Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity Web Appendix

Jacob Goldin   Tatiana Homonoff

June 7, 2012

Web Appendix A: Welfare Analysis Under Alternate Budget Adjustment Rules

Part I assumed that inattentive consumers who misperceive the price of x satisfy their budget constraints by reducing expenditures on y. This Appendix considers the robustness of our results to alternate rules for mapping infeasible intended consumption bundles into feasible final consumption bundles.

In addition to the rule that we employ, Chetty, Looney, and Kroft (2007) identify two other "intuitive" budget adjustment rules. First, consumers who misperceive the price of x may satisfy their budget constraints by reducing expenditures on x rather than y. This rule represents the other end of the spectrum from the one that we employ, and would be appropriate if consumers purchased x after completing their purchases of all other goods. Under this rule, it is easy to show that:

$$\frac{\partial x_B}{\partial t_r} = \frac{-x_B}{p+t_r+t_p} \tag{1}$$

$$\frac{\partial x_B}{\partial t_p} = \frac{-\left(\frac{\partial y_B}{\partial p}+x_B\right)}{p+t_r+t_p} \tag{2}$$

The second alternate budget adjustment considered by Chetty, Looney, and Kroft (2007) is for inattentive agents to reduce consumption of both x and y to make up the income lost to the register tax. Inattentive consumers ignore the register tax when making their consumption decisions, but recognize that their net-of-tax income is lower because of the tax. For example, consumers who purchase x and y repeatedly will eventually realize that they consistently have less money in their bank account than they had anticipated. Inattentive consumers whose behavior is described by this rule will fully account for the tax's income effect but fail to account for the tax's substitution effect.

As a result, we have:

$$\frac{\partial x_B}{\partial t_r} = -x_B \frac{\partial x_B}{\partial I} \tag{3}$$

$$\frac{\partial x_B}{\partial t_p} = \frac{\partial x_B}{\partial t_r} + \frac{\partial \tilde{x}_B}{\partial p} \tag{4}$$

where $\frac{\partial \tilde{x_B}}{\partial p}$ represents Hicksian (compensated) demand.

As before, we consider the welfare effects of a revenue-neutral shift from posted to register taxes. Because the attentive agent optimizes correctly, the welfare effect for that agent is the same as before:

$$\left.\frac{dV_A}{dt_r}\right|_{\overline{R}} = -U_y\left(x_A, y_A\right) x_A \left(1 + \left.\frac{\partial t_p}{\partial t_r}\right|_{\overline{R}}\right)$$

Totally differentiating the government's budget constraint yields an expression for the posted tax reduction associated with a revenue-neutral increase in the register tax:

$$\left.\frac{\partial t_p}{\partial t_r}\right|_{\overline{R}} = -\frac{x_A + x_B + (t_p + t_r)\left(\frac{\partial x_A}{\partial t_r} + \frac{\partial x_B}{\partial t_r}\right)}{x_A + x_B + (t_p + t_r)\left(\frac{\partial x_A}{\partial t_p} + \frac{\partial x_B}{\partial t_p}\right)}$$

A little algebra reveals that the welfare effect of the shift is positive for attentive consumers if and only if $\frac{\partial x_B}{\partial t_r} > \frac{\partial x_B}{\partial t_p}$, that is, when inattentive consumers reduce their demand for the taxed good by a larger amount in response to a posted tax increase than in response to a register tax increase. Intuitively, this condition ensures that the new register tax will be more effective at raising revenue than the old posted tax was. Consequently, the shift accommodates a reduction in the combined tax rate, thus generating a positive income effect. Using (1) - (4), it is easy to see that this condition is satisfied under the two alternate budget adjustment rules.[1]

