

## Estimating smoking-attributable mortality in the United States<sup>\*</sup>

Andrew Fenelon<sup>1†</sup>

Samuel H. Preston<sup>2</sup>

<sup>1</sup>Population Studies Center and Department of Sociology, University of Pennsylvania, 239 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104, [afenelon@sas.upenn.edu](mailto:afenelon@sas.upenn.edu), (805)-886-6586

<sup>2</sup>Population Studies Center, University of Pennsylvania, 239 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104, [spreston@sas.upenn.edu](mailto:spreston@sas.upenn.edu), (215)-746-5396

<sup>†</sup>To whom all correspondence should be sent

---

<sup>\*</sup> This research was supported by NICHD Training Grant HD007242, NIH fellowship 1-F31-AG-039188-01, and Grant from Social Security Administration and National Bureau of Economic Research. We are grateful to Andrew Noymer and Douglas Ewbank for comments and critiques. An earlier version of this paper was presented at the Population Association of America Annual Meeting in Dallas, TX, April 15-17, 2010.

## ABSTRACT

Tobacco is the largest single cause of premature death in the developed world. Two methods of estimating the number of deaths attributable to smoking use mortality from lung cancer as an indicator of the damage from smoking. We reestimate the coefficients of one of these, the Preston/Glei/Wilmoth model, using recent data from U.S. states. We calculate smoking attributable fractions for the 50 states and the U.S. as a whole in 2000 and 2004. We estimate that 21% of adult deaths among men and 17% among women were attributable to smoking in 2004. Across states, attributable fractions range from 11% to 30% among men and from 7% to 23% among women. Smoking related mortality also explains as much as 60% of the mortality disadvantage of Southern states. At the national level, our estimates are in close agreement with those of the Centers for Disease Control (CDC) and Preston/Glei/Wilmoth, particularly for men. But we find greater variability by state than does CDC. We suggest that our coefficients are suitable for calculating smoking-attributable mortality in contexts with relatively mature cigarette smoking epidemics.

## INTRODUCTION

Tobacco is the largest single cause of premature death in the developed world and is growing in importance throughout the developing world. At the individual level, cigarette smoking is strongly linked to lung cancer; but smoking also confers increased risk of death from other cancers, heart diseases, stroke, and chronic respiratory conditions (Doll et al. 2004). Although the CDC estimates that as many as 400,000 deaths annually in the United States are attributable to cigarette smoking, only about 30% of these deaths are caused by lung cancers (CDC 2008; Mokdad et al. 2004).

Cohort studies demonstrating the link between cigarette smoking and individual mortality track the mortality of individuals according to their smoking behavior. The prospective study of British doctors beginning in 1951 (Doll et al. 2004) and the American Cancer Society's Cancer Prevention Studies (CPS) Cohorts I and II, beginning in 1959 and 1982, respectively, provide rich data on the excess risks associated with cigarette smoking from a number of causes of death.

A second approach to estimating the amount of excess mortality attributable to smoking uses lung cancer mortality, rather than survey data, as the indicator of smoking. The most widely-known of these "indirect" approaches was designed by Peto, Lopez and colleagues (Peto et al. 1992). It uses the death rate from lung cancer as an indicator of the accumulated damage from smoking and combines that indicator with estimates of the relative risk of smokers compared with non-smokers of mortality from certain disease categories. Preston, Gleit, and Wilmoth (PGW, 2010a) recently developed another indirect method which relies on the statistical relationship between lung cancer and all other causes of death that was estimated across countries and time periods. Preston, Gleit, and Wilmoth (2010b) and Rostron (2010) use virtually the same data set but introduce small modifications of the estimation equation that

produce small changes in estimated attributable deaths for males but sizeable reductions for women above age 80. A recent report of the National Research Council (2011) relies upon the results of PGW (2010b) to conclude that international differences in smoking-attributable deaths are the principal explanation of shortfalls in life expectancy at age 50 in the United States relative to other OECD countries.

This paper has one substantive goal and one methodological goal. The substantive goal is to provide improved estimates of the extent to which smoking explains state inequality in adult life expectancy in the most recent time period. Available data indicate that states differ substantially in both all-cause mortality as well as in the prevalence of smoking among adults. The maps in Figure 1 present the patterns in 2004. Southern states exhibit considerable disadvantage with respect to mortality as well as relatively high rates of smoking in the population. Alternatively, states in the West and Northeast show low mortality and relatively low prevalence of smoking. Lung cancer mortality rates have also tended to follow a similar pattern, relatively high in the South and quite low in the West (Devesa et al. 1999). These patterns suggest that smoking may play an important role in accounting for the regional patterns of life expectancy within the United States. We intend to evaluate this linkage more precisely than has been done previously.

The methodological goal is to provide a test of the new PGW method on a new data set. To date, the three papers that use the basic PGW approach (Preston et al. 2010a; 2010b; Rostron 2010) have estimated the coefficients of the relation between all-cause mortality and lung cancer mortality on a data set pertaining to 20 or 21 countries over the period 1950-2006. If the approach is generalizable, the estimated relation should be much the same when different units of analysis are used. We reestimate the coefficients of the PGW model using data from US

states rather than international data. We use the results to calculate smoking-attributable mortality in the United States in 2004. We integrate the methodological and substantive goals in comparing our results for individual states and the U.S. as a whole to those of other methods for estimating smoking attributable-risk.

### **Methodological Approaches to Calculating Smoking Attributable Risk**

Studies calculating excess mortality due to cigarette smoking typically use an attributable-risk approach; they estimate the number of deaths that would not have occurred if smokers had experienced the same death rates as non-smokers (Peto et al. 1994). This calculation requires information about the increased risk conferred upon smokers by their behavior as well as information about the prevalence of smoking in the population. Two broad sets of methods have been developed and applied in a variety of different settings.

The first set of methods could be termed *direct* methods, because the mortality differential between smokers and non-smokers is directly observed. These studies require detailed data on the smoking behavior and mortality experience of cohort members followed over a number of years. In the United States, the most commonly cited study is the American Cancer Society Cancer Prevention Study Cohort II (CPS-II) which is composed of more than 1.2 million individuals followed from 1982 through 1988. The Centers for Disease Control and Prevention (CDC) issues regular estimates of smoking-attributable mortality in the U.S. using relative risks from CPS-II (Adhikari et al. 2009) and estimates of smoking prevalence from the National Health Interview Survey (NHIS) or from the Behavioral Risk Factor Surveillance System (BRFSS). Based on excess mortality among current smokers and former smokers relative to non-smokers, they calculate mortality attributable to cigarette smoking by applying relative

risks to current smoking status data. They find that 440,000 annual deaths can be attributed to cigarette smoking in the early 2000s, more than one-fifth of all adult mortality.

The CDC approach has been criticized on several grounds. First, the cohorts used to provide the relative mortality risk of smokers are not representative of the U.S. population (Thun et al. 1997). Second, the CDC assumes that the relative mortality level of smokers compared to non-smokers reflects only the effects of smoking, an assumption that ignores other behavioral and socioeconomic factors with which smoking may be correlated. Rogers et al. (2005) use the NHIS 1990 supplement to control for many such confounders, which decreases the estimated number of smoking-attributable deaths in the year 2000 to 338,000. Third, smoking behavior categories recorded at baseline do not reflect temporal changes in cohort smoking patterns, and many studies assume that baseline smoking status remains constant throughout the study; estimated risks will be attenuated if there are any changes during the period of observation. Finally, large-scale cohort studies require long periods of data collection and detailed demographic information that is unavailable for many relevant populations.

As noted above, Peto, Lopez, and colleagues (1992) developed another type of method that instead calculates the ‘impact’ of smoking indirectly from the lung cancer death rate in the population rather than through direct observation. Assuming that smoking behavior is the only factor which increases the risk of *lung cancer* death of smokers relative to non-smokers, they use CPS-II non-smoker lung cancer death rates to calculate age-specific ‘proportion exposed’ that reflects the prevalence of smoking-related damage. They then import relative risks for various disease categories from CPS-II and apply them to this ‘proportion exposed’ in the population of interest. In order to correct for confounding, they decrease the relative risks from causes of death

other than lung cancer by half. The method has been used to produce estimates of smoking attributable mortality for developed countries for the year 2000 (Peto et al. 2006).

