Quantifying Internalities, with an Application to Tobacco Regulations

Don Kenkel, PhD
Department of Economics
Brooks School of Public Policy
Cornell University

Alan Mathios, PhD
Department of Economics
Brooks School of Public Policy
Cornell University

Hua Wang, PhD
Brooks School of Public Policy
Cornell University

Sen Zeng
Department of Economics
Cornell University

February 2025

Preliminary draft prepared for possible presentation at the NBER conference Economic Analysis of Regulation, to be held May 1, 2025 – Cambridge, MA. Produced with the help of a grant to XXX University from Global Action to End Smoking (formerly known as Foundation for Smoke-Free World), an independent, U.S. nonprofit 501(c)(3) grantmaking organization. Global Action played no role in designing, implementing, data analysis, or interpretation of the research results, nor did Global Action edit or approve any presentations or publications from the study. The contents, selection, and presentation of facts, as well as any opinions expressed, are the sole responsibility of the authors and should not be regarded as reflecting the positions of Global Action. Through September 2023, Global Action received charitable gifts from PMI Global Services Inc. (PMI), which manufactures cigarettes and other tobacco products. To complement the termination of its agreement with PMI, Global Action's Board of Directors established a new policy to not accept or seek any tobacco or non-medicinal nicotine industry funding.

I. Introduction

Many traditional regulations are designed to correct market failures and improve economic efficiency. In the prominent case of environmental regulations, economists combine scientific research on the extent and impacts of pollution and pollution abatement with well-established economic valuation methods. The resulting estimates of the dollar value of the externalities reduced are key inputs into the cost-benefit analysis of major environmental regulations.

Over the past decade or so, internalities – where consumers' biased decision-making means that their choices fail to maximize their own welfare – are playing an increasingly important role in economic regulatory analysis. Examples include calory-posting requirements for chain restaurant menus, consumer finance regulations, energy and fuel efficiency standards, taxes on sugar-sweetened beverages, and tobacco product regulations. However, economic evaluation methods for internalities are less well-established and, as a result, there are fewer reliable estimates of the dollar value of internalities. As Weimer (2017) observes, the "gap between behavioral welfare economics and the craft of cost-benefit analysis remains large."

In this paper we contribute a new method to measure internalities due to present-biased and time-inconsistent preferences and apply the method to tobacco regulations. Present-biased and time-inconsistent preferences mean that internalities potentially arise in a wide range of consumer decisions about addictive goods or whenever consumers face tradeoffs between incurring a cost today for a benefit in the future. Our method is thus applicable to regulations that affect a sizeable subset of the goods and services sector. Tobacco regulations by themselves affect a sizeable sub-group of the U.S. population; for example, the graphic warning labels requirement that we analyze below will affect a product used regularly by almost 29 million smokers.

To implement our approach to measure internalities, we conducted a discrete choice experiment (DCE) to collect stated preference data about smokers' preferences for tobacco products versus quitting. We extended DCE methodology and asked subjects both about the immediate choices they would make today, and the choices they would make six months in the future. The difference between smokers' immediate choices and their stated intentions to quit in the future is commonly cited as evidence of smokers' time-inconsistency (Gruber and Koszegi 2001). We follow that logic and estimate mixed logit models of the subjects' immediate and future preference functions in willingness-to-pay space to provide estimates of the dollar value of the internalities from smoking.

Our DCE also embedded an experiment in which half of the subjects were exposed to one of the new required graphic warning labels while the control group saw one of the currently required text warning label. In work-in-progress, we combine the experimental results and our estimate of the internalities from smoking to complete an ex ante cost-benefit analysis of the graphic warning labels requirement. Our analysis of the mechanisms behind the estimated treatment effect concludes that graphic warning labels reduce smoking mainly by evoking fear/disgust rather than by providing an informational nudge. Our cost-benefit analysis contributes to a

controversial line of studies of the graphic warnings label requirement (FDA 2011, 2014, 2018; Ashley, Nardinelli, and Lavaty 2015; Chaloupka, Gruber, and Warner 2015; Levy, Norton, and Smith 2018).

II. Measuring Internalities by Comparing Immediate and Future Choices

Internalities arise when consumers fail to optimize. A useful framework assumes that the consumer's choices maximize their decision utility, which differs from the true experienced utility that will result from the choices. As a result, choices reflect the consumer's biased preferences, but their experience reflects their true unbiased preferences. The framework can be used to study specific behavioral biases or can follow a reduced-form approach to behavioral welfare economics and not specify the precise bias (Mullainathan, Schwartzstein and Congdon, 2012). Our approach measures internalities due to present-biased and time-inconsistent preferences, but we do not use a specific model such as semi-hyperbolic β/δ discounting.

