Jérôme Adda and Francesca Cornaglia

Taxes, cigarette consumption, and smoking intensity: reply

Article (Published version) (Refereed)

Original citation:
Adda, Jérôme and Cornaglia, Francesca (2013) *Taxes, cigarette consumption, and smoking intensity: reply*. American Economic Review, 103 (7). pp. 3102-3114. ISSN 0002-8282
DOI: 10.1257/aer.103.7.3102

American Economic Review 2013, 103(7): 3102–3114
http://dx.doi.org/10.1257/aer.103.7.3102

Taxes, Cigarette Consumption, and Smoking Intensity: Reply†

By Jérôme Adda and Francesca Cornaglia*

The economic literature has long established the positive relationship between cigarette taxes, prices, and cigarette consumption. Recently, attention has been brought to the fact that the behavior of smokers may offset the consumption effect of a tax increase (Adda and Cornaglia 2006). In a later contribution Abrevaya and Puzzello (2012)—henceforth, AP (2012)—re-examined Adda and Cornaglia's (2006)—henceforth, AC (2006)—analysis using a larger sample from the same source and have cast some doubts on the robustness of AC's (2006) analysis. In particular, they claim that the limited within-state tax variation observed in the dataset does not lead to precise cigarettes and cotinine tax-elasticities estimates. AP (2012) also argue that the use of the appropriate statistical inference leads to a further decrease in the precision of AC's (2006) estimates.

In this reply, we show that the intensity of smoking, defined as the ratio of cotinine levels to the number of cigarettes smoked, does respond to changes in excise taxes as previously found by AC (2006). We do so by using a dataset that spans from 1988 to 2006, allowing for more variations in taxes than in AC (2006) and in AP (2012) who both considered the period 1988–1994. We also stress the importance of considering the appropriate timing of taxes when analyzing their effect on smoking. AC (2006) and the replication of their results in AP (2012) used contemporaneous taxes, while we make the case for using lagged taxes.

We also make a number of other contributions to the literature on smoking behavior and tobacco control. First we show that smoking intensity responds to price changes over this period, and that consistent estimates require the use of instrumental variables because of endogeneity issues. We show that OLS estimates are biased toward finding no effects. We find that the tax elasticity of smoking intensity is significantly different from zero and equal to 0.07 and that the price elasticity is higher, at around one. We also find considerable heterogeneity in the response to tax increases across different groups, and notably across different race groups.

We then investigate whether biomarkers such as cotinine measures are informative of long-run outcomes. We provide supporting evidence using the Coronary Artery Risk Development in Young Adults (CARDIA) study. This dataset follows smokers over a period of 15 years, with information on cigarette consumption as well as cotinine levels, a biomarker of smoking. The panel nature of the CARDIA study

*Adda: Department of Economics, European University Institute, Via della Piazuolla 43, Firenze, 50133 Italy (email: jerome.adda@eui.eu); Cornaglia: School of Economics and Finance, Queen Mary University of London, Mile End Road, E1 4NS, London, UK (email: f.cornaglia@qmul.ac.uk). We are grateful to David Jacobs, Jared Reid, Jason Abrevaya, Laura Puzzello, and an anonymous referee for comments. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

† Go to http://dx.doi.org/10.1257/aer.103.7.3102 to visit the article page for additional materials and author disclosure statement(s).

allows us to investigate long-run effects that repeated cross-sectional data such as NHANES do not. We show that cotinine levels are a strong predictor of smoking cessation, over and above the number of cigarettes smoked. We finally use this data to shed further light on dynamic selection, and its potential to bias OLS regressions of smoking intensity on changes in prices and taxes.

We first present, in Section I, new evidence of compensatory behavior in response to tax changes. In Section II, we present effects of prices on smoking intensity. Section III shows that cotinine measures are significant predictors of long-run outcomes. Finally, Section IV investigates the potential for dynamic bias due to sample selection.

I. Smoking Intensity and Taxes

In this section we provide evidence of a link between taxes and smoking intensity. We first discuss four different issues relevant to the estimation, namely the need for an extended dataset, the importance of the timing of taxes, the use of weights when analyzing data from the NHANES dataset, and the issue of clustering of the standard errors.

