participation equation and a lower likelihood of consuming at least one pack per day in the consumption equation.

These ideas bear repeating. On average we expected to see smokers that had indicated an intention to quit (which we are taking as a reflection of the symmetric state) to be less likely to smoke in the current month if they also observed a signal to quit. We expected also that daily smokers would be less likely to consume a high quantity on a daily basis if they were in the symmetric state and had observed a signal to quit.

The addict's behavior in the complementary state is more interesting but perhaps less intuitive. A lack of attempts to quit is a sign that the smoker is in the complementary state. For smokers in the complementary state, we should observe the opposite behavior to that of smokers in the symmetric state. Specifically, we would expect complementary-state smokers to be more likely smoke in the current month if they had received a signal to quit. Moreover, we would expect that they would be likely to consume a higher-level per day. The increased probabilities associated should be reflected in odds ratios above one.

The signal variables in the table above do indeed exhibit the pattern of behavior that Bateson's model suggests. Individuals who have had a problem due to their marijuana consumption in the past year are less likely to smoke cigarettes in the current month (`PARTICIP`) given that they had attempted to quit. However, given no attempts to quit, they are actually *more* likely to smoke cigarettes in the current month when confronted with a signal to quit (`MRJPROBS`).

A similar pattern is observed with the other variables in the table and for the likelihood of increased consumption. The relationship, however, is modest. Additionally, no variables, other than ARGUMENT presented in the table, exhibit the expected pattern of odds ratios along with significant Q statistics. We note specifically that "problems with cigarettes" was not a signal to which smokers responded in the predicted fashion. We found however, signals from cigarette problems (CIGPROBS) did enter the regression model with the proper signs. We discuss in the next section.

Although we were able to identify suitable regressors to represent the observed signals, use of these regressors in the participation equation presented potential econometric issues. The variables used to identify the symmetric state (TRY_QUIT and QUIT_ANY) contain information about the choice to smoke. Specifically, a person can only attempt to quit if he or she already smokes. The consequence for the analysis was that estimation of the participation equations probably requires that use of instrumental variable or simultaneous equation methods to properly measure the effect of signals on the choice to smoke.

4.2.1.2 Participation

In Chapter 3, we explained the stochastic structure of the model. We proposed there that participation could be classified into three groups. For the analysis, we coded a dependent variable, PARTICIP, from the response to Question C-7 in the NHSDA. The question asks the respondent about his or her level of daily cigarette consumption in the past 30 days. The text of the question is reproduced below from the public use file documentation (SAMHSA, 1993b).

Table 4.2-3 Survey Question Used for Dependent Variables

LABEL	LEN	DESCRIPTION	FREQ	%
		C-7. How many cigarettes have you smoked per day, on the average, during the past 30 days? Give me the average number per day.		
AVCIG	2	1= Less than one cigarette a day	751	2.84
		2= One to five cigarettes a day	1540	5.81
		3= About 1/2 pack a day (6-15 cigarettes)	2640	9.97
		4= About a pack a day (16-25 cigarettes)	2397	9.05
		5= About 1 1/2 packs a day (26-35 cigarettes)	548	2.07
		6= About 2 packs or more a day (over 35	338	1.28
		89= LEGITIMATE SKIP Logically assigned	147	0.55
		91= NEVER USED CIGARETTES	10265	38.75
		94= DON'T KNOW	3	0.01
		97= REFUSED	1	0.00
		98= BLANK (NO ANSWER)	355	1.34
		99= LEGITIMATE SKIP	7504	28.33

If a legitimate skip or a never smoked response was indicated, then `PARTICIP` was coded as a zero (0). If the less than one cigarette response was indicated, then the dependent variable was coded as one (1). Higher levels of consumption, 1 or more cigarettes per day, were coded as a single value of two (2). A small number of records had invalid or missing information and so the dependent variable was not coded for those records. These records are those coded as responses 94, 97 and 98 in Questions C-7 from the 1993 NHSDA[6].

[6] The figures cited in the above figure are for the entire sample of the 1993 survey. They are a superset of the sample of records from the 1993 survey used in analysis.

The three levels of the dependent variable, 0, 1 and 2 correspond to the non-smoker, occasional smoker, and daily smoker classes we discussed in the section on the empirical model. Only 775 occasional smokers were observed in the entire analysis sample of 29976 respondents. This low prevalence created some statistical challenges in the ordered logistic participation equation. The two variables that were used to indicate the symmetric/complementary state of the smoker were near perfect classifiers of participation for small enough subsets of the analysis sample. This limited the types of validation that could be done on the regression results. In Table 4.2-4 we show the distribution of respondents as classified by the two dependent variables: participation and consumption.

Table 4.2-4 Distribution of Respondents by Consumption

Participation (PARTICIP)	Consumption (HALFPACK)					
	0	1	2	3	4	Total
Non-smoker	22848					22848
Occasional	775					775
Daily	1473	2237	2014	377	252	6353
Total	25096	2237	2014	377	252	29976

4.2.1.3 Consumption

The dependent variable for daily consumption, HALFPACK, was coded from the answer to the same question on NHSDA as used to code participation. It was coded in the ½ pack increments indicated by responses "3" to "6" from question. For example, if an individual marked the response "4" the answer to the question, then the dependent variable was coded as "2" indicating two halves of a package were consumed daily.

The choice to code the dependent variable for daily consumption as a count value was made based on the wording of the survey question. Although the question clearly specifies a range of cigarettes in the text of the questions, only an indicator of level is recorded. Moreover, the wording suggests that the respondent think in terms of ½ pack increments by use of the term "about". Such terminology leads the respondent to think in terms of aggregates rather than specifics.

The coding of the response labeled "2" presented a statistical problem. A specific range of cigarettes is given as the only reference consumption level rather than an aggregate as in the other responses. Moreover, the range is not a ½ pack increment.

Several solutions to the problem were considered. The first solution was to code all responses in ¼ pack increments. This solution would have maintained the count data nature of the responses but it arbitrarily recoded the four other responses in a scale that was not the natural scale clearly given in the survey wording. A second solution was to censor response "2" by recording with the same value as response "3". This second solution was attractive inasmuch as it provided clear separation between the zeros in consumption in and the non-zeros. In order to implement the solution, an alternative model specification would need to be estimated. We chose to consider this solution as a future research opportunity.

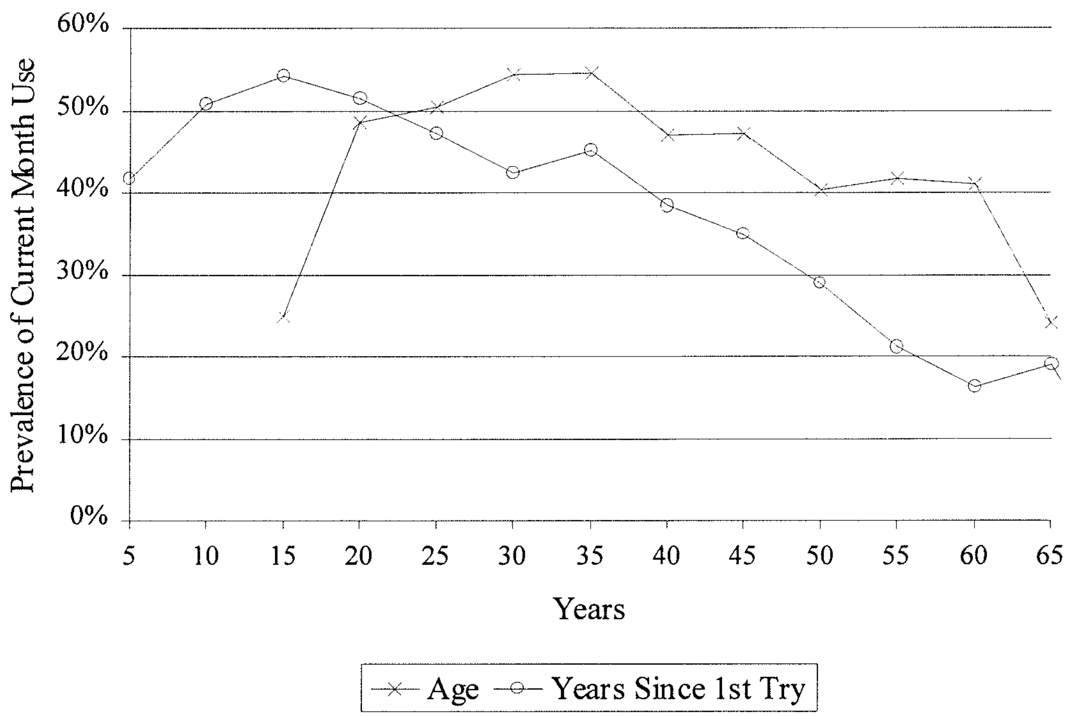
An additional solution, which we also rejected, was to estimate a non-count data model specification. In this last solution, the values for the dependent variable would have coded as the average of the ranges stated in the text of the survey question. We rejected this solution based on the wording of the questions. The solution we used was simply to round the value down to zero and treat response as an ordinal level with an equal distance between it and the next value.

4.2.1.4 Methods

We used factor analysis and regression diagnostics as tools to identify potential collinearity in the regressors. A number of the variables were jointly correlated. The factor structure identified collinearity between age, marital status, educational attainment, student status, veteran status and use of alcohol. Similarly, socioeconomic status, receipt of welfare payments, cost of cigarettes as a share of daily income, real household income and several of the minority race indicators were shown to be jointly correlated. The indicators of monthly and daily use of marijuana, cocaine and alcohol were found to be collinear. Judgement and univariate statistics were used to select the candidate variables for incorporation in the final regressions. In certain cases variables from the same factor were chosen for incorporation into the regressions so long as the collinearity remained small. An example is the pair of age and years of education. Both variables enter into the participation equation.

Univariate analysis provided insight into the relationship between the dependent and independent variables. The relationship between the any cigarette use in the past month (either occasional *or* daily) was studied. For example, age and years of education were found to have non-linear relationship with prevalence of use. We accounted for this in the estimation with additional polynomial terms in the regressions.

Figure 4.2-1 Prevalence of Cigarette Use



The relationship between age and use appears to reflect life-cycle behavior. We observed a very similar curve for prevalence by the number of years since the individual had first tried cigarettes. We suspect the relationship is both an unfortunate reflection of mortality rates and also a reflection of the movement of cohorts of smokers through the population. Cigarette use rates increased between the 1950s and 1980s. The cohorts of smokers would who began smoking during those years are those that have the highest prevalence on the graph. If this is the underlying causal relationship in the data, then the model results do not reflect a long-term stable relationship.

Finally, we used SAS software (version 6.12) published by the SAS Institute to perform our analysis. The ordered logistic model was estimated with the SAS PROC LOGISTIC procedure. The censored Poisson model was coded in SAS IML and we used the IML optimization packages to solve for the maximum likelihood. We applied the IML optimization routine NLPNRR that uses a modified Newton-Raphson algorithm.

4.3 Results

In this section we review the results of the estimation of the equations for participation and consumption. The two equations being estimated provide estimates for the conditional mean of a latent variable for participation, z^*, and the conditional mean of the amount consumed in ½ pack increments, y. We show these below.

$$E[z^* \mid x] = x\gamma$$ **EQ. 4.3-1 Participation**

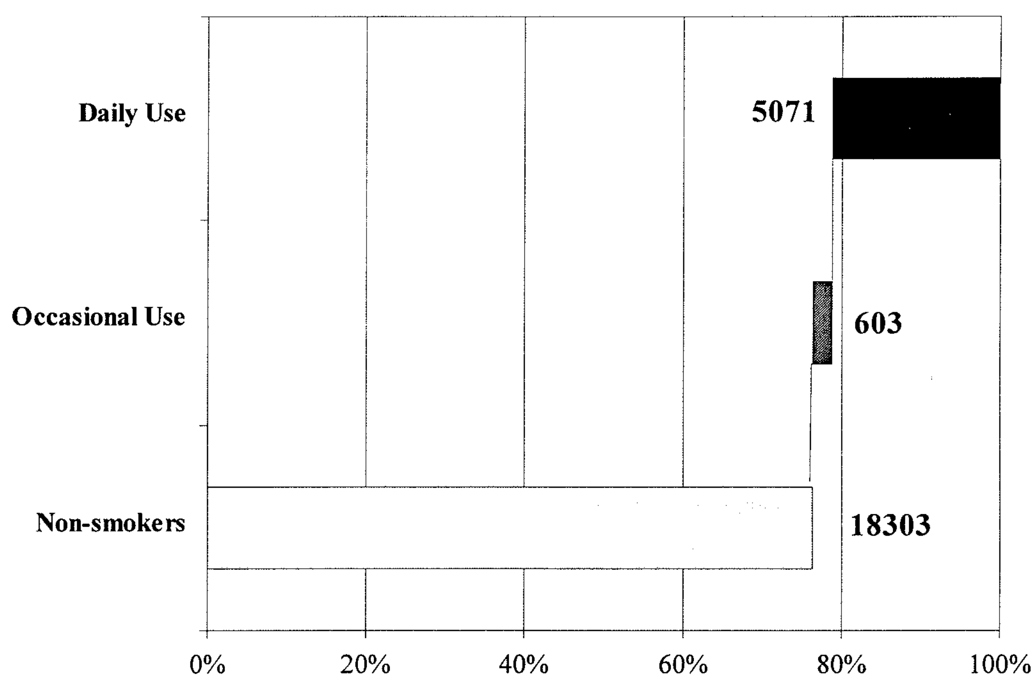
$$E[y \mid x] = e^{x\beta}$$ **EQ. 4.3-2 Consumption**

4.3.1 Participation Regressions

We first consider the results of the estimation of participation. As we discussed in the previous chapter, we used the ordered logit regression model to estimate the latent participation variable. We noted in the last section that we estimated our model on a subset of the data in order that we could perform some validation using the remainder of the sample. We also dropped a small number of records from the final estimation sample due to missing data in the regressors.

