.94 - 22.34). That is, individuals who smoke in 1988 are two times more likely to smoke in 1990 and 20 percent more likely to smoke in 1992. The smaller difference in 1992 than 1990 reflects the fact that some individuals, who were forced to smoke in 1988 and who otherwise would not have, will choose not to smoke in 1990 decreasing their probability of

smoking in 1992. The second simulation shows that light smoking in 1988 *and* 1990 increases the probability of smoking participation in 1992 by 17.5 percentage points. The table also reports the simulated participation behavior when individuals are forced to be moderate or heavy smokers in the early periods.²⁷

We report the same simulations using Model 4 to illustrate the biased estimated effect of previous behavior. When true state dependence is combined with unobserved heterogeneity, the simulated participation rates are larger and incorrect. The bottom panel of Table 7 shows that individuals who smoked lightly in 1988 as opposed to not smoking in 1988 are 4.4 times more likely to smoke in 1990 and 2.2 times more likely to smoke in 1992. Smoking lightly in 1988 and 1990 makes one 4.6 times more likely to smoke in 1992. The results from both models provide clear evidence that smoking behavior persists. More importantly, when we purge the estimated effects of unobserved heterogeneity, there is still a sizable effect of previous behavior. The Model 5 estimates allow us to measure the impact of previous smoking behavior exclusive of permanent unobserved tastes.

V. Conclusion

This paper highlights the importance of explicitly taking into consideration endogenous smoking history and unobserved individual heterogeneity when explaining smoking behavior, measuring price sensitivity, or predicting future smoking responses to price changes. By controlling for a wide range of observed and unobserved individual differences, we are able to show that behavior modification plays an important role in smoking persistence. Such state dependence, however, is overstated when the endogeneity of previous behavior is not considered. The paper also demonstrates that price increases can influence future behavior by reducing the current number

of smokers; such increases produce a population with a larger proportion of more price sensitive non-smokers. We show that ignoring smoking dynamics can give misleading estimates of the aggregate effect of various cigarette price increases. This bias is likely to be more severe over longer time horizons than the four years considered here if price increases continue to expand the number of non-smokers.

This dynamic framework can address questions of interest to policy-makers. For example, the model can forecast the immediate response of youth smoking to cigarette tax increases as well as predict how youth will be affected in the long run. We also show how tax changes can influence current smokers and non-smokers differently. The procedures we use can be extended to longer panels or an older population to obtain better estimates of long-run effects of tax increases and to determine the importance of dynamics among adults (that is, are adult smokers less price sensitive than adult non-smokers?). With a more complete understanding of lifetime smoking dynamics, we will be able to forecast how overall population smoking rates will vary in the short and long run in response to various price changes. Investigating these extensions will help inform the current debate regarding the impact of future cigarette taxes and other smoking policies.

References

- Angeles, Gustavo, David Guilkey, and Thomas Mroz. 1998. "Purposive Program Placement and the Estimation of Family Planning Program Effects in Tanzania." *Journal of the American Statistical Association* 93(443): 884-899.
- Becker, Gary and Kevin Murphy. 1988. "A Theory of Rational Addiction." *Journal of Political Economy* 96(4): 675-700.
- Becker, Gary, Michael Grossman, and Kevin Murphy. 1994. "An Empirical Analysis of Cigarette Addiction." *American Economic Review* 84(3): 396-418.
- Blau, David and Alison Hagy. 1998. "The Demand for Quality in Child Care." *Journal of Political Economy* 106(1): 104-146.
- Chaloupka, Frank. 1991. "Rational Addictive Behavior and Cigarette Smoking." *Journal of Political Economy* 99(4): 722-742.
- Chaloupka, Frank and Kenneth Warner. 2001. "The Economics of Smoking." In *The Handbook of Health Economics*, eds. J. Newhouse and A. Culyer. 1539-1627. New York: North-Holland.
- Copejans, Mark, Donna Gilleskie, Holger Sieg, and K Coleman Strumpf. 2003. "Price Uncertainty, Tax Policy, and Addiction: Evidence and Implications." Working Paper.
- Cutler, David. 1995. "The Incidence of Adverse Medical Outcomes under Prospective Payment." *Econometrica* 63(1): 29-50.

DeCicca, Philip, Donald Kenkel, and Alan Mathios. 2002. "Putting Out the Fires: Will Higher Taxes Reduce Youth Smoking?" *Journal of Political Economy* 110(1): 144-169.

Douglas, Stratford. 1998. "The Duration of the Smoking Habit." *Economic Inquiry* 36(1): 49-64.

