Online Appendix: Understanding the Demand-Side of an Illegal Market: A Case Study of the Prohibition of Menthol Cigarettes

In this Appendix we first present some additional results, then discuss several lines of evidence that shed light on the validity of the stated preference (SP) data we collected through our discrete choice experiment (DCE), and the implications for the empirical results reported in the text of the paper. In section A we report additional results including descriptive statistics and conditional logit and mixed logit models estimated for sub-groups by gender, age, and dual use status. In section B we review previous research that compares SP and revealed preference (RP) data in a range of applications. In section C we discuss public health research on menthol prohibitions. In section D we review previous research that conducts DCEs of tobacco product choices. In section E we provide additional empirical evidence on the validity of our SP data. Based on previous research and the empirical evidence in section B, we conclude that because tobacco products are familiar market goods, econometric models estimated using the SP data from our DCE are likely to provide reliable forecasts of consumer demand.

A. Additional Results

In table A1, we present some descriptive statistics of the sample characteristics, in table A2, we show the Average tobacco product choices across all 12 Scenarios in the experiment.

A.1. Descriptive Statistics

Table A1. Descriptive Statistics of the Sample

	Mean	Sd	Min	Max
Age	44.0	12.7	18	82
Male	0.33	0.47	0	1
Female	0.66	0.47	0	1
Non-Binary, Agender, Gender Nonconforming, etc.	0.0063	0.079	0	1
Non-Hispanic White	0.50	0.50	0	1
Non-Hispanic Black	0.24	0.43	0	1
Non-Hispanic Asian	0.025	0.16	0	1
Non-Hispanic other	0.049	0.22	0	1
Hispanic	0.18	0.39	0	1
Grade school/ some high school	0.064	0.25	0	1
Completed high school (With diploma or GED Certificate)	0.26	0.44	0	1
Technical/ trade school or community college	0.42	0.49	0	1
Completed university degree (Four-year bachelor degree)	0.26	0.44	0	1
Full-time employed	0.51	0.50	0	1
Self-employed	0.080	0.27	0	1
Part-time employed	0.12	0.32	0	1
Not employed	0.15	0.36	0	1
Student	0.020	0.14	0	1
Retired	0.12	0.32	0	1
Household income <$25,000	0.28	0.45	0	1
Household income $25,000-$49,999	0.30	0.46	0	1
Household income $50,000-$74,999	0.18	0.38	0	1
Household income $75,000+	0.25	0.43	0	1
Price ($) paid for the last pack of cigarettes	8.46	2.99	1.20	19.5
Smoking status: Every day	0.58	0.49	0	1
Smoking status: Some days	0.42	0.49	0	1
Vaping status: Every day	0.11	0.31	0	1
Vaping status: Some days	0.40	0.49	0	1
Vaping status: Not at all	0.24	0.43	0	1
Vaping status: Never	0.25	0.43	0	1
Use of other tobacco products: Every day	0.16	0.37	0	1
Use of other tobacco products: Some days	0.23	0.42	0	1
Use of other tobacco products: Never	0.61	0.49	0	1
Current vaper uses menthol flavored e-cigarettes	0.31	0.46	0	1
Tried quitting smoking in past 12 months	0.51	0.50	0	1
Considering quitting smoking in the next 6 months	0.65	0.48	0	1
Observations	639			

Source: Cornell online survey (4/26-5/9, 2022) of menthol smokers.

Table A2. Average DCE Responses across all 12 Scenarios

	Today	In 6 months
Non-menthol Cigarettes	0.089	0.082
Menthol Cigarettes	0.43	0.34
Tobacco-flavored E-cigarettes	0.080	0.077
Menthol-flavored E-cigarettes	0.18	0.17
I will quit smoking cigarettes and not use e-cigarettes	0.22	0.33
Observations	7,668	7,668

Source: Cornell online survey (4/26-5/9, 2022) of menthol smokers.

A.2. Subgroup Analysis by Race

Text Table 3 reports the mixed logit models of consumer tobacco product choices by race. Appendix Table A3 and Figure A1 report the model's predicted market shares and quitting by race under different policy scenarios.

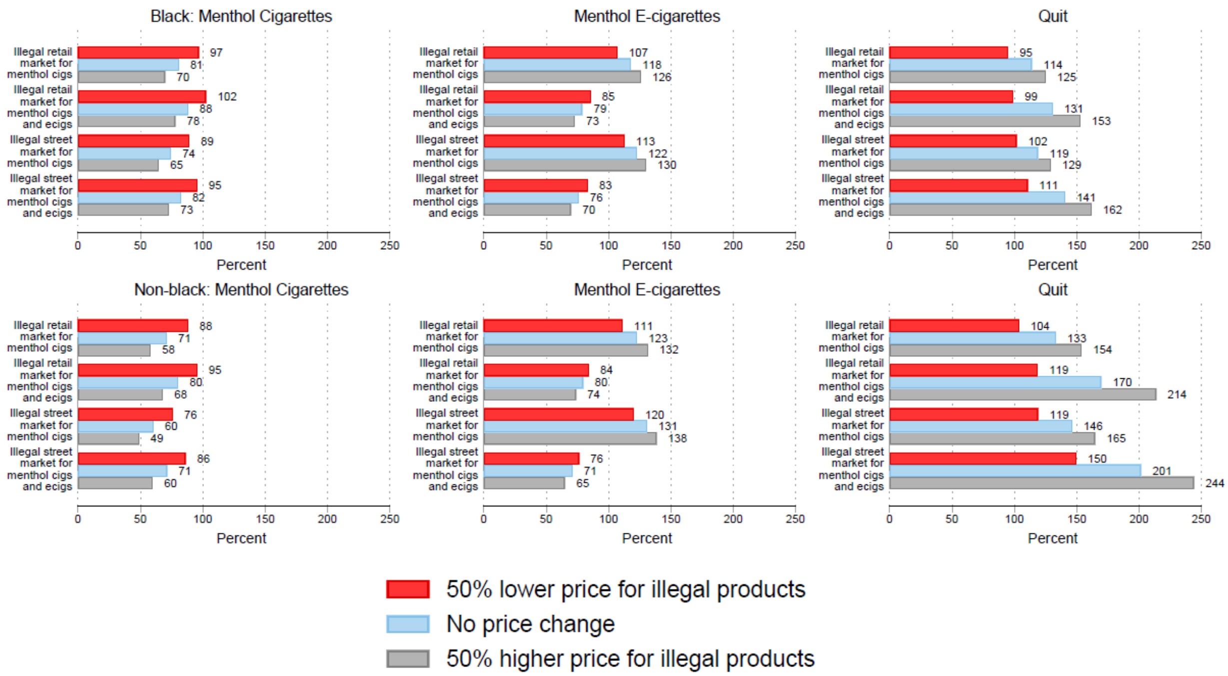
Table A3. Predicted Market Shares of Tobacco Products and Quitting, By Race

(A) Black Smokers	Non-menthol Cigs	Menthol Cigs	Tabacco-flavored E-cigs	Menthol-flavored E-cigs	Quitting
Status quo	0.034	0.467	0.065	0.225	0.209
Illegal Retail Market for Menthol Cigs					
1. 50% lower price for illegal products	0.041	0.451	0.070	0.240	0.198
2. No price change	0.044	0.376	0.077	0.265	0.238
3. 50% higher price for illegal products	0.048	0.326	0.083	0.283	0.261
Illegal Retail Market for Menthol Cigs & E-cigs					
4. 50% lower price for illegal products	0.047	0.478	0.076	0.192	0.206
5. No price change	0.053	0.411	0.087	0.177	0.272
6. 50% higher price for illegal products	0.058	0.363	0.096	0.164	0.319
Illegal Street Market for Menthol Cigs					
7. 50% lower price for illegal products	0.045	0.415	0.074	0.254	0.212
8. No price change	0.048	0.347	0.081	0.276	0.248
9. 50% higher price for illegal products	0.051	0.302	0.086	0.292	0.269
Illegal Street Market for Menthol Cigs & E-cigs					
10. 50% lower price for illegal products	0.054	0.445	0.083	0.187	0.231
11. No price change	0.059	0.383	0.093	0.171	0.293
12. 50% higher price for illegal products	0.064	0.339	0.102	0.157	0.337
(B) Non-black smokers	Non-menthol Cigs	Menthol Cigs	Tabacco-flavored E-cigs	Menthol-flavored E-cigs	Quitting
Status quo	0.086	0.481	0.056	0.270	0.106
Illegal Retail Market for Menthol Cigs					
1. 50% lower price for illegal products	0.102	0.423	0.065	0.300	0.110
2. No price change	0.112	0.341	0.074	0.332	0.141
3. 50% higher price for illegal products	0.121	0.279	0.081	0.355	0.163