The welfare analysis for inattentive consumers proceeds as in Part I. Under the first alternate rule,

$$\left.\frac{dV_B}{dt_r}\right|_{\overline{R}} = -\left(1 + \left.\frac{\partial t_p}{\partial t_r}\right|_{\overline{R}}\right) U_x(x_B, y_B)\left(\frac{x_{\mathbf{B}}}{p+t_r+t_p}\right) + \frac{\left(\frac{\partial y_B}{\partial p}\right)\left.\frac{\partial t_p}{\partial t_r}\right|_{\overline{R}}}{(p+t_r+t_p)}\left(U_y(x_A, y_A)(p+t_r+t_p) - U_x(x_A, y_A)\right)$$

Like the result in Part I, the welfare effect for inattentive consumers is ambiguous under this rule. Shifting to a register tax accommodates a reduction in the combined tax rate, generating a positive welfare effect (captured by the first term). Unlike before, however, the magnitude of this effect depends on the marginal utility of x rather than y because providing the consumer with additional income reduces the amount that the consumer must reduce her consumption of x to satisfy the budget constraint. The second term represents the cost of optimization error. Like before, this cost is zero when there are no register taxes and grows in size as register taxes push inattentive consumers further from their optimal bundle.

[1] Because y represents all goods other than x, it is reasonable to assume that $\frac{\partial y_B}{\partial p} > 0$.

Under the second alternate rule, the welfare effect of the shift for inattentive consumers is also similar to that found in Part I. Here the welfare effect is given by

$$\frac{dV_B}{dt_r}\bigg|_{\overline{R}} = -\left(1+\frac{\partial t_p}{\partial t_r}\bigg|_{\overline{R}}\right)U_y(x_B,y_B)x_B + \left(\frac{\partial x_B}{\partial p}\,\frac{\partial t_p}{\partial t_r}\bigg|_{\overline{R}} - x_B\frac{\partial x_B}{\partial I}\right)(U_x(x_B,y_B)-(p+t_r+t_p)U_y(x_B,y_B))$$

Again, the first term represents a positive income effect and the second term represents a negative welfare effect stemming from optimization error, which grows in size as register taxes increase.

Web Appendix B: A Cognitive Cost Model of Heterogeneous Attentiveness

How does attentiveness to register taxes vary by income? The model we develop in this Appendix does not make a uniform prediction for all goods, but rather highlights the factors that determine which income group will be more attentive for a particular good. We then consider those factors in the context of cigarettes to predict whether high- or low-income consumers are likely to be more attentive to cigarette register taxes.

Suppose all agents have the option of paying attention to register taxes, but that doing so carries with it some positive utility cost.[2] This "cognitive cost" could stem from the mental effort needed to remember and calculate a good's tax-inclusive price or might simply represent the opportunity cost of time spent on that task.

Assume that agents' final utility is additively separable between the cognitive cost and consumption so that we can write $W_i = U(x_i, y_i) - b_i c_i$ in which b_i is a binary choice variable indicating whether agent i pays the cognitive cost and c_i is the magnitude of the cost for agent i. We assume that the cognitive cost is fixed for a given individual in that it does not depend on the register tax rate (it requires just as much effort to take a 6 cent register tax into account as a 7 cent one).

The timing of the model with cognitive costs proceeds as in Part I, except here we add an initial step in which agents choose whether or not they will take register taxes into account when deciding on their consumption of x. As before, all agents choose an intended consumption bundle $(\hat{x}, \hat{y})$ subject to their perceived budget constraint, which we can now express as $\widehat{BC}: x_i(p + b_i t_r + t_p) + y_i \leq M_i$.

A few final pieces of notation will be helpful. Let (x_i^*, y_i^*) denote the (optimal) bundle that i would consume if she were to pay attention to the register tax and let $(\tilde{x}, \tilde{y})$ denote the (sub-optimal) bundle she would consume were she to ignore the register tax. Agents who fail to pay the cognitive cost misperceive the after-tax price of x as being lower than it actually is; as a result, they over-spend on x and under-spend on y. The net change in i's utility from taking the tax into account is therefore given by

$$W(x_i^*, y_i^*, 1) - W(\tilde{x}_i, \tilde{y}_i, 0) = G_i - c_i$$

[2] The cognitive cost model we use as our starting point follows the basic approach laid out in Chetty, Looney, Kroft (2007).

where $G_i \equiv U(x_i^*, y_i^*) - U(\tilde{x}, \tilde{y})$ represents the agent's utility gain from consuming the optimal feasible bundle.