Gilleskie, Donna and Koleman Strumpf. 2000. "The Behavioral Dynamics of Youth Smoking." National Bureau of Economic Research Working Paper No. 7838.

Goldman, Dana. 1995. "Managed Care as a Public Cost-Containment Mechanism." *Rand Journal of Economics* 26(2): 277-295.

Gruber, Jonathan. 2001. "Youth Smoking in the 1990s: Why Did it Rise and What are the Long Run Implications?" *American Economic Review: Papers and Proceedings* 91(2): 85-90.

Heckman, James. 1981. "Statistical Models for Discrete Panel Data." In *Structural Analysis of Discrete Data and Econometric Applications*, eds. C. Manski and D. McFadden. Cambridge: The MIT Press.

Heckman, James and Burton Singer. 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica* 52(2): 271-320.

Hoynes, Hilary. 1996. "Welfare Transfers in Two-parent Families: Labor Supply and Welfare Participation under AFDC-UP." *Econometrica* 64(2): 295-332.

Hsiao, Cheng. 1986. *Analysis of Panel Data*. Cambridge: Cambridge University Press.

Hu, W. 1999. "Child Support, Welfare Dependency, and Women's Labor Supply." *Journal of Human Resources* 34(1): 71-103.

Kenkel, Donald, Dean Lillard, and Alan Mathios. 2003. "Smoke or Fog? The Usefulness of Retrospectively Reported Information about Smoking." *Addiction* 98(9): 1307-1313.

Kreider, Brent and Regina Riphahn. 2000. "Explaining Applications to the US Disability System - A Semiparametric Approach." *Journal of Human Resources* 35(1): 82-115.

Labeaga, Jose. 1999. "A Double-Hurdle Rational Addiction Model with Heterogeneity: Estimating the Demand for Tobacco." *Journal of Econometrics* 93(1): 49-72.

Labeaga, Jose and Andrew Jones. 2003. "Individual heterogeneity and censoring in panel data estimates of tobacco expenditures." *Journal of Applied Econometrics* 18(2): 157-177.

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

Nickell, Stephen. 1981. "Biases in Dynamic Models with Fixed Effects." *Econometrica* 49(6): 1417-1426.

Tauras, John and Frank Chaloupka. 2001. "Determinants of Smoking Cessation: An Analysis of Young Adult Men and Women." In *The Economics of Substance Use and Abuse: The Experience of Developed Countries and Lessons for Developing Countries*, eds. M. Grossman and C. Hsieh. Edward Elgar Publishing Limited.

Tobacco Institute. 1997. *The Tax Burden on Tobacco: Historical Compilation*. vol. 32.

U.S. Department of Health and Human Services. 1994. "Preventing Tobacco Use Among Young People: A Report of the Surgeon General." Atlanta: Public Health Service, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Endnotes

1 In addition, smoking is initiated at an early age. Almost 42% of current or former adult smokers start before age 16 and 75% start before age 19 (Gruber, 2001). While individuals under age 18 consume only 2% of the total amount of cigarettes smoked in a year in the U.S., such teen smoking has long-term dynamic implications because youth smoking appears to persist into adulthood (U.S. DHHS 1994).

2 We do not use the term addiction, since there are many possible sources of state dependent behavior. In addition to the channels listed in the text, persistent smoking behavior might reflect convex adjustment costs.

3 Our panel data also provide recent and relatively accurate reports of behavior as opposed to the retrospective data that Chaloupka employs. Additionally, while he includes no controls for (unobserved) state-level heterogeneity, we show that this omission imparts a significant negative bias on price effects.

4 Gruber (2001) discusses other recent work on teen smoking behavior. These papers primarily use cross-sectional analysis.

5 Becker et al. (1994), an important paper that uses aggregate data, is among the first tests of the rational addiction model. Chaloupka (1991) is the first to use micro level data to test this model.

6 Because we do not estimate policy-invariant structural parameters, we only simulate policies that exist within the data.

7 Our approach differs from the empirical rational addiction literature where long-run behavior is defined as the behavior that would occur in steady state (for example, Chaloupka 1991). One difficulty with steady state predictions is that it may take several years to reach convergence and, as our results suggest, key variables like price sensitivity vary with time. Our approach predicts behavior along the adjustment trajectory.

8 A third follow-up was administered in 1994 but does not contain information on smoking behavior. A fourth follow-up was administered in the spring of 2000 but the restricted use file is unavailable as of summer 2004.

9 We do not have a state identifier for several dropouts in 1990. We retain the observation if the individual resided in the same state observed in 1988 and 1992 and assume that this was their state of residence in 1990. To be consistent with this assumption, we do not consider the behavior of any child who moves to a new state after 1988.