Illegal Retail Market for Menthol Cigs & E-cigs					
4. 50% lower price for illegal products	0.115	0.459	0.074	0.227	0.126
5. No price change	0.131	0.385	0.088	0.215	0.180
6. 50% higher price for illegal products	0.147	0.326	0.101	0.199	0.227
Illegal Street Market for Menthol Cigs					
7. 50% lower price for illegal products	0.112	0.365	0.072	0.324	0.126
8. No price change	0.122	0.289	0.080	0.353	0.155
9. 50% higher price for illegal products	0.130	0.235	0.087	0.374	0.175
Illegal Street Market for Menthol Cigs & E-cigs					
10. 50% lower price for illegal products	0.134	0.414	0.087	0.206	0.159
11. No price change	0.150	0.343	0.101	0.192	0.214
12. 50% higher price for illegal products	0.165	0.287	0.114	0.175	0.259

Notes: Predictions are derived from estimation results of mixed logit models.

Figure A1: Predicted Market Shares Relative to the Status Quo (Black smoker vs. Non-black smoker)



Notes: Upper panel features predicted shares for black smokers, lower panel features predicted shares for non-black smokers.
Source: 2022 online Cornell survey

A.3. Sub-Group Analysis by Gender, Age, and Dual Use Status

Table A4: Conditional Logit Models Subgroup Analysis of Consumer Tobacco Product Choices

Immediate Choice Today	Full sample	By gender		By age		By product usage status	
		Male	Female	Age < 42	Age >= 42	Pure smokers	Dual users
ASC (Non-menthol cigarettes)	-0.200*	0.246	-0.452***	-0.365**	-0.038	-0.319**	0.064
	(0.115)	(0.191)	(0.145)	(0.161)	(0.165)	(0.161)	(0.171)
ASC (Menthol cigarettes)	1.904***	2.011***	1.838***	1.640***	2.179***	2.150***	1.861***
	(0.095)	(0.169)	(0.115)	(0.135)	(0.135)	(0.131)	(0.153)
ASC (Tobacco-flavored e-cigarettes)	-0.667***	-0.544***	-0.760***	-0.533***	-0.837***	-1.742***	0.072
	(0.116)	(0.197)	(0.144)	(0.157)	(0.173)	(0.210)	(0.156)
ASC (Menthol-flavored e-cigarettes)	0.626***	0.834***	0.509***	0.651***	0.598***	-0.246	1.363***
	(0.098)	(0.176)	(0.119)	(0.139)	(0.140)	(0.152)	(0.146)
Price ($)	-0.079***	-0.102***	-0.068***	-0.054***	-0.107***	-0.099***	-0.070***
	(0.006)	(0.012)	(0.007)	(0.007)	(0.009)	(0.009)	(0.008)
Illegal Retail Market for Menthol Cigarettes	-0.616***	-0.522***	-0.667***	-0.667***	-0.586***	-0.773***	-0.517***
	(0.060)	(0.104)	(0.074)	(0.081)	(0.091)	(0.094)	(0.081)
Illegal Street Market for Menthol Cigarettes	-0.866***	-0.817***	-0.890***	-0.825***	-0.917***	-1.056***	-0.753***
	(0.066)	(0.116)	(0.082)	(0.091)	(0.098)	(0.102)	(0.089)
Illegal Retail Market for Menthol E-cigarettes	-0.740***	-0.695***	-0.776***	-0.789***	-0.686***	-0.939***	-0.751***
	(0.077)	(0.130)	(0.098)	(0.117)	(0.100)	(0.155)	(0.096)
Illegal Street Market for Menthol E-cigarettes	-0.945***	-0.933***	-0.969***	-0.982***	-0.896***	-1.004***	-1.011***
	(0.093)	(0.155)	(0.118)	(0.138)	(0.123)	(0.177)	(0.118)
Log-likelihood	-10349	-3470	-6773	-5235	-5056	-4321	-5580
Observations	7668	2544	5076	3720	3948	3756	3912

Notes: Standard errors clustered at the individual level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A5: Mixed Logit Models Subgroup Analysis of Consumer Tobacco Product Choices

Immediate Choice Today	Full sample		By gender				By age				By product usage status			
			Male		Female		Age < 42		Age >= 42		Pure smokers		Dual users	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
ASC (Non-menthol cigarettes)	-0.176	3.117***	0.584*	3.942***	-0.884***	3.304***	-0.604**	2.917***	0.072	3.874***	-1.649***	4.347***	-0.225	3.336***
	(0.222)	(0.223)	(0.328)	(0.419)	(0.250)	(0.210)	(0.246)	(0.230)	(0.251)	(0.297)	(0.464)	(0.397)	(0.291)	(0.250)
ASC (Menthol cigarettes)	4.472***	3.098***	5.086***	3.310***	4.467***	3.430***	3.518***	3.218***	5.247***	4.022***	5.802***	2.588***	3.941***	3.415***
	(0.165)	(0.205)	(0.305)	(0.294)	(0.201)	(0.202)	(0.213)	(0.205)	(0.253)	(0.345)	(0.280)	(0.157)	(0.211)	(0.207)
ASC (Tobacco-flavored e-cigarettes)	-2.053***	3.908***	-1.264***	3.128***	-1.894***	3.703***	-1.011***	3.083***	-1.746***	3.071***	-3.164***	2.706***	0.277	2.683***
	(0.322)	(0.276)	(0.340)	(0.308)	(0.248)	(0.229)	(0.254)	(0.223)	(0.332)	(0.223)	(0.338)	(0.177)	(0.220)	(0.184)
ASC (Menthol-flavored e-cigarettes)	1.737***	3.292***	2.277***	3.056***	1.305***	3.106***	1.726***	2.820***	1.072***	3.515***	-0.422	3.796***	3.304***	2.387***
	(0.150)	(0.182)	(0.278)	(0.233)	(0.171)	(0.158)	(0.196)	(0.273)	(0.287)	(0.208)	(0.344)	(0.337)	(0.178)	(0.168)
Price ($)	-0.384***	0.654***	-0.380***	0.487***	-0.274***	0.377***	-0.287***	0.599*	-0.364***	0.333***	-0.408***	0.566***	-0.347***	0.668
	(0.034)	(0.180)	(0.039)	(0.106)	(0.029)	(0.099)	(0.043)	(0.309)	(0.029)	(0.057)	(0.049)	(0.198)	(0.075)	(0.544)
Illegal Retail Market for Menthol Cigarettes	-1.544***	1.192***	-1.694*	3.847	-1.921***	2.222**	-1.543***	0.868***	-1.871	6.132	-2.814*	7.881	-1.254***	0.604*
	(0.226)	(0.403)	(0.806)	(6.309)	(0.340)	(0.877)	(0.230)	(0.323)	(1.324)	(24.087)	(1.605)	(23.762)	(0.214)	(0.313)
Illegal Street Market for Menthol Cigarettes	-2.157***	1.688**	-3.780*	12.077	-2.513***	2.666***	-2.092***	1.008***	-6.712	51.182	-2.813***	4.546	-1.798***	0.991**
	(0.338)	(0.715)	(2.193)	(40.447)	(0.341)	(0.770)	(0.227)	(0.338)	(5.675)	(716.428)	(1.095)	(5.935)	(0.256)	(0.400)
Illegal Retail Market for Menthol E-cigarettes	-1.534***	0.484***	-1.473***	0.214	-1.611***	0.527**	-1.916***	1.550**	-1.426***	0.326	-1.762***	0.675*	-1.468***	0.133
	(0.136)	(0.187)	(0.204)	(0.243)	(0.177)	(0.260)	(0.369)	(0.739)	(0.184)	(0.239)	(0.277)	(0.355)	(0.140)	(0.316)
Illegal Street Market for Menthol E-cigarettes	-2.547***	2.623**	-2.890***	2.935*	-3.460**	5.370	-4.677***	9.607	-2.107***	0.879**	-2.280***	1.793*	-2.789***	3.042*
	(0.378)	(1.022)	(0.642)	(1.679)	(1.652)	(9.181)	(1.408)	(11.680)	(0.242)	(0.353)	(0.487)	(0.962)	(0.554)	(1.663)
Log-likelihood	-6736.713		-2246.324		-4401.198		-3466.884		-3216.111		-2811.385		-3744.139	
Observations	7668		2544		5076		3720		3948		3756		3912	

Notes: ASC = alternative specific constant. ASCs are assumed to follow normal distributions, price and legality variables are assumed to follow lognormal distributions. All random coefficients are assumed to be correlated. 500 Halton draws are used for simulation. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Source: Cornell online Discrete Choice Experiments 4/26-5/9, 2022.