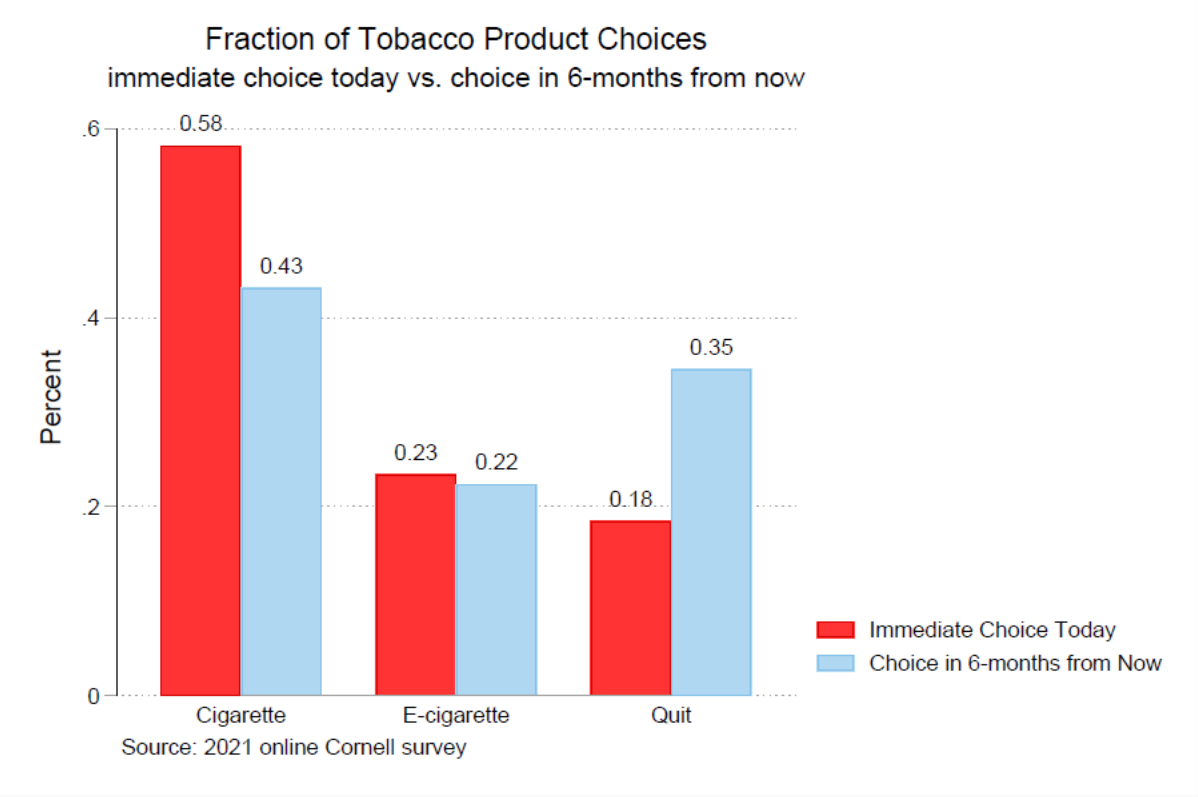
As is discussed in more detail below, we conducted a discrete choice experiment to collect subjects' stated preferences between cigarettes, e-cigarettes, and quitting. Each subject was given 12 choice tasks, where product prices and other attributes were experimentally varied across the tasks. In each task, subjects were asked about the immediate choice they would make today. The choice task was then repeated with the same prices and attributes, and subjects were asked about the choice they would make in six months from now.

Standard discrete choice experiments provide information about the consumer's preference function that predicts their immediate real-world choices.[1] In their seminal study of smoking internalities, Gruber and Koszegi (2001) cite differences between immediate choices and stated preferences to quit in the future as evidence that smokers make time-inconsistent choices due to present bias.[2] Following their logic, we assume that in our choice experiment, each subject's immediate choices today maximize their decision utility U_{now}^{SP} (biased preference), while their six-months-from-now-choices maximize their experienced utility U_{future}^{SP} (unbiased preference). Figure 1 shows the raw data on our discrete choice experiment subjects' immediate choices and six-month-from-now choices. Consistent with the evidence cited by Gruber and Koszegi (2001), 35 percent of the six-months-from-now choices are for quitting, almost twice the 18 percent of immediate choices for quitting. Compared to the 58 percent of immediate choices that are for

[1] McFadden (2017) concludes that: "Forecasts that are comparable in accuracy to RP [revealed preference] forecasts can be obtained from well-designed SP [stated preference] studies for familiar, relatively simple goods that are similar to market goods purchased by consumers…." 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.

[2] According to Gruber and Koszegi (2001): "A second feature that distinguishes time-consistent agents from time-inconsistent agents is an inability to actualize predicted or desired future levels of smoking….In fact, unrealized intentions to quit at some future date are a common feature of stated smoker preferences."

cigarettes, the fraction of six-months-from-now choices for cigarettes correspondingly drops by 15 percentage points to 43 percent.



We characterize subjects' internality I by comparing the utility derived from choice in 6 months (optimal choice) and utility derived from immediate choice today (suboptimal choice).

$$I = V^{SP}_{future}(y^{SP}_{future}, \hat{\theta}) - V^{SP}_{now}(y^{SP}_{now}, \hat{\theta})$$

where y indicates the choices made by the subject and $\hat{\theta}$ is the model estimated parameters, which represent the agent's preferences. Defined this way, the internality I measures the subject's welfare loss due to sub-optimal choices.

Because the choice tasks describe product attributes that do not change for the same individual making immediate choices and future choices, we can only learn about internality when an individual makes different choices for immediate choice and future choice (that is, $V^{SP}_{future}(y, \hat{\theta}) = V^{SP}_{now}(y, \hat{\theta}) = V$ for the same y). For each individual i in scenario t, her internality can be calculated according to the formula outlined in the following table 1; note that we normalize $V_{it}(quit)$ to zero.

In Table 1, the majority of choices are the same for immediate and future choices. In 41 percent choice patterns, both immediate and future choices are for cigarettes. In 17 percent of choice patterns both immediate and future choices are for e-cigarettes, and in another 17 percent of choice patterns both immediate and future choices are for quitting. Most of the remaining

patterns of choices are consistent with present-biased and time-inconsistent preferences, where future choices are for the healthier options of quitting or e-cigarettes. In 12 percent of choice patterns, the immediate choice is for cigarettes and the future choice is for quitting. In 5 percent of choice patterns, the immediate choice is for e-cigarettes and the future choice is for quitting. In 5 percent of choice patterns, the immediate choice is for cigarettes and the future choice is e-cigarettes. These choice patterns are the key sources for internality calculations. In small fractions of choice patterns, the future choices are less healthy than the immediate choices.

