Nick Pittman and Lauren DeRosa
Economics 312: Econometrics
Jeff Parker
May 4, 2012

The Effects of State Tobacco Taxes on Cigarette Consumption

Introduction

In this report we analyze how variation in state level taxes affects the consumption of tobacco between states. We look at 48 US states from 1991 to 2007 to determine how differences in tax policy and other relevant variables across states and time is related to the number of packs of cigarettes consumed in each state.

We use a fixed effects model and a fixed effects first-order autoregressive model to represent this relationship. These models allow the intercepts to change for each state, such that the effects of taxes and other variables are more discernable. Since the data are time-series, we use the autoregressive model to mitigate the problem of autocorrelation of error terms.

Data

Our data for this project come from three main sources: our data on consumption come from the 2007 report "The Tax Burden on Tobacco," issued by the economic consulting firm Orzechowski and Walker. Our data on all of the other tobacco-related variables come from the website impactteen.org, a group dedicated to "healthier youth behavior," and our demographic data, including GSP per capita, were provided by Jon Rork. A summary table of the variables that we ended up using is displayed below:

Variable	Obs	Mean	Std. Dev.	Min	Max
packs	816	87.61397	28.23571	31.8	186.8
stextpp	816	.5644363	.4403261	.03	2.76
ppp	816	3.379853	.9349012	1.949245	6.648579
alciati	768	13.65625	6.723759	0	31
pup	768	1.576823	1.067151	0	3
ctrlpr	722	47.83663	411.6242	.084	9185
ctrlfund	726	2.278209	3.289323	.01	20.71
pct85	816	.0157083	.0037783	.0062	.0272778
pct65	816	.127781	.0165227	.085	.1857
pcgsp	816	.0387475	.0074569	.0242069	.0901343
kids	816	.1931787	.0235755	.1387294	.3102661
unem	816	5.068382	1.371875	2.2	11.3
hs	815	83.24987	5.232464	66.14286	93
college	815	24.08287	5.028946	12.64286	40.4
tctpp	815	.936292	.4627497	.321	3.194
revenue	816	230.8458	280.6485	6.473	1493.286
freeair	816	.3566176	.4792943	0	1
lsttax	816	-.8713087	.8440629	-3.506558	1.015231
lttax	815	-.1638965	.4283902	-1.136314	1.161274

The descriptions of these various variables are given here:

packs: Consumption of packs of cigarettes per year per capita.
stextpp: State excise tax per pack, adjusted for inflation.
ppp: Average price per pack, adjusted for inflation.
alciati: A composite variable representing how many and how strict laws are in place regarding youth consumption of tobacco, with a higher number representing stricter laws.
pup: Possession Use Purchase index, represents the number of possession, use, and/or purchase laws for a given state in a given year.
ctrlpr: Total state tobacco control program funding, adjusted for inflation.
ctrlfund: Tobacco control funding per capita, adjusted for inflation.
pct85: Fraction of the population 85 or older.
pct65: Fraction of the population 65 or older.
pcgsp: Per capita, gross state product (GSP), measured in millions of dollars per person.
kids: Fraction of the population age 5-17.
unem: Unemployment rate.
hs: Percentage of population with a high school degree.
college: Percentage of population with a college degree.
tctpp: Total cigarette tax per pack (that is, including federal taxes), adjusted for inflation.
revenue: Tobacco tax revenue, adjusted for inflation.
freeair: Dummy variable indicating if a state has any smoke free preemption laws.
lsttax: The log of state excise tax, adjusted for inflation.
lttax: The log of total tax, adjusted for inflation.

Theory & Models

Theoretical Expectations

Theoretically, we expect that increases in tobacco taxes will reduce the number of cigarettes consumed in a given state. We expect that the average price per pack, the number of laws limiting the use of tobacco, unemployment, percent of the population between 5 and 17, and the average education level will be negatively correlated with tobacco consumption. We would expect that per capita GSP would be positively correlated with tobacco consumption since there is no reason to think that tobacco is an inferior good. It is unclear what the effect of the percent of the population over 65 and over 85 would be, if there is any effect.

