Evaluation of the economic impact of California's Tobacco Control Program: a dynamic model approach

Leonard S Miller,[1] Wendy Max,[2] Hai-Yen Sung,[2] Dorothy Rice,[2] Malcolm Zaretsky[3]

[1]School of Social Welfare, University of California, Berkeley, California, USA
[2]Institute for Health & Aging, Department of Social and Behavioral Sciences, University of California, San Francisco, California, USA
[3]Department of Molecular and Cell Biology, University of California, Berkeley, California, USA

Correspondence to
Wendy Max, Institute for Health & Aging, University of California, San Francisco, 3333 California Street, Suite 340, San Francisco, CA 94118, USA; wendy.max@ucsf.edu

Received 25 December 2008
Accepted 8 January 2010

ABSTRACT

Objective To evaluate the long-term net economic impact of the California Tobacco Control Program.

Methods This study developed a series of dynamic models of smoking-caused mortality, morbidity, health status and healthcare expenditures. The models were used to evaluate the impact of the tobacco control programme. Outcomes of interest in the evaluation include net healthcare expenditures saved, years of life saved, years of treating smoking-related diseases averted and the total economic value of net healthcare savings and life saved by the programme. These outcomes are evaluated to 2079. Due to data limitations, the evaluations are conducted only for men.

Results The California Tobacco Control Program resulted in over 700 000 person-years of life saved and over 150 000 person-years of treatment averted for the 14.7 million male California residents alive in 1990. The value of net healthcare savings and years of life saved resulting from the programme was $22 billion or $107 billion in 1990 dollars, depending on how a year of life is discounted. If women were included, the impact would likely be much greater.

Conclusions The benefits of California's Tobacco Control Program are substantial and will continue to accrue for many years. Although the programme has resulted in increased longevity and additional healthcare resources for some, this impact is more than outweighed by the value of the additional years of life. Modelling the programme's impact in a dynamic framework makes it possible to evaluate the multiple impacts that the programme has on life, health and medical expenditures.

INTRODUCTION

The California Tobacco Control Program (CTCP) was established in 1989[1–3] using a portion of the tax revenues generated as a result of the Tobacco Tax and Health Protection Act, Proposition 99. With an annual budget of roughly $100 million, the CTCP became the largest comprehensive tobacco control programme in the world.[1–3] The $0.25/pack increase in tobacco taxes, which funded the programme went into effect in January 1989. Several other components were launched in spring 1990: a statewide anti-tobacco media campaign, community-based interventions and school-based prevention programmes. From the beginning, the CTCP has emphasised a strategy of changing social norms to make tobacco use less desirable, less acceptable and less accessible.[4 5] The ultimate goal is to reduce tobacco-related diseases, poor health and deaths in California.[6]

A number of studies have been undertaken to evaluate the impact of the CTCP.[7–15] Most of them focused on process indicators such as amount of funding and the scope of programme implementation, smoking outcome measures such as per capita cigarette consumption and smoking prevalence, and percentage of population protected by smoke-free homes or workplaces. A few studies examined the health benefit of the CTCP. Fichtenberg and Glantz[16] found that the CTCP was associated with an immediate reduction in deaths from heart disease. Another study reported that from 1989 to 1999, the CTCP was associated with a 6% reduction in lung cancer incidence.[17] Only one study evaluated the economic effect of the CTCP, estimating that the programme saved $86 billion in 2004 dollars of healthcare expenditures between 1989 and 2004.[18] However, the long-term economic effect of the CTCP, including reduced smoking-related diseases (SRDs) and reduced smoking-related deaths, has not been documented.

To the extent that the CTCP successfully reduces the incidence of SRDs, it would save smoking-attributable healthcare expenditures (SAEs). The SAEs in California were estimated at $8.6 billion for 1999[19] and $8.7 billion for 1993.[20] These estimates, referred to as 'gross' SAEs,[21] were based on an annual cost of smoking approach.[22] The reduction in number of premature deaths may impose additional healthcare expenditures during the prolonged years of life for people with avoided premature death. The tobacco industry refers to the potential saving from premature death as the 'death benefit'.[23 24] Cost of smoking estimates which take into consideration the expenditures net of the death benefit are referred to as 'net' SAEs.[21]

The issue of the 'gross' versus 'net' SAEs was first raised by Leu and Schaub.[25] They estimated the lifetime cost of smoking by simulating the medical cost history of Swiss men with and without cigarette smoking. They concluded that the extra years of costs experienced by the longer-lived non-smoking cohort approximately balanced out the higher costs during each year of the smokers' shortened lives. Barendregt *et al*[26] used a dynamic method to estimate the effects of smoking cessation on healthcare costs over time in Finland and found that if all smokers quit, healthcare costs would be lower at first but after 15 years there would be a net increase in healthcare costs. Hodgson[27] used a life cycle approach to estimate the lifetime cost of smoking in the US and found contradictory evidence. His results showed that ever smokers incurred higher lifetime medical expenditures than never smokers even after adjusting for never smokers' additional years of life. In the debate over whether to use the 'gross' or 'net' SAEs, Warner *et al*[21] suggested that the net measure is the logically correct one if the question of interest is how much greater a nation's healthcare expenditures are with smoking compared with the absence of smoking.

It is important that evaluations of public health programmes such as the CTCP consider the impact of the programme over time and capture the impact on mortality or longevity. For example, smoking cessation even late in life has been shown to increase life expectancy.[28–30] Excluding this aspect of the programme from an evaluation is implicitly placing a value of 0 on life. Thus the value of lives saved and prolonged must be taken into account, in addition to changes in healthcare expenditures. The objective of this study is to evaluate the long-term net economic impact of the CTCP using models designed to capture both these effects.

Because it is virtually impossible to separate the impact of the tobacco tax increase from the impact of other tobacco control activities undertaken by the CTCP, we consider them together. Four outcome measures are considered: (1) years of life saved, (2) years of treating SRDs averted, (3) net healthcare expenditures saved after adjusting for additional healthcare expenditures for people who live longer due to not smoking and (4) total economic value of net healthcare expenditures saved and years of life saved. The evaluation is conducted on a cohort consisting of all men who resided in California in 1990. Since those who did not take up smoking or who quit smoking due to the CTCP would enjoy health benefits long into the future, we used an evaluation period from 1990 through 2079, the year when the youngest in 1990 would turn age 90.

METHODS

Data sources

This study relies on four data sources.

National Academy of Sciences—National Research Council (NAS-NRC) Twin Registry

This is the largest national twin registry in the US. It consists of adult male twins born between 1917 and 1927 both of whom served in the military, mostly during World War II. Two questionnaires were mailed to registry members in 1967–1969 and 1983–1985 to collect information on registrants' smoking habits at the time of survey. The registrants' mortality status was periodically obtained from the computerised records of the US Department of Veterans Affairs (DVA),[31 32] which was notified of the death of approximately 98% of World War II veterans by relatives or morticians who sought to claim a burial allowance. The Twin Registry data with mortality status followed-up through November 1999 was used to estimate the dynamic smoking-attributable mortality model. We did not use the cause of death information.

National Medical Expenditure Survey (NMES-2)

This is a national household survey conducted in 1987 which contains detailed data for 34 459 individuals on smoking history, healthcare utilisation and expenditures, reasons for service use (diagnosis), source of payment, health status and history of certain diseases.[33] The NMES-2 data were used to estimate the dynamic smoking-attributable morbidity, health status and healthcare expenditures models. We adjusted the expenditures to 1990 dollars using the medical care component of the Consumer Price Index (CPI).[35]

Tobacco Use Supplement to the Current Population Survey (TUS-CPS)

This is a national survey targeting adults aged 15 and older. It is sponsored by the National Cancer Institute and administered as part of the CPS, the US Census Bureau's continuing labour force survey.[34] It contains detailed information cigarette smoking history and other tobacco use. The sampling design allows producing state-specific and national estimates.[36 37] The 1992/93, 1995/96, 1998/99 and 2001/02 TUS-CPS data were used to estimate population smoking initiation and cessation rates for California and other states in the evaluation analysis.

California Tobacco Survey (CTS)

This is a telephone survey of California residents that collects information about tobacco use behaviour and tobacco-related beliefs, attitudes and knowledge.[38 39] The 1990 CTS Adult File (ages 18+), Youth File (ages 12–17) and the child sample (ages 0–11) from the Screener file were used to construct a cohort of all California male residents aged 0 and older for the evaluation analysis. The study cohort consisted of a weighted total of 14 711 966 males of age 0 and older.

Statistical analysis

Analyses in this study were conducted using several statistical software packages. Mathematica[40] was used to derive and estimate the dynamic smoking-attributable mortality model and to predict the four outcome measures. LIMDEP V.8.0[41] was used to estimate the dynamic smoking-attributable morbidity, health status and healthcare expenditures models. SAS/STAT V.8.2[42] was used to estimate the population smoking initiation and cessation rates.

Dynamic models of smoking

We developed a series of dynamic models to describe the impact of smoking on mortality, morbidity, health status and healthcare expenditures for men aged 40 and older. The lower boundary of age 40 was chosen because most SRDs begin to appear at this age. Figure 1 contains a flowchart showing the estimation process for these models.

The smoking-attributable mortality model

This describes the dynamic relationship between an individual's smoking history and his annual probability of death. It is at the core of all the other models because it yields an estimated index for an individual's expected tobacco exposure, given his smoking history (age started smoking, cigarettes smoked per day, age quit). Subsequent morbidity, health status and healthcare expenditures models are all functions of this tobacco exposure index. These models are dynamic in the sense that the tobacco exposure index changes as an individual's smoking behaviour changes over time.

The dynamic smoking-attributable mortality model begins by deriving a theoretical distribution of the tobacco exposure index, which is the solution to a two-equation system of stochastic differential equations describing the body's ability to accumulate and purge tobacco toxins in relationship to smoking behaviour and ageing over time.

$$\mathrm{d}[\mathrm{tox_c(t)}]/\mathrm{dt} = \delta \mathrm{p} - \nu_\mathrm{c}(\mathrm{t}) \tag{1}$$

$$\mathrm{d}[\nu_\mathrm{c}(\mathrm{t})]/\mathrm{dt} = -\gamma_0 - \gamma_1\,\mathrm{tox_c(t)} + \sigma_\mathrm{c}\,\mathrm{dW_t} \tag{2}$$

The first equation is an instantaneous accounting identity stating that the time rate of change of cumulative tobacco exposure for a current smoker (denoted by subscript c) at time t

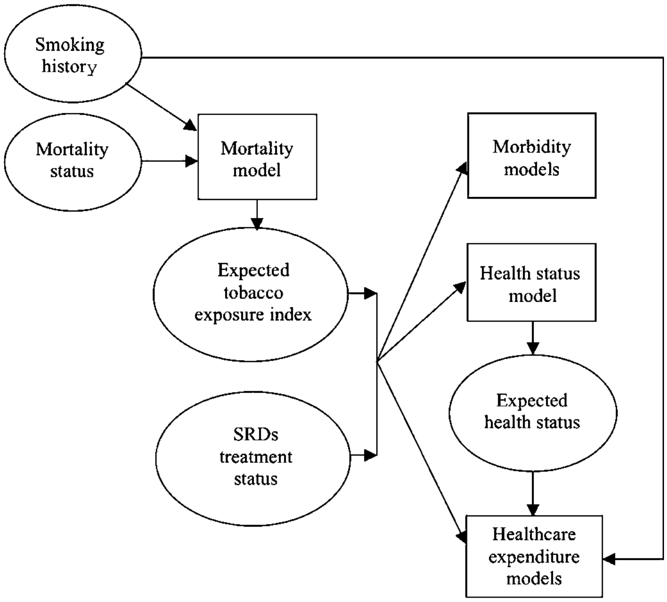


Figure 1 Flowchart of the estimation process for the dynamic models of smoking. Mortality model: the input includes two key variables for each respondent, (a) smoking history and (b) mortality status (including the date of death). From the mortality model, the parameters of the expected tobacco exposure index and the hazard rate are estimated. Given these parameters, the expected tobacco exposure index is derived. Morbidity models (including two models, one for high risk smoking-related diseases (SRDs) and another for low risk SRDs): the two key input variables are (a) the expected tobacco exposure index and (b) the SRDs treatment status. Health status model: the input is the same as that for the morbidity models. The output is the expected health status. Healthcare expenditure models (including three models separately for individuals with high risk SRDs, individual with low risk SRDs and individuals without SRDs): the input is the same as that for the morbidity models and for health status model plus two additional variables, (a) expected health status and (b) smoking history.

is equal to the difference between a smoker's momentary intake of tobacco exposure at time t (ie, the product of tobacco dosage per pack, δ, and packs of cigarettes smoked per day, p) and his momentary body purging of tobacco toxin at time t, $\nu_c(t)$. The second equation describes the time rate of change of a current smoker's body purging of tobacco toxin. It is specified as a function of: (1) a constant, which represents the reduction in purge ability due to ageing γ_0, (2) the cumulative tobacco exposure, γ_1 $tox_c(t)$, with the assumption that the body's purging ability declines with more tobacco exposure and (3) an instantaneous white noise term, W_t. To simplify the estimation process, we assume tobacco dosage per pack equals 1 and tobacco exposure at time 0 is 0.

This two-equation system is the same for former smokers except that, in equation 1, the momentary exposure term, δ p, is absent and the initial value of momentary body purging at the time when a former smoker quits has the same expected value as a current smoker with an identical smoking history. This specification implies that the tobacco exposure level accumulated in the body of former smokers would be diminished over time after they quit smoking.

The solution to equations 1 and 2 describes the expressions of the expected value and variance of the tobacco exposure index in the population with the same smoking history.[43] Supplementary appendix 1 derives the full theoretical distribution of this exposure index in detail.

The third equation in this analysis is a dynamic normal survival model specified as:

$$Die^*(t) = g(t) + \zeta(t) \qquad (3)$$

This equation states that the propensity to die by age t, Die* (t), is the sum of the expected propensity to die by age t, g(t), and a normally distributed random error term. The term g(t) is a function of an individual's age and the expressions of his expected tobacco exposure index at age t. Based on equation 3, and the expressions of the expected value and variance of the tobacco exposure index solved from equations 1 and 2, we derived the expressions of the hazard rate. Supplementary appendix 2 contains detailed description for the specification of equation 3 and the hazard rate formulation. We then estimated the parameters of the expected tobacco exposure index and the hazard rates with the maximum likelihood methods using the NAS-NRC Twin Registry data. Supplementary appendix 3 presents the detailed estimated parameters.

The smoking-attributable morbidity model

This includes two equations describing the propensity of being 'currently treated' for two groups of SRDs in a year. The first equation is for the group of high relative risk SRDs including lung cancer, laryngeal cancer and chronic obstructive pulmonary disease.[44] The second equation is for the group of low relative risk SRDs such as coronary heart disease, stroke and all other SRDs.[44] Both equations are specified as a function of individual's age and the expected tobacco exposure index. We estimated the morbidity model with a Probit model[45] using the NMES-2 data. See supplementary appendix 3 for the estimated parameters.

The smoking-attributable health status model

This describes the probability distribution of individual's self-reported health status (excellent, good, fair, poor) for individuals who are not currently treated for any SRDs. It is specified as a function of an individual's age and the expected tobacco exposure index. We estimated the health status model with an ordered Probit model using the NMES-2 data. See supplementary appendix 3 for the estimated parameters.

The smoking-attributable healthcare expenditures model

This describes the total healthcare expenditures of an individual in a year, and is estimated using the NMES-2 data for three groups of individuals stratified by disease status. For those currently treated for high relative risk SRDs, an individual's expected total expenditures are estimated as the average total expenditures of all individuals who have the same smoking status in this group. For those currently treated for low relative risk SRDs, an individual's annual total expenditures are modelled as a function of ever smoker status and his expected poor health status. This model was estimated using the ordinary least squares (OLS) method. For those not currently treated for any SRDs, a two-part model[46] is used to describe the propensity of having healthcare expenditures (first-part model) and the logarithm of the magnitude of annual expenditures for those with positive expenditures (second-part model). The first-part and second-part models are specified as functions of an individual's age, the expected tobacco exposure index, smoking status and expected poor health status. We estimated the first-part model with a Probit model and the second-part model with the OLS method. See supplementary appendix 3 for the model specification and estimated parameters.

Evaluation of the economic impact of the CTCP

We evaluated the economic impact of the CTCP over the full life of a cohort of all 1990 California male residents obtained from the 1990 CTS data. For each year, we estimated two sets of predictions for each outcome measure. The first set was estimated under the CTCP (the factual situation). The second set was estimated under the assumption that the CTCP did not exist (the counterfactual situation). The effects of the CTCP were measured as the difference between these two sets of predictions.[25] Specifically, this evaluation consists of three steps.

Simulate smoking initiation and cessation rates

In order to estimate the two sets of predictions, the population smoking behaviour of the cohort under the factual and counterfactual situations from 1990 to 2079 was simulated. We focused on two measures of smoking behaviour: smoking initiation and successful cessation.

We calculated the yearly smoking initiation and cessation rates during the period of 1981–1999 using the TUS-CPS data. Never smokers were defined as those who answered 'no' to the question: 'Have you smoked at least 100 cigarettes in your entire life?' Those who answer 'yes' were ever smokers. Ever smokers were asked, 'How old were you when you started smoking cigarettes fairly regularly?' Ever smokers were also asked whether they currently smoked. If not, they were defined as former smokers and were further asked: 'About how long has it been since you last smoked cigarettes fairly regularly?' We adopted previously developed techniques[47–52] to calculate yearly smoking initiation rates and cessation rates. First, we reconstructed each respondent's smoking status retrospectively for each year before the year of the survey, from 1981 to 1999. Consistent with another recent study, we assumed respondents' state of residence did not change over time during this reconstruction period.[53] Second, we calculated smoking initiation rates for three age groups (11–15, 16–18 and 19–22) separately for California and for all other states by dividing the weighted number of respondents who started smoking in a given year by the weighted number of non-smokers in the beginning of that year. Third, cessation rates were calculated by dividing the weighted number of long-time quitters with at least 6 months of abstinence who quit smoking in a given year by the weighted number of respondents who were current smokers in the beginning of that year. The cessation rates were calculated for four age groups (20–34, 35–44, 45–54 and 55+) for California and all other states. Finally, these crude rates were smoothed using a 3-year moving average.

For each age group, a time series model of California's smoothed initiation (or cessation) rates during 1981–1999 was specified as a function of all other states' smoothed initiation (or cessation) rates, a dummy variable measuring the effect of the CTCP (value of 1 since 1989; 0 otherwise), and a time trend, using a method similar to that employed by Fichtenberg and Glantz.[16] By including the rate for all other states in the model, we controlled for changes in the California rate due to national changes in risk factors. The simulated initiation (or cessation) rates under the factual situation were given by the predicted values from this model, and the simulated initiation (or cessation) rates under the counterfactual situation were also given by the predicted values from this model except that the dummy variable for the CTCP was assumed to be 0 in 1989 and later years. After 1999, age-specific factual and counterfactual initiation and cessation rates were assumed to be at their respective 1999 levels.

Simulate mortality, morbidity, health status and healthcare expenditures

The simulated smoking initiation and cessation rates and the estimated parameters from the dynamic models of smoking were applied to the California cohort to simulate their lifetime outcomes under the factual and counterfactual situations. For each year from 1990 to 2079, we began to simulate who dies or survives for individuals aged 40 and older. If an individual survives or is not yet 40 years old, we simulated who takes up or quits smoking and who remains at their previous year's smoking status, and estimated the expected tobacco exposure index. For individuals aged 40 and older who survives, we simulated who is currently treated for high or low relative risk SRDs and who is not, and for those not treated, what each individual's expected health status is. We then predicted each individual's healthcare expenditures. All of these simulations were performed under the factual and counterfactual situations. Supplementary appendix 4 contains details of the design of the simulations.

Estimate the effects of the CTCP on four outcome measures

Given the above simulation results, we used four different algorithms to estimate the effects of the CTCP on four outcome measures: (1) years of life saved, (2) years of treating SRDs averted, (3) net healthcare expenditures saved after adjusting for additional healthcare expenditures for people who live longer due to not smoking and (4) total economic value of net healthcare expenditures saved and years of life saved.

In the first algorithm, an individual is dropped from the factual and the counterfactual simulations when he dies in either simulation. Therefore, this algorithm derives 'gross' healthcare savings without considering the impact of potential prolonged years of life due to the CTCP. This is similar to what is assumed in the annual cost of smoking studies of national and state estimates of smoking-attributable expenditures.[19 20 44 54 55]

In the second algorithm, individuals who die in the factual or counterfactual simulation are still included in the other simulation until they die or reach age 90. Because more individuals live longer due to the health benefit of the CTCP, those additional years of life lead to additional healthcare expenditures. Therefore, this algorithm derives 'net' healthcare savings due to the CTCP, analogous to the lifetime cost of smoking studies,[25–27] by taking into account the reduced smoking-attributable healthcare expenditures during the years while people are alive and the additional non-smoking-related healthcare expenditures during the prolonged years of life.

The third and fourth algorithms consider the value of lives saved by the CTCP in addition to net healthcare savings. Because premature deaths from smoking usually occur among older people who have relatively low market earnings, we valued years of life using a willingness-to-pay (WTP) approach. While early WTP studies implied the value of life ranging from \$3 million to \$7 million,[56] Sloan *et al*[29] used a conservative value of \$100 000 per life year to estimate the economic losses from smoking-related mortality. We adopted \$100 000 per year to value the life year in 1990 with adjustments depending on each person's disease treatment and health status: \$100 000 for excellent health, \$80 000 for good health, \$50 000 for fair health and \$25 000 for poor health or being treated for SRDs. The only difference between the third and fourth algorithms is the discount rate used to calculate the present value of expected life years saved (see below).

An alternative approach for considering the value of life is to calculate disability adjusted life years (DALYs) or quality adjusted life years (QALYs). While formally calculating DALYs or QALYs was beyond the scope of our study due to the lack of data availability, we have taken into account the dimensions of the quantity and quality of life that DALYs and QALYs capture by assigning different values for a year of life based on disease and health status in our third and fourth algorithms.

Discounting

In all four algorithms, the present value of healthcare expenditures saved by the CTCP was estimated by taking into account discounting as performed in the lifetime costs of smoking literature.[27 29 57 58] First, considering the potential growth in future healthcare expenditures, we expressed each person's future stream of annual healthcare expenditures during his expected 'lifetime' from age 40 to age 90 by inflating the 1990 value of the predicted healthcare expenditure by 2% per year. The growth rate of 2% is approximately the difference in average annual growth rate between the CPI for medical care and the CPI for all urban consumers during 1990–1999.[35] Second, this future stream of predicted expenditures was discounted by the rate of time preference at 3% per year to derive the present value of the lifetime expenditures. For any person, the healthcare expenditures saved by the CTCP equalled the present value of the lifetime expenditures under the counterfactual situation minus the present value of the lifetime expenditures under the factual situation. Total healthcare expenditures saved for all male Californians were obtained by summing savings across individuals, taking into account sampling weights in the 1990 CTS data.

In the third algorithm, we discounted the value of future life years by the rate of time preference using 3% per year.[59] In the fourth algorithm, we discounted the value of future life years by discount rates that approximated the differential probabilities of death among individuals of different smoking statuses. As an approximation, we discounted a year of a current smoker's life by 2%, a year of a former smoker's life by 1.5% and a year of a never smoker's life by 1%.

RESULTS

The estimated probability of survival given age and smoking history is illustrated in figure 2. The more exposure to tobacco a person has had, as measured by number of decades smoked (or formerly smoked) and by packs per day smoked, the lower the probability that he will be alive.

Figure 3 shows the observed, predicted and simulated smoking cessation rates for California males during 1981–1999 for four age groups. The young adult group (20–34) had the highest increase in cessation rates—from about 2% in 1981 to over 5% in 1999—and was most responsive to the CTCP in cessation especially after 1995, as measured by the difference between the predicted cessation rates under the factual situation and the simulated cessation rates under the counterfactual situation. Figure 4 shows the observed, predicted and simulated smoking initiation rates for three age groups. The initiation rates for California males were lower than those for all other states, especially for the group aged 16–18. For the groups aged 11–15 and 19–22, their initiation rates declined noticeably after 1995 and the reduction was related to the implementation of the CTCP.

Table 1 shows the estimated economic benefits of the California CTCP for the 1990 cohort followed until death. Almost three-quarters of a million person-years of life are saved. In addition, 141 426 person-years of treatment for the high relative risk SRDs and 16 240 person-years of treatment for the low relative risk SRDs are averted.

Using our first algorithm, we estimate that the CTCP saved $1.438 billion dollars (in 1990 dollars) in healthcare costs over the period from 1990 through 2079. The estimate is statistically significant at p value <0.05, two-tailed test.

Our second algorithm yields an estimate of 'net' healthcare savings from the CTCP, including the additional healthcare expenditures associated with living longer due to the CTCP. The present value of the net savings for healthcare expenditures was estimated as −0.144 billion (in 1990 dollars), but is not statistically significant.

Based on the third and fourth algorithms, we derived two estimates for the total economic value of net healthcare savings and years of life saved due to the CTCP, valuing a year of life at $100 000 with adjustments for individual's disease treatment and health status. From the third algorithm, our estimated present value of the total net healthcare resources saved plus the value of years of life saved was $22.443 billion (in 1990 dollars). From the fourth algorithm, we estimated that the CTCP would generate $107.418 billion (in 1990 dollars) of total savings including net healthcare saving and the value of life saved. Both estimates are statistically significant at p value <0.05, two-tailed test.

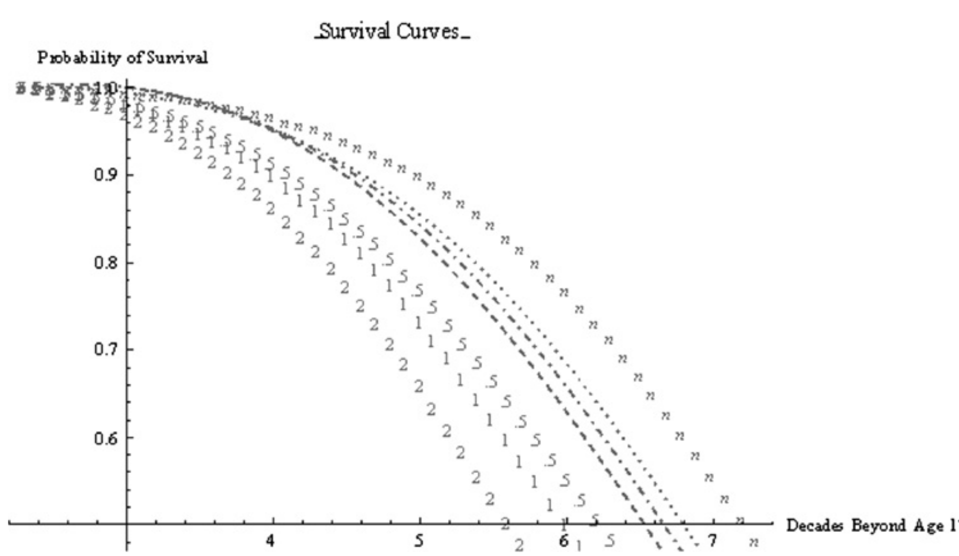


Figure 2 Probability of survival for men with different smoking histories. Seven survival curves denote different smoking histories: n, never smoker; …, former smoker who smoked 1 pack/day for 10 years since age 17 and quit at age 27; ·—·, former smoker who smoked 1 pack/day for 20 years since age 17 and quit at age 37; — —,former smoker who smoked 1 pack/day for 30 years since age 17 and quit at age 47; 5, current smoker who smoked 0.5 pack/day since age 17; 1, current smoker who smoked 1 pack/day since age 17; 2, current smoker who smoked 2 pack/day since age 17. Age 17 was chosen because it is the mean age when male smokers began to smoke.