PGW (2010a) developed an alternative method to Peto-Lopez which makes fewer assumptions and does not rely heavily on the generalizability of CPS-II relative risks. They develop a model estimating the statistical relationship between the lung cancer death rate and the death rate from other causes of death across developed countries between 1950 and 2006. This relation is then used to estimate the mortality impact of smoking on causes of death other than lung cancer. They use lung cancer death rates of non-smokers in CPS-II between 1982 and 1988 (Thun et al. 1997) to produce an estimate of lung cancer attributable-risk for each population of interest. The method produces results that are similar to those of Peto-Lopez while avoiding strong assumptions and a complex implementation procedure.

The key similarity of the two indirect methods, Peto-Lopez and PGW, is that the lung cancer death rate is interpreted as an indicator of the damage from smoking within a population. Lung cancer is a unique condition because it is so closely tied to one behavioral risk factor. While other causes of death are linked to smoking behavior, none is related as strongly as lung cancer. In CPS-II, smoking was responsible for more than 90% of lung cancer deaths among men and more than 70% among women (Thun et al. 1997). Since lung cancer mortality reflects current and past prevalence as well as intensity of cigarette smoking in a population, it is likely to be a more reliable measure of smoking's population-level impact than are direct cohort data derived from a single-round survey (Peto et al. 1992). The use of the lung cancer death rate in this way is further justified by evidence that differences in lung cancer mortality across place and time result almost exclusively from variation in cigarette smoking (see Preston et al. 2010b for a discussion of these issues).

## **State Variation in Mortality in the United States**

Geographic differences in mortality have been particularly longstanding within the United States (Devesa et al. 1999). States in the South region are at a clear disadvantage compared with their counterparts in other parts of the country (Figure 1). The cluster of high-mortality states in the South is striking, and in fact the 12 states with the highest death rates are geographically contiguous. According to vital statistics, these states exhibit all-cause death rates that are 30-40% higher than those of the low mortality states, which translates into around 4-5 years difference in life expectancy at birth. States that perform relatively well are slightly more dispersed; low-mortality pockets occur in the Upper Midwest (e.g. Minnesota, North Dakota), New England (Connecticut, Vermont), Mountain West (Arizona, Colorado), and Pacific (California, Hawaii). Although there have been some long-term changes in the size of state-to-state disparities in adult mortality (NCHS 1991), the general pattern of southern disadvantage has been remarkably stable over time.

The experience of the United States with respect to the smoking epidemic has been somewhat exceptional in comparison to its European counterparts. Smoking began early and remained quite heavy until relatively recently when the US experienced drastic declines in cigarette use (Forey et al. 2002). American women have shown particularly high rates of smoking compared to women in Europe, and the mortality burden of smoking is accordingly high among women in the United States (Peto et al. 2006). But large regional differences in cigarette smoking behavior and related mortality exist within the US (CDC 2009). The Southern states have the greatest numbers of smokers while states in the West and Northeast have few. Despite declining rates of cigarette smoking in the United States, many Southern states continue



to exhibit relatively high smoking prevalence (e.g., 30% in Kentucky compared with 10-17% in the West).

We merge our methodological with our substantive goal in applying the PGW model to US states. We reestimate the coefficients of the PGW model (Preston et al. 2010b) using annual mortality data at the U.S. state level between 1996 and 2004. Based on the results of this estimation, we calculate smoking-attributable mortality for the United States as a whole as well as for the fifty states and demonstrate the impact of smoking-related mortality on state-specific patterns of mortality in the U.S. Finally, we compare attributable fraction estimates for the U.S. produced by a variety of methods.

## DATA

We use vital statistics data on deaths for the fifty states annually between 1996 and 2004.<sup>1</sup> Death data are available through the Multiple Cause-of-Death (MCD) public-use micro-data files released annually by the National Center for Health Statistics (NCHS). MCD files contain demographic, geographic, and cause-of-death information about all deaths occurring in the United States. Population denominators for death rate calculations come from bridged-race files available from the NCHS.<sup>2</sup> Deaths are based on *state of residence*, rather than state of occurrence. There is little evidence that migration has a noticeable effect on geographic mortality patterns (Ezzati et al. 2008).

## METHOD

### Statistical Model

Following Preston et al. (2010a) we estimate the relationship between the age-specific lung cancer death rate and the log of the death rate from other causes of death annually between

---

<sup>1</sup> 2004 is the latest year for which geographic identifiers below the national level are available in the public-use version of the MCD files.

<sup>2</sup> Electronically from [http://www.cdc.gov/nchs/nvss/bridged\\_race.htm](http://www.cdc.gov/nchs/nvss/bridged_race.htm), accessed May, 2009.

1996 and 2004. We use negative binomial regression to predict the log the of death rate from causes other than lung cancer in five-year age groups from 50 – 54, 55 – 59, ..., 80 – 84 as a function of the death rate from lung cancer

$$\ln(M_O) = \beta_L M_L + \beta_a X_a + \beta_S X_S + \beta_t T + \beta_{aL}(M_L \times X_a) + e \quad (1)$$

where  $M_L$  and  $M_O$  are the death rate for lung cancer and other causes respectively in each state, year, and five-year age group.  $X_a$  and  $X_S$ , are dummy variables for age-group and state respectively, while  $\beta_a$  and  $\beta_S$  are their corresponding coefficients. We include a linear time trend ( $T$ ) as well as interactions between lung cancer mortality and age-group.  $e$  is a random disturbance term. In contrast to Preston et al. (2010b), we do not include interactions between lung cancer and time nor between state and time. The time period considered here is relatively short compared with that in PGW, and we do not expect substantial changes in the relationship between lung cancer and other causes over this period. We consider the impact of our choice of years on the estimated attributable fractions in our sensitivity analyses (below). We use age-specific population size as a statistical “offset” to control for exposure to mortality. We estimate separate models by sex to allow for distinct relationships between smoking and mortality for men and women. The coefficients of interest are  $\beta_L$  and  $\beta_{aL}$ , denoting the age-specific relationship between lung cancer and other causes of death ( $\beta'_L = \beta_L + \beta_{aL}$ ), which are used to calculate the attributable fraction.

Preston, Gleit, and Wilmoth have produced two sets of coefficients using the model: one which drops observations for age 85+ and one which maintains them (Preston et al. 2010a, 2010b). Data for this age group were subject to age misreporting and, as an open-ended interval, to extraneous influences resulting from differences in age distributions. These effects had produced a set of parameters that were implausible at the oldest ages. Dropping these

observations produced a smoother sequence of coefficients at older ages and reduced the fraction of deaths attributable to smoking among older women. The current estimation also drops data for ages 85+ and uses the results of PGW (2010b) for ages 50-84 in comparisons reported below.

### **Attributable Risk Calculation**

Lung cancer deaths attributable to cigarette smoking are estimated using values of lung cancer death rates among never-smokers, reported by Thun et al. (1997) from the CPS-II study between 1982 and 1988. The proportion of lung cancer deaths attributable to smoking is the ratio of smoking-related lung cancer death rate to the overall lung cancer death rate

$$A_L = \frac{M_L - M_L^*}{M_L}$$

where  $M_L^*$  is the lung cancer death rate among lifelong non-smokers, the expected death rate in the absence of smoking. While lung cancer mortality among never smokers does show some variation across populations (Thun et al. 2008), there is little evidence for long-term changes across periods (Rosenbaum, Sterling and Weinkam 1998).

Following Preston, Gleit, and Wilmoth (2010b), we calculate mortality attributable to smoking for causes of death other than lung cancer based on the relationship between lung cancer and other causes across states. The proportion of deaths from other causes attributable to smoking is found by comparing the actual number of deaths from other causes to the number that would be expected if mortality from lung cancer were set at the level for never-smokers. Given the model, this estimate is found by

$$A_O = \frac{e^{\beta'_L(M_L)} - e^{\beta'_L(M_L^*)}}{e^{\beta'_L(M_L)}}$$

where  $\beta'_L$  is the model coefficient for lung cancer including age interactions ( $\beta'_L = \beta_L + \beta_{aL}$ ).