Specifications:

Equation 1:

$$\boldsymbol{PACKS_{it} = \beta_{1i} + \beta_2 ln(STEXTPP)_{it} + \beta_3 PPP_{it} + \beta_4 CTRLFUND_{it} + \beta_5 ALCIATI_{it} + \beta_6 PUP_{it} + \beta_7 PCT85_{it} + \beta_8 PCT65_{it} + \beta_9 KIDS_{it} + \beta_{10} UNEM_{it} + \beta_{11} PCGSP_{it} + \beta_{12} HS_{it} + \beta_{13} COLLEGE_{it} + e_i.}$$

This is our first regression equation using a fixed effects model with US state as the cross-sectional variable, where individual states are denoted with the subscript *i* and time is denoted by the subscript *t*. We use the natural log of state excise tax per pack because the scatter plot of packs against tax shows that this is a more accurate description of this relationship.

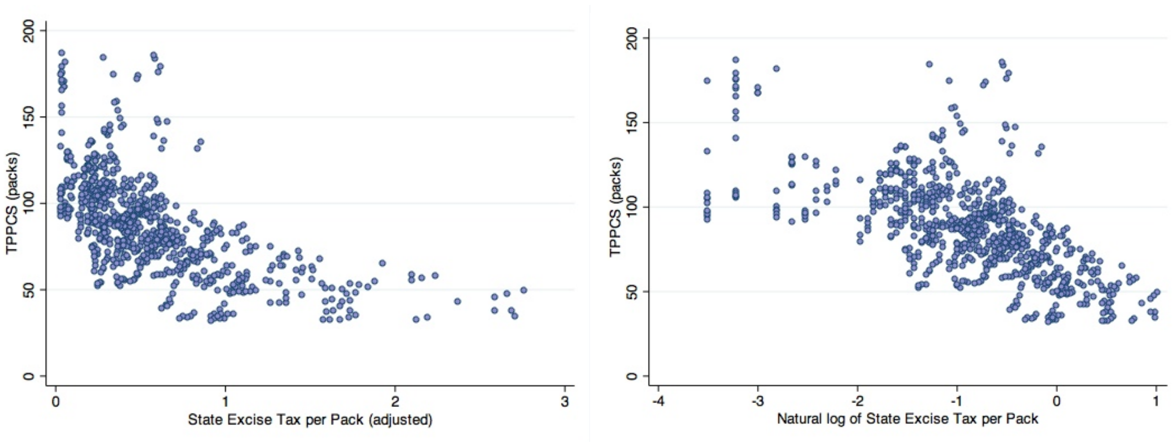

Figure 1: Scatter plots showing the relationship between tobacco consumption and taxes. The transformation of state excise tax per pack allows us to use linear regression to estimate coefficients.

Using this model (Equation 1), we do not find statistically significant results for a number of variables. The results from this regression are given in the first column of Table 1, and the Stata output can be found in Appendix 1.A. The effects of the percent of the population that are above 65, above 85, have graduated from high school, or are unemployed are not significantly different from zero, so we omit these variables in the next iteration of the model. Although it is just almost statistically significant, we include the variable *KIDS* in this model since the p-value in the first model for this variable is 0.102. We can say with 89.8% certainty that the higher the percentage of the population is between the ages of 5 and 17, the lower tobacco consumption will be.

The only unexpected result from this model is that state funding on tobacco control is positively correlated with tobacco consumption. This might point to the endogeneity problem that is an issue with our analysis (see Internal Validity). States which believe they have a relatively large amount of tobacco consumption may be directing more funding towards anti-smoking campaigns.

Using the results from the first model, we refine our specification:

Equation 2:

$$\boldsymbol{PACKS_{it} = \beta_{1i} + \beta_2 ln(STEXTPP)_{it} + \beta_3 PPP_{it} + \beta_4 CTRLFUND_{it} + \beta_5 ALCIATI_{it} + \beta_6 PUP_{it} + \beta_7 KIDS_{it} + \beta_8 PCGSP_{it} + \beta_9 COLLEGE_{it} + e_i.}$$

The results are presented in the second column of Table 1, and the Stata output can be found in Appendix 1.B. In this model we find that all the coefficients have the theoretically expected signs, but for *CTRLFUND*:

lsttax: a 10% increase in state sales tax will lead to approximately 0.93 less per capita consumption of packs.
ppp: a one dollar increase in the average price of a pack of cigarettes will lead to approximately 7.6 fewer packs per capita being consumed annually.
ctrlfund: Oddly, a one million dollar increase in tobacco funding leads to a rise in per capita pack consumption of 0.28.
freeair: The presence of any sort of smoke free preemption law corresponds to a decrease of 6.5 packs per capita annually.
pcgsp: Converting pcgsp to thousands per capita (from millions), a thousand dollar increase in per capita gross state product leads to an increase of .46 packs per capita annually.
college: A one percentage point increase in a population of people who have attended college corresponds to a drop of 1.3 in per capita pack consumption.