Table 1. Choice patterns and the internality calculations

Choice in 6 months (1)	Immediate choice today (2)	Internality (1) - (2)	Number of Choices	Percent %
Cigarette	Cigarette	0	5820	41.4%
Cigarette	E-cigarette	V(cig) - V(e-cig)	166	1.2%
Cigarette	Quit	V(cig)	89	0.6%
E-cigarette	Cigarette	V(e-cig) - V(cig)	661	4.7%
E-cigarette	E-cigarette	0	2398	17.1%
E-cigarette	Quit	V(e-cig)	86	0.6%
Quit	Cigarette	- V(cig)	1708	12.2%
Quit	E-cigarette	- V(e-cig)	726	5.2%
Quit	Quit	0	2416	17.2%

Notes: Total number of choices is number of respondents multiply by choice scenarios (1171 x 12 = 14052).

III. Data and Empirical Model

We conducted an online discrete choice experiment where subjects choose between three alternatives: cigarettes, e-cigarettes, or quitting both products. The design of the experiment consists of a 3 (cigarette price) X 3 (e-cigarette price) X 3 (nicotine levels) X 3 (flavor availability) X 4 (e-cigarette warning) design, for a total of 324 possible attribute combinations. Out of 324 possible scenarios, the DCE presented 120 scenarios to a total of 1202 respondents, which are all adult smokers (aged 18 and above).[3] Each respondent was shown with 12 scenarios, and different respondents were assigned different sets of scenarios. To study the impact of graphic warnings in a semi-realistic market context, we embedded a cigarette graphical warning label experiment into the discrete choice experiment. We split the sample into two groups, the first serves as the control group where subjects were presented with one of the currently required rotating text warning labels: "Smoking causes Lung Cancer, Heart Disease, Emphysema, and may Complicate Pregnancy." The second serves as the treatment group where subjects were presented with one of the newly proposed FDA graphic warning labels, it is an image of amputated toes with the accompanying message "Smoking Reduces Blood Flow to the Limbs, Which May Require Amputation". Note that each respondent sees the same graphic or

[3] In the analysis, we dropped respondents who have abnormal paid prices for a pack of cigarettes, this led to a sample of 1171 respondents.

text warning label in all 12 choice scenarios, this prevents the estimated impacts of the graphic warning labels from being contaminated by respondents' exposure to another label in a prior choice scenario. Figure 1 shows a choice task for a respondent in the treatment group that saw the graphic warning label on cigarette choice.

In each choice task, each respondent was asked about the most preferred choice for today and the most preferred choice for 6 months from now. This process is repeated 12 times (each time with a different combination of product attributes). For each respondent, we collected 12 immediate choices today and 12 choices in 6 months from now.

We used the survey firm SSRS to implement the surveys and experiments, the surveys include three sections: a background survey, the DCE choice tasks, and follow-up questions.

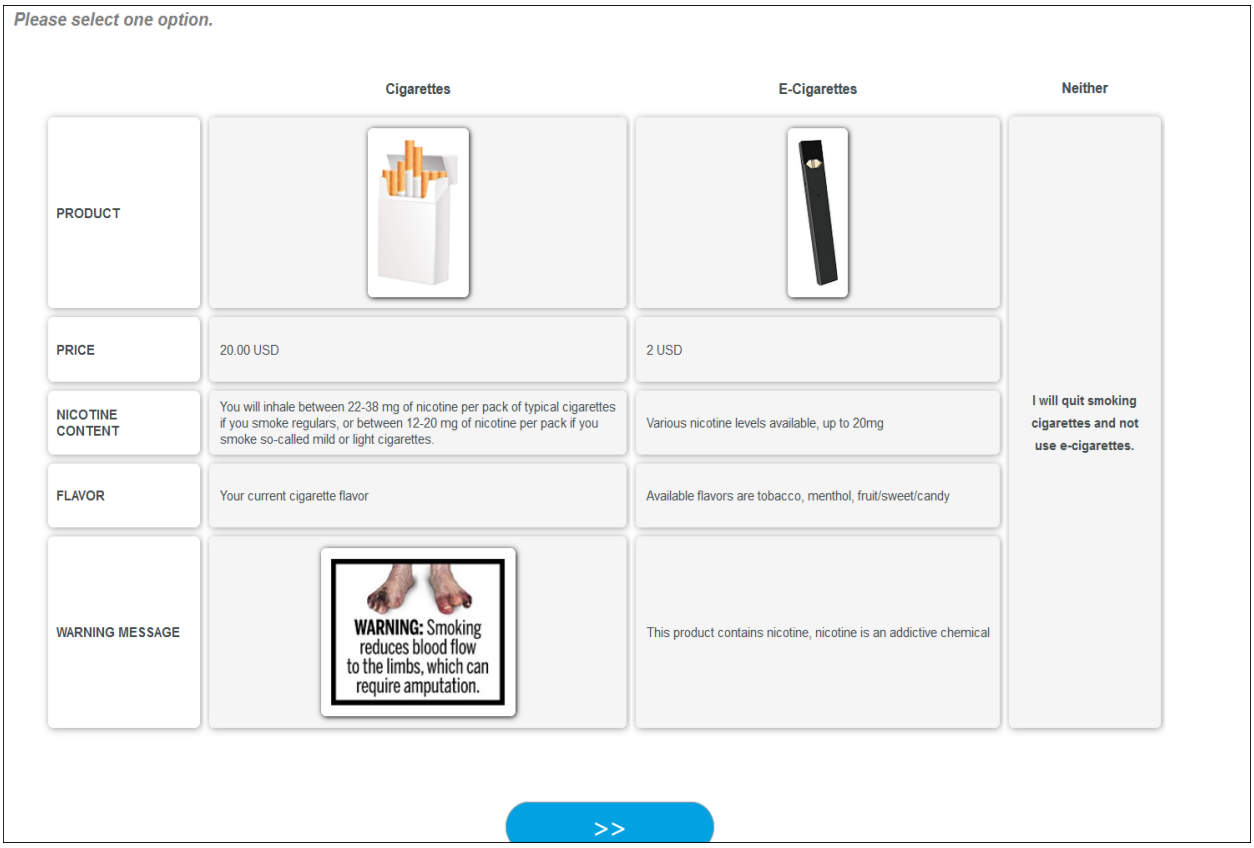


Figure 1. A Choice task Presented to Respondents in the Graphic Warning Label Group