A. *Extended Data 1988–2006*

AP (2012) expand the original NHANES III dataset (1988–1994) used by AC (2006) adding more states over the same period. A difference between the two datasets is the inclusion of tobacco states. These states are characterized by higher cigarette consumption, lower taxes, and little variation in excise taxes over the period considered. Essentially, when including these states, the effect of taxes on smoking intensity becomes much smaller.[1] AP (2012) observe that one needs sufficient variation in taxes within states to be able to identify the parameter of interest. This is certainly the case, and in a context of regressions which include state fixed effects, it is important to expand the data to add more time variation. The NHANES data have expanded over the years and more waves are now available for analysis, covering a period which has seen more variation in prices and excise taxes than the early 1990s. We supplement the NHANES III dataset used by AC (2006) and AP (2012) with later waves between 1999 and 2006. This data have been used by Adda and Cornaglia (2010)—henceforth, AC (2010)—to investigate the effect of smoking bans and excise taxes on nonsmokers. In this paper we analyze the effect of taxes on smoking intensity.

Table 1 presents some demographic characteristics for the samples 1988–1994 and 1988–2006. The extended sample is slightly older and there is a lower prevalence of white individuals as well as individuals who do not have a high school degree. The number of cigarettes consumed per day (conditional on smoking) is very similar in both samples. The same is true for cotinine concentration. The advantage of the expanded dataset is the greater coverage of American states, going from 26 to 35, and especially the increased number of state-year observations, on which relies the identification of the tax effects in the difference-in-difference method we employ. This number increases from 60 to 147. Adding the years to 2006 improves

[1] See AP (2012), Table SM10 in the online appendix.

Table 1—Descriptive Statistics: NHANES

	NHANES 1988–1994	NHANES 1988–2006
Observations	3,514	6,452
Age	42.1	43.3
	(16.0)	(15.8)
Male, percent	53.8	54.2
White, percent	61.4	58.1
African-American, percent	36.0	30.6
High school dropouts, percent	45.4	41.6
High school degree, percent	33.0	31.3
Number of cigarettes per day	16.2	16.2
	(11.3)	(10.9)
Cotinine level ng/ml	232.5	232.2
	(141.5)	(135.0)
Number of states observed	26	35
Number of years observed, per state	2.3	4.2
	(1.6)	(3.4)
Number of state × year observations	60	147
Number of observations, per state	218.9	329.5
	(165.5)	(256.0)

Notes: Standard deviations in parenthesis where appropriate. Unweighted means are displayed.

Source: Authors' calculations.

the analysis greatly as there has been lots of variation in taxes between 1988 and 2006, and allows, as pointed out by AP (2012) to better identify the parameter of interest. For a further description of the dataset, we refer the reader to AC (2010).

B. *Timing of Taxes*

A main difference between AC (2006) and our subsequent work on the effect of excise taxes on smoking behavior, as in AC (2010), is the issue of the timing of taxes. Our original work related smoking intensity in a given year to the contemporaneous tax level. However, the contemporaneous tax refers to the tax at the end of the fiscal year at the end of June. In addition, the tax measure we use is the real one, which is constructed using the nominal rate, deflated using the end of the year inflation rate. In contrast, measures of smoking are collected throughout the civil year in the NHANES survey. Using contemporaneous taxes is therefore problematic as a large fraction of the sample is imputed a tax, which does not apply to them yet. This is why AC (2010) used the lagged tax in their specifications. Acknowledging the short-comings of using contemporaneous taxation as in AC (2006), we use lagged taxes to correct for the mismatch between the interview date and the relevant tax measure.

C. *Sampling Weights*

In this section we address the issue of the use of sample weights in regressions involving NHANES data. AP (2012) follow DuMouchel and Duncan (1983) who show that differences between coefficients in weighted and unweighted regressions are a sign of mispecification. As pointed out in DuMouchel and Duncan (1983), this is however only true when weights are derived from exogenous variables. The weighting scheme in NHANES is more complex. The weights are a function of demographic variables

TABLE 2—CORRELATION BETWEEN SAMPLE WEIGHTS AND SMOKING BEHAVIOR

	NHANES 1988–1994		NHANES 1988–2006	
log number of cigarettes	0.27**	0.15 **	0.22**	0.07**
	(0.00)	(0.00)	(0.00)	(0.00)
log cotinine	0.10**	0.11**	0.10**	0.05**
	(0.00)	(0.00)	(0.00)	(0.00)
log smoking intensity	−0.24**	−0.05**	−0.13**	−0.02*
	(0.00)	(0.00)	(0.00)	(0.07)
Observations	3,514	3,514	6,318	6,318
Controls	No	Yes	No	Yes