Table 1: Column (1): fixed effects model with all possible explanatory variables. Column (2): fixed effects model with statistically significant variables. Column (3): fixed effects first-order autoregressive model. Column (4): random effects first-order autoregressive model.

VARIABLES	(1) packs	**(2) packs**	**(3) packs**	(4) packs
lsttax	-9.368***	**-9.331*** **	**-10.81*** **	-11.24***
	(0.921)	**(0.902)**	**(0.961)**	(0.931)
ppp	-7.598***	**-7.611*** **	**-7.283*** **	-7.690***
	(0.666)	**(0.599)**	**(0.731)**	(0.716)
ctrlfund	0.283**	**0.281** **	**0.133**	0.141
	(0.113)	**(0.111)**	**(0.120)**	(0.120)
alciati	-0.0957	**-0.0764**	**-0.0779**	-0.0999
	(0.0838)	**(0.0810)**	**(0.103)**	(0.0874)
pup	-0.700	**-0.719**	**-0.284**	-0.362
	(0.520)	**(0.504)**	**(0.626)**	(0.569)
freeair	-6.610***	**-6.490*** **	**-0.0530**	-0.981
	(1.280)	**(1.268)**	**(1.520)**	(1.385)
pct85	89.37			
	(249.1)			
pct65	-26.68			
	(65.88)			
kids	-27.03	**-22.76**	**-7.594**	-11.58
	(16.49)	**(16.00)**	**(11.75)**	(12.00)
unem	-0.366			
	(0.302)			
pcgsp	413.9***	**460.7*** **	**186.9**	298.6***
	(130.6)	**(118.7)**	**(114.0)**	(108.4)
hs	-0.0239			
	(0.169)			
college	-1.291***	**-1.291*** **	**-0.456*** **	-0.701***
	(0.203)	**(0.176)**	**(0.162)**	(0.157)
Constant	135.4***	**126.8*** **	**108.0*** **	113.4***
	(15.79)	**(5.058)**	**(2.004)**	(5.484)
Observations	678	**678**	**630**	678
R-squared	0.763	**0.763**		
Number of state_n	48	**48**	**48**	48

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

These initial models are insightful, and show that there is correlation between the policies that states use to curb tobacco consumption, but there is clearly a problem with autocorrelation since the smoking rate in a state in a given period will clearly be highly correlated with the smoking rates in previous periods.

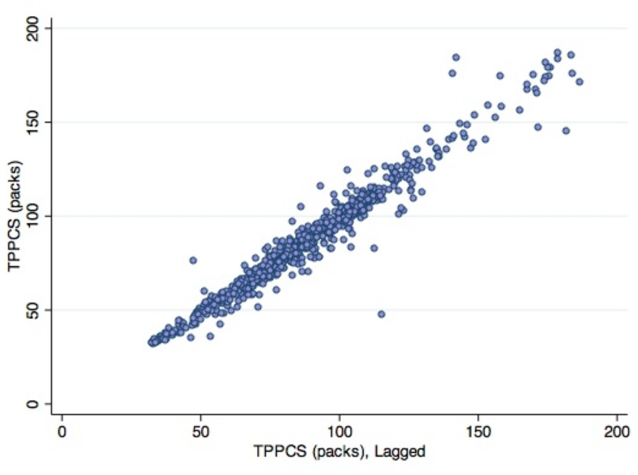

Figure 2: This scatter plot shows the correlation between packs and lagged packs, which indicates that autocorrelation is a source of error for our first two models.