We estimate a random utility based mixed logit model as follows:

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}$$

Where individual i chooses alternative j in choice scenario t that maximizes her utility U_{ijt}. The indirect utility U_{ijt} can be decomposed into two parts, V_{ijt} is the deterministic part of the utility and ε_{ijt} is the stochastic error term that follows an i.i.d. Type-1 extreme value distribution. Specifically

$$V_{ijt} = ASC_i^{cig} + ASC_i^{ecig} + \beta_i GWL_{ij} + X'_{ijt}\gamma_i + Z_i'\delta$$

Where ASC indicates Alternative-Specific-Constant, ASC_i^{cig} and ASC_i^{ecig} measure consumer valuation of cigarettes and e-cigarettes relative to quitting, respectively. GWL represents graphic warning label on cigarette packs, and β_i is the random coefficient that measures the marginal utility derived from facing GWL. X is a vector of product attributes including prices, flavor availability of e-cigarettes, nicotine availability of e-cigarettes, and warning message of e-cigarettes. Z is a vector of individual characteristics.[4] Because we are interested in the monetary measure of utility loss or gain, we estimate the model in willingness-to-pay (WTP) space. We estimate the model with Hierarchical Bayes, which avoids numerical convergence issues, and has more desirable properties compared to maximum simulated likelihood when the sample size is small (Train, 2009).

IV. Descriptives and Estimation Results

We first report the descriptive statistics of product choices in table 2. In the sample, for immediate choice today, the most often chosen option is cigarette, with a choice probability of 0.581, followed by 0.234 for e-cigarette and 0.184 for quitting both. Comparing the choices between GWL group and text warning group, we found that respondents are 6.7 percentage points less likely to choose cigarette, 2.81 percentage points more likely to choose e-cigarette and 3.96 percentage points more likely to choose quitting when exposed to the GWL. For choice in 6 months from now, we found similar patterns. Comparing to today's, we found that respondents choose cigarette less often while quitting more often for future choices. This implies that there are more stated preferences to quit in the future.

[4] Individual characteristics include smoking status (being every day smoker), age of smoking initiation, heavy smokers (heavy smoker index >= 3), smoke menthol flavored cigarettes, vaping status (ever used e-cigarette), quitting effort (tried to quit during last 12 months), quit intention (intend to quit in the next 6 months), gender (being male), age, household size, education (reference category is grade school/some high school), household income above 50k, full-time employed, white. We include these control variables to reduce bias from sample imbalances.

Table 2. Descriptive statistics of product choices in the DCE

	All	Mean graphic warning	text warning	Mean difference
Immediate choice today				
Cigarette	0.581	0.548	0.615	-0.0677***
E-cigarette	0.234	0.248	0.220	0.0281***
Quit	0.184	0.204	0.165	0.0396***
Choice of 6 months from now				
Cigarette	0.431	0.385	0.477	-0.0917***
E-cigarette	0.224	0.235	0.213	0.0221**
Quit	0.345	0.380	0.310	0.0697***
Respondents	1171	587	584	

Table 3 presents the estimation results of the mixed logit model in WTP space. We estimate models for today's choice and future choice separately and report the estimated mean and standard deviations of the distribution of random coefficients. The results show that comparing to quitting, respondents have strong preferences for cigarettes ($21.8 mean WTP for choice today and $15.5 mean WTP for choice in 6 months) while the preference for e-cigarette is weaker than that for cigarette ($7.9 mean WTP for choice today and $3.2 mean WTP for choice in 6 months). The mean WTP for the GWL is $-2.96 for choice today and $-12.2 for choice in 6 months. All the standard deviation estimates are statistically significant, which implies substantial heterogeneity in preferences for those product attributes.

Table 3. Estimation results of the mixed logit model

		Immediate Choice Today		Choice in 6 Months	
		Mean	SD	Mean	SD
Price ($)		-2.069***	5.390**	-1.902***	4.827**
		(0.333)	(2.661)	(0.330)	(2.159)
Alternative-Specific-Constant	Cigarettes	21.846***	15.354***	15.539***	17.795***
		(0.583)	(0.741)	(0.422)	(0.884)
Alternative-Specific-Constant	E-cigarettes	7.912***	16.501***	3.218***	18.865***
		(0.617)	(0.772)	(0.531)	(1.145)
Cigarette warning	Graphic warning of amputation	-2.962***	2.848***	-12.154***	13.365***
		(0.168)	(0.294)	(0.336)	(0.637)
E-cigarette available flavor	Tobacco, menthol, fruit/sweet/candy	-1.114**	4.314***	-0.776***	4.332***
		(0.520)	(0.606)	(0.244)	(0.247)
	Tobacco and menthol	-0.454**	2.810***	-1.018	3.275***
		(0.231)	(0.221)	(0.194)	(0.242)
E-cigarette available nicotine level	Up to 20mg	-0.786***	1.388***	-1.810***	2.463***
		(0.250)	(0.184)	(0.177)	(0.242)
	Up to 50mg	-0.920***	1.582***	-1.039***	1.777***
		(0.284)	(0.119)	(0.254)	(0.162)
E-cigarette warning	Are not complete risk free	-1.574***	2.011***	-2.013***	2.931***
		(0.280)	(0.350)	(0.175)	(0.202)
	Contain nicotine, which is addictive	-0.487***	0.847***	-1.040***	1.806***
		(0.056)	(0.077)	(0.140)	(0.164)
	May expose users to chemicals and toxins	-1.861***	3.214***	-2.522***	2.773***
		(0.364)	(0.599)	(0.164)	(0.197)
Log-likelihood at convergence		-6612		-6459	
Observations		14052		14052	