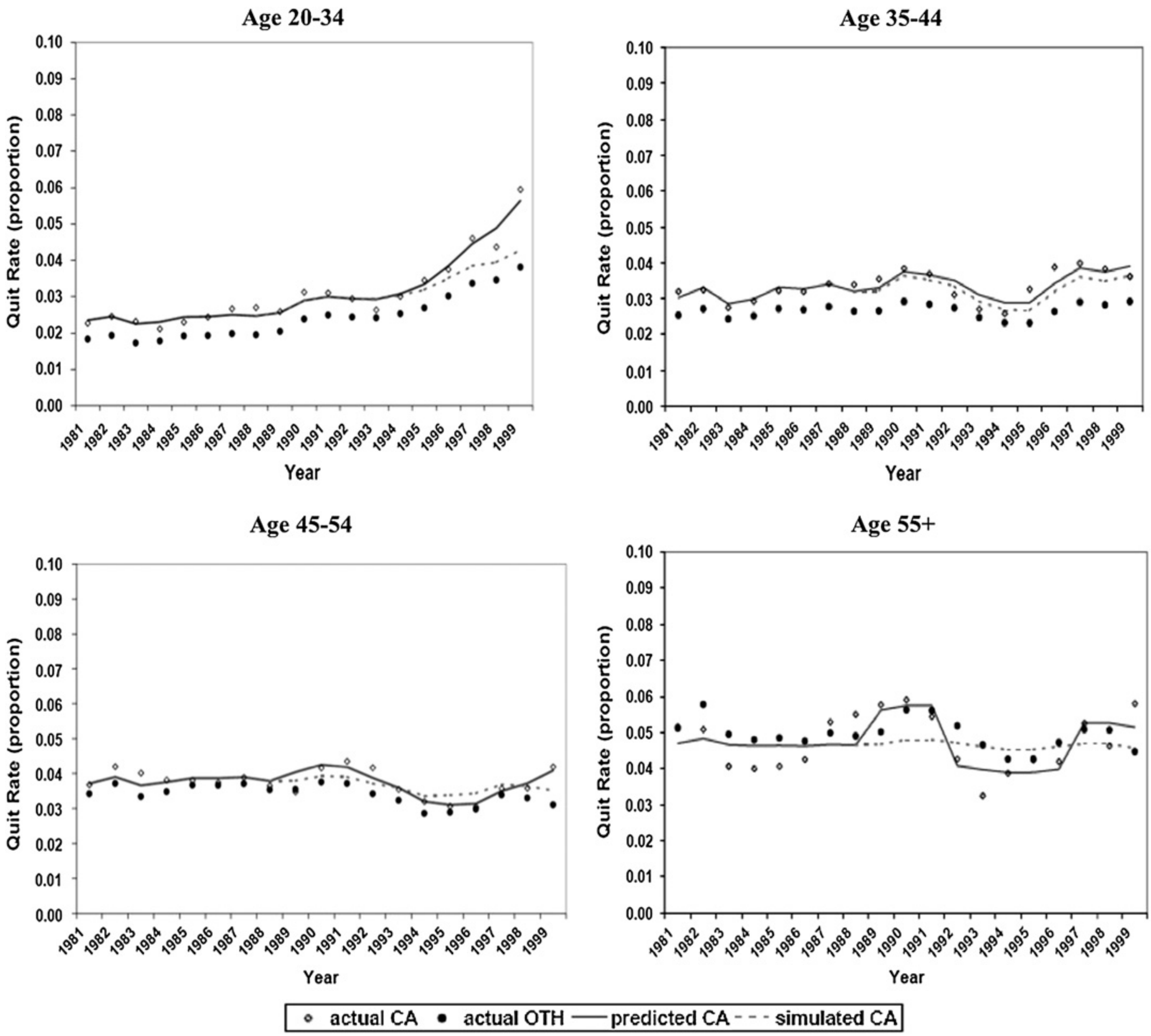


Figure 3 Smoking cessation rates for men in California and all other states by age, 1981–1999. The actual rates (dots) represent the 3-year moving average of the observed cessation rates for California (CA) and all other states (OTH). The predicted CA rates mean the predicted cessation rates from the time series model under the factual situation. The simulated CA rates mean the predicted cessation rates from the time series model under the counterfactual situation.

DISCUSSION

Our results highlight the importance of developing a comprehensive measure for evaluating the impact of a tobacco control programme that considers the value of healthcare resources used, and also the value of years of life saved and of improved health status associated with not smoking. A comparison of the 'gross' healthcare expenditures to the 'net' healthcare expenditures shows that when the healthcare costs resulting from longer life are considered, the healthcare savings from the CTCP disappear. However, these approaches ignore the value of having people live longer and healthier. When a value for life is included, the total economic value of the benefits from the CTCP amounts to \$22.4 billion in 1990 dollars. This is more than a 15-fold increase over the estimate of the 'gross' healthcare savings and a very different result from the 'net' healthcare savings, which ignore the value of life. This value is equivalent to \$35.5 billion in 2007 dollars (adjusted by the CPI). When an individual's probability of death is used to discount the years of life, the CTCP would generate \$107.4 billion in 1990 dollars, a 75-fold increase over the estimate of the 'gross' healthcare savings. This value is equivalent to \$170.2 billion in 2007 dollars. Given that a key public health outcome is improved health, the value of life saved and improved health should be central to evaluating the destructive effects of smoking, the single most important preventable public health hazard.

During the first decade of the programme, the CTCP spent about \$1.2 billion dollars (A Roeseler, California Department of Public Health, California Tobacco Control Program, personal communication, 2005). This is dwarfed by the total economic value of the net healthcare savings, lives saved and health improved due to the programme. However, it must be noted that our estimates result from the combined effect of the

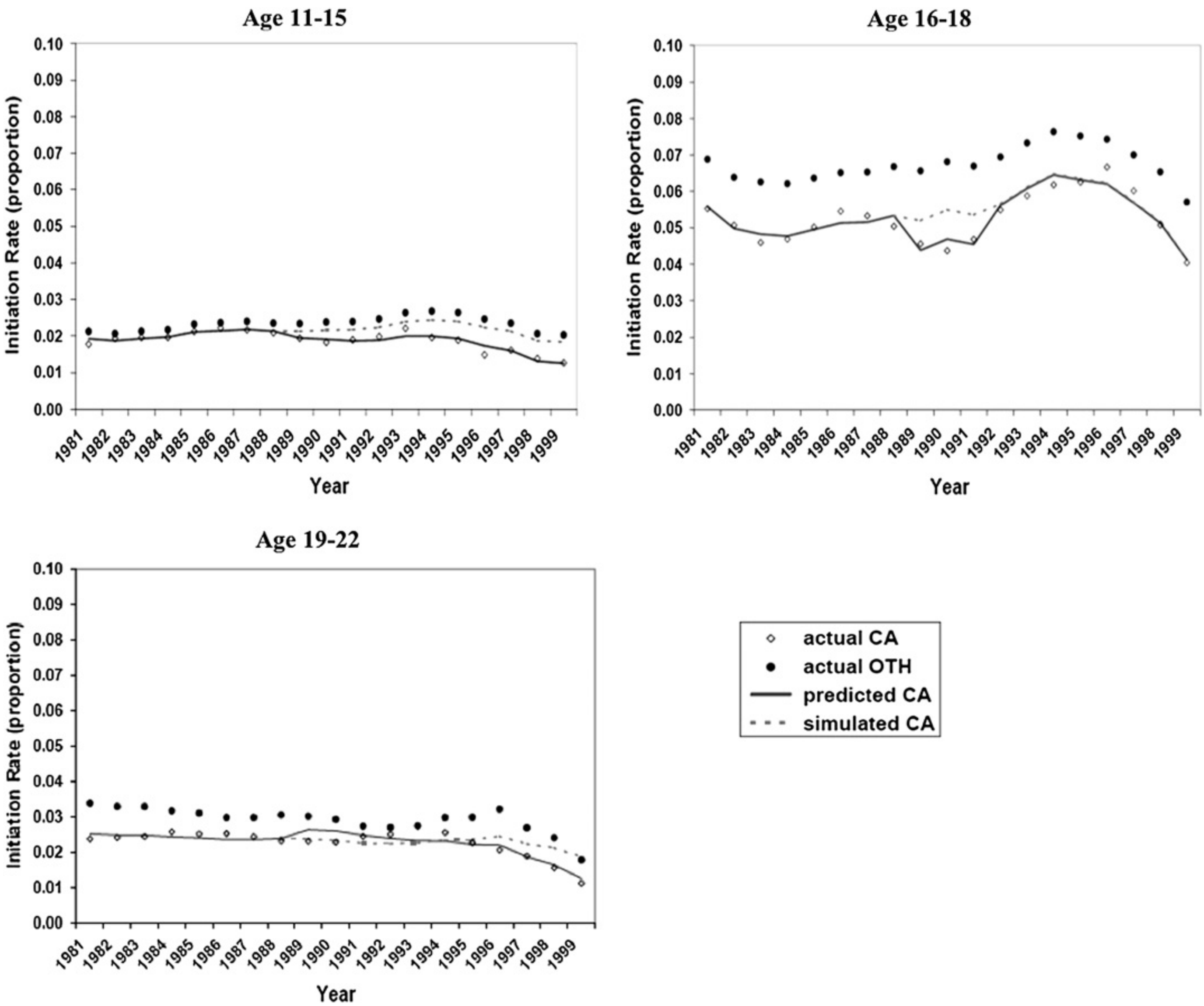


Figure 4 Smoking initiation rates for men in California and all other states by age, 1981–1999. The actual rates (dots) represent the 3-year moving average of the observed initiation rates for California (CA) and all other states (OTH). The predicted CA rates mean the predicted initiation rates from the time series model under the factual situation. The simulated CA rates mean the predicted initiation rates from the time series model under the counterfactual situation.

tobacco tax increase and other components of the CTCP including the statewide media campaigns, community-based interventions and school-based prevention programmes.

There are several limitations to this study. First, women were not included in the analysis because longitudinal data on female mortality and smoking were unavailable. However, we postulate that the economic effects of the CTCP for women would be on the order of two-thirds the size of the effects for men because the smoking prevalence rate for women was approximately 69% of the rate for men in California.[19] Thus, the total economic value of the CTCP including men and women may be considerably larger than our estimates. Further research is needed to include women in the evaluation so that a fuller understanding of the importance of tobacco control programmes can be obtained. Second, for the period from 2000 to 2079, we assumed that smoking initiation and cessation rates remain at their 1999 levels. For these rates to remain constant beyond 1999, tobacco control efforts must be sustainable. Future research could explore how the economic impact of the CTCP is sensitive to post 1999 smoking rates. Third, in the analyses of yearly smoking initiation rates and cessation rates, we assumed the state of residence for the respondents of the TUS-CPS data was unchanged. A recent study which compared the cessation rates between California and a group of comparison states pointed out that a large movement of former smokers from California to other states during the study period would artificially inflate the estimated cessation rate in other states, and vice versa.[53] Further research is needed to explore the smoking population's movement between states so as to determine the direction of potential bias due to such assumption. Fourth, the simple specification of a single dummy variable for the CTCP in the smoking initiation and cessation equations implies that the effect of the programme was constant over time. However, it has been reported that the impact of the CTCP on smoking prevalence rates was stronger during the early 1990s than during the late 1990s, implying that the impact of the CTCP on

Table 1 Estimated economic impact of the California Tobacco Control Program (CTCP) over a 90-year evaluation period from 1990 through 2079

Outcome measures	Predicted value	SE
A. Years of life saved (person-years)	712966*	60590
B. Years of treatment saved (person-years):		
High relative risk smoking-related diseases	141426*	5903
Low relative risk smoking-related diseases	16240	13617
C. Healthcare expenditures saved (in billions):		
Algorithm 1: 'gross' healthcare savings without accounting for the impact of prolonged years of life due to the CTCP	\$1.438*	\$0.227
Algorithm 2: 'net' healthcare savings after adjusting for additional healthcare expenditures associated with prolonged years of life due to the CTCP	−\$0.144	\$0.217
D. Total economic value of 'net' healthcare savings and years of life saved, assuming a year of life is valued at \$100 000 with adjustments for disease treatment and health status (in billions):		
Algorithm 3: present value of life years discounted at 3%	\$22.443*	\$1.118
Algorithm 4: present value of life years discounted at 2% for current smokers, 1.5% for former smokers and 1% for never smokers	\$107.418*	\$1.629

All monetary values are in 1990 dollars.
*Statistically significant at p value <0.05, two-tailed test.

smoking initiation and cessation rates might not be constant over time.[8] Nevertheless, even with this limitation, our results are in general consistent with the findings from a study by Messer and Pierce *et al*[53] which showed that from 1980 to 1999, cessation rates increased most for the young (age 20–34), and this age group also showed the greatest difference between California and the comparison states. Finally, we did not include any impact of the CTCP on secondhand smoke exposure, though data have shown a substantial decrease in exposure over time.

Tobacco control programmes are costly. However, the benefits of the programmes are substantial and continue to accrue for many years. Although those who are persuaded not to smoke will live longer, have better health status and require additional healthcare resources during their additional years of life, this impact is outweighed by the value of additional years of life and better health. Public health programmes need to be evaluated with healthcare costs, additional years of life and improved health considered as important outcomes.

What this paper adds

- This paper develops a series of dynamic models of smoking behavior and consequences that analyze the impact of smoking initiation and cessation on morbidity, mortality, health status, and healthcare expenditures over the lifetime of Californians.
- The models are used to evaluate the impact of the first decade of the California Tobacco Control Program on males.
- The findings indicate that when the value of increased longevity is included, the program saved \$22–\$107 billion, depending on how life is valued.

Funding This research was funded by the California Tobacco-Related Disease Research Program under grant 9RT-0157. The grant was awarded competitively and the funding agency was not involved in any aspect of the research.

Competing interests None.

Contributors All authors made contributions to this paper that justify authorship.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. **Tobacco Education and Research Oversight Committee (TEROC).** Toward a tobacco-free California: strategies for the 21st century 2000–2003, Jan 2000. http://www.cdph.ca.gov/programs/Tobacco/Documents/CTCPMasterPlan00-03.pdf (accessed 2 Aug 2009).
2. **Novotny T,** Siegel MB. California's tobacco control saga. *Health Aff* 1996;**60**:58–72.
3. **Farrelly MC,** Pechacek TF, Chaloupka FJ. The impact of tobacco control program expenditures on aggregate cigarette sales: 1981–2000. *J Health Econ* 2003;**22**:843–59.
4. **Tobacco Control Section.** *A model for change: the California experience in tobacco control*. Sacramento, CA: California Department of Health Services, 1998. http://www.cdph.ca.gov/programs/tobacco/Documents/CTCPmodelforchange1998.pdf (accessed 2 Aug 2009).
5. **Tobacco Education Research and Oversight Committee (TEROC).** *Confronting a relentless adversary: a plan for success toward a tobacco-free California 2006-2008*. Sacramento, CA: California Department of Health Services, Mar 2006 http://ww2.cdph.ca.gov/programs/Tobacco/Documents/CTCPMasterPlan06-08ExecSum.pdf (accessed 10 Aug 2009).
6. **California Department of Public Health, California Tobacco Control Program.** *California Tobacco Control Update 2009: 20 Years of Tobacco Control in California*: Sacramento, CA: California Department of Public Health, 2009. http://ww2.cdph.ca.gov/programs/tobacco/Documents/CTCPUpdate2009.pdf (accessed 10 Aug 2009).
7. **Pierce JP,** Gilpin EA, Emery SL, *et al*. *Tobacco control in California: Who's winning the war? An evaluation of the Tobacco Control Program, 1989-1996*. La Jolla, CA: University of California, San Diego, 1998.
8. **Gilpin EA,** Emery SL, Farkas AJ, *et al*. *The California tobacco control program: a decade of progress, results for the California tobacco surveys, 1990-1999*. La Jolla, CA: University of California, San Diego, 2001.
9. **Al-Delaimy WD,** White MM, Gilmer T, *et al*. *The California tobacco control program: can we maintain the progress? results from the California tobacco survey, 1990-2005*. **vol 1**. La Jolla, CA: University of California, San Diego, 2008.
10. **Pierce JP,** Gilpin EA, Emery SL, *et al*. Has the California tobacco control program reduced smoking? *JAMA* 1998;**280**:893–9. Erratum in: JAMA 281(1):37.
11. **Norman GJ,** Ribisl KM, Howard-Pitney B, *et al*. The relationship between home smoking bans and exposure to state tobacco control efforts and smoking behaviors. *Am J Health Promot* 2000;**15**:81–8.
12. **Rohrbach LA,** Howard-Pitney B, Unger JB, *et al*. Independent evaluation of the California tobacco control program: relationships between program exposure and outcomes, 1996-1998. *Am J Public Health* 2002;**92**:975–83.
13. **Siegel M.** The effectiveness of state-level tobacco control interventions: a review of program implementation and behavioral outcomes. *Annu Rev Public Health* 2002;**23**:45–71.
14. **Gilpin EA,** Messer K, White MM, *et al*. What contributed to the major decline in per capita cigarette consumption during California's comprehensive tobacco control programme? *Tob Control* 2006;**15**:308–16.
15. **McCarthy WJ,** Dietsch BJ, Dent C, *et al*. *Evaluation of the In-school tobacco use prevention education program, 2005-2006*. Sacramento, CA: California Department of Public Health, 2008. http://www.cdph.ca.gov/programs/tobacco/Documents/CTCPTupeReport05-06.pdf (accessed 2 Aug 2009).
16. **Fichtenberg CM,** Glantz SA. Association of the California Tobacco Control Program with declines in cigarette consumption and mortality from heart disease. *N Engl J Med* 2000;**343**:1172–5.
17. **Barnoya J,** Glantz SA. Association of the California Tobacco Control Program with declines in lung cancer incidence. *Cancer Causes Control* 2004;**15**:689–95.
18. **Lightwood JM,** Dinno A, Glantz SA. Effect of the California Tobacco Control Program on personal healthcare expenditures. *PloS Med* 2008;**5**:e178.
19. **Max W,** Rice DP, Sung HY, *et al*. The economic burden of smoking in California. *Tob Control* 2004;**13**:264–7.
20. **Miller LS,** Zhang X, Rice D, *et al*. State estimates of total medical expenditures attributable to cigarette smoking, 1993. *Public Health Rep* 1998;**113**:447–58.
21. **Warner KE,** Hodgson TA, Carroll CE. Medical costs of smoking in the United States: estimates, their validity, and their implications. *Tob Control* 1999;**8**:290–300.
22. **Max W.** The financial impact of smoking on health-related costs: A review of the literature. *Am J Health Promot* 2001;**15**:321–31.
23. **Zeger S,** Wyant T, Miller L, *et al*. Statistical testimony on damages in Minnesota versus the tobacco industry. In: Gastwirth J, ed. *Statistical science in the courtroom*. New York: Springer, 2000:19–55.
24. **Rubin DB.** Statistical issues. In: Gastwirth J, ed. *Statistical science in the courtroom*. New York: Springer, 2000:19–55.
25. **Leu RE,** Schaub T. Does smoking increase medical expenditures? *Soc Sci Med* 1983;**17**:1907–14.

26. **Barendregt JJ,** Bonneu L, Van Der Mass PJ. The health care costs of smoking. *N Engl J Med* 1997;**337**:1052—7.
27. **Hodgson TA.** Cigarette smoking and lifetime medical expenditure. *Milbank Q* 1992;**70**:81—125.
28. **Taylor DH,** Hasselblad V, Henley SJ, *et al*. Benefits of smoking cessation for longevity. *Am J Public Health* 2002;**92**:990—6.
29. **Sloan FA,** Ostermann J, Picone G, *et al*. *The price of smoking*. Cambridge, Mass: MIT Press, 2004.
30. **Doll R,** Peto R, Boreham J, *et al*. Mortality in relation to smoking: 50 years' observations on male British doctors. *BMJ* 2004;**328**:1519.
31. **Zaretsky M.** Communication between identical twins: Health behavior and social factors are associated with longevity that is greater among identical than fraternal U.S. World War II veteran twins. *J Gerontol A Biol Med Sci* 2003;**58A**:566—72.
32. **Hrubec Z,** Neel JV. Familial factors in early deaths: twins followed 30 years to ages 51-61 in 1978. *Hum Genet* 1978;**59**:39—46.
33. **U.S. Department of Health and Human Services (US DHHS).** *National medical expenditure survey, 1987: household survey data*. Rockville, MD: Agency for Health Care Policy and Research, 1991.
34. **Hartman A,** Willis G, Lawrence D, *et al*. *The 1998—1999 NCI tobacco use supplement to the current population survey (TUS-CPS): representative survey findings*. Bethesda, MD: US Department of Health and Human Services, National Institutes of Health, National Cancer Institute, 2002.
35. **U.S. Department of Labor, Bureau of Labor Statistics**, Consumer Price Index Detailed Report, Table 3A: Consumer Price Index for all Urban Consumers (CPI-U): U.S. city average, detailed expenditure categories. Various Years. http://www.bls.gov/cpi/tables.htm.
36. **Hansen RH.** *The current population survey: design and methodology*. Washington, DC: US Dept of Commerce, Bureau of the Census, 1995. Technical paper 40.
37. **Shopland DS,** Hartman AM, Gibson JT, *et al*. Cigarette smoking among US adults by state and region: Estimates from the Current Population Survey. *J Natl Cancer Inst* 1995;**88**:1748—58.
38. **Burns D,** Pierce JP. *Tobacco use in California, 1990-1991*. Sacramento, California: Department of Health Services, 1992.
39. **Pierce JP,** Goodman J, Gilpin EA, *et al*. *Technical report on analytic methods and approaches used in the Tobacco Use in California, 1990-1991 Report*. Sacramento, California: Department of Health Services, 1992.
40. **Wolfram S.** *The mathematica book*. 5th edn. Champaign, IL: Wolfram Media, 2003.
41. **Greene WH.** *LIMDEP, Version 8.0*. Plainview, NY: Econometric Software, Inc, 2002.
42. **Anon.** *SAS/STAT user's guide. Version 8*. Cary, NC: SAS Institute Inc, 1999.
43. **Oksendal B.** *Stochastic differential equations: An introduction with applications*. 5th edn. NY, NY: Springer, 2000:64—5.
44. **Johnson E,** Dominici F, Griswold M, *et al*. Disease cases and their medical costs attributable to smoking: an analysis of the national medical expenditure survey. *J Econom* 2002;**112**:139—45.
45. **McKelvey R,** Zavoina W. A statistical model for the analysis of ordinal level dependent variables. *J Math Sociol* 1975;**4**:103—20.
46. **Duan N.** Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc* 1983;**78**:605—10.
47. **Lee LL,** Gilpin EA, Pierce JP. Changes in the patterns of initiation of cigarette smoking in the United States: 1950, 1965, and 1980. *Cancer Epidemiol Biomarkers Prev* 1993;**2**:593—7.
48. **Pierce JP,** Lee L, Gilpin EA. Smoking initiation by adolescent girls, 1944 through 1988. An association with targeted advertising. *JAMA* 1994;**271**:608—11.
49. **Gilpin EA,** Lee L, Pierce JP. Smoking initiation rates in adults and minors: United States, 1944-1988. *Am J Epidemiol* 1994;**140**;535—43.
50. **US Centers for Disease Control and Prevention (US CDC).** Trends in smoking initiation among adolescents and young adults—United States, 1980—1989. *MMWR Morb Mortal Wkly Rep* 1995;**44**:521—5.
51. **Gilpin EA,** Pierce JP. Trends in adolescent smoking initiation in the United States: is tobacco marketing an influence? *Tob Control* 1997;**6**:122—7.
52. **Gilpin EA,** Pierce JP. Demographic differences in patterns in the incidence of smoking cessation: United States 1950-1990. *Ann Epidemiol* 2002;**12**:141—50.
53. **Messer K,** Pierce JP, Zhu SH, *et al*. The California Tobacco Control Program's effect on adult smokers: (1) Smoking cessation. *Tob Control* 2007;**16**:85—90.
54. **Bartlett JC,** Miller LS, Rice DP, *et al*. Medical expenditures attributable to cigarette smoking—United States, 1993. *MMWR Morb Mortal Wkly Rep* 1994;**43**:469—72.
55. **Miller VP,** Ernst C, Collin F. Smoking-attributable medical care costs in the USA. *Soc Sci Med* 1999;**48**:375—91.
56. **Viscusi WK.** The value of risks to life and health. *J Econ Lit* 1993;**31**:1912—46.
57. **Manning WG,** Keeler EB, Newhouse JP, *et al*. The taxes of sin: Do smokers and drinkers pay their way? *JAMA* 1989;**261**:1604—9.
58. **Manning WG,** Keeler EB, Newhouse JP, *et al*. *The costs of poor health habits*. Cambridge, Mass: Harvard University Press, 1991.
59. **Frederick S.** Valuing future life and future lives: a framework for understanding discounting. *J Econ Psychol* 2006;**27**:667—80.

Evaluation of the Economic Impact of California's Tobacco Control Program: A Dynamic Model Approach--Appendix 1: A theoretical distribution of the tobacco-exposure of cigarette smokers when cigarettes are assumed to be a fixed product.

Leonard S. Miller

> She rang under my feet like an empty Huntley & Palmer biscuit-tin kicked along a gutter; she was nothing so solid in make, and rather less pretty in shape, but I had expended enough hard work on her to make me love her. No influential friend would have served me better. She had given me a chance to come out a bit--to find out what I could do. No, I don't like work. I had rather laze about and think of all the fine things that can be done. I don't like work--no man does--but I like what is in the work,--the chance to find yourself. Your own reality--for yourself, not for others--what no other man can ever know. They can only see the mere show, and never can tell what it really means.
>
> Joseph Conrad,
> Heart of Darkness.

1. Introduction

The 25th Surgeon General's report proclaimed that "true scientific understanding of the health effects of tobacco" were achieved in the 20th century" (U.S. Department of Health and Human Services,1989). Major stepping stones in this understanding include Broders' (1920) link between tobacco use and lip cancer; Lombard and Doering's (1928) link between smoking and cancer; and Pearl's (1938) link between smoking and a shorter life span. By 1957, the national Study Group on Smoking and Health (1957) concluded that the relationship between smoking and lung cancer was causal. In a very short time, the Royal College of Physicians(1962) began to extended the adverse of effects of tobacco to a host of other diseases. The analysis reported here is a part of this unfolding of our scientific understanding of the relationship between smoking cigarettes and health. The act of smoking draws tobacco-toxin exposure into the lungs. The more packs-per-day smoked, the more years one smokes, the deeper one inhales, the higher the tars per cigarette the greater the amount of tobacco-exposure deposited in the body. Countervailing this smoker's ingestion process is a biological processes whose purpose is to expel extrinsic objects from the body. These two processes yield a resultant level of tobacco-toxin exposure in a smoker's body at any time. The purpose of this appendix is to derive the theoretical distri bution of these body resident tobacco-exposures for both current and former cigarette smokers, given particular smoking histories, and assuming that cigarettes have been a constant product (fixed level of tars per cigarette). I then present an outline of how knowledge of this distribution will be used to study the effect of smoking on health outcomes.

The tobacco-exposure distribution to be derived here arises from consideration of a stochastic dynamically described accumulation process. Formally, the process is described by a stochas tic differential equation system. Section 2 explicates the two equation accumulation process for current cigarette-smokers. The process is described and a summary of its solution is presented. The *Mathematica* program for the complete current-smoker solution is presented in Appendix 1 to this appendix. Section 3 parallels Section 2; its focus is on the stochastic dynamic tobacco-exposure accumulation process for former cigarette-smokers. Again, the process is described and a summary of its solution is presented. The *Mathematica* program for the complete former-smoker solution is presented in Appendix 2 to this appendix.

The analyses show that resident tobacco-exposures are normally distributed and that they have a heterogeneous variance. Closed form expressions for the expected value and the variance of these distributions are derived. The tobacco-exposure distri bution resulting from the analysis of a current cigarette-smoker is a function of five parameters; the distribution result ing from the analysis of a former-cigarette smoker is a function of one additional parameter. In section 4, I briefly outline how the parameters of these distributions are to be estimated and how the expressions of the moments of the tobacco-exposure distributions will be used to explain the effect of cigarette smoking on diseases caused by smoking and on health status.

Why engage in such an effort? As Lewontin (2003) suggests, the answer might come from consideration about the work such a theoretical distribution would provide? "Science and Simplicity", New York Review of Books, May 1, 2003, pp.39-42). Offering two classes of answers to this work question, he begins with, "Sometimes theoretical structures are nothing but calculat ing devices..." Indeed, that is precisely the principal use that will be made of the tobacco-exposure distribution derived here. As outlined in Section 4, in the chapters to follow, empirical exercises apply the expected value of this distribution to estimate the consequence of smoking on health outcomes. For each age and smoking history, I estimate the effect of smoking on the probability of death. Then, based on the estimated parameters, the distributions of exposure are predicted and used to estimate the effect of a particular smoking history on: (1) the probability of being currently treated for each of two classes of smoking related diseases (SRDs); (2) the distribution of self-reported health status for those not currently treated for SRDs; and on (3) the marginal cost of treatment. Recognizing that all of this has been done many times before, see Max () for a review of studies and for the range of obtained estimates, what can be learned from the effort to be constructed? First, because the full information about an individual's smoking history is not incorporated into existing cross sectionally based estimates of the consequence from smoking, the existing estimates contribute little to understanding the economic consequences of changes in smoking behavior. One of the principal benefits from such an effort is the knowledge gained from replacing the presently employed calculating devices with the calculating device that will be derived here.

In efforts to understand the economic costs of smoking, the usual "calculating device" allocates individuals into smoking history categories with current, former, and never-smoker being the categories most commonly employed. Since there is relatively little variation in the age when smoking is initiated, when an estimated fractional allocation of medical expenditures to smoking is age/gender and smoking status specific, estimates for current-smokers are probably reasonably accurate. However, the age when an individual quits smoking is not part of the existing specifications for former-smokers and the effect of this variation is not deducible from the estimates obtained for former-smokers. One consequence is that existing studies make almost no contribution to understanding the economic benefits arising from smoking cessation programs. The possible complications and the observed averages serve to further confuse. In fact, annual estimates of the level of expenditure or of the smoking attributable fraction for former smokers are often greater than annual estimates for current smokers. If one believes smoking is unambiguously detrimental to health, these findings are only understandable when recognition is made of the fact that smokers quit for different reasons. Some smokers become sick with a smoking related disease. Upon physician's advise, they quit. Their increased medical expenditures are associated with their smoking related disease. While these addi tional expenditures are appropriately allocatable to smoking, they do not reflect their newly initiated category of "former-smoker." Other smokers have a revelation about the importance of health on their own and their family's well-being and they quit as a means to an end. These individuals may also make greater use of discretionary medical expenditures, but these additional medical expenditures arise from a change in the quitter's demand for health services. These expenditures are not allocatable to tobacco-usage either.

In addition to the lack of a complete description (or of the major dimensions of) an individual's smoking history, another important dimension about smoking history that is omitted from most of the extant specifications focuses on dosage. If one is concerned with the economic benefits associated with smoking cessation programs, as the sponsors of this research are, one would want to be able to estimate the health benefits that accompany a reduction in some of the populations' daily consumption of cigarettes, and/or the health benefits associated with quitting for particularly critical periods of time, such as during the period that a woman is pregnant. Since dosage is not integrated into the current "computational devises", any benefits derived from dosage reduction are not addressable with the existing smoking computational devises.

Programs established to promote cessation in smoking behavior reap benefits when they reduce the smoking attributable phys ical outcomes requiring medical services. The economic evaluation of these programs require being able to estimate the reduction in medical services caused by smoking, given smoking history. To estimate such results is precisely the point of this effort. The theoretical "work" makes feasible estimates of the physical consequences of smoking on a full (perhaps fuller is a better way to put it) statement about an individual's smoking history.

