Web-Only Methods Section

The methods described in this section summarise the complex econometric approach that was used for estimating the cost of smoking in California. Further details are contained in a comprehensive report that includes estimates for the 58 counties of the state of California.[13] This study used a prevalence based approach, in which the annual economic burden of smoking was estimated for all smoking related illness and deaths that occurred in 1999 regardless of when a person first became ill. The burden of smoking consists of three components: direct cost, indirect cost of lost productivity from smoking related illness, and indirect cost of premature deaths caused by smoking related disease. The estimation of these three components relied on a common conceptual approach. In each case, a smoking attributable fraction (SAF) was estimated and applied to the total measure of interest. For example, the SAF for hospitalisation expenditures represents the proportion of hospitalisation expenditures that are attributable to smoking. This SAF was multiplied by the total hospitalisation expenditures to obtain smoking attributable hospital expenditures.

Direct costs

The direct costs of smoking in California are the health care expenditures for treatment of smoking related disease. Five types of health care services were included: hospitalisations, ambulatory care, nursing home care, prescription drugs, and home health care. Separate SAF models were developed for each type of heath care services for males and for females aged 18 years and older with the exception of the nursing home model, which was limited to those aged 55 and older because nearly all expenditures for nursing home care are limited to this older age group. These models were estimated using national data; California specific values of all independent variables were then used in the models to derive the California specific SAFs.

The national SAF models for hospitalisations, ambulatory care, prescription drugs, and home health care were based on the latest econometric approach developed by the authors[14–17] and were estimated by using the 1999 Aged National Medical Expenditure Survey (NMES). These models consist of multiple equations describing the effect of smoking history S (current, former, and never smokers) on the past history of five major smoking related diseases D (cancer, emphysema, arteriosclerosis, heart attack, and stroke), on self reported poor health status H, on the probability of having positive health care expenditures X, and on the magnitude of expenditures given that expenditures are positive. Demographic and socioeconomic status Y (age, race/ethnicity, geographic region, marital status, education, health insurance coverage, and family income) and other risk behaviours R (obesity and seatbelt use) were controlled in the model. The structural forms of these equations are:

$$D^* = f_1 (S, Y, R) \quad (1)$$

$$H^* = f_2 (S, Y, R, D^*|D) \quad (2)$$

$$Prob(X>0) = f_3 (S, Y, R, H^*|H) \quad (3)$$

$$Log(X|X>0) = f_4 (S, Y, R, H^*|H) \quad (4)$$

D is a binary variable that equals one if the respondent reported having one of the five major tobacco related diseases and zero otherwise. D* is an unobservable variable for the propensity for having major tobacco related diseases and was estimated as a probit model. H is self reported health status categorised as excellent, good, fair, or poor. H* is an unobservable variable for the propensity of having poor health and was estimated as an ordered probit model. $D^*|D$ denotes the expected propensity for having tobacco related disease conditional on self reported disease history. Likewise, $H^*|H$ denotes the expected propensity for having poor health conditional on self reported health status. Equation 3 was estimated as a probit model. Equation 4 is the logarithm of the magnitude of expenditures for those individuals with expenditures and was estimated using ordinary least squares.

The estimated coefficients of the national models were applied to 1999 California Tobacco Survey data to calculate predicted expenditures for each California male and female based on the actual values of all the independent variables. Another set of predicted expenditures was calculated for “hypothetical smokers” (current and former) who were identical to smokers in every way except that they were assumed to be never smokers. The difference between these two sets of predicted expenditures is the excess cost of smoking. The ratio of this difference to the aggregate predicted expenditures is the SAF.

Smokers have higher health expenditures because they are more likely to have smoking related diseases and poorer health (the “biological effect”). However, there is also a “mixed effect” that can increase or decrease expenditures for reasons not directly related to health status. For example, pregnant women do not generally describe themselves as being in poor health; yet pregnant smokers and their newborns often have higher medical expenses than their non-smoking counterparts. On the other hand, smokers might use fewer healthcare services than comparably healthy non-smokers due to their risk taking personalities and a resulting lower preference for seeking medical care. Because the mixed effects pathway was not causally related to tobacco use, and because 80–90% of the total effect results from the biological effect, mixed effects were not included here in the calculation of predicted excess costs of smoking and SAF. While the smoking variable remains in the model, for the final estimation it was replaced by mean current smoking prevalence, rather than 0 or 1.

The SAF for nursing home care was estimated using a conceptual model developed by Zhang.[18] According to this model, patients may be admitted to a nursing because they suffer from smoking related illness (the disability effect), or they may be forced to move to a nursing home when their caregiver dies from a smoking related illness and there is no one to care for them (the mortality effect). These effects combine to cause an increase in nursing home expenditures that is attributed to smoking. This model was estimated using the 1982–84 and 1987 National Health and Nutrition Examination Survey (NHANES) I epidemiologic followup study.