Notes: Alternative specific constants, coefficients associated with e-cigarette available flavor, and e-cigarette available nicotine level are assumed to follow normal distributions, the negative of coefficients associated with price and e-cigarette warnings are assumed to follow lognormal distributions, the negative of coefficient associated the graphic warning label is assumed to follow a truncated normal distribution. Additional control variables include smoking status (being every day smoker), age of smoking initiation, heavy smokers (heavy smoker index >= 3), smoke menthol flavored cigarettes, vaping status (ever used e-cigarette), quitting effort (tried to quit during last 12 months), quit intention (intend to quit in the next 6 months), gender (being male), age, household size, education (reference category is grade school/some high school), household income above 50k, full-time employed, white. The reference category of e-cigarette available flavor "tobacco only", the reference category of e-cigarette available nicotine level is "up to 5mg", the reference category of e-cigarette warning is "no warning". Standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We further use the model estimates to predict choice shares under status quo market conditions. The status quo market conditions are described as follows: the cigarette price is what the respondent paid for last pack, the e-cigarette price is set to be $8, cigarette warning is set to text warning only, e-cigarette flavor availability is set to tobacco, menthol, fruit/sweet/candy, e-cigarette available nicotine level is set to up to 50mg, e-cigarette warning message is set to the FDA warning message "Contain nicotine, which is addictive".

Figure 2 shows the plot of model predicted choice shares under status quo market conditions, the red bar indicates the predicted choice shares from the model of choice today and the blue bar shows the ones from the model of choice in 6 months. We see that model of choice today predicts much more cigarette choices than model of choice in 6 months (71% versus 55%) and predicts much less quitting choices than model of choice in 6 months (14% versus 30%).

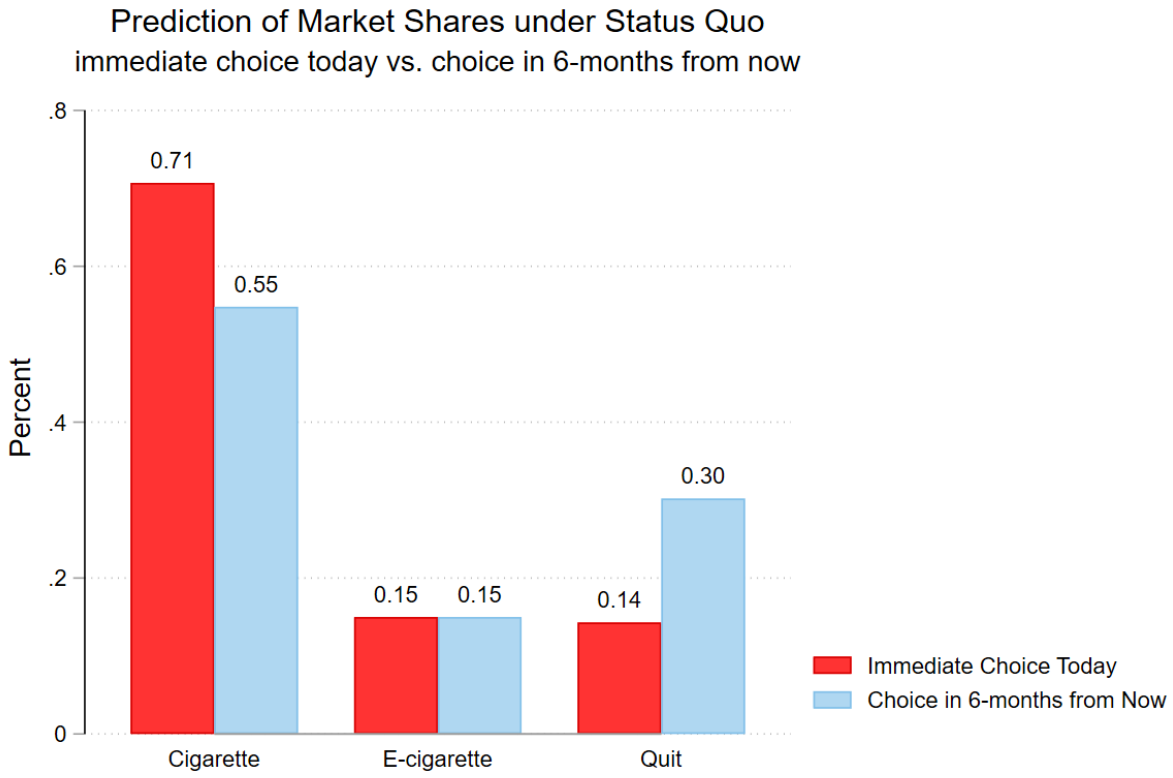


Figure 2. Predicted choice shares under status quo market conditions

V. Alternative Measurements of Internality

Our first estimate of the internality due to smoking implement the formula described above in Table 3. Specifically, for individual i in scenario t the utility associated with each choice is

$$V_{it}(cig) = \widehat{\alpha_i^1} + \widehat{\beta_i} GWL_{ij} + \widehat{p_i} Price_{it}(cig)$$
$$V_{it}(ecig) = \widehat{\alpha_i^2} + \widehat{p_i} Price_{it}(ecig) + \widehat{\gamma_i^1} Flavor_{it} + \widehat{\gamma_i^2} Nicotine + \widehat{\gamma_i^3} Warning$$
$$V_{it}(quit) = 0$$

Note that individual characteristics do not vary within individual, thus the utility associated with individual characteristics can be subtracted from V for all choices. And we normalize $V_{it}(quit)$ to zero.

