Cities and Smoking: ONLINE APPENDIX

Michael E. Darden*

Johns Hopkins University and NBER

June, 2020

Abstract

Appendix Tables and Figures

Table A1: Sample Construction

Unique Individuals	Person/Years	Description
12,484	71,026	Baseline Sample
10,534	58,371	- those with non-missing smoking waves.
9,315	52,381	- those with location ever missing.
9,163	51,375	- those with missing smoking or education.
5,260	42,080	- those with leaving through attrition .
4,626	37,008	- those with unobserved tax, local area, or Beale Code information

Notes: 4,626 unique individuals yields 37,008 person/year observations.

Table A2: Rural-Urban Continuum Codes, 2013

Code	Description
Metropolitan Counties	
1	Counties in metro areas of 1 million population or more
2	Counties in metro areas of 250,000 to 1 million population
3	Counties in metro areas of 50,000-250,000 population
Nonmetropolitan Counties	
4	Urban population of 20,000-49,999, adjacent to a metro area
5	Urban population of 20,000-49,999, not adjacent to a metro area
6	Urban population of 2,500 to 19,999, adjacent to a metro area
7	Urban population of 2,500 to 19,999, not adjacent to a metro area
8	Completely rural or less than 2,500 urban population, adjacent to a metro area
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area

Notes: Table presents the 2013 Rural-Urban Continuum Codes produced by the USDA Economic Research Service. These codes are merged to NLSY 1979 data at the county level. Data are available at https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/.

Table A3: Highest Level of Education

	High School		Some College		College or More	
Age	-0.079	(0.031)***	-0.081	(0.034)**	-0.164	(0.036)***
Female	0.106	(0.124)	0.569	(0.139)***	0.577	(0.148)***
Hispanic	-0.453	(0.179)***	0.052	(0.202)	-0.318	(0.221)
African-American	0.521	(0.160)***	1.574	(0.180)***	1.688	(0.197)***
AFQT						
Quartile 2	1.643	(0.168) ***	2.753	(0.192)***	3.519	(0.251)***
Quartile 3	2.162	(0.224)***	3.911	(0.244)***	5.308	(0.290)***
Quartile 4	3.693	(0.629)***	6.264	(0.638)***	8.715	(0.654)***
AFQT Missing	0.640	(0.322)**	1.378	(0.387)***	2.986	(0.421)***
Adjacent at 14	0.153	(0.207)	-0.084	(0.237)	-0.228	(0.252)
Remote at 14	-0.096	(0.310)	0.121	(0.353)	-0.215	(0.375)
Age 17 Economic Characteristics						
County Unemployment	0.000	(0.002)	0.000	(0.002)	-0.002	(0.002)
Log Wages	0.262	(0.417)	0.142	(0.471)	0.117	(0.500)
Tuition	0.147	(0.090)*	-0.012	(0.097)	-0.016	(0.103)
Presence of	0.148	(0.156)	0.408	(0.174)**	0.422	(0.184)**
4-Year University						
Constant	-0.952	(4.233)	-2.773	(4.774)	-2.163	(5.059)***
Unobserved Heterogeneity						
μ_1	0.000	.	0.000	.	0.000	.
μ_2	0.850	(0.229)***	1.802	(0.269)***	2.037	(0.290)***
μ_3	-0.194	(0.238)	0.311	(0.287)	0.408	(0.310)

Notes: Coefficients and standard errors from the dynamic system of equations are estimated jointly with parameters that dictate other behaviors and outcomes. The dependent variable is the highest level of schooling, and the omitted category is less than high school. The p-value on the test that all coefficients associated with economic characteristics at age 17 are zero is 0.0005. * $p < 0.1$, ** $p < 0.05$ *** $p < 0.01$. n=37,008.