Our final set of regression data for the ordered logit estimation included 23,977 observations. Figure 4.3—1 below shows the distribution of the values for the participation dependent variable. We note that a striking feature of the distribution is the small size of the occasional smoker segment of the population. Occasional smokers seem to compose about 2.5% of the unweighted sample. In the weighted sample, the proportional representation of occasional smokers is about the same – about 2.6%. The weighted representation conditioned on having smoked in the past year is 9.0%.

Figure 4.3-1 Distribution of Participation in Estimation Sample



We estimated several versions of the participation equation to compare the effect of inclusion of different variables. We estimated three different regressions: a regression including all the socioeconomic determinants participation, a regression with only the signal variables and a regression with all the variables.

We also estimated a regression on a sub-sample including only the ever-smoker population. This latter regression was to provide information about the potential impact on the regressions, which used the entire sample, of the endogeneity of the information variables. Jones (1989) also found that his results were affected by the difference in the populations of those individuals who had smoked and those who had never tried a cigarette. He suggests that future studies handle these populations separately. We report on this regression below with the other regressions.

Finally, in order to investigate the importance of the small representation of occasional, we also estimated a regression on the choice to smoke versus not smoke in the current month on the ever-smoker sample. The loglikelihood for that regression was the same as presented in EQ 3.2-16. We do not give the details of the results of that regression here. However, we found that the loglikelihood for intercepts only was –9464 and the loglikelihood for the intercepts and regressors was –6801. The likelihood ratio test for significance of this regression, 5325, was somewhat higher than that for ordered logistic with three levels of participation. Additionally, no important changes in the significance of the regressors were found.

Fifteen variables entered into the socioeconomic-only regression. These variables are detailed in Table 4.3-1 below. The coefficient estimates are shown with the associated χ^2 test of significance in parentheses below that coefficient. The marginal effect of each variable on the probability of smoking daily in the current month is given in the column labeled "Marginal". Also listed at the end of the table are the *c* statistic (a measure of fit), the loglikelihood for the regression and the intercept only, the χ^2 statistic for the proportional odds assumption (a specification test) and finally the number of observations in the sample.

All variables were strongly significant as predictors of participation. Age and the square root of age came in with the expected signs per our discussion in the previous section. We had noted there that age exhibited a non-linear relationship with participation where over most of its range age is negatively related to participation. However, for low years of age a positive relationship was seen.

The square root of age also had the largest marginal effect on the probability of smoking daily in the current month. Although inclusion of the square root of age makes the mathematical of the marginal effect of age untidy, the model fit was improved noticeably. Moreover, the square of age did not provide as good a fit as the square root. The next largest marginal effect, that on the indicator (`TREATMNT`) for attendance of a drug treatment program in the past 12 months, was about half that of the square root of age.

Table 4.3-1 Participation Regressions*

Variable	Non-signal		Signal	Both, Full Sample		Both, Ever Used	
	Coefficient	*Marginal*	*Coefficient*	*Coefficient*	*Marginal*	*Coefficient*	*Marginal*
Constant 1	-3.4008		-2.2130	-3.8123		-1.7840	
Constant 2	-3.2567		-2.0133	-3.5986		-1.5478	
AGE	-0.2571 (313.2)	-0.0348		-0.2470 (204.0)	-0.0276	-0.2481 (181.8)	-0.0528
ROOTAGE	3.1742 (299.9)	0.4296		3.0879 (199.0)	0.3453	3.0084 (167.1)	0.6405
YEARSEDU	-0.1137 (252.6)	-0.0154		-0.0944 (123.6)	-0.0106	-0.0988 (119.7)	-0.0210
FEMALE	-0.1081 (8.9)	-0.0147		-0.2107 (23.4)	-0.0164	-0.0677 (2.2)	-0.0102
MARRIED	-0.3605 (96.2)	-0.0488		-0.3598 (66.3)	-0.0278	-0.3543 (57.8)	-0.0538
LNINCOME	-0.0815 (17.2)	-0.0110		-0.0747 (10.1)	-0.0084	-0.0902 (12.6)	-0.0192
LOWSOCIO	0.1477 (15.3)	0.0203		0.2016 (19.8)	0.0160	0.2689 (31.4)	0.0420
STUDENT	-0.4467 (30.2)	-0.0542		-0.4925 (25.6)	-0.0329	-0.4440 (19.2)	-0.0594
MRJFLAG	0.7915 (427.5)	0.1189		0.4802 (105.2)	0.0403	0.0707 (2.0)	0.0107
ALCFLAG	1.0112 (270.9)	0.1122		0.7525 (118.4)	0.0489	-0.4844 (29.8)	-0.0830
ALCDLYF	0.7367 (208.7)	0.1210		0.7799 (166.9)	0.0789	0.7158 (136.3)	0.1263
TREATMNT	1.2069 (65.0)	0.2290		0.8425 (20.6)	0.0913	0.7542 (16.3)	0.1398
LORISK	0.8042 (79.9)	0.1377		0.9892 (98.9)	0.1120	0.8564 (60.3)	0.1614
HISPANIC	-0.2503 (32.9)	-0.0325		-0.2660 (26.2)	-0.0195	-0.2329 (17.5)	-0.0337
SERVICE	0.2628 (21.6)	0.0381		0.1890 (7.9)	0.0155	0.1331 (3.7)	0.0207
TRY_QUIT			3.5687 (5572.9)	3.3985 (4600.6)	0.5774	2.8239 (3073.9)	0.5572
MRJPROB0			1.2388 (49.3)	0.6746 (12.8)	0.0689	0.4515 (5.4)	0.0778
MRJPROBX			-1.8401 (41.9)	-1.5372 (26.0)	-0.0651	-1.1726 (14.9)	-0.1198
Ln(L)	-13069		-10824	-10043		-8526	
Ln(L) Intercept	-14640		-14640	-14640		-11157	
P.O. Test	268		90	351		348	
N	23997		23997	23997		14287	
Pr[daily smoker]	0.1614		0.1515	0.1283		0.3074	

* Wald statistics in parentheses. See Appendix C for definitions of the regressors. The P.O. test tests the proportional odds assumption: the assumption that regressors have identical coefficients in each logit. Alternatively, the coefficients of the regressors may differ across logits. For example, in the *true* model, the coefficient of age in the logit for non-smokers might be different than that in the logit for daily smokers.

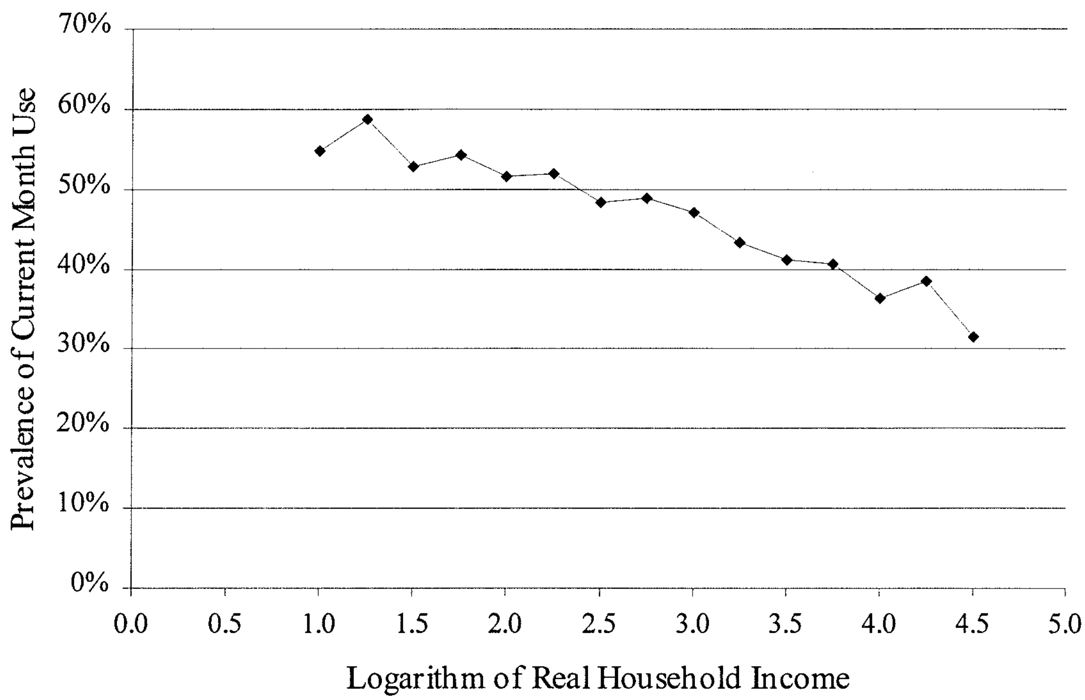
Years of education also entered with the expected sign – negative – which was expected given that the observed distribution was similar to that of age. We did not include the square root of years of education although we did consider the variable. The square root of education was strongly collinear with the square root of age and resulted in significantly biased results in another regression we have not presented here.

As in the study by the Farrelly and Bray (1998), Unger and Chen (1999) and others we found women less likely to be at risk of smoking in the current month. This risk is relative to our control group. Our control group is single non-student 40 year old men from an average socioeconomic status who have never used marijuana or alcohol, and have never served in the military. The control group also holds a moderate view of the risk of smoking one pack of cigarettes per day and has completed schooling through high school. In our fourth regression, the "Ever Used" regression, we modify the control group to be men of the same background who also have tried smoking a cigarette at least once.

Also similar to other studies (Ma and Shive (2000), Unger and Chen (1999)) we found that Hispanics were less likely to smoke in the current month. Student status was found to be an indicator of lower risk of smoking, and military service history was an indicator of increased risk.

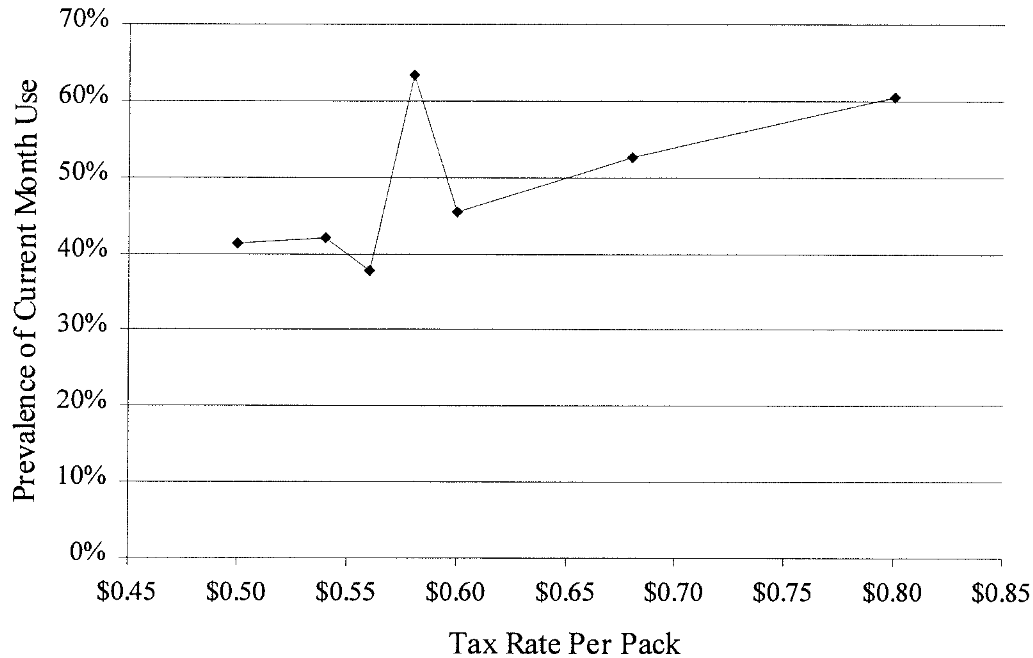
We found that household income was negatively related to risk of smoking in the current month. The relationship between risk (of either occasional or daily use) and the logarithm of income (1982-84=$100) is given in the figure below. The relationship is strongly linear.

Figure 4.3-2 Income and the Risk of Smoking



We considered other economic variables for the regression. Both real and nominal prices were considered. Generally, neither form of price was significant in the estimation. There were certain cases where price entered the regression but these were usually found to be cases where there was strong collinearity in a set of the regressors and thus were not considered further. We also considered excise taxes. These too were found to be insignificant in many regressions but there did appear to be a linear relationship between taxes and risk. However, the direction was counter to economic intuition. We actually found that increasing taxes increased risk. The relationship is given below.

Figure 4.3-3 Tax Rates and Risk of Smoking



We did not explore use of the taxes in the participation equation further given the apparent positive relationship with risk.