We assume that agents opt to pay the cognitive cost when doing so affords them greater utility: $b_i = 1\{G_i - c_i \geq 0\}$. Although a full-fledged comparison between the utility that would be achieved in the two scenarios would likely require more cognitive effort than simply taking the tax into account in the first place, it seems reasonable that the agents who decide to pay the cognitive cost tend to be the ones for whom doing so has the most benefit.[3]

Under the assumption that utility is additively separable in x and y, Chetty, Looney, and Kroft (2007) show that one can express G_i (the gain in consumption utility from taking the tax into account) as

$$G_i = \frac{1}{2} t^2 \varepsilon_{x,p} x_i^* v'(y_i^*) \left(\frac{1}{p+t} + \mu_i \gamma_i \right)$$

where $U(x,y) = u(x) + v(y)$, $\varepsilon_{x,p}$ is the elasticity (defined to be positive) of x_i^* with respect to its price, $\mu_i \equiv \frac{x_i^*}{y_i^*}$represents the optimal ratio of x to y, and γ_i measures the curvature of $v()$ at y_i: $\gamma_i \equiv \frac{-v''(y_i^*)}{v'(y_i^*)} y_i^*$.

CLK allow differences in the extent to which individuals take taxes into account by assuming heterogeneity in the cognitive costs that agents face (c_i), although they do not model the sources of that heterogeneity. Because our goal is to link differences in attentiveness to agents' income, we allow G_i to vary over individuals while abstracting from individual heterogeneity in cognitive costs: $c_i = \overline{c}$.[4] In particular, we focus on individual heterogeneity that arises from differences in agents' income. For a fixed tax rate and price, we can write G_i as a function of the agent's income(M_i)

$$G(M_i) = \frac{1}{2} t^2 \varepsilon_{x,p}(M_i) \left\{ \frac{x_i^*(M_i)}{p+t} + \mu_i(M_i) \gamma_i(M_i) \right\} v'(y_i^*(M_i))$$

The question we are interested in is whether low- or high-income individuals are more likely to take register taxes into account. Because agents are alike apart from their incomes, the question at hand is whether $G(.)$ is increasing or decreasing in M_i. Differentiating the above expression with respect to income yields:

$$\frac{\partial G_i}{\partial M_i} = \frac{1}{2} t^2 \left\{ \frac{\partial \varepsilon_{x,p}}{\partial M_i} x_i^* A v'(y_i^*) + \frac{\partial A}{\partial M_i} \varepsilon_{x,p} x_i^* v'(y_i^*) + \frac{\partial v'(y_i^*)}{\partial M_i} \varepsilon_{x,p} x_i^* A \right\} + \frac{\partial x_i^*}{\partial M_i} \varepsilon_{x,p} A v'(y_i^*)$$

[3] Another justification for this approach is that agents might make a one-time comparison between G_i and c_i to decide whether to pay the cognitive cost in future circumstance. A third possibility is that agents decide attentiveness tax by tax, rather than good by good (as assumed here). If so, low-income consumers may be particularly attentive to sales taxes because such taxes constitute a relatively high share of their expenditures.

[4] In reality, cognitive costs may also be correlated with income. The correlation may be positive, if high earners are better at cognitive tasks of this sort, or negative, if high earners have a greater opportunity cost of time. The extension to either of these cases is straightforward.

where $A = \frac{1}{p+t} + \mu_i(M_i)\gamma_i(M_i)$. Since A, x_i^*, $\varepsilon_{x,p}$ and $v'(y_i^*)$ are all positive, the key terms to sign are $\frac{\partial \varepsilon_{x,p}}{\partial M_i}$, $\frac{\partial A}{\partial M_i}$, $\frac{\partial x_i^*}{\partial M_i}$, and $\frac{\partial v'(y_i^*)}{\partial M_i}$.