Table A4: Location in 1984

	Adjacent		Remote	
Age	-0.031	(0.034)	0.040	(0.046)
Female	-0.013	(0.147)	-0.470	(0.200)**
Hispanic	-0.693	(0.224)***	-1.377	(0.482)***
African-American	-0.569	(0.205)***	-1.584	(0.335)***
AFQT				
Quartile 2	-0.449	(0.226)**	-0.347	(0.336)
Quartile 3	-0.580	(0.246)**	-0.466	(0.335)
Quartile 4	-0.660	(0.280)**	-0.830	(0.374)**
AFQT Missing	-1.868	(0.590)***	-1.576	(0.754)**
Education				
High School	-0.237	(0.266)	0.088	(0.436)
Some College	-0.539	(0.317)*	0.007	(0.480)
College or More	-0.926	(0.331)**	-0.001	(0.517)
Adjacent at 14	4.794	(0.158)***	2.207	(0.283)***
Remote at 14	2.478	(0.315)***	5.290	(0.232)***
Constant	-1.778	(0.846)**	-3.728	(1.159)***
Unobserved Heterogeneity				
μ_1	0.000	.	0.000	.
μ_2	-0.146	(0.240)	-0.197	(0.313)
μ_3	-0.490	(0.304)*	-0.635	(0.390)*

Notes: Coefficients and standard errors from the dynamic system of equations are estimated jointly with parameters that dictate other behaviors and outcomes. The dependent variable is an individual's location in 1984, which is the first wave of the smoking panel. The omitted category is metro. * $p < 0.1$, ** $p < 0.05$ *** $p < 0.01$. n=37,008.

Table A5: Smoking in 1984

|	Smokes in 1984	
Age	0.053	(0.025)**
Female	-0.043	(0.102)
Hispanic	-0.725	(0.158)***
African-American	-0.492	(0.137)***
AFQT		
Quartile 2	-0.082	(0.153)
Quartile 3	-0.255	(0.169)
Quartile 4	-0.645	(0.204)***
AFQT Missing	-0.162	(0.357)
Education		
High School	-0.430	(0.203)**
Some College	-0.239	(0.224)
College or More	-1.047	(0.264)***
Adjacent at 14	0.051	(0.194)
Remote at 14	-0.074	(0.301)
Tobacco Control Variables		
State Tax at 14	-0.726	(1.201)
Local Tax at 14	-0.333	(0.557)
State Tax at 14 Squared	0.151	(0.915)
Local Tax at 14 Squared	0.093	(0.314)
Pack Sales	0.000	(0.004)
per Capita at 14		
Gross Cigarette	0.001	(0.000)**
Tax Revenue (M $2013) at 14		
Adjacent 1984	-0.113	(0.214)
Remote 1984	-0.347	(0.329)
Constant	2.525	(0.974)***
Unobserved Heterogeneity		
μ_1	0.000	.
μ_2	-4.547	(0.203)***
μ_3	-1.554	(0.215)***

Notes: Coefficients and standard errors from the dynamic system of
equations are estimated jointly with parameters that dictate other
behaviors and outcomes. The dependent variable is whether and
individual smokes cigarettes in 1984, which is the first wave of the

Table A6: Per-Period Location 1992-2014

	Adjacent		Remote	
Age	-0.012	(0.012)	-0.009	(0.017)
Female	0.137	(0.074)*	0.093	(0.110)
Hispanic	-0.702	(0.117)***	-1.243	(0.227)***
African-American	-0.725	(0.103)***	-1.025	(0.162)***
AFQT				
Quartile 2	-0.244	(0.117)**	-0.079	(0.183)
Quartile 3	-0.220	(0.127)*	0.014	(0.184)
Quartile 4	-0.337	(0.144)**	-0.021	(0.208)
AFQT Missing	-0.591	(0.328)*	-0.076	(0.452)
Education				
High School	-0.187	(0.143)	-0.226	(0.232)
Some College	-0.427	(0.164)***	-0.255	(0.258)
College or More	-0.824	(0.170)***	-0.571	(0.265)***
Lagged Smoking	0.093	(0.132)	0.175	(0.195)
Lagged Local UR	-0.004	(0.001)***	-0.003	(0.002)*
Lagged Adjacent	6.153	(0.062)***	3.267	(0.179)***
Lagged Remote	3.248	(0.179)***	7.175	(0.108)***
Linear Trend	0.082	(0.056)	0.110	(0.081)
Constant	-3.136	(0.385)***	-4.400	(0.606)***
μ_1	0.000	.	0.000	.
μ_2	0.121	(0.191)	-0.128	(0.279)
μ_3	0.078	(0.176)	-0.158	(0.258)

Notes: Coefficients and standard errors from the dynamic system of equations are estimated jointly with parameters that dictate other behaviors and outcomes. The dependent variable an individual's location in each wave between years 1992 and 2014. * $p < 0.1$, ** $p < 0.05$ *** $p < 0.01$. n=37,008.