Smoking attributable expenditures in California were estimated for males and females, and for each of the five types of health care services, by multiplying the appropriate SAF by the corresponding state health care expenditure published by the Health Care Financing Administration (HCFA) for people aged 18 and older.[19] Expenditures were inflated to 1999 dollars using the appropriate component of the Consumer Price Index.

Indirect costs of lost productivity due to illness

Smokers with smoking related illnesses miss days of work and are unable to perform their usual activities. Two indicators of morbidity costs were considered here: smoking attributable work loss days and bed disability days. These were determined as the product of the SAF and the total number of days lost. We adapted the standard epidemiological formula to calculate SAFs for work loss days and bed disability days[20]:

$$SAF = \frac{[(p_n + p_c(RR_c) + p_f(RR_f)] - 1}{[(p_n + p_c(RR_c) + p_f(RR_f)]} \qquad (5)$$

where p_n, p_c, and p_f denote the percentage of never, current, and formers smokers; RR_c (RR_f) denotes the relative risk of the outcome measure of interest for current (former) smokers relative to never smokers. Relative risk for days lost was estimated using an econometric approach. First, work loss days or bed disability days were specified as a function of smoking status controlling for geographic region, demographic and socioeconomic variables and other risk behaviours and was estimated using a Tobit model. Work loss days were estimated separately for males and females 18 and older who were in the labour market using data from the 1999 National Health Interview Survey (NHIS). The model for bed disability days was estimated for females aged 18 and older who were not in the labour force but who were housekeeping using the 1999 NHIS data. The small sample size for males precluded estimation of smoking attributable bed disability days for them. Finally, the relative risk for current (or former) smokers was calculated as the ratio of predicted days for current (former) smokers to predicted days for "hypothetical current (former) smokers" with all the same characteristics of current (former) smokers except that they were assumed to be never smokers. The relative risk estimates were then used in equation 5 along with smoking prevalence rates estimated from the 1999 California Tobacco Survey data to determine the California specific SAFs.

Total days lost in California were extrapolated from the 1999 NHIS data, assuming that the work loss days and bed disability days in California were 11.96% of the US total since California represents 11.96% of the US total population aged 18 and older.[21] The SAF for days lost was applied to the total number of days lost to obtain smoking attributable days of lost productivity. For those in the labour market, work loss days were valued using California specific mean daily earnings plus an imputed value for household services by age and sex. For those not in the labour force, bed disability days were valued using a California specific mean daily imputed value for household services by age and sex. Labour market earnings included an adjustment for fringe benefits. The imputed value for household production was calculated by applying the mean wage rate for performance of tasks similar to those performed by housekeepers using the methodology developed by Douglass *et al*.[22]

Indirect costs of lost productivity due to premature death

We used the human capital approach to measure the value of lost productivity from lives lost due to smoking attributable diseases. The cost to society of smoking attributable premature death was calculated as the product of smoking attributable deaths and the present value of lifetime earnings (PVLE) for each person. We also estimated years of potential life lost due to smoking caused deaths.

The number of smoking attributable deaths was estimated by multiplying the SAFs by total deaths for each underlying cause of death reported as being causally linked to smoking in the 25th anniversary report of the US Surgeon General.[7] The SAF was determined for each age group and sex according to the epidemiological formula described above in equation 5. The relative risk of death by disease used was published by the Surgeon General.[7] Smoking prevalence rates were estimated using the 1999 California Tobacco Survey. Deaths for each smoking related diagnosis by sex and age (in five year category) were obtained from the 1999 California Death Statistical/Master file.[23] Only deaths for newborns and adults aged 35 and older were included.

The number of years lost from smoking caused death was estimated by sex and five year age group as the product of the number of smoking attributable deaths and the average number of years of life expectancy remaining at the age of death. Average remaining years of life for Californians was obtained from the California Department of Health Services.[24]

PVLE per person was estimated by five year age groups and sex using a computer program maintained at the University of California, San Francisco.[25] It takes into account life expectancy for different sex and age groups, varying rates of labour force participation or housekeeping, changing pattern of earnings at successive ages, an imputed value for housekeeping services, and a 3% discount rate to convert a stream of earnings into its current worth. To predict the future pattern of earnings, imputed value for household services, and labour force and housekeeping participation rates, it is assumed that people will be working and productive during their lifetimes in accordance with the current pattern of earnings and work experience for their sex and age groups. For this calculation, the proportions of the population participating in the labour force and housekeeping were estimated from the 1999 NHIS data.