The attributable fraction<sup>3</sup> for total mortality is a weighted average of the attributable fractions for lung cancer and other causes

$$A = \frac{A_L D_L + A_O D_O}{D}$$

where  $D_L$  and  $D_O$  are deaths from lung cancer and other causes respectively and  $D$  is total deaths.

In their application, Preston et al. (2010a) find that the estimated attributable fraction is generally robust to alternative specifications of age, time period, and interactions. We calculate standard errors for our attributable-fraction estimates by resampling within the parameter distributions.

We simulate 1,000 sets of  $\beta'_L$  coefficients, allowing them to vary based on the estimated variance-covariance matrix from the regression procedure. These simulated coefficients produce an artificial sample of attributable fractions for each state which allows us to calculate the standard error. We report 95% confidence intervals in Appendix Table A.1.

### **Variation in Mortality by U.S. State**

We estimate smoking-attributable mortality for ages 50-84 for the United States as well as the fifty states. We calculate the expected number of years lived in this age range and age-adjusted death rates both including and excluding smoking-related deaths. Age-specific death rates in the absence of smoking ( $M_{abs}$ ) include only those deaths not attributed smoking by our model

$$M_{abs} = \frac{D - D_A}{P}$$

---

<sup>3</sup> Since the lung cancer death rate is the chief input for the calculation of the attributable fraction, the correlation between the age-adjusted lung cancer death rate and the attributable fraction across states is very high (0.97 among women, 0.99 among men). The attributable fraction is a more meaningful measure of the burden of smoking than simple lung cancer mortality since it accounts for various other causes of death for which smoking is a risk factor (Preston et al 2010a).

where  $D_A$  the number of deaths attributed to smoking and  $P$  is the number of person-years of exposure. We then recalculate life tables and age-adjusted mortality for each state with smoking-related deaths removed.<sup>4</sup>

## RESULTS

Figure 2 shows the age-pattern of lung cancer mortality for men and women in the United States for ages 50-84. The death rate rises with age for both sexes, with men experiencing substantially higher death rates than women at all ages. These death rates reflect the accumulated damage from smoking for the cohorts in each respective age group in the United States.

Table 1 presents estimated coefficients from the model in Equation 1. If exponentiated, they can be interpreted as the proportional increase in the death rate of causes other than lung cancer associated with an increase in the lung cancer death rate of one per thousand, all else being equal. Coefficients are smaller at higher ages, reflecting both higher death rates overall and more varied factors influencing mortality. Lung cancer death rates for the US population in 2004 and among lifelong non-smokers from CPS-II are presented in Table 2. Given that we assume smoking to be the sole source of population variation in lung cancer death rates, the non-smoker rates in Table 2 are intended to represent conditions in which smoking were eliminated. The difference between these rates and observed lung cancer rates is used to calculate lung cancer attributable-risk. As shown in Table A.1, we estimate that smoking was responsible for 21% of deaths among men and 17% among women aged 50-84 in the U.S. in 2004. The maps in Figure 3 display estimated attributable fractions by state for females and males, respectively. Darker shades represent a greater proportion of attributable deaths. For both sexes there is substantial

---

<sup>4</sup> We elect to simply ‘remove’ smoking deaths from the life table calculation as opposed to using ‘cause-deleted’ life tables in order to preserve the simplicity of interpretation. The results do not change substantively.

geographic variation in the burden of smoking-attributable mortality (Figure 3 and Table A.1.). The highest attributable fractions among women are found in Alaska, Kentucky, and Nevada (around 22%), states notorious for relatively high rates of smoking among women (CDC 1996; Remington et al. 1989). The lowest fractions are found in Utah, New Mexico, and Hawaii. Utah has an exceptionally low mortality burden, with only 7% of deaths attributed to smoking in 2004. Among men, there is a strong concentration of smoking-related mortality in the Southern states. Kentucky, Mississippi, and Tennessee all exhibit attributable fractions close to 30%. States in the Mountain West like Utah, New Mexico, and Colorado have fractions lower than 15%.

To estimate the extent to which smoking explains variation in life expectancy across states in the United States we compare variance in state-specific age-adjusted mortality before and after removing smoking-related deaths. The proportional reduction in variance represents the fraction explained by smoking-related mortality. We find that smoking accounts for 35% of state variation in mortality among women in 2004. Among men, it is even more important, explaining 65%. The sex difference reflects greater overall importance of smoking as well as a stronger correlation with state-specific mortality experience among men. Differences in smoking patterns are evidently a huge source of variance in life expectancy among states.

To assess more specifically the role of smoking in the very high mortality in the South, we compare the mortality experience of the South relative to other regions in the presence and absence of mortality related to smoking. Table 3 reports the proportion of the Southern disadvantage that is attributable to smoking. High mortality related to smoking is an important factor in each regional comparison. For women, it explains 18-20% of the Southern disadvantage relative to the Pacific states, the Central Midwest, and the Northeast, 25% relative to the Upper Midwest, and 35% relative to the Mountain states. Smoking is responsible for 23%

of the difference between the Southern states and all states outside the South. Among men, smoking-attributable mortality is even more important. It explains 43-48% of the disadvantage relative to the Pacific states, and the Central and Upper Midwest, 50% relative to the Northeast, and 60% relative to the Mountain. Overall, we estimate that the difference in male mortality between the South and all other states would be cut in half in the absence of smoking.

### **Sensitivity Analyses**

Our data apply to very recent years, in order to produce estimates that best reflect the mortality burden of smoking in the current stage of the smoking epidemic in the United States. PGW estimations cover a much longer time period and introduce a linear trend in the coefficient relating lung cancer mortality to mortality from other causes and in country coefficients. To see whether their approach would change our results, we perform the above analyses using data for the period 1970 – 2004 and include interactions between lung cancer and year and between state and year in order to capture changes in the impact of smoking that occur over the longer period. We find that this model produces attributable fractions virtually identical to those from the original model, indicating that our estimates are not sensitive to the length of the period considered or to the treatment of trends. This specification produces attributable fractions of 0.21 for men and 0.17 for women for the U.S. as a whole in 2004, identical to the fractions produced by the model without time trends.

Additionally, to ensure that our results are not driven by state differences in racial composition, we estimate parameters of our basic model using exclusively data on the white population. The estimated attributable fractions are slightly lower than those for the total population, but attributable fractions for the white population are correlated with those for the total population at 0.95 for men and 0.99 for women.

## COMPARISON WITH ALTERNATIVE METHODS

Researchers have developed a number of methods for estimating the number of deaths in a population attributable to cigarette smoking. Table 4 shows estimated smoking-attributable fractions for the United States using six different procedures. The first row shows estimates of smoking-attributable fraction for U.S. men and women in 2000 and 2004 using the present procedure. Row 2 shows estimates obtained using coefficients found by PGW (2010b) using the same estimation model with different coefficients estimated across a sample of 21 developed countries. Their attributable risk estimates for ages 50-84 are very similar to ours, especially for men. Their estimates for women are somewhat higher (0.20 vs 0.17). Row 3 presents estimates from Rostron's (2010) modification of the PGW estimation procedure. Again, male estimates are very similar but female estimates are lower than ours. Some of the disparity is a result of the inclusion of ages 85+ in the Rostron estimates but not in ours, since he finds a low attributable risk above age 85. Row 4 shows estimates using the Peto-Lopez method reported in Peto et al. (2006) for ages 35+, which are somewhat higher than ours. Estimates from the CDC (Row 5) are slightly higher than ours for men (0.24) and lower among women (0.15). The estimates made by Rogers et al. (2005) (Row 6) using smoking-status data from the NHIS are substantially lower than our estimates for females (0.13) and quite similar to ours for males (0.21). As noted earlier, relative risks derived from baseline smoking data would be downwardly biased if status at baseline is misclassified or if changes in smoking status occurred during the seven-year follow-up period. Both the Rogers et al. (2005) and CDC estimates suffer from this limitation.<sup>5</sup>

Table 4 indicates that there is considerably more uncertainty about estimates for women than estimates for men. The male attributable fractions in the Table have a range of only 0.03,

---

<sup>5</sup> However, CDC estimates use the current prevalence of smoking to make attributable-risk estimates, which does not accurately reflect the mortality burden of smoking. However, depending on yearly changes in the prevalence of smoking, this may offset some of the downward bias from the use of baseline relative risks.



whereas the range for women is 0.08. Estimates based on smoking behavior (CDC and Rogers et al.) occupy the lower end of the range for women. The current estimates are in the middle of the range. We can also compare our estimates to state-specific estimates made by the Center for Disease Control (2009). CDC estimated smoking prevalence at the state level from the Behavioral Risk Factor Surveillance Survey (BRFSS). This data source is based on telephone surveys and has a response rate that differs by state, in part because states have control over how the BRFSS is executed (e.g., with respect to questionnaire length, whether data collection is in-house or contracted out, and sampling design) The national response rate in 2004 was 52.7% (Schneider and Lapane 2007). CDC combined these estimated prevalences with estimates of the proportion of deaths from various causes that is attributable to smoking, estimates that were drawn from deaths for 1982-88 (CDC 2009). Data used in the CDC estimates is thus somewhat dated and subject to reporting biases.