The calculation proceeds as follows: we first estimate a mixed logit model in WTP space with the choices in 6 months as the outcome variable to obtain the true preference function estimates, then we use the true preference function estimates to evaluate the utility associated with the choices, i.e., choice today and choice in 6 months separately. For example, if the person chooses cigarette for today but chooses quit in 6 months, then the internality is calculated as I_{it} = V(cig) - V(quit) = V(cig) (because V(quit) is normalized to zero). In this way, we obtain internality measures for each individual for each choice task, and this gives us a distribution of the internality measures over the sample. Calculated this way, the internality due to smoking has a mean of \$4.15/pack with a standard deviation of \$10.15. By comparison, previous estimates of the internality due to smoking range from \$3.61/pack estimated by Jin et al. (2015) to \$9.55/pack used by Allcott and Rafkin (2022).[5]

Because the data we use is from a DCE, which measures stated preference (SP), there is always a concern of hypothetical bias associated with SP data. Here we explore an alternative measurement to check the robustness of our internality calculations.

Suppose SP choices of 6-months from now maximize consumers' true experienced utility U_{future}^{SP}, and consumers' observed RP choices (which can be constructed from the background survey by a probabilistic rule) maximize consumers' decision utility U^{RP}, then we can calculate the internalities as the difference in utility derived from the two choices.

$$I = V_{future}^{SP}(y_{future}^{SP}, \hat{\theta}) - V^{RP}(y^{RP}, \hat{\theta})$$

Note that each respondent made 12 SP choices for 6 months and one RP choice, also the market conditions are different between SP choices and RP choices. To make an apple-to-apple comparison, we first use the model of 6 months choices to predict choices under the status quo

[5] An unpublished working paper by Chaloupka, Levy, and White (2019) yields an outlier internality estimate of \$57/pack. Their estimated structural model also yields an estimate that the short-term costs of smoking withdrawal are worth \$8100 per week. The estimated internality and withdrawal costs are large because the structural model yields a very small marginal utility of income.

market conditions, then compare the predicted choices with the RP choices, in this way, we are comparing one predicted SP choice and one RP choice under the same market conditions. For each individual i in scenario t, her internality can be calculated according to the formula outlined in the following table 4. We see that the choice patterns are very similar to those in table 3. With this method, the calculate internality has a mean of \$4.71 with a standard deviation of 7.87, which is very close to the results under previous measurement. This indicates the robustness of our methods to calculate internality of smoking.

Table 4. Choice patterns and the internality calculations under an alternative measurement

Choice in 6 months (1)	Revealed preference choice (2)	Internality (1) - (2)	Number of choices	Percent %
Cigarette	Cigarette	0	591	50.5%
Cigarette	E-cigarette	V(cig) - V(e-cig)[RP]	42	3.6%
Cigarette	Quit	V(cig)	11	0.9%
E-cigarette	Cigarette	V(e-cig) - V(cig)[RP]	106	9.1%
E-cigarette	E-cigarette	0	66	5.6%
E-cigarette	Quit	V(e-cig)	1	0.1%
Quit	Cigarette	- V(cig)[RP]	309	26.4%
Quit	E-cigarette	- V(e-cig)[RP]	35	3.0%
Quit	Quit	0	10	0.9%

Notes: Total number of choices is the number of respondents (1171).

VI. Cost-Benefit Analysis of GWL

With the calculated internalities, we can proceed to the cost-benefit analysis of implementing GWL on cigarette packs. First, we need to know whether the GWL has impacts on smokers' choices between cigarettes, e-cigarettes and quitting, from table 1, we can see clearly that the GWL reduces the probability of choosing cigarettes while increases the probability of choosing e-cigarettes and quitting. To formally explore the causal effects, we estimate linear probability models as follows:

$$y_{it} = \beta_0 + \beta_1 GWL_i + X'_{it}\gamma_i + Z'_i\delta + \epsilon_{it}$$

Where y_{it} is the choice of individual i made in scenario t, GWL refers the graphic warning label on cigarette packs, X is a vector of product attributes including prices, flavor availability of e-cigarettes, nicotine availability of e-cigarettes, and warning message of e-cigarettes. Z is a vector of individual characteristics. ϵ is the stochastic error term.

We estimate six versions of the above equation. In the first three versions, the dependent variables are binary indicators of whether the subject's immediate choice was to purchase a cigarette, to purchase an e-cigarette, or to quit. In the next three versions, the dependent variables measure the subject's stated choice for 6 months from now. Each equation is estimated separately, and we do not impose any cross-equation restrictions. However, because the three choices are mutually exclusive and exhaustive, empirically the parameter estimates in the first

three immediate choice equations and the next three six-months-from-now equations sum to zero. For example, if the graphic warning label is estimated to decrease the probability of choosing cigarettes, there will be an equal and opposite increase in the probability of choosing e-cigarettes and/or quitting. In the equation the parameter β_1 provide reduced-form estimates of the total impact of the graphic warning label on the probability of choosing cigarettes, e-cigarettes, or quitting.