Notes: p-values in parenthesis. Controls include age, age squared, sex, race, education, state of residence, and year indicators.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations.

because the survey oversampled certain categories such as age groups or racial groups. In addition, the weights were constructed to take into account non-participation, especially for the medical exam from where the cotinine measure is taken. Hence, the weights are also a function of the endogenous variables. This is why the National Center of Health Statistics (NCHS) strongly recommends to use weights in the analysis (we refer the reader to the NHANES guidelines (NCHS 1996) for a detailed description). AC (2006) overlooked this issue, but AC (2010) use them in their regressions.

Table 2 shows empirical evidence that the sample weights are indeed correlated with the outcome variable, over and above demographic characteristics. The table displays the correlation between the outcome variable of our regressions and the weights (p-values in parenthesis). We distinguish between three outcome variables, the log number of cigarettes smoked per day, log cotinine concentration, and log smoking intensity. The correlation between these three outcome variables and the sample weights varies between 0.1 and 0.27 in the sample 1988–1994, and between 0.1 and 0.22 in the extended sample (1988–2006). Once we control for a set of observable characteristics, which include age, sex, race, education, region of residence, and year of examination, the correlation is closer to zero but statistically different from zero. The correlation ranges between 0.05 and 0.15 in the sample 1988–1994 and from −0.02 to 0.07 in the extended sample. The results show that the sample weights are indeed correlated with the endogenous variable, even when a set of demographic controls are included. The problem is particularly severe in the sample used by AP (2012), but less so with the extended dataset we use in this article. In the presence of endogenous stratification, the assumptions in DuMouchel and Duncan (1983) are violated. As discussed in Maddala (1983), the use of weights is recommended instead. However, the right way to adjust for endogenous weighting is not obvious and we leave this to future work. When using data from NHANES in this reply, we present the results with and without weights for comparison with previous results.

D. *Computation of the Standard Errors*

AP (2012) rightly point out that the standard errors should be clustered at state level only, rather than at state times year level, following the findings of Bertrand, Duflo, and Mullainathan (2004). This is the case because of potential serial correlation in the error term within states. AP (2012) show that clustering has an effect on the precision of the estimates. Going from a state-year clustering to state clustering alone, tends to increase the standard errors of the coefficients, although a subset of them decreases (see Table 1 in AP 2012). In contrast, AC (2006) used clustering at state-year level. We adopt the more robust methodology below, clustering at state level only.[2]

E. *Estimation Results*

We now turn to the behavioral effect of taxes. We first note that there is a difference in focus between AC (2006) and AP (2012). AC (2006) are interested in the existence of compensatory behavior, which amounts to testing whether the ratio of cotinine to cigarettes is significantly related to taxes. In other words, the focus is on whether the tax elasticity of cigarettes is larger than the tax elasticity of cotinine. The fact that smoking intensity responds to public policies is an important finding for the design of health policies. As argued in AC (2006), this also has consequences on the estimation of popular models such as the rational addiction model. AP (2012), on the other hand, mainly test whether the elasticity of cotinine or cigarettes with respect to taxes or prices is significantly different from zero. It is of course possible that both the elasticities of cotinine and of cigarettes smoked are insignificant—perhaps because of lack of variability in the data—and that the elasticity of the intensity of smoking is significantly different from zero.

AC (2006) provide results for various specifications, including conditioning on onset of smoking or cotinine levels, to address the issue of dynamic selection. We present the regressions of the baseline specification, for the subset of smokers who started smoking before age 17. We return to the dynamic selection issue in Section IV.

Table 3 displays the results of OLS regressions of smoking intensity (defined as the ratio of cotinine and the number of cigarettes smoked), the number of cigarettes smoked and cotinine levels, on log state taxes as well as demographic variables, state indicators and year indicators. All left-hand-side variables are log transformations for ease of interpretation.

The table has four columns of results, the first two using the contemporaneous tax as a regressor, the following ones using the lagged tax instead. In both cases, the table displays results for the sample 1988–1994 and the extended sample 1988–2006. In addition, the table has two panels, where we use unweighted and weighted regressions.