The second regression we estimated for participation was the estimation of the effects of the signal variables. Again, we note that these results are problematic in that the regressor, `TRY_QUIT`, is likely to be endogenous by definition. We provide the regression results as descriptive. We found the signs of the variables to be as expected. The sign was negative on coefficient of the indicator (`MRJPROBX`) of individuals who had had problems (the signal) with marijuana and who had also attempted to abstain. The sign for the coefficient of the indicator (`MRJPROBO`) of individuals who had not reported a quit attempt was positive.

All the results given so far hold for the combined regression including the socioeconomic variables and the signal variables. The signs of the coefficients are the same and the variables all enter as significant. Most of the marginal effects change slightly except for the effects of the indicators of alcohol and marijuana use. These both decrease by about fifty percent, as does the indicator of participation in a treatment program. This change is probably due to the effect of the signal variable that captures related information about the problems associated with the use of marijuana.

The final regression, that of all variables over a sub-sample of individuals who had tried cigarettes at least once, gives some different results. Notably, the gender indicator is found to be insignificant. The same is found for the indicator of ever use of marijuana. Additionally, the indicator of ever-use of alcohol changes sign. The indicator of military service is found to be barely significant at the 95% confidence level.

These results suggest the following. First, gender does not help predict future smoking once initiation has occurred. This result is consistent with Patton et al. (1998) who found that females were no more likely than males to exhibit an incidence of daily smoking when conditioned on prior daily use.

Second, indicators of ever use of either alcohol or marijuana appear to be predictors strictly of initiation but not repeat use of cigarettes. The daily use of alcohol remains a strongly significant variable whereas the other two indicators do not; repeat use of an alternative substance is predictive of repeat use on cigarettes. A similar conclusion might be drawn about the indicator of military service – that military service is predictive of initiation but not repeat use.

The marginal effects are stronger in the "Ever Used" regression. The pattern between the regressors is similar, however. The square root of age still has the largest effect on the probability of daily smoking. The signal variables all exhibit strong effects on the probability of daily smoking. The indicator of the symmetric state (i.e., the indicator of quit attempts) is has the second largest effect on the probability and the indicator of those individuals with marijuana problems *and* quit attempts has the third largest effect. None of the other variables exhibit significant sign changes (or even large changes in the magnitude of the coefficients) in the results of the estimation the sub-sample of ever users.

The negative relationship between income and participation is interesting. We might be led to conclude that cigarettes are therefore an inferior good. However, we can draw that conclusion only by examining the relationship of income and consumption. We considered that the relationship was due to possibility that income was associated with educational attainment and that educational attainment indicated a level of information about the risks of smoking hence the choice to abstain. Upon reflection, however, we would expect that the incorporation of the years of education variable in the model would render the income variable insignificant. We also considered the same idea for age. With both age and educational status in the model and the income variable still strongly significant, we conclude that there is some unique information about the choice to smoke inherent in the income variable.

We calculated the participation elasticity of income from the regression results. The elasticity, ε_{daily}, is derived from the following relationship:

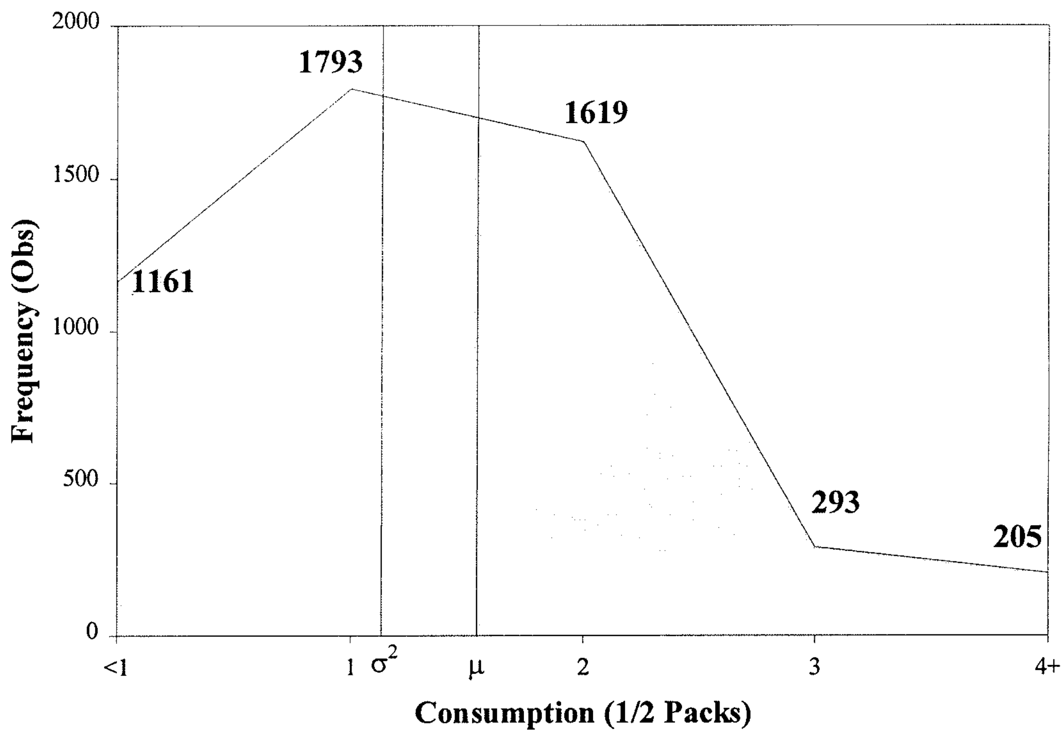
$$\varepsilon_{daily} = \frac{\partial p_{daily}}{\partial Y}\frac{Y}{p_{daily}} = \left(Y\frac{\partial p_{daily}}{\partial Y}\right)\frac{1}{p_{daily}} = \left(\frac{\partial p_{daily}}{\partial \ln Y}\right)\frac{1}{p_{daily}} \qquad \textbf{EQ. 4.3-3}$$

where Y is income and p_{daily} is the probability of smoking daily. The change in p_{daily} for a given change in Y is simply the marginal effect given on Table 4.3-1 above. For the "Ever Used" regression this value -0.0210 when evaluated at the mean of the regressors. The associated risk of daily smoking at the means is 30.74%. The ratio of these two numbers gives an elasticity of –0.068. The elasticity is for changes in income near the mean income value of $20,5000.

4.3.2 Consumption

We estimated three separate regressions for consumption. There was no problem with endogeneity in the regressors so all three results are consistent (maintaining the assumed conditional mean specification). We tested the model specification using a test cited by Greene (1999) that tests the specification of the conditional mean. The test compares the null hypothesis of equidispersion of the data to the alternative of overdispersion. The test was run for a regression for which results are not presented here. Greene's test is a fairly general test of overdispersion. Under the null, the conditional mean is specified as stated above, namely $E[y \mid x] = \mu = \exp(x\beta)$. Under the alternative, the conditional mean is $E[y \mid x] = \mu + g(\mu)$. Given that μ is strictly positive in the Poisson regression model this test can only test positive departures from equidispersion.

Figure 4.3-4 Histogram of Consumption



The figure above shows the relationship of the variance σ^2 (1.08) to the mean μ (1.47) in the sample. The values for these two moments are given as the weighted values using the sample weights. The unweighted values maintain the relationship of mean greater than variance. This is suggestive of underdispersion. The statistic value from Greene's test strongly rejected the null. However, we do not conclude that the data is overdispersed because the test is not strictly applicable to underdispersion. An estimation of consumption using the negative binomial distribution was performed. The results showed the over dispersion parameter α to be not significantly different from zero. These results all suggest that overdispersion is not a problem in the data. Cameron and Trivedi (1998) suggest a test of the alternative that the model was specified as a Katz system. The Katz system is a generalization of the Poisson that allows for both over- and underdispersion.

Table 4.3-2 Consumption Regressions*

Variable	Non-Signal		Signal		Both	
	Coefficien	*Marginal*	*Coefficien*	*Marginal*	*Coefficient*	*Marginal*
Intercept	0.3731 (34.9)		0.4410 (39.8)		0.3899 (37.5)	
AGETRIED	-0.0079 (7.7)	-0.0117			-0.0083 (8.4)	-0.0123
HOW_LONG	0.0099 (106.7)	0.0147			0.0096 (100.4)	0.0143
FEMALE	-0.1246 (25.0)	-0.1795			-0.1230 (24.2)	-0.1771
HAS_KIDS	0.0602 (6.2)	0.0875			0.0606 (6.2)	0.0880
STUDENT	-0.3485 (25.5)	-0.4340			-0.3588 (26.9)	-0.4445
BLACK	-0.2884 (84.4)	-0.3860			-0.2825 (80.6)	-0.3783
HISPANIC	-0.3261 (97.0)	-0.4332			-0.3265 (97.2)	-0.4334
MRJFLAG	0.0975 (12.4)	0.1411			0.0928 (11.1)	0.1341
ALCFLAG	0.2041 (13.6)	0.2711			0.2093 (14.3)	0.2772
ALCDLYF	0.1772 (35.3)	0.2723			0.1798 (36.0)	0.2763
PHLEGMYR	0.2044 (29.0)	0.3219			0.1979 (26.7)	0.3106
CA	-0.2417 (61.4)	-0.3346			-0.2333 (56.4)	-0.3231
IL	-0.1553 (22.8)	-0.2165			-0.1502 (21.2)	-0.2095
TAXES	-0.4022 (6.7)	-0.5973			-0.4001 (6.6)	-0.5941
TRY_QUIT			-0.0760 (9.9)	-0.1147	-0.0541 (4.7)	-0.0785
CIGPROB0			0.3658 (15.7)	0.6519	0.3470 (13.7)	0.5901
CIGPROBX			-0.2181 (4.3)	-0.2972	-0.2560 (5.8)	-0.3306
Ln(L)	-6742		-7023		-6732	
Ln(L) Intercept	-7125		-7125		-7125	
N	5071		5071		5071	

* Wald statistics are in parentheses. See Appendix C for the definitions of the regressors. Note, the marginal effects for dummy variables are calculated slightly differently from those for continuous variables. See Cameron and Trivedi (1998, p. 80-82) equation (3.44) versus the equations on p. 82.

The results for the estimates of consumption are shown in the table above. As noted, we estimated three separate regressions: one with the socioeconomic variables, one with the signals and one with both sets of variables. We found that the signal variables (CIGPROBO, CIGPROBX) based on the experience (by the smoker) of at least one problem with cigarettes in the past year, entered significantly into the regressions with the correct sign. This result stands in contrast to the univariate analysis that showed a positive relationship between consumption and cigarette problems for symmetric smokers. The indicator (TRY_QUIT) for the symmetric state was also significant in both regressions. The marginal effects of the signal variables were comparable in magnitude to the other variables.

Other than the signal variables, the variable with the largest marginal effect was level of taxes (TAXES). This result was also unexpected. Univariate analysis of the relationship between the two variables showed no discernable pattern or possibly even a slight positive relationship. We believe that the relationship became significant when we controlled for the other factors in the model. We believe that regional dummy variables for the Los Angeles, California MSA (CA) and the Chicago, Illinois MSA (IL) were particularly important in this change in significance because the variation in tax rates was in part regional.

The results for taxes allow us to calculate an elasticity of response in consumption, ε_{tax}, to tax changes. We use a similar derivation from the income result given in the discussion about the participation equation.

$$\varepsilon_{Tax} = \frac{\partial E[y]}{\partial Tax}\frac{Tax}{E[y]} = \left(\frac{\partial E[y]}{\partial Tax}\frac{1}{E[y]}\right)Tax = (\gamma_{Tax})Tax \qquad \textbf{EQ. 4.3-4}$$

The elasticity is simply the coefficient of the tax variable from the regression, γ, times the tax rate, *Tax*, evaluated in this case at the mean in the sample. For our data, the tax coefficient from the third regression results in Table 4.3-2 is -0.4001 and the sample mean of the tax rate[7] is 0.0923. The tax elasticity is therefore approximately –0.037.

Earlier we noted that Cameron (1998) compared own-price elasticities from roughly twenty-five studies of demand that were static in nature. The reported elasticities range from zero to 1.44 with many in the range of 0.20-0.40. These figures somewhat larger than the result we have from our regression.

Yen (1999) found household size was positively related to consumption and education was negatively related to consumption. We also found this in our study although education (YEARSEDU) entered the participation rather than the consumption equation. We did not find household size to be significant but we suspect that this is due correlation with other more predictive regressors.

The rest of the variables all enter the regressions as significant. The second largest marginal effect on consumption came from student status. Students (STUDENT) appear to smoke significantly less than the average smoker. This may be due to personal income level in that population which is likely lower than the average in the sample. Women

---

[7] The tax rate was centered for use in estimation. The value 0.0923 results from the subtracting off the original sample mean, 58.1¢, from the value on each observation during data preparation. The result is not zero due to the weighting.

(FEMALE) are also seen to smoke less daily. Two race indicators enter the regression: one for Hispanics (HISPANIC) and one for African-Americans (BLACK). Both variables have negative signs indicating lower average daily consumption. This result is consistent with other studies (MMWR, 1994, No. SS-3).

