ECON298: Econometrics Replication Group Project

Group 2 – Paper 7

Do Workplace Smoking Bans Reduce Smoking?

by LONG TRAN, AARYAN KAPHLEY, AND LUKA AGULASHVILI

1. Purpose:

The purpose of this report is to replicate the research findings of the publication by Evans et al. titled, "Do Workplace Smoking Bans Reduce Smoking". This paper was published in the American Economic Review, volume 89, issue 4, pages 728-747 on September 1999.

a) Motivation:

This paper investigates the effects of workplace smoking policies which were becoming very prevalent and restrictive during the 90s. There was an increased public awareness towards the adverse effects of second hand or environmental tobacco smoke (ETS). In response to the public's growing concern and intolerance towards ETS, many firms started to voluntarily adopt workplace smoking restrictions. In 1985, only 25% of workers worked in establishments that banned smoking in work areas. By 1993, this number shot up to 70%. The public response to smoking in the workplace reached its apex during March 1994, when Occupational Health and Safety Administration proposed a complete ban in smoking over 6 million workplaces! Besides reducing exposure to ETS, workplace smoking restrictions greatly affect smoking behavior because of the number of hours that workers are subject to these restrictions. Previous literature produced conflicting results regarding this topic because they generally lacked control

groups and investigated restrictions in only one location over relatively short periods of time. These studies also tended to share a common methodological problem of self-selection bias. This paper advances the literature by addressing these prevalent issues.

b) Data and Methods:

The primary data for this paper comes from the National Health Interview Survey (NHIS) conducted in 1991 and 1993. The supplement section from both years contain detailed questions about workplace smoking policies. After deleting observations with missing values, 9,704 observations from 1991 and 8,386 samples from 1993 were collected for a total of 18,090 observations of workers who work indoors and are not self-employed. The paper uses three measures of smoking. The first is dummy variable that equals 1 if a worker is a current smoker. The second is a continuous variable which average smoking consumption in cigarettes per day. The third is a composite variable that equals daily consumption for all smokers and 0 for nonsmokers. The model for current smoker variable is estimated by a Probit model and the cigarettes per day equations by OLS. The study uses three different approaches to address the potential for Omitted Variable Bias (OVB). First, it controls for additional covariates that may signal underlying healthiness of the individual or firm since smokers may be attracted to firms without workplace smoking bans. Second, it attempts to generate results that are consistent with a causal interpretation by showing that the impact of workplace smoking ban is greatest for workers with the longest work weeks. Third, it uses two-stage least squares(2SLS) to control for unobserved differences between workers and non-workers.

c) Results and Conclusion:

The sample mean for the three variables are presented in a table along with mean values of these variables for workers who are and are not exposed to work area smoking bans, as well as the differences in means and finally, the multivariate models that control for a number of covariates. Some of the covariates included are age, family size, income, region, education, ethnicity, metropolitan area and marital status, among others. These statistics show that workers are 8.2% less likely to smoke in firms with a smoking ban and cigarette consumption per day decreases by more than 3 cigarettes a day. In the multivariate model, the results indicate that workplace smoking bans reduce smoking by more than 5.7% and smoking intensity drops by 2.5 cigarettes a day, which is more than 10% of daily consumption. The paper also replicates the basic results using data from the special smoking supplements to the Central Population Survey (CPS), to make sure that the estimates are not a product of using the NHIS data. The results from the two surveys are nearly identical. The paper also addresses the problem of OVB. First, the results are practically unchanged while controlling for additional covariates. Second, workplace smoking bans have the largest impact on workers who have longer work weeks. Finally, there appears to be no correlation between the health habits of workers and establishment size. Thus, the paper concludes that the drop in smoking among workers is a result of workplace smoking bans.

2. Replication Data and Methods:

The data used to replicate the results of the original paper is a cross-sectional data set with observations on 10,000 indoor workers, which is a subset of an 18,090-observation data set collected as part of the National Health Interview Survey in 1991 and again in 1993 with different respondents. The survey collected information on whether individuals' workplace had enacted a smoking ban, whether or not the individuals were smokers and other characteristics like race and gender. The original dataset was provided by Professor William Evans of the University of Maryland and was used in his study along with Matthew Farrelly and Edward Montgomery.