10 Because the response “smoke less than 1 cigarette a day” is available only in 1990, we group this outcome with “smoke 1 to 5 cigarettes.” Due to the small number of responses in the 40+ category, our top category is defined as “smoke more than half a pack.”

11 Because very few children younger than age 14 smoke, because most individuals are age 14 or older in 8th grade, and because our data do not indicate first age smoking if initiation occurred before 8th grade, we assume that no teens in our sample smoked prior to 8th grade.

12 There is noticeable price variation even after controlling for unobserved state-level heterogeneity. After regressing prices on state and year fixed effects, the residuals move considerably with several large states (California, New Jersey, and New York) having ranges in

excess of 15 cents. Gilleskie and Strumpf (2000) document the wide variation in prices, as well as in state tax rates and tax rate changes during the sample period.

13 See DeCicca et al. (1997) and Labeaga and Jones (2003) for exceptions. Retrospective data provides a different route for analysis, as in Chaloupka (1991), but also presents different issues of data reliability (Kenkel et al. 2003).

14 A significant effect of lagged prices on current behavior might also suggest, however, that behavior today is a function of future price expectations which are formed using price histories (Coppejans et al. 2004).

15 We also explore the endogeneity of an important and self-selected group: individuals who drop out of school. We include an equation explaining the probability of leaving school between observed waves in the data and estimate it jointly with the smoking decisions that include the endogenous dropout status as an explanatory variable. Details on our approach can be found in Gilleskie and Strumpf (2000).

16 We technically need six individuals (two in each of three different states) to achieve identification in this example, since there are also state and year fixed effects.

17 Full parameter estimates from each of the models we discuss are available in appendices on the *Journal of Human Resources* web page.

18 In order to investigate price sensitivity variation over time we include interactions of price with previous use and year indicators. In unreported specifications, we interact price with the continuous age variable and found similar price sensitivities suggesting that elasticities differ by year but not by age within a particular year or grade.

19 Overall, the Model 4 estimates give a price elasticity of total consumption of -0.51. The predicted total number of cigarettes consumed for any set of prices is $[p(smoke) \sum_i p(i / smoke) \times (number\ of\ cigarettes\ in\ i)]$ where i is a conditional smoking intensity category. We assigned the following number of cigarettes smoked to these categories (which are reported as a range in the data): 3 for category 1 (1-5 cigarettes), 8 for category 2 (6-10 cigarettes), and 15 for category 3 (11+ cigarettes).

20 The first stage instruments, or variables included in the first stage that are excluded from the second stage, are the lagged price variables of the previous period. Note that Chaloupka (1991) uses this two-stage procedure for modeling the endogeneity of previous smoking levels.

21 Results are reported from a model with two factors (μ) representing unobserved heterogeneity and three points of support on each factor; four points of supports on either factor did not provide significant improvement in the likelihood function. Because the estimated factor loadings (ρ) are negative, individuals with unobserved characteristics at the right of the each distribution are less likely to smoke. The estimated points of support are (0.00, 0.38, and 1.00) with estimated weights of (0.11, 0.51, and 0.36) for one factor (factor a) and (0.00, 0.06, and 1.00) with estimated weights of (0.42, 0.47, and 0.11) for the second factor (factor b). Note that the asymmetric distribution exhibited by the second factor and the large negative factor loadings on this factor (displayed in Table 4) help explain why some individuals never smoke.

22 The significance of the reported elasticities is obtained by simulating the behavior of the sample at numerous random draws of the full set of parameters based on point estimates and the estimated variance-covariance matrix. The standard deviation of the $n = 100$ simulations approximates the standard error of the calculated statistic.

23 Additional details on the updating procedure are available from the authors.

24 By construction, updating has no effect on behavior in 1988, the first year of the sample, because previous behavior is unchanged.

25 We also performed the simulations in this section on the two-stage model that instruments for endogenous previous smoking participation (discussed at the end of section IV.A). The results are similar to those from Model 3. The instrumental variable approach relies only on observed covariate variation; the joint modeling approach we use in our preferred Model 5 considers both observed and unobserved influences.

26 Predictions of these future responses are not feasible without modeling behavior over time. Similarly, it is necessary to update behavior in response to a price change because predictions based on actual behavior would indicate no change in long-run behavior under this temporary increase.

27 We simulate smaller probabilities of future smoking among those forced to smoke heavily than those forced to be moderate smokers. Interestingly, we see the same behavior in the observed data displayed in Table 1, and this is also reflected in the estimated coefficients on lagged behavior in Table 4.