Figure A1: Smoking by Type over Time
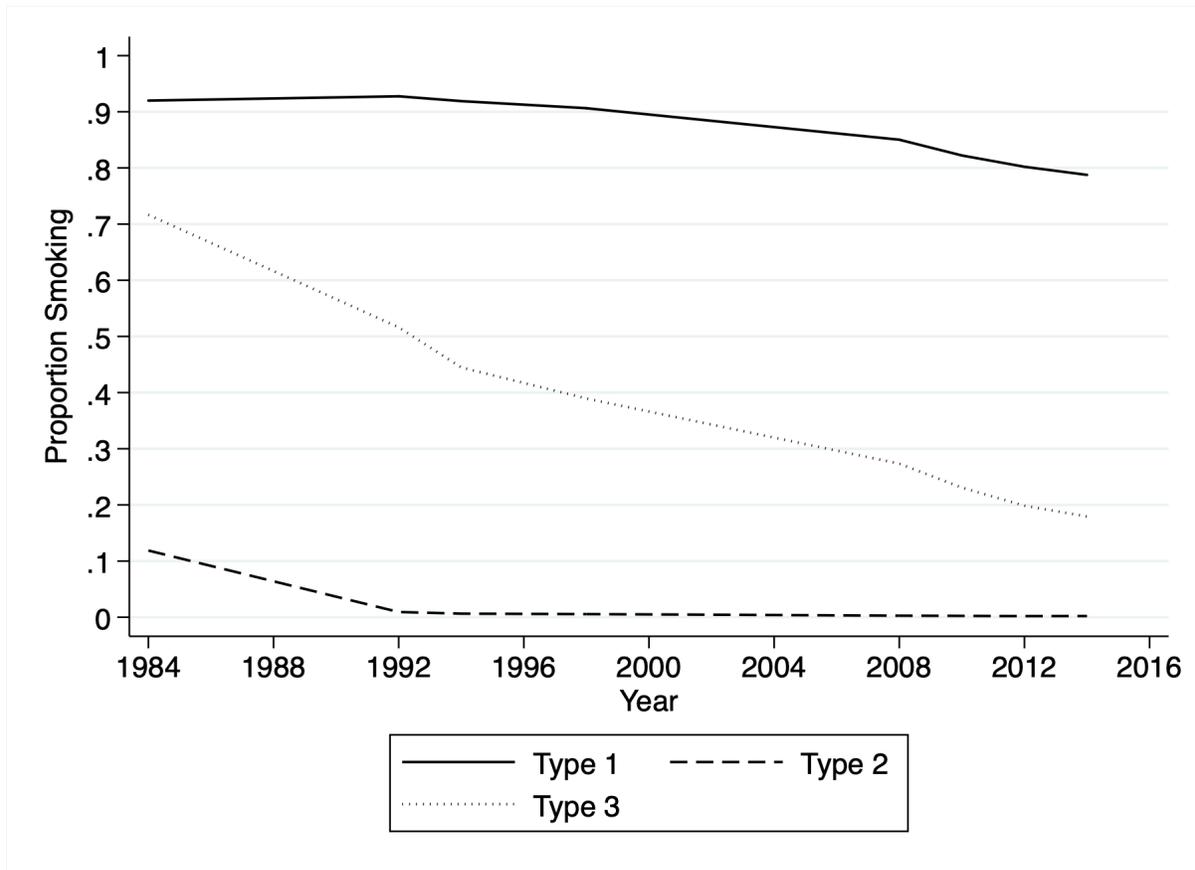
Notes:Figure presents the baseline simulation of smoking behavior by unobserved heterogeneity type.