NHANES III and the subsequent NHANES surveys differ in the amount of demographic variables which were recorded. NHANES III has a wider set of explanatory variables. To make our results consistent across the table, we control for a set of variables which are present throughout all waves. These controls are age, sex, race, and education levels. AC (2006) also controlled for occupation, household size, and passive smoking, and in some cases for the time of examination and height. However, there is little reason why these additional controls should be correlated with state taxes, especially when the regressions include state and year fixed effects.

[2] Note that AC (2010) rely on such inferences as well. Furthermore, AP (2012) also use a bootstrap based method to correct the standard errors along the lines of Cameron, Gelbach, and Miller (2008), because the number of clusters in their data is small. In this paper, this approach is of less relevance as the number of clusters is considerably higher.

TABLE 3—TAX ELASTICITY OF SMOKING INTENSITY, NUMBER OF CIGARETTES, AND OF COTININE: BASELINE

	NHANES 1988–1994	NHANES 1988–2006	NHANES 1988–1994	NHANES 1988–2006
	Contemporaneous tax		Lagged tax	
Unweighted regressions				
Elasticity smoking intensity	0.034	0.051	0.062	0.074**
	(0.13)	(0.046)	(0.055)	(0.035)
Elasticity number of cigarettes	0.072	−0.035	−0.001	−0.070
	(0.20)	(0.046)	(0.088)	(0.051)
Elasticity, cotinine	0.105	0.016	0.060	0.005
	(0.13)	(0.026)	(0.057)	(0.030)
Weighted regressions				
Elasticity smoking intensity	0.13	0.04	0.227**	0.0691**
	(0.14)	(0.04)	(0.089)	(0.029)
Elasticity number of cigarettes	0.059	−0.005	−0.059	−0.089**
	(0.22)	(0.044)	(0.132)	(0.044)
Elasticity, cotinine	0.19	0.036	0.168*	−0.020
	(0.19)	(0.028)	(0.095)	(0.038)
Observations	3,514	6,318	3,514	6,318

Notes: All regressions control for age, sex, race, education, year, and state effects. Robust standard errors clustered at state level. Regression results displayed in columns 1 and 2 use the contemporaneous tax as a regressor, whereas those in columns 3 and 4 use lagged taxes.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations.

Indeed, the results in AC (2006) show that the coefficients of interest are not significantly different when including a fuller list of controls.

Table 3, column 1, confirms the results obtained by AP (2012).[3] Unweighted regressions for the period 1988–1994 show that a one percent increase in taxes leads to a 0.03 percent increase in smoking intensity, but this elasticity is not statistically significant. When using the extended data (column 2) the elasticity of smoking intensity increases to 0.05, but is still not significant. Column 3 displays the results when lagged taxes are used instead. Unsurprisingly, the standard errors are smaller and the point estimates larger. The elasticity of smoking intensity is 0.06 in the 1988–1994 sample and 0.07 in the extended sample, significantly different from zero at the 5 percent confidence level. This number is smaller than the one obtained in the NHANES sample in AC (2006), but in line with the results in the same paper when using data from NHANES 1999–2000. The second panel uses the weights provided in the NHANES dataset. The use of weights appears to be of particular importance in NHANES III, as the elasticity of smoking intensity is larger than the unweighted one, equal to 0.13 and 0.23 respectively, using either contemporaneous or lagged taxes. The second elasticity is significantly different from zero. Finally, on the larger sample, the use of weights appears to be less important, as we find a very similar tax elasticity of smoking intensity, equal to about 0.07 percent and statistically

[3] The results in Table 3 differ somewhat from those in AP (2012) as we use a different set of controls. For instance, the dataset does not record occupation in all years, nor the time of examination.