Lewontin's (2003) second category for theory work is that it "...help(s) us "understand" a process whose outcome has been observed but whose dynamical details are not known from experiment or observation." For the topic at hand, Lewontin is discussing science as understood by scientists. There has been a great deal of work of late understanding the dynamic process of smoking induced cellular abnormality development (REF to latest surgeons general report). For example, exposure to a number of things in everyday life, from sunlight to cigarette smoke, can degrade DNA, but our bodies have developed mechanisms to mitigate this damage (Sarah Graham, Scientific American.com, News, September 03, 2003). Livneh and colleagues (Journal of the National Cancer Institute) studied the role of a repair enzyme known as OGG1 in preventing lung cancer. OGG1 deletes DNA parts that have been damaged by oxygen radicals. The theoretical tobacco-exposure distribution derived here neither makes, nor is intended to make any contribution to the scientific under standing of the dynamic biologic process leading to disease at the heart of the material under discussion. Here, the theory offered is merely a metaphor for the biological process.

However, the diminishment in health, the diseases and the deaths caused by smoking are the single most preventable public health hazard. The population of smokers is the group most in need of understanding the true consequences of smoking behavior. Some argue that the negative effects of smoking are common knowledge and smoking is an expression of rational consumer choice (REF). Yet studies show that few have good estimates about the details of these negative effects (REF). It could very well be that biological metaphors effectively transmit the essence of the biological process and contribute toward conveying a general sense of understanding. Computation devises based on believable/understandable metaphors are more likely to lead to believable results. The closer the metaphor's structure is to the true underlying scientific process, the more cred ible the computational device, the more general "understanding" can be derived from the computed knowledge. In addition to the work of understanding, belief serves as a source for judgements. Judges, juries, legislatures, public health bureaucrats, individuals, everyone of us needs to understand the health destructive consequences of smoking. It is toward this understanding that the work of the theory derived here is addressed.

2. A description of the tobacco-exposure accumulation process of current cigarette-smokers.

For current-smokers, the level of tobacco-exposure and the time rate of change of this level are denoted $tox_C[t]$ and $tox_C'[t]$, respectively. The first truth about the postulated tobacco-exposure accumulation process (alternatively, read equation describing the process) is an accounting identity that describes the time rate of change of the level of tobacco-exposure in the body of a current-smoker. At each moment t, the change in tobacco-exposure level is simply the difference between the tobacco-exposure ingested through smoking cigarettes and the tobacco-exposure purged from the body through the body's natural process to rid itself of foreign mate rial. It is helpful to think of the moment of time denoted by t as a day in the life of a current-smoker. In such a temporal framework, the ingestion of tobacco-exposure is given by the product of the exposure per pack of cigarettes, denoted by δ, and the number of packs of cigarettes the current-smoker smokes in a day, denoted by p. Thus, at time t, tobacco-exposure inges tion is given by δ p.

While it is clear that neither δ, nor p have necessarily been constant over the smoking history of any individual, to simplify I assume that both the exposure-per-pack, δ, and the packs-per-day smoked have been constant over an individual's smoking history.

At time t, the rate any current-smoker is able to purge him or herself of tobacco-exposures is denoted ν_c[t].

Given this notation describing ingestion and purging, the accounting identity describing the time rate of change of the accumulation of tobacco-exposures for a current-smoker at time t is given by equation [2.1],

[2.1] $tox_c'[t] = \delta\ p - \nu_c[t]$.

Factors that affect the purge rate constitute the second half of the tobacco-exposure accumulation process. The second truth (equation) about the exposure accumulation process describes the time rate of change of the tobacco-exposure purge rate. For current smokers, the time rate of change of the purge rate is denoted by $\nu_c'[t]$. I assume three factors effect the time rate of change of the purge rate. The first factor is aging. Aging causes a decline in all the body's somatic functioning; aging causes the body to be less efficient. More to the point, aging causes a decline in the body's ability to purge itself of exposure, including tobacco-exposures. Accordingly, I assume that the time rate of change in the purge rate declines at a constant rate with aging. This rate is denoted by $\gamma 0$.

Second, I assume that the level of accumulated tobacco-exposures in the body negatively affects the efficiency with which the body expels tobacco-exposures. Increasing tobacco-exposure levels cause a decline in the purge rate. A unit change in accumulated tobacco-exposures causes a $\gamma 1$ decline in the body's purge rate.

Third, I assume that individuals have different somatic reactions to tobacco-exposures. This variability in reaction (think allergic variability) is captured by incorporating a random process into the description of the time rate of change of an individual's purge rate.

The random elements incorporated into this description of the change in the purge rate are instantaneous. I assume that this instantaneous randomness is described by a Wiener process, which is a standardized Brownian motion process. If σ_c denotes the standard deviation of a Brownian motion process at time t, and if $d\omega_c$ denotes the time rate of change of a Wiener process at time t, $\sigma_c \, d\omega_c$ denotes the magnitude of the resolution of the instantaneous random shocks occurring at t. Incorporating these three effects results in the description of the time rate of change in the current smoker's purge rate given by equation [2.2],

[2.2] $\nu_c'[t] = -\gamma_0 - \gamma_1 \text{tox}[t] + \sigma_c \, d\omega_c$, with $\gamma_0>0$, $\gamma_1>0$, and $\sigma_c>0$.

There are two initial conditions on this accumulation system. First, the body's tobacco-exposure level when smoking is initiated, $\text{tox}_c[0]$, equals zero. This analysis assumes away second-hand smoke effects. Second, the initial purge rate, $\nu_c[0]$, is an unknown parameter of the problem and denoted ν_{c0}. These initial conditions are described by equation [2.3]

[2.3] $X0 = \begin{pmatrix} \text{tox}_c[0] \\ \nu_c[0] \end{pmatrix} = \begin{pmatrix} 0 \\ \nu_{c0} \end{pmatrix}$.

- **A Solution to the stochastic description of the current-smokers tobacco accumulation process**

Here, I present a solution for the dynamic stochastic differential equation system given by equations [2.1] through [2.3]. The solution yields knowledge about the distribution of tobacco-exposures and purge rate implied by the tobacco-exposure accumulation process described above. I begin by representing equations [2.1]-[2.3] in matrix form. For notational purposes, let $dX[t]$ denote a vector describing the first derivatives of the principal variables in the accumulation process of the current-cigarette smoker. Its first element is the time rate of change of the body's tobacco-exposures; its second element is the time rate of change of the purge rate,

$$[2.4] \quad dX[t] = \begin{pmatrix} tox_c'[t] \\ \nu_c'[t] \end{pmatrix}.$$

Let A denote the matrix relating the derivatives of the variables to their magnitudes,

$$[2.5] \quad A = \begin{pmatrix} 0 & -1 \\ -\gamma 1 & 0 \end{pmatrix}.$$

Let H denote a matrix of the constants in the description of the system,

$$[2.6] \quad H = \begin{pmatrix} p\,\delta \\ -\gamma_0 \end{pmatrix}.$$

Let K denote the matrix of constants multiplying the Wiener process for each equation in the system,

$$[2.7] \quad K = \begin{pmatrix} 0 \\ \sigma_c \end{pmatrix}.$$

X0, the initial values of the system, were given by equation [2.3] above. And let ϖ_t indicate a Wiener process at time t.

For a very different framework, Oksendal (2000) presented a solution to a problem with same mathematical structure as the stochastic differential system under analysis here.(Stochastic Differential Equations, An Introduction with Applications, 5th Edition,Springer,pp.64-65). If the exposure accumulation process unfolds between time t0 and time t, X[t], the magnitudes of the accumulated tobacco-exposures and the purge rate, is given by equation [2.8],

$$[2.8]\quad X[t] = \text{MatrixExp}[A\,(t - t_0)].X0 +$$

$$\text{MatrixExp}[A\,(t - t_0)].\text{MatrixExp}[-A\,(t - t_0)].K.\varpi[t] +$$

$$\text{MatrixExp}[A\,(t - t_0)].\int_{t_0}^{t} \text{MatrixExp}[-A\,s].H\,ds +$$

$$\text{MatrixExp}[A\,(t - t_0)].\int_{t_0}^{t} \text{MatrixExp}[-A\,s].A.K.\varpi[s]\,ds.$$

Note that MatrixExp[<arg>] evaluates the power series for the exponential function with ordinary powers replaced by the matrix <arg> (Wolfram, 1996,p.846). Recognize that $t_0 = 0$ and $tox_c[0] = 0$, the closed form solutions for the magnitude of tobacco-exposures and the purge rate, respectively, simplify to equations [2.9a] and [2.9b] (the *Mathematica* derivation of this answer is contained in Appendix 1 to this appendix).

$$[2.9a]\,tox_c[t] = \left\{\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1 + e^{t\sqrt{\gamma 1}}\right)^2 \gamma_0 + \left(-1 + e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta - \nu_{c0})\right)\right)\right\} +$$
$$\left\{-\left(\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(\omega_t - \omega_0)\right)\Big/\left(\sqrt{\gamma 1}\right)\right\}$$

$$[2.9b]\,\nu_c[t] = \left\{\frac{1}{\sqrt{\gamma 1}}\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0 + \sqrt{\gamma 1}\left(p\,\delta + \text{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta + \nu_{c0})\right)\right)\right\} +$$
$$\left\{\text{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(\omega_t - \omega_0)\right\}$$

The instantaneous random white noise of the Wiener processes integrates over time to a Normally distributed random variable with a heterogeneous variance. In these expressions, the term $(\omega_t - \omega_0)$ represents a Wiener process between time 0 and time t. Wiener processes of duration t have expected values equal to zero, and variances equal to σ^2t, where σ^2 in this analysis is given by σ_c^2.

For expository purposes, I have separated the expressions on the right-hand side (RHS) of equations [2.9a] and [2.9b], with curled brackets, "{}". The first term (on the RHS) of each of these expressions is the expected value of the magnitude described at time t, the body's tobacco-exposures at time t in equation [2.9a]; the body's purge rate at time t in equation [2.9b]. Each expected value describes its value after a smoking duration of length t, with a dosage of p packs of cigarettes smoked per day. The second term in each of the above expressions is a random variable, the difference between an individual's true tobacco-exposure level and his expected tobacco-exposure level ([2.9a]; the difference between an individual's true purge rate and his expected purge rate, given the smoking history described by t and p.

Given values for $\gamma 1$, $\gamma 0$, δ, ν_{c0}, and σ_c^2, the parameters of the tobacco-exposure accumulation system, tobacco-exposure and purge-rate levels are linear transformations of a normal random variable. Accordingly, the tobacco-exposure and purge rate levels are normally distributed. Since the variance of a constant times a random variable is equal to the product of the square of the constant and the variance of the random variable, the variance of the tobacco-exposure level, denoted by $\sigma^2{}_{toxc}[t]$, is given by equation [2.10a],

$$[2.10a] \quad \sigma^2{}_{toxc}[t] = \left\{\frac{t\,\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]^2 \sigma_c^2}{\gamma 1}\right\},$$

and the variance of the purge rate, denoted by $\sigma^2{}_{\nu' c}[t]$, is given by equation [2.10b],

$$[2.10b] \quad \sigma^2{}_{\nu c}[t] = \left\{t\,\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]^2 \sigma_c^2\right\}.$$

Sinh[<arg>] and Cosh[<arg>] are respectively the hyperbolic sine and hyperbolic cosine of the argument <arg> (Wolfram, 1996, p.731).

3. A description of the tobacco-exposure accumulation process of former cigarette-smokers.

The subscript c was used in the presentation of Section 2 above to denote the current smoker's model. Here, the subscript f will be used to denote the former-smokers model. Fundamentally, the model describing a former-smoker is based on the same mathematical structure as the model describing a current smoker. However two variable values are different, one parameter is allowed to be different, and the initial values of the magnitudes of the system, an individual's level of tobacco-exposures and his purge rate are different. We begin with the changes in the variable values. First, time for the former smoker is a count of the time of abstention, not time of smoking. The variable u denotes the length of time of abstinence. The time when a current-smoker quit is denoted by te (think e for end). When a current-smoker transitions to a former-smoker, t=te, u=0. For a former-smoker, the current-smoker's model applies for the period in which he smoked, that is, from t=0 to t=te. Take two individuals who are identical with the single exception that one stopped smoking at time te, 0 < te < t. The value t for the current smoker equals te + u for the former smoker. The second variable value that changes in the former-smoker's model is the dosage measure, the packs-per-day smoked. During the period of abstention zero packs-per-day are consumed. Thus, the ingestion described in the time rate of change of the tobacco-exposure accounting identity during the former-smoker's abstention period has a value of zero. Since the body's purging process continues to operate, similar to Section 2 above, equation [3.1], describes the time rate of change in the tobacco-exposure level without ingestion, but with a continuing purging process. Recall, time, denoted by u, measures the duration of the abstention period,

[3.1] $tox_f'[u] = -\nu_f[u]$.

The biological metaphor does not change when the individual changes smoking status. The time rate of change in the former-smoker's purge-rate, $\nu_f'[u]$, remains affected by the same three factors that affect the time rate of change in the purge-rate of a current-smoker. However, I allow for the possibility that the instantaneous random process during the period of abstinence can have a different standard deviation than the standard deviation operating during the period of cigarette consumption. Equation [3.2] specifies the time rate of change of the purge rate of a former smoker,

[3.2] $\nu_f'[u] = -\gamma_0 - \gamma_1 \, tox_f[u] + \sigma_f \, d\omega_u$, with $\gamma_0 > 0$, $\gamma_1 > 0$, and $\sigma_f > 0$.

The third difference between the current and former smoker's models is the description of the values of the magnitudes of the variables of the system when the system begins; the initial conditions of the system. The initial condition for a current smoker was the description of the stochastic differential system at time t equal to zero. The initial condition for a former-smoker is the description of the stochastic differential system at time t equal to te (u=0). When a current-smoker transitions to a former smoker, his expected level of tobacco-exposures and his expected purge-rate equals its value as a current smoker, see equations [2.9a] and [2.9b] above. Equation [3.3] describes the initial conditions at the moment a current-smoker transitions to a former smoker,

[3.3] X0 =

$$tox_f[u = 0] = \left\{\frac{1}{2\gamma 1}\left(e^{-te\sqrt{\gamma 1}}\left(\left(-1 + e^{te\sqrt{\gamma 1}}\right)^2 \gamma_0 + \left(-1 + e^{2\,te\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta - \nu_{c0})\right)\right)\right\} + \left\{-\frac{1}{\sqrt{\gamma 1}}\left(\mathrm{Sinh}\left[te\sqrt{\gamma 1}\right]\sigma_c\,(\omega_t - \omega_0)\right)\right\},$$

$$\nu_f[u=0] \quad \left\{\frac{1}{\sqrt{\gamma 1}}\left(-\text{Sinh}\left[te\sqrt{\gamma 1}\right]\gamma_0 + \sqrt{\gamma 1}\left(p\,\delta + \text{Cosh}\left[te\sqrt{\gamma 1}\right](-p\,\delta + \nu_{c0})\right)\right)\right\} + \left\{\text{Cosh}\left[te\sqrt{\gamma 1}\right]\sigma_c\,(\omega_t - \omega_0)\right\}.$$

Similar to equation [2.4], the vector of first derivatives of the variables of the system, evaluated at time u, is denoted by $dX[u]$. It is given by equation [3.4],

[3.4] $$dX[u] = \begin{pmatrix} tox'_f[u] \\ \nu'_f[u] \end{pmatrix}.$$

As in the description of a current-smoker, the matrix A is the matrix of coefficients relating the derivatives of the variables to their magnitudes, the matrix H is the matrix of constants relating the derivatives to their magnitudes. These matrices are the same in both systems. The matrix K is the matrix of constants multiplying the Wiener processes associated with each of the two equations in the system. The only difference between the matrix K in a current-smoker's model and the matrix K in a former-smoker's model is the subscript on the stan dard deviation of the white noise process. Equation [3.5] is the appropriate representation of the K matrix for a former-smoker's model.

[3.5] $$K = \begin{pmatrix} 0 \\ \sigma_f \end{pmatrix}.$$

The essence of the solution to this system is again given by equation [2.7]. Equations [3.6a] and [3.6b] present the simplified closed form solutions for the magnitude of the tobacco-exposures and the purge-rate, respectively, for a former-smoker. The full *Mathematica* derivation of this answer is contained in Appendix 2 to this appendix).

[3.6a] $tox_f[u,te] =$

$$\left\{\frac{1}{2\,\gamma 1}\left(e^{-(te+u)\sqrt{\gamma 1}}\left(-2\,e^{(te+u)\sqrt{\gamma 1}}\,p\,\sqrt{\gamma 1}\;\delta\,\mathrm{Sinh}\left[u\,\sqrt{\gamma 1}\,\right]+\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\;(p\,\delta-\nu_{c0})\right)\right)\right\}+$$

$$\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[u\,\sqrt{\gamma 1}\,\right]\sigma_f\,(\omega_u-\omega_0)\right)\right\}+\left\{\mathrm{Cosh}\left[u\,\sqrt{\gamma 1}\,\right]\epsilon_{toxc}[te]\right\}$$

$$+\left\{-\frac{1}{\sqrt{\gamma 1}}\left(\mathrm{Sinh}\left[u\,\sqrt{\gamma 1}\,\right]\epsilon_{\nu c}[te]\right)\right\}.$$

[3.6b] $\nu_f[u,te] =$

$$\left\{\frac{1}{\sqrt{\gamma 1}}\left(p\,\sqrt{\gamma 1}\;\delta\,\mathrm{Cosh}\left[u\,\sqrt{\gamma 1}\,\right]-\mathrm{Sinh}\left[(te+u)\,\sqrt{\gamma 1}\,\right]\gamma_0+\sqrt{\gamma 1}\;\mathrm{Cosh}\left[(te+u)\,\sqrt{\gamma 1}\,\right](-p\,\delta+\nu_{c0})\right)\right\}+$$

$$\left\{\mathrm{Cosh}\left[u\,\sqrt{\gamma 1}\,\right]\sigma_f\,(\omega_u-\omega_0)\right\}+$$

$$\left\{-\sqrt{\gamma 1}\;\mathrm{Sinh}\left[u\,\sqrt{\gamma 1}\,\right]\epsilon_{toxc}[te]\right\}$$

$$\left\{\mathrm{Cosh}\left[u\,\sqrt{\gamma 1}\,\right]\epsilon_{\nu c}[te]\right\}.$$

Completely analogous to the solution equations in Section 2 above, I have separated the expressions on the RHS of equations [3.6a] and [3.6b] with curled brackets, "{}". The first expression in each equation is the expected value of the respective magnitude, given estimated parameter values and values for how long the former-smoker smoked, te, and how long a respondent abstained from smoking, u. The second expressions are the random errors resulting from the Wiener process for a former-smoker. The third expressions are the consequence of the deviation between the true tobacco-exposure level and the expected tobacco-exposure level when the current-smoker transitions to a former-smoker. This deviation was then acted upon by the former-smoker's stochastic dynamic tobacco-exposure accumulation process. The fourth (last) expressions are the consequence of the deviation between the true purge-rate level and the expected purge-rate level, again when the current-smoker transitions to a former-smoker. This deviation, again, was subsequently acted upon by the former-smoker's stochastic dynamic tobacco-exposure accumulation process.

Given values for $\gamma 1$, $\gamma 0$, δ, ν_{c0}, σ_c^2, and σ_f^2, the parameters of a former-smoker's tobacco-exposure accumulation system, the tobacco-exposure and purge-rate levels are linear transformations of normal random variables. Accordingly, these tobacco-exposure and purge-rate levels are normally distributed. The variances of these respective levels are denoted $\sigma^2{}_{\text{toxf}}[u, te]$, and $\sigma^2{}_{\text{toxf}}[u]$. Again, for identification purposes, the right-hand terms of equations [3.7a] and [3.7b], expressions for these variances, are separated with curly brackets. The first term on the RHS of equations [3.7a] and [3.7b], respectively, is the variance of the Wiener process acting on the exposure level and purge rate, respectively. The second terms are the variance induced by the initial difference between the true and expected tobacco-exposure levels. The third terms are the variance induced by the initial difference between the true and expected purge rate,

$$[3.7a]\quad \sigma^2{}_{\text{toxf}}[u, te] = \left\{\frac{u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_f^2}{\gamma 1}\right\} + \left\{\frac{1}{\gamma 1}\left(te\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2 \text{Sinh}\left[te\sqrt{\gamma 1}\right]^2 \sigma_c^2\right)\right\} + \left\{\frac{1}{\gamma 1}\left(te\,\text{Cosh}\left[te\sqrt{\gamma 1}\right]^2 \text{Sinh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_c^2\right)\right\},$$

$$[3.7b]\quad \sigma^2{}_{\nu f}[u, te] = \left\{u\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_f^2\right\} + \left\{te\,\text{Sinh}\left[te\sqrt{\gamma 1}\right]^2 \text{Sinh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_c^2\right\} + \left\{te\,\text{Cosh}\left[te\sqrt{\gamma 1}\right]^2 \text{Cosh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_c^2\right\}.$$

4. Towards achieving a calculation device based on these theo-

retic tobacco-exposure distributions.

Incorporating the theoretic tobacco-exposure distribution into behavioral specifications

All of the empirical analyses to follow in this study estimate the effect of smoking on health and cost related outcomes. The derived distribution of accumulated tobacco-exposures is used as a tobacco-exposure index to portray a history of smoking behavior. In all of the analyses, the propensity for the occurrence of the health or cost outcome under analysis is set equal to parameter weighted measures of the observation's age and his accumulated tobacco-exposure level. The coefficient weighted expected level of tobacco-exposures are a RHS variable in the specification of the expected propensity under analysis; the coefficient weighted difference between the true and expected level of tobacco-exposures is, in effect, part of the random error of the model.

The feasibility of this plan requires (1) recognition that at least a part of the random error term in each model has a normal distribution; and (2) estimates of the parameters of the exposure accumulation model, so that the expected exposure level and its variance can be estimated and incorporated into the analyses. In the next appendix to follow a survival model is proposed that has the appropriate error distributions. I call it a Normally Distributed Survival Model. The model is difficult to estimate and starting values are quite critical. In Chapter Three I suggest methods that one can employ to achieve starting values to use with estimating the proposed Probit Survival Model. Chapter Four presents a simple use of the Probit Survival Model. The propensity to die is estimated for two classes of never-smokers, those who have had no higher education and those who have had some higher education. Higher education is being used as a proxy for social class, which is a proxy for access to health care services. Chapter Five presents estimates of the propensity to die for current smokers based on the Probit Survival Model specification. The estimates for the never-smokers are used to characterize the effect of age and health care access. The model estimates the parameters of a simplified version of the current-smokers model. Chapter Six presents estimates of the propensity to die for former-smokers, again based on the Probit Survival Model and again with estimates for the never-smokers characterizing the effect of age and health care access. The somewhat simplified version of the current smokers model is extended to represent the former smokers.

With the obtained tobacco-exposure parameter estimates, given an individual's smoking history, it is possible to calculate expected levels of tobacco-exposures and standard deviations in those levels. The next three analyses are based on these calculated values. In Chapter Seven, I estimate the probability of being currently treated for a class of smoking related diseases (lung cancer, esophageal cancer, and chronic obstructive pulmonary disease) that compared to never-smokers have a high relative risk due to smoking. In Chapter Eight the exercise is repeated for the class of smoking related diseases (all of the remaining smoking related diseases) that have a relatively low relative risk due to smoking. In Chapter Nine, for a sample that is not currently treated for a smoking related disease I estimate the effect of smoking on self-reported poor health status.

These analyses allow comparisons between a wider selection of smoking histories than is usually made. They also allow correction for sample selection due to death, obtaining a more accurate (and unbiased) estimate of the relative risk of current treatment statuses induced by various intensities of smoking behavior. The set of analyses also allows correction for the additional contributor to sample selection, current treatment status, to yield an unbiased estimate of the effect of smoking on the probability distribution of self-reported poor health status.

Chapter Ten estimates a medical expenditure model for people who are not currently treated for a smoking related disease. Again, smoking status is based on the distribution of tobacco-exposures. The specification is able to differentiate among the tobacco-exposure effect of the demand for medical services and any change in demand for medical services that accompany a shift in smoking status from current to former.

Chapter Eleven uses all of the derived models to estimate the expected deaths, the distribution of expected smoking related disease treatment, and the distribution of the self-reported health status of California's population for a considerable period into the future. Chapter Eleven estimates the smoking prevalence rates and quit rates that California might have had, if its Tobacco Prevention Program had not existed, and Chapter Twelve compares the simulations performed on California's population to the simulations performed on California's population without its tobacco prevention program to estimate the economic and physical benefits from the program over the decade of the nineteen nineties.

Appendix 1.1: Solution for Current Smoker

The model to be solved, representing the generalized tobacco-exposure accumulation model for the current-smoker is as follows:

$tox_c'[t] = \delta\ p - \nu_c[t]$,
$\nu_c'[t] = -\gamma_0 - \gamma_1\ tox[t] + \sigma_c\ d\omega_t$,

where: tox[t]= accumulated level of tobacco originating exposures in the body;
$\nu_c[t]$ = body purge rate;
γ_0= drift rate in the body's purging ability due to aging;
γ_1 = drift rate in the body's purging ability due to exposure accumulation;
σ_c= standard deviation of the Wiener stochastic process;
$d\omega_t$= Wiener process at time t;
δ = density of exposures per pack of cigarettes;
p = packs of cigarettes smoked per day.

The vector of first derivatives evaluated at time t is dX[t],

```
dX[t] = {{tox_c'[t]}, {ν_c'[t]}}
```

$\{\{tox_c'[t]\}, \{\nu_c'[t]\}\}$

```
MatrixForm[%]
```

$$\begin{pmatrix} tox_c'[t] \\ \nu_c'[t] \end{pmatrix}$$

A is the matrix relating the derivatives of the variables in the model to the variables in the model,

```
A = {{0, -1}, {-γ1, 0}}
```

$\{\{0, -1\}, \{-\gamma 1, 0\}\}$

```
MatrixForm[%]
```

$$\begin{pmatrix} 0 & -1 \\ -\gamma 1 & 0 \end{pmatrix}$$

H is a matrix of constants,

H = {{δ p}, {-γ0}}

$\{\{p\,\delta\}, \{-\gamma_0\}\}$

K is a matrix of constants multiplying the Wiener processes associated with each equation,

K = {{0}, {σc }}

$\{\{0\}, \{\sigma_c\}\}$

X0 denotes the matrix of starting values,

X0 = {{toxc0}, {νc0}}

$\{\{tox_{c0}\}, \{\nu_{c0}\}\}$

W indicates the Wiener process (Standardized Brownian Motion)

W = {{ω}}

$\{\{\omega\}\}$

The next four operations are the four parts of the solution.

