On-line Appendix

In this section we present results which address, to some extent, potential threats to the validity of the results we report in the main paper. In most cases we report results overall and by race/ethnicity in this section. While we also reported results by gender and race/ethnicity-by-gender in the main paper we feel the results by race/ethnicity are most likely to be influenced by the factors assessed in this section due to either the uneven racial/ethnic distribution of race/ethnicity across states or the racial/ethnic differences in the timing of smoking initiation.

First, accounting for potential unobserved differences across states which could be related to both smoking rates and taxes might bias our results for tax and price. In our main model we accounted for this using the adult smoking prevalence for each state in 1997 as a proxy for a measure of a state's smoking sentiment. We also estimated several additional models where we used alternative measures of a state's smoking (or anti-) sentiment: using adult smoking prevalence averaged over the years for which each state has data in the BRFSS to replace the BRFSS adult smoking prevalence at baseline measure; using an average BRFSS measure (1991–2005) and an average SASS measure (1991–2005) instead of the BRFSS adult smoking prevalence were modeled separately; including state fixed effects as a proxy for SASS (Note that when estimating models with state fixed effects by race/ethnicity, convergence of the models were sensitive to inclusion of those who had moved. We thus estimated our models with and without those who moved and found this did not influence our results. The results presented below when state fixed effects are included excludes movers). We also estimated models with and without our measure of tobacco control funding. These results are presented below in table A.1. The results are mostly robust to these alternative specifications though the tax results are sensitive to inclusion of state fixed effects in the overall sample.

Table A-1. Alternative Specifications of Model with regional and state fixed effects—Overall and by Race/ethnicity

Odds Ratio Standard Error p-value	Region FE								State FE							
	Overall		White		Black		Hispanic		Overall		White		Black		Hispanic	
	Tax	Price	Tax	Price	Tax	Price	Tax	Price	Tax	Price	Tax	Price	Tax	Price	Tax	Price
BRFSS at 1997 + funding (main model)	0.88*	0.88**	0.96	0.94	0.83**	0.85**	0.84	0.83**	0.92	0.86**	1.08	0.90*	0.72**	0.78**	0.92	0.86*
	0.060	0.033	0.070	0.045	0.069	0.047	0.126	0.065	0.061	0.036	0.117	0.057	0.104	0.067	0.115	0.071
	0.063	0.001	0.612	0.174	0.021	0.004	0.236	0.020	0.229	0.000	0.487	0.100	0.024	0.005	0.492	0.061
BRFSS at 1997 without funding.	0.87**	0.88**	0.97	0.94	0.82**	0.84**	0.78*	0.80**	0.90	0.85**	1.08	0.91	0.71**	0.78**	0.84	0.81**
	0.054	0.030	0.069	0.043	0.066	0.045	0.115	0.066	0.059	0.034	0.113	0.055	0.103	0.065	0.107	0.065
	0.024	0.000	0.667	0.203	0.012	0.002	0.096	0.006	0.120	0.000	0.493	0.111	0.020	0.003	0.172	0.009
Average BRFSS + funding	0.88**	0.88**	0.95	0.94	0.80**	0.80**	0.81**	0.85**	0.92	0.86**	1.08	0.90*	0.72**	0.78**	0.92	0.86*
	0.031	0.033	0.046	0.046	0.038	0.037	0.066	0.065	0.061	0.036	0.117	0.057	0.104	0.067	0.115	0.071
	0.000	0.001	0.260	0.208	0.000	0.000	0.010	0.033	0.229	0.000	0.487	0.100	0.024	0.005	0.492	0.061
Average BRFSS without funding	0.87**	0.88*	0.99	0.98	0.71**	0.72**	0.81	0.86	0.90	0.85**	1.08	0.91	0.71**	0.78	0.84	0.81**
	0.058	0.061	0.074	0.073	0.041	0.041	0.124	0.131	0.059	0.034	0.113	0.055	0.103	0.065	0.107	0.065
	0.040	0.059	0.872	0.791	0.000	0.000	0.169	0.336	0.120	0.000	0.493	0.111	0.020	0.003	0.172	0.009
Average SASS + funding	0.88**	0.88**	0.94	0.93	0.83**	0.83**	0.82	0.84**	0.92	0.86**	1.08	0.90*	0.72**	0.78**	0.92	0.86*
	0.034	0.034	0.047	0.048	0.050	0.049	0.07	0.067	0.061	0.036	0.117	0.057	0.104	0.067	0.115	0.071
	0.001	0.001	0.202	0.172	0.002	0.001	0.019	0.034	0.229	0.000	0.487	0.100	0.024	0.005	0.492	0.061
Average SASS without funding	0.88*	0.88*	0.98	0.97	0.75**	0.75**	0.83	0.86	0.90	0.85**	1.08	0.91	0.71**	0.78**	0.84	0.81**
	0.066	0.067	0.081	0.081	0.059	0.059	0.131	0.134	0.059	0.034	0.113	0.055	0.103	0.065	0.107	0.065
	0.093	0.100	0.788	0.746	0.000	0.000	0.233	0.327	0.120	0.000	0.493	0.111	0.020	0.003	0.172	0.009