Table 5 reports the estimation results of the linear probability models. For immediate choice today, the GWL reduced the probability of choosing cigarette by 5.4 percentage points, this is statistically significant at 1% level, and GWL increased the probability of choosing quitting by 2.8 percentage points. For choice in 6 months, the GWL reduced the probability of choosing cigarette by 7.8 percentage points and GWL increased the probability of choosing quitting by 5.5 percentage points. The results are consistent with the descriptive statistics of the raw data reported in table 1.

Table 5. Estimation results of linear probability models

Variables		Immediate choice today			Choice of 6 months from now		
		Cigarette	E-cigarette	Quit	Cigarette	E-cigarette	Quit
Cigarette warning	GWL	-0.054***	0.027	0.028*	-0.078***	0.024	0.055***
		(0.019)	(0.017)	(0.016)	(0.019)	(0.018)	(0.021)
Price	Cigarette price	-0.018***	0.010***	0.008***	-0.015***	0.007***	0.008***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	E-cigarette price	0.017***	-0.017***	-0.000	0.013***	-0.014***	0.002
		(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
E-cigarette available flavor	Tobacco, menthol, fruit/sweet/candy	-0.022***	0.027***	-0.005	-0.027***	0.036***	-0.009
		(0.008)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)
	Tobacco and menthol	-0.026***	0.010	0.016**	-0.023***	0.019***	0.003
		(0.008)	(0.007)	(0.006)	(0.007)	(0.006)	(0.006)
E-cigarette available nicotine level	Up to 20mg	-0.004	0.003	0.002	-0.004	-0.001	0.006
		(0.006)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)
	Up to 50mg	0.012*	-0.009	-0.003	-0.000	-0.001	0.001
		(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
E-cigarette warning	Are not completely risk free	0.005	-0.002	-0.003	0.002	0.003	-0.005
		(0.008)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
	Contain nicotine, which is addictive	0.032***	-0.017**	-0.015**	0.018*	-0.005	-0.013
		(0.009)	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
	May expose users to chemicals and toxins	0.012	-0.018**	0.005	0.005	-0.011	0.007
		(0.009)	(0.008)	(0.007)	(0.008)	(0.008)	(0.008)
	Constant	0.628***	0.214***	0.158**	0.731***	0.193***	0.075
		(0.074)	(0.061)	(0.068)	(0.078)	(0.066)	(0.087)
Observations		14,052	14,052	14,052	14,052	14,052	14,052
Adjusted R-squared		0.208	0.144	0.098	0.239	0.099	0.169

Notes: Additional control variables include smoking status (being every day smoker), age of smoking initiation, heavy smokers (heavy smoker index >= 3), smoke menthol flavored cigarettes, vaping status (ever used e-cigarette),

quitting effort (tried to quit during last 12 months), quit intention (intend to quit in the next 6 months), gender (being male), age, household size, education (reference category is grade school/some high school), household income above 50k, full-time employed, white. The reference category of e-cigarette available flavor "tobacco only", the reference category of e-cigarette available nicotine level is "up to 5mg", the reference category of e-cigarette warning is "no warning". Standard errors clustered at respondent level are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

It is also important to distinguish the mechanisms of how does the GWL work. Although we do not fully specify a structural model of consumer utility, in our approach we conceptualize the graphic warning label's impact on utility as being through the mechanisms of fear/disgust and knowledge, which we observe in our data, as well as unobserved mechanisms. To understand the mechanisms, we estimate models of the impact of the graphic warning label on the observed antecedents of choice. We examine how the graphic warning label influences emotions such as fear and disgust, and knowledge of the health risk of smoking featured in the label:

$$Fear/Disgust_i = \alpha_{f0} + \alpha_{f1} GWL_i + Z_i'\delta + \eta_i$$

$$Knowledge_i = \alpha_{k0} + \alpha_{k1} GWL_i + Z_i'\delta + \eta_i$$

We then re-estimate the choice equations given by the linear probability models, but we drop the GWL indicator and replace it with the Fear/Disgust and Knowledge variables:

$$y_{it} = \beta_0 + \beta_{f1a} Fear/Disgust_i + \beta_{k1b} Knowledge_i + X'_{it}\gamma_i + Z_i'\delta + \epsilon_{it}$$

The estimates α_{f1} and α_{k1} show the impacts of the GWL on the subject's fear/disgust and knowledge measures, respectively, the estimates β_{f1a} and β_{k1b} show the impact of fear/disgust and knowledge on the choice probabilities. Combining these estimates provides estimates of the impacts of the GWL on choices through the mechanisms of fear/disgust and knowledge $\delta_1 = \beta_{f1a} \times \alpha_{f1}$ and knowledge $\delta_2 = \beta_{k1a} \times \alpha_{k1}$. The shares of the total impact of the GWL on choices accounted for by each of the mechanisms are given by δ_1/β_1 and δ_2/β_1.

In addition to fear/disgust and knowledge we also examine other possible mechanisms, including how the GWL influence perceptions of the relative risk of e-cigarettes and cigarettes, how the GWL influences expected years of life lost from smoking and vaping, and how the GWL influences perceptions of the harm of nicotine.

Table 6 shows the shares of the total impact of the GWL on choices accounted for by the mechanism. The impact of the GWL on fear/disgust accounts for approximately 17% of the total impact of the labels on cigarette choices. The impact of the GWL on knowledge of the specific health risk mentioned in the label accounts for approximately 3% of the total impact on cigarette choices. In general, the effects of the other changes in knowledge are also small. With respect to the e-cigarettes, the results suggest that the cigarette GWL's effects on both fear/disgust and

perceptions of the relative risks of cigarettes versus e-cigarettes both play an important role to increase e-cigarette choices.