MatrixExp[A (t – t0)].X0

$$\left\{\left\{\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)tox_{c0}-\frac{e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}}{2\sqrt{\gamma 1}}\right\},\right.$$
$$\left.\left\{-\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\sqrt{\gamma 1}\;tox_{c0}+\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}\right\}\right\}$$

MatrixExp[A (t – t0)]. MatrixExp[–A (t – t0)].K.W

$$\left\{\{0\},\ \left\{\left(-\frac{1}{4}\, e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2+\frac{1}{4}\, e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2\right)\omega\,\sigma_c\right\}\right\}$$

MatrixExp[A (t – t0)].
Integrate[MatrixExp[–A s].H , {s, t0, t}]

$$\left\{\left\{-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]-\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0\right)\right)+\frac{1}{\gamma 1}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)\right\},\right.$$

$$\left.\left\{\frac{1}{2}\, e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]-\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\frac{\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0}{\sqrt{\gamma 1}}\right)-\frac{1}{\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)\right\}\right\}$$

```
MatrixExp[A (t - t0)].
 Integrate[MatrixExp[-A s] . A . K . W, {s, t0, t}]
```

$$\left\{\left\{-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right)+\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right)\right\},\left\{\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c-\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right\}\right\}$$

The solution of the magnitudes is the sum of the four parts given directly above,

```
X[t] = %9 + %10 + %11 + %12
```

$$\left\{\left\{\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\mathrm{tox}_{c0}-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]-\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0\right)\right)+\frac{1}{\gamma 1}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\mathrm{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)-\frac{e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}}{2\sqrt{\gamma 1}}-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right)+\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]+\mathrm{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right)\right\},$$

$$\left\{-\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\sqrt{\gamma 1}\,tox_{c0}+\right.$$
$$\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\text{Cosh}\left[t\sqrt{\gamma 1}\right]-\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\right.$$
$$\left.\frac{\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0}{\sqrt{\gamma 1}}\right)-$$
$$\frac{1}{\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\right.$$
$$\left.\left(p\sqrt{\gamma 1}\,\delta\,\text{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)+$$
$$\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}+$$
$$\left(-\frac{1}{4}\,e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2+\right.$$
$$\left.\frac{1}{4}\,e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2\right)\omega\,\sigma_c+$$
$$\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega\left(-\text{Cosh}\left[t\sqrt{\gamma 1}\right]+\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c-$$
$$\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\omega$$
$$\left.\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\right\}\}$$

At time t, ω would denote the difference between the Wiener process at t and the Wiener process at 0

%13 /. $\omega \to (W_t - W_0)$

$$\left\{\left\{\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)tox_{c0}-\frac{1}{2\sqrt{\gamma 1}}\right.\right.$$
$$\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\text{Cosh}\left[t\sqrt{\gamma 1}\right]-\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\right.\right.$$
$$\left.\left.\frac{1}{\sqrt{\gamma 1}}\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0\right)\right)+$$
$$\frac{1}{\gamma 1}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\right.$$
$$\left.\left(p\sqrt{\gamma 1}\,\delta\,\text{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)-$$
$$\frac{e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}}{2\sqrt{\gamma 1}}-$$

$$\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\right.$$
$$\left.\left(-\text{Cosh}\left[t\sqrt{\gamma 1}\right]+\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)\right)+$$
$$\frac{1}{2\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\right.$$
$$\left.\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)\right)\Big\},$$
$$\Big\{-\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\sqrt{\gamma 1}\,\text{tox}_{c0}+$$
$$\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(p\,\delta\left(\text{Cosh}\left[t\sqrt{\gamma 1}\right]-\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)+\right.$$
$$\left.\frac{\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\gamma_0}{\sqrt{\gamma 1}}\right)-$$
$$\frac{1}{\sqrt{\gamma 1}}\left(e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t-t_0)\right]\right.$$
$$\left.\left(p\sqrt{\gamma 1}\,\delta\,\text{Cosh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]-\text{Sinh}\left[\frac{1}{2}\sqrt{\gamma 1}\,(t+t_0)\right]\gamma_0\right)\right)+$$
$$\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\nu_{c0}+$$
$$\left(-\frac{1}{4}e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2+\right.$$
$$\left.\frac{1}{4}e^{-2\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)^2\right)\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)+$$
$$\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)\left(-\text{Cosh}\left[t\sqrt{\gamma 1}\right]+\text{Cosh}\left[\sqrt{\gamma 1}\,t_0\right]\right)$$
$$\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)-\frac{1}{2}e^{-\sqrt{\gamma 1}\,(t-t_0)}\left(-1+e^{2\sqrt{\gamma 1}\,(t-t_0)}\right)$$
$$\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]+\text{Sinh}\left[\sqrt{\gamma 1}\,t_0\right]\right)\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)\Big\}\Big\}$$

Setting the starting values for tox_c and t equal to their known values, the solution equations become:

{{tox$_c$[t]}, {ν_c[t]}} = %14 /. {tox$_{c0}$ → 0, t$_0$ → 0}

$$\left\{\left\{\frac{1}{\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)\mathrm{Sinh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]-\mathrm{Sinh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]\gamma_0\right)\right)-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)\left(p\,\delta\left(-1+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\right)-\frac{\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0}{\sqrt{\gamma 1}}\right)\right)-\frac{e^{-t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)\nu_{c0}}{2\sqrt{\gamma 1}}-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)\left(1-\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\right)\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)\right)-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)\right)\right\},$$

$$\left\{-\frac{1}{\sqrt{\gamma 1}}\left(e^{-t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)\mathrm{Sinh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]-\mathrm{Sinh}\left[\frac{t\sqrt{\gamma 1}}{2}\right]\gamma_0\right)\right)+\frac{1}{2}e^{-t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)\left(p\,\delta\left(-1+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\right)-\frac{\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0}{\sqrt{\gamma 1}}\right)+\frac{1}{2}e^{-t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)\nu_{c0}+\left(-\frac{1}{4}e^{-2t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)^2+\frac{1}{4}e^{-2t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)^2\right)\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)+\frac{1}{2}e^{-t\sqrt{\gamma 1}}\left(1+e^{2t\sqrt{\gamma 1}}\right)\left(1-\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\right)\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)+\frac{1}{2}e^{-t\sqrt{\gamma 1}}\left(-1+e^{2t\sqrt{\gamma 1}}\right)\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)\right\}\right\}$$

FullSimplify[%15]

$$\left\{\left\{\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0}+\sigma_c\,(\{\{\omega\}\}_0-\{\{\omega\}\}_t))\right)\right)\right\},\right.$$
$$\left.\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\text{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0}+\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t))\right)\right)\right\}\right\}$$

Note that in the expression for $tox_c[t]$the Wiener process is going backward and needs to be changed. First identify the coefficient on the Wiener process. What follows is the expected value of $tox_c[t]$,

%16[[1]] /. ({{ω}}$_0$ - {{ω}}$_t$) → 0

$$\left\{\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)\right\}$$

The difference between the whole expression for $tox_c[t]$ and the expected value of $tox_c[t]$ is the random error in the expression for $tox_c[t]$

%16[[1]] - %17

$$\left\{-\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\right.$$
$$\left.\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0}+\sigma_c\,(\{\{\omega\}\}_0-\{\{\omega\}\}_t))\right)\right)\right\}$$

FullSimplify[%]

$$\left\{\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(\{\{\omega\}\}_0-\{\{\omega\}\}_t)}{\sqrt{\gamma 1}}\right\}$$

Multiply this term by -1 and change the signs in the Wiener processes.

-1 %19

$$\left\{-\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(\{\{\omega\}\}_0-\{\{\omega\}\}_t\right)}{\sqrt{\gamma 1}}\right\}$$

$$\left\{-\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)}{\sqrt{\gamma 1}}\right\}$$

$$\left\{-\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)}{\sqrt{\gamma 1}}\right\}$$

Accordingly, (from %17 and %19), the following is the correct statement for $\text{tox}_c[t]$, (see Out[26] below),

$$\left\{\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\left(p\,\delta-\nu_{c0}\right)\right)\right)\right\}+$$

$$\left\{-\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)}{\sqrt{\gamma 1}}\right\}$$

$$\left\{\frac{1}{2\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\left(p\,\delta-\nu_{c0}\right)\right)\right)-\frac{\text{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_t\right)}{\sqrt{\gamma 1}}\right\}$$

Returning to the expression for $\nu_c[t]$, its expected value is:

%16[[2]] /. ({{ω}}$_t$ − {{ω}}$_0$) → 0

$$\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\text{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\text{Cosh}\left[t\sqrt{\gamma 1}\right]\left(-p\,\delta+\nu_{c0}\right)\right)\right)\right\}$$

The difference between the whole expression for $\nu_c[t]$ and the expected value of $\nu_c[t]$ is the random error in the expression for $\nu_c[t]$,

%16[[2]] - %23

$$\left\{-\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0}+\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t))\right)\right)\right\}$$

FullSimplify[%]

$$\left\{\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)\right\}$$

The expression for the solution to the problem is as follows:

{%22, %23 + %25}

$$\left\{\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)-\frac{\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)}{\sqrt{\gamma 1}}\right\},\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_t)\right\}\right\}$$

We can express this solution as the sum of its expected values and its error. The expected value results from setting the random error equal to zero. Taking the result from above, (where in the statement to follow the first parts are the expected values and the second parts are the error terms,

$$\left\{\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^{2}\gamma_{0}+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[t]\right\},\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_{0}+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu_c}[t]\right\}\right\}$$

— *General::spell1 : Possible spelling error: new symbol name "toxc" is similar to existing symbol "tox". More…*

$$\left\{\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^{2}\gamma_{0}+\left(-1+e^{2t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[t]\right\},\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_{0}+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu_c}[t]\right\}\right\}$$

The random errors are given by

$$\left\{\left\{-\frac{\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_{c}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{t})}{\sqrt{\gamma 1}}\right\},\left\{\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_{c}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{t})\right\}\right\}$$

$$\left\{\left\{-\frac{\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_{c}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{t})}{\sqrt{\gamma 1}}\right\},\left\{\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_{c}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{t})\right\}\right\}$$

The following expressions (Out[27] evaluated at t->te describe the starting value for $tox_f[0]$, $\nu_f[0]$, in equation [1.14]

$$\left\{\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\;(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[t]\right\},\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu_c}[t]\right\}\right\}$$

$$\left\{\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\;(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[t]\right\},\left\{\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu_c}[t]\right\}\right\}$$

The variances of $tox_c[t]$ and $\nu_c[t]$ are the variance of these error terms. This expression is the variance squared times t.

$$\left\{\left\{\left(-\frac{1}{\sqrt{\gamma 1}}\left(\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,\mathrm{Sqrt}[t]\right)\right)^2\right\},\left\{\left(\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]\sigma_c\,\mathrm{Sqrt}[t]\right)^2\right\}\right\}$$

$$\left\{\left\{\frac{t\,\mathrm{Sinh}\left[t\sqrt{\gamma 1}\right]^2\sigma_c^2}{\gamma 1}\right\},\left\{t\,\mathrm{Cosh}\left[t\sqrt{\gamma 1}\right]^2\sigma_c^2\right\}\right\}$$

This is a check to insure that the expected value of $tox_{c0}=0$.

$$\left\{\frac{1}{2\,\gamma 1}\left(e^{-t\sqrt{\gamma 1}}\left(\left(-1+e^{t\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,t\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\;(p\,\delta-\nu_{c0})\right)\right)\right\}\;/.\;t\to 0$$

$$\{0\}$$

Appendix 1.2:Solution for a Former-Smoker

The model to be solved represents the generalized tobacco-exposure accumulation model for a former smoker is as follows:

$tox_f'[u] = -\nu_f[u]$,

$\nu_f'[u] = -\gamma_0 - \gamma_1\, tox[u] + \sigma_f d\omega_u$,

where:

$\nu_f[u]$ = body purge rate for former smokers at time u into abstinence;

$tox_f[u]$= the accumulated level of tobacco originating exposures in the body at time u;

γ_0= drift rate in the body's purging ability due to aging;

γ_1 = drift rate in the body's purging ability due to exposure accumulation;

σ_f= standard deviation of the Wiener stochastic process;

$d\omega_u$= Wiener process at time u;

δ = density of exposures per pack;

p = packs of cigarettes smoked per day.

The vector of first derivatives evaluated at time u is $dX[u]$

```
dX[u] = {{tox_f'[u]}, {ν_f'[u]}}
```

$\{\{tox'_f[u]\}, \{\nu'_f[u]\}\}$

```
MatrixForm[%]
```

$$\begin{pmatrix} tox'_f[u] \\ \nu'_f[u] \end{pmatrix}$$

A is the matrix relating the variable to its derivative.

```
A = {{0, -1}, {-γ1, 0}}
```

```
{{0, -1}, {-γ1, 0}}
```

```
MatrixForm[%]
```

$$\begin{pmatrix} 0 & -1 \\ -\gamma 1 & 0 \end{pmatrix}$$

H is a matrix of constants relating the derivatives to the magnitudes of the variables.magnitudes

```
H = {{ 0}, {-γ0}}
```

$\{\{0\}, \{-\gamma_0\}\}$

K is a matrix of constants multiplying the Wiener processes associated with each equation.

```
K = {{0}, {σf }}
```

$\{\{0\}, \{\sigma_f\}\}$

X0 denotes the matrix of starting values.

```
X0 = {{toxf0}, {νf0}}
```

$\{\{tox_{f0}\}, \{\nu_{f0}\}\}$

W indicates the Wiener process (Brownian Motion).

```
W = {{ω}}
```

$\{\{\omega\}\}$

The next four operations are the four parts of the solution

MatrixExp[A (u - u0)].X0

$$\left\{\left\{\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\,tox_{f0}+\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}\right)\,\nu_{f0}\right\},\left\{\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\,tox_{f0}+\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\,\nu_{f0}\right\}\right\}$$

MatrixExp[A (u - u0)]. MatrixExp[-A (u - u0)].K.W

$$\left\{\left\{\left(\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}\right)+\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(-\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}+\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}\right)\right)\,\omega\,\sigma_f\right\},\left\{\left(\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)^2+\left(-\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}+\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\,\sqrt{\gamma 1}}\right)\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\right)\,\omega\,\sigma_f\right\}\right\}$$

```
MatrixExp[A (u - u0)].
 Integrate[MatrixExp[-A s].H , {s, u0, u}]
```

$$\left\{\left\{-\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\,\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\,\gamma 1}\right)\gamma 0-\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma 0\right\},\left\{-\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma 0-\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\,\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\,\gamma 1}\right)\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}\right)\gamma 0\right\}\right\}$$

MatrixExp[A (u − u_0)].
Integrate[MatrixExp[−A s] . A . K . W, {s, u_0, u}]

$$\left\{\left\{\left(-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\omega\,\sigma_f+\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(-\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}+\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\omega\,\sigma_f\right\},\right.$$

$$\left.\left\{\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)\omega\,\sigma_f+\left(-\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}+\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}\right)\omega\,\sigma_f\right\}\right\}$$

The solution of the magnitudes is the sum of the four parts given directly above

X[t] = %9 + %10 + %11 + %12

$$\left\{\left\{\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\mathrm{tox}_{f0}-\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\right.\right.$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\,\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\,\gamma 1}\right)\gamma_0 -$$

$$\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma_0 +$$

$$\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\nu_{f0} +$$

$$\left(\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\right.\right.$$

$$\left.\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)+\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left.\left(-\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}+\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\right)\omega\,\sigma_f +$$

$$\left(-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)$$

$$\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\omega\,\sigma_f +$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(-\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}+\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)$$

$$\left.\omega\,\sigma_f\right\},$$

$$\Bigg\{\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\mathrm{tox}_{f0}-$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma_0-$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\,\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\,\gamma 1}\right)$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\gamma_0+$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\nu_{f0}+$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)$$

$$\omega\,\sigma_f+\Bigg(\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\right)^2+$$

$$\left(-\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}+\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\Bigg)\,\omega\,\sigma_f+$$

$$\left(-\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}+\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}-\frac{1}{2}\,e^{\sqrt{\gamma 1}\,(u-u_0)}\,\sqrt{\gamma 1}\right)\omega\,\sigma_f\Bigg\}\Bigg\}$$

At time u, ω denotes the difference between the Wiener process at u and the Wiener process at u=0

%13 /. ω → (W_u − W_0)

$$\Bigg\{\Bigg\{\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\mathrm{tox}_{f0}-\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\gamma 1}\right)\gamma_0-\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma_0+\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\nu_{f0}+\left(\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)+\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\left(-\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}+\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\right)\sigma_f\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_u\right)+\left(-\frac{1}{2}e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)$$

$$\left(\frac{e^{-\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}-\frac{e^{\sqrt{\gamma 1}\,(u-u_0)}}{2\sqrt{\gamma 1}}\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+$$

$$\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(-\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}+\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)$$

$$\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\Big\},$$

$$\Big\{\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}-\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}\right)\mathrm{tox}_{f0}-$$

$$\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)}{2\sqrt{\gamma 1}}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(-1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\sqrt{\gamma 1}}\right)\gamma_0-$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\gamma 1}-\frac{e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)}{2\gamma 1}\right)$$

$$\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}-\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\sqrt{\gamma 1}\right)\gamma_0+$$

$$\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)\nu_{f0}+$$

$$\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)$$

$$\left(-\frac{1}{2}e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)+\frac{1}{2}e^{-\sqrt{\gamma 1}\,u_0}\left(1+e^{2\sqrt{\gamma 1}\,u_0}\right)\right)$$

$$\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+$$

$$\Bigg(\left(\frac{1}{2}e^{-\sqrt{\gamma 1}\,(u-u_0)}+\frac{1}{2}e^{\sqrt{\gamma 1}\,(u-u_0)}\right)^2+$$

$$\left(-\frac{e^{-\sqrt{\gamma1}\,(u-u_0)}}{2\sqrt{\gamma1}}+\frac{e^{\sqrt{\gamma1}\,(u-u_0)}}{2\sqrt{\gamma1}}\right)\left(\frac{1}{2}\,e^{-\sqrt{\gamma1}\,(u-u_0)}\sqrt{\gamma1}-\right.$$

$$\left.\left.\frac{1}{2}\,e^{\sqrt{\gamma1}\,(u-u_0)}\sqrt{\gamma1}\right)\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+$$

$$\left(-\frac{e^{-u\sqrt{\gamma1}}\left(-1+e^{2u\sqrt{\gamma1}}\right)}{2\sqrt{\gamma1}}+\frac{e^{-\sqrt{\gamma1}\,u_0}\left(-1+e^{2\sqrt{\gamma1}\,u_0}\right)}{2\sqrt{\gamma1}}\right)$$

$$\left(\frac{1}{2}\,e^{-\sqrt{\gamma1}\,(u-u_0)}\sqrt{\gamma1}-\frac{1}{2}\,e^{\sqrt{\gamma1}\,(u-u_0)}\sqrt{\gamma1}\right)$$

$$\left.\left.\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right\}\right\}$$

%14 /. {$u_0 \to 0$}

$$\left\{\left\{\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)tox_{f0}-\right.\right.$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-\frac{1}{\gamma1}+\frac{e^{-u\sqrt{\gamma1}}\left(1+e^{2u\sqrt{\gamma1}}\right)}{2\,\gamma1}\right)\gamma_0-$$

$$\frac{e^{-u\sqrt{\gamma1}}\left(-1+e^{2u\sqrt{\gamma1}}\right)\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\gamma_0}{2\sqrt{\gamma1}}+$$

$$\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\nu_{f0}+$$

$$\left(\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)+\right.$$

$$\left.\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\right)\sigma_f$$

$$(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\left(1-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)\right)$$

$$\left(\frac{e^{-u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}-\frac{e^{u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)-$$

$$\frac{1}{2\sqrt{\gamma 1}}\left(e^{-u\sqrt{\gamma 1}}\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)\right.$$

$$\left.\left(-1+e^{2u\sqrt{\gamma 1}}\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right)\Big\},$$

$$\Big\{\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}\,e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\mathrm{tox}_{f0}-$$

$$\left(-\frac{1}{\gamma 1}+\frac{e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)}{2\,\gamma 1}\right)$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}\,e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\gamma_0-$$

$$\frac{e^{-u\sqrt{\gamma 1}}\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)\left(-1+e^{2u\sqrt{\gamma 1}}\right)\gamma_0}{2\sqrt{\gamma 1}}+$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)\nu_{f0}+\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)$$

$$\left(1-\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\left(1+e^{2u\sqrt{\gamma 1}}\right)\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+$$

$$\left(\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)^2+\left(-\frac{e^{-u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}\right)\right.$$

$$\left.\left(\frac{1}{2}\,e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}\,e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\right)$$

$$\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)-\frac{1}{2\sqrt{\gamma1}}\left(e^{-u\sqrt{\gamma1}}\left(-1+e^{2u\sqrt{\gamma1}}\right)\left(\frac{1}{2}e^{-u\sqrt{\gamma1}}\sqrt{\gamma1}-\frac{1}{2}e^{u\sqrt{\gamma1}}\sqrt{\gamma1}\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right)\}\}$$

Set the starting values for tox_f[u,te] and ν_f[u,te] equal to the values the individual had when he/she stopped smoking

$$\{\{\mathbf{tox_f[u,te]}\},\{\nu_f[u,te]\}\}=\%15\;/.\;\Big\{tox_{f0}\to\Big\{\frac{1}{2\gamma1}\left(e^{-te\sqrt{\gamma1}}\left(\left(-1+e^{te\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2te\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[te]\Big\},$$

$$\nu_{f0}\to\Big\{\frac{1}{\sqrt{\gamma1}}\left(-\mathrm{Sinh}\left[te\sqrt{\gamma1}\right]\gamma_0+\sqrt{\gamma1}\left(p\,\delta+\mathrm{Cosh}\left[te\sqrt{\gamma1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu c}[te]\Big\}\Big\}$$

— *General::spell1 :*
Possible spelling error: new symbol name "toxc" is similar to existing symbol "tox".

— *General::spell1 :*
Possible spelling error: new symbol name "νc" is similar to existing symbol "ν".

$$\Big\{\Big\{\Big\{-\left(\frac{1}{2}e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-\frac{1}{\gamma1}+\frac{e^{-u\sqrt{\gamma1}}\left(1+e^{2u\sqrt{\gamma1}}\right)}{2\gamma1}\right)\gamma_0-$$

$$\frac{e^{-u\sqrt{\gamma1}}\left(-1+e^{2u\sqrt{\gamma1}}\right)\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\gamma_0}{2\sqrt{\gamma1}}+$$

$$\left(\left(\frac{1}{2}e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)+\right.$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\right)\sigma_f$$

$$(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\left(1-\frac{1}{2}\,e^{-u\sqrt{\gamma1}}\left(1+e^{2u\sqrt{\gamma1}}\right)\right)$$

$$\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)-$$

$$\frac{1}{2\sqrt{\gamma1}}\left(e^{-u\sqrt{\gamma1}}\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-1+e^{2u\sqrt{\gamma1}}\right)\right.$$

$$\left.\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right)+\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)$$

$$\left(\frac{1}{2\,\gamma1}\left(e^{-te\sqrt{\gamma1}}\left(\left(-1+e^{te\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2\,te\sqrt{\gamma1}}\right)\right.\right.\right.$$

$$\left.\left.\left.\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[te]\right)+$$

$$\left(\frac{e^{-u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}-\frac{e^{u\sqrt{\gamma1}}}{2\sqrt{\gamma1}}\right)\left(\frac{1}{\sqrt{\gamma1}}\left(-\mathrm{Sinh}\left[te\sqrt{\gamma1}\right]\gamma_0+\right.\right.$$

$$\left.\left.\sqrt{\gamma1}\left(p\,\delta+\mathrm{Cosh}\left[te\sqrt{\gamma1}\right](-p\,\delta+\nu_{c0})\right)\right)+\right.$$

$$\left.\epsilon_{\nu c}[te]\right)\}\},\ \left\{\left\{-\left(-\frac{1}{\gamma1}+\frac{e^{-u\sqrt{\gamma1}}\left(1+e^{2u\sqrt{\gamma1}}\right)}{2\,\gamma1}\right)\right.\right.$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}\sqrt{\gamma1}-\frac{1}{2}\,e^{u\sqrt{\gamma1}}\sqrt{\gamma1}\right)\gamma_0-$$

$$\frac{e^{-u\sqrt{\gamma1}}\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(-1+e^{2u\sqrt{\gamma1}}\right)\gamma_0}{2\sqrt{\gamma1}}+$$

$$\left(\frac{1}{2}\,e^{-u\sqrt{\gamma1}}+\frac{e^{u\sqrt{\gamma1}}}{2}\right)\left(1-\frac{1}{2}\,e^{-u\sqrt{\gamma1}}\left(1+e^{2u\sqrt{\gamma1}}\right)\right)$$

$$\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+$$

$$\left(\left(\frac{1}{2}e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)^2+\left(-\frac{e^{-u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2\sqrt{\gamma 1}}\right)\left(\frac{1}{2}e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\right)$$

$$\sigma_f\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_u\right)-\frac{1}{2\sqrt{\gamma 1}}\left(e^{-u\sqrt{\gamma 1}}\left(-1+e^{2u\sqrt{\gamma 1}}\right)\left(\frac{1}{2}e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\sigma_f\left(-\{\{\omega\}\}_0+\{\{\omega\}\}_u\right)\right)+$$

$$\left(\frac{1}{2}e^{-u\sqrt{\gamma 1}}\sqrt{\gamma 1}-\frac{1}{2}e^{u\sqrt{\gamma 1}}\sqrt{\gamma 1}\right)\left(\frac{1}{2\gamma 1}\left(e^{-te\sqrt{\gamma 1}}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2te\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{toxc}[te]\right)+\left(\frac{1}{2}e^{-u\sqrt{\gamma 1}}+\frac{e^{u\sqrt{\gamma 1}}}{2}\right)$$

$$\left(\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[te\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\left(p\,\delta+\mathrm{Cosh}\left[te\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)\right)+\epsilon_{\nu c}[te]\right)\}\}\}$$

FullSimplify[%]

$$\left\{\left\{\left\{\frac{1}{2\,\gamma1}\left(e^{-(te+u)\sqrt{\gamma1}}\left(\left(-1+e^{(te+u)\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})+2\,e^{(te+u)\sqrt{\gamma1}}\,\gamma1\,\mathrm{Cosh}\left[u\sqrt{\gamma1}\right]\epsilon_{toxc}[te]-2\,e^{(te+u)\sqrt{\gamma1}}\sqrt{\gamma1}\,\mathrm{Sinh}\left[u\sqrt{\gamma1}\right](p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\epsilon_{\nu c}[te])\right)\right)\right\}\right\},$$

$$\left\{\left\{\frac{1}{\sqrt{\gamma1}}\left(-\mathrm{Sinh}\left[(te+u)\sqrt{\gamma1}\right]\gamma_0+\sqrt{\gamma1}\,\mathrm{Cosh}\left[(te+u)\sqrt{\gamma1}\right](-p\,\delta+\nu_{c0})-\gamma1\,\mathrm{Sinh}\left[u\sqrt{\gamma1}\right]\epsilon_{toxc}[te]+\sqrt{\gamma1}\,\mathrm{Cosh}\left[u\sqrt{\gamma1}\right](p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\epsilon_{\nu c}[te])\right)\right\}\right\}\right\}$$

The expected value of $tox_f[u, te]$ and $\nu_f[u,te]$ are found by setting all Wiener processes and random errors equal to zero.

$$\textbf{\% /. } \{(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\to 0,\ \epsilon_{toxc}[te]\to 0,\ \epsilon_{\nu c}[te]\to 0\}$$

$$\left\{\left\{\left\{\frac{1}{2\,\gamma1}\left(e^{-(te+u)\sqrt{\gamma1}}\left(-2\,e^{(te+u)\sqrt{\gamma1}}\,p\sqrt{\gamma1}\,\delta\,\mathrm{Sinh}\left[u\sqrt{\gamma1}\right]+\left(-1+e^{(te+u)\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})\right)\right)\right\}\right\},$$

$$\left\{\left\{\frac{1}{\sqrt{\gamma1}}\left(p\sqrt{\gamma1}\,\delta\,\mathrm{Cosh}\left[u\sqrt{\gamma1}\right]-\mathrm{Sinh}\left[(te+u)\sqrt{\gamma1}\right]\gamma_0+\sqrt{\gamma1}\,\mathrm{Cosh}\left[(te+u)\sqrt{\gamma1}\right](-p\,\delta+\nu_{c0})\right)\right\}\right\}\right\}$$

The Wiener process here is given by all of the terms, minus the expected values, and then setting the other two errors to zero

(%17 - %18) /. {ϵ_{toxc}[te] → 0, ϵ_{vc}[te] → 0}

$$\Bigg\{\Bigg\{\Bigg\{-\frac{1}{2\,\gamma 1}\left(e^{-(te+u)\sqrt{\gamma 1}}\left(-2\,e^{(te+u)\sqrt{\gamma 1}}\,p\sqrt{\gamma 1}\,\delta\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]+\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\frac{1}{2\,\gamma 1}\left(e^{-(te+u)\sqrt{\gamma 1}}\left(\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})-2\,e^{(te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]\,(p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u))\right)\right)\Bigg\}\Bigg\},$$

$$\Bigg\{\Bigg\{-\frac{1}{\sqrt{\gamma 1}}\left(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]-\mathrm{Sinh}\left[(te+u)\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\,\mathrm{Cosh}\left[(te+u)\sqrt{\gamma 1}\right]\,(-p\,\delta+\nu_{c0})\right)+\frac{1}{\sqrt{\gamma 1}}\left(-\mathrm{Sinh}\left[(te+u)\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\,\mathrm{Cosh}\left[(te+u)\sqrt{\gamma 1}\right]\,(-p\,\delta+\nu_{c0})+\sqrt{\gamma 1}\,\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]\,(p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u))\right)\Bigg\}\Bigg\}\Bigg\}$$

FullSimplify[%19]

$$\left\{\left\{\left\{\frac{\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(\{\{\omega\}\}_0-\{\{\omega\}\}_u)}{\sqrt{\gamma 1}}\right\}\right\},\ \left\{\left\{\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right\}\right\}\right\}$$

In the above statement of the Wiener process for tox_f[u, te], the Wiener process is going backward. Multiply this term by -1 and change the signs in the Wiener processes. Call these errors $\epsilon_{\text{toxf}}[u]$ and $\epsilon_{\gamma f}'[u]$, respectively.

$$\left\{\frac{-\text{Sinh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)}{\sqrt{\gamma 1}},\ \text{Cosh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right\}$$

$$\left\{-\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)}{\sqrt{\gamma 1}},\ \text{Cosh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)\right\}$$

The coefficient on the $\epsilon_{\text{toxc}}[te]$error is found by taking the whole expression, subtracting off the expected values, the Wiener process, and setting $\epsilon_{\gamma c}[te]\to 0$

(%17 - %18 - %21) /. ϵ_{vc}[te] → 0

$$\Bigg\{\Bigg\{\Bigg\{-\frac{1}{2\,\gamma 1}\Big(e^{-(te+u)\sqrt{\gamma 1}}\Big(-2\,e^{(te+u)\sqrt{\gamma 1}}\,p\,\sqrt{\gamma 1}\,\delta\,\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big] + \left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0 + \left(-1+e^{2\,(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\Big)\Big) + \frac{\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]\,\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)}{\sqrt{\gamma 1}} + \frac{1}{2\,\gamma 1}\Big(e^{-(te+u)\sqrt{\gamma 1}}\Big(\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0 + \left(-1+e^{2\,(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0}) - 2\,e^{(te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]\,(p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)) + 2\,e^{(te+u)\sqrt{\gamma 1}}\,\gamma 1\,\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]\,\epsilon_{toxc}[te]\Big)\Big)\Bigg\}\Bigg\},$$

$$\Bigg\{\Bigg\{-\frac{1}{\sqrt{\gamma 1}}\Big(p\,\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big] - \mathrm{Sinh}\big[(te+u)\sqrt{\gamma 1}\big]\,\gamma_0 + \sqrt{\gamma 1}\,\mathrm{Cosh}\big[(te+u)\sqrt{\gamma 1}\big]\,(-p\,\delta+\nu_{c0})\Big) - \mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]\,\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u) + \frac{1}{\sqrt{\gamma 1}}\Big(-\mathrm{Sinh}\big[(te+u)\sqrt{\gamma 1}\big]\,\gamma_0 + \sqrt{\gamma 1}\,\mathrm{Cosh}\big[(te+u)\sqrt{\gamma 1}\big]\,(-p\,\delta+\nu_{c0}) + \sqrt{\gamma 1}\,\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]\,(p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)) - \gamma 1\,\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]\,\epsilon_{toxc}[te]\Big)\Bigg\}\Bigg\}\Bigg\}$$

FullSimplify[%22]

$$\left\{\left\{\left\{\text{Cosh}\left[u\sqrt{\gamma 1}\right]\epsilon_{\text{toxc}}[\text{te}]\right\}\right\},\left\{\left\{-\sqrt{\gamma 1}\ \text{Sinh}\left[u\sqrt{\gamma 1}\right]\epsilon_{\text{toxc}}[\text{te}]\right\}\right\}\right\}$$

The coefficient on the $\epsilon_{\gamma\text{c}}[\text{te}]$ is found by taking the whole expression, subtracting off the expected values, the Wiener process, and setting $\epsilon_{\text{toxc}}[\text{te}] \to 0$

(%17 - %18 - %21) /. ϵ_{toxc}[te] → 0

$$\Bigg\{\Bigg\{\Bigg\{-\frac{1}{2\,\gamma 1}\Bigg(e^{-(te+u)\sqrt{\gamma 1}}\Bigg(-2\,e^{(te+u)\sqrt{\gamma 1}}\,p\sqrt{\gamma 1}\,\delta\,\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]+\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^{2}\gamma_{0}+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\Bigg)\Bigg)+$$

$$\frac{\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]\,\sigma_{f}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{u})}{\sqrt{\gamma 1}}+$$

$$\frac{1}{2\,\gamma 1}\Bigg(e^{-(te+u)\sqrt{\gamma 1}}\Bigg(\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^{2}\gamma_{0}+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})-2\,e^{(te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\mathrm{Sinh}\big[u\sqrt{\gamma 1}\big]\,(p\,\delta+\sigma_{f}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{u})+\epsilon_{\nu c}[te])\Bigg)\Bigg)\Bigg\}\Bigg\},$$

$$\Bigg\{\Bigg\{-\frac{1}{\sqrt{\gamma 1}}\Big(p\sqrt{\gamma 1}\,\delta\,\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]-\mathrm{Sinh}\big[(te+u)\sqrt{\gamma 1}\big]\,\gamma_{0}+\sqrt{\gamma 1}\,\mathrm{Cosh}\big[(te+u)\sqrt{\gamma 1}\big]\,(-p\,\delta+\nu_{c0})\Big)-$$

$$\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]\,\sigma_{f}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{u})+$$

$$\frac{1}{\sqrt{\gamma 1}}\Big(-\mathrm{Sinh}\big[(te+u)\sqrt{\gamma 1}\big]\,\gamma_{0}+\sqrt{\gamma 1}\,\mathrm{Cosh}\big[(te+u)\sqrt{\gamma 1}\big]\,(-p\,\delta+\nu_{c0})+\sqrt{\gamma 1}\,\mathrm{Cosh}\big[u\sqrt{\gamma 1}\big]\,(p\,\delta+\sigma_{f}\,(-\{\{\omega\}\}_{0}+\{\{\omega\}\}_{u})+\epsilon_{\nu c}[te])\Big)\Bigg\}\Bigg\}\Bigg\}$$

FullSimplify[%]

$$\left\{\left\{\left\{-\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right]\epsilon_{\nu c}[te]}{\sqrt{\gamma 1}}\right\}\right\},\left\{\left\{\text{Cosh}\left[u\sqrt{\gamma 1}\right]\epsilon_{\nu c}[te]\right\}\right\}\right\}$$

This is a check. Do all the components add up to the whole?

