



Are Graphic Warning Labels Stopping Millions of Smokers? A Comment on Huang, Chaloupka, and Fong¹

Trinidad Beleche², Nellie Lew³, Rosemarie
L. Summers⁴, and J. Laron Kirby⁵

[LINK TO ABSTRACT](#)

Over the past several decades, the prevalence of smoking has declined in much of the developed world (OECD 2012), including in the United States and Canada. During this time, as certain health consequences of cigarette use have become more well-known, governments have implemented policies that impose disincentives or restrictions on the consumption of cigarettes in the form of taxes, indoor smoking bans, and other policy measures. One approach is to require manufacturers to place warning labels on tobacco products. The warning labels traditionally have taken the form of black-and-white text boxes on packs of cigarettes explaining the potential health consequences of smoking. More recently, beginning in 2001 in Canada,⁶ governments have begun mandating that warning labels contain colorful, graphic depictions of those consequences, depictions that can be gruesome, depressing, and often ghastly. As of October 2016, at least 105

1. This article reflects the views of the authors and should not be construed to represent the views or policies of the United States Food and Drug Administration (FDA).

2. FDA, Silver Spring, MD 20993. (Corresponding author.) We thank Don Kenkel, Clark Nardinelli, and Ian Irvine for their comments on an earlier version of the manuscript. All errors are our own.

3. FDA, Silver Spring, MD 20993.

4. FDA, Silver Spring, MD 20993.

5. FDA, Silver Spring, MD 20993.

6. In 2012 Canada revised its policy to require new messages and new and larger graphics.

countries have adopted such policies.⁷

Figure 1. Number of countries adopting graphic warning labels on cigarette packs



While interest in graphic warning labels (GWLs) has grown around the world, their effect on smoking rates is not well-established. Most studies have focused on the impact of graphic warning labels on quit intentions rather than actual behavioral changes. Although reported intent may be correlated with actual behavior, in the context of addictive behavior such as smoking, the strength of the correlation is unclear. Recent efforts by the U.S. Food and Drug Administration (FDA) to implement graphic warning labels in the United States were blocked by the courts for violation of First Amendment rights. The battle continues to generate debate over the impact on actual smoking behavior. A lawsuit filed in October 2016 against the FDA argues that, for its not issuing yet another rule to require graphic warning labels, the FDA is in violation of the Family Smoking Prevention and Tobacco Control Act's mandate.⁸

A prominent 2014 study in the journal *Tobacco Control* has attempted to measure the effect of GWLs on smoking rates; it uses a difference-in-difference

7. See Appendix Table A1 for a full list of these countries ([link to appendices](#)), and see Figure 1 for an illustration of the growth in policy implementation. For examples of package images by country, including Canada, see the Tobacco Labelling Resource Centre website ([link](#)).

8. See Wilson (2011) for the cigarette package images that were selected but not implemented in the United States, and see CBS News (2011) for other images that were considered.

methodology to estimate the impact of the 2001 Canadian GWL policy on smoking rates. The authors are Jidong Huang, Frank Chaloupka, and Geoffrey Fong (Huang et al. 2014), and their results have been cited widely, often in support of implementation of GWLs in the United States (e.g., New York Times 2013; Myers 2013; Robert Wood Johnson Foundation 2017; Bach 2018). In a 2015 article also in *Tobacco Control*, a select set of authors—the economists Frank Chaloupka, Kenneth Warner, Daron Acemoglu, Jonathan Gruber, Fritz Laux, Wendy Max, Joseph Newhouse, Thomas Schelling, and Jody Sindelar—strongly endorsed the Huang et al. (2014) findings (Chaloupka et al. 2015). The Huang et al. (2014) estimates have also been cited as existing evidence of the estimated impacts of GWLs on smoking rates (e.g., Van Minh 2016; Jung 2016; Gibson et al. 2015) or as the yardstick for comparing other estimates (e.g., Chaloupka et al. 2015; Starr and Drake 2017). Other researchers have used the Huang et al. (2014) estimates as key inputs in simulation models to make predictions about similar policies in different settings (see, e.g., Levy et al. 2017; Kang 2017; Tauras et al. 2017), or to incorporate them in meta-analyses (e.g., Noar et al. 2016). Since its publication, Huang et al. (2014) has accumulated 66 Google Scholar citations.

In this paper, we elaborate three major reasons to discount the findings of Huang et al. (2014). First, the magnitude of the effect found is simply implausible, and prudence counsels skepticism. Second, after replicating the results of Huang et al. (2014), we show that by improving the functional form the estimated effect size shrinks by about half. Third, based on a battery of robustness checks we find cause to believe that there is omitted-variable bias and also that a key assumption underlying the analysis is not satisfied. We suggest that the study likely overstates the impact of GWLs.

Implausibility of effect size in Huang et al. (2014)

Immediate effects would equal the effect from the first decade of tobacco policy in the United States

The release of the first report of the Surgeon General’s Advisory Committee on Smoking and Health (“Surgeon General’s Report”) on January 11, 1964, marks the beginning of modern tobacco control efforts. In 2006, the Centers for Disease Control and Prevention (CDC) described the Surgeon General’s Report as “the first in a series of steps, still being taken more than 40 years later, to diminish the impact of tobacco use on the health of the American people” (CDC 2006).

Lawrence Jin, Don Kenkel, Feng Liu, and Hua Wang (2015) study the cumulative impact of tobacco control efforts from 1964, when the Surgeon General's Report was released, to 2010. They simulate how smoking prevalence would have evolved over time in the absence of the Surgeon General's Report and all the policy measures that followed—the “no-policy” counterfactual.” Comparing that to observed smoking prevalence at any point in time provides a measure of the cumulative impact of the post-Report anti-smoking movement, and the policies it encompasses, on smoking prevalence. Comparing the magnitude of these results to Huang et al.'s estimate of the effectiveness of the GWL policy is instructive. Huang et al. (2014) estimate a 12.1 percent to 19.6 percent immediate reduction in smoking prevalence in response to GWLs. Jin et al. (2015) estimate that in 1980, 16 years after the Surgeon General's report, observed smoking prevalence was 14.85 percent lower than the no-policy counterfactual. In other words, the cumulative effect of the first 16 years of post-Report progress was within the range of the immediate impact of GWLs as estimated by Huang et al. (2014).

The risks conveyed by GWLs are largely widely known in countries such as the United States and Canada. Furthermore, although the images are gruesome and elicit emotions of disgust, there is evidence suggesting that smokers may respond by covering up the label, keeping the pack out of sight, or transferring the cigarettes to a different container to avoid viewing the images every time they reach for a cigarette (e.g., Gibson et al. 2015; Guillaumier et al. 2014). In the study by Amy McQueen et al. (2015), 30 percent of survey respondents reported keeping the label out of sight some, most, or all of the time (while a low percentage reported using a case or cover to hide the label). Similarly, David Hammond et al. (2004) find that at least 30 percent of their study participants attempted to avoid the warnings. Rachel McCloud et al. (2017) find that avoidance behavior is associated with reported addiction to cigarettes, age, and the intensity of the reaction to the GWL.

While it may seem intuitive that shocking images should lead to a behavioral change, there is evidence that adults' smoking behavior in terms of actual quitting is not even responsive to news of their own parent's cardiovascular or cancer diagnoses, which may be smoking-related (Darden and Gilleskie 2016).⁹ Thus it is surprising to see results that predict an immediate and sustained effectiveness that is more than that achieved in the entire first decade of progress after the Surgeon General's Report, a time when knowledge of the major smoking risks was diffusing throughout the general population.

9. Margolis et al. (2014) examine the impact of health shocks on individuals' health behavior and find that smokers who undergo more invasive surgeries such as coronary artery bypass graft are more likely to improve health behavior such as by quitting smoking.

Effects would be larger than those estimated in settings where GWLs are expected to have a relatively large impact

There are different ways to measure the effectiveness of graphic warning labels. For example, the evidence from the literature unequivocally supports that pictorial cigarette pack warnings in Canada and other countries are more effective than text-only warnings as measured by cognitive and emotional reactions, increased motivation to quit smoking, and greater awareness of risks and warnings (see Fong, Hammond, and Hitchman 2009; Hammond 2011; Noar et al. 2016). These associations have been confirmed in more recent studies (e.g., Partos et al. 2013; Rousu et al. 2014; Cameron et al. 2015; Nagelhout et al. 2016; Yong et al. 2016; Brewer et al. 2016; Gravely et al. 2016; Green et al. 2016), although there is some evidence that their effectiveness declines over time (Hitchman et al. 2014). On the other hand, studies examining changes in smoking behavior that can be interpreted as causal are more limited. A recent systematic review concludes that existing evidence of pictorial warnings on behavioral change is inconclusive (Monarrez-Espino et al. 2014), while other studies discuss some of the underlying difficulties (e.g., Hammond 2011; Harris et al. 2015; Irvine 2015). As will be discussed below, these studies provide a wide range of estimates of the potential effect of GWLs on smoking behavior.