Finally, we found that the indicators of use of other substances and the signal variables were strong predictors of the level of consumption. These are important results. Only one other known economic study (Dee, 1999) explicitly accounts for the complementary effects of multiple substance use. Dee (1999) examines the complementarity of alcohol and cigarette use. Dee approaches the problem of joint estimation of demand for the two substances and uses instrumental variable (IV) methods to perform the regressions. He notes that the IV approach with simultaneous equations is motivated by the observation that there are probably some unobserved determinants that effect both alcohol and cigarette demand. His results, however, are qualified due to the low significance levels.

We conclude that, even after accounting for the complementary effects of multiple substance use, there is evidence in our study for the type of behavior predicted by Bateson's theory and our associated model. The signal variables all entered the regressions with significant coefficients and with the expected signs. In the next chapter, we provide our conclusions from the study. We also discuss areas for future research.

5 CONCLUSIONS

Overall, we have found modest evidence for the relationship that we hypothesized between the symmetric and complementary states. The weak evidence for the relationship may be for various reasons. First, the model itself may not be correct. We feel that there is enough empirical evidence from this initial study to warrant further research before discarding the model entirely. Below we discuss the areas for future investigation of the model.

5.1 Summary

An area that warrants further investigation is the causality in the model. We have estimated pooled cross-sectional equations for the model. Unfortunately, such a methodology does not allow a full examination of causation. Granger (1969) and Sims (1972) provide methods to measure causation in behavioral equations. However, their methods require times series information. Specific to our research, we would need information about the timing of the observed events in the model. We would need at the least to know when a signal was received by the addict (i.e., when did the argument with the spouse occur, or when did the addict hear about health risks from his or her medical provider) in relation to the time the consumption was observed.

A better understanding of causation is also suggested by the variables that were identified as signal variables. Our prior hypothesis was that the main signal variable in the NHSDA dataset would be the variable indicating problems due to cigarettes. The data showed a strong univariate relationship with the indicator of cigarette problems but not the one expected by the theoretical model. The odds ratios for the cigarette problem indicator are given in Table 5.1-1 below.

The odds ratio for "Tried" population (those that tried to reduce consumption) in the table should be less than one under the hypotheses of the addiction model. Instead, we find that the occurrence of problems attributed to smoking is in all cases is associated with a higher level of consumption. Oddly, the indicator was found to be significant in the regression and with the correct signs. We suspect this occurred because the regression included other confounding variables that had not been controlled for in the univariate analysis. These conflicting pieces of evidence leads us to conclude that further work needs to be done in the area of the causal relationships in the model. A new data source would be required in order to pursue the Granger measure of causation – the NHSDA is not currently administered and collected in way to allow this type of research.

Table 5.1-1 Association of Problems and Consumption

HalfPack	Odds Ratio[1]		Q_{MH} [2]		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
CIGPROBS	2.384	1.213	20.421	9.681	22.839

[1] Odds of smoking at least 1 pack per day (HALFPACK) for smokers that tried and did not try to quit in the 12 months prior to the survey interview.
[2] Mantel-Haenszel χ^2 statistic for testing the association between HALFPACK (4 d.f.) and CIGPROBS.

Another related area of potential research, is the refinement of the econometric techniques used on the current data. There are at least two ways the current econometric study could be refined. First, the method of instrumental variables could be used to correct for any bias inherent in the endogeneity of the regressor that identifies the symmetric state. With proper application of this method more confidence could be placed on the results.

Second, an empirical model form that allows for more flexibility in the conditional mean also may improve the strength of the results. Cameron and Trivedi (1998) note that count data models in the same family as the Poisson regression are particularly sensitive in their variance estimates to improperly specified conditional mean and variance functions. The authors do note however that the given the correct specification of the conditional mean, the estimator we used here will give consistent results even if the dependent variable is not distributed Poisson. In this case, the estimator is referred to as the Pseudo-Poisson Maximum Likelihood estimator (PMLE).

We commented in the previous chapter that Cameron and Trivedi (1998) suggest use of a test for the Katz system in order to identify the existence of underdispersion (and hence the appropriateness of the conditional function.) Both Murphy (1996) and Weiss (1997) tests for various departures from likelihood function used in the current study. A refinement of our results would report the values of the tests given in Murphy and/or Weiss.

The current study did not include information about governmental restrictions or tobacco control efforts except indirectly through the tax rate variable. An enhancement of the work done in our study would be the incorporation of such data. During the years we considered in our study, 1991-1993, nearly all states had enacted legislation related to smoking. Jacobsen, Wasserman and Anderson (1997) document that all but nine states as of 1995 prevented sales to minors. The authors also document that the majority of states had some regulation restricting smoking at worksites. The National Cancer Institute (2000) documents the large number of local regulations that restrict smoking activities. These references provide a rich source of information about the regulatory environment faced by smokers. Inasmuch as regulatory restrictions are signals to the smoker that smoking poses risk for themselves and society, these signals should be investigated in the framework we propose here. The challenge is how best to incorporate this information into an econometric analysis.

Several authors (Brown, 1995; Hu et al., 1995; Chaloupka, 1997) have attempted this effort in various ways. We initially considered the data used by Chaloupka from the National Cancer Institute (NCI, 1993). The NCI document Chaloupka used is an earlier release of the monograph we have cited elsewhere in this study. Our initial analysis using dummy variables for the restrictions found that the dummy variables were no more significant than dummy variables for the MSA of residence. Given the likelihood of additional unobserved regional differences in the regulatory environment we continued the study using only the dummy variables. A larger dataset with more regional variation might allow us to determine the effect of different regulations using dummy variables.

In the early 1990s, very few states had undertaken large-scale anti-smoking media campaigns. Two notable cases are Massachusetts and California. Given the limitations in our data (only four states were represented one of which was California) we did not pursue the use of data on media campaigns. Hu, Sung and Keeler (1995) did however and found lagged expenditures on media to be significant. They used a different approach to estimating demand however. They modeled aggregate demand in a time series model where media expenditures entered the model in a polynomial lag. Nonetheless, as more states initiate media campaigns there may be data available to study the effect at an individual level.

Brown (1995) also used measures of regulatory restrictions. Brown also studied macroeconomic demand in a demand and supply model of the U.S. tobacco market. He used an index of regulatory stringency that was produced by Wasserman et al. (1991) to predict the effect on demand. The index was a state by state measure that could be used in a cross-sectional analysis like the one that we have performed here.

In summary, the analysis of the microeconomic determinants of demand may be developed along the lines we pursued here. However, both technical challenges and data constraints confound the effort. We believe that there is a great deal that economists can glean from the research to date in the allied field of psychology.

APPENDICES

APPENDIX A. Derivations for the Model in Section 3.1

Total Differential of the Addiction Model: Unconscious Choice

The smoker has two choice problems: one that is perceived consciously and another that is, in part, unconscious. Behavior in the short-run is governed by unconscious choice. Planning, however, is based on the conscious choice problem. The short-run dichotomy between planned behavior and observed behavior stems from the differences in the two choice problems. We refer to the conscious choice problem as the *normative* problem. The unconscious two-period utility maximization problem is stated below:

$$\max U = \begin{array}{l} U_1\left(c_1,\, a_1,\, X\left(\Theta(a_1, s),\, \theta(a_1),\, z\left(\breve{U}_2, \hat{U}_2\right)\right)\right) + \\ \left(1 - \Theta(a_1, s)\right) \cdot \left[\left(1 - \theta(a_1)\right) \cdot \breve{U}_2(c_2, a_2) + \theta(a_1) \cdot \hat{U}_2(c_2, a_2)\right] \end{array}$$

EQ. A-1

$$\begin{aligned} &\textit{subject to}: \\ &\quad p \cdot a_1 + c_1 = W \\ &\quad p \cdot a_2 + c_2 = W \end{aligned}$$

where a is consumption of the addictive good, c is the consumption of the non-addictive good, X is the anxiety, p is the price of good a relative to good c, and W are the available assets. Prices and quantities of assets both are assumed fixed over time. $\breve{U}_2$ is utility obtained when no trauma occurs in the second period and $\hat{U}_2$ is the utility obtained when a trauma occurs. The values Θ and θ are the probabilities of death and trauma, respectively, and the quantity z is the amount of utility at stake, $\breve{U}_2 - \hat{U}_2$, in the event a trauma were to occur.

We assume that utility is quasi-concave or strictly concave and therefore both constraints on consumption will bind at the optimal solution of the EQ. A-1. We can therefore substitute our constraints into the utility function. Upon substitution we can write:

$$\max U = \begin{aligned} & U_1\left(W - p \cdot a_1,\ a_1,\ X\left(\Theta(a_1,s),\ \theta(a_1),\ z\left(\breve{U}_2, \hat{U}_2\right)\right)\right) + \\ & (1-\Theta(a_1,s)) \cdot \left[(1-\theta(a_1)) \cdot \breve{U}_2(W - p \cdot a_2, a_2) + \theta(a_1) \cdot \hat{U}_2(W - p \cdot a_2, a_2)\right] \end{aligned} \qquad \textbf{EQ. A-2}$$

where we have replaced the consumption, c, of the non-addictive good in each period with its optimal value, $W\text{-}p\cdot a_t$, which is a function of total assets and the consumption of the addictive good.

From EQ. A-2, we derive the total differential of utility as:

$$\begin{aligned} dU = & \frac{\partial U_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial X}\left[\frac{\partial X}{\partial \Theta}\left(\frac{\partial \Theta}{\partial a_1} da_1 + \frac{\partial \Theta}{\partial s} ds\right) + \frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1} da_1\right] + \\ & \frac{\partial U_1}{\partial a_1}\frac{\partial X}{\partial z}\left[\frac{\partial z}{\partial \breve{U}_2}\left(\frac{\partial \breve{U}_2}{\partial a_2} + \frac{\partial \breve{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right) - \frac{\partial z}{\partial \hat{U}_2}\left(\frac{\partial \hat{U}_2}{\partial a_2} + \frac{\partial \hat{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)\right]\frac{\partial a_2}{\partial a_1} da_1 + \\ & (1-\Theta)\left[(1-\theta)\left(\frac{\partial \breve{U}_2}{\partial a_2} + \frac{\partial \breve{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)\frac{\partial a_2}{\partial a_1} da_1 - \breve{U}_2\frac{\partial \theta}{\partial a_1} da_1\right] + \\ & (1-\Theta)\left[\theta\left(\frac{\partial \hat{U}_2}{\partial a_2} + \frac{\partial \hat{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)\frac{\partial a_2}{\partial a_1} da_1 + \hat{U}_2\frac{\partial \theta}{\partial a_1} da_1\right] - \left(\frac{\partial \Theta}{\partial a_1} da_1 + \frac{\partial \Theta}{\partial s} ds\right)\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right] \end{aligned} \qquad \textbf{EQ. A-3}$$

where X – the measure of anxiety – depends on two probabilities and, z, the difference between $\breve{U}_2$ and $\hat{U}_2$. We have allowed for a tolerance effect in the addictive good, a, as represented by the quantity $\frac{\partial a_2}{\partial a_1}$. The intra-period budget relation is reflected in, $\frac{\partial c_t}{\partial a_t}$, the substitution effect between the two goods, which in this case is $\frac{\partial c_t}{\partial a_t} = \frac{\partial}{\partial a_t}(W - p \cdot a_t) = -p < 0$.

The Policy Effect of Increasing Signals

We wish to study the effect of signals on the consumption of good a. We solve the total differential for $\frac{\partial a_1}{\partial s}$ the marginal effect of increasing signals. We will consider the case only of no tolerance although we could use the above result to derive the case including a tolerance effect. Under the case with no tolerance effect, the quantity $\frac{\partial a_2}{\partial a_1}$ is set to zero and from EQ. A-3 above we obtain:

$$
\begin{aligned}
0= \quad & \frac{\partial U_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial X}\left[\frac{\partial X}{\partial \Theta}\left(\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial \Theta}{\partial s}ds\right) + \frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}da_1\right] + \\
& \frac{\partial U_1}{\partial a_1}\frac{\partial X}{\partial z}\left[\frac{\partial z}{\partial \breve{U}_2}\left(\frac{\partial \breve{U}_2}{\partial a_2} + \frac{\partial \breve{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right) - \frac{\partial z}{\partial \hat{U}_2}\left(\frac{\partial \hat{U}_2}{\partial a_2} + \frac{\partial \hat{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)\right](0)da_1 + \\
& (1-\Theta)\left[(1-\theta)\left(\frac{\partial \breve{U}_2}{\partial a_2} + \frac{\partial \breve{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)(0)da_1 - \breve{U}_2\frac{\partial \theta}{\partial a_1}da_1\right] + \\
& (1-\Theta)\left[\theta\left(\frac{\partial \hat{U}_2}{\partial a_2} + \frac{\partial \hat{U}_2}{\partial c_2}\frac{\partial c_2}{\partial a_2}\right)(0)da_1 + \hat{U}_2\frac{\partial \theta}{\partial a_1}da_1\right] - \left(\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial \Theta}{\partial s}ds\right)\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]
\end{aligned}
$$