Despite considerable differences between our method and that of CDC, the geographic patterns implied by both methods are relatively consistent across states. Figure 4 shows the comparability of CDC (2009) state-specific attributable risk estimates for the period 2000-2004 and those based on our method for 2004. The correlation between the two series is a relatively high 0.81 for both males and females. However, CDC's estimates are consistently lower than ours for women and higher for men. These discrepancies may reflect the crudeness of the CDC method, specifically its effort to model the mortality impact of smoking through contemporary surveys of smoking status and its uses of dated estimates of the relative risk of smoking. The relative mortality risk of smoking depends on duration (number of years smoked), intensity (number of cigarettes per day), inhaling practices, and type of cigarette (Flanders et al. 2003). As the composition of smokers changes over time, so does the observed relative risk of death

among smokers (Thun et al. 1997). This risk rose between the American Cancer Society's Cancer Prevention Studies I and II (Thun et al. 1995) and also during the major study of British doctors (Doll et al. 2004). CDC uses relative risks estimated in the CPS-II for current and former smokers during the period 1982 – 1988. Unpublished analyses of NHIS data by Mehta and Preston (2011) indicate that the relative risk of death among smokers has continued to rise among women since 1988, which may account for an underestimate of smoking effects among women by CDC. Because of the lag between smoking behavior and mortality outcomes, smoking prevalence may say more about the burden of smoking in the future than in the current period (Peace 1985; Preston et al. 2010a).

## **METHODOLOGICAL IMPLICATIONS AND LIMITATIONS**

The PGW model, unlike the approach used by the CDC and by Peto and Lopez (1992), does not borrow relative risk estimates from prospective studies of smokers and non-smokers. It uses lung cancer mortality as an indicator of the damage from smoking and assumes that such damage can be identified in other causes of death by modeling the relation between lung cancer mortality and mortality from other causes. Parameters of that model have been estimated using international and intertemporal data in PGW (2010a, 2010b) and in Rostron (2010).

The present paper applies the PGW model to recent cross-state data in the US. Coefficients for males and females from the current estimation and from PGW (2010b) are presented in Table 1 and graphed in Figure 5. Several patterns are clearly evident:

- 1) The sets of coefficients estimated on the basis of data in the contemporary US are quite similar for men and women, suggesting that lung cancer mortality is functioning in the US as stable indicator of the incremental mortality risk, presumably associated with smoking, for other causes of death. On the other hand, female coefficients are

- much larger than male coefficients in the international data set investigated by PGW (2010b) and Rostron (2010). No explanation of this sex difference has been provided.
- 2) Male coefficients estimated using US data are remarkably similar to those estimated from the international/intertemporal data. This similarity provides an encouraging indication that the overall approach to estimating the impact of smoking is reliable for males.
  - 3) The outlier series is the set of coefficients for females estimated from the international/intertemporal data. Coefficients for this series are generally higher than those from the other three series, and substantially so at younger ages.

We suspect that the high coefficients for women in the PGW (2010b) series are a result of the recency of the smoking epidemic for women in their data set. The data set begins with observations from the early 1950's for all 21 countries. In most of these countries, few older women were smoking during that era. PGW's (2010a) estimates of attributable risk from smoking for women in 1955 are above 0.01 in only two of 21 countries. In contrast, the median value for males was already 0.07 by 1955.

The maturity of the smoking epidemic may, for example, affect the relation between lung cancer mortality and mortality from other causes of death by virtue of different lags in the relation between smoking and different causes of death. It is possible that the damage inflicted by smoking takes longer to manifest itself in mortality from lung cancer than from other causes of death. If so, this different pattern of lags would explain why the coefficient relating mortality from other causes to that from lung cancer is higher among newly-smoking countries than in the US, connoting fewer excess lung cancer deaths per excess death from other causes. Support for this possibility comes from a comparison of relative risks of death between CPS I (1959-65) and

CPS II (1982-88). The relative risks of death from lung cancer among female smokers compared to non-smokers rose dramatically from 2.7 in the former to 12.8 in the latter (Thun et al. 1995). The increases for coronary heart disease (1.4 to 1.8) and for “other smoking-related cancers” (1.8 to 2.6) were much smaller.

The US has a relatively mature smoking epidemic among both men and women (Forey et al. 2002; Pampel 2010). In contrast to the data used by PGW, our sample of geographic units is relatively homogeneous with respect to the stage of the epidemic.<sup>6</sup> Thus we do not expect any noticeable distortion resulting from state-to-state variation in the maturity of the epidemic since smoking is longstanding throughout the United States. We suggest that the US coefficients for women from the current estimation in Table 1 may be more appropriate for countries such as the US with a mature tradition of women’s smoking, whereas the PGW estimates for women may be more appropriate for relative newcomers. Indeed, our coefficients imply an attributable fraction that is slightly lower than that given by PGW. Yet the difference is relatively small because women’s coefficients become closer at ages 70-85 where deaths are heavily concentrated. And of course lung cancer deaths are treated the same way in both methods. For men, the choice between the two series is basically immaterial because they are so similar to one another.

Our analysis has several limitations. First, we use mortality data only from the most recent period, unlike PGW who investigate data from 1950-2006. Since our focus is on the burden of smoking-related mortality across US states during the current period, we chose to include only the previous 10 years. Our sensitivity analysis indicates that our estimates do not

---

<sup>6</sup> In the PGW sample, lung cancer death rates among men range from less than 3 to more than 40 per 10,000. This compares with a range between 12 and 34 per 10,000 in states in our estimation. Among women, rates in the PGW sample range from 4 to more than 140 per 100,000 compared with a range between 60 and 190 per 100,000 among states. Furthermore, the variance in lung cancer mortality is nearly 4 times as large in the PGW sample as in our sample among men and nearly twice as large among women.

change when we include data for the period 1970 – 2004, which connotes considerable robustness for the present estimates.

The second limitation is our assumption that, in the absence of smoking, individuals would have the death rate from lung cancer recorded among lifelong never smokers in the CPS-II. Although smoking is the primary factor driving differences in lung cancer mortality over time and space, we cannot be certain that never-smoker death rates are the same across states. Research by Thun and others suggests, for example, that non-smoker lung cancer rates are quite different in Asian and non-Asian populations (Thun et al. 2008). However, because smoking has been found to cause the vast majority of lung cancer deaths in heavy-smoking populations (Ezzati and Lopez 2003), slight differences in the non-smoker rates will not greatly affect our conclusions regarding the level or geographic pattern of the burden of smoking.

Third, we are unable to completely account for the consistent differences between our results and those of CDC for states. It is likely that the discrepancies primarily reflect the crudeness of CDC's procedure and the drawbacks of direct methods described previously. In the absence of ideal cohort smoking data, each method must make some assumptions about the relationship between smoking and mortality at the individual level. CDC assumes that all current smokers have the same mortality risk, and applies this risk to the observed smoking status composition of each state population. PGW uses the lung cancer death rate as an indicator of the accumulated damage from smoking in the population and assumes a constant relationship (in the form of the coefficients) between mortality from lung cancer and mortality from other causes. Fortunately, both methods identify approximately the same geographic pattern of smoking-related mortality.