Variables	**Replication Data**	**Original Data**
Mean age	38.7	37.8
Percent male	43.6%	44.1%
Percent black	7.7%	10.5%
Percent hispanic	11.3%	6.7%
Percent work are smoking ban	61.0%	67.0%
Percent smoker	24.2%	24.2%
Percent hsdrop	9.1%	
Percent hsgrad	32.7%	
Percent colsome	28.0%	
Percent colgrad	19.7%	

TABLE 1. ORIGINAL AND REPLICA SAMPLE CHARACTERISTICS

Summary statistics are very similar but there are slight deviations in variables black, hispanic and work area smoking-ban. This change is due to the nature of the replication data,

which only consists of 10,000 observations randomly selected from the original 18,090-observation data set.

To determine the effect of different variables on the smoking status we will conduct Ordinary Least Squares as well as Probit regression models with smoking status as the dependent variable. Independent variables include work area smoking ban, age in years and age squared, high school dropout, high school graduate, attended some college, black, Hispanic, female. All variables above except age are binary variables and the educational indicators refer to the highest level attained and therefore are mutually exclusive.

The replicated model differs from the model used in the original paper, which used additional covariates like industry and occupation as well as log income, indicator variables for the region and type of metropolitan area, marital status, real cigarette tax, and a year effect. Analysis in the original paper also examined the number of cigarettes smoked per day by workers as a dependent variable in addition to smoking status. However, while it included them as controlled covariates in the regression models, the smoking ban study did not provide the coefficients as well as other statistics of the four educational variables — high-school drop-out, high-school graduate, person having some college, and college graduate — which were included in our data set. The workplace smoking restriction is also categorized into indoor public area smoking ban and work area smoking ban, which is just simplified to a workplace smoking ban for our analysis.

3. Replication Results:

In our replication project, we use a simple indicator that equals to 1 if a worker is a current smoker and equals to 0 if a worker is not a current smoker. To offer a brief understanding of the relationship between smoking ban and smoking behavior, or in other words, smoking behaviors of workers before and after smoking ban is implemented, we report the difference in mean values of the smoking indicators for workers whose firms adopted smoking ban or not in the bar graph below:

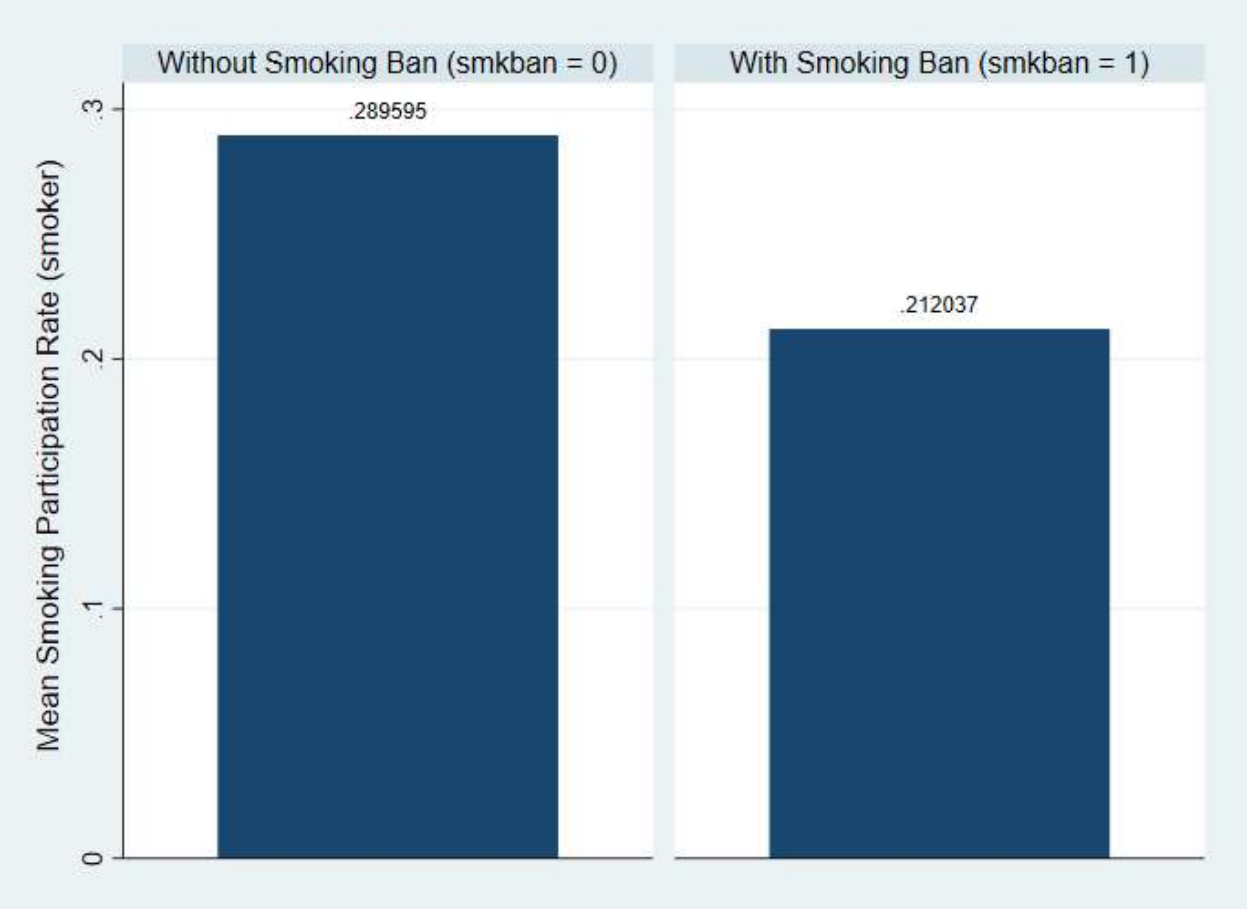

FIGURE 1. SMOKING PARTICIPATION RATES, WITH AND WITHOUT WORKPLACE SMOKING BAN

The graphs indicate that a worker in a firm with smoking ban is likely to have 7.8 percentage points less behavior to smoke than that of one working in a firm without ban. This is near consistent with the result from the original paper, which reports the difference of 8 percentage points between behavior of the two circumstances; taking into account that the replicated data set is just a subset of the original data set, this is understandable. However, as correlation does not imply causation, this result does not suggest that there is a possible causality between smoking bans and smoking behaviors. Considering that there might be presence of omitted variables bias that affects smoking participation rate besides smoking bans, it is necessary to estimate multivariable models with other independent variables that might be significant.

To achieve this, in the replication, LPM and Probit regressions will be performed to report the changes in the likelihood of smoking when working in a firm with smoking ban compared to working in a firm without one. For each of the modelling approaches, the first models (1 and 4) would not be influenced by other covariates, leaving only *smkban* involved in; in the second models (2 and 5), some new implementations are age, its square, and whether the worker is black, Hispanic, or female; lastly, in the last models (3 and 6), we added the rest of the covariates: four educational variables. Because due to the nature of the coefficients on covariates in the Probit model, we made columns 7, 8, and 9, which shows the normalized estimates that measure the “marginal effect”, or the change in the probability that an individual smokes, given the adoption of a workplace smoking ban, from a change in an independent variable when its value is at its mean, from columns 4, 5, and 6.

VARIABLES	(1) LPM smoker	(2) LPM smoker	(3) LPM smoker	(4) Probit smoker	(5) Probit smoker	(6) Probit smoker	(7) Probit smoker (margins)	(8) Probit smoker (margins)	(9) Probit smoker (margins)
Workplace with smoking ban	-0.0776***	-0.0764***	-0.0472***	-0.245***	-0.241***	-0.159***	-0.0762	-0.0767	-0.0503
	(0.0090)	(0.0090)	(0.0090)	(0.0279)	(0.0281)	(0.0291)	(0.0087)	(0.0090)	(0.0093)
Age in years		0.0040**	0.0097***		0.0144**	0.0345***		-0.0008	-0.0005
		(0.0019)	(0.0019)		(0.0066)	(0.0069)		(0.0004)	(0.0004)
Age in years (squared)		-0.00006***	-0.0001***		-0.0002***	-0.0005***			
		(0.00002)	(0.00002)		(0.00008)	(0.00008)			
Black		-0.00001	-0.0276*		0.0014	-0.0843		0.0004	-0.0268
		(0.0161)	(0.0161)		(0.0519)	(0.0535)		(0.0165)	(0.0169)
Hispanic		-0.0367***	-0.105***		-0.120***	-0.338***		-0.0381	-0.1074
		(0.0132)	(0.0140)		(0.0447)	(0.0494)		(0.0142)	(0.0157)
Female		-0.0187**	-0.0333***		-0.0593**	-0.112***		-0.0189	-0.0355
		(0.0087)	(0.0086)		(0.0278)	(0.0288)		(0.0089)	(0.0091)
High School Dropout			0.323***			1.142***			0.3625
			(0.0195)			(0.0730)			(0.0234)
High School Graduate			0.233***			0.883***			0.2802
			(0.0126)			(0.0604)			(0.0193)
Some College			0.164***			0.677***			0.215
			(0.0126)			(0.0614)			(0.0196)
College Graduate			0.0448***			0.235***			0.0745
			(0.0120)			(0.0654)			(0.0207)
Constant	0.290***	0.250***	-0.0141	-0.555***	-0.708***	-1.735***			
	(0.0073)	(0.0391)	(0.0414)	(0.0212)	(0.132)	(0.152)			
Observations	10,000	10,000	10,000	10,000	10,000	10,000			
R-squared	0.008	0.011	0.057						