First, consider $\frac{\partial v'(y_i^*)}{\partial M_i}$. We know that $\frac{\partial v'(y_i^*)}{\partial M_i} = v''(y_i^*)\frac{\partial y_i^*}{\partial M_i} < 0$ assuming concave utility and that y is a normal good. Intuitively, when the marginal utility of income declines rapidly with wealth, consumers who have little income to begin with are made much worse off by accidentally over-spending on x.

Second, consider $\frac{\partial x_i^*}{\partial M_i}$. This term will be positive as long as x is a normal good, but will be smaller in magnitude for goods for which consumption does not much change as income rises. In words, consumers who consume more will gain more from optimizing correctly simply because the consumption difference caused by the optimization error will be larger in magnitude. When demand for x is relatively insensitive to income, contribution of this term will be small.

Next consider $\frac{\partial \varepsilon_{x,p}}{\partial M_i}$. Are high- or low-income consumers more price sensitive in their demand for x? In general, theory is ambiguous as to whether elasticities rise or fall with income (the sign depends upon the magnitude of the third derivative of the utility function with respect to x).

Finally, consider $\frac{\partial A}{\partial M_i} = \frac{\partial \mu_i}{\partial M_i}\gamma_i + \frac{\partial \gamma_i}{\partial M_i}\mu_i$. Let's take the two pieces in turn. $\frac{\partial x_i^*}{\partial M_i}$ is clearly positive as long as x is a normal good. $\frac{\partial \mu_i}{\partial M_i}$ refers to how the optimal ratio of x to y changes with income. This term is zero when preferences are homothetic and negative for consumption goods that constitute a larger share of expenditures for poor consumers than for rich consumers. The second term, $\frac{\partial \gamma_i}{\partial M_i}$, captures change in the curvature of utility from wealth as income rises; it will be weakly negative when consumers exhibit constant or decreasing relative risk aversion.

We have highlighted the factors that determine whether attentiveness to a register tax is increasing or decreasing by income. What does the analysis imply for the case of cigarettes? Regardless of the good in question, low-income consumers suffer more from lost consumption of other goods when they accidentally overspend on the taxed good. The key determinants that vary between goods are $\frac{\partial x}{\partial M_i}$, $\frac{\partial \varepsilon_{x,p}}{\partial M_i}$, and $\frac{\partial \mu}{\partial M_i}$.

For the case of cigarettes, all three of these factors suggest that attentiveness to register taxes should decrease by income. The income elasticity of cigarettes is generally found to be quite small (or even negative), implying a low value for $\frac{\partial x}{\partial M_i}$. Similarly, on average, poor households spend a substantially larger fraction of their income on cigarettes compared to rich households (Chaloupka and Warner 2000), which implies that $\frac{\partial \mu}{\partial M_i} < 0$. Finally, the sign of $\frac{\partial \varepsilon_{x,p}}{\partial M_i}$ hinges on whether low- or high-income consumers are more sensitive to cigarette prices. The empirical literature on this question is mixed, with most studies concluding that low-income smokers are slightly more price sensitive and other studies finding the opposite. In our data, we find the differences in price-sensitivity between rich and poor smokers to be small, implying that $\frac{\partial \varepsilon_{x,p}}{\partial M_i}$ is small in magnitude.

As a whole, our model suggests that attentiveness to cigarette register taxes should decline by income. Low-income consumers suffer more when they over-spend on y because their marginal utility of wealth is greater than that of high-income consumers. Although the magnitude of the optimization error will in general be larger for high-income consumers (the difference between their intended and realized bundles is bigger), this factor is mitigated in the case of cigarettes by the fact that smoking demand is relatively insensitive to income and by the fact that low-income consumers spend a substantially higher fraction of their income on cigarettes compared to high-income consumers.

Web Appendix C: Additional Robustness Checks

This Appendix investigates the sensitivity of our analysis to additional robustness checks.