A.5 Sensitivity Analysis of Excluding Subjects Living in a Place Where Menthol Cigarettes are Already Banned

As noted in the text, menthol cigarettes are currently banned in two states and 190 localities in the U.S. Footnote 5 explains that in our analysis sample of 639 subjects, we are able to identify 44 subjects who lived in a state or locality where menthol cigarettes were banned in April 2022 when our online survey was conducted. At that time, menthol cigarettes were banned in Massachusetts, but the California state-side ban had not yet been enacted.

Our online survey used a screening question to limit our sample to current adult smokers who indicated that they usually smoke menthol cigarettes. It is not clear why our sample includes subjects from places where menthol cigarettes are already banned. Many of them might be able to legally avoid the bans by making cross-border purchases of menthol cigarettes. Some might be making illegal menthol purchases. Either way, their DCE responses might be systematically different than other subjects' responses.

Tables A6 and A7 report conditional logit and mixed logit models for the full analysis sample and for the sub-sample when subjects who already faced a ban are dropped. The coefficient estimates are not sensitive.

Table A6: Sensitivity Analysis of Excluding Respondents Living in a Place where Menthol Cigarettes are already Banned (Conditional Logit Models)

Immediate Choice Today	**Full sample**	**Drop respondents facing a menthol ban**
ASC (Non-menthol cigarettes)	-0.200*	-0.174
	(0.115)	(0.120)
ASC (Menthol cigarettes)	1.904***	1.923***
	(0.095)	(0.098)
ASC (Tobacco-flavored e-cigarettes)	-0.667***	-0.687***
	(0.116)	(0.120)
ASC (Menthol-flavored e-cigarettes)	0.626***	0.629***
	(0.098)	(0.103)
Price ($)	-0.079***	-0.080***
	(0.006)	(0.006)
Illegal Retail Market for Menthol Cigarettes	-0.616***	-0.625***
	(0.060)	(0.063)
Illegal Street Market for Menthol Cigarettes	-0.866***	-0.891***
	(0.066)	(0.070)
Illegal Retail Market for Menthol E-cigarettes	-0.740***	-0.710***
	(0.077)	(0.080)
Illegal Street Market for Menthol E-cigarettes	-0.945***	-0.934***
	(0.093)	(0.097)
Log-likelihood	-10349	-9621
Observations	7668	7140

Notes: Standard errors clustered at the individual level are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table A7: Sensitivity Analysis of Excluding Respondents Living in a Place where Menthol Cigarettes are already Banned (Mixed Logit Models)

Immediate Choice Today	Full sample		Drop respondents facing a menthol ban	
	Mean	SD	Mean	SD
ASC (Non-menthol cigarettes)	-0.176	3.117***	-0.579***	3.142***
	(0.222)	(0.223)	(0.218)	(0.221)
ASC (Menthol cigarettes)	4.472***	3.098***	4.913***	3.295***
	(0.165)	(0.205)	(0.174)	(0.174)
ASC (Tobacco-flavored e-cigarettes)	-2.053***	3.908***	-0.938***	3.525***
	(0.322)	(0.276)	(0.207)	(0.196)
ASC (Menthol-flavored e-cigarettes)	1.737***	3.292***	1.646***	3.021***
	(0.150)	(0.182)	(0.148)	(0.137)
Price ($)	-0.384***	0.654***	-0.323***	0.443***
	(0.034)	(0.180)	(0.025)	(0.083)
Illegal Retail Market for Menthol Cigarettes	-1.544***	1.192***	-1.826***	2.511
	(0.226)	(0.403)	(0.429)	(1.579)
Illegal Street Market for Menthol Cigarettes	-2.157***	1.688**	-3.260***	5.282
	(0.338)	(0.715)	(0.852)	(4.873)
Illegal Retail Market for Menthol E-cigarettes	-1.534***	0.484***	-1.645***	1.532**
	(0.136)	(0.187)	(0.313)	(0.696)
Illegal Street Market for Menthol E-cigarettes	-2.547***	2.623**	-4.389***	9.692
	(0.378)	(1.022)	(1.262)	(12.331)
Log-likelihood	-6736.713		-6263.993	
Observations	7668		7140	

Notes: ASC = alternative specific constant. ASCs are assumed to follow normal distributions, price and legality variables are assumed to follow lognormal distributions. All random coefficients are assumed to be correlated. 500 Halton draws are used for simulation. Standard errors are in parentheses. * p<0.1, ** p<0.05, *** p<0.01.
Source: Cornell online Discrete Choice Experiments 4/26-5/9, 2022.

B. Literature Review of Studies that Compare SP and RP Data[1]

DCEs and the related contingent valuation method are used to collect SP data in a range of applications. DCEs are commonly used in marketing research and economics to provide predictions of consumer demand in scenarios that are not yet observed in actual markets. In addition to the tobacco product DCEs discussed below in section C, examples of the use of DCEs to study hypothetical market situations include studies of electricity markets (Blass, Lach, and Manski, 2010), health insurance markets (Kesternich, Hiess, McFadden, and Winter, 2013), labor markets (Mas, Alexandre and Pallais, 2017; Maestas et al. 2023), and firearms markets (Moshary, Shapiro, and Drango 2023). DCEs are also widely used in health economics to evaluate existing or prospective pharmaceutical products and health care treatment interventions (Ryan

[11] Sections B and D of the Online Appendix are mainly the same as sections of an Online Appendix of another paper by the same authors which is currently under review.

et al. 2007). Another large body of research uses DCEs and the related contingent valuation method to estimate willingness to pay for non-market goods like environmental quality.

Research that compares SP and RP data concludes that the external validity of SP data is much stronger in applications similar to familiar market goods. In a narrative review of DCE research, McFadden (2017) concludes that there is a "sharp reliability gradient":

> Forecasts that are comparable in accuracy to RP forecasts can be obtained from well-designed SP studies for familiar, relatively simple goods that are similar to market goods purchased by consumers, particularly when calibration to market benchmarks can be used to correct experimental distortions. However, studies of unfamiliar, complex goods give erratic, unreliable forecasts.

Asking smokers about tobacco products involves asking about familiar market goods. However, McFadden is skeptical about SP data on complex and unfamiliar environmental public goods. Although he does not discuss health care applications, by the same reasoning DCEs might not provide reliable data on unfamiliar pharmaceutical and health care treatment interventions.

