Smoking Cessation and Cigarette Tax Avoidance

Margaret Loudermilk
Institute for Health Research and Policy
University of Illinois at Chicago

January 17, 2008

American Statistical Association
International Conference on Health Policy Statistics

Send correspondence to: mlouderm@uic.edu

Objectives

- Examine smoking cessation and cigarette tax avoidance behaviors using the International Tobacco Control Survey (ITCS)

- Do cigarette taxes affect smoking participation?
  - Who is affected?
  - What is the magnitude of the effect?

- How do cigarette taxes affect consumer purchasing behavior and are the effects economically meaningful?

Along the way: Does survey attrition create bias toward heavy smokers and those least likely to quit?

Send correspondence to: mlouderm@uic.edu

Previous Study Methodologies and Results

Evidence has been mixed on the ability of cigarette taxes to affect smoking participation and whether the magnitude of any effects are economically meaningful.

- Evans & Farrelly (1998) - Investigate the compensating behavior of smokers
  - Use NHIS, 1979-1987 - little change in taxes over this period, prior to tobacco settlements
  - Higher taxes may lead to reduced cigarette consumption but the same tar and nicotine intake
  - Find evidence of consumption underreporting and expect the largest underreporting by heaviest smokers
  - Estimates show compensating behavior is so large that average daily intake is unaffected by taxes

- Tauras & Chaloupka (1999) - Investigte the effect of prices and tobacco control policies on cessation for young adults
  - Use Monitoring the Future Surveys - longitudinal data
  - Find price and restricting smoking in the workplace increase the probability of cessation

- Sheu, et al. (2004) - Investigate the effect of a large cigarette price change in California
  - Use BRFSS, 1996-1999 - 3 metro areas in CA
  - Find price did not have an effect on reducing estimated prevalence of smoking - includes both smokers and nonsmokers

- DeCicca & McLeod (2007) - Expolit large tax increases after 2001 to identify tax effects on older adult smoking
  - Use BRFSS, 2000-2005
  - Find price did not have an effect on reducing estimated prevalence of smoking - includes both smokers and nonsmokers
  - Find older adult smoking fell substantially in response to large tax increases

With one exception, these are repeated cross-section data sets and can't account for

- **dynamics, as in the rational addiction framework,**

- **individual specific time invariant effects.**

Data - International Tobacco Control Survey (ITCS)

1. All 5 waves of the U.S. sample, 2002-2006
2. Annual telephone survey of adults (18-92, mean age 44)
3. 79% of those recruited have been surveyed
4. Individuals are added to the survey in each wave to offset attrition (a.k.a. - "replenishment samples")

"Demographics"	Percent of Respondents
Female	56%
Married	45%
Employed	53%
Daily smokers	86%
Intend to quit	71%
Ever tried to quit	82%
Consider themselves addicted	100%
Attempted to quit during survey	8%
Have a negative opinion of smoking	52%

Data - Monitoring the Future (MTF) Tobacco Policy Data

- Clean/smoke-free air index - measure of restriction on smoking in public and private facilities
- State tobacco prices - 6 month average price of a pack of cigarettes (cents per pack)
- Tobacco control expenditures - combined state and federal expenditures per capita
- State excise tax on tobacco - 6 month average excise tax rate (cents per pack)

Year	Tax	Expenditure	Price
2002	51.93	2.92	378.77
2003	70.12	2.71	413.75
2004	75.78	2.21	417.16
2005	83.04	2.21	426.86
2006	91.74	2.18	440.31

Cross-Sectional Analysis: Pooled Ordinary Least Squares & Probit

Table 1: Pooled Ordinary Least Squares Estimates with Demographics

Quit Smoking	Coeff. Estimate
State tax	**-0.0005** (0.0002)
Tobacco control expenditure	0.0019 (0.0014)
Infrequent smoker	**-0.103** (0.005)
Health index	**0.009** (0.003)
Number of friends that smoke	**-0.019** (0.002)
Smoke at home	**-0.091** (0.007)
Price	**0.0002** (0.0002)

Red indicates significance at the 99% level, blue at 95%, bold at 90%.

Income variables were all statistically insignificant.
Year and Northeast region were positive & statistically significant at 99%.
Divorced and high school education were negative & statistically significant at 90%.
Married was negative & statistically significant at 99%.

Table 2: Cross-Section Probit Estimates

Quit Smoking	Coeff. Est.	Marginal Effect
State tax	**-0.004**	**-0.0004**
	(0.001)	(0.0002)
Tobacco control expenditure	0.013	0.001
	(0.010)	(0.001)
Health index	**0.064**	**0.008**
	(0.022)	(0.003)
Number of friends that smoke	**-0.141**	**-0.017**
	(0.013)	(0.002)
Smoke at home	**-0.599**	**-0.084**
	(0.044)	(0.007)
Price	**0.002**	**0.0002**
	(0.001)	(0.0001)

Red indicates significance at the 99% level, blue at 95%, bold at 90%.

Education and income variables were all statistically insignificant.
Year and Northeast region were positive & significant at 99%.
Married and divorced were negative & statistically significant.