## **CONCLUSION**

Despite recent declines in the prevalence of smoking in many developed countries including the United States, the mortality burden of smoking remains large among both men and women. The US has been characterized by early onset of the smoking epidemic and by relatively heavy smoking in comparison to many European countries. At the same time, data have suggested that individual states differ greatly in the prevalence of smoking as well as mortality from smoking-related cancers. One goal of this paper was to apply the recently developed PGW model to data from the United States in order to provide detailed estimates of the contribution of smoking to geographic disparities in adult mortality. We simultaneously evaluate the robustness of the PGW indirect estimation technique and identify smoking as a key factor determining regional variation in adult mortality within the United States.

The key substantive pattern we attempt to explain is the Southern mortality disadvantage relative to other regions. Among women, smoking-related mortality is responsible for 23% of the South's disadvantage relative to the rest of the country and 35% of its excess relative to the Mountain region. For men, it explains 50% of the South's excess mortality relative to the rest of the US and 60% of the disparity with the Mountain region. Given the lag in the relation between smoking and mortality, these disparities reflect both historical and contemporary state-to-state differences in smoking behavior.

Such persistent differences in the burden of smoking across states to some extent reflect local tobacco policy environments and cultures surrounding smoking. Since the mid 1990s, state tobacco control programs have been rather effective at promoting smoking cessation and preventing others from taking up the habit (Cokkinides et al. 2009; Farrelly et al. 2008). Statewide workplace smoking bans may be beneficial not just for individuals at work, but also for the acceptability of smoking in the state context (Farrelly, Evans and Sfeekas 1999). Indeed,

states with no statewide smoking ban show a higher prevalence of smoking compared with those states banning smoking from all workplaces (CDC 2005, 2010). Along with smoking bans, cigarette excise taxes may also be an important factor in determining local tobacco cultures and are likely to be key policy interventions responsible for declines in cigarette consumption (Franks et al. 2007; Pierce et al. 2010). States vary widely in the amount of per-pack tax levied on cigarettes, from less than \$0.25 to more than \$4.00, which produce large differences in the price of a pack of cigarettes (CDC 2010). However, the correlation between tax level and smoking prevalence appears to be relatively weak (CDC 2010).

Estimating mortality attributable to cigarette smoking is important for informing public health policies aimed at limiting avoidable deaths. Direct methods, such as that used by the CDC, require extensive data collection, make numerous assumptions about the impact of smoking on mortality, and are subject to a variety of potential biases. The use of lung cancer mortality as the indicator of damage from smoking bypasses many of the attendant difficulties. We have provided a set of estimates of the impact of smoking using lung cancer mortality and its empirical correlation with other causes of death. This correlation was estimated based exclusively on interstate data in the US. Results suggest that smoking is continuing to play a major role in the level of and regional variation in American mortality.

In the course of this investigation, we have estimated the parameters of a model proposed by Preston, Gleit and Wilmoth on an entirely new data set than the one that they employ. We find that the relation between lung cancer mortality and mortality from other causes of death is remarkably similar for males across 50 states of the United States to the one they identify across 21 countries. For women, however, a unit change in lung cancer mortality is associated at most ages with a smaller increment in other causes of death when estimated on data for US states than

when estimated on international data. We believe that this difference reflects a greater maturity of the smoking epidemic in the contemporary US than in the sample of countries on which the international estimates were based. Accordingly, we suggest that the coefficients estimated here are more appropriate for countries like the US where smoking has been pervasive for many decades.



## References

- Adhikari, B., J. Kahende, A. Malarcher, C. Husten, and K. Asman. 2009. "State-Specific Smoking-Attributable Mortality and Years of Potential Life Lost-United States, 2000-2004 (Reprinted from MMWR, vol 58, pg 29-33, 2009)." *Journal of the American Medical Association* 301(9):928-929.
- CDC. 1996. "State-specific prevalence of cigarette smoking--United States, 1995." *MMWR Morb Mortal Wkly Rep* 45(44):962-966.
- . 2005. "State-specific prevalence of cigarette smoking and quitting among adults--United States, 2004." *Mortality and Morbidity Weekly Report* 54(44):1124-1127.
- . 2008. "Smoking-attributable mortality, years of potential life lost, and productivity losses--United States, 2000-2004." *MMWR Morb Mortal Wkly Rep* 57(45):1226-1228.
- . 2009. "State-specific smoking-attributable mortality and years of potential life lost--United States, 2000-2004." *Morb Mortal Wkly Rep* 58(2):29-33.
- . 2010. *Tobacco Control State Highlights*. Atlanta: Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.
- Cokkinides, V., P. Bandi, C. McMahon, A. Jemal, T. Glynn, and E. Ward. 2009. "Tobacco Control in the United States-Recent Progress and Opportunities." *Ca-a Cancer Journal for Clinicians* 59(6):352-365.
- Devesa, S.S., D.J. Grauman, W.J. Blot, and J.F. Fraumeni. 1999. "Cancer surveillance series: Changing geographic patterns of lung cancer mortality in the United States, 1950 through 1994." *Journal of the National Cancer Institute* 91(12):1040-1050.
- Doll, R., R. Peto, J. Boreham, and I. Sutherland. 2004. "Mortality in relation to smoking: 50 years' observations on male British doctors." *British Medical Journal* 328(7455):1519-1528.

- Ezzati, M., A.B. Friedman, S.C. Kulkarni, and C.J.L. Murray. 2008. "The reversal of fortunes: Trends in county mortality and cross-county mortality disparities in the United States." *Plos Medicine* 5(4):557-568.
- Ezzati, M. and A.D. Lopez. 2003. "Measuring the accumulated hazards of smoking: global and regional estimates for 2000." *Tobacco Control* 12(1):79-85.
- Farrelly, M.C., W.N. Evans, and A.E.S. Sfekas. 1999. "The impact of workplace smoking bans: results from a national survey." *Tobacco Control* 8(3):272-277.
- Farrelly, M.C., T.F. Pechacek, K.Y. Thomas, and D. Nelson. 2008. "The impact of tobacco control programs on adult smoking." *American Journal of Public Health* 98(2):304-309.
- Flanders, W.D., C.A. Lally, B.P. Zhu, S.J. Henley, and M.J. Thun. 2003. "Lung cancer mortality in relation to age, duration of smoking, and daily cigarette consumption: Results from Cancer Prevention Study II." *Cancer Research* 63(19):6556-6562.
- Forey, B.A., J. Hamling, P. Lee, and N. Wald. 2002. *International smoking statistics: a collection of historical data from 30 economically developed countries, 2nd Edition*. London: Oxford University Press.
- Franks, P., A.F. Jerant, J.P. Leigh, D. Lee, A. Chiem, I. Lewis, and S. Lee. 2007. "Cigarette prices, smoking, and the poor: Implications of recent trends." *American Journal of Public Health* 97:1873-1877.
- Mehta, N. and S.H. Preston. 2011. "Continued Increases in the Relative Risk of Death from Smoking." Manuscript, University of Michigan.
- Mokdad, A.H., J.S. Marks, D.F. Stroup, and J.L. Gerberding. 2004. "Actual causes of death in the United States, 2000." *Journal of the American Medical Association* 291(10):1238-1245.

- National Research Council. 2011. "Explaining Divergent Levels of Longevity in High Income Countries." in *Panel on Divergent Trends in Longevity*, edited by E. Crimmins, S.H. Preston, and B. Cohen. Washington, D.C.: National Academy Press.
- Pampel, F. 2010. "Divergent Patterns of Smoking Across High-Income Nations." in *International Differences in Mortality at Older Ages: Dimensions and Sources*, edited by E. Crimmins, S.H. Preston, and B. Cohen. Washington, DC: National Academy Press.
- Peace, L.R. 1985. "A TIME CORRELATION BETWEEN CIGARETTE-SMOKING AND LUNG-CANCER." *Statistician* 34(4):371-381.
- Peto, R., A.D. Lopez, J. Boreham, and M. Thun. 2006. "Mortality From Smoking in Developed Countries 1950–2000, 2nd Edition."
- Peto, R., A.D. Lopez, J. Boreham, M. Thun, and C. Heath. 1992. "Mortality From Smoking in Developed Countries - Indirect Estimation From National Vital-Statistics." *Lancet* 339(8804):1268-1278.
- . 1994. *Mortality from Smoking in Developed Countries 1950-2000*. Oxford: Oxford University Press.
- Pierce, J.P., K. Messer, M.M. White, S. Kealey, and D.W. Cowling. 2010. "Forty Years of Faster Decline in Cigarette Smoking in California Explains Current Lower Lung Cancer Rates." *Cancer Epidemiology Biomarkers & Prevention* 19(11):2801-2810.
- Preston, S.H., D.A. Gleij, and J.R. Wilmoth. 2010a. "A new method for estimating smoking-attributable mortality in high-income countries." *International Journal of Epidemiology* 39(2):430-438.