Penn and Hu (2018) report a meta-analysis that provides quantitative evidence consistent with McFadden's (2017) conclusion that SP data are more reliable for familiar market goods. The meta-analysis used estimates from 132 studies that provided 908 observations of comparisons of SP and RP data. For studies including choice experiments that did not provide estimates of willingness to pay but did provide proportions of responses, Penn and Hu inferred lower-bound estimates of willingness to pay. Each observation is an estimate of the "calibration factor" (CF) which shows the ratio of willingness to pay estimated from SP data to the willingness to pay estimated from RP data. When SP and RP estimates are similar, the CF will be close to one. In the meta-analysis, about one quarter of the CFs are between 0.81 and 1.2. The distribution of CFs is skewed right showing a tendency for SP willingness-to-pay estimates to be larger than RP estimates, sometimes to a large extent. The median CF is 1.94, implying that for almost half of the observations the SP estimate is over twice as large as the RP estimate. Penn and Hu estimate regression models of the effects of study characteristics on CFs. The results imply that compared to studies of public goods, studies of private goods find lower CFs. The results also imply that compared to other hypothetical elicitation methods, CFs are lower for studies that used choice experiments. The meta-analysis empirical results are consistent with McFadden's conclusion that although SP data from DCEs are reliable for private goods, there is a sharp reliability gradient for SP data on willingness to pay for public goods.

In the remainder of this section, we discuss examples of studies that compare SP and RP data on health-related choices. Quaife et al (2018) review a number of studies of health-related choices that focused on the external validity of DCEs (by comparing SP and RP estimates) and conclude that DCEs provide moderate levels of external validity in terms of matching actual choices. de Bekker-Grob et al (2020) find that when measured at the individual level, stated preferences in a DCE about vaccinations predict 91 percent of actual choices. Telser and Zweifel (2007) examine the external validity of a DCE focused on decisions about a harm reduction product (hip protectors for accidental falls). They compare the willingness to pay for risk reduction that was derived from the DCE to other measures of willingness to pay for the same risk reduction derived from established alternatives that used revealed preference data. The comparison supports a high level of convergent validity. Linley and Hughes (2013) examine hypothetical decisions about new medicine approvals and find that the predicted probabilities of recommending new medicines derived from the DCE match well with the cumulative probability of actual positive recommendations (though the ability of the DCE to discriminate between individual new medicines was limited). Mahammad et al (2017) use a DCE and estimate a mixed logit based on the hypothetical choices with respect to type of tuberculosis treatment (or none) in response to each treatment having six treatment attributes. They compare these choices with actual choices and find strong external validity and the degree of accuracy depends on the distributional assumptions used in the mixed logit models with some models. Kesternich et al (2013) implement a DCE to analyze Medicare part D choices and compare these results to those that emerge from analysis of actual choices. They conclude that hypothetical choice experiments are useful in studying insurance choices as hypothetical behavior is related to actual behavior. They find that the coefficients that emerge in the DCE experiment are of the same sign as the coefficients that are estimated from market behavior. They note that the magnitudes of the coefficients are quite similar and do not find significant differences between hypothetical and real choices between different attributes of the insurance scenarios. They do find a higher willingness to pay for insurance in the hypothetical market and thus higher insurance take-up rates but attribute this to the nature of the default option in the DCE.

C. Discussion of Public Health Research on Menthol Prohibitions

Mills et al. (2024) reports the results of a systematic review and meta-analysis of public health research on menthol prohibitions. The systematic review covers 78 studies; the quantitative meta-analysis uses results from 16 of those studies. The systematic review of public health research includes studies of the impacts of menthol prohibitions on real-world tobacco use behavior and studies of hypothesized impacts of menthol prohibitions on tobacco use behavior. Almost all of the studies of hypothesized impacts asked about smokers' intentions in the event of a menthol prohibition. Two studies included in the systematic review that conducted discrete choice experiments are discussed in section D.

In the quantitative meta-analysis, estimates from the pooled results are that after a menthol prohibition: 24 percent of smokers quit, 50 percent of smokers switch to non-menthol cigarettes, 12 percent of smokers switch to other flavored tobacco products, and 24 percent of smokers continue to smoke menthol cigarettes (Mills et al. 2024). The estimated impacts sum to over 100 percent because not all of the studies included in the quantitative meta-analysis measured all four outcomes. The meta-analysis uses three estimates (all hypothesized impacts) for the outcome of switching to flavored e-cigarettes or other flavored tobacco products and nine estimates (three real-world impacts and six hypothesized impacts) for the outcome of continuing to smoke menthol cigarettes. The systematic review also suggests that real-world and hypothesized quitting and switching rates were similar, and that national prohibitions appear to be more effective than local or state menthol bans.

Studies of the impacts on real-world smoking behavior use data from state and local menthol prohibitions in the U.S. and national bans in Canada and the E.U. In the main text we note several general limitations to these studies. First, as pointed out by Carpenter and Nguyen (2021), most public health research on the prohibition of menthol cigarettes in Canada and the E.U. does not use quasi-experimental methods required for causal inference. Second, estimates of the impacts of state and local prohibitions might not generalize to national prohibitions because of the much different opportunities to avoid the prohibition through cross-border purchases. This concern is reinforced by the conclusion from Mills et al. (2024) that state and local prohibitions are not as effective as national prohibitions. Third, estimates of the impacts of the E.U. prohibitions might not generalize to the U.S. because the E.U. prohibition is more limited than the FDA proposal.

Public health research on the potential impact of prohibitions to create illegal markets is especially limited. As we note in the text, research findings on prohibitions in Canada and the E.U. might not generalize to U.S. illegal markets for menthol. The pre-prohibition menthol market shares in Canada and the E.U. were lower than the U.S. menthol market share and far lower than the 85 percent menthol share among U.S. Black smokers. As a result, Canadian and E.U. illegal menthol markets may be thin, with their size limited by high prices and high transactions costs (Jacobson 2004, Cook et al. 2007, Cutler and Donohoe 2024).

The evidence from public health research on menthol prohibitions in the U.S. provides very limited evidence about the potential for illegal menthol markets. The narrative systematic review discusses three U.S. studies of real-world tobacco use behaviors and 10 U.S. studies of hypothesized behaviors after a menthol prohibition (Mills et al. 2024 references 26, 28, and 68 study real-world behaviors and references 31-40 study hypothesized behaviors). The studies of real-world behaviors conduct before-and-after analysis of small samples (menthol smokers N = 14 in reference 26, N = 81 in reference 28, and N = 120 in reference 68) of special populations in Boston and San Francisco. None of these estimates is included in the quantitative meta-analysis. In six of the 10 studies of hypothesized behaviors, respondents were not given the opportunity to indicate that they would try to continue using menthol cigarettes after a prohibition. In two of the other four studies, respondents were first asked questions that did not allow for an illegal market, but subsequent questions allowed for the possibility. The question sequence and framing in these surveys might have primed respondents against reporting intentions to make illegal purchases. In one study, the response options included "none of the above" but did not specifically mention continuing to smoke menthols. In one study, when asked how they might respond to a menthol prohibition, one of the response options was that they would find a way to continue menthol cigarette use, which the researchers interpreted as an intention to make purchases from an illegal market. Estimates from three of the U.S. studies of hypothesized behaviors are included in the quantitative meta-analysis. In comparison to the pooled estimate that 24 percent of menthol smokers would continue to smoke menthol cigarettes after a prohibition, the three estimates from the included U.S. studies were 24 percent (reference #32), 34 percent (reference #39), and 55 percent (reference #40).

As a final comment on public health research on menthol prohibitions, we emphasize the distinction between public health studies of hypothesized behaviors versus economics-based discrete choice experiments. Public health studies of hypothesized behaviors use broad questions about intentions, sometimes with qualifications such as asking what respondents "might" do or "most likely" would do. In contrast, DCEs are designed to create a hypothetical but realistic market and ask consumers about their choices in such a market. In section B we review evidence on the external validity of stated preference data from DCEs compared to revealed preference data on behavior in real-world markets. Because of the differences between public health studies of hypothesized behaviors and DCEs, the evidence in section B does not necessarily apply to the public health studies of hypothesized behaviors. However, Mills et al. (2024) also note that the estimates of hypothesized and real-world quitting and switching rates were similar.

D. DCE Studies of Tobacco Product Choices

A growing body of research conducts DCEs to study the determinants of consumer choices about tobacco products. Table D1 lists recent DCE studies of tobacco product choices published in economics journals. Table D1 also include two DCE studies published in a public health journal, because these two studies explore the role of menthol in tobacco product choices.