One of the earliest attempts to estimate the effect of GWLs on smoking prevalence was that of Nikolay Gospodinov and Ian Irvine (2004). Although it was one of the first published studies to use smoking rates as an outcome measure, Gospodinov and Irvine (2004) has received much less attention (only 25 Google Scholar citations as of March 3, 2018) than has Huang et al. (2014). Using Canadian Tobacco Use Monitoring Survey data to compare smoking rates in the six months before and after graphic warning labels were in effect, Gospodinov and Irvine (2004) estimate a 0.34 percentage-point reduction in smoking rates in Canada that is not statistically significant.

Using a longer time period (1998–2008, before and after implementation of the policy) and longitudinal data, Sunday Azagba and Mesbah Sharaf (2013) find that graphic warning labels in Canada decreased the odds of being a smoker and increased the odds of quit attempts.¹⁰ Azagba and Sharaf (2013) use a longitudinal panel and control for a set of individual demographic characteristics as well as tobacco prices and smoke-free policies.¹¹ However, Azagba and Sharaf's empirical

10. The estimated odds ratio cannot be directly interpreted as the marginal effect of GWLs on smoking prevalence, although it appears that Chaloupka et al. (2015, 113) may have misinterpreted Azagba and Sharaf's results as such.

11. Azagba and Sharaf (2013) control for a set of variables including gender, age, education, household income, marital status, household size, employment status, immigrant status, workplace smoking bans,

approach does not include a control group. Since the GWL policy was implemented at a national level in Canada, there remains doubt as to whether one can make causal inferences from those estimates.

One of the motivations for the Huang et al. (2014) study was to reassess FDA's analysis of the potential effect of graphic warning labels in the United States. In FDA's 2010 and 2011 prospective regulatory impact analyses of the GWLs rule—which is not currently in effect—FDA extrapolated from Canada's experience to generate a prediction of the effect of graphic warning labels if they were to be implemented in the United States (75 FR 69523; 76 FR 36627). FDA's approach compared the difference between the actual 1994–2009 smoking rates and rates predicted by pre-2001 trends in Canada and the United States, accounting for changes in cigarette excise taxes, and calculated how the average difference for 2001–2009 compared with the average difference for 1994–2000. Extrapolating from the resulting point estimate of the effect of GWLs in Canada, FDA predicted that the rule would reduce U.S. smoking rates by 0.4 percent, or 0.088 percentage points. Consistent with Gospodinov and Irvine's findings, FDA's estimate of the effect of Canadian cigarette graphic warning labels on smoking rates, based on the available evidence in 2011, was not statistically significant. By contrast, Huang et al. (2014) analyze the same aggregate-level data on Canada and United States smoking rates used in the FDA's regulatory impact analysis and report that Canadian GWLs led to a statistically significant 12.1 percent to 19.6 percent reduction in smoking rates (which would be a 2.87 to 4.68 percentage point decrease using pre-2001 US smoking rates of 23.9 percentage points as the baseline).¹² Huang et al. (2014, i7) indicate that their estimates would have translated into a reduction of 5.3 to 8.6 million smokers between 2012 and 2013 had the U.S. government implemented the GWL policy in 2012. As mentioned above, these results would suggest an immediate impact greater in magnitude than the reduction in U.S. smoking prevalence in the first ten years following the Surgeon General's Report.

A study by Jeffrey E. Harris, Ana Ines Balsa, and Patricia Triunfo (2015) provides estimates that are informative regarding the plausibility of the estimates in Huang et al. (2014). Harris et al. (2015) find that a GWL policy in Uruguay was associated with a 3 percentage-point increase in the cessation rate among pregnant women. Considering the context in which Uruguay's policy was implemented and evaluated, it is surprising how large the implied effect of Huang et al. (2014)'s estimate on cessation would be, as discussed later in this subsection. First, it is

providence of residence and cigarette prices.

12. Starr and Drake (2017) use an event study methodology to estimate abnormal stock rates of return resulting from the stock market reaction to the publication of the FDA graphic warning label rule to infer smokers' demand for cigarettes. Starr and Drake then conclude that smokers' demand for cigarettes would decrease by a similar order of magnitude as Huang et al. (2014)'s estimates.

important to note that Uruguay is an upper-middle-income country where the mean number of cigarettes smoked per day among daily smokers is about the same if not greater than the corresponding mean for smokers in the United States and Canada (ITC 2014). Second, to the extent to which pregnant women in Uruguay had little knowledge of the risks of smoking during pregnancy, one would expect to see a large reduction in smoking behavior. The estimated effects found by Harris, Balsa, and Triunfo (2015) also include the effects of other tobacco health measures known to be important to all smokers (not just pregnant women) that could not be separately estimated.

More recently, Ce Shang et al. (2017) examine cross-sectional data from surveys in 18 countries that had adopted the World Health Organization Framework Convention on Tobacco Control. Shang et al. (2017) report that GWLs were associated with 10 percent lower smoking prevalence among low-education individuals and 2.3 percent among all adults—a figure well below the Huang et al. (2014) lower bound of 12.1 percent. The Shang et al. (2017) estimates are correlations that cannot be interpreted as causal, but their study suggests that developing countries might experience greater impacts from tobacco policies than developed countries. In addition, the results of another cross-country study that compares trends before and after a GWL policy among more than 60 unnamed countries shows that smoking prevalence decreased by 0.40 percent after enacting a GWL policy, but this change is not statistically significant (Ngo et al. 2018).¹³

Studies of the adoption of plain packaging in combination with enlarged graphic health warnings in Australia are also instructive. One such study found that adopting these policies in Australia was associated with a 0.55 percentage-point reduction in smoking prevalence in a period of 34 months (AGDH 2016).¹⁴ That study does not address certain confounding factors and only compares the impact before and after the policy within Australia (i.e., there is no control group), which leaves questions regarding the causal interpretation of the results. A study by Pascal Diethelm and Timothy Farley (2015), seeking to refute that plain packaging had no effect in Australia, estimates that adopting plain packaging in combination with enlarged graphic health warnings was associated with a statistically significant 3.7 percent reduction in smoking prevalence. Diethelm and Farley do not report the exact smoking rate measured in their dataset prior to policy implementation, but by using an approximate smoking rate of 19 percent as reported above, the 3.7

13. Huang and Chaloupka are among the authors of Ngo et al. (2018), yet that article does not reconcile, or discuss the contrast of, its findings with those of Huang et al. (2014).

14. The model predicted that without the packaging changes, smoking prevalence would have been 17.77 percent rather than 17.21 percent. The average smoking prevalence 34 months preceding the policy was 19.4 percent, and the study concluded that the plain packaging accounted for about 25 percent of the 2.2 percentage-point decline between the pre- and post-policy periods.

percent reduction would imply an approximately 0.7 percentage-point reduction in smoking. The study controls for a linear time trend and certain other tobacco-control policies selected by stepwise regression, but it does not use a control in estimating the before-and-after impact, again leaving questions regarding causal interpretation of the results. Still, the size of the reduction in smoking prevalence associated with these joint policies in a high-income country such as Australia raises doubts about the large and almost immediate impact attributed to only GWLs in Huang et al. (2014).

When placed in context of the existing literature, the estimates of Huang et al. (2014) appear unrealistic. Consider the upper-bound estimate of a 4.68 percentage-point decrease and what this would mean if we apply the estimate given the current rates of initiation, cessation, and smoking in the United States. Even if smoking initiation was *eradicated* as a result of GWLs, the 4.68 percentage-point estimate would still imply that cessation rates would have to at least double the current U.S. cessation rate of about 6 percent in order to generate the predicted change in the United States population smoking rate (see CDC 2011). That would require an extraordinary change in population behavior given that cessation rates in the United States have remained relatively unchanged since 2001. According to CDC, there has been little progress made towards “increasing receipt for advice to quit, use of counseling and/or medication” to meet one of the Healthy People 2020 objectives to increase U.S. cessation rates to 8.0 percent by 2020 (Babb et al. 2017).¹⁵ Using the Current Population Survey Tobacco Use Supplement, Shu-Hong Zhu et al. (2017) report that the U.S. overall population cessation rate was about 4.5 percent throughout the period 2001–2010, and it increased to 5.6 percent in 2014–2015. These data indicate a 1.1 percentage-point increase in a period covering 15 years, or about a 0.08 percentage-point increase per year. Data show that the implication of the estimates would be surprising for Canada, too. Data show that quit rates among ‘ever’ smokers (i.e., current and former smokers) in Canada went from 52.3 percent to 59.7 percent from 2001 to 2005 and that they hovered around 60 percent until 2011; the quit rate in 2015 was 67.7 percent and increased by 13 percentage points between 2004 and 2015, suggesting an average annual increase of about 1.1 percentage points (Reid et al. 2017).

15. Studies have estimated a 15 percent successful quit rate without relapse among those who participate in nicotine replacement therapy or participation in smoking cessation programs (Warner et al. 2004; Cummings and Hyland 2005).