- . 2010b. "Contribution of Smoking to International Differences in Life Expectancy." in *International Differences in Mortality at Older Ages: Dimensions and Sources*, edited by E. Crimmins, S.H. Preston, and B. Cohen. Washington, D.C.: National Academy Press.
- Remington, P.L., T.E. Novotny, D.F. Williamson, and R.F. Anda. 1989. "State-Specific Progress Toward the 1990 Objective for the Nation for Cigarette-Smoking Prevalence." *American Journal of Public Health* 79(10):1416-1419.
- Rogers, R.G., R.A. Hummer, P.M. Krueger, and F.C. Pampel. 2005. "Mortality attributable to cigarette smoking in the United States." *Population and Development Review* 31(2):259-+.
- Rosenbaum, W.L., T.D. Sterling, and J.J. Weinkam. 1998. "Use of multiple surveys to estimate mortality among never, current, and former smokers: Changes over a 20-year interval." *American Journal of Public Health* 88(11):1664-1668.
- Rostron, B. 2010. "A modified new method for estimating smoking-attributable mortality in high-income countries." *Demographic Research* 23:397-420.
- Schneider, K.L. and K.L. Lapane. 2007. "Adult health behaviors." *Public Health Reports* 122(4):432-432.
- Thun, M., C. Day-Lally, D.G. Myers, E.E. Calle, W.D. Flanders, and B.P. Zhu. 1997. "Trends in tobacco smoking and mortality from cigarette use in cancer prevention studies I (1959 through 1965) and II (1982 through 1988)." Pp. 305-382 in *Changes in cigarette-related disease risks and their implications for prevention and control, smoking and tobacco control. Smoking and tobacco control monograph no. 8, NIH publication no. 97-4213*, edited by D. Burns, L. Garfinkel, and J.M. Samet. Bethesda, MD.: Cancer Control and Population Sciences, National Cancer Institute, U.S. National Institutes of Health.

Thun, M.J., C.A. Daylally, E.E. Calle, W.D. Flanders, and C.W. Heath. 1995. "Excess Mortality Among Cigarette Smokers - Changes in a 20-Year Interval." *American Journal of Public Health* 85(9):1223-1230.

Thun, M.J., L.M. Hannan, L.L. Adams-Campbell, P. Boffetta, J.E. Buring, D. Feskanich, W.D. Flanders, S.H. Jee, K. Katanoda, L.N. Kolonel, I.M. Lee, T. Marugame, J.R. Palmer, E. Riboli, T. Sobue, E. Avila-Tang, L.R. Wilkens, and J.M. Samet. 2008. "Lung Cancer Occurrence in Never-Smokers: An Analysis of 13 Cohorts and 22 Cancer Registry Studies." *Plos Medicine* 5(9):1357-1371.

**Table 1: Estimated coefficients for lung cancer death rate by age and sex**

Age	Male	Female
50-54	0.297	0.207
55-59	0.186	0.175
60-64	0.111	0.087
65-69	0.073	0.085
70-74	0.046	0.069
75-79	0.027	0.056
80-84	0.016	0.039

Estimated using negative binomial regression in Equation (1). Includes controls and age interactions. The exponential of the above coefficients represents the proportional increase in the death rate for other causes associated with a one-per-thousand increase in the lung cancer death rate.

**Table 2: Age-specific lung cancer death rates (per 1,000)**

Age	Observed (2004)		Lifelong nonsmokers <sup>1</sup>	
	Male	Female	Male	Female
50-54	0.31	0.50	0.06	0.06
55-59	0.61	0.96	0.05	0.07
60-64	1.13	1.80	0.12	0.12
65-69	1.75	2.86	0.22	0.17
70-74	2.39	4.01	0.35	0.31
75-79	2.75	5.08	0.52	0.33
80-84	2.80	5.31	0.89	0.58

<sup>1</sup>From Thun et al. (1997) for death rates of never smokers in the Cancer Prevention Study, Cohort II 1982 - 1988.

**Table 3: Contribution of Smoking to Southern Mortality Disadvantage**

<b>Regional Comparison</b>	Fraction of difference in all-cause mortality due to smoking in 2004	
	Women	Men
Pacific <sup>1</sup>	17.6%	47.3%
Mountain <sup>2</sup>	35.5%	60.0%
Upper Midwest <sup>3</sup>	24.5%	42.8%
Central Midwest <sup>4</sup>	18.1%	47.5%
Northeast <sup>5</sup>	20.0%	50.2%
All non-Southern states	23.3%	50.3%

Note: Refers to death rate ages 50+ standardized to the 2000 U.S. population.  
 Southern states considered are: Alabama, Arkansas, Georgia, Kentucky,  
 Louisiana, Mississippi, Tennessee, West Virginia

<sup>1</sup>Alaska, California, Hawaii, Oregon, Washington

<sup>2</sup>Arizona, Colorado, Idaho, Montana, New Mexico, Utah, Wyoming

<sup>3</sup>Iowa, Minnesota, North Dakota, South Dakota, Wisconsin

<sup>4</sup>Illinois, Indiana, Michigan, Ohio

<sup>5</sup>Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York,  
 Rhode Island, Vermont



**Table 4: Mortality attributable to cigarette smoking in the U.S.:  
A comparison of estimates**

	Females		Males	
	2000	2004	2000	2004
Current Model <sup>1</sup>	0.17	0.17	0.22	0.21
Preston, Glei, Wilmoth (2010) <sup>2</sup>	0.19	0.2	0.23	0.22
Rostron (2010) <sup>3</sup>	0.14	—	0.22	—
Peto-Lopez <sup>4</sup>	0.21	—	0.24	—
CDC Method <sup>5</sup>	—	0.15*	—	0.23*
Rogers <sup>6</sup>	0.13 <sup>†</sup>	—	0.21 <sup>†</sup>	—

<sup>1</sup>Coefficient estimates across 50 U.S. states using negative binomial regression, ages 50-84

<sup>2</sup> Estimates pertain to ages 50-84 across countries using negative binomial regression

<sup>3</sup>Ages 50+ based on negative binomial regression including age-period interaction term

<sup>4</sup>Ages 35+. Peto-Lopez estimates from (<http://www.ctsu.ox.ac.uk/deathsfromsmoking>)

<sup>5</sup>Estimates reported by the Centers for Disease Control and Prevention (2008), ages 35+