EQ. A-4

$$
\begin{aligned}
0= \quad & \frac{\partial U_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial X}\left[\frac{\partial X}{\partial \Theta}\left(\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial \Theta}{\partial s}ds\right) + \frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}da_1\right] + (0) + \\
& (1-\Theta)\left[(0) - \breve{U}_2\frac{\partial \theta}{\partial a_1}da_1\right] + (1-\Theta)\left[(0) + \hat{U}_2\frac{\partial \theta}{\partial a_1}da_1\right] - \left(\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial \Theta}{\partial s}ds\right)\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]
\end{aligned}
$$

$$0 = \frac{\partial U_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial X}\left[\frac{\partial X}{\partial \Theta}\left(\frac{\partial \Theta}{\partial a_1} da_1 + \frac{\partial \Theta}{\partial s} ds\right) + \frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1} da_1\right] + (1-\Theta)\left[\hat{U}_2 \frac{\partial \theta}{\partial a_1} da_1 - \breve{U}_2 \frac{\partial \theta}{\partial a_1} da_1\right] - \left(\frac{\partial \Theta}{\partial a_1} da_1 + \frac{\partial \Theta}{\partial s} ds\right)\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]$$

EQ. A-5

$$0 = \frac{\partial U_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1} da_1 + \frac{\partial U_1}{\partial X}\left[\left(\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1} da_1 + \frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial s} ds\right) + \frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1} da_1\right] + (1-\Theta)\left[\hat{U}_2 \frac{\partial \theta}{\partial a_1} da_1 - \breve{U}_2 \frac{\partial \theta}{\partial a_1} da_1\right] - \left(\frac{\partial \Theta}{\partial a_1} da_1\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right] + \frac{\partial \Theta}{\partial s} ds\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]\right)$$

$$0= \frac{\partial U_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}da_1 + \left[\left(\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial s}ds\right) + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}da_1\right] +$$

$$(1-\Theta)\left[\hat{U}_2\frac{\partial \theta}{\partial a_1}da_1 - \breve{U}_2\frac{\partial \theta}{\partial a_1}da_1\right] - \left(\frac{\partial \Theta}{\partial a_1}da_1\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right] + \frac{\partial \Theta}{\partial s}ds\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]\right)$$

$$0= \frac{\partial U_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}da_1 + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1}da_1 + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}da_1 - (1-\Theta)\left[\breve{U}_2 - \hat{U}_2\right]\frac{\partial \theta}{\partial a_1}da_1$$

$$-\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]\frac{\partial \Theta}{\partial a_1}da_1 - \left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]\frac{\partial \Theta}{\partial s}ds + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial s}ds$$

EQ. A-6

$$\left[\frac{\partial U_1}{\partial a_1} + \frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1} + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1} + \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1} - (1-\Theta)\left[\breve{U}_2 - \hat{U}_2\right]\frac{\partial \theta}{\partial a_1} - \left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right]\frac{\partial \Theta}{\partial a_1}\right]da_1 =$$

$$\left[\left[(1-\theta)\breve{U}_2 + \theta\hat{U}_2\right] - \frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\right]\frac{\partial \Theta}{\partial s}ds$$

The marginal effect is therefore:

$$\frac{\partial a_1}{\partial s}=\frac{\left[\left((1-\theta)\breve{U}_2+\theta\hat{U}_2\right)-\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\right]\frac{\partial \Theta}{\partial s}}{\left[\frac{\partial U_1}{\partial a_1}+\frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}+\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1}+\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}-(1-\Theta)\left[\breve{U}_2-\hat{U}_2\right]\frac{\partial \theta}{\partial a_1}-\left[(1-\theta)\breve{U}_2+\theta\hat{U}_2\right]\frac{\partial \Theta}{\partial a_1}\right]}$$

which may also be written as:

$$\frac{\partial a_1}{\partial s}=\frac{\left[\left((1-\theta)\breve{U}_2+\theta\hat{U}_2\right)-\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\right]\frac{\partial \Theta}{\partial s}}{\left[\frac{\partial U_1}{\partial a_1}+\frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}-\left[\left((1-\theta)\breve{U}_2+\theta\hat{U}_2\right)-\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\right]\frac{\partial \Theta}{\partial a_1}-\left[(1-\Theta)\left[\breve{U}_2-\hat{U}_2\right]-\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\right]\frac{\partial \theta}{\partial a_1}\right]} \quad \textbf{EQ. A-7}$$

We now consider the conditions under which the smoker would decrease current consumption under an increase in the signals to abstain. For a decrease, we must have $\frac{\partial a_1}{\partial s}<0$. Upon considering the signs of the terms in the numerator (see

Chapter 3), we find the numerator to be strictly positive. The denominator, however, can vary in sign. Therefore we require the following condition for a decrease in consumption motivated by the signals,

$$\frac{\partial U_1}{\partial a_1}+\frac{\partial U_1}{\partial c_1}\frac{\partial c_1}{\partial a_1}+\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \Theta}\frac{\partial \Theta}{\partial a_1}+\frac{\partial U_1}{\partial X}\frac{\partial X}{\partial \theta}\frac{\partial \theta}{\partial a_1}<\left((1-\Theta)\left[\breve{U}_2-\hat{U}_2\right]\frac{\partial \theta}{\partial a_1}+\left[(1-\theta)\breve{U}_2+\theta\hat{U}_2\right]\frac{\partial \Theta}{\partial a_1}\right) \quad \textbf{EQ. A-8}$$

APPENDIX B. Derivation of the Likelihood Equations

The current appendix provides the derivations of the first and second order conditions of the likelihood equations used in the estimation of the participation and consumption equations in our model.

The Participation Equation Derivatives

The participation equation was specified in Chapter 3 as an ordered logit function. The ordered logit measures the probability of occurrence, $\Pr[I = i]$, of one of several (more than two) mutually exclusive events, *i*. Furthermore, the events should have a natural ordering. The probability associated with each event may be constructed as the difference of two cumulative logistic distributions, namely $\Lambda(\tau_{i+1} - x\gamma) - \Lambda(\tau_i - x\gamma)$, where $\Lambda(\cdot)$ is the cumulative logistic distribution, *z* is the vector of regressors, and τ_i, τ_{i+1} and γ are constants to be estimated. The likelihood function for the ordered is the function of the products of the probabilities, however, in practice we use the loglikelihood. The loglikelihood, ℓ, for a single event is simply the natural logarithm of the probability.

To derive the maximum likelihood estimator for the ordered we solve for the values of τ and γ that solve the set of first derivative equations of loglikelihood set equal to zero. The first derivative of the loglikelihood of event *i* with respect to τ_{i+1} is:

$$\begin{aligned}\frac{\partial}{\partial \tau_{i+1}}\Pr[I=i] &= \frac{\partial}{\partial \tau_{i+1}}\ln\left(\Lambda(\cdot)-\Lambda(\cdot)\right)\\ &= \frac{1}{\left(\Lambda(\cdot)-\Lambda(\cdot)\right)}\frac{\partial}{\partial \tau_{i+1}}\left[\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right]\\ &= \frac{1}{\left(\Lambda(\cdot)-\Lambda(\cdot)\right)}\left[\frac{\partial}{\partial \tau_{i+1}}\Lambda(\tau_{i+1}-x\gamma)-0\right]\end{aligned} \qquad \textbf{EQ. B-1}$$

and,

$$\begin{aligned}\frac{\partial}{\partial \tau_{i+1}}\Pr[I=i] &= \frac{\partial \ell}{\partial \tau_{i+1}} = \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\left[\frac{\partial}{\partial \tau_{i+1}}\frac{1}{1+e^{-(\tau_{i+1}-x\gamma)}}\right]\\ &= \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\left[\frac{\partial}{\partial \tau_{i+1}}\left[1+e^{-(\tau_{i+1}-x\gamma)}\right]^{-1}\right]\\ &= \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\left[(-1)\left[1+e^{-(\tau_{i+1}-x\gamma)}\right]^{-2}(-1)e^{-(\tau_{i+1}-x\gamma)}\right]\\ &= \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\left[\frac{1}{1+e^{-(\tau_{i+1}-x\gamma)}}\frac{e^{-(\tau_{i+1}-x\gamma)}}{1+e^{-(\tau_{i+1}-x\gamma)}}\right]\\ &= \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\left[\frac{1}{1+e^{-(\tau_{i+1}-x\gamma)}}\right]\left[1-\frac{1}{1+e^{-(\tau_{i+1}-x\gamma)}}\right]\\ &= \frac{1}{\left(\Lambda(\tau_{i+1}-x\gamma)-\Lambda(\tau_i-x\gamma)\right)}\Lambda(\tau_{i+1}-x\gamma)\left(1-\Lambda(\tau_{i+1}-x\gamma)\right)\end{aligned} \qquad \textbf{EQ. B-2}$$

The first derivatives for τ_i and γ can be derived in a similar fashion.

The Consumption Equation Derivatives

The consumption equation was estimated under the assumption that consumption was distributed as a censored Poisson random variable. In Chapter 3, we defined the cumulative Poisson distribution as *H* where the density of *H* was:

$$h(y_k;x_k) = \frac{e^{-\mu_k}\mu_k^{y_k}}{y_k!} = \frac{e^{-e^{(x_k\beta)}}e^{(x_k\beta)^{y_k}}}{y_k!} = \frac{e^{(x_k\beta)^{y_k}}}{\exp(e^{(x_k\beta)})y_k!} \qquad \textbf{EQ. B-3}$$

where h is the density and μ is the mean function, $e^{x\beta}$, conditional on x, the vector of regressors, y is consumption, and β is the coefficient vector to be estimated. The probability of consuming a certain quantity, j, when the data is censored at some fixed point is therefore:

$$\Pr[y_k = j] = [h(y_k)]^{(1-d_k)} \cdot [1 - H(c-1)]^{d_k} \qquad \textbf{EQ. B-4}$$

where d is an indicator equal to one (1) when j is greater than or equal to c, the censoring point, and zero (0) otherwise. In the above equation, k is the observation number.

We also use MLE for estimating the equation for consumption. As in the participation we use the loglikelihood in the estimation. To find the solution we again use the first derivatives of the loglikelihood, ℓ. The first derivative of loglikelihood is:

$$\frac{\partial}{\partial\beta}\ln(\Pr[y_k = j]) = \frac{\partial\ell}{\partial\beta} = \frac{\partial}{\partial\beta}\ln([h(y_k)]^{(1-d_k)} \cdot [1 - H(c-1)]^{d_k})$$

$$= \frac{\partial}{\partial\beta}((1-d_k)\ln[h(y_k)] \cdot d_k \ln[1 - H(c-1)])$$

and,

$$\frac{\partial}{\partial\beta}\ell = \left((1-d_k)\frac{\partial}{\partial\beta}\ln[h(y_k)]\cdot d_k\frac{\partial}{\partial\beta}\ln[1-H(c-1)] \right)$$
$$= (1-d_k)\frac{\partial}{\partial\beta}\ln[h(y_k)]\cdot d_k\frac{\partial}{\partial\beta}\ln[1-H(c-1)] \qquad \textbf{EQ. B-5}$$

We simplify the expression in EQ. B-5 above by solving for the derivatives of ln(h) and ln(H) separately. The derivative of ln(h) is:

$$\frac{\partial}{\partial\beta}\ln(h) = y!\frac{-xe^{x\beta}e^{-e^{x\beta}}\left(\frac{e^{yx\beta}}{y!}\right)+e^{-e^{x\beta}}e^{yx\beta}y\left(\frac{x}{y!}\right)}{e^{-e^{x\beta}}e^{yx\beta}}$$
$$= y!\frac{-xe^{x\beta}\left(\frac{e^{yx\beta}}{y!}\right)+e^{yx\beta}y\left(\frac{x}{y!}\right)}{e^{yx\beta}} \qquad \textbf{EQ. B-6}$$
$$= y!\frac{-xe^{x\beta}\left(\frac{1}{y!}\right)+y\left(\frac{x}{y!}\right)}{1}$$
$$= x\left(-e^{x\beta}+y\right)$$

The derivative of ln(H) is:

$$\frac{\partial}{\partial\beta}\ln[1-H(c-1,x,\beta)] = -\frac{\frac{\partial}{\partial\beta}H(c-1,x,\beta)}{1-H(c-1,x,\beta)} \qquad \textbf{EQ. B-7}$$

where,

$$\frac{\partial}{\partial\beta}H(c-1,x,\beta)=\sum_{Y=0}^{c-1}\frac{\partial}{\partial\beta}h(Y,x,\beta)$$
$$=\sum_{Y=0}^{c-1}\left[-xe^{x\beta}e^{-e^{x\beta}}\left(\frac{e^{Yx\beta}}{Y!}\right)+e^{-e^{x\beta}}e^{Yx\beta}Y\left(\frac{x}{Y!}\right)\right]$$
$$=\sum_{Y=0}^{c-1}\left[\left(\frac{xe^{Yx\beta}}{\exp(e^{x\beta})Y!}\right)(-e^{x\beta}+Y)\right]$$
$$=-\left(\frac{xe^{cx\beta}}{\exp(e^{x\beta})c!}\right)(c)$$
$$=-xe^{x\beta}\left(\frac{(e^{x\beta})^{c-1}}{(c-1)!\exp(e^{x\beta})}\right)$$
$$=-xe^{x\beta}h(c-1,x,\beta)$$