Comparison of GWL estimates to other policy interventions (tobacco and non-tobacco)

There are other tobacco control policies that have been implemented for which comparisons of their relative effectiveness help to better understand the implications of the purported effects estimated in Huang et al. (2014). Kevin Callison and Robert Kaestner (2014) show that tobacco taxes, which would have direct financial effects on smokers, have little impact on cigarette consumption, and they estimate that a doubling of all tobacco taxes would be required to obtain a 5 percent reduction in smoking.¹⁶ On the other hand, Christopher Carpenter and Philip J. Cook's study (2008) suggests that a one-dollar increase in the tax per pack of cigarettes would reduce smoking participation by about 9–20 percent (or 2.7–5.9 percentage points). Carpenter and Cook acknowledge the limitations of the self-reported nature of the data they utilize and indicate that “[p]erhaps the best interpretation of our results is that they reflect a reduced form of direct and indirect influences on youth decisions” (2008, 23). There are also mixed results with respect to the impact of indoor smoking bans. Thomas Carton et al. (2016) show that indoor smoking bans can reduce smoking between 2.35 percent and 3.29 percent, while Carpenter, Sabina Postolek, and Casey Warman (2011) find that public smoking bans have no impact on smoking behavior.

A comparison to non-tobacco policies governing the provision of health information about consumable products shows how large the Huang et al. (2014) estimate is in this broader context. In a study that included 20 countries over a period of 26 years, the authors conclude that a ban on alcohol advertising led to a 5–8 percent reduction in alcohol consumption (Saffer and Dave 2002). The results of these studies differed from previous studies that found little to no effect of alcohol advertising on its consumption. Another health policy is that of posting the calorie count of foods to inform consumers of the caloric content. Bryan Bollinger and colleagues (2011) found that calorie posting has been associated with a 6 percent reduction in average calories purchased per transaction. On the other hand, another study found little evidence of the labeling on calorie purchases (Elbel et al. 2009).

The results of Bollinger et al. (2011) are consistent with a prominent study by Pauline Ippolito and Alan Mathios (1990), who examined the impact of advertising and labeling campaigns that for the first time cited the National Cancer Institute's statements on the link between fiber and cancer. The first such campaign was launched in 1984 by the Kellogg Company and was followed by other companies

16. This size effect may in part be due to reduced responsiveness to taxes by youth (Hansen, Sabia, and Rees 2015).

in the industry. The setting in which that study occurred was important because during that period of time only government and general nutrition sources provided health information to the public. In the mid-1970s, research had already suggested a link between the consumption of fiber and a lower incidence of colon cancer. Although the information about the link had been known, this campaign represented the first time consumers received this information from industry. In other words, this case represents a similar setting in which knowledge about the risk of consuming a certain product is known already but is first made available to the consumer in a new and more visible format. In this case, Ippolito and Mathios (1990) estimated that health warnings led to a 5 percent increase in consumption of high-fiber cereal, at the expense of low-fiber cereal.

Most of the studies discussed here (tobacco and non-tobacco control) provide a range of policy-effect estimates that goes from none to 10 percent, with most estimates being around 5 percent. Thus, the results from these studies suggest that the estimates in Huang et al. (2014) are implausibly large and call for further investigation. In the next section we present additional reasons for doubt, focusing on the empirical methods and data used in Huang et al. (2014).

Replication and assessment of Huang et al. (2014) model and results

Data and methods

We begin by replicating the results of Huang et al. (2014) using their preferred specification and data provided in their appendix. Our analysis also uses data from other sources to conduct sensitivity and robustness checks. The data provided in the Huang et al. (2014) appendix include information on smoking rates, as well as measures of cigarette prices for Canada and the United States. Data on Canada's smoking rates for the population 15 years and older in the years 1994 to 2009 come from multiple sources, including the General Social Survey, Survey on Smoking in Canada, National Population Health Survey, and Canadian Tobacco Use Monitoring Survey. The data on the smoking rates for the U.S. population ages 18 years and older for years 1994 to 2009 come from the National Health Interview Survey (NHIS).

The analysis draws data on cigarette prices and excise taxes for Canada and the United States from several sources. U.S. cigarette price information is compiled by the Bureau of Labor Statistics and is adjusted by the overall consumer price index. Canada's cigarette price index is drawn from the Canadian monthly

consumer price index components for cigarettes and is also adjusted using Canada's overall consumer price index. The Canadian price index is constructed as the average over the months covered by the Canadian smoking surveys while the U.S. price index is defined as the average over the month specific to the NHIS surveys. We also use data on actual prices paid by smokers, as constructed from self-reported prices in the International Tobacco Control Policy Evaluation Project (ITC) surveys. As described by Huang et al., the tax and price measures were normalized and indexed to 1 in November 2002, and the U.S. tax and price variables were normalized to a Canadian scale using the exchange rate between the U.S. dollar and the Canadian dollar. We use data on vehicle car accidents made available from Transport Canada ([link](#)), and the U.S. Department of Transportation ([link](#)) to conduct a battery of robustness checks; these data were not utilized in Huang et al. (2014).

Finally, we note that although more recent data are available, the period of analysis ends in 2009 not only to be able to replicate Huang et al. (2014)'s study but to minimize two confounding factors. First, the Family Smoking Prevention and Tobacco Control Act was signed into law on June 22, 2009, granting FDA the authority to regulate the manufacture, distribution, and marketing of tobacco products. Second, in 2012 Canada completed implementation of a new policy ([link](#)) that required all cigarette and little cigar packages to have health warnings covering 75 percent of the front and 75 percent of the back of the package.

The difference-in-difference method is captured by equation (1) below.

$$\ln(\text{SmokingRate}) = \alpha + \beta_1 \text{Canada} + \beta_2 \text{PostGWL} + \beta_3 \text{Canada} * \text{PostGWL} + \beta_4 (\ln \text{Excise Tax or ln Price Index}) + \beta_5 \ln(\text{MonthlyTrend}) + \text{error} \quad (1)$$

The outcome variable of interest is the natural log of the annual smoking rate in each country. In equation (1), a binary variable, *Canada*, is equal to 1 if the observation is from Canada and equal to 0 if the observation is from the U.S. The coefficient on this variable is interpreted as the difference in the average smoking rates in Canada relative to the U.S. The binary variable *PostGWL* is equal to 1 if the smoking rate is for the post-2001 (post-treatment) period and equal to 0 otherwise. The coefficient of this variable is interpreted as the difference in average smoking rates before and after the treatment. The interaction between *Canada* and *PostGWL* takes on a value of 1 if the observation comes from Canada in the post-treatment period and 0 otherwise. The coefficient of *Canada*PostGWL*, $\hat{\beta}_3$, is interpreted as the difference in average smoking rates in Canada before and after policy adoption minus the difference in average smoking rates in the U.S. before and after policy adoption (the difference-in-difference estimator, henceforth referred to as the

GWL policy effect). In other words, assuming that the required conditions hold, $\hat{\beta}_3$ represents the estimated impact of the GWL policy on smoking rates.

Following Huang et al. (2014)'s approach, the variable *lnExcise Tax* or *lnPrice Index* is defined as the natural log of the tax or the natural log of a price index, and is included to control for prices. Huang et al. use two separate price indices, one that adjusts for prices captured in surveys conducted by the ITC (*lnPrice Index w/ ITC*) and one that does not (*lnPrice Index w/o ITC*). Thus there are three separate variations of the price or excise tax variable: *lnExcise Tax*, *lnPriceIndex w/ ITC*, and *lnPriceIndex w/o ITC*.¹⁷ Equation (1) also includes the variable *ln(MonthlyTrend)*, which is the natural log of the monthly trend variable. In other specifications, equation (1) includes country-specific linear trends, *Canada*MonthlyTrend*.

Following Huang et al. (2014), equation (1) is estimated using ordinary least squares and with non-robust standard errors. In total, there are six specifications (referred to as Models 1–6) that vary in the inclusion of a measure of cigarette price, excise tax, a monthly trend, and country-specific linear trends. In the sections that follow, we describe our results in replicating Huang et al. (2014)'s estimates, as well as our results from conducting a battery of tests and robustness checks to shed further light on the sensitivity and validity of the results.

Replication of Huang et al. (2014) estimates

We replicate Huang et al. (2014)'s estimates within a small margin of error. The results are presented in panel A of Table 1 for Models 1 through 6. In panel B of Table 1 we also show the results using robust standard errors to correct for heteroskedasticity. In subsequent corrections to Huang et al. (2014)'s estimation procedure, we present robust standard errors but note that our findings are not sensitive to this adjustment. Among the six specifications, Models 4 to 6 are the preferred specifications as those control for country-specific trends (later we discuss the importance of these variables). Consistent with Huang et al. (2014)'s results for Models 4 to 6, the estimated coefficients in panel A and panel B of Table 1 are statistically significant and suggest a reduction of 12.1 percent to 19.6 percent (or a 2.87 to 4.68 percentage point decrease using pre-2001 U.S. smoking rates of 23.9 percent as the baseline)

17. In the "Assessment of...model specification and assumptions" subsection below, we discuss some of the issues that may arise when using an endogenous variable such as price as a control variable.