Table D1. Recent DCE Studies of Tobacco Product Choices, Published in Economics Journals

Authors	Year	Journal
Buckell, Hensher, and Hess	2021	Health Economics
Buckell and Hess	2019	Journal of Health Economics
Buckell, Marti, and Sindelar	2019	Tobacco Control
Buckell, et al.	2023	Tobacco Control
Kenkel, Peng, Pesko, and Wang	2020	Health Economics
Marti, Buckell, Maclean, and Sindelar	2019	Economic Inquiry

Table D1 includes two studies that report results from exercises that use a combination of RP and SP data to develop a calibrated model that is grounded in real-world behavior. Kenkel, Peng, Pesko, and Wang (2020) report in an Online Appendix the results of a conditional logit model estimated using a combination of SP data from a DCE and RP data from the DCE subjects' responses about their prior use of e-cigarettes, combustible cigarettes, and nicotine replacement products. The estimated scale parameter is close to 1, suggesting that the scales in the RP and SP data are similar. The estimated coefficients on the tobacco product attributes show the same patterns as in the model based on SP data only.

Buckell and Hess (2019) report the results of a more in-depth investigation of combining SP and RP data on tobacco product choices. In a model estimated using combined SP and RP data they estimate a scale parameter greater than 1, consistent with the argument that in SP data subjects overstate the impact of interventions which leads to elasticities that are biased upwards. In terms of forecasts from the models, they find that compared to calibrated models the uncalibrated forecasts under-predict cigarette choices and over-predict e-cigarette choices. They conclude that appropriately calibrated choice models "provide better quality empirical evidence for policymakers." (Buckell and Hess 2019, p. 100)

The remaining tobacco product DCEs listed in Table D1 do not provide in-depth discussions of external validity. However, the results of the studies are consistent with predictions from health economic models of consumer behavior, which broadly supports their external validity. For example, Marti, Buckell, Maclean, and Sindelar (2019) conduct a DCE to study how smokers' product choices are affected by variations in the perceived healthiness and cessation effectiveness of e-cigarettes, as well as by bans on smoking in public places and prices. Buckell, Hensher, and Hess (2021) use SP data from a DCE combined with a latent variable approach to model addiction. They find that more addicted smokers are unwilling to switch to e-cigarettes. The studies in Table D1 all find that smokers' product choices respond to cigarette and e-cigarette prices; the estimated price-responsiveness is generally consistent with the large body of econometric estimates from observational (RP) data (DeCicca, Kenkel, and Lovenheim 2022).

DCE studies of tobacco product choices are also published in inter-disciplinary public health journals, including journals focusing on tobacco such as *Tobacco Control* and the *Journal of Nicotine and Tobacco Research*.[2] Most of these articles are not included in our Table D1 review because they have a different focus for a different audience. For example, Shang, Huang, Chaloupka, and Emery (2018) focus on the roles flavors, e-cigarette device type, and e-cigarette warning labels play in youth stated preferences to try

[2] Regmi, Kaphle, Timilsina, and Tuha (2018) report a systematic review of peer-reviewed studies published from 2000 – 2016 that used DCE methods in tobacco control. Of the 12 studies included in their review, 4 were published in health economics journals. Because these 4 studies focused on pharmaceutical smoking cessation products, they are not included in Table D1.

e-cigarettes. Subjects were presented with e-cigarette products with varying attributes but were not given the alternative of choosing combustible cigarettes.

Table D1 includes two DCE studies published in the specialty public health journal *Tobacco Control* – Buckell, Marti, and Sindelar (2019) and Buckell et al. (2023) – that use data from DCEs to predict the impact of a menthol prohibition in the U.S. However, the choice sets in these experiments did not include an alternative corresponding to an illegal purchase; by construction the studies' policy simulations do not allow for illegal menthol markets. Buckell, Marti, and Sindelar (2019) report results from a DCE conducted with subjects who were U.S. adult smokers or recent quitters. In the choice tasks, subjects were asked to indicate their first choice and second choice of cigarettes, e-cigarettes, or none of these. Four attributes of cigarettes and e-cigarettes were experimentally varied: flavor, nicotine level, health impact, and price. In the econometric choice model, the estimated alternative specific constant for menthol cigarettes is negative, implying that for the average subject menthol is an undesirable product attribute. This is not surprising because most smokers in the U.S. prefer non-menthol cigarettes. The impact of a menthol prohibition was simulated by using the estimated model to predict choice shares when menthol cigarettes are not in the choice set. Using this method, menthol cigarette prohibition is predicted to reduce the share of cigarette choices by 5.2 percentage points, with most of the choices shifting to e-cigarettes. If menthol e-cigarettes are also prohibited, the reduction in the cigarette share falls to 3.5 percentage points. In the policy simulations, the model-based approach of removing menthol cigarettes and/or menthol e-cigarettes from the choice set is equivalent to assuming that the menthol prohibition is complete and there are no illegal menthol markets.

Buckell et al. (2023) report results from a DCE conducted with subjects who were U.S. adult smokers who reported a low interest in quitting. In the choice tasks, subjects were shown the alternative of their usual cigarette, pod e-cigarettes, and disposable e-cigarettes. Five attributes of e-cigarettes were experimentally varied: flavors, nicotine levels, healthier than cigarettes, helpful for quitting smoking, and price. The attributes of the cigarette alternative did not vary; in particular, the flavor was always described as "your usual cigarette flavor." The online data supplement to Buckell et al. (2023) explains the policy simulations: "First, a baseline simulation is estimated to represent the current state of the world under current policies. Then policy impacts are simulated and compared to the baseline. Simulations use the estimated parameters from the model and manipulate the attribute values to mimic the impact of a given policy. For example, changing the cigarette flavor from menthol to tobacco simulates a menthol cigarette flavor ban." This model-based approach is equivalent to assuming that the menthol prohibition is complete and there are no illegal menthol marketa.

As a final comment, we note that outside economics journals the term "discrete choice experiment" is used in both a broad and narrow sense. In the broad sense, DCE has been used to describe various surveys that asks subjects to make choices with random assignment of the descriptions of the alternatives. In contrast, Louviere, Flynn, and Carson (2010) define DCEs as being necessarily grounded in random utility theory. Some of the studies published in public health journals do not fit this narrow definition of DCEs, even though the studies use the term to describe their research method. For example, Reynolds, Popova, Ashley et al. (2022) report a DCE about very low nicotine cigarettes (VLNCs) that asked respondents which message would most motivate them and least motivate them to quit smoking; the message attributes varied in terms of content about VLNCs and the source of the message. Subjects did not make choices between products. This is a study of consumer perceptions of message effectiveness which cannot be grounded in random utility theory. A related concern is that in some DCEs where subjects make choices between tobacco products, the product attributes are described in terms of consumer perceptions rather than observable characteristics of the products and/or the product marketplace. For example, Shang, Weaver, White, et al.

(2020) report a DCE about e-cigarettes that included “less harmful to health than cigarettes” and “effective for helping people quit” as product attributes. The Marti et al. (2019, 2023) studies included in Table D1 also use this approach. Although the results of these studies provide information about the relative importance of these perceptions on tobacco product choices, the results are less useful for policy analysis because perceptions are not directly policy manipulable. For this reason, in our DCE we describe the policy-manipulable attribute of warning labels which mention or do not mention health effects.

E. Empirical Evidence on the Internal and External Validity of SP Data Collected through the Cornell DCE

E.1. Validity Checks of the Quality of the Stated Preference Data

In sub-sections E.1 and E.2 we present empirical evidence on the internal and external validity of the SP data we collected through our DCE. Like other experimental research designs, the randomly assigned variation in product attributes in DCEs provides an internally valid research design to estimate the causal treatment effects of product attributes on subjects' stated preferences for tobacco products and quit attempts. However, because SP data are the subjects' responses about hypothetical choices, subjects might not provide thoughtful and meaningful responses that provide useful information about the actual choices they would make in real-world markets. In this section E.1, we report the results of validity checks on the quality of our SP data and the implications for the empirical results reported in the text of the paper.