Our results suggest that fear/disgust is the main mechanism through which one of the new graphic warning labels influences stated preferences for cigarettes. We estimate that improvements in knowledge about the specific risk featured in the label – the risks that smoking plays in reducing blood flow to the limbs – play a minimal role. From these results, we conclude that graphic warning labels are mainly not a nudge intervention that reduces the smoking internalities but instead are a direct source of utility loss due to emotional effects such as fear/disgust.

Table 6. Mechanism explained impacts.

	Percent of Total Impact of The Pictorial Warning on Choice δ/β					
	Immediate Choice Today			Choice of 6 months from now		
	Cigarette	E-cigarette	Quit	Cigarette	E-cigarette	Quit
Feel Disgust/Fear as a Result of seeing Cigarette Warning Label	16.62%	21.80%	12.96%	12.21%	17.46%	10.55%
Perceived Harm of E-cigarettes Relative to Cigarettes	8.97%	31.34%	-6.82%	3.99%	35.34%	-5.88%
Believe that "I smoke more than I should"	-2.87%	0.57%	-5.29%	0.29%	-1.79%	0.95%
Believe that Nicotine in Cigarettes is the Substance that Causes Most of the Cancer Caused by Smoking	4.70%	3.03%	5.88%	0.19%	7.31%	-2.06%
Nicotine, on Its Own, Can Harm the Developing Brain of Youth and Young Adults	-0.05%	0.77%	-0.63%	-0.60%	-1.72%	-0.25%
Believe that Smoking Cigarettes Reduces Blood Flow to the Lower Limbs	3.19%	1.35%	4.50%	2.51%	-1.61%	3.81%
Life Expectancy Loss from Smoking	0.11%	-2.08%	1.66%	-0.16%	-7.23%	2.06%
Life Expectancy Loss from Vaping	0.13%	6.78%	-4.56%	-0.19%	6.55%	-2.31%

Based on the mixed logit model of choices in six months, we estimate that the mean willingness to pay for the graphic warning label (GWL) is -$12.20 (Table 3, above). Given the evidence that GWLs decrease smoking by imposing a utility loss, cost-benefit analysis of the graphic warning label requirement parallels the analysis of a cigarette tax of $12.20/pack.[6] Our cost-benefit analysis remains work-in-progress, but we offer two preliminary perspectives.

[6] An important difference between the social cost of GWLs versus its tax equivalent is that GWLs do not generate tax revenues.

The first perspective is to compare the estimated GWL tax-equivalent of $12.20/pack to the estimated average internality due to smoking of $4.15/pack. Allcott and Rafkin (2022) develop formulas for the optimal taxes on cigarettes and e-cigarettes. Under simplifying assumptions, the optimal cigarette tax should be set equal to the sum of the uncorrected internalities and externalities due to smoking. DeCicca, Kenkel, and Mathios (2022) review evidence that on average current cigarette taxes are at least high enough to cover the externalities due to smoking, i.e. the uncorrected externalities are zero or negative. The simple comparison that the estimated GWL tax-equivalent of $12.20/pack is almost three times larger than the estimated internality of $4.15/pack suggests that the graphic warning labels requirement does not pass the cost-benefit test.

The second perspective is to consider the impacts of the GWL tax-equivalent on different groups of smokers. For many smokers, the GWL will be inframarginal; they will continue to smoke despite the GWLs. GWLs impose a utility loss on continuing smokers with no corresponding benefits. Some smokers will respond to GWLs by making the healthier choices to quit smoking entirely or switch to vaping e-cigarettes. To the extent these smokers' original choices imposed internalities on themselves, the GWLs improve their welfare. However, because for some smokers the GWL tax-equivalent will be too large compared to the internalities they impose on themselves, the GWL requirement will create net welfare losses even though their health improves. When our work-in-progress cost-benefit analysis is completed, we will be able to more fully characterize and quantify the GWL impacts on smokers' welfare.

References

Allcott, H., & Rafkin, C. (2022). Optimal regulation of e-cigarettes: Theory and evidence. *American Economic Journal: Economic Policy*, *14*(4), 1-50.

Ashley, Elizabeth M., Clark Nardinelli, and Rosemarie A. Lavaty. 2015. "Estimating the Benefits of Public Health Policies that Reduce Harmful Consumption." *Health Economics* 24 (5): 617–24.

Chaloupka, Frank J., IV, Matthew R. Levy, and Justin S. White. 2019. "Estimating Biases in Smoking Cessation: Evidence from a Field Experiment." NBER Working Paper 26522.

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

Gruber, Jonathan, and Botond Ko˝szegi. 2001. "Is Addiction 'Rational'? Theory and Evidence." *Quarterly Journal of Economics* 116 (4): 1261–1303.

Jin, Lawrence, Don Kenkel, Feng Liu, and Hua Wang. 2015. "Retrospective and Prospective Benefit-Cost Analyses of U.S. Anti-Smoking Policies." *Journal of Benefit-Cost Analysis* 6 (1): 154–86.

McFadden, D. (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.

Levy, Helen G., Edward C. Norton, and Jeffrey A. Smith. 2018. "Tobacco Regulation and Cost-Benefit Analysis: How Should We Value Foregone Consumer Surplus?" *American Journal of Health Economics* 4 (1): 1–25.

Mullainathan, S., Schwartzstein, J., & Congdon, W. J. (2012). A reduced-form approach to behavioral public finance. *Annual Review of Economics* 4, 511–540. https://doi.org/10.1146/annurev-economics-111809-125033

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

David Weimer (2017). *Behavioral Economics for Cost-Benefit Analysis: Benefit Validity When Sovereign Consumers Seem to Make Mistakes*. Cambridge University Press.