TABLE 1. Replication of Huang et al. (2014)'s estimates

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Panel A. Huang et al.'s specification: logged trends, non-robust standard errors						
CA*GWL	-0.145*** (0.037)	-0.163*** (0.042)	-0.180*** (0.045)	-0.128* (0.064)	-0.181** (0.072)	-0.219*** (0.075)
R-squared	0.942	0.921	0.909	0.942	0.921	0.911
Panel B. Huang et al.'s specification: logged trends, robust standard errors						
CA*GWL	-0.145*** (0.036)	-0.163*** (0.040)	-0.180*** (0.041)	-0.128** (0.056)	-0.181** (0.075)	-0.219*** (0.070)
R-squared	0.942	0.921	0.909	0.942	0.921	0.911
Panel C. Huang et al.'s specification: non-logged monthly trends, robust standard errors						
CA*GWL	-0.203*** (0.028)	-0.202*** (0.028)	-0.202** (0.029)	-0.091** (0.035)	-0.099** (0.037)	-0.098** (0.038)
R-squared	0.964	0.964	0.964	0.976	0.975	0.975
<i>Controls:</i>						
Ln(Monthly Trend) or Monthly Trend	Yes	Yes	Yes	Yes	Yes	Yes
Ln(Index Excise Tax)	Yes			Yes		
Ln(Price Index w/o ITC Price)		Yes			Yes	
Ln(Price Index w ITC Price)			Yes			Yes
Canada & Trend Interaction				Yes	Yes	Yes
Observations	29	29	29	29	29	29
<i>Notes:</i> Unless specified, robust standard errors are reported in parentheses. *p<0.10, **p<0.05, ***p<0.01. The list of Controls applies to panels A to C. Panels A and B control for logged monthly trends. Panel C controls for non-logged monthly trends.						

Assessment of Huang et al. (2014) model specification and assumptions

Difference-in-difference estimates only produce valid estimates under certain assumptions, i.e., parallel trends, correct model specification, and no omitted variables. If any of these assumptions does not hold, the estimated effect is biased. In some cases, the validity of certain assumptions cannot be directly verified because the assumptions are made about unobservable data. But in general one can test the plausibility of the causal identification assumptions. In this subsection, we discuss the results from a battery of tests to assess the validity and robustness of the estimates.

Functional form of the time trends and potential serial correlation

An important assumption in statistical inference is that the model is correctly specified. Misspecification of the functional form can result in biased estimates, and in this subsection we discuss how these estimates are sensitive to the choice of

certain model specifications. Log transformation is common when there is a non-linear relationship between two variables, when there is an underlying theory that suggests a log form is needed (e.g., estimating elasticities), or when it is necessary to change the scale of a variable because its distribution is skewed. Upon review of the data, we do not find evidence to support use of the log monthly trend as specified in equation (1). We therefore re-estimate the specifications presented by Huang et al. (2014) replacing the logged monthly trends with the linear (non-logged) monthly trends, and we present the results in panel C of Table 1. The results show that the magnitude of the coefficient of interest in Models 4 to 6, which include the country-specific trends, decreases significantly compared to panels A and B, while the magnitude for Models 1 to 3 increases compared to panels A and B. The estimated coefficients for Models 4 to 6 suggest a reduction of 9.5 to 10.3 percent (or a 2.27 to 2.46 percentage-point decrease using pre-2001 U.S. smoking rates of 23.9 percent as the baseline). These estimates are about half of what Huang et al. (2014) estimate, but are still suspect as will be discussed later in this subsection.

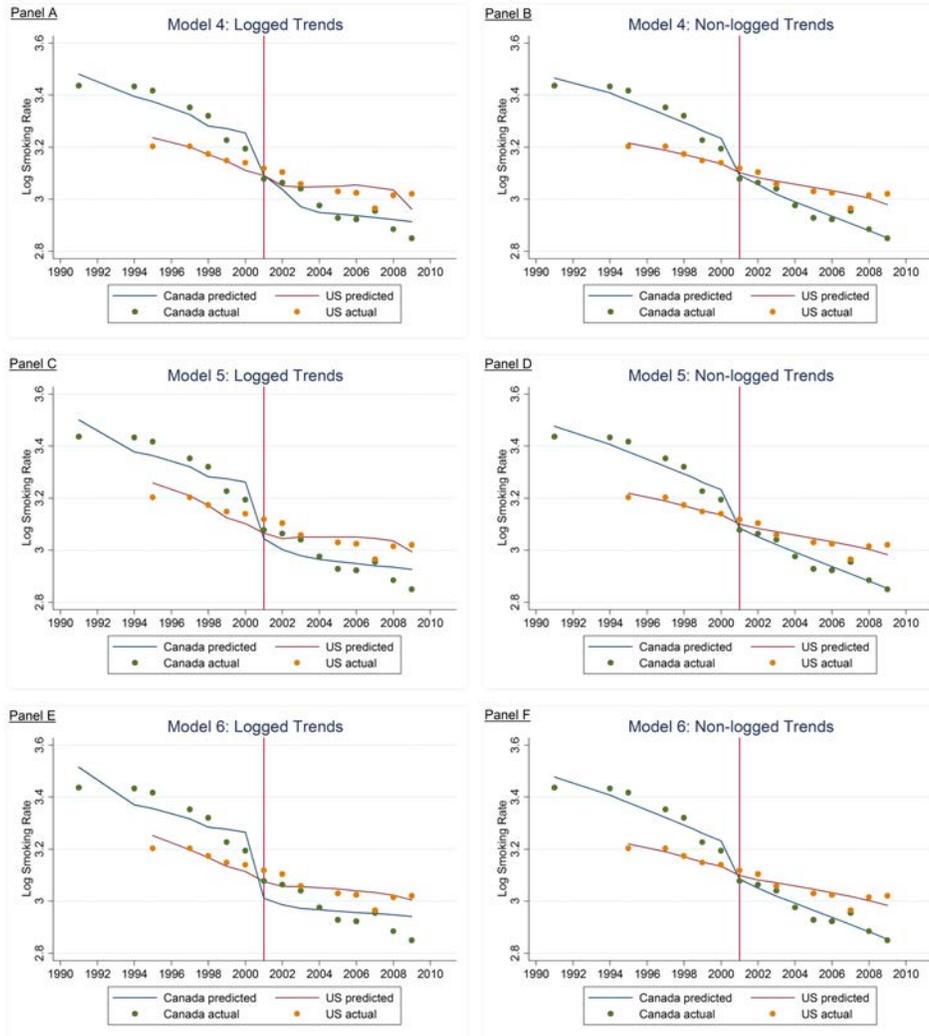
We provide further evidence against the use of logged trends with the following two exercises. First, we illustrate how misspecification of the trend variable affects the estimates by comparing a graph of the actual and the predicted outcomes using the logged trends with an analogous graph produced using non-logged trends for Models 4 to 6 (Figure 2, panels A to F). Second, we plot the residuals (the difference between the predicted smoking rate and the actual smoking rate) from the regression specifying logged trends and compare those to the residuals obtained from the corresponding regression using non-logged trends (Appendix Figure A1 panels A to F; appendices are available [here](#)).

Figure 2 shows that specifying a log functional form for the time trend variable causes the regression to fit a flatter trend with a much larger break at the time of policy implementation, compared with the non-logged functional form. Comparing the predicted trends for the U.S. and Canada in Model 6, we see that the functional form of the unlogged time trend (panel F) results in a smaller difference compared to the predicted values where logged trends are used (panel E). As a result, using logged time trends produces a much larger estimate of the GWL policy's effect. These results illustrate the sensitivity of the estimates to the functional form of the time trend included as an independent variable.

Similarly, Appendix Figure A1 panels A to F show that the residuals from the logged time trend models exhibit a downward trend, suggesting the presence of serial correlation—a violation of a fundamental assumption required for correct interpretation of the estimated parameters. Furthermore, the downward trend in the residuals is steeper for Canada than for the U.S. Note that there should be no pattern in the residuals if there is no serial correlation. In other words, no serial correlation means that the error in a given time period should not influence the

error in subsequent time periods. Therefore, the residuals should appear to be more randomly distributed and their mean should be zero.

Figure 2. Predicted versus actual smoking rates: Models 4–6



In a separate analysis presented in Appendix Table A2, we estimated Models 1 to 6 controlling for year fixed effects to account for annual trends that may be common to both countries. Adding year fixed effects reduces the estimated coefficients in magnitude further. However, we exercise caution in interpreting these results given that the sample includes only 29 observations and adding year fixed effects reduces the degrees of freedom significantly. Thus, although the results from Appendix Table A2 suggest that year fixed effects matter, for the rest

of the analysis we focus on the model without year fixed effects. Taken together, the results show that controlling for monthly trends in levels reduces the serial correlation; however, concerns about the unbiasedness of the estimates remain given that the errors in some of the specifications do not seem to be randomly distributed and symmetric around zero, and that there are other factors that we may not be controlling for that may explain both smoking rates and policy adoption tendencies.

We remain concerned about the estimated effects even after controlling for linear rather than logged trends. The next subsections discuss the additional data and empirical challenges in more detail.