To mitigate this problem we use a fixed effects autoregressive model, which is also represented by Equation 2. The results of this model are presented in column (3) of Table 1, and the Stata output can be found in Appendix 1.C. With this model we find that the only statistically significant factors that seem to affect the consumption of tobacco are the average price per pack, the state excise tax per pack and the percentage of the population, which has graduated from college. The results from this model are interpreted below:

lsttax: a 10% increase in state sales tax will lead to a drop in per capita packs consumption of 1.08.
ppp: a one dollar increase in the average price of a pack of cigarettes leads to a drop in per capita consumption of 7.28.
college: a one percentage point increase in a population of people who have graduated college leads to a per capita packs consumption drop of 0.46.
Looking at these, it is important, as always, to keep in mind the difference between statistical and economic significance. However, the variables described above appear to be economically as well as statistically significant.

We estimate the random effects first order autoregressive model (Equation 2), since this would be more efficient that the fixed effects model. The results from this model are given in column (4) of Table 1, and the Stata output can be found in Appendix 1.D. We perform a Hausman test with the null hypothesis that there is no systematic difference between the fixed and random effects models and the alternative hypothesis that there is a difference. The fixed effects model is consistent under both hypotheses, while the random effects model is not consistent under the alternative hypothesis. We compute a chi^2 value of 115.84, and reject the null hypothesis at the 5% level (See Appendix 2 for Stata output). This means that the random effects model is inconsistent, so we do not use it to model this relationship.

External Validity

With respect to other time periods in America, we expect this model to be fairly valid, given that we were able to look at 48 of the 50 states, and over a period of more than 15 years. A possible complication here is that the data end in 2007, right before the financial meltdown of 2008. It's possible that in the ensuing years, given the more extreme values for income, consumption habits and reactions to taxes may be different.

This analysis likely only applies to the United States since smoking is an activity that is affected by cultural factors. In other countries, such as Russia, with greater tobacco consumption, taxes may have less of an effect. Likewise, a country having a different "preferred method" of tobacco consumption, such as hookah instead of cigarettes, would change the effect of tobacco taxes.

Internal Validity

Autocorrelation

Autocorrelation across time is a problem with this type of analysis, but we have tried to mitigate this through the use of a first-order autoregressive model. It is also possible that there is autocorrelation within regions of the country. The fixed effects model should lessen this issue, but it may still be a source of error.

Omitted Variables

Our greatest threat to internal validity in this model seems to be omitted variable bias: the omission of variables that are correlated with any of the explanatory variables we do include will bias our coefficients, as these pick up some of the effects of the omitted variables. We are missing a couple of key demographic variables including the gender balance of states and racial composition of a state. Additionally, if we were not using a fixed effects model, it may have been informative to include a variety of US region dummies such as "South" and "Northwest" to attempt to get at the question of varying culture causing some of these differences. Another variable that would be useful to include would be the percentage of the GSP that is from the production of tobacco products, since this may be positively correlated with tobacco consumption.

A problem with our data is that consumption is determined by the number of packs sold in a state. It's possible that residents of a state with high cigarette taxes living near a state with low ones routinely cross the border to purchase cigarettes. This exaggerate the effect of state excise tax per pack on consumption, since states with low taxes will have higher numbers of packs sold, but not necessarily consumed by residents. It might be possible to correct for this using dummy variables for states that have very different tax levels than their neighbors, but it would be difficult to determine a valid procedure for producing this variable.

Endogeneity

Endogeneity is certainly a problem with this model that we unfortunately did not correct for in this analysis. It is likely that states with relatively high tobacco consumption may direct more funding towards anti-smoking campaigns. This is probably the reason for the positive coefficients on tobacco control funding per capita. However, it is also possible that states might also raise tobacco taxes in response to high tobacco consumption. We do find negative coefficients for state excise tax on tobacco, but it is still a problem that possibly affects our results.

It may be possible to deal with this endogeneity through instrumental variables, although we do not perform this analysis. Possible instrumental variables could be political. The assumption would be that democrats are more likely to fund anti-smoking campaigns and raise taxes. It is also necessary to assume that states with high tobacco consumption are not more likely to elect democrats. This would allow us to see the effect of changes in cigarette taxes that are only the result of changes between political parties.