1. Alternative Specifications

So far, we have followed the approach taken by much of the smoking literature by separately modeling the extensive and intensive margins of cigarette consumption. This approach has the advantage of providing information about the mechanism by which tax changes reduce cigarette demand, in particular whether higher prices reduce demand by motivating smokers to quit or cut back. However, a drawback of this approach is that the intensive margin results may be biased by changes to the composition of the smoking population.[5]

As a robustness check, we estimate smoking demand using a linear regression and a Tobit model censored at zero. The dependent variable in these regressions is the number of cigarettes smoked per day, with the variable assigned a value of zero when the individual in question is not a smoker. Because the entire population of respondents is used, these approaches avoid the problem that tax rate changes affect selection into the smoking population. The flip side of the coin is that these models do not allow variables to differ in how they affect smoking demand on the intensive and extensive margins. Moreover, the Tobit specification relies on the normality of the unobservables and the linear functional form is probably unrealistic for an application in which so many of the observations have a dependent variable equal to zero. The results of the linear and Tobit specifications are presented in Appendix Table 1 and are consistent with the results from the two-part model used in the rest of the paper.

2. Delayed Responses to Tax Changes

So far we have assumed that smoking demand depends only upon current cigarette taxes, but it could be that tax changes affect consumer behavior with a lag. For example, higher prices might

[5] For example, suppose that smokers' demand for cigarettes were completely insensitive to price changes, but that light smokers quit when the price became too high. In such a world, a tax increase would appear to raise the intensity of smoking demand on the intensive margin merely by raising the fraction of heavy smokers in the smoking population.

motivate smokers to quit, but the quitting process itself could take several months. Alternatively, it could be that consumers take some time to learn about sales tax changes, only gradually incorporating them into their behavior. If these lags were different for high- and low-income consumers, it could provide an alternative explanation for our results.[6] To investigate this issue, we examine the sensitivity of our results to using various lags of the tax rates instead of the current rate.

$$\begin{aligned} y_{ismt} &= \alpha + \beta_1 \tau^e_{sm,t-k} + \beta_2 \tau^s_{sm,t-k} + \rho_1 \tau^e_{sm,t-k} LI_{ismt} + \rho_2 \tau^s_{sm,t-k} LI_{ismt} + \eta LI_{ismt} + \\ &\quad \gamma x_{smt} + \delta z_{ismt} + \mu_s + \lambda_t + \pi_m + \varepsilon_{ismt} \end{aligned} \tag{5}$$

where k, is three, six, or twelve months. The results are reported in Appendix Table 2 and suggest that our results are not being driven by differences in the time it takes high- and low-income consumers to respond to cigarette tax changes.

3. State Specific Trends

Although including state fixed-effects accounts for unobserved factors that affect the levels of smoking demand by state, it could be that changes in a state's tax rates are correlated with trends in that state's cigarette demand, such as anti-smoking sentiment. To reduce the influence of any such omitted third factors, we add state-specific year trends to the econometric model.[7] Appendix Table 3 shows that the estimated coefficients are largely unchanged by this addition.

$$\begin{aligned} y_{ismt} &= \alpha + \beta_1 \tau^e_{smt} + \beta_2 \tau^s_{smt} + \rho_1 \tau^e_{smt} LI_{ismt} + \rho_2 \tau^s_{smt} LI_{ismt} + \eta LI_{ismt} + \\ &\quad \gamma x_{smt} + \delta z_{ismt} + \mu_s + \lambda_t + \xi_s * t + \pi_m + \varepsilon_{ismt} \end{aligned} \tag{6}$$

[6]For example, high-income consumers may be better able to afford top of the line smoking-cessation products.

[7]Although our tax rate data is probably largely free of measurement error, including state trends could still cause substantial attenuation bias in the current context. Suppose that smoking demand depends upon a function of current and past tax rates, $x_t = x(a(L)x_t)$, where *a(L)* is some lag polynomial. The situation here is analogous to the standard measurement error problem: although the original tax variable x_t may be highly correlated with the "true" tax variable $a(L)x_t$, the new tax measure after including state trends may only be weakly correlated with the "true" tax rate, causing an attenuation bias.