Panel Construction and Attrition

The panel is constructed assuming that the first survey for each individual is period 1 & attrition is an absorbing state $\Rightarrow$ 25,940 observations on 5,188 individuals

Construct an indicator variable, s_{it}, that is equal to 1 if an individual has data for year t and is 0 otherwise.

Table 3: ICTS Attrition

Number of Waves Surveyed	Percent of Respondents
1	49.25
2	24.61
3	12.45
4	6.36
5	7.32

- In the existing literature, usually balanced subpanels, aggregate level data, or time averages are used.
  - Using a balanced subpanel may be a problem if individual attrition is related to cessation and other smoking behaviors.
  - Even if selection is random or ignorable, using a balanced subpanel is inefficient since it is, in effect, throwing away data.

- To produce consistent estimates on an unbalanced panel, selection may be related to exogenous explanatory variables or unobserved individual characteristics but may not be correlated with the error term.

- Variable addition tests can be used to test for selection. For attrition, a lead of the selection indicator, $s_{i,t+1}$, is included as a regressor and tested for significance.

- Testing for sample selection was not performed using POLS because if unobserved effects are omitted but are correlated with selection, inference may be misleading.

Table 4: First-Difference Estimates Including Lead Selection Indicator

Quit Smoking	Coeff. Estimate
State tax	-0.0004 (0.0005)
Tobacco control expenditure	**-0.010** (0.005)
Infrequent smoker	**-0.292** (0.023)
Health index	**0.015** (0.006)
Number of friends that smoke	**-0.016** (0.003)
Smoke at home	**-0.122** (0.014)
Price	.0001 .0004
Lead Selection Indicator	**-0.018** (0.010)

Red indicates significance at the 99% level, blue at 95%, bold at 90%.
Education, marital status, and region variables were all statistically insignificant.
Income variables all had positive coefficients but most were insignificant.
Year had a positive coefficient and was statistically significant.

$\Rightarrow$ **Selection is NOT "ignorable" & those remaining in the sample are less likely to quit.**

A Model of Correlated Outcomes

$$
\begin{aligned}
y_{it} &= 1(y_{it}^* > 0) && (1)\\
y_{it}^* &= x_{it}\beta + z_{it}\gamma + c_i + e_{it}; \quad t = 1, ..., T_i && (2)\\
s_{it} &= 1(s_{it}^* > 0) && (3)\\
s_{it}^* &= y_{it}^*\delta + w_{it}\phi + u_{it}; \quad t = 2, ..., T_i && (4)
\end{aligned}
$$

y = quitting/tax avoidance behavior
x = taxation, z = outcome controls
w = "cost of survey"

The timing of outcome and selection leads to a different densities for t=1 and t $\geq$ 2.

For t=1

$$
\begin{aligned}
e_{it}|(x_{it}, z_{it}, c_i) &\sim N(0, \sigma_e^2) && (5)\\
P(y_{i1} = 1|x_{i1}, z_{i1}, c_i) &= \Phi(x_{i1}\beta_e + z_{i1}\gamma_e + c_i) && (6)\\
f_1(y_{i1}|x_{i1}, z_{i1}, c_i) &= \Phi(x_{i1}\beta_e + z_{i1}\gamma_e + c_i)^{y_{i1}} \left[1 - \Phi(x_{i1}\beta_e + z_{i1}\gamma_e + c_i)\right]^{(1-y_{i1})} && (7)
\end{aligned}
$$

For t $\geq$ 2

$$\begin{aligned} P(s_{it}=1|x_{it},z_{it},c_i,w_{it}) &= P(y_{it}^{*}\delta + w_{it}\phi + u_{it} > 0) \\ &= P((x_{it}\beta + z_{it}\gamma + c_i + e_{it})\delta + w_{it}\phi + u_{it} > 0) \\ &= P((x_{it}\beta + z_{it}\gamma + c_i)\delta + w_{it}\phi + v_{it} > 0) \end{aligned} \tag{8}$$

where $v_{it} = \delta e_{it} + u_{it}$. Let $\begin{pmatrix} e_{it} \\ u_{it} \end{pmatrix} \sim N\left(0,\ \begin{pmatrix} \sigma_e^2 & \rho\sigma_e\sigma_u \\ \rho\sigma_e\sigma_u & \sigma_u^2 \end{pmatrix}\right)$

$$\Rightarrow \begin{pmatrix} e_{it} \\ v_{it} \end{pmatrix} \sim N\left(0,\ \begin{pmatrix} \sigma_e^2 & \delta\sigma_e^2 + \rho\sigma_e\sigma_u \\ \delta\sigma_e^2 + \rho\sigma_e\sigma_u & \delta^2\sigma_e^2 + \sigma_u^2 + 2\delta\rho\sigma_e\sigma_u \end{pmatrix}\right)$$

$$\begin{aligned} P(y_{it}=1|x_{it},z_{it},c_i) &= P(s_{it}=1|x_{it},z_{it},c_i,w_{it})P(y_{it}=1|x_{it},z_{it},c_i) \\ &\quad + P(s_{it}=0|x_{it},z_{it},c_i,w_{it})P(y_{it}=1|x_{it},z_{it},c_i) \end{aligned} \tag{9}$$