Heteroskedasticity

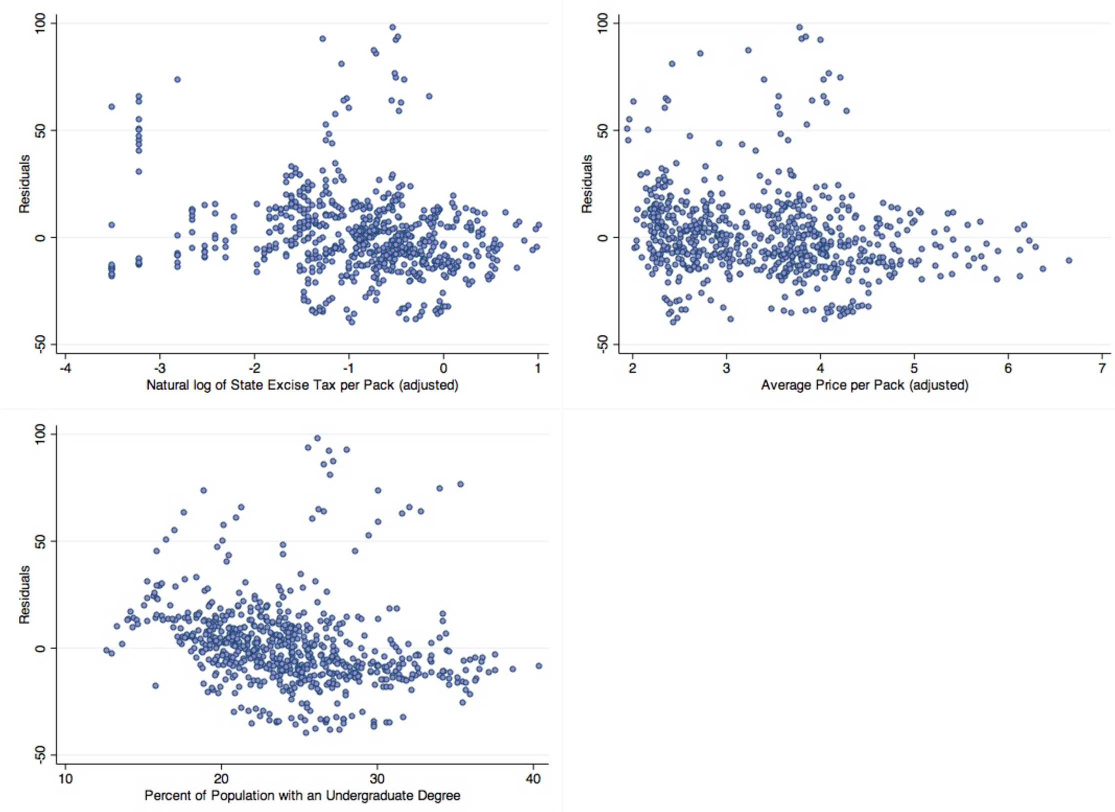

Figure 3: These scatter plots show the distribution of error terms in relation to the three most significant explanatory variables. Heteroskedasticity does not seem to be a large problem. The plots of residuals against average price may appear to be heteroskedastic, but this is most likely due to the small number of observations where the average price is higher than $4.50.

Conclusion

Using these models, we determine that taxes on tobacco are negatively correlated with cigarette consumption. The three other important determinants of cigarette consumption that we find are GSP per capita, the price of cigarettes, and the percent of the population that has graduated from college. This analysis answers some preliminary questions about the determinants of tobacco consumption, but there are still many interesting topics that could be explored in this area. Although our results vary slightly between the different models we use, the coefficients from these four main explanatory variables remain consistent. This is promising for the validity of our analysis since despite changes in our assumptions, we arrive at similar conclusions.

If we were able to find individual level data for tobacco consumption, this sort of analysis would benefit particularly from the method of quantile regression; that is, we assume that the effects of tobacco taxes vary greatly depending on what section of the sample with respect to, say, income we are looking at. Any tax analysis is conducive to using quantile regression, in part to determine whether a particular tax is progressive or regressive. We suppose that if we had more time even without individual data a quantile regression could be useful, that is to see whether the poorer states react to cigarette tax changes differently from the richer states.