EQ. B-8

We substitute EQ. B-8 into EQ. B-7. We then take that result and along with the result from EQ. B-6 we substitute into EQ. B-5 to obtain:

$$\frac{\partial}{\partial\beta}\ell=(1-d_k)\frac{\partial}{\partial\beta}\ln[h(y_k)]\cdot d_k\frac{\partial}{\partial\beta}\ln[1-H(c-1)]$$
$$=(1-d_k)[x(-e^{x\beta}+y)]\cdot d_k\left[\frac{xe^{x\beta}h(c-1,x,\beta)}{1-H(c-1,x,\beta)}\right]$$

EQ. B-9

We define the quantity $e^{x\beta}h/(1-H)$ as δ and the first derivative of the loglikelihood becomes:

$$\frac{\partial}{\partial\beta}\ell=[(1-d_k)[-e^{x\beta}+y]\cdot d_k\delta]x'$$

EQ. B-10

The second derivative of the loglikelihood follows immediately as:

$$\frac{\partial}{\partial\beta\partial\beta'}\ell = \left[(1-d_k)\left[-xe^{x\beta}\right] \cdot d_k \frac{\partial}{\partial\beta'}\delta \right] x' \qquad \textbf{EQ. B-11}$$

where we can show (using some lengthy algebra and the results from EQ. B-8) that the derivative of δ is:

$$\begin{aligned} \frac{\partial}{\partial\beta'}\delta &= \frac{\partial}{\partial\beta'}\frac{e^{x\beta}h(\cdot)}{1-H(\cdot)} = \frac{\frac{\partial}{\partial\beta'}\left[e^{x\beta}h(\cdot)\right]}{1-H(\cdot)} + \left[e^{x\beta}h(\cdot)\right]\frac{\partial}{\partial\beta'}[1-H(\cdot)]^{-1} \\ &= x\frac{e^{x\beta}h(\cdot)}{1-H(\cdot)}\left[(-e^{x\beta}+c) - \frac{e^{x\beta}h(\cdot)}{1-H(\cdot)}\right] \\ &= x\delta\left[(-e^{x\beta}+c)-\delta\right] \end{aligned} \qquad \textbf{EQ. B-12}$$

Substituting the results from the above equation into EQ. B-11 we find the second derivative of the loglikelihood to be:

$$\begin{aligned} \frac{\partial}{\partial\beta\partial\beta'}\ell &= \left[(1-d_k)\left[-xe^{x\beta}\right] + d_k\left(x\delta\left[(-e^{x\beta}+c)-\delta\right]\right)\right]x' \\ &= \left[(1-d_k)\left[-e^{x\beta}\right] + d_k\left(\delta\left[(-e^{x\beta}+c)-\delta\right]\right)\right]xx' \\ &= -\left[(1-d_k)e^{x\beta} + d_k\left(\delta\left[e^{x\beta}+\delta-c\right]\right)\right]xx' \end{aligned} \qquad \textbf{EQ. B-13}$$

APPENDIX C. Data Dictionary

The following table contains descriptions and reference information about the data used in the present study. The first two sections of the table describe the variables as they were used in the econometric regressions detailed in the chapters above. The remainder of the table describes the source data and gives a reference (where applicable) to the page number in the 1993 NHSDA Codebook where the reader can find a full description of the variable.

Ty	Variable	Description	NHSDA Source Element	NHSDA 1993
Participation Equation Regressors				
P	AGE	Age in Years Centered at 30		
P	ROOTA	Square Root of Age Centered at 5		
P	YEARS	Years of Education Centered at 12		
P	FEMAL	Indicator of Gender, Female=1		
P	MARRI	Indicator of Marital Status, Married=1		
P	LNINC	Natural Log of Real Household Income in $000		
P	LOWSO	Indicator of Residence in Low Socio Area, Yes=1		
P	STUDE	Indicator of Full-time Student Status, Student=1		
P	MRJFL	Indicator of Ever Marijuana Use, Use=1		
P	ALCFL	Indicator of Ever Alcohol Use, Use=1		
P	ALCDL	Indicator of Current Daily Alcohol Use, Use=1		
P	TREAT	Indicator of Attendance in Treatment Program, Yes=1		

Typ	Variable	Description	NHSDA Source Element	NHSDA 1993
P	LORISK	Indicator of Belief Smoking is Low Risk, Yes=1		
P	HISPAN	Indicator of Hispanic Origin, Hispanic=1		
P	SERVIC	Indicator of Ever Military Service, Military=1		
P	TRY_Q	Attempted to Quit Cigarettes in Past Year, Yes=1		
P	MRJPR	Indicator of Problems with Marijuana, Problems=1		
P	MRJPR	TRY_QUIT times MRJPROB0		
Consumption Equation Regressors				
C	AGETR	Age 1st Tried Cigarettes in Years Centered at 15.2		
C	HOW_L	AGE minus AGETRIED Centered at 20		
C	FEMAL	Indicator of Gender, Female=1		
C	HAS_KI	Indicator of Children, Yes=1		
C	STUDE	Indicator of Full-time Student Status, Student=1		
C	BLACK	Indicator of African-American Race, Afro-Am=1		
C	HISPAN	Indicator of Hispanic Origin, Hispanic=1		
C	MRJFL	Indicator of Ever Marijuana Use, Use=1		
C	ALCFL	Indicator of Ever Alcohol Use, Use=1		
C	ALCDL	Indicator of Current Daily Alcohol Use, Use=1		
C	PHLEG	Indicates had Bad Cough w/ Phlegm Past Year, Yes=1		
C	CA	Indicator of Residence in LA MSA, Yes=1		
C	IL	Indicator of Residence in Chicago MSA, Yes=1		
C	TAXES	Observed Excise Tax Per Pack		
General Dictionary				
A	HIRISK	Belief That Smoking 1 Pack/Day Is Very risky	rskpkcig	171
A	LORISK	Belief That Smoking 1 Pack/Day Is Not Risky	rskpkcig	171
A	ANAL	Final Person–Level Sample Weight	analwt	260
A	DIVISI	Census Major Division	division	225
A	METRO	Metropolitan Area – Raw Code	metro	225
A	MONT	Interview Month	intmonth	218

Ty	Variable	Description	NHSDA Source Element	NHSDA 1993
R	REGIO	Census Major Region	region	225
A	SURVE	Name And Year Of Survey	{data separated by	
A	YEAR	Year Of Survey	{data separated by	
D	HALFP	# ½ Packs Per Day In Past 30	avcig	4
D	OCCAS	Smoked, But Less Than Daily In The Past 30 Days	avcig	4
D	PARTIC	Participation For Ord. Logit: 0=Non,1=Occ,2=Daily	avcig	4
D	PASTM	Smoked At Least 1 Cigarette In Past Month	cigmon	250
D	PASTY	Used Cigarettes In The Past Year	cigyr	250
D	PSYST	Psychological State For Ord. Logit: 1=Sym 2=Quit	cigyr,ircdcig	250,307
D	TRY_Q	Tried To Cut Down On Cigarettes In Past Year? Y=1	ircdcig	307
I	ALCDL	Alcohol Daily/Almost Daily Flag	alcdlyf	256
I	ALCFL	Alcohol – Ever Used	alcflag	250
I	ALCMO	Alcohol – Past Month Use	alcmon	250
I	ALCYR	Alcohol – Past Year Use	alcyr	250
I	COCDL	Cocaine Daily/Almost Daily Flag	cocdlyf	257
I	COCFL	Cocaine – Ever Used	cocflag	251
I	COCM	Cocaine – Past Month Use	cocmon	252
I	COCYR	Cocaine – Past Year Use	cocyr	251
I	MRJDL	Marijuana Daily/Almost Daily Flag	mrjdlyf	257
I	MRJFL	Marijuana – Ever Used	mrjflag	251
I	MRJMO	Marijuana – Past Month Use	mrjmon	251
I	MRJYR	Marijuana – Past Year Use	mrjyr	251
I	SMKDL	Smokeless Tobacco Daily/Almost Daily Flag	smkdlyf	256
I	SMKFL	Smokeless Tobacco – Ever Used	smkflag	250
I	SMKM	Smokeless Tobacco – Past Month Use	smkmon	251
I	SMKYR	Smokeless Tobacco – Past Year Use	smkyr	251
I	ARGU	Arguments With People About Substance Use? Y=1	drargum	129
I	ASTHM	Ever Diagnosed W/ Lung/Chest Condition?	asthma	178

Typ	Variable	Description	NHSDA Source Element	NHSDA 1993
	BOOKE	Ever Arrested/Booked In Past 12 Months?	nobookyr	112
I	EFFCT	Needed More 4 Same Effect? Y=1	irefcig	307
I	JOBLO	Almost Lost Job To Drink In Past Year?	alcjob	121
I	PHLEG	Cough W/ Phlegm 3+ Weeks In Past Year?	phlegmyr	180
I	PLZ_Q	Someone Asked To Stop Drink:Y=1	iralcfrd,iralcmat,iralcre	312,313,
M	ALCDE	Felt Dependent On Alcohol?	irdpalc	308
M	ALCPR	Had Problems Due To Alcohol?	prb1alc	304
M	ALC_SI	Was Sick When Cut Down On Alcohol?	irwsalc	309
M	CIGDEP	Felt Dependent On Cigarettes?	irdpcig	308
M	CIGPR	Had Problems Due To Cigarettes?	prb1cig	304
M	COCDE	Felt Dependent On Cocaine?	irdpcoc	306
M	COCPR	Had Problems Due To Cocaine?	prb1coc	304
M	COC_SI	Was Sick When Cut Down On Cocaine?	irwscoc	309
M	FELTD	Felt Dependent On Alcohol,Cocaine,Marijuana?	irdpalc,irdpcoc,irdpmrj	308,306,
M	FELTSI	Was Sick When Cut Down On Alcohol,Coke,Pot?	irwsalc,irwscoc,irwsmj	309,309,
M	HADPR	Had Problems Due To Alcohol,Cocaine,Marijuana?	prb1alc,prb1coc,prb1m	304,304,
M	LOHEA	Fair/Poor Self–Reported Health	healthyr	180
M	MRJDE	Felt Dependent On Marijuana?	irdpmj	308
M	MRJPR	Had Problems Due To Marijuana?	prb1mrj	304
M	MRJ_SI	Was Sick When Cut Down On Marijuana?	irwsmj	309
M	TRTPS	Ever Treated For Psychological/Emotional Condition?	trpsych	178
M	WITHD	Was Sick When Cut Down On Cigarettes?	irwscig	309
P	AVGPR	Nominal $/Pack Reported From Tobacco Council		
P	CPI	CPI–U For MSA From BLS		
P	PCTTA	Tax As Percent Of Price		
P	PRICEC	Nominal 1Year $/Pack Change In Price		
P	RELCH	Change In $/Pack Deflated By Current CPI		
P	RELINC	Income $000 Deflated By CPI	totinc,irafood	247,235

Ty	Variable	Description	NHSDA Source Element	NHSDA 1993
P	RELPRI	$/Pack Deflated By CPI		
P	RELTA	$Tax/Pack Deflated By CPI		
P	RELTX	Change In $Tax/Pack Deflated By Current CPI		
P	TAXCH	Nominal 1Year $Tax/Pack Change In Tax		
P	TAXES	Nominal $Tax/Pack Reported From Tobacco Council		
R	EVERU	Ever Used Cigarettes? Y=1	cigflag	250
R	LNCIG	Log Of 1 Plus # Of Packs	avcig	4
R	QUITY	Stopped Smoking During Current Year? Y=1	cigyr,cigmon	250,250
R	QUIT_A	Tried To Cut Down On Alcohol? Y=1	ircdalc	307
R	QUIT_A	Tried To Cut Down On Cig,Alc,Coc, or Mrj? Y=1	ircdcig,ircdalc,ircdcoc,i	307
R	QUIT_C	Tried To Cut Down On Cocaine? Y=1	ircdcoc	307
R	QUIT_	Tried To Cut Down On Marijuana? Y=1	ircdmj	307
S	DROPC	Was Respondent Hostile To Interviewer?	coopint	301
S	DROPD	Was The # Of ½ Packs Missing Or Bad Data?	avcig	4
S	DROPF	Was The Data Unusable?	avcig,coopint,metro	4,301,22
S	DROP_	State/MSA Was Not CA, FL, IL, NY	metro	225
S	AGE	Age In Years	irage	226
S	AMERI	Is IRRACE=Native American? Y=1	irrace,race	226,244
S	ASIAN	Is IRRACE=Asian? Y=1	irrace,race	226,244
S	BLACK	Is Race=Black? Y=1	race	244
S	BUDGE	Ratio of $/Pack to Daily Income Per Person	indinc,totinc,totpeop,{	246,247
S	CA	Is Respondent from LA MSA? Y=1	metro	225
S	CHILD	Were Children present in Interview? Y=1	cpres	220
S	FEMAL	Is Respondent Female? Y=1	irsex	226
S	FL	Is Respondent from MIAMI MSA? Y=1	metro	225
S	HIGHS	Completed >= 12 Years Of School	irhsgrad	228
S	HISPAN	Is Race=Hispanic? Y=1	race	244
S	IL	Is Respondent from CHICAGO MSA? Y=1	metro	225