%17 - (%18 + %21 + %23 + %25)

$$\left\{\left\{\left\{-\frac{1}{2\gamma 1}\left(e^{-(te+u)\sqrt{\gamma 1}}\left(-2e^{(te+u)\sqrt{\gamma 1}}p\sqrt{\gamma 1}\,\delta\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]+\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)}{\sqrt{\gamma 1}}-\text{Cosh}\left[u\sqrt{\gamma 1}\right]\epsilon_{toxc}[te]+\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right]\epsilon_{\nu c}[te]}{\sqrt{\gamma 1}}+\frac{1}{2\gamma 1}\left(e^{-(te+u)\sqrt{\gamma 1}}\left(\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})+2e^{(te+u)\sqrt{\gamma 1}}\gamma 1\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]\epsilon_{toxc}[te]-2e^{(te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\text{Sinh}\left[u\sqrt{\gamma 1}\right](p\,\delta+\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\epsilon_{\nu c}[te])\right)\right)\right\}\right\},$$

$$\left\{\left\{-\frac{1}{\sqrt{\gamma 1}}\left(p\sqrt{\gamma 1}\,\delta\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]-\text{Sinh}\left[(te+u)\sqrt{\gamma 1}\right]\gamma_0+\sqrt{\gamma 1}\,\text{Cosh}\left[(te+u)\sqrt{\gamma 1}\right](-p\,\delta+\nu_{c0})\right)-\text{Cosh}\left[u\sqrt{\gamma 1}\right]\sigma_f\,(-\{\{\omega\}\}_0+\{\{\omega\}\}_u)+\right.\right.$$

$$\sqrt{\gamma 1}\ \text{Sinh}\left[u\sqrt{\gamma 1}\right] \epsilon_{toxc}[te] - \text{Cosh}\left[u\sqrt{\gamma 1}\right] \epsilon_{\gamma c}[te] + \frac{1}{\sqrt{\gamma 1}} \Big(-\text{Sinh}\left[(te+u)\sqrt{\gamma 1}\right] \gamma_0 + \sqrt{\gamma 1}\ \text{Cosh}\left[(te+u)\sqrt{\gamma 1}\right] (-p\,\delta + \nu_{c0}) - \gamma 1\ \text{Sinh}\left[u\sqrt{\gamma 1}\right] \epsilon_{toxc}[te] + \sqrt{\gamma 1}\ \text{Cosh}\left[u\sqrt{\gamma 1}\right] (p\,\delta + \sigma_f(-\{\{\omega\}\}_0 + \{\{\omega\}\}_u) + \epsilon_{\gamma c}[te])\Big)\}\}\}$$

FullSimplify[%26]

{{{0}}, {{0}}}

We now derive the variance of the Wiener process. Start with the process

%21

$$\left\{-\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right] \sigma_f \left(-\{\{\omega\}\}_0 + \{\{\omega\}\}_u\right)}{\sqrt{\gamma 1}},\ \text{Cosh}\left[u\sqrt{\gamma 1}\right] \sigma_f \left(-\{\{\omega\}\}_0 + \{\{\omega\}\}_u\right)\right\}$$

The variance of the Wiener process is

$$\left\{\left(-\frac{\text{Sinh}\left[u\sqrt{\gamma 1}\right] \sigma_f}{\sqrt{\gamma 1}}\right)\text{^}2\, u,\ \left(\text{Cosh}\left[u\sqrt{\gamma 1}\right] \sigma_f\right)\text{^}2\, u\right\}$$

$$\left\{\frac{u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_f^2}{\gamma 1},\ u\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2 \sigma_f^2\right\}$$

The whole error term that includes $\epsilon_{toxc}[te]$ is

%23

$$\left\{\left\{\left\{\text{Cosh}\left[u\sqrt{\gamma 1}\right] \epsilon_{toxc}[te]\right\}\right\},\ \left\{\left\{-\sqrt{\gamma 1}\ \text{Sinh}\left[u\sqrt{\gamma 1}\right] \epsilon_{toxc}[te]\right\}\right\}\right\}$$

The variance of these terms are

$$\left\{\left(\mathbf{Cosh}\left[u\sqrt{\gamma 1}\right]\right)\text{^}2\,V[\epsilon_{toxc}[te]],\ \left(-\sqrt{\gamma 1}\ \mathbf{Sinh}\left[u\sqrt{\gamma 1}\right]\right)\text{^}2\,V[\epsilon_{toxc}[te]]\right\}$$

$$\left\{\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]^2 V[\epsilon_{toxc}[te]],\ \gamma 1\ \mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2 V[\epsilon_{toxc}[te]]\right\}$$

Next is the expression of $V[\epsilon_{toxc}[te]$ from Appendix 1 :

$$\left\{-\frac{\mathbf{Sinh}\left[te\sqrt{\gamma 1}\right]\sigma_c}{\sqrt{\gamma 1}}\right\}\text{^}2\ te$$

$$\left\{\frac{te\,\mathrm{Sinh}\left[te\sqrt{\gamma 1}\right]^2\sigma_c^2}{\gamma 1}\right\}$$

Putting them together

$$\left\{\mathbf{Cosh}\left[u\sqrt{\gamma 1}\right]^2\left(\frac{te\,\mathbf{Sinh}\left[te\sqrt{\gamma 1}\right]^2\sigma_c^2}{\gamma 1}\right),\ \gamma 1\ \mathbf{Sinh}\left[u\sqrt{\gamma 1}\right]^2\left(\frac{te\,\mathbf{Sinh}\left[te\sqrt{\gamma 1}\right]^2\sigma_c^2}{\gamma 1}\right)\right\}$$

$$\left\{\frac{te\,\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]^2\mathrm{Sinh}\left[te\sqrt{\gamma 1}\right]^2\sigma_c^2}{\gamma 1},\ te\,\mathrm{Sinh}\left[te\sqrt{\gamma 1}\right]^2\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_c^2\right\}$$

The $\epsilon_{vc}[te]$ error term is

%25

$$\left\{\left\{\left\{-\frac{\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]\epsilon_{vc}[te]}{\sqrt{\gamma 1}}\right\}\right\},\ \left\{\left\{\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]\epsilon_{vc}[te]\right\}\right\}\right\}$$

The variance of this term is

$$\left\{\left(-\frac{\mathbf{Sinh}\left[\mathbf{u}\sqrt{\gamma\mathbf{1}}\right]}{\sqrt{\gamma\mathbf{1}}}\right)\mathbf{\wedge 2\ V[\ \epsilon_{\nu c}[te]],}\right.$$
$$\left.\left(\mathbf{Cosh}\left[\mathbf{u}\sqrt{\gamma\mathbf{1}}\right]\right)\mathbf{\wedge 2\ V[\ \epsilon_{\nu c}[te]]}\right\}$$

$$\left\{\frac{\text{Sinh}\left[u\sqrt{\gamma1}\right]^2 V[\epsilon_{\nu c}[te]]}{\gamma1}, \text{Cosh}\left[u\sqrt{\gamma1}\right]^2 V[\epsilon_{\nu c}[te]]\right\}$$

Next is the expression of $V[\epsilon_{\nu c}[te]$ from TRDRP1.nb :

$$\left\{\mathbf{te\ Cosh}\left[\mathbf{te}\sqrt{\gamma\mathbf{1}}\right]^{\mathbf{2}}\sigma_{\mathbf{c}}^{\mathbf{2}}\right\}$$

$$\left\{\text{te Cosh}\left[te\sqrt{\gamma1}\right]^2\sigma_c^2\right\}$$

Putting them together

$$\left\{\left(-\frac{\mathbf{Sinh}\left[\mathbf{u}\sqrt{\gamma\mathbf{1}}\right]}{\sqrt{\gamma\mathbf{1}}}\right)\mathbf{\wedge 2}\left\{\mathbf{te\ Cosh}\left[\mathbf{te}\sqrt{\gamma\mathbf{1}}\right]^{\mathbf{2}}\sigma_{\mathbf{c}}^{\mathbf{2}}\right\},\right.$$
$$\left.\left(\mathbf{Cosh}\left[\mathbf{u}\sqrt{\gamma\mathbf{1}}\right]\right)\mathbf{\wedge 2}\left\{\mathbf{te\ Cosh}\left[\mathbf{te}\sqrt{\gamma\mathbf{1}}\right]^{\mathbf{2}}\sigma_{\mathbf{c}}^{\mathbf{2}}\right\}\right\}$$

$$\left\{\left\{\frac{\text{te Cosh}\left[te\sqrt{\gamma1}\right]^2\text{Sinh}\left[u\sqrt{\gamma1}\right]^2\sigma_c^2}{\gamma1}\right\},\right.$$
$$\left.\left\{\text{te Cosh}\left[te\sqrt{\gamma1}\right]^2\text{Cosh}\left[u\sqrt{\gamma1}\right]^2\sigma_c^2\right\}\right\}$$

The variance of {tox_f[te,u], ν_f[te,u]} is equal to the sum of the variances of the three error terms in its expression (%29 + %33 + %37)

%29 + %33 + %37

$$\left\{\left\{\frac{\text{te}\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2\,\text{Sinh}\left[\text{te}\sqrt{\gamma 1}\right]^2\,\sigma_c^2}{\gamma 1}+\frac{\text{te}\,\text{Cosh}\left[\text{te}\sqrt{\gamma 1}\right]^2\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_c^2}{\gamma 1}+\frac{u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_f^2}{\gamma 1}\right\},\right.$$
$$\left.\left\{\text{te}\,\text{Cosh}\left[\text{te}\sqrt{\gamma 1}\right]^2\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_c^2+\text{te}\,\text{Sinh}\left[\text{te}\sqrt{\gamma 1}\right]^2\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_c^2+u\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_f^2\right\}\right\}$$

FullSimplify[%]

$$\left\{\left\{\frac{1}{\gamma 1}\left(\frac{1}{4}\,\text{te}\left(-2+\text{Cosh}\left[2\,(\text{te}-u)\sqrt{\gamma 1}\right]+\text{Cosh}\left[2\,(\text{te}+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2+u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_f^2\right)\right\},\right.$$
$$\left.\left\{\frac{1}{4}\,\text{te}\left(2+\text{Cosh}\left[2\,(\text{te}-u)\sqrt{\gamma 1}\right]+\text{Cosh}\left[2\,(\text{te}+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2+u\,\text{Cosh}\left[u\sqrt{\gamma 1}\right]^2\,\sigma_f^2\right\}\right\}$$

Evaluation of the Economic Impact of California's Tobacco Control Program: A Dynamic Model Approach--Appendix 2: A dynamic, normally distributed survival analysis of the relationship between aging, smoking history, and the mortality of men.

Leonard S. Miller

Introduction

In this appendix, I derive a survival model that makes use of the expressions for the index of tobacco-exposure resulting from an individual's smoking history-derived in Appendix 1. Based on this survival model, expressions for the probability of living and dying are derived for never-smokers, current smokers, and former-smokers. The likelihood function for a sample of individuals based on these probabilitieis serves to estimate the parameters of the survival model, which include the parameters of the tobacco-exposure index..

Generally, summarizing smoking history with the tobacco exposure index, and the calculations that estimate the effect of smoking on morbidity, health status and medical costs that derive from usage of this index to summarize smoking history make three improvements over smoking status as the operative description of the effect of smoking behavior on health outcomes. The first improvement focuses on the level of information about an individual's smoking history. This exercise allows for greater detail about the relationship between variations in smoking behavior and about their causal effect on health outcomes. Details about an individual's smoking history can be incorporated into the measure used to summarize an individual's smoking behavior, the level of accumulated tobacco-exposure of an ever-smoker. The measure permits any combination of starting and stopping smoking times and any daily dossage level, measured as packs of cigarettes smoked per day.

The second improvement focuses on the causal effect of smoking on the deterioration of health outcomes. This improvement is meant to address the fact that estimates of the smoking attributable medical services are often greater for former smokers than they are for current smokers. In theory, this should not be the case. In this analysis, the derived measure of smoking's ability to damage health, the index of tobacco-exposure, incorporates theoretical distinctions between current and former-smokers that cause the expected damage to be less for former-smokers compared to current smokers, given all other dimensions of smoking history are the same. The effects due to smoking status, especially current-smoker versus former-smoker, of the relationship between smoking behavior and health outcomes that are to be estimated based on this model are not a "curve fit" exercise that best describes the smoking status-health outcome data. Rather, the estimates best fit the relationship between the effects of smoking behavior and health outcome expressed by the theory expressed in Appendix 1. In that theory the process describing tobacco-exposure implies that toxin levels fall when a current-smoker quits his smoking habit. If a former-smoker has his actual costs greater than a current-smoker, it results from the randomness in life, or the randomness in response to tobacco.

The third improvement focuses on the sample selection bias that is always present in analyses of the effect of smoking on health outcomes. Recognize that analyses of the health effects of smoking are performed on living populations. Death causes sample selection bias among living cohorts--alive responders are always the stronger members of any original cohort because they are the group least affected by smoking behavior. Consequently, the "all other things equal" assumption between never-smokers and ever-smokers is never met. Because the propensity to die for smokers is higher, the sample of smokers who remain alive is always inherently stronger than the sample of alive never-smokers. Thus, the estimated negative effects of smoking on health outcomes are always understated.

The method developed here is best described as a dynamic normally distributed survival analysis; or, perhaps, a dynamic Probit model. Rather than estimate the probability of an event occurring over a defined period of time, as in the Probit model, the dynamic normal survival model estimates the probability of an event occurring over an open ended, unfolding period of time. In this survival model: (1) the event of interest--in this case death--either occurs or it does not occur; (2) the propensity for the event to have occurred by time w is specified as equal to the expected value of the propensity of the event plus a random error term; (3) the variables specifying the expected value can vary continuously with time; (4) the error term at time w has a normal distribution, with (5) an expected value equal to zero, and (6) a variance that can vary with time.

If a respondent is an ever-smoker, his tobacco-exposure level is specified as equal to its expected value plus a random error. The expected value of tobacco-exposure, and the distribution of the random errors (the difference between the true value and the expected value) were derived in Appendix 1. Since the random error has a Normal distribution, the method and specification of the empirical analysis explained here is built on survival analyses that are based on random errors that have a Normal distribution.

A survival analysis (Kalbfleisch & Prentice, 1980) is developed in this Normal framework. The particular survival analysis developed here is of particular interest because it melds two historic lines of quantitative methods: limited dependent variable methods, which have been extensively developed by econometricians (Maddala, 1983), and survival analysis methods, which have been extensively developed by demographers, biostatisticians/epidemiologists, and engineers. In Section 2, I analytically construct a dynamic survivor model from a Probit like model describing the propensity to be dead at a particular time w in the random life span indicated by the variable T of a respondent. The propensity to die is specified as a linear sum of the expected value of an individual's propensity to experience the criterion event and a normally distributed random error. As in a survival model, the model describes the distribution of a respondent's life span ("time to failure"). The dynamic character of the resulting analysis is apparent in two ways. Rather than focusing on whether death {occurred, did not occur} over a defined, fixed period of time, as in the Probit model, the period under analysis is increasing with the passage of time, as in a survival analysis. Thus the Probit like specification of the propensity to be dead at each moment of time is transformed into a survival analysis describing the random life span variable T. This transformation is accomplished in the relationship between the propensity to be dead and the hazard rate, the instantaneous rate of failure (also known as the force for mortality and the failure rate) at each moment w.

Variations among parametric survival models focus on the functional form translating a model's hazard rate into its survival function, the models description of the probability that a respondent will live at least until time T. In the various models used in practice, hazard rates are either constants (such as in the exponential model, (Chiang, 1980)), functions of constants and powers of time (such as in the Weibull model,1939), multivariate--weighted linear sums of fixed characteristics (Tuma, Hannan, & Groeneveld, 1980), characteristics that vary at discrete points of time (Petersen, 1986a, 1986b), or, to a limited degree, characteristics that can be functions of time (Cox, 1972). All of the standard models (that I am aware of) yield closed form expressions for survivor functions and probability density functions of T. The analysis developed here makes use of technological and software developments. The analytically challenging event probability expressions are derived using *Mathematica* (*Mathematica*, Version 7.1, 2008). In the present analysis, the probability expressions for the never-smoker are closed form expressions-,but the probability expressions for current and former smokers are not and numerical integration methods must be used in the estimation of the model's coefficients.

In the analysis to follow, the determinants of the expected value of the propensity die at moment w, denoted by g[w], and the standard error of the random term of this propensity, denoted by σ[w], are functions of time and of parameters describing the cigarette smoking tobacco-exposure process. The analysis in Section 2 focuses on melding the Probit and Survival analyses. General probabilistic expressions for the observed sample events are obtained; that is, for the survivor function--the probability that a life span exceeds the time of data collection (a right censored event), and the probability density function of the life span T at moment of death t. To render these probability expressions applicable to the problem at hand, more detailed specifications are required before it is possible to construct the likelihood function for the observed sample. Section 3 presents a set of background comments that relate to how the specifications are to be made.

While the age of a respondent is observed, if the respondent is an ever-smoker, his accumulated tobacco-exposure is not observed. In Appendix 1, I presented the development of expressions for the theoretical distribution of tobacco-exposure of ever smokers. To render this Appendix "self-contained", a summary of the relevant closed form expressions is contained in Section 3. The tobacco-exposure distribution depends on: (1) an individual's smoking behavior (when smoking was initiated, what was its intensity (packs per day smoked), if and when did a respondent quit); (2) on parameters describing the distribution of tobacco-exposure, which require estimation; and (3) on randomness that is internal to the smoking process (depth of inhalation, an individual's inherent reaction to tobacco-toxin ingestion, variation in toxins per pack by brand, etc.). These tobacco-exposure effects are present in the propensity to die for ever-smokers. Second-hand smoke is not considered in this study. The expected level of tobacco-exposure is incorporated into the specification of the expected propensity to die by time w (g[w]); the randomness associated with a respondent's smoking history is incorporated into the random error of the propensity to die for ever-smokers, and consequently, effects the standard error of the propensity to die, σ[w], of respondants in ever-smoker groups. The random errors in the propensity to die for ever-smokers include both the random error describing everyone's random chance in life (the random error in the never-smokers propensity to die equation) and the random error describing an individual's random response to smoking. For every smoking history group, the resulting random error in the propensity to die has a Normal distribution (Kotz, Balakrishnan, & Johnson, 2000).

The specifications assume that never-smokers form the basis of the description between age and death for smokers and never-smokers. Based on the general probabilistic expressions developed in Section 2 and the specification of the model, the probability expressions developed for never-smokers, current-smokers, and former-smokers in Sections 4, 5, and 6. For never-smokers, the expected propensity to die by time w is specified in Section 4 as a linear function of age and age-squared as well as a random variable that increases with time. The basic randomness in the propensity to die for never-smokers is the randomness representing the vicissitudes of life. This randomness is also present in the normal random variable of the propensity to die for respondents who are ever-smokers. For current-smokers, the expected propensity to die by time w is specified in Section 5 as equal to the expected propensity to die by time w for the never-smoker plus a linear function of the current-smoker's tobacco-exposure. The random variable is equal to the random variable of the never-smoke, plus the product of the coefficient on the expected tobacco-toxin and the difference between the actual tobacco toxin level for the individual and the expected value of his tobacco exposure. This difference is a random variable whose variance was derived in Appendix 1. For former-smokers, the expected propensity to die by time w is specified in Section 6 as equal to the expected propensity to die by time w for the never-smoker, plus the expected propensity to die for current smokers by time te--the time the individual ended smoking--plus the expected value of the propensity to die for former smokers who have babstained from smoking for time u. The random variable has a component from each these expressions.

Section 2: A dynamic Normal survival model.

Let T represent a random variable denoting the life span of a respondent (time to failure) and let F[T<=w] denote the probability that a respondent will die prior to time w. F[T<=w] is the probability distribution of T. Let h[w] + O[△] denote the probability that an individual will die within the interval [w, w + △). h[w] denotes the rate of dying at time w. In the older literature h[w] is known as the "force of mortality" (Gompertz, 1825; Makeham, 1860); in later literature h[w] is known as the hazard rate or the failure rate (Kalbfleisch & Prentice, 1980). O[△] represents second order effects. O[△] is a function of △; it tends to zero faster than △ tends to zero (Chiang, 1980).

The modern theory of survivor analysis derives from the construction of the differential equation describing how the distribution of the life span T changes over time. To the best of my knowledge, this approach was first offered for the Poisson process by Feller (1957). If an individual dies prior to the time w + △, the probability of this event can be expressed by F[T<=w+△]. The respondent must either have died prior to w, with probability F[T<=w], or if he lived to time w, the event has a probabilty (1 - F[T<=t]), then he must have died between w and w + △, with probability (h[w] + O[△]). The probabilistic statement detailing these possibilities is given by equation [2.1],

[2.1]   F[T<=w+△] = F[T<=w] + (1 - F[T<=w])(h[w] + O[△]).

Rearranging terms (moving F[T<=w] to the left side of the equality), dividing through by △, and taking the limit as △ goes to zero yields the differential equation describing the time rate of change of the distribution of T. The probability density function of T (denoted by f[w]) follows from these operations and is given by equation [2.2a], where the distribution function is subject to the initial condition that it is equal to 0 when the process begins, F[T=0]=0. Equation [2.2.2b] represents this initial condition,

[2.2a] f[w] = d/dw F[T<=w] = (1 - F[T<=w]) h[w],

subject to

[2.2b]   F[T=0]=0.

The solution to equations [2.2a], subject to [2.2b], defines the survival function, the probability that time to death exceeds time w. This probability, denoted by G[T>w] is given by equation [2.3],

[2.3] $G[T>w] = (1 - F[T<=w]) = \mathrm{Exp}[-\int_0^w h[\tau]\, d\tau]$.

We begin by constructing the propensity of a respondent to be dead at some time w, 0 ⩽ w ⩽ t. The propensity to be dead at w is denoted death*[w]. Assume that the propensity to be dead at w is the sum of the expected value of the propensity evaluated at time w, denoted by g[w], and a random error at time w, denoted by ς[w]. Whether the individual is dead or alive at time w (1 or 0, respectively) is a measure of the observable event "the observation is dead or alive at time w", respectively. If the propensity to be dead is greater than zero, an observed measure will be one, and vise-a-versa. Equation [2.4a], defines the propensity to be dead at time w. Equation [2.4b] defines the relationship between an individual's propensity score and his observable measure death[w]; equation [2.4c] defines the distribution of the random variable at time w, [2.4a] death*[w] = g[w] + ς[w];

where:

g[w] is the expected value at time w of the respondent's propensity to have died by time w;

ς[w] is a random variable at time w;

[2.4b] death*[w] {>, $\leqslant$} 0, death[w] = {1,0},

and

[2.4c] ς[w] ~ Normal[0, σ^2[w]].

With the exception that a Probit model expresses equations [2.4a] through [2.4c] for a fixed interval of time rather than for a particular time w, equations [2.4a] through [2.4c] describe the Probit model, which perhaps suggests the Probit name for the survival model under development.

In survival analyses, the hazard rate is defined as the ratio of the rate of change of the probability of dying to the probability of being alive. With this propensity score, the maximum probability of being alive is measured by the distribution function evaluated at a propensity to die equal to the value zero. Time rates of change in this probability will also occur at this propensity value. The description of the propensity to be dead by time w implies that the propensity score has a normal distribution with a mean g[w] and a variance σ^2[w]. This distribution implies that equation [2.1] can be stated in Normal distribution terms as equation [2.5],

[2.5] (1 - Φ[(death*[w + $\triangle$]-g[w + $\triangle$])/σ[w + $\triangle$]) =

(1 - Φ[(death*[w]-g[w])/σ[w]]) +

Φ[(death*[w]-g[w])/σ[w]] (h[w] + O[$\triangle$]),

where Φ[] is the normal distribution function. Replicating the steps that led from equation [2.1] to equation [2.2] yields an expression for the hazard rate of this problem; that is--rearrange terms, divide by $\triangle$, and take the limit as $\triangle$ goes to zero--and then (1) evaluate the expressions at death*[w]=0, and (2) solve for the hazard rate, h[w]. Equation [2.6] describes the hazard rate at time w for this problem,

[2.6] $h[w] = \partial_w(\Phi[g[w]/\sigma[w]])/(1-\Phi[g[w]/\sigma[w]])$

$= \{(1/\sigma[w])\ \varphi[g[w]/\sigma[w]]\ (\partial_w\, g[w]\ /\ \sigma[w])\}\ /$

$(1-\Phi-g[w]/\sigma[w]]),$

where $\varphi[]$ is the normal probability density function and ∂_w denotes the partial derivative with respect to w.

The survival function, G[T>w], and the probability density function, f[T=w], of the random life-span variable T are, respectively, the probability that a respondent was alive when the data were collected at time w, and the probability that a respondent lived until time w, and then died at time w. These are the probabilities of the observed events that are associated with the life and death of the respondents under analysis. Based on the survival function and the hazard rate (equations [2.3] and [2.2a], above) the probability of survival and the probability density function expressions are given by equations [2.7a] and [2.7b],

[2.7a] $G[T>t] = (1 - F[T<=t]) = Exp[-\int_0^t h[w]\, dw],$

and

[2.7b] $f[t] = d/dt\ F[T<=t] = G[T>t]\ h[t].$

The likelihood expression for a sample is the product of the probabilities associated with each of the observed events in a sample. Explicit development of the likelihood function for this problem requires further specification, which will begin to be made in Section 4. Section 3 presents background considerations that affect the specification of the model.

Section 3: Background considerations about time, tobacco-exposure, and heterogeneity.

It is useful to begin a discussion of the specific implementation of the model with background considerations about how time is notated and treated in the model.

The zero point of time is taken to be the mean age that American male's begin to smoke, 17 years of age (REF to NMES). Moreover, time is measured in decades. Thus the age of a 40 year old is measured with a time value of 2.3 decades, ((40 - 17)/10).

Prior to age 35 or 40 (depending on the specific disease) epidemiologists do not generally attribute negative effects of smoking behavior on health, especially its effect on smoking related diseases (Sammet, ????). Consistent with this framework, parameter estimations, both in the mortality model under discussion, and in the smoking related disease models (see Appendix 3), are based on respondents who are at least 40 years of age (2.3 decades in the age units used in the study's time measure).