Appendix

Appendix 1.A
Stata output for fixed effects model using equation 1:

```
. xtreg packs lsttax ppp ctrlfund alciati pup freeair pct85 pct65 kids unem pcgsp hs college, fe

Fixed-effects (within) regression               Number of obs      =       678
Group variable: state_n                         Number of groups   =        48

R-sq:  within  = 0.7634                         Obs per group: min =        13
       between = 0.4599                                        avg =      14.1
       overall = 0.5191                                        max =        16

                                                F(13,617)          =    153.17
corr(u_i, Xb)  = 0.1957                         Prob > F           =    0.0000

------------------------------------------------------------------------------
       packs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsttax |   -9.36778   .9206478   -10.18   0.000    -11.17576   -7.559797
         ppp |  -7.598407   .6664554   -11.40   0.000    -8.907203   -6.289611
    ctrlfund |   .2832291   .1128842     2.51   0.012     .0615453    .5049129
     alciati |  -.0956618   .0838455    -1.14   0.254     -.260319    .0689953
         pup |  -.6996964   .5196333    -1.35   0.179    -1.720161     .320768
     freeair |   -6.60956   1.280484    -5.16   0.000    -9.124196   -4.094925
       pct85 |   89.36973   249.0547     0.36   0.720     -399.728    578.4675
       pct65 |  -26.67913   65.88095    -0.40   0.686    -156.0572     102.699
        kids |  -27.03198   16.49138    -1.64   0.102    -59.41802    5.354054
        unem |  -.3658644   .3024867    -1.21   0.227    -.9598927    .2281639
       pcgsp |   413.8875   130.5566     3.17   0.002     157.4984    670.2767
          hs |  -.0238531   .1694291    -0.14   0.888    -.3565806    .3088745
     college |  -1.291238    .203213    -6.35   0.000    -1.690311   -.8921651
       _cons |   135.4394   15.78691     8.58   0.000     104.4368     166.442
-------------+----------------------------------------------------------------
     sigma_u |  19.348318
     sigma_e |  6.4480165
         rho |  .90003971   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47, 617) =    90.51             Prob > F = 0.0000
```

Appendix 1.B
Stata output for fixed effects model using equation 2:

```
. xtreg packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, fe

Fixed-effects (within) regression               Number of obs      =       678
Group variable: state_n                         Number of groups   =        48

R-sq:  within  = 0.7628                         Obs per group: min =        13
       between = 0.4651                                        avg =      14.1
       overall = 0.5210                                        max =        16

                                                F(9,621)           =    221.89
corr(u_i, Xb)  = 0.2038                         Prob > F           =    0.0000

------------------------------------------------------------------------------
       packs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsttax |  -9.330977   .9023114   -10.34   0.000    -11.10293   -7.559026
         ppp |  -7.611413     .59921   -12.70   0.000    -8.788137    -6.43469
    ctrlfund |   .2805946   .1112882     2.52   0.012     .0620479    .4991413
     alciati |  -.0764231   .0809562    -0.94   0.346    -.2354043    .0825581
         pup |  -.7188776   .5037048    -1.43   0.154    -1.708049    .2702934
     freeair |   -6.48967   1.268317    -5.12   0.000     -8.98038    -3.99896
        kids |  -22.76429   16.00215    -1.42   0.155    -54.18918    8.660601
       pcgsp |   460.7052   118.6971     3.88   0.000     227.6089    693.8015
```

```
     college |  -1.290982   .1764167    -7.32   0.000    -1.637428   -.9445367
       _cons |   126.7696   5.058271    25.06   0.000     116.8362     136.703
-------------+----------------------------------------------------------------
     sigma_u |  19.323767
     sigma_e |  6.4358305
         rho |  .90015157   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47, 621) =    97.88             Prob > F = 0.0000
```

Appendix 1.C
Stata output for fixed effects AR(1) model using equation 2:

```
. xtregar packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, fe

FE (within) regression with AR(1) disturbances  Number of obs      =       630
Group variable: state_n                         Number of groups   =        48

R-sq:  within  = 0.4842                         Obs per group: min =        12
       between = 0.4524                                        avg =      13.1
       overall = 0.4951                                        max =        15

                                                F(9,573)           =     59.76
corr(u_i, Xb)  = 0.2108                         Prob > F           =    0.0000

------------------------------------------------------------------------------
       packs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsttax |  -10.80558   .9608775   -11.25   0.000    -12.69286   -8.918313
         ppp |  -7.283408   .7312701    -9.96   0.000    -8.719705   -5.847111
    ctrlfund |   .1333471   .1199202     1.11   0.267    -.1021897    .3688838
     alciati |  -.0778951    .102636    -0.76   0.448    -.2794837    .1236936
         pup |  -.2844226   .6255585    -0.45   0.650     -1.51309    .9442447
     freeair |  -.0530195   1.519942    -0.03   0.972    -3.038357    2.932318
        kids |  -7.593953   11.74682    -0.65   0.518    -30.66604    15.47813
       pcgsp |   186.9121   114.0403     1.64   0.102     -37.0758    410.9001
     college |  -.4558748   .1618757    -2.82   0.005     -.773817   -.1379327
       _cons |    108.035   2.004435    53.90   0.000     104.0981     111.972
-------------+----------------------------------------------------------------
      rho_ar |  .65907536
     sigma_u |   20.08889
     sigma_e |   5.120859
     rho_fov |  .93898556   (fraction of variance because of u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47,573) =    22.49              Prob > F = 0.0000
```