Note: BRFSS = Behavioral Risk Factor Surveillance System; SASS = state anti-smoking sentiment; FE = fixed effects

* Significant at 10%.
** Significant at 5%.

Another potential source of bias stems from the discrete-time methodology we employ. We report results that impose a proportional odds assumption (i.e., that tax and price have the same impact on the hazard at each age). It is possible to relax this assumption by including interaction terms between the tax/price and each age indicator. This model specification, although possible, is complicated, especially in the context of a logistic regression model, and thus we chose not to report results from such a model. However, we did explore this issue and found that a model including the interaction terms did not improve the fit of the model. We also estimated a model using age groupings (5–11, 12–17, and 18+) rather than single year age dummies (to reduce the number of interaction terms) and found that the proportional odds assumption is violated—suggesting that older youth are more sensitive to tax or price (see table A-2 below).

Table A-2. Regression Results for Model Specification Testing Proportional Odds Assumption (coefficient, std. error, p-value)

	Tax	Tax+Funding	Price	Price+Funding
Tax	3.28	3.31		
	0.719	0.731		
	0.000	0.000		
Price			2.57	2.56
			0.254	0.255
			0.000	0.000
Tax*Age2 (12–16)	0.40	0.40		
	0.085	0.085		
	0.000	0.000		
Tax*Age3 (17–20)	0.267	0.267		
	0.057	0.057		
	0.000	0.000		
Tax*Age4 (21+)	0.309	0.307		
	0.071	0.07		
	0.000	0.000		
Price*Age2 (12–16)			0.42	0.42
			0.044	0.044
			0.000	0.000
Price*Age3 (17–20)			0.31	0.31

			0.032	0.032
			0.000	0.000
Price*Age4 (21+)			0.38	0.38
			0.049	0.049
			0.000	0.000
Age2 (12–16)	14.89	14.92	61.55	61.52
	2.222	2.23	15.554	15.542
	0.000	0.000	0.000	0.000
Age3 (17–20)	21.54	21.68	168.95	169.02
	3.411	3.45	45.258	45.266
	0	0.000	0.000	0.000
Age4 (21+)	6.88	6.98	32.26	32.02
	1.356	1.39	14.111	14.027
	0.000	0.000	0.000	0.000

Another potential limitation is that our unweighted sample, which we used in estimating our results, is not very representative of the US population (see table A-3). To partially address this limitation we re-estimated our model using baseline weights (see table A-4 below). Using this model specification, our results remained qualitatively similar.

Table A-3. Race/ethnicity Distribution for NLSY97 Unweighted, NLSY97 Weighted, and 2000 Census

Race/Ethnicity	NLSY (non-weighted)	NLSY (weighted)	2000 Census (10–19 year olds)
Non-Hispanic White	49%	65%	63%
Black	26%	17%	15%
Hispanic**	21%	13%	17%
Other Race	4%	5%	5%

**Hispanic ethnicity is over counted for the Census numbers, as it is for the total number of hispanic of any race. Also, Black and Other Races (for the Census) are not exclusive of Hispanic ethnicity. Only, white race is exclusive of hispanic origin.