<sup>6</sup>Figures reported in Rogers et al. (2005) for the year 2000, ages 35+

<sup>†</sup> Estimates pertain to ages 35+ in 2000

\* Estimates based on data for the period 2000-2004

**Table A.1: Estimated attributable fraction by state: 2004**

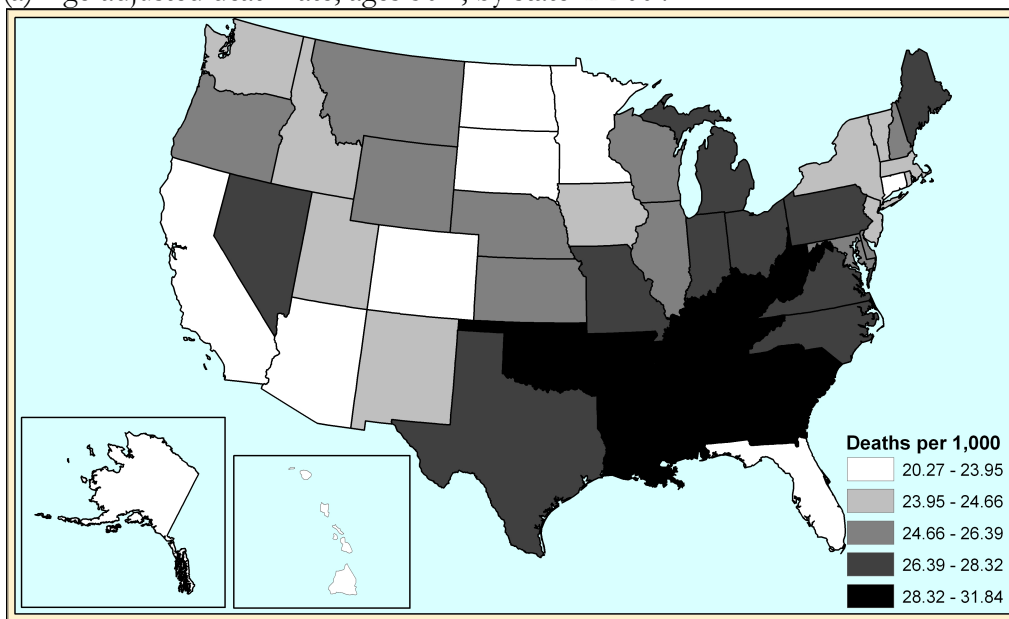
	Female		Male	
Alabama	0.16	(0.143,0.180)	0.25	(0.242,0.269)
Alaska	0.23	(0.207,0.248)	0.20	(0.191,0.211)
Arizona	0.16	(0.142,0.179)	0.17	(0.164,0.184)
Arkansas	0.20	(0.176,0.221)	0.27	(0.255,0.283)
California	0.15	(0.133,0.168)	0.16	(0.149,0.167)
Colorado	0.14	(0.125,0.156)	0.13	(0.127,0.143)
Connecticut	0.17	(0.149,0.189)	0.19	(0.182,0.204)
Delaware	0.21	(0.189,0.235)	0.25	(0.236,0.264)
Florida	0.18	(0.161,0.200)	0.20	(0.199,0.216)
Georgia	0.16	(0.139,0.175)	0.25	(0.238,0.263)
Hawaii	0.13	(0.113,0.138)	0.15	(0.146,0.162)
Idaho	0.15	(0.128,0.163)	0.15	(0.142,0.158)
Illinois	0.17	(0.153,0.192)	0.21	(0.203,0.227)
Indiana	0.18	(0.159,0.199)	0.24	(0.224,0.252)
Iowa	0.17	(0.148,0.184)	0.22	(0.206,0.231)
Kansas	0.17	(0.147,0.187)	0.21	(0.200,0.224)
Kentucky	0.22	(0.200,0.247)	0.30	(0.281,0.312)
Louisiana	0.18	(0.159,0.201)	0.26	(0.248,0.275)
Maine	0.20	(0.179,0.224)	0.23	(0.216,0.243)
Maryland	0.18	(0.159,0.200)	0.20	(0.189,0.211)
Massachusetts	0.18	(0.160,0.202)	0.19	(0.179,0.203)
Michigan	0.18	(0.163,0.205)	0.22	(0.205,0.228)
Minnesota	0.17	(0.150,0.186)	0.18	(0.172,0.193)
Mississippi	0.17	(0.152,0.190)	0.28	(0.268,0.296)
Missouri	0.19	(0.172,0.215)	0.25	(0.241,0.269)
Montana	0.20	(0.177,0.224)	0.17	(0.163,0.185)
Nebraska	0.14	(0.122,0.152)	0.20	(0.190,0.214)
Nevada	0.22	(0.194,0.245)	0.18	(0.175,0.195)
New Hampshire	0.20	(0.175,0.221)	0.21	(0.195,0.218)
New Jersey	0.17	(0.149,0.190)	0.18	(0.169,0.190)
New Mexico	0.12	(0.110,0.138)	0.13	(0.125,0.140)
New York	0.16	(0.138,0.172)	0.17	(0.166,0.185)
North Carolina	0.17	(0.147,0.184)	0.25	(0.235,0.261)
North Dakota	0.13	(0.121,0.149)	0.20	(0.190,0.215)
Ohio	0.18	(0.163,0.206)	0.23	(0.215,0.242)
Oklahoma	0.18	(0.159,0.200)	0.24	(0.229,0.255)
Oregon	0.20	(0.176,0.222)	0.20	(0.185,0.208)
Pennsylvania	0.16	(0.144,0.184)	0.20	(0.189,0.213)
Rhode Island	0.18	(0.156,0.198)	0.21	(0.201,0.226)
South Carolina	0.15	(0.135,0.169)	0.24	(0.232,0.257)
South Dakota	0.15	(0.131,0.162)	0.20	(0.193,0.218)
Tennessee	0.19	(0.164,0.206)	0.28	(0.263,0.292)
Texas	0.16	(0.143,0.180)	0.21	(0.195,0.217)
Utah	0.07	(0.059,0.076)	0.11	(0.102,0.114)

Vermont	0.16	(0.146,0.181)	0.19	(0.183,0.204)
Virginia	0.18	(0.157,0.199)	0.22	(0.207,0.231)
Washington	0.19	(0.171,0.215)	0.19	(0.177,0.199)
West Virginia	0.19	(0.164,0.208)	0.25	(0.233,0.263)
Wisconsin	0.16	(0.144,0.179)	0.20	(0.185,0.209)
Wyoming	0.14	(0.119,0.151)	0.15	(0.141,0.159)
<b>United States</b>	<b>0.17</b>	<b>(0.151,0.188)</b>	<b>0.21</b>	<b>(0.195,0.218)</b>

Note: 95% Confidence intervals in parentheses

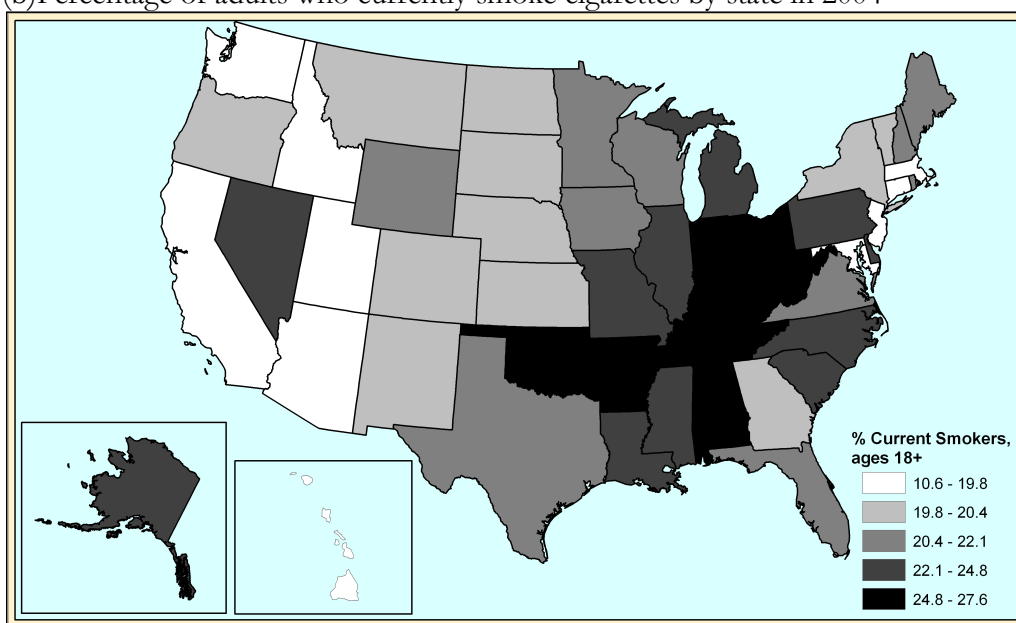
Figure 1: All-cause mortality and smoking prevalency by state: 2004