As an overview of the sensitivity of the empirical results to the validity checks, Table E1 reports the sensitivity of conditional logit models of tobacco product choices to alternative approaches to improve SP data validity. Column (0) reports a baseline conditional logit model estimated over the same sample used in estimation of the main text models. Columns (1) – (2) report conditional logit models estimated using sample restrictions to improve the quality of the SP data. Column (3) reports a conditional logit model estimated using a combination of SP and revealed preference (RP) data. In columns (1) – (2) the point estimates of the alternative specific constants (ASCs) and the product attribute parameters tend to be very similar to the baseline model parameter estimates in column (0). The combined SP + RP model reported in column (3) yields estimated ASCs that are slightly different than the estimated ASCs in the baseline column (0) model. We will discuss the SP + RP data model results in more detail in the next sub-section E.2 of this Appendix.

In this sub-section we focus on sample restrictions that might improve the quality of our SP data. First, we examine data on the length of time subjects spent answering the DCE choice tasks, to identify possible "speedsters" who provided lower-quality responses. Figure E1 shows the distribution of time spent on the choice tasks. The median and mode times spent on the choice tasks are 5.7 and 4.7 minutes (there are multiple modes, the minimum mode is 4.2 and the maximum mode is 5.2), respectively. Each subject completed 12 immediate choice tasks and 12 six-months-from-now choice tasks; the six-months-from-now choice tasks might be easier to complete quickly because they presented the subject with the same choices as in the preceding immediate choice task. For the model reported in column (1) of Table E1, we drop responses from 8 speedster subjects who completed the choice tasks in under 2 minutes.

Second, we examine data on the extent to which subjects paid attention to attribute variation across choice tasks. After subjects completed the choice tasks, we asked subjects which attributes varied across the tasks. Figure E2 shows the fraction of subjects who correctly indicated that the attribute in question varied. Although substantial fractions of the responses about attribute variation were incorrect, the results suggest that subjects paid the most attention to the price attribute of menthol cigarettes, which was the most common tobacco product choice. The patterns of attentiveness across attributes and products are consistent with rational decisions to pay the most attention to the attributes and products that matter to their preferences. We also note that there is an ambiguity in our measure of attentiveness. As noted above, the six-months-from-now choice task was always identical to the preceding immediate choice task, i.e., in those pairs of tasks the attributes did not vary. For the model reported in column (2) of Table E1, we drop responses from 106 subjects who were inattentive to variation in the attributes of menthol cigarettes.

E.2. Improving Data Quality by Combining SP and RP Data

In this sub-section we discuss the approach reported in column (3) of Table E1, where we estimate a conditional logit model of tobacco product choices using a combination of SP and RP data. In his monograph on econometric analysis of discrete choice data, Train (2002, pp. 174-175) discusses the advantages and disadvantages of SP and RP data:

> Revealed preference data have the advantage that they reflect actual choices…. However, RP data are limited to the choice situations and attributes of alternatives that currently exist or have existed historically. Often a researcher will want to examine people's responses in situations that do not currently exist, such as the demand for a new product. RP data are simply not available for these new situations.
>
> Stated-preference data complement revealed-preference data…. The limitations of SP data are obvious: what people say they will do is often not the same as what they actually do. People might not know what they would do if a hypothetical situation were real. Or they might not be willing to say what they would do.

Train suggests that by combining RP and SP data, "the advantages of each can be obtained while mitigating the limitations. The SP data provide the needed variation in attributes, while the RP data ground the predicted shares in reality." He outlines the approach we take, where we use our DCE's subjects' responses about their tobacco product use and quit attempts over the past year as RP data to calibrate our model.

To construct RP choices, we propose a probabilistic classification rule that uses information from the background survey of respondents' smoking behaviors. The process and classification rule are described in table E3. The construction takes three steps, first, we classify respondents as either choosing cigarettes or e-cigarettes. Second, we classify respondents as either choosing menthol flavor or non-menthol flavor. Third, we classify respondents as either choosing quit or not quit. To classify the choices of cigarettes and e-cigarettes, we use information of subjects' vaping history, vaping status, and smoking status. Among all the subjects, those who have never vaped are classified as choosing cigarettes, among those who have ever vaped, if they currently do not vape at all, they are classified as choosing cigarettes, for those who vape ever day and someday, we classify their choices according to their smoking status. Specifically, among those who vape ever day, if they also smoke every day, then their probabilities of choosing cigarettes versus e-cigarettes are 50% versus 50%, if they smoke someday (say m days out of 30 days), then their probabilities of choosing cigarettes versus e-cigarettes are $m/(30+m)$ versus $30/(30+m)$. For those who vape someday, the probability rules are similar, and the choice of quitting is constructed following a similar rule, details are reported in table E2.

After obtaining RP choices, we jointly estimate the model with SP and RP data. We assume that the random error term for the RP data follows an i.i.d. type I extreme value distribution with scale parameter normalized to 1, and the scale parameter for the SP data is given by λ. The choice probabilities for individual i chooses alternative j in the RP and SP data can then be written as:

$$P_{ij}^{RP} = \frac{e^{(\alpha_j + X_{ij}^{RP}\beta)}}{\sum_k e^{(\alpha_k + X_{ik}^{RP}\beta)}}$$

$$P_{ij}^{SP} = \frac{e^{[\lambda(\alpha_j + X_{ij}^{SP}\beta)]}}{\sum_k e^{[\lambda(\alpha_k + X_{ik}^{SP}\beta)]}}$$

In the joint estimation we maximize the joint likelihood function:

$$L(\alpha, \beta, \lambda) = \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij}^{RP} \log P_{ij}^{RP} + \sum_{i=1}^{N} \sum_{j=1}^{J} y_{ij}^{SP} \log P_{ij}^{SP}$$

The results are presented in column (4) of table E1. The estimated scale parameter λ is 0.8, suggesting that the scales in the RP and SP data are similar. The estimated coefficients on the product attributes show similar patterns as in the column (0) model based on SP data only. We further use the model to predict the choice shares under a status quo scenario and various counterfactual policy scenarios considered in the paper. We report the predicted product shares and the size of the market relative to status quo in table E3. We see that the calibrated model predicts larger shares of menthol cigarettes and smaller shares of quitting, while the represented size of market relative to status quo is fairly similar to those predicted by the model using only SP data. Overall, we interpret the results from the calibrated model as supporting the usefulness of our SP data.

Table E1. Estimation Results from Conditional Logit Models

Variables	(0)	(1)	(2)	(3)
ASC (Non-menthol cigarettes)	-0.200*	-0.184	-0.054	-0.166
	(0.115)	(0.116)	(0.125)	(0.277)
ASC (Menthol cigarettes)	1.904***	1.914***	2.100***	2.429***
	(0.095)	(0.095)	(0.101)	(0.315)
ASC (Tobacco-flavored e-cigarettes)	-0.667***	-0.668***	-0.607***	-0.749***
	(0.116)	(0.116)	(0.126)	(0.217)
ASC (Menthol-flavored e-cigarettes)	0.626***	0.637***	0.762***	0.813**
	(0.098)	(0.099)	(0.106)	(0.395)
Price ($)	-0.079***	-0.080***	-0.091***	-0.102***
	(0.006)	(0.006)	(0.006)	(0.008)
Illegal Retail Market for Menthol Cigarettes	-0.616***	-0.628***	-0.708***	-0.781
	(0.060)	(0.061)	(0.069)	(0.549)
Illegal Street Market for Menthol Cigarettes	-0.866***	-0.880***	-0.992***	-1.087***
	(0.066)	(0.067)	(0.075)	(0.081)
Illegal Retail Market for Menthol E-cigarettes	-0.740***	-0.745***	-0.847***	-0.911***
	(0.077)	(0.078)	(0.086)	(0.263)
Illegal Street Market for Menthol E-cigarettes	-0.945***	-0.958***	-1.140***	-1.170***
	(0.093)	(0.094)	(0.106)	(0.300)
Scale Parameter of SP Data				0.808***
				(0.224)
Log-likelihood at convergence	-10349	-10224	-8512	-10959
Respondents	639	631	533	639
Observations	38340	37860	31980	38340

Notes: Standard errors are clustered at the individual level (except for column (4)). *** p<0.01, ** p<0.05, * p<0.1.