Time has different relevant meanings within the different smoking statuses. The notation to be developed will account for all of these differences. More specifically, in the specification to be developed for never-smokers, time represents age; in the specification to be developed for current-smokers, time represents both age and time smoked; and in the specification to be developed for former-smokers, time represents age, the duration of time smoked, and the duration of time a respondent abstained from smoking. As equation [2.4.3] below will show, the expected value of the propensity to be dead at time w for never-smokers is specified as a linear function of age and age-squared. For current and former smokers, the specification of the propensity to be dead includes these same never-smoker terms. Additionally, the specification includes a coefficient weighted expected level of tobacco-exposure, which estimates the effect of smoking history on the propensity to die. The propensity to be dead also has a random variable and the variance of this random variable affects an ever-smoker's probability of dying. In every respondent's propensity to be dead, the random variable includes a term associated with the random variable in the never-smokers propensity to be dead. This term represents the general vicissitudes of life. For ever-smokers, additionally, the error term includes the product of the coefficient on the tobacco-exposure variable in the expected propensity to be dead and a random variable measuring the difference between a recipient's true tobacco-exposure level and his expected tobacco-exposure level, given his smoking history. Thus the variance of the random variable in an ever-smoker's propensity to be dead includes the square of the coefficient on the tobacco-exposure measure in the expected value of the ever-smoker's propensity and the variance of the difference between the true and expected tobacco-exposure in the body of the ever-smoker.

As depicted in equations [2.7a] and [2.7b] above, the notation used in the survival function and probability density functions, respectively, describe the observed events--lived between time 0 and w (where w represents the final observation), or died at time w, (here w represents time of death, which is after the acquisition of smoking history, but before the final observation about death in 1999. These probability expressions include exponentials of the integral of the negative of the hazard function over the relevant time period. Based on this data set, the integration is actually over the age of each respondent from the initial smoking history acquisition, to either the respondent's age at death, or his age when the final accumulation of death data was completed.

To represent both age, and duration of smoking (and for a former-smoker, period of abstention) in the same integration over observed time, I created a recipient specific coefficient "α" to represent a transformation of a recipient's decades of age into his decades of smoking duration. That is, "α" equals the difference between a respondent's duration of smoking and his age. Consequently, age + α equals duration of smoking. For current-smokers, the integration in the survival function occurs over the recipient's age (w) to his age at death or age at the time of final data collection (w + follow-up time). However, the time dimensions in a current-smoker's tobacco-toxin expression are measuring smoking time. Thus during the period under analysis the levels of tobacco-exposure are being evaluated for the years the respondent smoked; from (age + α) to (age + follow-up time + α). Similarly, the time dimensions in a former-smoker's tobacco-toxin expression are measuring decades of abstention, given decades smoked, and the levels of tobacco-exposure are being evaluated between the years the respondent had abstained at his age when observation started to the years the respondent had abstained when observation ceased. The decades a former-smoker smoked are denoted by te. The decades he subsequently abstained from smoking are denoted by u. The integration for former-smokers is over an abstention period (respondents are classified as former-smokers on their base-line interview). Thus the time they smoked, te, is a given, and the duration of abstention from smoking variable, u, expressed in terms of age as u = w + α - te, is integrated over the age of the respondent during his smoking abstention and either his age at time at death or at final data collection.

With temporal notation explained, it is possible to understand equations [3.1a] and [3.1b], closed form expressions describing the expected level and the variance, respectively, of the level of tobacco-exposure for a current-smoker at age w during the observation period in the NAS-NRC data. Equations [3.2a] and [3.2b] report these same expressions for former-smokers. The derivations of these expressions were made in Appendix 1. Here, these expressions are to be taken as given.

[3.1 a] $\text{tox}_c[w, \alpha]$ =

$$= \left\{\frac{1}{2\gamma 1}\left(e^{-(w+\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)\right\}+\epsilon_{\text{toxc}}[w+\alpha];$$

where:

[3.1b] $\epsilon_{\text{toxc}}[w+\alpha]\sim N\left[0, \left\{\frac{1}{\gamma 1}\left((w+\alpha)\,\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2\sigma_c^2\right)\right\}\right] = N[0, \sigma_{\text{toxc}}{}^2[t]].$

[3.2 a] $\text{tox}_f[u, te]$ =

$$= \frac{1}{2\gamma 1}\left(e^{(-(te+u))\sqrt{\gamma 1}}\left(-2\,e^{(te+u)\sqrt{\gamma 1}}\,p\sqrt{\gamma 1}\,\delta\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]+\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2(te+u)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right)+\epsilon_{\text{tox}_f}[u, te];$$

where:

[3.2b] $\epsilon_{\text{toxf}}[u, te]\sim$

$$N\left[0, \frac{1}{\gamma 1}\left(\frac{1}{4}\,te\left(-2+\text{Cosh}\left[2(te-u)\sqrt{\gamma 1}\right]+\text{Cosh}\left[2(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2+u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\right] =$$

$N[0, \sigma_{\text{toxf}}{}^2[u,te]].$

In a Probit model with a homogenous variance, the propensity equation is implicitly "standardized". The assumed error term's unit variance is achieved by implicitly dividing the propensity expression by the (unknown) standard error of the random error term. The implicit division renders the coefficients in the expected value "standardized" and the model with a random error that has a variance equal to one. In a Probit model with heterogenous variance, implicitly a similar step is taken. The heterogenous variance might be specified as the exponential of a weighted linear sum of characteristics, say Exp[Zγ]. Feasibility of parameter estimation requires that Z not have a column of one's, which would be multiplying an intercept term in the vector γ (Green, 1990). If γ0 were the coefficient on an intercept, and if the remaining part of the variance's Zγ description were partitioned to separate γ0 from the remaining products of gamma coefficients and their Z variables, the latter of which will be denoted by Z1γ1, then [Zγ]= Exp[γ0] Exp[Z1γ1]. The absence of Exp[γ0] is equivalent to having divided the specification of the propensity score by the square root of Exp[γ0] (i.e., Exp[γ0/2], to remove Exp[γ0] from the variance specification.

For the never-smokers in this study, the variance arising from integrating white noise over time, for example from age 40, measured as 2.3, to age 46, measured as 2.9, is $\sigma^2{}_n$ 0.6. The never-smoker's death propensity equation is divided by σ_n and the variance of the random error in the propensity of a never-smoker by age w is expressed as the value of his age, w. For current and former-smokers, additionally, the variance includes a term associated with the coefficient weighted variance in the distribution of tobacco-toxins in the body. Thus the coefficients on the variables in the expected level of the propensity to be dead by age w are "standardized" by the standard deviation in the vicissitudes of life. Additionally, the coefficients indicating the constants in the propensity to be dead for every smoking status, whose description is yet to be made, are similarly standardized.

Section 4: The probability that a never-smoker lives longer than the final data collection date, or that he dies between the initial base-line and the final data collection date.

A never-smoker's age (measured in the units of the problem--decades, with zero equal to 17 years of age)--is represented by the variable w. I specify the expected value of the propensity to be dead by age w (or what would be age w if the person were alive) as the sum of the product of a constant, η1, and the individual's age (w), and the product of a second constant, η2, and the square of the individual's age$\left(w^2\right)$.

Equation [4.1] presents the relevant equation,

[4.1] $\text{death}^*{}_n[w] = \eta 1\ w + \eta 2\ w^2 + \varsigma_n[w]$,

where $\varsigma_n[w] \sim N[0, w]$.

A point to emphasize here, which is true for the specifications of the model for all smoking status groups, is that this is a dynamic model. All of the variables in the expected value of the propensity to be dead, as well as the variance of the error in the propensity to be dead are changing continuously as the respondent ages. With respect to the error term, $\varsigma_n[w]$, I have assumed that a white noise process underlies the random error expressed in the propensity to be dead. As commented on in Section 3 above, I have also assumed that the propensity specification has been standardized by the size of the standard deviation of the Brownian motion (white-noise) process. Hence, the eta coefficients are to be understood as standardized. It follows that the hazard-rate for the never-smoker at age w is given as follow:

The hazard-rate for a never-smoker, the ratio of the time rate of change in the probability of dying at age w, divided by the probability of living at age w, is given (in *Mathematica* notation) by equation [4.2],

[4.2]

$$\mathbf{h_n = D\left[CDF\left[NormalDistribution\left[0, \sqrt{w}\right], (\eta 1\ w + \eta 2\ w^2)\right], w\right] \Big/ \left(1 - CDF\left[NormalDistribution\left[0, \sqrt{w}\right], (\eta 1\ w + \eta 2\ w^2)\right]\right)}$$

$$\frac{e^{-\frac{(w\,\eta 1 + w^2\,\eta 2)^2}{2\,w}} \left(\frac{\eta 1 + 2\,w\,\eta 2}{\sqrt{2}\,\sqrt{w}} - \frac{w\,\eta 1 + w^2\,\eta 2}{2\,\sqrt{2}\,w^{3/2}}\right)}{\sqrt{\pi}\,\left(1 + \frac{1}{2}\left(-1 - \mathrm{Erf}\left[\frac{w\,\eta 1 + w^2\,\eta 2}{\sqrt{2}\,\sqrt{w}}\right]\right)\right)}$$

The survival function for a never - smokes at age "ulag" (which is the upper limit of the integration over age), is the probability that the never - smoker will live beyond age = ulage, given that he was alive at a base - line age, llage (lower - limit of the integration over age). It is equal to the exponential of the integral of the negative of the hazard - rate over the period of observation, from llage, the age of the respondent at base - line, to ulage, the age of the respondent when observation is complete, which is either (1) when follow - up is completed; or (2) age at death. The Mathematica expression for the survival function is given by equation [4 .3],

[4 .3]

$$\mathbf{g_n =}$$

$$\mathbf{Exp\left[Integrate\left[-\frac{e^{-\frac{(w\,\eta 1 + w^2\,\eta 2)^2}{2\,w}} \left(\frac{\eta 1 + 2\,w\,\eta 2}{\sqrt{2}\,\sqrt{w}} - \frac{w\,\eta 1 + w^2\,\eta 2}{2\,\sqrt{2}\,w^{3/2}}\right)}{\sqrt{\pi}\,\left(1 + \frac{1}{2}\left(-1 - Erf\left[\frac{w\,\eta 1 + w^2\,\eta 2}{\sqrt{2}\,\sqrt{w}}\right]\right)\right)}, \{w, llagein66, ulage\}, Assumptions \to \{Element[\{llagein66, ulage\}, Reals]\ \&\&\ llagein66 < ulage\ \&\&\ llagein66 > 0\}\right]\right]}$$

$$\frac{-1+\mathrm{Erf}\left[\frac{\sqrt{\mathrm{ulage}}\,(\eta 1+\mathrm{ulage}\ \eta 2)}{\sqrt{2}}\right]}{-1+\mathrm{Erf}\left[\frac{\sqrt{\mathrm{llagein66}}\,(\eta 1+\mathrm{llagein66}\ \eta 2)}{\sqrt{2}}\right]}$$

The hazard-rate evaluated at age of death is given by equation [4.4],

[4.4]

$$\mathrm{ht}_n = \mathrm{h}_n \ /.\ \mathrm{w} \to \mathrm{ulage}$$

$$\frac{e^{-\frac{(\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2)^2}{2\,\mathrm{ulage}}}\left(\frac{\eta 1+2\,\mathrm{ulage}\ \eta 2}{\sqrt{2}\ \sqrt{\mathrm{ulage}}}-\frac{\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2}{2\sqrt{2}\ \mathrm{ulage}^{3/2}}\right)}{\sqrt{\pi}\left(1+\frac{1}{2}\left(-1-\mathrm{Erf}\left[\frac{\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2}{\sqrt{2}\ \sqrt{\mathrm{ulage}}}\right]\right)\right)}$$

The probability density function of the random life span for a never smoker is the product of the hazard-rate evaluated at time of death and the survival function evaluated to the time of death, equation [4.5],

[4.5]

$$\mathrm{f}_n = \mathrm{ht}_n\ \mathrm{g}_n$$

$$\left(e^{-\frac{(\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2)^2}{2\,\mathrm{ulage}}}\left(\frac{\eta 1+2\,\mathrm{ulage}\ \eta 2}{\sqrt{2}\ \sqrt{\mathrm{ulage}}}-\frac{\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2}{2\sqrt{2}\ \mathrm{ulage}^{3/2}}\right)\left(-1+\mathrm{Erf}\left[\frac{\sqrt{\mathrm{ulage}}\ (\eta 1+\mathrm{ulage}\ \eta 2)}{\sqrt{2}}\right]\right)\right)\Big/$$
$$\left(\sqrt{\pi}\left(-1+\mathrm{Erf}\left[\frac{\sqrt{\mathrm{llagein66}}\ (\eta 1+\mathrm{llagein66}\ \eta 2)}{\sqrt{2}}\right]\right)\left(1+\frac{1}{2}\left(-1-\mathrm{Erf}\left[\frac{\mathrm{ulage}\ \eta 1+\mathrm{ulage}^2\ \eta 2}{\sqrt{2}\ \sqrt{\mathrm{ulage}}}\right]\right)\right)\right)$$

Section 5: The dynamic Normal survival model for Current-Smokers.

The nomenclature in the distribution of tobacco-exposure, is as follows: The variable w is a measure of a respondent's age (in decades after age 17). The distribution of tobacco-exposure of current smokers at age w after smoking is initiated is Normal. Its expected value and variance are given by equation [5.1], (see Appendix 1 for its derivation):

[5.1]

$$\mathbf{tox_c[w + \alpha] \sim N\left[\left(\frac{1}{2\gamma 1}\left(e^{-(w+\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu c0)\right)\right)\right),\left(\frac{(w+\alpha)\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2 \sigma c^2}{\gamma 1}\right)\right]}$$

where:

w is the age of the respondent;

α is the difference between the time the respondant initiated smoking and his age (measured in the units of the problem, so that w + α is the dura tion a respondent smoked;

$\gamma 0$ is the trend in the time rate of change of the purge rate;

$\gamma 1$ is the marginal effect of a unit of tobacco-exposure on the time rate of change of the purge rate;

$\nu c0$ is the purge rate when smoking is initiated,

p is the packs of cigarettes smoked per day;

δ is the toxins per pack smoked, and

σc^2 is the square of the standard deviation of the Brownian motion process of the random variable in the specification of the time rate of change of the purge-rate. This Brownian motion process has been "standardized" by the standard deviation of the Brownian motion process of describing the vicitudes of life, which is the dynamic process leading to the propensity to die by time w for a never-smoker, and applies to all respon dents.

Equation [5.2.a] is the expression for the expected value of $tox_c[w + \alpha]$,

[5.2a]

$$\mathbf{tox_c[w + \alpha] = \left(\frac{1}{2\gamma 1}\left(e^{-(w+\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu c0)\right)\right)\right)}$$

$$\frac{e^{(-w-\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu c0)\right)}{2\gamma 1}$$

and equation [5.2.b] is the expression for the time rate of change of the expected value of $\text{tox}_c[w+\alpha]$,

[5.2b]

$\text{tox}_c{}'[w+\alpha] = D[\text{tox}_c[w+\alpha], w]$

$$-\frac{e^{(-w-\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu c0)\right)}{2\sqrt{\gamma 1}}+\frac{1}{2\,\gamma 1}$$

$$e^{(-w-\alpha)\sqrt{\gamma 1}}\left(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)\gamma 0\sqrt{\gamma 1}+2\,e^{2(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p\,\delta-\nu c0)\right)$$

Equation [5.2.c] is the expression for the standard deviation of the error term in the latent index of death, denoted $\sigma\text{tox}_c[w+\alpha]$,

[5.2c]

$\sigma\text{tox}_c[w+\alpha] = \text{Sqrt}\left[w+\eta 3^2\left(\frac{(w+\alpha)\,\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2\sigma c^2}{\gamma 1}\right)\right]$

$$\sqrt{w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}}$$

and equation [5.2.d] is the expression for the time rate of change in the standard deviation of latent index of death, denoted $\sigma\text{tox}_c[w+\alpha]$,

[5.2d]

$\sigma\text{tox}_c{}'[w+\alpha] = D[\sigma\text{tox}_c[w+\alpha], w]$

$$\left(1+\frac{1}{\sqrt{\gamma 1}}2\,(w+\alpha)\,\eta 3^2\,\sigma c^2\,\text{Cosh}\left[(w+\alpha)\sqrt{\gamma 1}\right]\right.$$
$$\left.\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]+\frac{\eta 3^2\,\sigma c^2\,\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)\Bigg/$$
$$\left(2\sqrt{w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\text{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}}\right)$$

Equation [5.2.e] is the expression for the time rate of change in the probability of dying,

[5.2e]

$$\mathbf{D}\left[\mathbf{CDF}\left[\mathbf{NormalDistribution}[0,\ \sigma\mathbf{tox}_{\mathbf{c}}[w+\alpha]],\ \left(\eta\mathbf{1}\ w\ +\eta\mathbf{2}\ w^{2}+\eta\mathbf{3}\ \mathbf{tox}_{\mathbf{c}}[w+\alpha]\right)\right],\ w\right]$$

$$\frac{1}{\sqrt{\pi}}\ e^{-\frac{\left(w\,\eta1+w^{2}\,\eta2+\frac{e^{(-w-\alpha)\sqrt{\gamma1}}\,\eta3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma1}}\right)^{2}\gamma0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu c0)\right)}{2\,\gamma1}\right)^{2}}{2\left(w+\frac{(w+\alpha)\,\eta3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma1}\right]^{2}}{\gamma1}\right)}}\ \Bigg(-\Bigg(\Bigg(w\,\eta1+w^{2}\,\eta2+$$

$$\frac{1}{2\,\gamma1}\,e^{(-w-\alpha)\sqrt{\gamma1}}\,\eta3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma1}}\right)^{2}\gamma0+\right.$$

$$\left.\left(-1+e^{2(w+\alpha)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu c0)\right)\Bigg)$$

$$\Bigg(1+\frac{1}{\sqrt{\gamma1}}\,2\,(w+\alpha)\,\eta3^{2}\,\sigma c^{2}\,\mathrm{Cosh}\left[(w+\alpha)\sqrt{\gamma1}\right]$$

$$\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma1}\right]+$$

$$\frac{\eta3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma1}\right]^{2}}{\gamma1}\Bigg)\Bigg)\Bigg/$$

$$\left(2\sqrt{2}\left(w+\frac{1}{\gamma 1}(w+\alpha)\,\eta 3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^{2}\right)^{3/2}\right)+$$

$$\left(\eta 1+2\,w\,\eta 2-\frac{1}{2\sqrt{\gamma 1}}\,e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^{2}\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu c0)\right)+\frac{1}{2\,\gamma 1}\,e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)\gamma 0\sqrt{\gamma 1}+2\,e^{2(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p\,\delta-\nu c0)\right)\right)\Big/\left(\sqrt{2}\sqrt{\left(w+\frac{1}{\gamma 1}(w+\alpha)\,\eta 3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^{2}\right)}\right)$$

Equation [5.3] is the expression for the hazard rate for current smokers. In the operation to follow, the component expressions, equations 5.2a-e, developed directly above, as well as the estimated parameters for never-smokers, are substituted into the hazard rate expression,

[5.3]

$$
\begin{aligned}
h_c = &\ \mathrm{D}\Big[\mathrm{CDF}\Big[\mathrm{NormalDistribution}[0, \sigma tox_c[w+\alpha]], \ (\eta 1\, w + \eta 2\, w^2 + \eta 3\, tox_c[w+\alpha])\Big], w\Big] \Big/ \\
&\Big(1 - \mathrm{CDF}\Big[\mathrm{NormalDistribution}[0, \sigma tox_c[w+\alpha]], \ (\eta 1\, w + \eta 2\, w^2 + \eta 3\, tox_c[w+\alpha])\Big]\Big) \;/. \\
&\Big\{tox_c[w+\alpha] \to \frac{1}{2\,\gamma 1}\Big(e^{-(w+\alpha)\sqrt{\gamma 1}} \\
&\quad \Big(\Big(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\Big)^2 \gamma 0 + \Big(-1 + e^{2(w+\alpha)\sqrt{\gamma 1}}\Big)\sqrt{\gamma 1}\,(p\,\delta - \nu c0)\Big)\Big),\ tox_c{}'[w+\alpha] \to \\
&\quad -\frac{1}{2\sqrt{\gamma 1}}\Big(e^{-(w+\alpha)\sqrt{\gamma 1}}\Big(\Big(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\Big)^2 \gamma 0 + \Big(-1 + e^{2(w+\alpha)\sqrt{\gamma 1}}\Big)\sqrt{\gamma 1}\,(p\,\delta - \nu c0)\Big)\Big) + \frac{1}{2\,\gamma 1} \\
&\quad \Big(e^{-(w+\alpha)\sqrt{\gamma 1}}\Big(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\Big(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\Big)\gamma 0\sqrt{\gamma 1} + 2\,e^{2(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p\,\delta - \nu c0)\Big)\Big), \\
&\quad \sigma tox_c[w+\alpha] \to \sqrt{w + \frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\big[(w+\alpha)\sqrt{\gamma 1}\big]^2}{\gamma 1}}, \\
&\quad \sigma tox_c{}'[w+\alpha] \to \frac{1 + \frac{2(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Cosh}\big[(w+\alpha)\sqrt{\gamma 1}\big]\,\mathrm{Sinh}\big[(w+\alpha)\sqrt{\gamma 1}\big]}{\sqrt{\gamma 1}} + \frac{\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\big[(w+\alpha)\sqrt{\gamma 1}\big]^2}{\gamma 1}}{2\sqrt{w + \frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\big[(w+\alpha)\sqrt{\gamma 1}\big]^2}{\gamma 1}}}\Big\} \;/. \\
&\{\delta \to 1,\ \eta 1 \to -1.5681,\ \eta 2 \to 0.2205\}
\end{aligned}
$$

$$\Bigg(e^{-Y}\Bigg(-\Bigg(\Bigg(-1.5681\grave{}\ w+0.2205\grave{}\ w^2+\frac{1}{2\,\gamma 1}e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)\Bigg)$$

$$\left(1+\frac{2\,(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Cosh}\left[(w+\alpha)\sqrt{\gamma 1}\right]\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]}{\sqrt{\gamma 1}}+\frac{\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)\Bigg)\Bigg/\left(2\sqrt{2}\left(w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)^{3/2}\right)+$$

$$\Bigg(-1.5681\grave{}+0.441\grave{}\ w-\frac{e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)}{2\sqrt{\gamma 1}}+$$

$$\frac{1}{2\,\gamma 1}e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)\gamma 0\sqrt{\gamma 1}+2\,e^{2(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p-\nu c0)\right)\Bigg)\Bigg/$$

$$\left(\sqrt{2}\sqrt{w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}}\right)\Bigg)\Bigg)\Bigg/$$

$$\Bigg(\sqrt{\pi}\Bigg(1+\frac{1}{2}\Bigg(-1-\mathrm{Erf}\Bigg[\Bigg(-1.5681\grave{}\ w+0.2205\grave{}\ w^2+\frac{1}{2\,\gamma 1}e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3$$

$$\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)\Bigg)\Bigg/$$

$$\left(\sqrt{2}\sqrt{w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}}\right)\Bigg]\Bigg)\Bigg)\Bigg)$$

where : Y =

$$\left(-1.5681`\, w + 0.2205`\, w^2 + \frac{1}{2\,\gamma 1}\, e^{(-w-\alpha)\sqrt{\gamma 1}}\, \eta 3\, \left(\left(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0 + \left(-1 + e^{2\,(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\, (p - \nu c0)\right)\right)^2 \Big/ \left(2\left(w + \frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)\right)$$

The survival function for a current-smoke at age "ulage" (which is the upper limit of the integration over age), is the probability that the current-smoker will live beyond ulage (upper-limit of the integration over age), given that the respondent was alive at the base-line--his age in 1966, llagein66 (lower - limit of the integration over age--the respondent's age in 1966). The survival function is equal to the exponential of the integral of the negative of the hazard - rate over the period of observation, from llage to the age of the respondent when observation is complete, which is either: (1) the respondent's age when follow-up is completed (his age in 1999); or (2) his age at death. The Mathematica expression for the survival function is given by equation [5.4]. The Hold[] function tells Mathematica not to evaluate the expression. It will be evaluated when the individuals age, w, and his adjustment for when he started smoking, α, are substituted in.

[5.4]

$$\mathbf{g_c = Exp\Big[Hold\Big[NIntegrate\Big[}$$

$$-\left(e^{-Y}\left(-\left(\left(-1.5681`\, w + 0.2205`\, w^2 + \frac{1}{2\,\gamma 1}\, e^{(-w-\alpha)\sqrt{\gamma 1}}\, \eta 3\, \left(\left(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0 + \left(-1 + e^{2\,(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\, (p - \nu c0)\right)\right)\right.\right.\right.$$

$$\left(1+\frac{1}{\sqrt{\gamma 1}}\,2\,(w+\alpha)\,\eta 3^{2}\,\sigma c^{2}\,\mathrm{Cosh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]\right.$$

$$\left.\left.\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]+\frac{\eta 3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]^{2}}{\gamma 1}\right)\right)\Big/$$

$$\left(2\sqrt{2}\left(w+\frac{(w+\alpha)\,\eta 3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]^{2}}{\gamma 1}\right)^{3/2}\right)+$$

$$\left(-1.5681\grave{}+0.441\grave{}\,w-\frac{1}{2\sqrt{\gamma 1}}\,e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\right.$$

$$\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^{2}\gamma 0+\left(-1+e^{2\,(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)+$$

$$\frac{1}{2\,\gamma 1}\,e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\,\left(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)\right.$$

$$\left.\left.\gamma 0\,\sqrt{\gamma 1}+2\,e^{2\,(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p-\nu c0)\right)\right)\Big/$$

$$\left.\left.\left(\sqrt{2}\,\sqrt{w+\frac{(w+\alpha)\,\eta 3^{2}\,\sigma c^{2}\,\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]^{2}}{\gamma 1}}\right)\right)\right)\Big/$$

$$\left(\sqrt{\pi}\,\left(1+\frac{1}{2}\,\left(-1-\mathrm{Erf}\left[\left(-1.5681\grave{}\,w+0.2205\grave{}\,w^{2}+\frac{1}{2\,\gamma 1}\right.\right.\right.\right.\right.$$

$$e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0 + \left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)\Bigg/$$

$$\left(\sqrt{2}\sqrt{w+\frac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}}\;\right]\Bigg)\Bigg)\Bigg)\Bigg),$$

$\{w,\ \mathrm{llagein66},\ \mathrm{ulage}\}\big]\big]\big]$

$$\mathrm{Hold}\Big[\mathrm{NIntegrate}\Big[-\frac{e^{-Y}\left(-\dfrac{\left(-1.5681\,w+0.2205\,w^2+\dfrac{e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)}{2\,\gamma 1}\right)\left(1+\right.}{2\sqrt{2}\left(w+\dfrac{(w+\alpha)\,\eta 3^2\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\right.}{\gamma 1}\right.}\right.}{}$$

e

$g_c =$
$\mathrm{Exp}[\mathrm{Hold}\,[\mathrm{NIntegrate}\,[-(\mathrm{Exp}[-Y]\,((-A\,B)\,/\,C+(D\,/\,E)))\,/\,F,\ \{w,\ \mathrm{llagein66}\ ,\ \mathrm{ulage}\}]]$

$$A=\left(-1.5681`\,w+0.2205`\,w^2+\frac{1}{2\,\gamma 1}\,e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)\right)$$

$$B=\left(1+\frac{2\,(w+\alpha)\,\sigma c^2\,\mathrm{Cosh}\left[(w+\alpha)\sqrt{\gamma 1}\right]\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]}{\sqrt{\gamma 1}}+\frac{\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)\Bigg)$$

$$C=\left(2\sqrt{2}\left(w+\frac{(w+\alpha)\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\sqrt{\gamma 1}\right]^2}{\gamma 1}\right)^{3/2}\right)$$

$$D=-1.5681`+0.441`\,w+\eta 3\left(-\frac{1}{2\sqrt{\gamma 1}}e^{(-w-\alpha)\sqrt{\gamma 1}}\left(\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2\gamma 0+\left(-1+e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p-\nu c0)\right)+\frac{1}{2\,\gamma 1}e^{(-w-\alpha)\sqrt{\gamma 1}}\left(2\,e^{(w+\alpha)\sqrt{\gamma 1}}\left(-1+e^{(w+\alpha)\sqrt{\gamma 1}}\right)\gamma 0\sqrt{\gamma 1}+2\,e^{2(w+\alpha)\sqrt{\gamma 1}}\,\gamma 1\,(p-\nu c0)\right)\right.$$

$$\mathbf{E} = \sqrt{2}\,\sqrt{w + \frac{(w+\alpha)\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]^2}{\gamma 1}}$$

$$\mathbf{F} = \sqrt{\pi}\left(1 + \frac{1}{2}\left(-1 - \mathrm{Erf}\left[\left(-1.5681\grave{}\ w + 0.2205\grave{}\ w^2 + \frac{1}{2\,\gamma 1}\, e^{(-w-\alpha)\sqrt{\gamma 1}}\,\eta 3\left(\left(-1 + e^{(w+\alpha)\sqrt{\gamma 1}}\right)^2 \gamma 0 + \left(-1 + e^{2(w+\alpha)\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p - \nu c0)\right)\right) \Big/ \left(\sqrt{2}\,\sqrt{w + \frac{(w+\alpha)\,\sigma c^2\,\mathrm{Sinh}\left[(w+\alpha)\,\sqrt{\gamma 1}\right]^2}{\gamma 1}}\right)\right]\right)\right)$$

To construct the probability density function of the life-span, the probability of living until time ulage and then dying at time ulage, we need the hazard-rate evaluated at time of death. This value is given by equation [5.5],

[5.5]

$$\mathbf{ht_c = h_c\ /.\ w \to ulage}$$

$$\left(e^{-Y}\left(-\left(\left(-1.5681`\ \text{ulage}+0.2205`\ \text{ulage}^2+\frac{1}{2\,\gamma1}e^{(-\text{ulage}-\alpha)\sqrt{\gamma1}}\,\eta3\left(\left(-1+e^{(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)^2\gamma0+\left(-1+e^{2\,(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p-\nu c0)\right)\right)\right.\right.$$

$$\left(1+\frac{1}{\sqrt{\gamma1}}2\,(\text{ulage}+\alpha)\,\eta3^2\,\sigma c^2\,\text{Cosh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]\text{Sinh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]+\frac{\eta3^2\,\sigma c^2\,\text{Sinh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]^2}{\gamma1}\right)\Bigg)\Bigg/$$

$$\left(2\sqrt{2}\left(\text{ulage}+\frac{(\text{ulage}+\alpha)\,\eta3^2\,\sigma c^2\,\text{Sinh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]^2}{\gamma1}\right)^{3/2}\right)+$$

$$\left(-1.5681`+0.441`\ \text{ulage}-\frac{1}{2\sqrt{\gamma1}}e^{(-\text{ulage}-\alpha)\sqrt{\gamma1}}\,\eta3\right.$$

$$\left(\left(-1+e^{(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)^2\gamma0+\left(-1+e^{2\,(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p-\nu c0)\right)+\frac{1}{2\,\gamma1}e^{(-\text{ulage}-\alpha)\sqrt{\gamma1}}$$

$$\left.\eta3\left(2\,e^{(\text{ulage}+\alpha)\sqrt{\gamma1}}\left(-1+e^{(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)\gamma0\sqrt{\gamma1}+2\,e^{2\,(\text{ulage}+\alpha)\sqrt{\gamma1}}\,\gamma1\,(p-\nu c0)\right)\right)\Bigg/$$

$$\left.\left.\left(\sqrt{2}\sqrt{\text{ulage}+\frac{(\text{ulage}+\alpha)\,\eta3^2\,\sigma c^2\,\text{Sinh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]^2}{\gamma1}}\right)\right)\right)\Bigg/$$

$$\left(\sqrt{\pi}\left(1+\frac{1}{2}\left(-1-\text{Erf}\left[\left(-1.5681`\ \text{ulage}+0.2205`\ \text{ulage}^2+\right.\right.\right.\right.\right.$$

$$\frac{1}{2\,\gamma1}e^{(-\text{ulage}-\alpha)\sqrt{\gamma1}}\,\eta3\left(\left(-1+e^{(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)^2\gamma0+\left(-1+e^{2\,(\text{ulage}+\alpha)\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p-\nu c0)\right)\Bigg)\Bigg/$$

$$\left.\left.\left.\left.\left(\sqrt{2}\sqrt{\text{ulage}+\frac{(\text{ulage}+\alpha)\,\eta3^2\,\sigma c^2\,\text{Sinh}\left[(\text{ulage}+\alpha)\sqrt{\gamma1}\right]^2}{\gamma1}}\right)\right]\right)\right)\right)$$

The probability density function of the random life span for a current smoker is the product of the hazard-rate evaluated at time of death and the survival function evaluated to the time of death, equation [5.6],

[5.6]

$$\mathbf{f_c = ht_c\, g_c}$$

f_c can be obtained by substitution.