Table 4—Tax Elasticity of Smoking Intensity, Number of Cigarettes, and of Cotinine: Early Starters

	NHANES 1988–1994	NHANES 1988–2006
Unweighted regressions		
Elasticity smoking intensity	0.081	0.097**
	(0.067)	(0.037)
Elasticity number of cigarettes	−0.022	−0.105**
	(0.093)	(0.050)
Elasticity, cotinine	0.059	−0.009
	(0.086)	(0.031)
Weighted regressions		
Elasticity smoking intensity	0.205**	0.109**
	(0.077)	(0.033)
Elasticity number of cigarettes	0.084	−0.126**
	(0.134)	(0.060)
Elasticity, cotinine	0.280**	−0.016
	(0.118)	(0.046)
Observations	2,060	3,611

Notes: All regressions control for age, sex, race, education, year, and state effects. The regressions exclude smokers who started after age 17 and use lagged taxes as a regressor. Robust standard errors clustered at state level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations.

significant at the 5 percent confidence level. Note that we also find a significant tax elasticity for the number of cigarettes smoked (−0.09 percent), but cotinine levels do not respond much to changes in taxes. In the remainder of the paper, the analysis uses lagged taxes as an explanatory variable for the reasons detailed above.

As discussed in AC (2006), an OLS regression could overstate the effect of taxes on smoking intensity because of a dynamic selection bias, where light smokers could be more responsive to changes in taxes and quit smoking at a higher rate. As a robustness check AC (2006) used a subsample of smokers who started early (before age 17) and who are therefore less likely to quit. We repeat this analysis in Table 4. Column 1 uses the sample period used in AP (2012). None of the elasticities using unweighted regressions are significantly different from zero. With weights, we find evidence of a significant effect of taxes on smoking intensity. With the larger dataset (column 2), the results are in line with Table 3 and the original results of AC (2006). A 1 percent increase in excise taxes increases smoking intensity by about 0.1 percent (this is the case in both weighted and unweighted regressions). The increase in smoking intensity is the outcome of a significant decrease in the number of cigarettes smoked, and not the effect of taxes on cotinine concentration. These results suggest that the findings on the whole sample of smokers is not due to a change in the composition of smokers. We present further evidence on this issue below in Section IV.

F. *Heterogeneous Effects*

Table 5 presents evidence of heterogenous effects. We estimate the tax elasticity of smoking intensity for various subgroups of smokers. We find evidence that

Table 5—Smoking Intensity: Tax Elasticities for Different Subsamples

	Coefficient	SE	Sample size
Full sample, 1988–2006	0.069**	(0.029)	6,318
Men	0.105**	(0.039)	3,423
Women	0.049	(0.045)	2,895
White	0.053	(0.039)	3,690
Black	0.193**	(0.075)	1,943
Ages 17–29	0.085	(0.06)	1,464
Ages 30–44	0.075*	(0.041)	2,167
Ages 45+	0.062	(0.043)	2,687
Below median income	0.110**	(0.055)	2,919
Above median income	0.058*	(0.034)	3,399

Notes: Sample: Regressions control for age, sex, race, education, state of residence, and year of interview. The results reported use the lagged excise tax. All regressions are weighted using NHANES sample weights. Standard errors in parenthesis are clustered by state. Median 2000 income is $26,500.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations, NHANES 1988–2006.

compensatory behaviors are more important for men than for women, and in particular more substantial for African-Americans than for whites. AC (2006) found evidence that African-Americans smoke cigarettes more intensively than other racial groups. It appears that they are also more reactive to tax changes. There is not much evidence of a change in the elasticity by age groups. None are significantly different from zero at the 5 percentage confidence level. One reason for this is that sample sizes are becoming smaller when one stratifies the sample in such a way, showing the limits of such an exercise.

Finally, smokers with income below the median income (defined as $26,500 annual income in 2000 dollars) have a higher elasticity than those above the median. The elasticities vary from 0.1 to 0.05. This result is in line with the predictions of the model in AC (2006).

II. Smoking Intensity and Prices

AC (2006) did not provide price elasticities for two reasons. First, there is an endogeneity issue to which we will come back below. Second, from a public health point of view, states and governments can only manipulate prices through taxes, so the knowledge of the elasticity with respect to taxes is important to shape policy.