Price is endogenous, and the coefficients go against established empirical evidence of price effects

Huang et al. (2014) highlight the importance of controlling for cigarette prices or cigarette excise taxes. Whether it is more appropriate to use the price of cigarettes or the excise tax on cigarettes as an explanatory variable is an example of one of the oldest problems in econometrics: the non-independence of market price. Because price changes with shifts in demand caused by other market factors, one cannot treat price as independent when estimating changes in quantity demanded. In other words, the same market forces could be causing both the change in price and the change in quantity. According to Jonathan Gruber and Michael Frakes (2006), excise taxes are preferable to using cigarette prices because cigarette prices may be linked to changes in cigarette demand. If changes in cigarette demand are driven by factors that also influence the dependent variable, it becomes difficult to untangle whether the results are caused by the change in price or by factors that caused the change in demand. That is, the price of cigarettes can change the amount that people smoke, but changes in people's preferences for the amount they want to smoke can also cause changes in prices; thus, standard statistical and inferential analysis that does not use an appropriate methodological approach cannot distinguish which is the cause and which the effect and in turn leads to biased estimators. Moreover, generally the estimator is inconsistent, which means that the problem does not go away by increasing the sample.

If the changes in the non-tax portion of prices in the U.S. and Canada followed roughly the same pattern post-2001 as they did pre-2001, and if the relationship between smoking status and cigarette prices was also relatively constant between the two time periods, then smoking rate trends would adequately control for the effect of non-tax price changes on smoking rates. In Appendix Figure A2 (panels A to C), we observe that the trends for each of the (normalized) price measures examined in Huang et al. (2004) are similar and that excise taxes

exhibit more variation in the period 2003–2009. We would then expect the coefficient on the GWL policy to vary little when either of the price measures is used.

The results of Models 4 to 6 (those including the country-specific trends) in panel C of Table 1 confirm our prediction that the magnitude of the coefficient remains relatively similar in Models 5 and 6 where the price index is used (-0.099 and -0.098 , respectively), and smaller in Model 4 (-0.091) where excise taxes are used. The results of Models 4 to 6 show that controlling for the price indices rather than excise taxes does not drive substantial variation in the magnitude of the estimated coefficient as Huang et al. assert. Moreover, the coefficients of the price indices range from -0.03 to -0.02 and none of them is statistically significant (these results are not shown but are available upon request).¹⁸ We are not saying that the results are evidence that prices are not important in this context. On the contrary, prices are important, and the lack of statistical significance on the coefficient of price measures is concerning because it contradicts empirical evidence of substantial price elasticities in both the U.S. and Canada (see, e.g., Gruber et al. 2003; Adda and Cornaglia 2006; 2010).¹⁹ There are econometric techniques to address endogeneity of price (for example, finding an instrumental variable). Addressing this issue empirically is beyond the scope of this paper, which adds to the uncertainty regarding the unbiasedness of the estimates.

Parallel trends and pre-existing trends

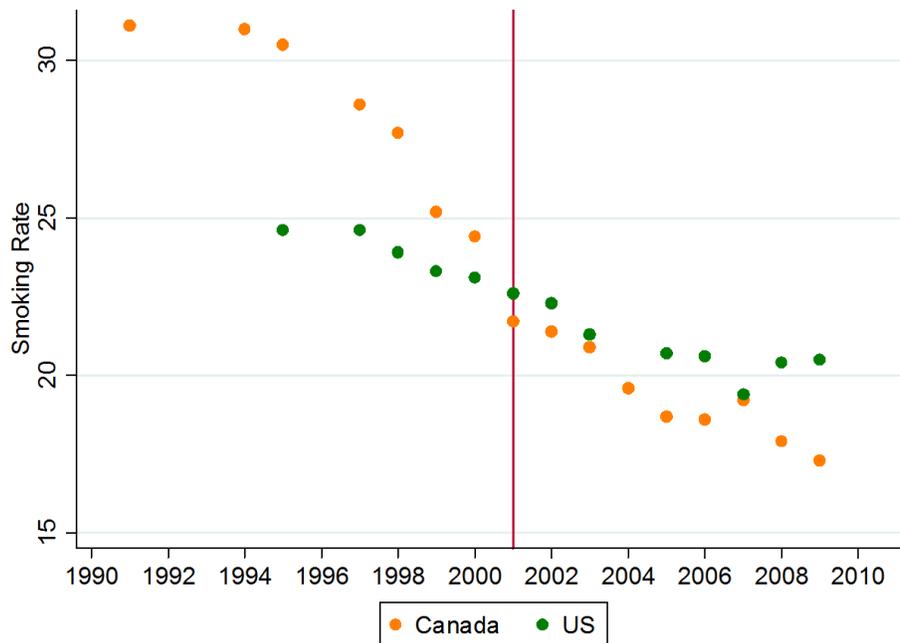
Difference-in-difference estimation is a popular quasi-experimental approach for policy evaluation because of its simplicity and its potential to control for secular trends that may confound simple pre-post comparisons. The idea behind this method is that the unobserved counterfactual—what would have happened to the outcome in the treatment group in the absence of the policy—would follow the observed evolution of the outcome for the control group. Differences in the evolution of the outcome between the treated and control group after the intervention period would be attributed to the effect of the policy intervention. But one of the critical assumptions, known as the parallel trends assumption, is that trends in the outcome variable are similar for the control and treatment groups before the policy intervention and that the only other differences between the two groups are constant over time.

18. Huang et al. (2004) report negative coefficients but not all of them are statistically significant for Models 4 to 6.

19. Irvine (2015) raises the same concern about how the results of Huang et al. (2014) would suggest that prices would have no effect, in contrast to well-established findings.

While smoking rates can be different in absolute terms between the two countries, the parallel trends assumption requires that their paths mimic each other in such a way that the difference between the two paths remains constant over the pre-treatment period. As seen in Figure 3, smoking rates in Canada and the U.S. prior to 2001 were on different trajectories, suggesting that the parallel trends assumption does not hold. Specifically, the smoking rate was decreasing faster in Canada than in the U.S. from 1995 to 2001.²⁰ If the parallel trends assumption were valid, the difference in smoking rates would remain constant in the pre-treatment period.²¹

Figure 3. Smoking rates in Canada and the United States, 1991–2009



Another point that is salient from Figure 3 is that smoking rates are going down in both the U.S. and Canada even before the GWL policy was implemented in 2001. When applying difference-in-difference estimation methods,²² it is standard practice to account for pre-existing trends that differ across countries by

20. This remains true even when examining the logged smoking rates; see Appendix Figure A3.

21. Notably, Models 1 to 3, which restrict the trend to be the same for Canada and the U.S., contradict the evidence showing that the trends differ.

22. It is customary for difference-in-differences analyses to address pre-existing trends to convince the reader that the estimated impact is due to the intervention rather than to pre-existing trends (see, e.g., Moser and Voena 2012; Gaynor et al. 2013; Bleakley 2010; Oreopoulos 2006; Carpenter 2006).

including country-specific time trends in the estimating equation.²³ Failing to account for differences in the evolution of the outcome variable in the pre-intervention period would result in biased estimates of the impact of the treatment. In other words, the differential decrease in the dependent variable may be falsely attributed to the policy of interest because one would have observed a decrease even without the policy. One can see the nature of the biasedness in the results presented in Table 1 where the estimates are sensitive to the inclusion of country-specific trends.

While controlling for country-specific trends reduces the bias in the estimates presented in panel C of Table 1, there are still concerns based on the results of an event history analysis (we refer to it as a dynamic specification and discuss it in more detail in Appendix B). In the event history analysis, we estimated a variation of equation (1) that includes lags ($t + x$ years) and leads ($t - x$ years) of the GWL policy as independent variables. If the GWL policy caused a reduction in smoking rates, then we should not expect to see an impact of the policy on lead years, otherwise, it suggests that the parallel trend assumption is violated. The results of Appendix B indicate that the policy affected smoking rates in the years preceding the actual policy implementation.

These results provide evidence that the results are biased, given that the parallel trends assumption is violated, and should not be interpreted as causal.

Estimating the effect of GWLs on an outcome that should not be affected

A widely used falsification test to examine the validity of difference-in-difference estimates is the use of an outcome that should not be affected by the policy—a placebo (Gertler et al. 2011, 99–101). The idea behind this approach is that by examining an outcome that is unrelated to GWLs, one should find zero impact of the policy on the placebo outcome. In this section, we use the number of fatal car accidents in both countries as the unrelated outcome and estimate Models 1 to 6 using unlogged trends in panel A of Table 2). Above, we discussed that in this analysis, the more appropriate specifications are those in Models 4 to 6. The associated coefficients of interest are negative but not statistically significant in Models 4 to 6; the estimated coefficients range between -0.062 and -0.053 . This result of statistical insignificance is what one would expect. Appendix Table A3 panel A shows the results of this specification using the logged monthly trends, and

23. See, for example, Abouk and Adams (2013), Leive and Stratmann (2015), Fortson (2009), Di Tella et al. (2003), and Ruhm (1998), all of which discuss the danger in not controlling for geographic differences in their respective analyses that use a difference-in-differences approach.

the results show a negative and statistically significant estimate for Models 5 and 6. These results yield further evidence that the logged functional form yields biased results.