(a) Age-adjusted death rate, ages 50+, by state in 2004



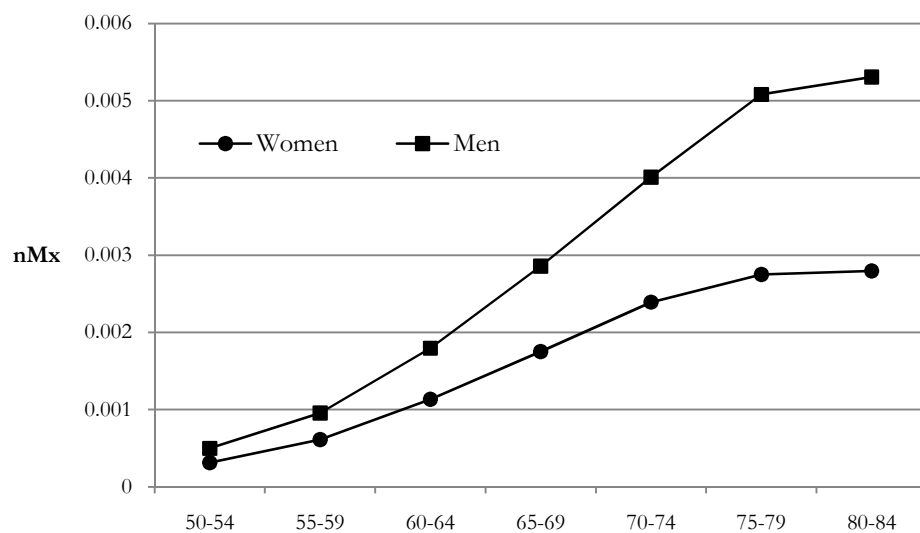
Source: Author's calculations from National Center for Health Statistics

(b) Percentage of adults who currently smoke cigarettes by state in 2004



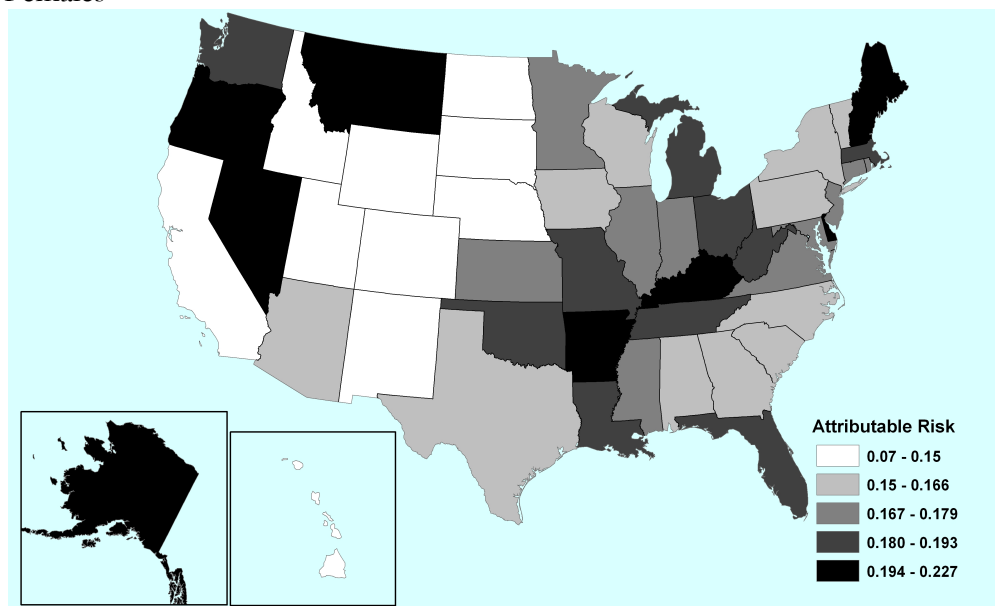
Source: Centers for Disease Control and Prevention

Figure 2: Age-specific lung cancer death rates for men and women in the United States in 2004



Source: National Center for Health Statistics

Figure 3: State-specific smoking-attributable mortality by sex: 2004  
Females



Males

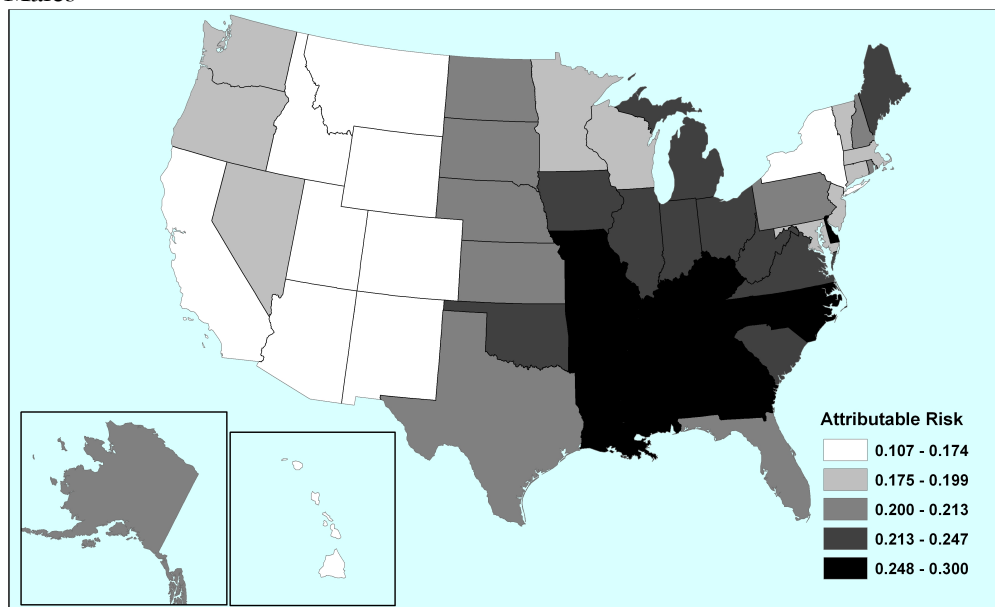


Figure 4: Comparability of Attributable fraction based on our estimates and CDC across 50 states

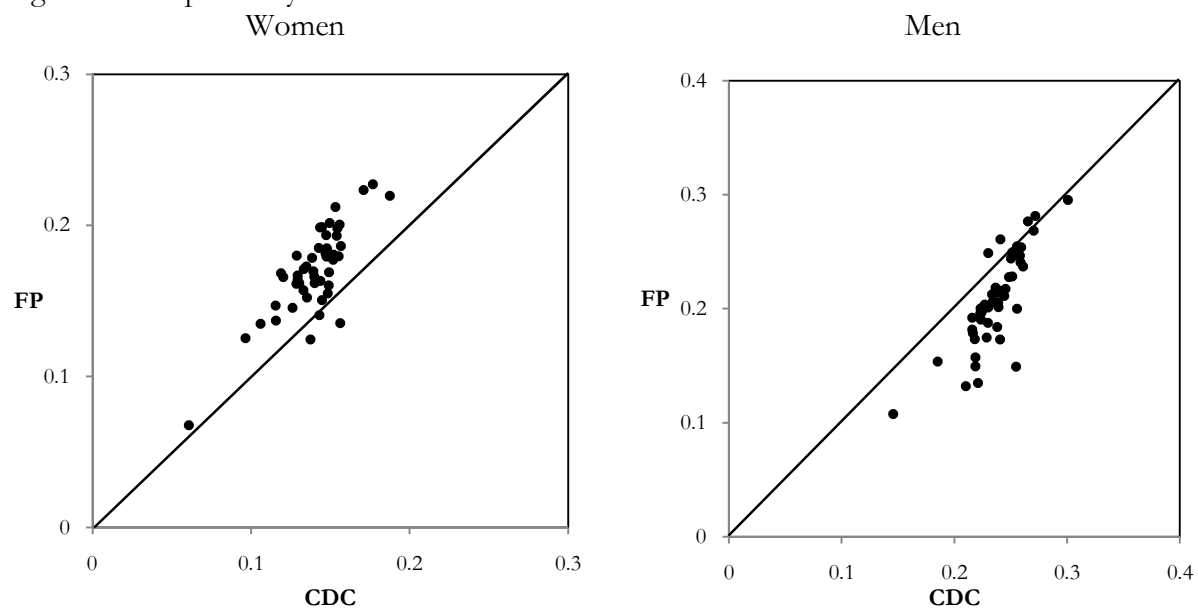


Figure 5: Estimated model coefficients based on the current model and Preston, Glei, and Wilmoth (2010)