AP (2012) write that an OLS regression can recover consistent estimates of price elasticities, arguing that endogeneity is not a problem with micro data. We dispute this argument. First, the regression is indeed using micro data, but the real variation is at the state times year level, as argued by AP (2012) in their discussion about standard errors. Second, in the presence of aggregate (state) shock to demand, it is likely that tobacco companies change their prices to respond to such shocks. The fact that the influence of individuals is too small to affect prices is not the issue. In the presence of endogeneity, an OLS regression would tend to produce coefficients which are biased toward zero. A positive demand shock would induce an endogenous increase in prices, which would counteract the causal effect of prices on demand. To solve the issue of endogeneity, we instrument prices with arguably

Table 6—Price Elasticity of Smoking Intensity: Number of Cigarettes and Cotinine

	OLS	IV
Unweighted regressions		
Elasticity smoking intensity	0.062	0.761*
	(0.336)	(0.418)
Elasticity number of cigarettes	0.101	−0.245
	(0.501)	(0.577)
Elasticity, cotinine	0.164	0.517
	(0.377)	(0.418)
F-statistic first stage; (p-value)	—	57.12; (0.00)
F-test of endogeneity; (p-value)	—	10.99; (0.02)
Weighted regressions		
Elasticity smoking intensity	0.393	1.056**
	(0.397)	(0.455)
Elasticity number of cigarettes	−0.514	−0.891*
	(0.313)	(0.492)
Elasticity, cotinine	−0.121	0.165
	(0.365)	(0.407)
Number of observations	4,870	4,870
F-statistic first stage; (p-value)	–	57.12; (0.00)
F-test of endogeneity; (p-value)	–	10.99; (0.002)

Notes: All regressions control for age, sex, race, education, year, and state effects. The results reported use lagged prices or taxes. Prices are instrumented with state tax levels. Robust standard errors clustered at state level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations, NHANES 1988–2006.

exogenous tax shocks. Taxes are often changed to raise revenue and not to counter demand shocks.

In Table 6 we report estimates of price elasticities. The first column displays OLS results, without sample weights in the first panel, and with weights in the second. The OLS results are similar to those found in AP (2012), and consistent with the intuition detailed above. None of the elasticities are significantly different from zero. The second column displays instrumental variable estimates. We also report the F-statistic for the first stage, which has a value of 57. This indicates that state excise taxes are a significant predictor of prices, over and above state and year indicators. The instrumental variable estimates show that smoking intensity responds to price changes. Without sample weights, we find that a 1 percent increase in prices leads to a 0.76 percent increase in smoking intensity, significant at the 10 percent level. When sample weights are used, we find a price elasticity of smoking intensity of 1.06, significantly different from zero at the 5 percent confidence level.[4] The table also displays an F-test for the endogeneity of prices, which we carried out for the regressions involving smoking intensity. They show that the null of no endogeneity is strongly rejected, at a confidence level of 2 percent.

[4] The price elasticity of the number of cigarettes varies from 0.1 to −0.9 depending on the estimation method and the weighting we use. Some of our estimates are outside the range commonly found in the literature, which is typically between −0.25 and −0.5. Note however, that given the precision of our estimates, they are not significantly different from any number within that range.

III. Long-Run and Short-Run Measures of Smoking

It has been argued (AP 2012) that the number of cigarettes smoked per day is a long-run measure of smoking whereas cotinine levels are a short-run measure. It is not clear what justifies such a categorization, especially when only cross-sectional data is at hand in NHANES. Following this view, AP (2012) dismiss the measure of smoking intensity constructed by AC (2006), that is, the ratio of cotinine extracted per cigarette smoked. We dispute this claim and present two arguments in favor of considering cotinine measures a long-term measure of smoking (and health).

The first one is the one developed in AC (2006). They present evidence of long-run health outcomes which are linked to the way cigarettes are smoked. We refer the reader to the discussion about lung cancer rates by race in the United States, in Section IIB in AC (2006).

Here we provide further evidence using data from the Coronary Artery Risk Development in Young Adults (CARDIA) study. These data allow us to follow the same smokers over time, in contrast with the cross-sectional nature of the NHANES dataset.

The CARDIA data were collected between 1985 and 2001 in four locations in the United States (Alabama, Illinois, Minnesota, and California). A group of 5,115 individuals aged 18–30, were followed over 15 years, which provides ample longitudinal variation. In each wave the survey asks a number of questions on smoking behavior, whether the individual is still smoking at the time, and if not, the age at which smoking cessation took place. We define age at smoking cessation as the age of the first recorded quit, whether the individual relapses after that period or not. We therefore only consider one smoking spell per smoker. In addition, the survey records the number of cigarettes smoked in all waves, cotinine levels in the first wave as well as the age of smoking initiation. Table 7 presents key descriptives of the dataset, and we refer the reader to Friedman et al. (1988) for a detailed description of this dataset.