Typ	Variable	Description	NHSDA Source Element	NHSDA 1993
S	INCOM	Household Income In $000s	totinc,irafood	247,235
S	INCOM	Personal Income In $000s	indinc	246
S	INSUR	Has Private Health Insurance?	irpinsur	229
S	HAS_KI	Indicates Respondent has Children Y=1	ownchild,stepchld	193,246
S	LOWSO	Low Socioeconomic Census Block? Y=1	chicago,la,miami,newy	224,225
S	MARRI	Is Respondent Married? Y=1	irmarit	227
S	NY	Is Respondent From NEW YORK MSA? Y=1	metro	225
S	OWNM	Is Personal Income > 0? Y=1	indinc	246
S	PAREN	Were Parents present in Interview? Y=1	ppres	220
S	SERVIC	Ever In Military? Y=1	service	195
S	STATE	Imputed State Of Residence	chicago,la,miami,newy	224,225
S	STUDE	Full–Time Student? Y=1	student	190
S	TOTPE	Total People in Household	totpeop	191
S	UNDER	Is Respondent < 18 Yrs Old? Y=1	irage	226
S	WAGE	Wage Rate For Wage Earners	hours ,irawag1	203,230
S	WELFA	Is Someone in HH On Welfare? Y=1	welfare	248
S	WHITE	Is Race=White? Y=1	race	244
S	YEARS	Years Of Education	ireduc	228
T	CIGHE	Received Treatment For Cigarettes Past Year	cighelp	156
T	DRNKH	Received Treatment For Alcohol Past Year	drnkhelp	156
T	TREAT	Received Treatment For Alcohol/Drug In Past Year	drnkhelp,trmthelp	156,157
T	TRMTH	Received Treatment For Other Drug Past Year	trmthelp	157
U	ABSTAI	Ever Quit For 1+ Years?	{many incl. cigage}	2
U	AGEDA	Age 1st Used Cigarettes Daily	cigage	2
U	AGETR	Age 1st Tried Cigarettes	cigtry	2
U	AVCIG	# Cigarettes Smoked Per Day Last 30 Days – Raw	avcig	4
U	CANTS	Try Quit But Cannot In Past Year?	ucdcig	83
U	ILLICIT	Used Illicit Drug In Past Year (Other Than Coke/Pot)?	{many incl. halyr}	249

Type	Variable	Description	NHSDA Source Element	NHSDA 1993
U	LEVEL	Ever Had A Change In Daily Level Of Smoking?	avcig,packs	4,5
U	LIFEMT	# Of Months Smoked Daily In Lifetime	{many incl. cigtime}	4
U	LIFEPC	# Of Packs Smoked In Lifetime	{many incl. cigtime}	4

* Type Legend:

i) ATT: Attitudinal Variables
ii) AUX: Auxiliary Variables
iii) DEP: Dependent variables
iv) IND: Indicators of use
v) INF: Information/Signal Variables
vi) MOR: Morbidity Indicators
vii) PRC: Price and Income Measures
viii) REL: Dependent variable-related Indicators
ix) SEL: Sample selection Indicators
x) SOC: Socio-demographic Variables
xi) TRT: Treatment Indicators
xii) USE: Usage History Indicators

APPENDIX D. Association between Signal and Use Variables

The tables in this appendix detail the association between the signal variables, and the occurrence of use (occasional or daily) in the last 30 days (PASTMNTH) and daily consumption (HALFPACK), controlling for attempts to quit. The odds ratios are the proportional odds of either smoking in past month or consuming at least one pack per day depending on the table. The Q statistics are all Mantel-Haenszel (1959) measures of association distributed χ^2 with 1 degree of freedom except where noted. The null hypothesis for a test of the Q statistic is that there is no association. The critical value of for $\chi^2(1)$ is 3.84 with a 95% confidence level. The element PASTMNTH is coded from PARTICIP.

The tables below document the direction and strength of the average smoker's response to various signals. A column heading titled "Tried" refers to the individuals that attempted to reduce (as measured by either TRY_QUIT or QUIT_ANY) use of a substance in the past year. The *Q* statistic for all strata measures the average association across individuals who tried and those that did not.

See the data dictionary in Appendix C for definitions of the variables discussed below.

Table D-2 Association with PASTMNTH **Controlling for** TRY_QUIT**, All Obs.**

	Odds Ratio		Q_{MH}		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
ALCPROBS	3.072	1.243	219.827	2.388	154.180
ARGUMENT	4.204	0.996	273.708	0.001	132.241
BOOKED	2.714	1.197	49.515	0.573	32.949
CIGPROBS	18.402	1.036	352.132	0.064	63.782
MRJPROBS	2.875	0.631	44.525	4.486	12.087
PHLEGMYR	1.263	0.853	9.676	1.672	3.742
PLZ_QUIT	3.827	1.351	420.054	5.841	293.245
TAXRISE	0.900	1.030	6.833	0.112	4.971

Table D-3 Association with PASTMNTH **Controlling for** QUIT_ANY**, All Obs.**

	Odds Ratio		Q_{MH}		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
ALCPROBS	3.469	1.256	173.076	9.037	77.934
ARGUMENT	4.540	1.743	183.048	44.367	133.521
BOOKED	2.453	1.766	26.414	16.051	36.159
CIGPROBS	14.661	5.013	248.636	172.630	292.204
MRJPROBS	3.264	1.012	32.474	0.007	5.834
PHLEGMYR	1.272	1.294	8.830	10.800	19.628
PLZ_QUIT	4.802	1.291	335.150	15.558	125.632
TAXRISE	0.894	0.951	7.011	1.012	7.123

Table D-4 Association with PASTMNTH Controlling for TRY_QUIT, Ever Used[1]

	Odds Ratio		Q_{MH}		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
ALCPROBS	1.963	1.243	66.778	2.388	56.711
ARGUMENT	2.797	0.996	111.977	0.001	61.856
BOOKED	1.966	1.197	18.913	0.573	14.840
CIGPROBS	8.431	1.036	140.542	0.064	38.824
MRJPROBS	1.574	0.631	7.242	4.486	0.752
PHLEGMYR	1.106	0.853	1.599	1.672	0.136
PLZ_QUIT	2.389	1.351	145.624	5.841	123.849
TAXRISE	0.818	1.030	21.737	0.112	16.156

Table D-5 Association with HALFPACK Controlling for TRY_QUIT

	Odds Ratio[2]		Q_{MH}		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
ALCPROBS	1.020	0.882	0.009	0.568	6.622
ARGUMENT	1.309	0.965	2.753	12.964	15.196
BOOKED	0.751	0.811	0.366	0.213	0.001
CIGPROBS	2.384	1.213	20.421	9.681	22.839
MRJPROBS	1.257	0.782	0.006	0.317	0.157
PHLEGMYR	0.749	1.491	0.010	37.708	23.378
PLZ_QUIT	1.061	1.220	2.753	12.644	14.612
TAXRISE	1.033	1.176	0.783	2.849	0.395

[1] This sample was composed of only those respondents who indicated they tried cigarettes at some point – regardless of their current smoking status.

Table D-6 Association with HALFPACK **Controlling for** QUIT_ANY

	Odds Ratio[3]		Q_{MH}		Q_{MH}
	Not Tried	Tried	Not Tried	Tried	All Strata
ALCPROBS	0.890	0.932	0.185	1.049	0.431
ARGUMENT	0.974	1.067	0.397	17.755	16.168
BOOKED	0.508	0.877	0.034	0.020	0.002
CIGPROBS	2.714	1.213	23.780	9.796	22.696
MRJPROBS	1.322	0.878	0.042	0.069	0.108
PHLEGMYR	0.651	1.555	0.713	41.235	48.971
PLZ_QUIT	1.363	1.091	6.745	10.064	16.333
TAXRISE	1.112	1.118	0.035	0.939	0.378

[2] Odds of smoking at least 1 pack per day. The "All Strata" Q_{MH} has 4 degrees of freedom in this table.

[3] See note 1 above.

REFERENCES

Akerloff, George A., and William T. Dickens (1982), "The Economic Consequences of Cognitive Dissonance," The American Economic Review, June 1982, 72(3), 307-319.

American Psychiatric Association (1994), Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Washington, DC: American Psychiatric Association.

Anda, R.F. et al. (1999), "Adverse Childhood Experiences and Smoking During Adolescence and Adulthood," The Journal of the American Medical Association, 3 November 1999, 282(17), 1652-8.

Barendregt, Jan J., Luc Bonneux and Paul J. van der Maas (1997), "The Health Care Costs of Smoking," The New England Journal of Medicine, 7 October 1997, 337(15), 1052-57.

Barthold, Thomas A. and Harold M. Hochman (1988), "Addiction as Extreme-Seeking," Economic Inquiry, January 1988, 26, 89-106.

Bartlett, J. C. et al. (1994), "Medical-care Expenditures Attributable to Cigarette Smoking – United States, 1993," Morbidity and Mortality Weekly Report, 8 July 1994, 43(26).

Bass, Frank M. (1969), "A Simultaneous Equation Regression Study of Advertising and Sales of Cigarettes," Journal of Marketing Research, August 1969, 6, 291-300.

Bateson, Gregory (1987), "The Cybernetics of Self: A Theory of Alcoholism," in Steps Towards an Ecology of Mind, Second Edition. London: Jason Aronson Inc.

Baltagi, Badi H. and Rajeev K. Goel (1990), "Quasi-Experimental Price Elasticity of Liquor Demand in the United States: 1960-83," American Journal of Agricultural Economics, May 1990, 451-454.

Becker, Gary S., Michael Grossman and Kevin M. Murphy (1994), "An Empirical Analysis of Cigarette Addiction," American Economic Review, June 1994, 84, 396-417.

Becker, Gary S. and Kevin M. Murphy (1988), "A Rational Theory of Addiction," Journal of Political Economy, 96, 675-700.

Bickel, W. K., A. L. Odum and G. J. Madden (1999), "Impulsivity and Cigarette Smoking: Delay Discounting in Current, Never, and Ex-smokers," Psychopharmacology, October 1999, 146(4), 447-54.

Black, Robert A. (1991), "Endogenous Demand for Alcohol and Self-Command: A Model of the Temperance Solution," in Handbook of Behavioral Economics. JAI Press.

Borland, Ron (1997), "Tobacco Health Warning and Smoking-related Cognitions and Behaviors," Addiction, 92(11), 1427-35.

Boyer, Marcel (1983), "Rational Demand and Expenditure Patterns under Habit Formation," Journal of Economic Theory, 31, 27-53.

Brentville-Jensen, A. L. (1999), "Addiction and Discounting," Journal of Health Economics, 18, 393-407.

Brown, A. Blake (1995), "Cigarette Taxes and Smoking Restrictions: Impacts and Policy Implications," American Journal of Agricultural Economics, November 1995, 77, 946-951.

Bureau of Labor Statistics (2000), "Consumer Price Index – All Urban Consumer," from the Department of Labor Internet site, Summer 2000.

Cameron, A. Colin and Pravin K. Trivedi (1998), Regression Analysis of Count Data. New York: Cambridge University Press.

Cameron, Samuel (1998), "Estimation of the Demand for Cigarettes: A Review of the Literature," Economic Issues, September 1998, 3(2), 51-71.

Cameron, Samuel (2000), "Nicotine Addiction and Cigarette Consumption: A Psycho-Economic Model," Journal of Economic Behavior and Organization, 41, 211-219.

Chaloupka, Frank (1991), "Rational Addictive Behavior and Cigarette Smoking," The Journal of Political Economy, August 1991, 99(4), 722-742.

Chaloupka, Frank (1992), "Clean Air Laws, Addiction and Cigarette Smoking," Applied Economics, 24(2), 193-205.

Chaloupka, Frank (1998), "A Review of Economic Models of Habitual and Addictive Behavior and Their Empirical Applications to Cigarette Smoking," in Advances in Behavioral Economics, Vol. 3, L. Green and J. H. Kagel, Eds. Norwood, NJ: Ablex Publishing Corporation.

Chaloupka, Frank and Henry Weschler (1997), "Price Tobacco Control Policies and Smoking Among Young Adults," Journal of Health Economics, 16, 359-373.

Centers for Disease Control (1996), "Cigarette Smoking Before and After an Excise Tax Increase and an Anti-Smoking Campaign – Massachusetts, 1990-1996," Morbidity and Mortality Weekly Report, 8 November 1996, 45(44), 966-970.

Davidson, Russell and James G. MacKinnon (1993), Estimation and Inference in Econometrics. New York: Oxford University Press.

Davis, R. M. (1987), "Current Trends in Cigarette Advertising and Marketing," New England Journal of Medicine, March 1987, 316(12), 725-732.

Deaton, A. and J. Muellbaeur (1980), Economics and Consumer Behavior. New York: Cambridge University Press.

Dee, Thomas S. (1999), "The Complementarity of Teen Smoking and Drinking," Journal of Health Economics, 18, 769-783.

Dockner, Englebert J. and Gustav Feichtinger (1993), "Cyclical Consumption Patterns and Rational Addiction," American Economic Review, March 1993, 83, 256-263.

Earl, Peter E. (1990), "Economics and Psychology: A Survey," The Economic Journal, September 1990, 100, 718-755.

Elder, John P. et al. (1996), "Indepedent Evaluation of the California Tobacco Education Program," Public Health Reports, August 1996, 111, 353-358.

Farrelly, M. C. and J. W. Bray (1998), "Response to Increases in Cigarette Prices by Race/Ethnicity, Income, and Age Groups – United States, 1976-1993," Morbidity and Mortality Weekly Report, 31 July 1998, 47(29), 605-609.