■ Section 6: The dynamic Normal survival model for Former-Smokers.

In the analysis of the survival of former smokers, as with the other smoking statuses all time is measured in decades. te is the time the individual smoked, u is the time he abstained from smoking. w is his age, past 17 and α is the time adjustment to convert age into time smoked. Moreover, te + u = w + α. The distribution of tobacco-exposure of former-smokers at age w, (u + te - α) in decades beyound age 17, is Normal. Its expected value and variance are given by equation [6.1], (see Appendix 1 for derivation). Note that both the expected value and the variance are composed of two terms. The first term (in both) is the expected tobacco-exposure and variance that occured while the former-smoker was a current smoker. The second term is the "contribution" of the former-smoker's abstention to his tobacco exposure and its variance,

[6.1]

$$\mathrm{tox}_f[u \mid te] \sim$$

$$N\Bigg[\Bigg(\left(\frac{e^{-te\sqrt{\gamma 1}}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,te\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)}{2\,\gamma 1}\right)+$$
$$\left(\frac{1}{2\,\gamma 1}e^{-(te+u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0+\right.\right.$$
$$\left.\left.\sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)\right)\Bigg),$$

$$\Bigg(\left(\frac{1}{\gamma 1}\left(\frac{1}{4}te\left(-2+\mathrm{Cosh}\left[2\,(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2\right)+\right.$$
$$\left.u\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\Bigg)\Bigg]$$

In addition to the age and smoking related terms discussed immediately above,

$\gamma 0$ is the trend in the time rate of change of the purge rate;

$\gamma 1$ is the marginal effect of a unit of tobacco-toxin on the time rate of change of the purge rate;

$\nu c0$ is the purge rate when smoking is first initiated,

p is the packs of cigarettes smoked per day;

δ is the toxins per pack smoked, and

σc^2 & σf^2 are, respectively, the square of the standard deviation of the Brownian motion process of the random variable in the specification of the time rate of change of the purge-rate for current and former smokers. This Brownian motion process has been "standardized" by the standard deviation of the Brownian motion process of describing the vicitudes of life, which is the dynamic process leading to the propensity to die by time w for a never-smoker, and applies to all respondents.

The propensity for a former-smoker to be dead by age w (in decades), given he smoked for te decades is specified by equation [6.2],

[6.2] $$\text{death}^*[u|te,\alpha] = \eta 1\,(u+te-\alpha) + \eta 2\,(u+te-\alpha)^2 + \eta 3\,E[\text{tox}_c[te]] + \eta 4\,E[\text{tox}_f[u\,|\,te]] + \varsigma[u|\alpha,te],$$

where:

E[] is the expectation operator;

$\varsigma[u|\alpha,te]$ is a random variable with a Normal distribution whose expected value equals zero; and whose variance at age w, for a former-smoker who smoked a duration te decades equals the quantity (See Appendix 1),

$$V[\varsigma[u\,|\,\alpha,te]] = (u+te-\alpha) + \eta 3^2\left(\frac{1}{\gamma 1}\left(\frac{1}{4}te\left(-2+\text{Cosh}\left[2\,(te-u)\sqrt{\gamma 1}\right] + \text{Cosh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\right)\right)\sigma_c^2 + \eta 4^2\left(u\,\text{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)$$

As in Section 5 above, we now detail expressions for the components of the hazard rate. From equation [6.1] the expected value of the former-smokers tobacco exposure is given by equation [6.3],

[6.3a]

$$\mathbf{tox}_f = \left(\frac{1}{2\gamma 1}e^{-te\sqrt{\gamma 1}}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)^2\gamma_0 + \left(-1+e^{2\,te\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)\right) + \left(\frac{1}{2\gamma 1}e^{-(te+u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0 + \sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta - \left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)\right)$$

$$\frac{e^{-te\sqrt{\gamma 1}}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)^2\gamma_0 + \left(-1+e^{2\,te\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(p\,\delta-\nu_{c0})\right)}{2\gamma 1} + \frac{1}{2\gamma 1}e^{(-te-u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0 + \sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta - \left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)$$

Equation [6.3b] is the expression for the time rate of change of the expected value of tox_f [u|te],
[6.3b]

$tox_f' = D[tox_f, u]$

$$\frac{1}{2\sqrt{\gamma 1}} e^{(-te-u)\sqrt{\gamma 1}+u\sqrt{\gamma 1}}\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0 + \sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right) - \frac{1}{2\sqrt{\gamma 1}} e^{(-te-u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0 + \sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right) + \frac{1}{2\gamma 1} e^{(-te-u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\left(e^{(2\,te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\gamma_0 + \sqrt{\gamma 1}\left(e^{(te+u)\sqrt{\gamma 1}}\left(-1+e^{te\sqrt{\gamma 1}}\right)p\sqrt{\gamma 1}\,\delta - e^{(2\,te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\nu_{c0}\right)\right)$$

From the remarks for equation [6.2], equation [6.3c] is the expression for the standard deviation of the error term in the latent index of death, ς[w, α | te],denoted $\sigma tox_f[w+\alpha]$,

[6.3c]

$$\boldsymbol{\sigma\, tox_f = \mathrm{Sqrt}\Big[(u+te-\alpha) + \eta 3^2\left(\frac{1}{\gamma 1}\left(\frac{1}{4}te\left(-2+\mathrm{Cosh}\left[2(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2\right)\right)+\eta 4^2\left(u\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\Big]}$$

$$\sqrt{\left(te+u-\alpha+\frac{te\,\eta 3^2\left(-2+\mathrm{Cosh}\left[2(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2}{4\gamma 1}+u\,\eta 4^2\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)}$$

Equation [6.3d] is the expression for the time rate of change in the standard deviation of the error term in the latent index of death, denoted $\sigma tox_f'$,

[6.3d]

$$\sigma\mathrm{tox}_f{}' = \mathrm{D}\left[\sqrt{\left((\mathrm{te}+\mathrm{u}-\alpha) + \frac{\mathrm{te}\,\eta 3^2\left(-2+\mathrm{Cosh}\left[2\,(\mathrm{te}-\mathrm{u})\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}\right]\right)\sigma_c^2}{4\,\gamma 1} + \mathrm{u}\,\eta 4^2\,\mathrm{Sinh}\left[\mathrm{u}\sqrt{\gamma 1}\right]^2\sigma_f^2\right)},\ \mathrm{u}\right]$$

$$\left(1+\frac{1}{4\,\gamma 1}\mathrm{te}\,\eta 3^2\left(-2\sqrt{\gamma 1}\,\mathrm{Sinh}\left[2\,(\mathrm{te}-\mathrm{u})\sqrt{\gamma 1}\right]+2\sqrt{\gamma 1}\,\mathrm{Sinh}\left[2\,(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}\right]\right)\sigma_c^2 + 2\,\mathrm{u}\sqrt{\gamma 1}\,\eta 4^2\,\mathrm{Cosh}\left[\mathrm{u}\sqrt{\gamma 1}\right]\mathrm{Sinh}\left[\mathrm{u}\sqrt{\gamma 1}\right]\sigma_f^2+\eta 4^2\,\mathrm{Sinh}\left[\mathrm{u}\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\Big/\left(2\sqrt{\left(\mathrm{te}+\mathrm{u}-\alpha+\frac{\mathrm{te}\,\eta 3^2\left(-2+\mathrm{Cosh}\left[2\,(\mathrm{te}-\mathrm{u})\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}\right]\right)\sigma_c^2}{4\,\gamma 1}+\mathrm{u}\,\eta 4^2\,\mathrm{Sinh}\left[\mathrm{u}\sqrt{\gamma 1}\right]^2\sigma_f^2\right)}\right)$$

Equation [6.3e] is the expression for the time rate of change in the probability of dying,

[6.3e] $\partial\mathrm{Pr}[\mathrm{death}]/\partial\mathrm{u} =$

$$\mathrm{D}\Big[\mathrm{CDF}\Big[\mathrm{NormalDistribution}\Big[0,\ \mathrm{Sqrt}\left[(\mathrm{u}+\mathrm{te}-\alpha)+\eta 3^2\left(\frac{1}{\gamma 1}\left(\frac{1}{4}\mathrm{te}\left(-2+\mathrm{Cosh}\left[2\,(\mathrm{te}-\mathrm{u})\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}\right]\right)\right)\right)\sigma_c^2+\eta 4^2\left(\mathrm{u}\,\mathrm{Sinh}\left[\mathrm{u}\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\right]\Big],$$

$$\Bigg(\eta 1\,(\mathrm{te}+\mathrm{u}-\alpha)+\eta 2\,(\mathrm{te}+\mathrm{u}-\alpha)^2+\eta 3\left(\frac{e^{-\mathrm{te}\sqrt{\gamma 1}}\left(\left(-1+e^{\mathrm{te}\sqrt{\gamma 1}}\right)^2\gamma_0+\left(-1+e^{2\,\mathrm{te}\sqrt{\gamma 1}}\right)\sqrt{\gamma 1}\,(\mathrm{p}\,\delta-\nu_{c0})\right)}{2\,\gamma 1}\right)+$$

$$\eta 4\left(\frac{1}{2\,\gamma 1}e^{-(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}}\left(-1+e^{\mathrm{u}\sqrt{\gamma 1}}\right)\left(\left(-1+e^{(2\,\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}}\right)\gamma_0+\sqrt{\gamma 1}\left(\left(-1+e^{\mathrm{te}\sqrt{\gamma 1}}\right)\left(-1+e^{(\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}}\right)\mathrm{p}\,\delta-\left(1+e^{(2\,\mathrm{te}+\mathrm{u})\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)\right)\Bigg)\Big],\ \mathrm{u}\Big]$$

$$= \frac{1}{\sqrt{\pi}} e^{-(A/B)}$$

$$\Bigg(-\Bigg(\Bigg((te+u-\alpha)\,\eta1 + (te+u-\alpha)^2\,\eta2 + \frac{1}{2\,\gamma1} e^{-te\sqrt{\gamma1}}\,\eta3\,\Big(\big(-1+e^{te\sqrt{\gamma1}}\big)^2\,\gamma_0 + \big(-1+e^{2\,te\sqrt{\gamma1}}\big)\sqrt{\gamma1}$$

$$(p\,\delta - \nu_{c0})\Big) + \frac{1}{2\,\gamma1} e^{(-te-u)\sqrt{\gamma1}}\,\big(-1+e^{u\sqrt{\gamma1}}\big)\,\eta4\,\Big(\big(-1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\gamma_0 +$$

$$\sqrt{\gamma1}\,\Big(\big(-1+e^{te\sqrt{\gamma1}}\big)\big(-1+e^{(te+u)\sqrt{\gamma1}}\big)\,p\,\delta - \big(1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\nu_{c0}\Big)\Big)\Bigg)$$

$$\Bigg(1+\frac{1}{4\,\gamma1}\,te\,\eta3^2\,\Big(-2\sqrt{\gamma1}\,\mathrm{Sinh}\big[2\,(te-u)\sqrt{\gamma1}\big] + 2\sqrt{\gamma1}\,\mathrm{Sinh}\big[2\,(te+u)\sqrt{\gamma1}\big]\Big)\,\sigma_c^2 +$$

$$2\,u\sqrt{\gamma1}\,\eta4^2\,\mathrm{Cosh}\big[u\sqrt{\gamma1}\big]\,\mathrm{Sinh}\big[u\sqrt{\gamma1}\big]\,\sigma_f^2 + \eta4^2\,\mathrm{Sinh}\big[u\sqrt{\gamma1}\big]^2\,\sigma_f^2\Bigg)\Bigg)\Bigg/$$

$$\Bigg(2\sqrt{2}\,\Big(te+u-\alpha+\frac{1}{4\,\gamma1}\,te\,\eta3^2\,\Big(-2+\mathrm{Cosh}\big[2\,(te-u)\sqrt{\gamma1}\big]+\mathrm{Cosh}\big[2\,(te+u)\sqrt{\gamma1}\big]\Big)\,\sigma_c^2 +$$

$$u\,\eta4^2\,\mathrm{Sinh}\big[u\sqrt{\gamma1}\big]^2\,\sigma_f^2\Big)^{3/2}\Bigg) +$$

$$\Bigg(\eta1 + 2\,(te+u-\alpha)\,\eta2 + \frac{1}{2\sqrt{\gamma1}} e^{(-te-u)\sqrt{\gamma1}+u\sqrt{\gamma1}}\,\eta4\,\Big(\big(-1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\gamma_0 +$$

$$\sqrt{\gamma1}\,\Big(\big(-1+e^{te\sqrt{\gamma1}}\big)\big(-1+e^{(te+u)\sqrt{\gamma1}}\big)\,p\,\delta - \big(1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\nu_{c0}\Big)\Big) -$$

$$\frac{1}{2\sqrt{\gamma1}} e^{(-te-u)\sqrt{\gamma1}}\,\big(-1+e^{u\sqrt{\gamma1}}\big)\,\eta4\,\Big(\big(-1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\gamma_0 +$$

$$\sqrt{\gamma1}\,\Big(\big(-1+e^{te\sqrt{\gamma1}}\big)\big(-1+e^{(te+u)\sqrt{\gamma1}}\big)\,p\,\delta - \big(1+e^{(2\,te+u)\sqrt{\gamma1}}\big)\,\nu_{c0}\Big)\Big) +$$

$$\frac{1}{2\,\gamma1} e^{(-te-u)\sqrt{\gamma1}}\,\big(-1+e^{u\sqrt{\gamma1}}\big)\,\eta4\,\Big(e^{(2\,te+u)\sqrt{\gamma1}}\sqrt{\gamma1}\,\gamma_0 +$$

$$\sqrt{\gamma1}\,\Big(e^{(te+u)\sqrt{\gamma1}}\big(-1+e^{te\sqrt{\gamma1}}\big)\,p\sqrt{\gamma1}\,\delta - e^{(2\,te+u)\sqrt{\gamma1}}\sqrt{\gamma1}\,\nu_{c0}\Big)\Big)\Bigg)\Bigg/$$

$$\Bigg(\sqrt{2}\,\sqrt{\Big(te+u-\alpha+\frac{1}{4\,\gamma1}\,te\,\eta3^2\,\Big(-2+\mathrm{Cosh}\big[2\,(te-u)\sqrt{\gamma1}\big]+\mathrm{Cosh}\big[2\,(te+u)\sqrt{\gamma1}\big]\Big)\,\sigma_c^2 +}$$

$$u\,\eta4^2\,\mathrm{Sinh}\big[u\sqrt{\gamma1}\big]^2\,\sigma_f^2\Big)\Bigg)\Bigg)$$

where (A / B) is defined as follows :

$$\left((te+u-\alpha)\,\eta1+(te+u-\alpha)^2\,\eta2+\frac{1}{2\,\gamma1}\,e^{-te\sqrt{\gamma1}}\,\eta3\,\left(\left(-1+e^{te\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2\,te\sqrt{\gamma1}}\right)\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})\right)+\right.$$
$$\frac{1}{2\,\gamma1}\,e^{(-te-u)\sqrt{\gamma1}}\left(-1+e^{u\sqrt{\gamma1}}\right)\eta4\,\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma1}}\right)\gamma_0+\right.$$
$$\left.\left.\sqrt{\gamma1}\,\left(\left(-1+e^{te\sqrt{\gamma1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma1}}\right)\nu_{c0}\right)\right)\right)^2\Big/$$
$$\left(2\left(te+u-\alpha+\frac{1}{4\,\gamma1}\,te\,\eta3^2\left(-2+\mathrm{Cosh}\left[2\,(te-u)\sqrt{\gamma1}\right]+\mathrm{Cosh}\left[2\,(te+u)\sqrt{\gamma1}\right]\right)\sigma_c^2+\right.\right.$$
$$\left.\left.u\,\eta4^2\,\mathrm{Sinh}\left[u\sqrt{\gamma1}\right]^2\sigma_f^2\right)\right)\ /.$$

$$\{\delta\to 1,\ \eta1\to -1.5681,\ \eta2\to 0.2205,\ \eta3\to 0.1107,\ \gamma_0\to 0,$$
$$\gamma1\to \mathrm{Exp}[-21.5052],$$
$$\nu_{c0}\to -1.7248,\ \sigma_c\to \mathrm{Exp}[-27.1608]\}$$

$$(A\,/\,B)=$$
$$\left(2587.691333548374`\ e^{-0.000021389722677666222`\,te}\left(-1+e^{0.000042779445355332445`\,te}\right)(1.7248`+p)-1.5681`\right.$$
$$(te+u-\alpha)+0.2205`\ (te+u-\alpha)^2+23375.71213684168`\ e^{0.000021389722677666222`\,(-te-u)}$$
$$\left(-1+e^{0.000021389722677666222`\,u}\right)\left(1.7248`\ \left(1+e^{0.000021389722677666222`\,(2\,te+u)}\right)+\right.$$
$$\left.\left.\left(-1+e^{0.000021389722677666222`\,te}\right)\left(-1+e^{0.000021389722677666222`\,(te+u)}\right)p\right)\eta4\right)^2\Big/$$
$$\left(2\left(te+u-\alpha+1.714959287456654`{*}\hat{}{-17}\ te\ (-2+\mathrm{Cosh}[0.000042779445355332445`\ (te-u)]+\mathrm{Cosh}[\right.\right.$$
$$\left.\left.0.000042779445355332445`\ (te+u)])+u\,\eta4^2\,\mathrm{Sinh}[0.000021389722677666222`\ u]^2\,\sigma_f^2\right)\right)$$

Set::write : Tag Times in $\dfrac{(\ll 1\gg)^2}{B}$ is Protected. ≫

The hazard rate for former smokers, h_f, is the ratio of the time rate of change in the probability of dying, divided by the probability of living, equation [6.4],

[6.4]

$$h_f=\frac{1}{\sqrt{\pi}}\,e^{-A/B}$$
$$\left(-\left(\left((te+u-\alpha)\,\eta1+(te+u-\alpha)^2\,\eta2+\frac{1}{2\,\gamma1}e^{-te\sqrt{\gamma1}}\,\eta3\,\left(\left(-1+e^{te\sqrt{\gamma1}}\right)^2\gamma_0+\left(-1+e^{2\,te\sqrt{\gamma1}}\right)\right.\right.\right.\right.$$
$$\left.\sqrt{\gamma1}\,(p\,\delta-\nu_{c0})\right)+\frac{1}{2\,\gamma1}\,e^{(-te-u)\sqrt{\gamma1}}\left(-1+e^{u\sqrt{\gamma1}}\right)\eta4\,\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma1}}\right)\right.$$

$$\gamma_0 + \sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right) p\,\delta - \left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\Bigg)\Bigg)$$

$$\left(\frac{1}{4\,\gamma 1}\,te\,\eta 3^2\left(-2\sqrt{\gamma 1}\,\mathrm{Sinh}\left[2\,(te-u)\sqrt{\gamma 1}\right]+2\sqrt{\gamma 1}\,\mathrm{Sinh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2 + 2\,u\sqrt{\gamma 1}\,\eta 4^2\,\mathrm{Cosh}\left[u\sqrt{\gamma 1}\right]\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]\sigma_f^2+\eta 4^2\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\Bigg)\Bigg/$$

$$\left(2\sqrt{2}\left((u+te-\alpha)+\frac{1}{4\,\gamma 1}te\,\eta 3^2\left(-2+\mathrm{Cosh}\left[2\,(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2+u\,\eta 4^2\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)^{3/2}\right)+$$

$$\Bigg(\eta 1+2\,(te+u-\alpha)\,\eta 2+\frac{1}{2\sqrt{\gamma 1}}\,e^{(-te-u)\sqrt{\gamma 1}+u\sqrt{\gamma 1}}\,\eta 4\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0+\sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)-$$

$$\frac{1}{2\sqrt{\gamma 1}}\,e^{(-te-u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\eta 4\left(\left(-1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\gamma_0+\sqrt{\gamma 1}\left(\left(-1+e^{te\sqrt{\gamma 1}}\right)\left(-1+e^{(te+u)\sqrt{\gamma 1}}\right)p\,\delta-\left(1+e^{(2\,te+u)\sqrt{\gamma 1}}\right)\nu_{c0}\right)\right)+$$

$$\frac{1}{2\,\gamma 1}\,e^{(-te-u)\sqrt{\gamma 1}}\left(-1+e^{u\sqrt{\gamma 1}}\right)\eta 4\left(e^{(2\,te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\gamma_0+\sqrt{\gamma 1}\left(e^{(te+u)\sqrt{\gamma 1}}\left(-1+e^{te\sqrt{\gamma 1}}\right)p\sqrt{\gamma 1}\,\delta-e^{(2\,te+u)\sqrt{\gamma 1}}\sqrt{\gamma 1}\,\nu_{c0}\right)\right)\Bigg)\Bigg/$$

$$\left(\sqrt{2}\sqrt{\left((u+te-\alpha)+\frac{1}{4\,\gamma 1}te\,\eta 3^2\left(-2+\mathrm{Cosh}\left[2\,(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\sigma_c^2+u\,\eta 4^2\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)}\right)\Bigg/$$

$$\Bigg(1-\mathrm{CDF}\Bigg[\mathrm{NormalDistribution}\Bigg[0,\ \mathrm{Sqrt}\Bigg[(u+te-\alpha)+\eta 3^2\left(\frac{1}{\gamma 1}\left(\frac{1}{4}\,te\left(-2+\mathrm{Cosh}\left[2\,(te-u)\sqrt{\gamma 1}\right]+\mathrm{Cosh}\left[2\,(te+u)\sqrt{\gamma 1}\right]\right)\right)\right)\sigma_c^2+\eta 4^2\left(u\,\mathrm{Sinh}\left[u\sqrt{\gamma 1}\right]^2\sigma_f^2\right)\Bigg]\Bigg],$$

$$\Bigg(\eta 1\ (te + u - \alpha) + \eta 2\ (te + u - \alpha)^2 +$$

$$\eta 3\ \left(\frac{1}{2\,\gamma 1} e^{-te\sqrt{\gamma 1}} \left(\left(-1 + e^{te\sqrt{\gamma 1}}\right)^2 \gamma_0 + \left(-1 + e^{2\,te\sqrt{\gamma 1}}\right) \sqrt{\gamma 1}\ (p\,\delta - \nu_{c0})\right)\right) +$$

$$\eta 4\ \left(\frac{1}{2\,\gamma 1}\ e^{-(te+u)\sqrt{\gamma 1}} \left(-1 + e^{u\sqrt{\gamma 1}}\right) \left(\left(-1 + e^{(2\,te+u)\sqrt{\gamma 1}}\right) \gamma_0 +\right.\right.$$

$$\left.\left.\sqrt{\gamma 1}\ \left(\left(-1 + e^{te\sqrt{\gamma 1}}\right) \left(-1 + e^{(te+u)\sqrt{\gamma 1}}\right) p\,\delta - \left(1 + e^{(2\,te+u)\sqrt{\gamma 1}}\right) \nu_{c0}\right)\right)\right)\Big]\Bigg)\ /.$$