We measure the propensity to quit as a function of both the number of cigarettes smoked and the cotinine levels at baseline in 1985. We estimate a duration model, using a Cox proportional form, where we stratify by geographical location, sex, race, education, and age at smoking onset. The data provide us with information on 11,073 observations following 1,459 smokers until they quit or are right censored. We normalized both the number of cigarettes and the cotinine levels to have mean zero and variance one, so that we can interpret the coefficients in a straightforward way.

Table 8 displays the results. We find that the number of cigarettes smoked per day is not significant at the 5 percent level, but interestingly, cotinine levels are a highly significant predictor of quitting behavior (column 1). Controlling for the number of cigarettes, a one standard deviation increase in cotinine levels at baseline decreases the likelihood to quit by 49 percent. Moreover, holding cotinine levels constant, a one standard deviation increase in cigarettes is associated with a decrease of 2 percent in the likelihood of quitting. Hence the statement that cotinine is only a short-run measure of smoking does not appear to be grounded in facts as it significantly predicts quitting over a period of 15 years.

We also explored heterogenous effects. Columns 2 and 3 of Table 8 present the effect of cotinine and the number of cigarettes on quitting behavior by sex. Men and

Table 7—Descriptive Statistics: CARDIA Sample

	All individuals at baseline (year 0)	Year 15	Year 15 individuals still smoking
Observations	1,546	923	633
Smoking prevalence, percent	100	68.5	100
Mean number of cigarettes	13.1	9.2	13.4
	(9.1)	(10.0)	(9.5)
Mean number cigarettes at baseline	13.1	13.4	13.8
	(9.1)	(9.5)	(8.9)
Mean cotinine level at baseline ng/ml	224.4	222.2	241.1
	(158.4)	(153.8)	(152.1)
Male, percent	47.2	45.3	47.0
White, percent	57.1	52.9	60.5
African-American, percent	42.8	47.1	39.5
Mean age	25.0	40.0	40.0
	(3.6)	(3.6)	(3.6)
Mean years of schooling	12.8	13.0	12.8
	(2.0)	(2.0)	(1.9)

Note: Standard deviations in parenthesis where appropriate.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Authors' calculations, CARDIA data.

Table 8—Hazard Rates for Quitting Smoking

	All	Men	Women
Number of cigarettes	−0.021	0.183	−0.153
	(0.077)	(0.115)	(0.11)
Cotinine levels	−0.493**	−0.598**	−0.445**
	(0.093)	(0.162)	(0.114)
Observations	1,459	691	768
Time at risk	11,073	5,418	5,655

Notes: Cox proportional regression. Age is the analysis time. Stratified by sex (first column), race, education, state, year and age at smoking onset. Number of cigarettes and cotinine levels have been normalized to mean zero and variance one for ease of interpretation.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Source: Authors' calculations, CARDIA data.

women differ in their propensity to quit. In particular, cotinine levels play a bigger role for men than for women.

IV. Dynamic Selection

An important issue refers to dynamic selection, whereby smokers who quit following an increase in taxes may come disproportionally from a low smoking intensity group. This issue has been raised by both AC (2006) and AP (2012). If this is the case, an OLS regression of smoking intensity on excise taxes may find a spurious positive effect due to a change in composition in the pool of smokers. AC (2006) investigate this point in two ways. They first include in their sample individuals who

TABLE 9—EFFECT OF STATE TAX ON QUITTING BY LEVEL OF SMOKING INTENSITY

	All	Intensity < median	Intensity > median
Controlling for year and year square			
log tax	0.629**	0.557**	0.930**
	(0.230)	(0.295)	(0.409)
Controlling for year fixed effects			
log tax	0.380*	0.231	0.693
	(0.239)	(0.311)	(0.434)
Observations	1,477	857	616
Time at risk	11,185	6,475	4,698

Notes: Cox proportional regression, stratified by sex, race, education, state, and age at smoking onset. Regressions include year either through a quadratic specification or through year fixed effects.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Source: Authors' calculations, CARDIA data.

are less likely to quit, for instance individuals who started smoking at a young age, to explore the effect of taxes on smoking intensity. Second, they use an econometric technique developed by Manski (1994), to use worst case bounds. However, with only cross-sectional data this point could not be fully addressed as the bounds tend to be large. Using the panel data from CARDIA, we now present new evidence on this issue, which helps to interpret the results in AC (2006), and those in Table 3 in this article.