Festinger, L. (1957), A Theory of Cognitive Dissonance. Evanston: Row, Peterson.

Flynn, Brian S. et al. (1992), "Prevention of Cigarette Smoking Through Mass Media Interventions and School Programs," American Journal of Public Health, 82(6), 827-834.

Flynn, Brian S. et al. (1994), "Mass Media and School Interventions for Cigarette Smoking Prevention: Effects 2 Years after Completion," American Journal of Public Health, 84(7), 1148-1150.

Galbraith, John W. and Murray Kaiserman (1997), "Taxation, Smuggling and the Demand for Cigarettes in Canada: Evidence from Time-series Data," Journal of Health Economics, 16, 287-301.

Gilpin, Elizabeth, Shirley W. Cavin and John P. Pierce (1997), "Adult Smokers Who Do Not Smoke Daily," Addiction, 92(4), 473-480.

Giovino, Gary A. et al. (1994), "Surveillance for Selected Tobacco-Use Behaviors – United States, 1900-1994," Morbidity and Mortality Weekly Report, 18 November 1994, 43(SS-3).

Glatt, Max M., Christopher C. H. Cook (1987), "Pathological Spending as a Form of Dependence," British Journal of Addiction, November 1987, 82, 1257-1258.

Granger, C. W. J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," Econometrica, 37, 424-438.

Grapendaal, Martin (1992), "Cutting Their Coat According to Their Cloth: Economic Behavior of Amsterdam Opiate Users," The International Journal of the Addictions, 27, 487-501.

Greene, William H. (1999), Econometric Analysis, Fourth Edition. Paramus: Prentice Hall.

Hamermesh, Daniel S. and Francis W. Hamermesh (1983) "Does Perception of Life Expectancy Reflect Health Knowledge?," American Journal of Public Health, August 1983, 73(8), 911-913.

Hamilton, James L. (1972), "The Demand for Cigarettes: Advertising, the Health Scare, and the Cigarette Advertising Ban," The Review of Economics and Statistics, November 1972, 54(4), 401-411.

Hammond, Peter J. (1975), "Changing Tastes and Coherent Dynamic Choice," Review of Economic Studies, 43(1), 159-173.

Houthakker, H. and L. D. Taylor (1966), Consumer Demand in the United States, Analyses and Projections. Cambridge, MA: Harvard University Press.

Hu, Teh-Wei, Sung, Hai-Yen, and Theodore Keeler (1995), "Reducing Cigarette Consumption in California: Tobacco Taxes versus an Anit-Smoking Media Campaign," American Journal of Public Health, September 1995, 85(9), 1218-1222.

Hughes, John R., K. Michael Cummings and Andy Hyland (1999), "Ability of Smokers to Reduce Their Smoking and Its Association with Future Smoking Cessation," Addiction, 94(1), 109-114.

Jones, Andrew M. (1989), "A Double-Hurdle Model of Cigarette," Journal of Applied Econometrics, March 1989, 4, 23-39.

Jones, Andrew M. (1999), "Adjustment Costs, Withdrawal Effects, and Cigarette Addiction," Journal of Health Economics, 18, 125-137.

Jacobsen, Peter D., Jeffrey Wasserman and John R. Anderson (1997), "Historical Overview of Tobacco Legislation and Regulation," Journal of Social Issues, 53(1), 75-95.

Kelly, G. A. (1955), The Psychology of Personal Constructs. New York: W. W. Norton & Co.

Khantzian, E. J. (1985), "The Self-Medication Hypothesis of Addictive Disorders: Focus on Heroin and Cocaine Dependence," American Journal of Psychiatry, 142, 1259-1264.

Kozlowski, Lynn T. et al. (2000), "Massachusetts' Advertising Against Light Cigarettes Appears to Change Beliefs and Behavior," American Journal of Preventive Medicine, 18(4), 339-342.

Labeaga, José M. (1993), "Individual Behavior and Tobacco Consumption: A Panel Data Approach," Health Economics, 2, 103-112.

Labeaga, José M. (1999), "A Double-hurdle Rational Addiction Model with Heterogeneity: Estimating the Demand for Tobacco," Journal of Econometrics, 93, 49-72.

Lewin, Shira B. (1996), "Economics and Psychology: Lessons For Our Own Day From the Early Twentieth Century," Journal of Economic Literature, September 1996, 34, 1293-1323.

Lewinsohn, Peter M., Paul Rohde and Richard Brown (1999), "Level of Current and Past Adolescent Cigarette Smoking as Predictors of Future Substance Use Disorders in Young Adulthood," Addiction, 94(6), 913-921.

Lewitt, E. M. and D. Coate (1982), "The Potential for Using Excise Taxes to Reduce Smoking," Journal of Health Economics, 1(1), 121-145.

Luce, B. R. and S. O. Schweitzer (1977), "The Economic Costs of Smoking Induced Illness," in National Institute on Drug Abuse Research Monograph Series 17, M. E. Jarvik et al., Eds. Rockville, MD: National Institute on Drug Abuse.

Ma, Grace Xuequin and Steve Shive (2000), "A Comparative Analysis of Perceived Risks and Substance Abuse Among Ethnic Groups," Addictive Behaviors, 25(3), 361-371.

Madden, Gregory and Warren K. Bickel (1999), "Abstinence and Price Effects on Demand for Cigarettes: A Behavioral Economic Analysis," Addiction, 94(4), 577-588.

Mantel, N. (1963), "Chi-square Tests with One Degree of Freedom: Extensions of the Mantel-Haenszel Procedure," Journal of the American Statistical Association, 58, 690-700.

Mantel, N. and W. Haenszel (1959), "Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease," Journal of the National Cancer Institute, 22, 719-748.

Marvel, Gregory A. and Barbara Reed Hartmann (1986), "An 'Economic' Theory of Addiction, Hypomania, and Sensation Seeking," The International Journal of the Addictions, 21, 495-508.

Michaels, Robert J. (1988), "Addiction, Compulsion, and the Technology of Consumption," Economic Inquiry, January 1988, 26, 75-88.

Morgenstern, Jon and Jeremy Leeds (1993), "Contemporary Psychoanalytic Theories of Substance Abuse: A Disorder in Search of a Paradigm," Psychotherapy, 30(2), 194-206.

Mullahy, John (1985), "Cigarette Smoking: Habits, Health Concerns, and Heterogeneous Unobservables in a Microeconometric Analysis of Consumer Demand," Ph.D. Dissertation, University of Virginia.

Mullahy, John (1997a), "Heterogeneity, Excess Zeros, and the Structure of Count Data Models," Journal of Applied Econometrics, 12(3), 337-350.

Mullahy, John (1997b), "Instrumental Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior," Review of Economics and Statistics, 586-593.

Murphy, Anthony (1996), "Simple LM Tests of Mis-specification for Ordered Logit Models," Economics Letters, 52, 137-141.

National Cancer Institute (2000), State and Local Legislative Action to Reduce Tobacco Use. Smoking and Tobacco Control Monograph No. 11. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute, NIH Pub. No. 00-4804.

National Institutes of Health (1993), Major Local Tobacco Control Ordinances in the United States. Smoking and Tobacco Control Monograph 3. Bethesda, MD: U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, NIH Pub. No. 93-3532.

National Institute on Drug Abuse (1990), National Household Survey on Drug Abuse, Main Findings 1990. Rockville, MD: National Institute on Drug Abuse, 1990.

Nelson, David E. et al. (1994), "Surveillance for Smoking Attributable Mortality and Years of Potential Life Lost, by State – United States, 1990," Morbidity and Mortality Weekly Report, 10 June 1994, 43(SS-1).

Oster, G., G. A. Colditz and N. L. Kelly (1984), "The Economic Costs of Smoking and Benefits of Quitting for Individual Smokers," Preventive Medicine, 13(4), 377-389.

Parrot, Andrew C. (1995), "Stress Modulation over the Day in Cigarette Smokers," Addiction, 90, 233-244.

Patton, G. C. et al. (1998), "The Course of Early Smoking: A Population-Based Cohort Study Over Three Years," Addiction, 93(8), 1251-1260.

Phlips, L. (1983), Applied Consumption Analysis. New York: Elsevier Science.

Pollack, R. A. (1968), "Consistent Planning," Review of Economic Studies, 35(2), 201-208.

Pucci, L. G. and M. Siegel (1999), "Features of Sales Promotion in Cigarette Magazine Advertisements, 1980-1993: An Analysis of Youth Exposure in the United States," Tobacco Control, 8(1), 29-36.

Salandha, Fernando B. and Pierre-Jean Lancry (1991), "Psychological Opponent Processes and Consumer Behavior," in Handbook of Behavioral Economics, Volume 2B. Stanford: JAI Press.

Schaef, Anne W. (1987), When Society Becomes An Addict. San Francisco: Harper Collins.

Schechter, Mordechai (1988), "Incorporating Anxiety Induced by Episodes in Life Valuation," in Applied Behavioral Economics, Vol. 2, S. Maital, Ed. New York: NYU Press.

Schneider, L., B. Klein and K. M. Murphy (1981), "Government Regulation of Cigarette Health Information," Journal of Law and Economics, 24, 575-612.

Schoenberg, E. H. (1933), "The Demand Curve for Cigarettes," Journal of Business, 6(1), 15-35.

Shelton, Dana M. et al. (1998), "State Laws on Tobacco Control – United States, 1995," Morbidity and Mortality Weekly, 44(SS-6), 1-28.

Siegel, Michael and Lois Biener (1997), "Evaluating the Impact of Statewide Anti-Tobacco Campaigns: The Massachusetts and California Tobacco Control Programs," Journal of Social Issues, 53(1), 147-168.

Sims, C. A. (1972), "Money, Income and Causality," American Economic Review, 62, 540-552.

Solomon, R. L. and J. D. Corbit (1974), "An Opponent Process Theory of MotivationL I. Temporal Dynamics Affect," Psychological Review, 81(2), 119-135.

Stavrinos, V. G. (1987), "The Effects of an Anti-Smoking Campaign on Cigarette Consumption: Empirical Evidence from Greecce," Applied Economics, 19(3), 323-329.

Stigler, George J. and Gary S. Becker (1977), "De Gustibus Non Est Disputandum," American Economic Review, March 1977, 67, 76-90.

Substance Abuse and Mental Health Services Administration (1991), National Household Survey on Drug Abuse: 1991. Rockville, MD: U. S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration (1992), National Household Survey on Drug Abuse: 1992. Rockville, MD: U. S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration (1993a), National Household Survey on Drug Abuse: 1993. Rockville, MD: U. S. Department of Health and Human Services.

Substance Abuse and Mental Health Services Administration (1993b), National Household Survey on Drug Abuse, Public Release Codebook 1993. Rockville, MD: U. S. Department of Health and Human Services.

Suranovic, Steven M., Robert S. Goldfarb and Thomas C. Leonard (1999), "An Economic Theory of Cigarette Addiction," Journal of Health Economics, 18, 1-29.

Taylor, Lester D. (1986), "Opponent Processes and the Dynamics of Consumption," in Economic Psychology: Intersections in Theory and Practice, Alan J. and Heather W. MacFadyen, Eds. New York: Elsevier Science.

Taylor, Lester D. (1988), "A Model of Consumption and Demand Based on Pyschological Opponent Processes," in Psychological Foundations of Economic Behavior, Paul J. Albanese, Ed. Westport: Greenwood Publishing Group, Inc.

Tegene, A. (1991), "The Kalman Filter and the Demand for Cigarettes," Applied Economics, 23(7), 1175-1182.

The Tobacco Institute (1995), The Tax Burden on Tobacco: Historical Compilation 1995. Washington, D.C.: The Tobacco Institute.

Unger, Jennifer B. and Xinguang Chen (1999), "The Role of Social Networks and Media Receptivity in Predicting Age of Smoking Initiation: A Proportional Hazards Model of Risk and Protective Factors," Addictive Behaviors, 24(3), 371-381.

United States Census Bureau (1999), Statistical Abstract of the United States: 1999. Washington, D. C.: U. S. Department of Commerce.

Viscusi, W. Kip (1979), Employment Hazards. Cambridge, MA: Harvard College.

Viscusi, W. Kip (1992), Smoking: Making the Risky Decision. New York: Oxford University Press.

Warner, K. E., Hodgson T. A. and C. E. Carroll (1999), "Medical Costs of Smoking in the United States: Estimates, Their Validity, and Their Implications," Tobacco Control, Autumn 1999, 8(3), 290-300.

Wasserman, J., W. G. Manning and J. P. Newhouse (1991), "The Effects of Excise Tax and regulations on Cigarette Smoking," Journal of Health Economics, 10, 43-64.

Weinstein, Milton C., Donald S. Shepard and Joseph S. Pliskin (1980), "The Economic Value of Changing Mortality Probabilities: A Decision-Theoretic Approach," Quarterly Journal of Economics, March 1980, 94(2), 373-396.

Weiss, Andrew A. (1997), "Specification Tests in Ordered Logit and Probit Models," Econometric Reviews, 16(4), 361-391.

Yen, Steven T. (1999), "Gaussian versus Count-Data Hurdle Models: Cigarette Consumption by Women in the U.S.," Applied Economic Letters, 6, 73-76.