Appendix Table 1: Alternative Demand Models

	(1) Linear	(2) Tobit
Excise Tax	-2.831***	-17.409***
	(0.612)	(3.102)
Sales Tax	-0.696	-0.327
	(2.542)	(12.036)
Excise*Low-income	0.628	4.772
	(1.396)	(5.405)
Sales*Low-income	-10.338**	-39.285***
	(4.464)	(15.199)
F-stat	5.55	7.39
prob>F	0.02	0.01
N	1,281,525	1,281,525

Standard errors clustered at the state level in parentheses.
274,137 observations have non-zero demand.
All specifications include individual demographic characteristics and state, year, and calendar month fixed effects.
The F-stat is associated with the test for equality between the excise*low-income and sales*low-income interaction coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 2: Timing

	Extensive Margin			Intensive Margin			Combined Effect		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	3 Month	6 Month	12 Month	3 Month	6 Month	12 Month	3 Month	6 Month	12 Month
Excise Tax	-0.116***	-0.104***	-0.112***	-0.349***	-0.332***	-0.259***	-3.633***	-3.353***	-3.187***
	(0.027)	(0.026)	(0.027)	(0.067)	(0.061)	(0.075)	(0.635)	(0.629)	(0.639)
Sales Tax	-0.057	-0.156	-0.133	0.363	0.417	0.574	0.079	-1.998	-0.976
	(0.109)	(0.105)	(0.085)	(0.352)	(0.308)	(0.386)	(2.836)	(2.252)	(2.050)
Excise*LI	0.046	0.034	0.038	0.048	0.042	0.072	0.932	0.681	0.866
	(0.059)	(0.065)	(0.069)	(0.113)	(0.113)	(0.126)	(1.417)	(1.547)	(1.665)
Sales*LI	-0.460**	-0.470***	-0.471***	-1.420**	-1.427*	-1.360*	-11.099***	-11.200***	-11.268***
	(0.179)	(0.167)	(0.154)	(0.672)	(0.721)	(0.717)	(4.215)	(4.078)	(3.871)
F-stat	7.63	8.51	10.86	5.51	4.55	3.94	9.09	9.23	10.66
prob>F	0.01	0.01	0.00	0.02	0.04	0.05	0.00	0.00	0.00
N	1,285,448	1,282,765	1,277,073	273,406	272,636	271,088	1,285,448	1,282,765	1,277,073

Standard errors clustered at the state level in parentheses.

All specifications include individual demographic characteristics and state, year, and calendar month fixed effects.

Outcome variables: probability of smoking (extensive), log cigarette demand (intensive), and cigarette demand in levels (combined).

The F-stat is associated with the test for equality between the excise*low-income and sales*low-income interaction coefficients.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Appendix Table 3: State Trends

	Extensive Margin	Intensive Margin	Combined Effect
	(1)	(2)	(3)
Excise Tax	-0.017	-0.348***	-1.357**
	(0.026)	(0.095)	(0.685)
Sales Tax	0.122	0.342	3.436
	(0.131)	(0.425)	(3.336)
Excise*Low-income	0.060	0.019	1.078
	(0.057)	(0.112)	(1.379)
Sales*Low-income	-0.501***	-1.378**	-11.814***
	(0.182)	(0.665)	(4.380)
F-stat	8.62	5.09	8.98
prob>F	0.01	0.03	0.00
N	1,288,031	274,137	1,288,031

Standard errors clustered at the state level in parentheses.
All specifications include individual demographic characteristics and state, year, and calendar month fixed effects.
Outcome variables: probability of smoking (extensive), log cigarette demand (intensive), and cigarette demand in levels (combined).
The F-stat is associated with the test for equality between the excise*low-income and sales*low-income interaction coefficients.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$