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

TABLE 2. LPM AND PROBIT ESTIMATES OF CURRENT SMOKER REGRESSIONS

From the models, we can see that enforcing a smoking ban in the workplace negatively affects a worker's probability of smoking and this pattern continues with different methods and controlled variables. The coefficients are somewhat similar for both the Linear Probability Model and Probit model at -0.0472 and -0.0503, when controlled for all variables, respectively. In addition, all six coefficients on smoking ban variable in different controlled models are all

statistically significant at the 1% level, strongly suggesting a correlation between having workplace smoking ban and the reduction of smoking participation rate. According to the third model of Probit regression result, when all external factors in the data are accounted for, workplace smoking ban reduce the probability of smoking by 5.03 percentage points. To put this result in perspective, the result is 5.7 percentage points in the original paper. There exists a small difference between the two, but the near similarity of them suggests a causal relationship between workplace smoking ban and the probability of smoking. Inspecting further into the influences of more independent covariates adopted in each model on the coefficient of interest, we can see that age, age-squared, black, Hispanic, and female do not have much effect on workplace smoking ban coefficient (-0.0776 to -0.0764 in LPM model and -0.0762 to -0.0767 in Probit model), while the fantastic four educational variables have strong impact (-0.0764 to -0.0472 in LPM model and -0.0767 to -0.0503 in Probit model).

While the findings from both the original data and the replicated data indicate a strong correlation between the two covariates, there still remain the possibility of internal validity problems, such as Omitted Variables Bias, that undermine the legitimacy of the outcome of the study. Self-selection bias is one possible issue that pose threat to the causal relationship between workplace smoking ban and smoking participation. Smokers, due to their personal preferences, might choose to work in workplaces that do not implement the ban on smoking, while non-smokers, caring for their health, would choose the opposite. That being said, the estimates from our model might overstating the effect of smoking ban. The original study included more additional controlled variables like health habits and use establishment size as an instrument to address these problems. Nevertheless, the results show no correlation

between health habits of workers and establishment size, and in the end, most of the evidence we provided is consistent with the hypothesis that simple cross-sectional estimates represented earlier successfully reflect the causal impact of bans on smoking.

4. Potential Extensions:

The study found correlation between workplace smoking bans and reduction in smoking rate in workers, but while single-equation estimate shows a difference of means between percentage of smokers with and without smoking bans to be 8 percentage points, this number shrinks down to 5.7 percent when controlled for covariates. The study explores omitted variable bias by using same covariates to model workers health instead. The Probit estimates showed that people working in places with smoking bans were more likely to have never smoked a cigarette, less likely to salt their food, more likely to wear seatbelt and more likely to have smoke detectors at home. These results suggest that what was assumed to be correlation between smoking bans and reduction in smoking rate, could be a direct result of workplace selection, simply people with more focus on health and fitness tend to choose workplaces that ban smoking. But when controlled for these variables the results were not significantly different. But this does not rule out the possibility of omitted variable bias. Study might be missing confounding variables, like difficulty of work and stress. For an extension the study should also include vaporizers and electronic cigarettes as well, as they can be still used secretly inside of buildings and places where smoking is banned and examine weather smoking restrictions effect electronic cigarette users. Study should also be extended geographically, exploring smoking restrictions in other countries where different cultural factors might play a large role.