We estimate the effect of tax changes on quitting behavior and whether the effects vary with smoking intensity, measured as the ratio of cigarettes to cotinine levels at baseline. Using geographical information on the center of examination, we merge information on excise taxes to the original CARDIA data. We therefore have variation on taxes across years and geographical location, which we exploit to estimate a model of the duration to quitting.

Table 9 presents the effect of taxes on quitting, for the whole sample, as well as for smokers divided into two groups, below or over the median smoking intensity measured at baseline in 1985. The first panel displays the results, controlling for age, sex, race, education, state of residence, and a quadratic time trend. We find a significant effect of taxes on quitting behavior. A doubling of the excise tax increases the likelihood of quitting by 63 percent. Interestingly, the point estimate of this effect is larger for individuals with a high smoking intensity than for a low smoking intensity. The former have a likelihood of quitting of 56 percent and the latter of 93 percent. However, the estimation is not precise enough to conclude that the difference between the two is significant. The second panel includes year fixed effects rather than a quadratic trend. We get qualitatively similar results, although with much less precision.

The results of this exercise suggest that the dynamic bias may not be of such importance, and that OLS regressions do not overstate the effect of taxes on smoking intensity. Understanding why individuals with higher smoking intensity may be more likely to quit when taxes increase is an interesting question that we leave to future research.

V. Conclusion

Considering a dataset that spans from 1988 to 2006 we show that the intensity of smoking, defined as the ratio of cotinine levels to the number of cigarettes smoked, responds to changes in excise taxes as previously found by Adda and Cornaglia (2006).

We also show that smoking intensity responds to price changes and that consistent estimates can be obtained using instrumental variables because of endogeneity issues. We find a significant effect of the tax elasticity of smoking intensity (0.07) and a price elasticity of around one. We also find considerable heterogeneity in the response to tax increases across different groups, and notably across different race groups.

We then provide supporting evidence that biomarkers such as cotinine are informative of long-run outcomes. Using a dataset that follows smokers for over 15 years we show that cotinine levels are a strong predictor of smoking cessation, over and above the number of cigarettes smoked. Finally we use this longitudinal dataset to shed light on the issue of dynamic selection, and the potential bias of OLS regressions of smoking intensity on changes in prices and taxes.

REFERENCES

Abrevaya, Jason, and Laura Puzzello. 2012. "Taxes, Cigarette Consumption, and Smoking Intensity: Comment." *American Economic Review* 102 (4): 1751–63.

Adda, Jérôme, and Francesca Cornaglia. 2006. "Taxes, Cigarette Consumption, and Smoking Intensity." *American Economic Review* 96 (4): 1013–28.

Adda, Jérôme, and Francesca Cornaglia. 2010. "The Effect of Bans and Taxes on Passive Smoking." *American Economic Journal: Applied Economics* 2 (1): 1–32.

Adda, Jérôme, and Francesca Cornaglia. 2013. "Taxes, Cigarette Consumption, and Smoking Intensity: Reply: Dataset." *American Economic Review.* http://dx.doi.org/10.1257/aer.103.7.3102.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (1): 249–75.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90 (3): 414–27.

DuMouchel, William H., and Greg J. Duncan. 1983. "Using Sample Survey Weights in Multiple Regression Analyses of Stratified Samples." *Journal of the American Statistical Association* 78 (383): 535–43.

Friedman, Gary D., Gary R. Cutter, Richard P. Donahue, Glenn H. Hughes, Stephen B. Hulley, David R. Jacobs Jr., Kiang Liu, and Peter J. Savage. 1988. "CARDIA: Study Design, Recruitment, and Some Characteristics of the Examined Subjects." *Journal of Clinical Epidemiology* 41 (11): 1105–16.

Maddala, G. S. 1983. *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge, UK: Cambridge University Press.

Manski, Charles F. 1994. "The Selection Problem." In *Advances in Econometrics: Sixth World Congress*, Vol. 1, edited by Christopher Sims, 143–70. Cambridge, UK: Cambridge University Press.

National Center for Health Statistics (NCHS). 1996. "Analytic and Reporting Guidelines: The Third National Health and Nutrition Survey, NHANES III (1988–94)." Hyattsville, MD: Centers for Disease Control and Prevention.