TABLE 2. Robustness checks

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Panel A. Falsification test of Huang et al.'s specification: outcome is fatal car accidents						
CA*GWL	-0.103** (0.041)	-0.104** (0.039)	-0.104*** (0.037)	-0.062 (0.063)	-0.053 (0.064)	-0.054 (0.065)
Panel B. Outcome is smoking rate, treatment is 1999–2000, control is 1995–1998						
CA*GWL	-0.093** (0.026)	-0.059 (0.052)	-0.059 (0.052)	-0.058 (0.025)	-0.079 (0.038)	-0.079 (0.038)
Panel C. Outcome is smoking rate, treatment is 1996–2000, control is 1991–1995						
CA*GWL	-0.093** (0.026)	-0.101** (0.033)	-0.101** (0.033)	0.031 (0.044)	0.056 (0.046)	0.056 (0.046)
<i>Controls:</i>						
Monthly Trend	Yes	Yes	Yes	Yes	Yes	Yes
Ln(Index Excise Tax)	Yes			Yes		
Ln(Price Index w/o ITC Price)		Yes			Yes	
Ln(Price Index w ITC Price)			Yes			Yes
Canada & Trend Interaction				Yes	Yes	Yes
<i>Sources:</i> Canadian Council of Motor Transport Administrators (2011); National Highway Traffic Safety Administration (1995; 2017). <i>Notes:</i> Robust standard errors in parentheses. *p<0.10, **p<0.05, ***p<0.01. Outcome is log of annual number of fatal car accidents.						

Testing the effect of GWLs when the policy had not been implemented

Another commonly used method to test the validity of the estimated results is to perform the difference-in-difference estimation restricting the analysis to a period in which the program has not been implemented and therefore the policy should have no effect on the outcome—a fake policy (Gertler et al. 2011, 99–101). We conduct this exercise by simply examining the pre-intervention period and redefining the pre-treatment and post-treatment periods in several ways. First, we select only the period 1995–2000 and define the pre-treatment period as 1995–1998 and the post-treatment period as 1999–2000 (panel B of Table 2). Since the Canadian GWL policy was enacted in 2001, we would expect to find no effect of the policy on smoking rates in this timeframe. Panel B of Table 2 shows that the coefficient on the GWL policy is negative and statistically insignificant for Models 4 to 6, which control for country-specific linear trends.

Next, in panel C of Table 2, we present results from limiting the analysis period to 1991–2000 and defining the pre-treatment period as 1991–1995 and post-treatment as 1996–2000. The estimated coefficients are positive albeit not statistically significant for Models 4 to 6. While a negative coefficient in the pre-

treatment period may lead to an upward bias, a positive coefficient on the placebo policy may lead to bias in the opposite direction. Irrespective of the sign, however, the idea behind this test is that we should find a zero or statistically insignificant effect of the policy when it had not been implemented. The corresponding results of panels B and C when the trend is logged in Appendix Table A3 show the coefficients of interest being negative and statistically significant in panel B but positive and marginally statistically significant in panel C.

In one set of specifications presented in Appendix Table A3, the estimated coefficients suggest that during the fake-policy treatment, smoking rates decreased, but in Table 2 (using unlogged trends), the results suggest there is no association. These results again cast doubts on the unbiasedness of the specifications that use logged trends.

Discussion and conclusion

In this study, we discuss reasons to be skeptical about the large estimates found in Huang et al. (2014), and we show that their difference-in-difference model, as implemented, leads to biased (and potentially inconsistent) estimates. That casts doubt on whether their approach produces sufficient statistical evidence to identify the causal effects of Canadian graphic warning labels on smoking rates. In addition, we show that the estimated effects do not withstand a battery of robustness checks.

While we are able to improve upon Huang et al. (2014)'s methodology, we highlight the following limitations. First, we find that given the available data and the results in Table 2 and Appendix B, we are still unable to address the important assumptions that the parallel trends assumption is valid and that there is no omitted variable bias or endogeneity bias. Omitted variable bias may be present when the specification does not control for other factors that are correlated with the tobacco control policy and smoking behavior that we are evaluating in this paper. Examples of these factors, some of which are discussed by Huang et al. (2014), include smoke-free air policies, marketing restrictions, anti-smoking media campaigns, consumer attitudes towards smoking, and health shocks. Huang et al. (2014) acknowledge that "the strength and implementation of these other policies in the USA were as strong as, if not stronger than, those in Canada during the post-2000 period" (2014, i11). This difference in policy implementation could also violate the assumptions of the difference-in-difference methodology and result in biased estimates.

Similarly, the data do not permit us to examine participation in the black market (e.g., cigarette packs smuggled from the U.S. into Canada that do not

contain graphic warnings), which has been expanding since 2002 (Physicians for a Smoke-Free Canada 2010). Furthermore, the methodology does not address the issue of endogeneity of prices and simply assumes that prices are independent of smoking rates. All of these factors suggest that there are other omitted time-varying factors that confound the effect of the Canadian GWL policy as modeled. Besides what we have examined here, Irvine (2015) suggests that the data patterns in the period of analysis utilized in Huang et al. (2014)'s paper may be due to differences in survey implementation. Thus, Huang et al. (2014)'s analysis and the revised estimates we present in this study are insufficient to support a definitive conclusion about causal effects or the magnitude of the policy effect.

Our commentary and assessment should not be interpreted as saying that GWL policies have no impact. What the results indicate is that we cannot answer with certainty what the magnitude of the impact is, if any, given the analytical framework and the data utilized in this setting. Estimates could be improved with additional data, better-specified instruments for problematic data, or alternative methods that would address the known limitations. Lacking confidence that the estimated effects of the GWL policy are unbiased, we caution users of these estimates to avoid portraying results that may be spurious correlations as causal effects.

Appendices, data, and code

Appendix A and Appendix B are available for download from the *Econ Journal Watch* website ([link](#)), as are data and code used in our analysis ([link](#)).

References

- Abouk, Rahi, and Scott Adams.** 2013. Texting Bans and Fatal Accidents on Roadways: Do They Work? Or Do Drivers Just React to Announcements of Bans? *American Economic Journal: Applied Economics* 5(2): 179–199.
- Adda, Jérôme, and Francesca Cornaglia.** 2006. Taxes, Cigarette Consumption, and Smoking Intensity. *American Economic Review* 96(4): 1013–1028.
- Adda, Jérôme, and Francesca Cornaglia.** 2010. The Effect of Bans and Taxes on Passive Smoking. *American Economic Journal: Applied Economics* 2(1): 1–32.
- Australian Government Department of Health (AGDH).** 2016. Post-Implementation Review: Tobacco Plain Packaging 2016. Department of the Prime Minister and Cabinet (Canberra). [Link](#)
- Azagba, Sunday, and Mesbah F. Sharaf.** 2013. The Effect of Graphic Cigarette Warning Labels on Smoking Behavior: Evidence from the Canadian Experience. *Nicotine &*

Tobacco Research 15(3): 708–717.

- Babb, Stephen, Ann Malarcher, Gillian Schauer, Kat Asman, and Ahmed Jamal.** 2017. Quitting Smoking Among Adults—United States, 2000–2015. *Morbidity and Mortality Weekly Report* 65(52): 1457–1464. [Link](#)
- Bach, Laura.** 2018. Tobacco Health Warnings: Evidence of Effectiveness. January 18. Campaign for Tobacco-Free Kids (Washington, D.C.). [Link](#)
- Bleakley, Hoyt.** 2010. Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure. *American Economic Journal: Applied Economics* 2(2): 1–45.
- Bollinger, Bryan, Phillip Leslie, and Alan Sorensen.** 2011. Calorie Posting in Chain Restaurants. *American Economic Journal: Economic Policy* 3(1): 91–128.
- Brewer, Noel T., Marissa G. Hall, Seth M. Noar, Humberto Parada, Al Stein-Seroussi, Laura E. Bach, Sean Hanley, and Kurt M. Ribisl.** 2016. Effect of Pictorial Cigarette Pack Warnings on Changes in Smoking Behavior: A Randomized Clinical Trial. *JAMA Internal Medicine* 176(7): 905–912. [Link](#)
- Callison, Kevin, and Robert Kaestner.** 2014. Do Higher Tobacco Taxes Reduce Adult Smoking? New Evidence of the Effect of Recent Cigarette Tax Increases on Adult Smoking. *Economic Inquiry* 52(1): 155–172.
- Cameron, Linda D., Jessica K. Pepper, and Noel T. Brewer.** 2015. Responses of Young Adults to Graphic Warning Labels for Cigarettes. *Tobacco Control* 24: e14–e22.
- Canadian Cancer Society.** 2016. *Cigarette Package Health Warnings: International Status Report*, 5th ed. Canadian Cancer Society (Toronto). [Link](#)
- Canadian Council of Motor Transport Administrators (CCMTA).** 2011. Canadian Motor Vehicle Traffic Collision Statistics: 2009. Transport Canada (Ottawa). [Link](#)
- Carpenter, Christopher.** 2006. Did Ontario’s Zero Tolerance & Graduated Licensing Law Reduce Youth Drunk Driving? *Journal of Policy Analysis and Management* 25(1): 183–195.
- Carpenter, Christopher, and Philip J. Cook.** 2008. Cigarette Taxes and Youth Smoking: New Evidence from National, State, and Local Youth Risk Behavior Surveys. *Journal of Health Economics* 27(2): 287–299.
- Carpenter, Christopher, Sabina Postolek, and Casey Warman.** 2011. Public-Place Smoking Laws and Exposure to Environmental Tobacco Smoke (ETS). *American Economic Journal: Economic Policy* 3(3): 35–61.
- Carton, Thomas W., Michael Darden, John Levendis, Sang H. Lee, and Iben Ricket.** 2016. Comprehensive Indoor Smoking Bans and Smoking Prevalence: Evidence from the BRFSS. *American Journal of Health Economics* 2(4): 535–556.
- CBS News.** 2011. 27 Cigarette Warning Labels Nixed by the FDA. CBSNews.com, June 21. [Link](#)
- Centers for Disease Control (CDC).** 2006. History of the Surgeon General’s Reports on Smoking and Health. Centers for Disease Control (Atlanta). [Link](#)
- Centers for Disease Control (CDC).** 2011. Quitting Smoking Among Adults—United States, 2001–2010. *Morbidity and Mortality Weekly Report* 60(44): 1513–1519. [Link](#)
- Chaloupka, Frank J., Kenneth E. Warner, Daron Acemoglu, Jonathan Gruber, Fritz Laux, Wendy Max, Joseph Newhouse, Thomas Schelling, and Jody Sindelar.** 2015. An Evaluation of the FDA’s Analysis of the Costs and Benefits of the Graphic Warning Label Regulation. *Tobacco Control* 24(2): 112–119. [Link](#)