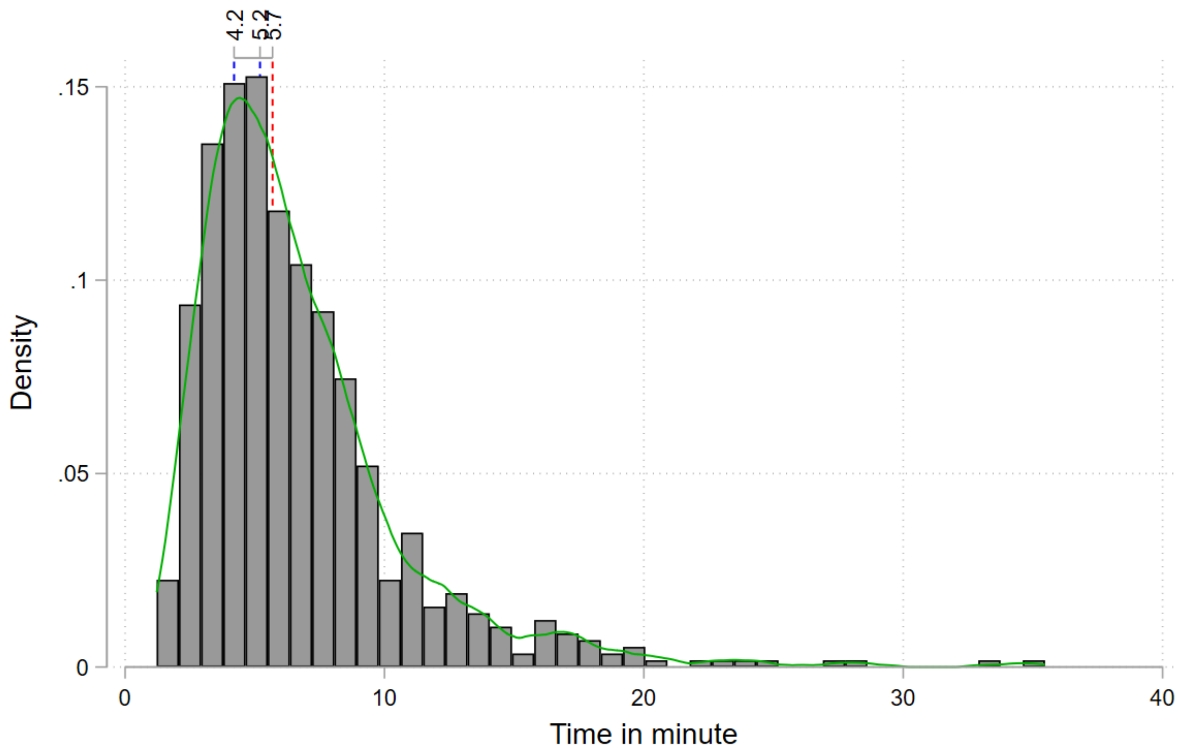


Figure E1. Distribution of Subject's Time Spent on the Choice Tasks

Notes: The blue and red dash lines indicate the modes and median of time spent on the DCE section respectively. There are multiple modes in the distribution, we show the minimum mode (4.2) and maximum mode (5.2).

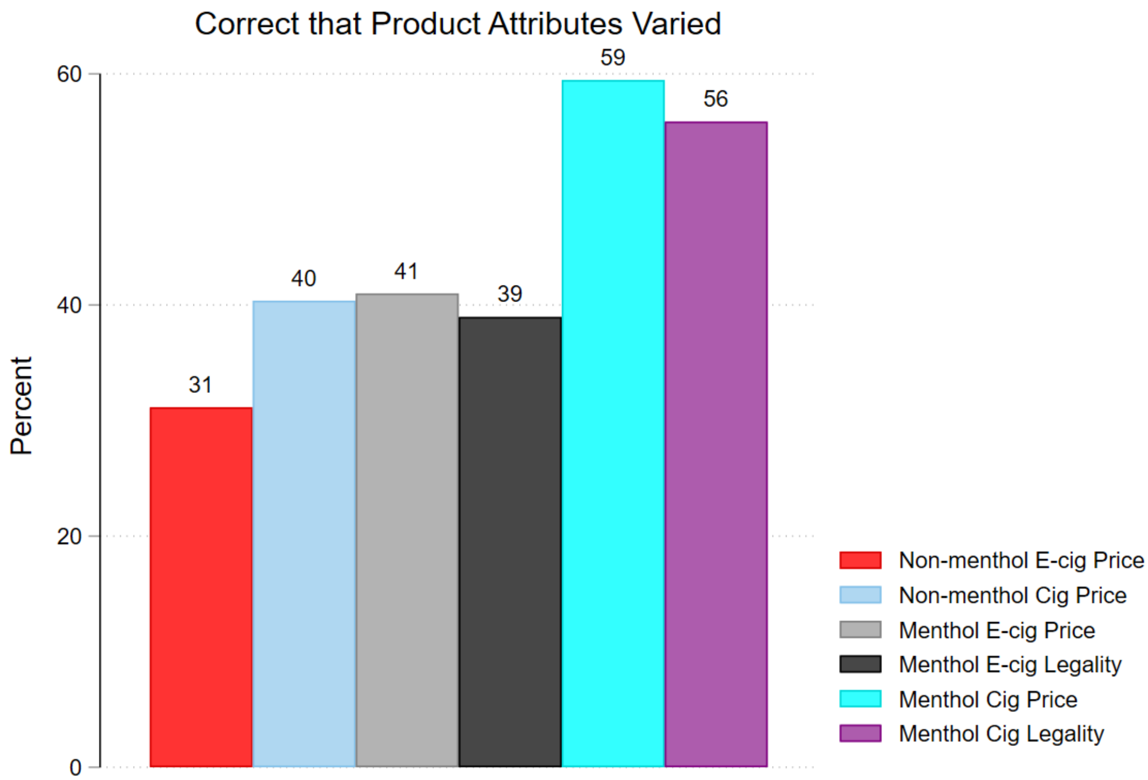


Figure E2. Fraction of Subjects that Were Correct that the Attribute in Question Varied.

Notes: Subjects are allowed to give correct but inconsistent answers, for example, they could select both price varied across scenarios and price was always the same. We use a strict criterion that the subject was correct that one attribute in question varied if she only selects the attribute varied and does not select the attribute was always the same.

Table E2. A Probabilistic Classification Rule to Construct Revealed Preference Choices

Step 1: classification of choices between cigarettes and e-cigarettes					
Subject	Vaping history	Vaping status	Smoking status	Probability of choosing cigarettes	Probability of choosing e-cigarettes
N=639	Vape ever (Yes) N=478	Vape every day (N=71)	Smoke every day (N=35)	50%	50%
			Smoke someday (m out of 30 days) (N=36)	$m/(30+m)$	$30/(30+m)$
		Vape someday (n out of 30 days) (N=255)	Smoke every day (N=115)	$30/(30+n)$	$n/(30+n)$
			Smoke someday (m out of 30 days) (N=140)	$m/(m+n)$	$n/(m+n)$
		Not at all (N=152)	Smoke every day (N=111) Smoke someday (N=41)	100%	0
	Vape every (No) N=161		Smoke every day (N=109) Smoke someday (N=52)		

Step 2: classification of choices between menthol flavor and non-menthol flavor		
Among those classified as choosing cigarettes		
During last 12 months, did you ever switch to non-menthol cigarettes to try to quit?	Yes	Choose non-menthol cigarettes
	No	Choose menthol cigarettes
Among those classified as choosing e-cigarettes		
Usually vape menthol flavored e-cigarettes	Yes	Choose menthol e-cigarettes
	No	Choose non-menthol e-cigarettes

Step 3: classification of choices of quitting			
Subject	Quitting intention	Quitting history	Probability of choosing quitting
N=639	Plan to quit (Yes) N=416	Tried quitting (N=289) Didn't try to quit (N=127)	1/26
	Plan to quit (No) N=223	Tried quitting (N=38)	1/52
		Didn't try to quit (N=185)	0