Appendix 1.D
Stata output for random effects AR(1) model using equation 2:

```
. xtregar packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, re

RE GLS regression with AR(1) disturbances       Number of obs      =       678
Group variable: state_n                         Number of groups   =        48

R-sq:  within  = 0.7461                         Obs per group: min =        13
       between = 0.4596                                        avg =      14.1
       overall = 0.5131                                        max =        16

                                                Wald chi2(10)      =    828.26
corr(u_i, Xb)      = 0 (assumed)                Prob > chi2        =    0.0000

------------------- theta --------------------
  min      5%       median        95%      max
0.7750   0.7750     0.7750     0.7919   0.7919

------------------------------------------------------------------------------
       packs |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      lsttax |  -11.24066   .9305375   -12.08   0.000    -13.06448   -9.416843
```

```
         ppp |  -7.689891   .7164187   -10.73   0.000    -9.094046   -6.285736
    ctrlfund |   .1410929   .1200319     1.18   0.240    -.0941652    .3763511
     alciati |  -.0998862   .0873968    -1.14   0.253    -.2711808    .0714083
         pup |  -.3615051   .5686716    -0.64   0.525    -1.476081    .7530706
     freeair |  -.9807887     1.3849    -0.71   0.479    -3.695142    1.733565
        kids |  -11.57611   11.99633    -0.96   0.335    -35.08849    11.93627
       pcgsp |   298.6233   108.3598     2.76   0.006       86.242    511.0046
     college |  -.7014027   .1571281    -4.46   0.000    -1.009368   -.3934373
       _cons |    113.362   5.484079    20.67   0.000     102.6134    124.1106
-------------+----------------------------------------------------------------
      rho_ar |  .65907536   (estimated autocorrelation coefficient)
     sigma_u |  17.006135
     sigma_e |  5.4972002
     rho_fov |  .90539581   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Appendix 2
Hausman test for fixed and random effects models using equation 2:

```
. hausman fe2 re2

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      fe2          re2         Difference          S.E.
-------------+----------------------------------------------------------------
      lsttax |   -10.80558    -11.24066        .4350787         .239553
         ppp |   -7.283408    -7.689891        .4064831        .1466294
    ctrlfund |    .1333471     .1410929       -.0077459               .
     alciati |   -.0778951    -.0998862        .0219912        .0538141
         pup |   -.2844226    -.3615051        .0770825        .2606455
     freeair |   -.0530195    -.9807887        .9277693        .6263205
        kids |   -7.593953    -11.57611        3.982157               .
       pcgsp |    186.9121     298.6233       -111.7111        35.54348
     college |   -.4558748    -.7014027        .2455279        .0389167
------------------------------------------------------------------------------
                          b = consistent under Ho and Ha; obtained from xtregar
           B = inconsistent under Ha, efficient under Ho; obtained from xtregar

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =      115.84
                Prob>chi2 =      0.0000
                (V_b-V_B is not positive definite)
```

Appendix 3
All Stata commands:

```
xtreg packs lsttax ppp ctrlfund alciati pup freeair pct85 pct65 kids unem pcgsp hs college, fe
outreg2 using tobacco, word

xtreg packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, fe
outreg2 using tobacco, word

xtregar packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, fe
estimate store fe2
outreg2 using tobacco, word

xtregar packs lsttax ppp ctrlfund alciati pup freeair kids pcgsp college, re
estimate store re2
outreg2 using tobacco, word

hausman fe2 re2

scatter packs stextpp
scatter packs lsttax

predict uehat, ue
```

```
scatter uehat lsttax
scatter uehat ppp
scatter uehat college
```