- Cummings, K. Michael, and Andrew Hyland.** 2005. Impact of Nicotine Replacement Therapy on Smoking Behavior. *Annual Review of Public Health* 26: 583–599. [Link](#)
- Darden, Michael, and Donna Gilleskie.** 2016. The Effects of Parental Health Shocks on Adult Offspring Smoking Behavior and Self-Assessed Health. *Health Economics* 25(8): 939–954.
- Diethelm, Pascal A., and Timothy M. Farley.** 2015. Refuting Tobacco-Industry Funded Research: Empirical Data Shows a Decline in Smoking Prevalence Following the Introduction of Plain Packaging in Australia. *Tobacco Prevention & Cessation* (European Publishing, Heraklion, Greece) 1(Nov.): 6. [Link](#)
- Di Tella, Rafael, Robert J. MacCulloch, and Andrew J. Oswald.** 2003. The Macroeconomics of Happiness. *Review of Economics and Statistics* 85(4): 809–827.
- Elbel, Brian, Rogan Kersh, Victoria L. Brescoll, and L. Beth Dixon.** 2009. Calorie Labeling and Food Choices: A First Look at the Effects on Low-Income People in New York City. *Health Affairs* 28(6): w1110–w1121.
- Food and Drug Administration (FDA).** 2010. Required Warnings for Cigarette Packages and Advertisements; Proposed Rule. 21 CFR 1141. *Federal Register* 75(218): 69523–69565. [Link](#)
- Food and Drug Administration (FDA).** 2011. Required Warnings for Cigarette Packages and Advertisements; Final Rule. 21 CFR 1141. *Federal Register* 76(120): 36627–36777. [Link](#)
- Fong, Geoffrey T., David Hammond, and Sara C. Hitchman.** 2009. The Impact of Pictures on the Effectiveness of Tobacco Warnings. *Bulletin of the World Health Organization* 87(8): 640–643. [Link](#)
- Fortson, Jane G.** 2009. HIV/AIDS and Fertility. *American Economic Journal: Applied Economics* 1(3): 170–194.
- Gaynor, Martin, Rodrigo Moreno-Serra, and Carol Propper.** 2013. Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service. *American Economic Journal: Economic Policy* 5(4): 134–166.
- Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch.** 2011. *Impact Evaluation in Practice*. Washington, D.C.: World Bank. [Link](#)
- Gibson, Laura, Emily Brennan, Ani Momjian, Dina Shapiro-Luft, Holli Seitz, and Joseph N. Cappella.** 2015. Assessing the Consequences of Implementing Graphic Warning Labels on Cigarette Packs for Tobacco-Related Health Disparities. *Nicotine & Tobacco Research* 17(8): 898–907.
- Gospodinov, Nikolay, and Ian J. Irvine.** 2004. Global Health Warnings on Tobacco Packaging: Evidence from the Canadian Experiment. *Topics in Economic Analysis & Policy* 4(1): 1–23.
- Gravelly, Shannon, Geoffrey T. Fong, Pete Driezen, Mary McNally, James F. Thrasher, Mary E. Thompson, Marcelo Boado, Eduardo Bianco, Ron Borland, and David Hammond.** 2016. The Impact of the 2009/2010 Enhancement of Cigarette Health Warning Labels in Uruguay: Longitudinal Findings from the International Tobacco Control (ITC) Uruguay Survey. *Tobacco Control* 25(1): 89–95. [Link](#)
- Green, Adam E., Darren Mays, Emily B. Falk, Donna Vallone, Natalie Gallagher,**

- Amanda Richardson, Kenneth P. Tercyak, David B. Abrams, and Raymond S. Niaura.** 2016. Young Adult Smokers' Neural Response to Graphic Cigarette Warning Labels. *Addictive Behaviors Reports* 3: 28–32. [Link](#)
- Gruber, Jonathan, Anindya Sen, and Mark Stabile.** 2003. Estimating Price Elasticities When There Is Smuggling: The Sensitivity of Smoking to Price in Canada. *Journal of Health Economics* 22(5): 821–842.
- Gruber, Jonathan, and Michael Frakes.** 2006. Does Falling Smoking Lead to Rising Obesity? *Journal of Health Economics* 25(2): 183–197.
- Guillaumier, Ashleigh, Billie Bonevski, and Christine Paul.** 2014. Tobacco Health Warning Messages on Plain Cigarette Packs and in Television Campaigns: A Qualitative Study with Australian Socioeconomically Disadvantaged Smokers. *Health Education Research* 30(1): 57–66.
- Hammond, David.** 2011. Health Warning Messages on Tobacco Products: A Review. *Tobacco Control* 20(5): 327–337. [Link](#)
- Hammond, David, Geoffrey T. Fong, Paul W. McDonald, K. Stephen Brown, and Roy Cameron.** 2004. Graphic Canadian Cigarette Warning Labels and Adverse Outcomes: Evidence from Canadian Smokers. *American Journal of Public Health* 94(8): 1442–1445.
- Hansen, Benjamin, Joseph J. Sabia, and Daniel I. Rees.** 2015. Cigarette Taxes and Youth Smoking: Updated Estimates Using YRBS Data. *NBER Working Paper* 21311. June. National Bureau of Economic Research (Cambridge, Mass.). [Link](#)
- Harris, Jeffrey E., Ana Inés Balsa, and Patricia Triunfo.** 2015. Tobacco Control Campaign in Uruguay: Impact on Smoking Cessation During Pregnancy and Birth Weight. *Journal of Health Economics* 42: 186–196.
- Hitchman, Sara C., Pete Driezen, Christine Logel, David Hammond, and Geoffrey T. Fong.** 2014. Changes in Effectiveness of Cigarette Health Warnings Over Time in Canada and the United States, 2002–2011. *Nicotine & Tobacco Research* 16(5): 536–543.
- Huang, Jidong, Frank J. Chaloupka, and Geoffrey T. Fong.** 2014. Cigarette Graphic Warning Labels and Smoking Prevalence in Canada: a Critical Examination and Reformulation of the FDA Regulatory Impact Analysis. *Tobacco Control* 23(supp.1): i7–i12. [Link](#)
- International Tobacco Control Project (ITC).** 2014. *ITC Uruguay National Report: Findings from the Wave 1 to 4 Surveys (2006–2012)*. August. University of Waterloo (Waterloo, Canada); Centro de Investigación para la Epidemia del Tabaquismo and Universidad de la República (Montevideo, Uruguay). [Link](#)
- Irvine, Ian.** 2015. Graphic Warning Labels on Cigarette Packaging in Canada: A Targeted Commentary on Our Limited State of Knowledge. *Concordia University Department of Economics Working Paper* 15002. June. Concordia University (Montreal).
- Ippolito, Pauline M., and Alan D. Mathios.** 1990. Information, Advertising and Health Choices: A Study of the Cereal Market. *RAND Journal of Economics* 21(3): 459–480.
- 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): 154–186.
- Jung, Minsoo.** 2016. Implications of Graphic Cigarette Warning Labels on Smoking