Table E3. Predicted Choice Shares under Status Quo and Counterfactual Scenarios

Policy Scenario	Non-menthol Cigs	Menthol Cigs	Tabacco-flavored E-cigs	Menthol-flavored E-cigs	Quitting
Status quo	0.045	0.599	0.040	0.190	0.126
Illegal Retail Market for Menthol Cigs					
1. 50% lower price for illegal products	0.056	0.514	0.048	0.229	0.153
2. No price change	0.067	0.410	0.059	0.279	0.186
3. 50% higher price for illegal products	0.077	0.316	0.068	0.323	0.215
Illegal Retail Market for Menthol Cigs & E-cigs					
4. 50% lower price for illegal products	0.063	0.582	0.054	0.128	0.173
5. No price change	0.080	0.490	0.071	0.135	0.224
6. 50% higher price for illegal products	0.098	0.400	0.087	0.137	0.278
Illegal Street Market for Menthol Cigs					
7. 50% lower price for illegal products	0.064	0.439	0.056	0.265	0.176
8. No price change	0.075	0.339	0.065	0.312	0.208
9. 50% higher price for illegal products	0.084	0.256	0.074	0.352	0.234
Illegal Street Market for Menthol Cigs & E-cigs					
10. 50% lower price for illegal products	0.077	0.524	0.067	0.121	0.212
11. No price change	0.095	0.430	0.084	0.124	0.266
12. 50% higher price for illegal products	0.114	0.343	0.101	0.122	0.320

Policy Scenario	Size of market relative to status quo		
	Menthol Cigs	Menthol-flavored E-cigs	Quitting
Illegal Retail Market for Menthol Cigs			
1. 50% lower price for illegal products	0.858	1.208	1.208
2. No price change	0.684	1.471	1.471

3. 50% higher price for illegal products	0.528	1.704	1.704
Illegal Retail Market for Menthol Cigs & E-cigs			
4. 50% lower price for illegal products	0.971	0.675	1.368
5. No price change	0.817	0.713	1.774
6. 50% higher price for illegal products	0.668	0.721	2.196
Illegal Street Market for Menthol Cigs			
7. 50% lower price for illegal products	0.732	1.396	1.396
8. No price change	0.566	1.644	1.644
9. 50% higher price for illegal products	0.427	1.853	1.853
Illegal Street Market for Menthol Cigs & E-cigs			
10. 50% lower price for illegal products	0.875	0.636	1.673
11. No price change	0.718	0.654	2.106
12. 50% higher price for illegal products	0.573	0.642	2.534

References

Blass, A. A., Lach, S., & Manski, C. F. (2010). Using elicited choice probabilities to estimate random utility models: Preferences for electricity reliability. *International Economic Review*, *51*(2), 421-440.

Buckell, J., & Hess, S. (2019). Stubbing out hypothetical bias: improving tobacco market predictions by combining stated and revealed preference data. *Journal of Health Economics*, *65*, 93-102.

Buckell, John, Joachim Marti, Jody L. Sindelar (2019). "Should Flavours Be Banned in Cigarettes and E-cigarettes? Evidence on Adult Smokers and Recent quitters from a Discrete Choice Experiment. *Tobacco Control* 28:168–75.

Buckell, J., Hensher, D. A., & Hess, S. (2021). Kicking the habit is hard: A hybrid choice model investigation into the role of addiction in smoking behavior. *Health Economics*, 30(1), 3-19.

Buckell, John, Lisa M. Fucito, Suchitra Krishnan-Sarin, Stephanie O'Malley, and Jody L. Sindelar (2022). "Harm Reduction for Smokers with Little to No Quit Interest: Can Tobacco Policies Encourage Switching to E-cigarettes?" *Tobacco Control*

de Bekker-Grob, E. W., Donkers, B., Bliemer, M. C., Veldwijk, J., & Swait, J. D. (2020). Can healthcare choice be predicted using stated preference data? *Social Science & Medicine*, *246*, 112736.

DeCicca, P., Kenkel, D., & Lovenheim, M. F. (2022). The economics of tobacco regulation: a comprehensive review. *Journal of Economic Literature*, *60*(3), 883-970.

Lancsar, E., & Louviere, J. (2006). Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics*, *15*(8), 797-811.

Linley, W. G., & Hughes, D. A. (2013). Decision-makers' preferences for approving new medicines in wales: a discrete-choice experiment with assessment of external validity. *Pharmacoeconomics*, *31*, 345-355.

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling*, *3*(3), 57-72.

Kenkel, D. S., Peng, S., Pesko, M. F., & Wang, H. (2020). Mostly harmless regulation? Electronic cigarettes, public policy, and consumer welfare. *Health Economics*, *29*(11), 1364-1377.

Kesternich, I., Heiss, F., McFadden, D., & Winter, J. (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, *32*(6), 1313-1324.

Maestas, N., Mullen, K. J., & Powell, D. (2023). The Effect of Population Aging on Economic Growth, the Labor Force, and Productivity. *American Economic Journal: Macroeconomics*, *15*(2), 306-332.

Marti, J., Buckell, J., Maclean, J. C., & Sindelar, J. (2019). To "vape" or smoke? Experimental evidence on adult smokers. *Economic Inquiry*, *57*(1), 705-725.

Mas, A., & Pallais, A. (2017). Valuing alternative work arrangements. *American Economic Review*, *107*(12), 3722-3759.

McFadden, Daniel (2017). “Stated Preference Methods and Their Applicability to Environmental Use and Non-use Valuations.” *In Contingent Valuation of Environmental Goods: A Comprehensive Critique.* Daniel McFadden and Kenneth Train, editors. Edward Elgar Publishing. Elgaronline.com.

Mohammadi, T., Bansback, N., Marra, F., Khakban, A., Campbell, J.R., FitzGerald, J.M., Lynd, L.D., & Marra,C.A. (2017) Testing the External Validity of a Discrete Choice Experiment Method: An Application to Latent Tuberculosis Infection Treatment, Value in Health, 20 (7), 969-975.

Moshary, S., Shapiro, B., & Drango, S. (2023). *Preferences for firearms and their implications for regulation* (No. w30934). National Bureau of Economic Research.

Penn, J. M., & Hu, W. (2018). Understanding hypothetical bias: An enhanced meta-analysis. *American Journal of Agricultural Economics*, *100*(4), 1186-1206.

Quaife, M., Terris-Prestholt, F., Di Tanna, G. L., & Vickerman, P. (2018). How well do discrete choice experiments predict health choices? A systematic review and meta-analysis of external validity. *The European Journal of Health Economics*, *19*, 1053-1066.

Regmi, K., Kaphle, D., Timilsina, S., Tuha, N. A. A. (2018). Application of discrete-choice experiment methods in tobacco control: A systematic review. *PharmacoEconomics Open*, 2, 5-17.

Reynolds, R. M., Popova, L., Ashley, D. L., Henderson, K. C., Ntansah, C. A., Yang, B., ... & Thrasher, J. (2022). Messaging about very low nicotine cigarettes (VLNCs) to influence policy attitudes, harm perceptions and smoking motivations: a discrete choice experiment. *Tobacco Control*.

Rousu, M.C., Marette, S., Thrasher, J. F., Lusk, J. L. (2014). The economic value to smokers of graphic warning labels on cigarettes: Evidence from combining market and experimental auction data. J*ournal of Economic Behavior and Organization,* 108, 123-134.

Ryan, M., Gerard, K., & Amaya-Amaya, M. (Eds.). (2007). *Using discrete choice experiments to value health and health care* (Vol. 11). Springer Science & Business Media.

Shang, C., Huang, J., Chaloupka, F. J., and Emery, S. L. (2018). The impact of flavor, device type, and warning messages on youth preferences for electronic nicotine delivery systems: evidence from an online discrete choice experiment. *Tobacco Control* 27: e152-e159.

Shang, C., Weaver, S. R., White, J. S., Huang, J., Nonnemaker, J., Cheng, K. W., & Chaloupka, F. J. (2020). E-Cigarette product preferences among adult smokers: a discrete choice experiment. *Tobacco Regulatory Science*, *6*(1), 66.

Telser, H., & Zweifel, P. (2007). Validity of discrete-choice experiments evidence for health risk reduction. *Applied Economics*, *39*(1), 69-78.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.