Table A-4. Weighted Results Overall, by Race/ethnicity and Race-by-Gender

		With Regional Fixed Effects			
		Tax	Tax+Funding	Price	Price+Funding
Overall	OR	0.91	0.92	0.89**	0.89**
	Std Err.	0.051	0.051	0.036	0.037
	P-value	0.109	0.119	0.004	0.006

	Elasticity	−0.06	−0.06	−0.29	−0.29
White	OR	1.00	1.00	0.95	0.95
	Std Err.	0.073	0.074	0.047	0.048
	P-value	0.959	0.972	0.337	0.284
	Elasticity	0.00	0.00	−0.12	−0.13
Black	OR	0.73**	0.74**	0.83**	0.84**
	Std Err.	0.083	0.083	0.065	0.065
	P-value	0.007	0.008	0.018	0.023
	Elasticity	−0.21	−0.21	−0.47	−0.44
Hispanic	OR	0.75**	0.78*	0.78**	0.81**
	Std Err.	0.108	0.114	0.069	0.074
	P-value	0.044	0.089	0.006	0.022
	Elasticity	−0.24	−0.21	−0.65	−0.57

* Significant at 10%.
** Significant at 5%.

An additional source of bias arises as a result of cases that had missing data for some period of the survey years. In our sample, 819 cases have missing data but are not censored observations (attrition or loss to follow-up is handled via right censoring in the discrete-time models); that is, they are missing for several waves (and were a nonsmoker prior to missing) but then came back into the sample at which time we again observed their smoking status. For these individuals, we have several options including dropping them or imputing when the event occurred (or smoking status during the missing period). Table A-5 shows the sensitivity of our overall results to different assumptions about when the event occurred during the missing period. The options are (1) drop (i.e., delete the entire respondent record), (2) assume the event happened at the mid-point of the range of missing waves—this is what we report in the paper, (3) assume the event happened at the beginning of the period of missing, and (4) assume the event happens at the end of the period of missingness. Assuming the event happens at the beginning of the period of missingness shifts the hazard profile to the left (earlier ages), whereas assuming the event happens at the end of the period of missingness shifts the hazard to the right (older ages). These assumptions do have a small impact on our results (see Table A-5) – in

general the later the timing of event is made by assumption for missing cases the less significant the effects for tax and price become.

Table A-5. Results Showing Sensitivity of Results Overall and by Race to Different Assumptions about When Event Occurred for Those with a Missing Period of Data (but not censored)

		First Period		Last Period		No Imputation		Dropping Those w/ Missing		Midpoint	
		Tax	Price	Tax	Price	Tax	Price	Tax	Price	Tax	Price
Overall	OR	0.86**	0.84**	0.96	0.96	0.96	0.97	0.86**	0.82**	0.88*	0.88**
	Std Err.	0.060	0.030	0.060	0.040	0.060	0.040	0.060	0.040	0.060	0.030
	P-value	0.020	0.000	0.549	0.388	0.532	0.453	0.025	0.000	0.065	0.001
	Elasticity	−0.111	−0.454	−0.029	−0.091	−0.031	−0.080	−0.11	−0.505	-0.093	-0.311
White	OR	1.00	0.92*	1.03	1.00	1.03	1.01	0.99	0.90*	0.99	0.94
	Std Err.	0.080	0.050	0.080	0.050	0.080	0.050	0.090	0.050	0.080	0.050
	P-value	0.970	0.079	0.730	0.959	0.736	0.875	0.902	0.069	0.852	0.244
	Elasticity	−0.002	−0.218	0.02	0.007	0.02	0.02	−0.008	−0.259	-0.010	-0.145
Black	OR	0.72**	0.77**	0.85**	0.90**	0.84**	0.91**	0.72**	0.74**	0.74**	0.82**
	Std Err.	0.050	0.040	0.050	0.040	0.050	0.040	0.050	0.050	0.050	0.040
	P-value	0.000	0.000	0.003	0.041	0.001	0.041	0.000	0.000	0.000	0.000
	Elasticity	−0.224	−0.65	−0.116	−0.251	−0.121	−0.249	−0.23	−0.749	-0.207	-0.503
Hispanic	OR	0.77**	0.75**	1.01	0.98	1.01	0.98	0.78*	0.71**	0.86	0.85**
	Std Err.	0.100	0.050	0.140	0.090	0.140	0.090	0.110	0.050	0.130	0.070
	P-value	0.049	0.000	0.949	0.792	0.964	0.848	0.072	0.000	0.326	0.033
	Elasticity	−0.212	−0.754	0.008	−0.067	0.005	−0.049	−0.202	−0.901	-0.125	-0.445

* Significant at 10%.
** Significant at 5%.

40