- Behavior: An International Perspective. *Journal of Cancer Prevention* (Korean Society of Cancer Prevention, Suwon, South Korea) 21(1): 21–25. [Link](#)
- Kang, Eunjeong.** 2017. Assessing Health Impacts of Pictorial Health Warning Labels on Cigarette Packs in Korea using DYNAMO-HIA. *Journal of Preventive Medicine & Public Health* (Korean Society for Preventive Medicine, Seoul) 50(4): 251–261. [Link](#)
- Leive, Adam, and Thomas Stratmann.** 2015. Do National Cancer Screening Guidelines Reduce Mortality? *Journal of Population Economics* 28(4): 1075–1095.
- Levy, David T., Darren Mays, Zhe Yuan, David Hammond, and James F. Thrasher.** 2017. Public Health Benefits from Pictorial Health Warnings on US Cigarette Packs: A SimSmoke Simulation. *Tobacco Control* 26(6): 649–655.
- Margolis, Jesse, Jason Hockenberry, Michael Grossman, and Shin-Yi Chou.** 2014. Moral Hazard and Less Invasive Medical Treatment for Coronary Artery Disease: The Case of Cigarette Smoking. *NBER Working Paper* 20373. August. National Bureau of Economic Research (Cambridge, Mass.). [Link](#)
- McCloud, Rachel F., Cassandra Okechukwu, Glorian Sorensen, and K. Viswanath.** 2017. Cigarette Graphic Health Warning Labels and Information Avoidance Among Individuals from Low Socioeconomic Position in the U.S. *Cancer Causes & Control* 28(4): 351–360.
- McQueen, Amy, Matthew W. Kreuter, Sonia Boyum, Vetta S. Thompson, Charlene A. Caburnay, Erika A. Waters, Kimberly A. Kaphingst, Suchitra Rath, and Qiang Fu.** 2015. Reactions to FDA-Proposed Graphic Warning Labels Affixed to U.S. Smokers' Cigarette Packs. *Nicotine & Tobacco Research* 17(7): 784–795.
- Monárrez-Espino, Joel, Bojing Liu, Felix Greiner, Sven Bremberg, and Rosaria Galanti.** 2014. Systematic Review of the Effect of Pictorial Warnings on Cigarette Packages in Smoking Behavior. *American Journal of Public Health* 104(10): e11–e30.
- Moser, Petra, and Alessandra Voena.** 2012. Compulsory Licensing: Evidence from the Trading with the Enemy Act. *American Economic Review* 102(1): 396–427.
- Myers, Matthew L.** 2013. New Study: FDA Vastly Underestimated How Much Graphic Warnings Would Reduce Smoking in U.S. [press release]. November 25. Campaign for Tobacco-Free Kids (Washington, D.C.). [Link](#)
- Nagelhout, Gera E., Marc C. Willemsen, Hein de Vries, Ute Mons, Sara C. Hitchman, Anton E. Kunst, Romain Guignard, Mohammad Siahpush, Hua-Hie Yong, Bas van den Putte, Geoffrey T. Fong, and James F. Thrasher.** 2016. Educational Differences in the Impact of Pictorial Cigarette Warning Labels on Smokers: Findings from the International Tobacco Control (ITC) Europe Surveys. *Tobacco Control* 25(3): 325–332.
- National Highway Traffic Safety Administration (NHTSA).** 1995. *Traffic Safety Facts 1994: A Compilation of Motor Vehicle Crash Data from the Fatal Accident Reporting System and the General Estimates System*. Washington, D.C.: U.S. Department of Transportation.
- National Highway Traffic Safety Administration (NHTSA).** 2017. Fatality Analysis Reporting System (FARS) Encyclopedia. July 7. National Highway Traffic Safety Administration (Washington, D.C.). [Link](#)
- New York Times.** 2013. The Power of Cigarette Warning Labels [editorial]. *New York Times*, November 28. [Link](#)

- Ngo, Anh, Kai-Wen Cheng, Ce Shang, Jidong Huang, and Frank J. Chaloupka.** 2018. Global Evidence on the Association Between Cigarette Graphic Warning Labels and Cigarette Smoking Prevalence and Consumption. *International Journal of Environmental Research and Public Health* 15(3): 421. [Link](#)
- Noar, Seth M., Diane B. Francis, Christy Bridges, Jennah M. Sontag, Kurt M. Ribisl, and Noel T. Brewer.** 2016. The Impact of Strengthening Cigarette Pack Warnings: Systematic Review of Longitudinal Observational Studies. *Social Science & Medicine* 164: 118–129.
- Oreopoulos, Philip.** 2006. Estimating Average and Local Average Treatment Effects of Education When Compulsory Schooling Laws Really Matter. *American Economic Review* 96(1): 152–175.
- Organisation for Economic Co-operation and Development (OECD).** 2012. *Health at a Glance: Europe 2012*. Paris: OECD Publishing. [Link](#)
- Partos, Timea Reka, Ron Borland, Hua-H Yong, James Thrasher, and David Hammond.** 2013. Cigarette Packet Warning Labels Can Prevent Relapse: Findings from the International Tobacco Control 4-Country Policy Evaluation Cohort Study. *Tobacco Control* 22(e1): e43–e50.
- Physicians for a Smoke-Free Canada.** 2010. Estimating the Volume of Contraband Sales of Tobacco in Canada. April. Physicians for a Smoke-Free Canada (Ottawa): [Link](#)
- Reid, Jessica L. David Hammond, Vicki L. Rynard, Cheryl L. Madill, and Robin Burkhalter.** 2017. *Tobacco Use in Canada: Patterns and Trends*, 2017 ed. Propel Centre for Population Health Impact, University of Waterloo (Waterloo, Canada). [Link](#)
- Robert Wood Johnson Foundation.** 2017. RWJF Comments on the Experimental Study on Warning Statements for Cigarette Graphic Health Warnings. May 30. Robert Wood Johnson Foundation (Princeton, N.J.). [Link](#)
- Rousu, Matthew C., Stéphan Marette, James F. Thrasher, and Jayson L. Lusk.** 2014. The Economic Value to Smokers of Graphic Warning Labels on Cigarettes: Evidence from Combining Market and Experimental Auction Data. *Journal of Economic Behavior & Organization* 108: 123–134.
- Ruhm, Christopher J.** 1998. The Economic Consequences of Parental Leave Mandates: Lessons from Europe. *Quarterly Journal of Economics* 113(1): 285–317.
- Saffer, Henry, and Dhaval Dave.** 2002. Alcohol Consumption and Alcohol Advertising Bans. *Applied Economics* 34(11): 1325–1334.
- Shang, Ce, Jidong Huang, Kai-Wen Cheng, Yanyun He, and Frank J. Chaloupka.** 2017. The Association Between Warning Label Requirements and Cigarette Prevalence by Education—Findings from the Global Adult Tobacco Survey (GATS). *International Journal of Environmental Research and Public Health* 14(1): 98. [Link](#)
- Starr, Martha A., and Keith Drake.** 2017. Graphic Warning Labels and the Demand for Cigarettes. *Tobacco Control* 26(2): 169–174.
- Surgeon General’s Advisory Committee on Smoking and Health.** 1964 (Surgeon General’s Report). *Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service*. Washington, D.C.: U.S. Government Printing Office. [Link](#)
- Tauras, John A., Richard M. Peck, Kai-Wen Cheng, and Frank J. Chaloupka.** 2017.

Graphic Warning Labels and the Cost Savings from Reduced Smoking Among Pregnant Women. *International Journal of Environmental Research and Public Health* 14(2): 164. [Link](#)

Van Minh, Hoang, Le Hong Chung, Kim Bao Giang, Duong Minh Duc, Nguyen Duc Hinh, Vu Quynh Mai, Nguyen Manh Cuong, Pham Duc Manh, Ha Anh Duc, and Jui-Chen Yang. 2016. Potential Impact of Graphic Health Warnings on Cigarette Packages in Reducing Cigarette Demand and Smoking-Related Deaths in Vietnam. *Asian Pacific Journal of Cancer Prevention* (Shahid Beheshti University of Medical Sciences, Tehran) 17(S1): 85–90. [Link](#)

Warner, Kenneth E., David Mendez, and Dean G. Smith. 2004. The Financial Implications of Coverage of Smoking Cessation Treatment by Managed Care Organizations. *Inquiry: The Journal of Health Care Organization, Provision, and Financing* 41(Spring): 57–69. [Link](#)

Wilson, Duff. 2011. U.S. Releases Graphic Images to Deter Smokers. *New York Times*, June 21. [Link](#)

Wolfers, Justin. 2006. Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. *American Economic Review* 96(5): 1802–1820.

Yong, Hua-Hie, Ron Borland, David Hammond, James F. Thrasher, K. Michael Cummings, and Geoffrey T. Fong. 2016. Smokers' Reactions to the New Larger Health Warning Labels on Plain Cigarette Packs in Australia: Findings from the ITC Australia Project. *Tobacco Control* 25(2): 181–187.

Zhu, Shu-Hong, Yue-Lin Zhuang, Shiushing Wong, Sharon E. Cummins, and Gary J. Tedeschi. 2017. E-Cigarette Use and Associated Changes in Population Smoking Cessation: Evidence from US Current Population Surveys. *BMJ* 358: j3262. [Link](#)

About the Authors



Trinidad Beleche is an Economist at FDA. She received her Ph.D. from the University of California, Riverside. Her email address is tbeleche@gmail.com.



Nellie Lew is an Assistant Director of the Economics Staff at FDA. She received her Ph.D. from the University of California, Santa Barbara.



Rosemarie Lavaty Summers is an Economist at FDA. She received her Ph.D. in Economics from the University of California, Santa Barbara.



Joseph Laron Kirby is an Economist at FDA. He received his Ph.D. from Middle Tennessee State University.

[Go to archive of Comments section](#)
[Go to May 2018 issue](#)



Discuss this article at Journaltalk:
<https://journaltalk.net/articles/5964>