$\{\ \delta \to 1,\ \eta 1 \to -1.5681,\ \eta 2 \to 0.2205\grave{},\ \eta 3 \to 0.1107,\ \gamma_0 \to 0,$
$\gamma 1 \to \mathrm{Exp}[-21.50524],\ \nu_{c0} \to -1.7248,\ \sigma_c \to \mathrm{Exp}[-27.1608]\}$

$$
\begin{aligned}
h_f = \ & \Big(e^{-(A/B)} \\
& \Big(-\Big(\Big(2587.7430878925893` \, e^{-0.00002138929488749056` \mathrm{te}} \left(-1+e^{0.00004277858977498112` \mathrm{te}}\right) (1.7248` + p) - \\
& \quad 1.5681` (\mathrm{te}+u-\alpha) + 0.2205` (\mathrm{te}+u-\alpha)^2 + \\
& \quad 23376.179655759614` \, e^{0.00002138929488749056` (-\mathrm{te}-u)} \left(-1+e^{0.00002138929488749056` u}\right) \\
& \quad \Big(1.7248` \left(1+e^{0.00002138929488749056` (2\,\mathrm{te}+u)}\right) + \left(-1+e^{0.00002138929488749056` \mathrm{te}}\right) \\
& \quad \left(-1+e^{0.00002138929488749056` (\mathrm{te}+u)}\right) p\Big) \eta 4\Big) \Big(1.7150278872001417`\text{*^-17} \, \mathrm{te} \\
& \quad (-0.00004277858977498112` \, \mathrm{Sinh}[0.00004277858977498112` (\mathrm{te}-u)] + \\
& \quad 0.00004277858977498112` \, \mathrm{Sinh}[0.00004277858977498112` (\mathrm{te}+u)]) + \\
& \quad 0.00004277858977498112` \, u \, \eta 4^2 \, \mathrm{Cosh}[0.00002138929488749056` u] \\
& \quad \mathrm{Sinh}[0.00002138929488749056` u] \, \sigma_f^2 + \eta 4^2 \, \mathrm{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2\Big)\Big) \Big/ \\
& \Big(2\sqrt{2} \, \Big(\mathrm{te}+u-\alpha+1.7150278872001417`\text{*^-17} \, \mathrm{te} \, (-2+\mathrm{Cosh}[0.00004277858977498112` \\
& \quad (\mathrm{te}-u)] + \mathrm{Cosh}[0.00004277858977498112` (\mathrm{te}+u)]) + \\
& \quad u \, \eta 4^2 \, \mathrm{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2\Big)^{3/2}\Big) + \\
& \Big(-1.5681` + 0.441` (\mathrm{te}+u-\alpha) + 23376.179655759614` \, e^{0.00002138929488749056` (-\mathrm{te}-u)} \\
& \quad \left(-1+e^{0.00002138929488749056` u}\right) \Big(0.00003689225582194372` \, e^{0.00002138929488749056` (2\,\mathrm{te}+u)} + \\
& \quad 0.00002138929488749056` \, e^{0.00002138929488749056` (\mathrm{te}+u)} \left(-1+e^{0.00002138929488749056` \mathrm{te}}\right) p\Big) \\
& \quad \eta 4 + 0.5` \, e^{0.00002138929488749056` (-\mathrm{te}-u)+0.00002138929488749056` u} \\
& \quad \Big(1.7248` \left(1+e^{0.00002138929488749056` (2\,\mathrm{te}+u)}\right) + \left(-1+e^{0.00002138929488749056` \mathrm{te}}\right) \\
& \quad \left(-1+e^{0.00002138929488749056` (\mathrm{te}+u)}\right) p\Big) \eta 4 - 0.5` \, e^{0.00002138929488749056` (-\mathrm{te}-u)} \\
& \quad \left(-1+e^{0.00002138929488749056` u}\right) \Big(1.7248` \left(1+e^{0.00002138929488749056` (2\,\mathrm{te}+u)}\right) + \\
& \quad \left(-1+e^{0.00002138929488749056` \mathrm{te}}\right) \left(-1+e^{0.00002138929488749056` (\mathrm{te}+u)}\right) p\Big) \eta 4\Big) \Big/ \\
& \Big(\sqrt{2} \, \sqrt{\Big(\mathrm{te}+u-\alpha+1.7150278872001417`\text{*^-17} \, \mathrm{te} \, (-2+\mathrm{Cosh}[0.00004277858977498112`} \\
& \quad (\mathrm{te}-u)] + \mathrm{Cosh}[0.00004277858977498112` (\mathrm{te}+u)]) + \\
& \quad u \, \eta 4^2 \, \mathrm{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2\Big)\Big)\Big)\Big) \Big/ \\
& \Big(\sqrt{\pi} \, \Big(1+\frac{1}{2}\Big(-1-\mathrm{Erf}\Big[\Big(2587.7430878925893` \, e^{-0.00002138929488749056` \mathrm{te}} \\
& \quad \left(-1+e^{0.00004277858977498112` \mathrm{te}}\right) (1.7248` + p) - 1.5681` (\mathrm{te}+u-\alpha) + \\
& \quad 0.2205` (\mathrm{te}+u-\alpha)^2 + 23376.179655759614` \, e^{0.00002138929488749056` (-\mathrm{te}-u)} \\
& \quad \left(-1+e^{0.00002138929488749056` u}\right) \Big(1.7248` \left(1+e^{0.00002138929488749056` (2\,\mathrm{te}+u)}\right) + \\
& \quad \left(-1+e^{0.00002138929488749056` \mathrm{te}}\right) \left(-1+e^{0.00002138929488749056` (\mathrm{te}+u)}\right) p\Big) \eta 4\Big) \Big/ \\
& \quad \Big(\sqrt{2} \, \sqrt{\Big(\mathrm{te}+u-\alpha+1.7150278872001417`\text{*^-17} \, \mathrm{te} \, (-2+\mathrm{Cosh}[} \\
& \quad 0.00004277858977498112` (\mathrm{te}-u)] + \mathrm{Cosh}[0.00004277858977498112` \\
& \quad (\mathrm{te}+u)]) + u \, \eta 4^2 \, \mathrm{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2\Big)\Big)\Big]\Big)\Big)\Big)
\end{aligned}
$$

$$
\begin{aligned}
&(e^{-(A/B)} (-((2587.7430878925893` \, e^{-0.00002138929488749056` te} (-1 + e^{0.00004277858977498112` te}) (1.7248` + p) - \\
&\quad 1.5681` (te + u - \alpha) + 0.2205` (te + u - \alpha)^2 + \\
&\quad 23376.179655759614` \, e^{0.00002138929488749056` (-te-u)} (-1 + e^{0.00002138929488749056` u}) \\
&\quad (1.7248` (1 + e^{0.00002138929488749056` (2 te+u)}) + (-1 + e^{0.00002138929488749056` te}) \\
&\quad (-1 + e^{0.00002138929488749056` (te+u)}) p) \eta 4) (1.7150278872001417`*\text{^}{-17} te \\
&\quad (-0.00004277858977498112` \operatorname{Sinh}[0.00004277858977498112` (te - u)] + \\
&\quad 0.00004277858977498112` \operatorname{Sinh}[0.00004277858977498112` (te + u)]) + \\
&\quad 0.00004277858977498112` \, u \, \eta 4^2 \operatorname{Cosh}[0.00002138929488749056` u] \\
&\quad \operatorname{Sinh}[0.00002138929488749056` u] \sigma_f^2 + \eta 4^2 \operatorname{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2)) / \\
&\quad (2 \sqrt{2} (te + u - \alpha + 1.7150278872001417`*\text{^}{-17} te (-2 + \operatorname{Cosh}[0.00004277858977498112` \\
&\quad (te - u)] + \operatorname{Cosh}[0.00004277858977498112` (te + u)]) + \\
&\quad u \, \eta 4^2 \operatorname{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2)^{3/2}) + \\
&\quad (-1.5681` + 0.441` (te + u - \alpha) + 23376.179655759614` \, e^{0.00002138929488749056` (-te-u)} \\
&\quad (-1 + e^{0.00002138929488749056` u}) (0.00003689225582194372` \, e^{0.00002138929488749056` (2 te+u)} + \\
&\quad 0.00002138929488749056` \, e^{0.00002138929488749056` (te+u)} (-1 + e^{0.00002138929488749056` te}) p) \eta 4 + \\
&\quad 0.5` \, e^{0.00002138929488749056` (-te-u)+0.00002138929488749056` u} \\
&\quad (1.7248` (1 + e^{0.00002138929488749056` (2 te+u)}) + \\
&\quad (-1 + e^{0.00002138929488749056` te}) (-1 + e^{0.00002138929488749056` (te+u)}) p) \eta 4 - \\
&\quad 0.5` \, e^{0.00002138929488749056` (-te-u)} (-1 + e^{0.00002138929488749056` u}) \\
&\quad (1.7248` (1 + e^{0.00002138929488749056` (2 te+u)}) + \\
&\quad (-1 + e^{0.00002138929488749056` te}) (-1 + e^{0.00002138929488749056` (te+u)}) p) \eta 4) / \\
&\quad (\sqrt{2} \sqrt{(te + u - \alpha + 1.7150278872001417`*\text{^}{-17} te (-2 + \operatorname{Cosh}[0.00004277858977498112`} \\
&\quad (te - u)] + \operatorname{Cosh}[0.00004277858977498112` (te + u)]) + \\
&\quad u \, \eta 4^2 \operatorname{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2)))) / \\
&(\sqrt{\pi} (1 + \frac{1}{2} (-1 - \operatorname{Erf}[(2587.7430878925893` \, e^{-0.00002138929488749056` te} \\
&\quad (-1 + e^{0.00004277858977498112` te}) (1.7248` + p) - 1.5681` (te + u - \alpha) + \\
&\quad 0.2205` (te + u - \alpha)^2 + 23376.179655759614` \, e^{0.00002138929488749056` (-te-u)} \\
&\quad (-1 + e^{0.00002138929488749056` u}) (1.7248` (1 + e^{0.00002138929488749056` (2 te+u)}) + \\
&\quad (-1 + e^{0.00002138929488749056` te}) (-1 + e^{0.00002138929488749056` (te+u)}) p) \eta 4) / \\
&\quad (\sqrt{2} \sqrt{(te + u - \alpha + 1.7150278872001417`*\text{^}{-17} te (-2 + \operatorname{Cosh}[} \\
&\quad 0.00004277858977498112` (te - u)] + \operatorname{Cosh}[0.00004277858977498112` \\
&\quad (te + u)]) + u \, \eta 4^2 \operatorname{Sinh}[0.00002138929488749056` u]^2 \sigma_f^2)])))
\end{aligned}
$$

The survival function for a former-smoker, the exponential of the integral of the negative of the hazard-rate over the period of observation, is given by equation [6.5]]. Note that this equation requires numerical methods to carry out.

[6.5]

$$g_f = \operatorname{Exp}[\operatorname{Hold}[\operatorname{NIntegrate}[-h_f, \{u, 0, ulage\}]]]$$

The hazard-rate for former-smokers evaluated at time of death is given by equation [6.6],

[6.6]

$$\mathbf{ht_f = h_f \text{ /. } u \to ulage}$$

$$
\begin{aligned}
&\Big(e^{-(A/B)}\Big(-\Big(\Big(2587.7430878925893` \, e^{-0.00002138929488749056` \,\text{te}}\left(-1+e^{0.00004277858977498112` \,\text{te}}\right)(1.7248`+p)-\\
&\quad 1.5681`\,(\text{te}+\text{ulage}-\alpha)+0.2205`\,(\text{te}+\text{ulage}-\alpha)^2+\\
&\quad 23376.179655759614`\, e^{0.00002138929488749056`\,(-\text{te}-\text{ulage})}\left(-1+e^{0.00002138929488749056`\,\text{ulage}}\right)\\
&\quad \Big(1.7248`\left(1+e^{0.00002138929488749056`\,(2\,\text{te}+\text{ulage})}\right)+\left(-1+e^{0.00002138929488749056`\,\text{te}}\right)\\
&\quad \left(-1+e^{0.00002138929488749056`\,(\text{te}+\text{ulage})}\right)p\Big)\eta 4\Big)\Big(1.7150278872001417`\text{*^-17}\,\text{te}\\
&\quad (-0.00004277858977498112`\,\text{Sinh}[0.00004277858977498112`\,(\text{te}-\text{ulage})]+\\
&\quad 0.00004277858977498112`\,\text{Sinh}[0.00004277858977498112`\,(\text{te}+\text{ulage})])+\\
&\quad 0.00004277858977498112`\,\text{ulage}\,\eta 4^2\,\text{Cosh}[0.00002138929488749056`\,\text{ulage}]\\
&\quad \text{Sinh}[0.00002138929488749056`\,\text{ulage}]\,\sigma_f^2+\\
&\quad \eta 4^2\,\text{Sinh}[0.00002138929488749056`\,\text{ulage}]^2\,\sigma_f^2\Big)\Big)\Big/\\
&\quad \Big(2\sqrt{2}\,\Big(\text{te}+\text{ulage}-\alpha+1.7150278872001417`\text{*^-17}\,\text{te}\,(-2+\text{Cosh}[0.00004277858977498112`\\
&\quad (\text{te}-\text{ulage})]+\text{Cosh}[0.00004277858977498112`\,(\text{te}+\text{ulage})])+\\
&\quad \text{ulage}\,\eta 4^2\,\text{Sinh}[0.00002138929488749056`\,\text{ulage}]^2\,\sigma_f^2\Big)^{3/2}\Big)+\\
&\quad \Big(-1.5681`+0.441`\,(\text{te}+\text{ulage}-\alpha)+23376.179655759614`\,e^{0.00002138929488749056`\,(-\text{te}-\text{ulage})}\\
&\quad \left(-1+e^{0.00002138929488749056`\,\text{ulage}}\right)\Big(0.00003689225582194372`\,e^{0.00002138929488749056`\,(2\,\text{te}+\text{ulage})}+\\
&\quad 0.00002138929488749056`\,e^{0.00002138929488749056`\,(\text{te}+\text{ulage})}\left(-1+e^{0.00002138929488749056`\,\text{te}}\right)p\Big)\\
&\quad \eta 4+0.5`\,e^{0.00002138929488749056`\,(-\text{te}-\text{ulage})+0.00002138929488749056`\,\text{ulage}}\\
&\quad \Big(1.7248`\left(1+e^{0.00002138929488749056`\,(2\,\text{te}+\text{ulage})}\right)+\\
&\quad \left(-1+e^{0.00002138929488749056`\,\text{te}}\right)\left(-1+e^{0.00002138929488749056`\,(\text{te}+\text{ulage})}\right)p\Big)\eta 4-\\
&\quad 0.5`\,e^{0.00002138929488749056`\,(-\text{te}-\text{ulage})}\left(-1+e^{0.00002138929488749056`\,\text{ulage}}\right)\\
&\quad \Big(1.7248`\left(1+e^{0.00002138929488749056`\,(2\,\text{te}+\text{ulage})}\right)+\\
&\quad \left(-1+e^{0.00002138929488749056`\,\text{te}}\right)\left(-1+e^{0.00002138929488749056`\,(\text{te}+\text{ulage})}\right)p\Big)\eta 4\Big)\Big/\\
&\quad \Big(\sqrt{2}\,\sqrt{}\Big(\text{te}+\text{ulage}-\alpha+1.7150278872001417`\text{*^-17}\,\text{te}\,(-2+\text{Cosh}[0.00004277858977498112`\\
&\quad (\text{te}-\text{ulage})]+\text{Cosh}[0.00004277858977498112`\,(\text{te}+\text{ulage})])+\\
&\quad \text{ulage}\,\eta 4^2\,\text{Sinh}[0.00002138929488749056`\,\text{ulage}]^2\,\sigma_f^2\Big)\Big)\Big)\Big)\Big/\\
&\Big(\sqrt{\pi}\,\Big(1+\frac{1}{2}\Big(-1-\text{Erf}\Big[\Big(2587.7430878925893`\,e^{-0.00002138929488749056`\,\text{te}}\left(-1+e^{0.00004277858977498112`\,\text{te}}\right)\\
&\quad (1.7248`+p)-1.5681`\,(\text{te}+\text{ulage}-\alpha)+0.2205`\,(\text{te}+\text{ulage}-\alpha)^2+\\
&\quad 23376.179655759614`\,e^{0.00002138929488749056`\,(-\text{te}-\text{ulage})}\left(-1+e^{0.00002138929488749056`\,\text{ulage}}\right)\\
&\quad \Big(1.7248`\left(1+e^{0.00002138929488749056`\,(2\,\text{te}+\text{ulage})}\right)+\left(-1+e^{0.00002138929488749056`\,\text{te}}\right)\\
&\quad \left(-1+e^{0.00002138929488749056`\,(\text{te}+\text{ulage})}\right)p\Big)\eta 4\Big)\Big/\\
&\quad \Big(\sqrt{2}\,\sqrt{}\Big(\text{te}+\text{ulage}-\alpha+1.7150278872001417`\text{*^-17}\,\text{te}\,(-2+\text{Cosh}[\\
&\quad 0.00004277858977498112`\,(\text{te}-\text{ulage})]+\text{Cosh}[0.00004277858977498112`\\
&\quad (\text{te}+\text{ulage})])+\text{ulage}\,\eta 4^2\,\text{Sinh}[0.00002138929488749056`\,\text{ulage}]^2\,\sigma_f^2\Big)\Big)\Big]\Big)\Big)\Big)
\end{aligned}
$$

The probability density function of the random variable "life span" for former-smokers is the product of the hazard-rate evaluated at time of death and the survival function evaluated at time of death, equation [6.7],

[6.7]

$$ft_f = ht_f \; g_f$$

which can be obtained by substitution.

Bibliography

Broders, A.C., 1920. "Squamous-cell epithelioma of the lip. A study of five hundred and thirty-seven cases," Journal of the American Medical Association 74(10):656-664, March 6, 1920.

Chiang, Chin Long, An Introduction to Stochastic Processes and their Applications. Robert E.Krieger Pub.Co., Inc, Huntington, NY, Chapter 8.0.

Cox, D., 1972. "Regression Models and Life Tables," Journal of the Royal Statistical Society, Series B, 34:187-220.

Feller, William, 1957. An Introduction to Probability Theory and Its Applications, Volume I, Second Edition, John Wiuley & Sons, Inc. NY.

Gillespie, Brenda W., Michael T. Halpern, Kenneth E. Warner, "Patterns of Lung Cancer Risk in Ex-Smokers," Chapter 19 of Case Studies in Biometry,....

Gompertz, B., 1825. On the nature of the function expressive of the law of human mortality. Phil. Trans. Royal Soc. (London), 115:513-583.

Greene, William H., 1990. Econometric Analysis, MacMillan Publishing Company, New

York.

Heckman, J., 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," The Annals of Economic and Social Measurement, 5:475-92, 1976.

1979. "Sample Selection Bias as a Specification Error," Econometricia, 47:153-161,1979.

Kalbfleisch, J., and R. Prentice, The Statistical Analysis of Failure Time Data, New York: Wiley, 1980.

Kotz, Samuel, N. Balakrishnan, and Normal L. Johnson, Continuous Multivariate Distributions, Volume 1: Models and Applications, Second Edition, John Wiley and Sons, Inc., N.Y.

Lee, L., 1976. "Estimation of Limited Dependent Variable Models by Two-Stage Methods," Ph.D. dissertation, University of Rochester.

Lombard, H.L. and C.R. Doering, 1928. "Cancer studies in Massachusetts. 2. Habits, characteristics and environment of individuals with and without cancer," New England Journal of Medicine, 198(10):481-487, April 26, 1928.

Makeham, W.M., 1860. On the law of mortality and the construction of annuity tables, J. Inst. Actuaries, 8.

Maddala, G.S., 1983. Limited Dependent and Qualitative Variables in Econometrics, Cambridge University Press, New York.

Mathematica, 2002. *Mathematica*, Version 4.2.

Pearl, R., 1938. "Tobacco smoking and longevity," Science 87(2253):216-217, March 4, 1938.

Petersen, T., 1986a. "Fitting Parametric Survival Models with Time-Dependent Covariates," Journal of the Royal Statistical Society, Series C (Applied Statistics), 35,3:281-288.

1986b. "Estimating Fully Parametric Hazard Rate Models with Time Dependent Covariates,"
Sociological Methods and Research, 14:219-246.

Rose, Colin and Murray D. Smith, 2001. *Mathematica*l Statistics and *Mathematica*, Springer, NY, Chapter 12.

Royal College of Physicians, 1962. Smoking and Health. Summary and Report of the Royal College of Physicians of London on Smoking in Relation to Cancer of the Lung and Other Diseases. New York:Pitman Publishing, 1962.

Study Group on Smoking and Health, 1957. Smoking and Health. Joint report of the Study Group on Smoking and Health. Science 125(3258):1129-1133, June7,1957.

Tuma, Nancy Brandon, Michael T. Hannan, and Lyle P. Groeneveld, 1980. "Dynamic Analysis of Event Histories," in Ernst W. Stromsdorfer and George Farkas (eds.), Evaluation Studies Review Annual, Volume 5, Berverly Hills, CA. Sage Publications, Inc.

U.S.Department of Health and Human Services, 1989. Reducing the Health Consequences of Smoking:25 Years of Progress. A Report of the Surgeon General.U.S.Depratment of Health and Human Services, Public Health Service,Centers for Disease Control,Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Helth,DHHS Publication No.(CDC)89-8411,1989.

Evluation of the Economic Impact of California's Tobacco Control Program: A Dynamic Model Approach--Appendix 3 : Parameter Estimates for the Mortality, Morbidity, Health Status and Expenditure Models.

Leonard S.Miller

Section 1 : The Mortality Model.

See Appendix 2 for the specification of the mortality model and for the method used to estimate the parameters of the model.The order of estimation of the mo9rtality model’ s parameters is as follows : the never smoker parameters are estimated first; then, the current smoker’ s parameters are estimated, given the never smoker’ s parameter estimates; finally, the single former smoker parameter, the standard deviation of the former smoker’ s variance is estimated, given the never smoker’ s and the current smoker’ s parameters.Table 1 presents the full information (for never smokers) and limited information (for current and former smokers) maximum likelihood estimates of the parameter estimates for the mortality model.Figure 1 illustrates the survival model, given its estimates.Figure 1 depicts the probability of survival for men with six different smoking histories : (1, tan) a current - smoker 2 packs per day; (2, lime) a current - smoker 1 pack per day; (3, blue) a former - smoker, 20 years, 1 pack per day; (4, green) a current - smoker, 1/2 pack per day; (5, purple) a former - smoker, 10 years, 1 pack per day; and (6, red) a never - smoker.One of the products of the mortality model is the estimation of the parameters of the tobacco - exposure index which is used to represent the smoking history of individuals.Figure 2 illustrates the model’ s tobacco - exposure index for current and former smokers with habits of ½ and 2 packs a day.Expected index measures summarize an individual’ s smoking history in the morbidity, health status, and cost models to follow.

Table1.Parameter Estimates for Mortality Model

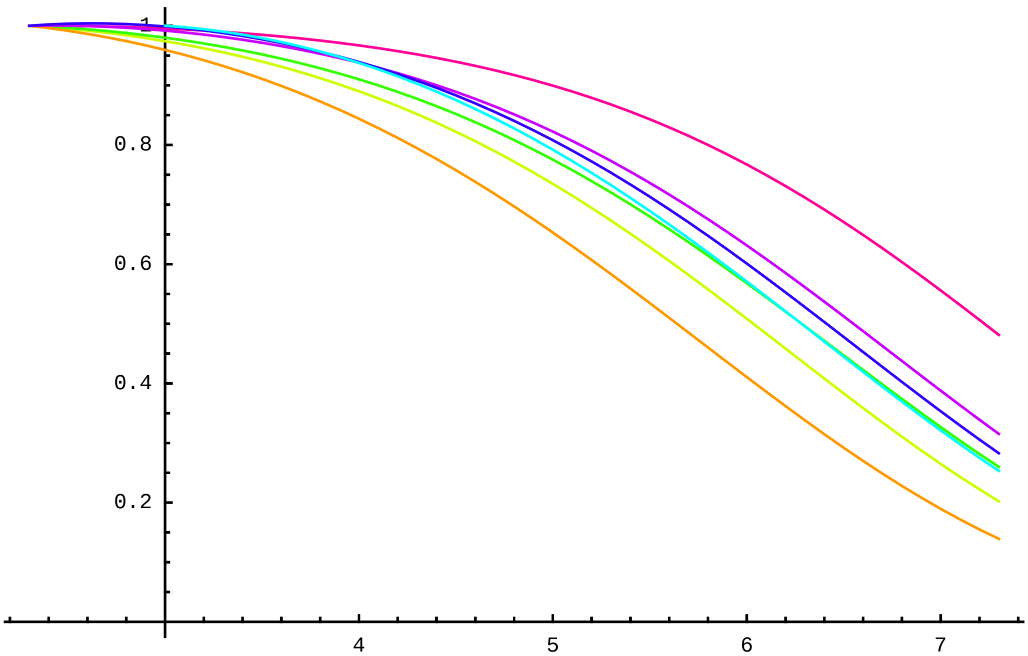
	Parameter	Estimate	Stdev	t – value	Source
Never – Smokers	$\eta 1$	1.5681	.0787	19.917	a.
	$\eta 2$	0.2205	.0114	19.275	
Current – Smokers	$\eta 3$	0.1107	0.0111	9.951	b.
$\gamma 1 = \mathrm{Exp}[\beta 1]$	$\beta 1$	–21.5052	0.113	–190.2	
	ν_{c0}	–1.7248	0.183	–9.405	

σ_c = Exp[β2]	β2	−27.1608	0.182	−149.3	
Former − Smokers	η4	0.0394	0.00796	4.950	c.
σ_f = Exp[β3]	β3	−9.891	10 989.8	−0.0009	

a.= mathematica files//Twinsnev3.nb ;
b.= mathematica files//Twinscur61.nb;
c.= mathematica files//Twinsfor31.nb

- **Figure 1 : Probability of Survival, Given Age and Smoking History**

{Decades of age>17--Males 2.3 to 7. decades}

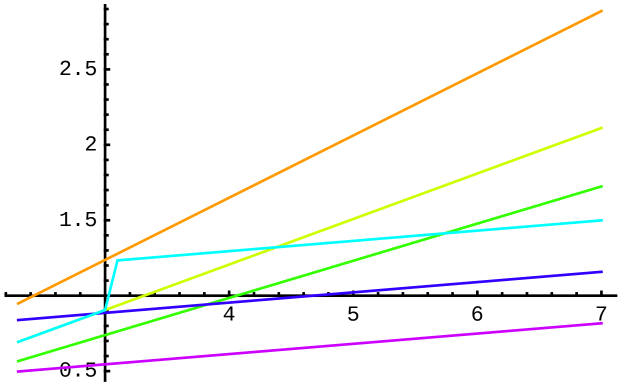


Probability of Survival for Males vs.Decades after age 17
(Source=Twinsfor31.nb)

The order of the survival curves Figure 1 from bottom to top, evaluated at age 77 (6 decades), is as follows :

```
orange     = current - smoker 2 packs per day
lime       = current - smoker 1 pack per day
teal blue  = former - smoker, 30 years, 1 pack per day
green      = current - smoker, 1/2 pack per day
light blue = former - smoker, 30 years, 1 pack per day
dark blue  = former - smoker, 20 years, 1 pack per day
purple     = former - smoker, 10 years, 1 pack per day
red        = never - smoker
```

Figure 2: Tobacco-Exposure, Current and Former smokers (20 years), ½ and 2 packs/day



Tobacco Exposure Index for Males vs.Decades of age after age 17

The order is Figure 2, from bottom to top, is as follows :
Evaluated at age 77 (6 decades), from bottom to top :

purple	=	former - smoker, 10 years, 1 pack per day
dark blue	=	former - smoker, 20 years, 1 pack per day
light blue	=	former - smoker, 30 years, 1 pack per day
green	=	current - smoker, 1/2 pack per day
lime	=	current - smoker 1 pack per day
orange	=	current - smoker 2 packs per day

- **Section 2. Models of the Probability of being Currently Treated for Two Classes of Smoking Caused Diseases.**

Section 2 presents parameter estimates in the models to predict current treatment status (within a year) for smoking caused diseases.We have separated the smoking caused diseases into groups based on their comparable relative odds ratios.The group labeled LC5 has a relative odds ratio of around ten—the relative odds of a smoker being currently treated for an LCF disease is ten times the relative odds for a never - smoker being treated.The group labeled CHD5 has a relative odds ratio of around two.Table 2 lists the ICD - 9 codes for the two groups.

- **Table 2 : Smoking Caused Diseases with their ICD - 9 designations.**

Disease Group	**Disease Name and ICD - 9**
CodeClass 1 : LC5	lung cancer (162), laryngeal cancer (161), chronic obstructive pulmonary disease (491 - 2, 496)
Class 2 : CHD5	atherosclerosis/aortic aneurysm (440 - 441, 444), bladder cancer (188),cerebrovascular disease(430 - 438), with sequelae : hemiplegia and hemiparesis (342), coronary heart disease (410 - 414, 427 - 428), with sequelae : cardiomyopathy and congestive heart failure (425), esophageal cancer (150), kidney cancer (189), oral cancer (140 - 141, 143 - 149), other arterial disease, buerger’ s disease (443.1), peripheral vascular disease (443.9), pancreatic cancer (157), and stomach cancer (151)

The probability of being currently treated in any year for each class is specified by the respective equations,

ProbLC5 = CDF[NormalDistribution[0, (Exp[φ1 curr + φ2 form]) 1/2], (β0 + β1 Age + β3 toxc + β4 toxff + β5 toxfu)]

ProbCHD5 = CDF[NormalDistribution[0, (Exp[φ1 curr + φ2 form]) 1/2], (β0 + β1 Age + β2 Age2 + β3 toxc + β4 toxff + β5 toxfu)]

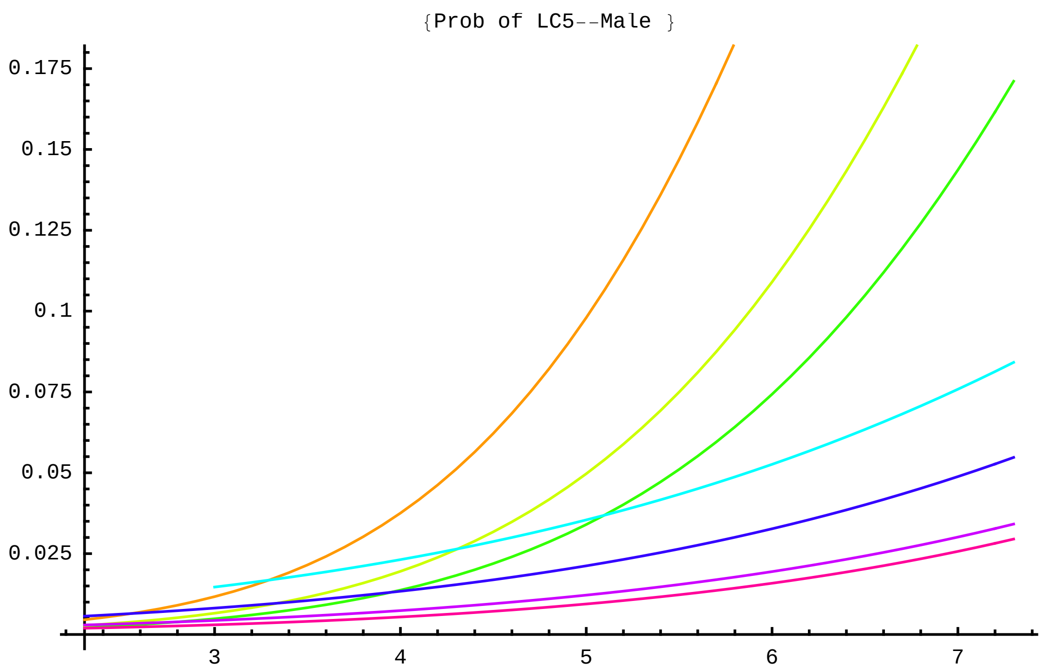
where CDF is the normal distribution function;
curr is 1 if a current smoker, 0 otherwise;
form is 1 if a former smoker, 0 otherwise;
age is measured as decades after age 17;
toxc is the expected toxin exposure index level of the current smoker;
toxff is the expected toxin exposure index level of the former smoker at the time he stopped smoking; and
toxfu is the expected toxin exposure index level of the former smoker in decades after quitting.Table 3 presents parameter estimates for the LC5 model and Table 4 presents parameter estimates for the CHD5 model.

- **Table3: Parameter Estimates for the Current Annual Treatment of LC5 Diseases (Source 719200lc55.nb)**

Variable	Parameter	Estimate	Standard Error	t - value
Current smoker in var	$\varphi 1$	- 0.214	0.184	- 1.163
Former smoker in var	$\varphi 2$	- 0.034	0.208	- 0.164
Constant	$\beta 0$	- 3.349	0.272	- 12.314
Age in decades >17	$\beta 1$	0.200	0.0489	4.091
toxc	$\beta 3$	0.0637	0.00905	7.040
toxff	$\beta 4$	0.0839	0.0345	2.435
toxfu	$\beta 5$	- 0.00978	0.0360	- 0.273

(Source 719200lc55.nb)

- **Figure 3: The probability of annual current treatment (that is, being treated within a year) for LC5, given age and smoking behavior. (Source:/Male Diseases/lc5graphsfinal55.nb)**



- **Bottom to top (evaluated at 6 decades after age 17):**
  never-smoker, (red)
  former-smoker, 10 yr, 1 pk/day (purple)
  former-smoker, 20 yr, 1 pk/day (dark blue)
  former-smoker, 30 yr, 1 pk/day (light blue)
  current-smoker, 1/2 pk day (green)
  current-smoker, 1 pk/day (lime)
  current-smoker, 2 pks/day (tan)

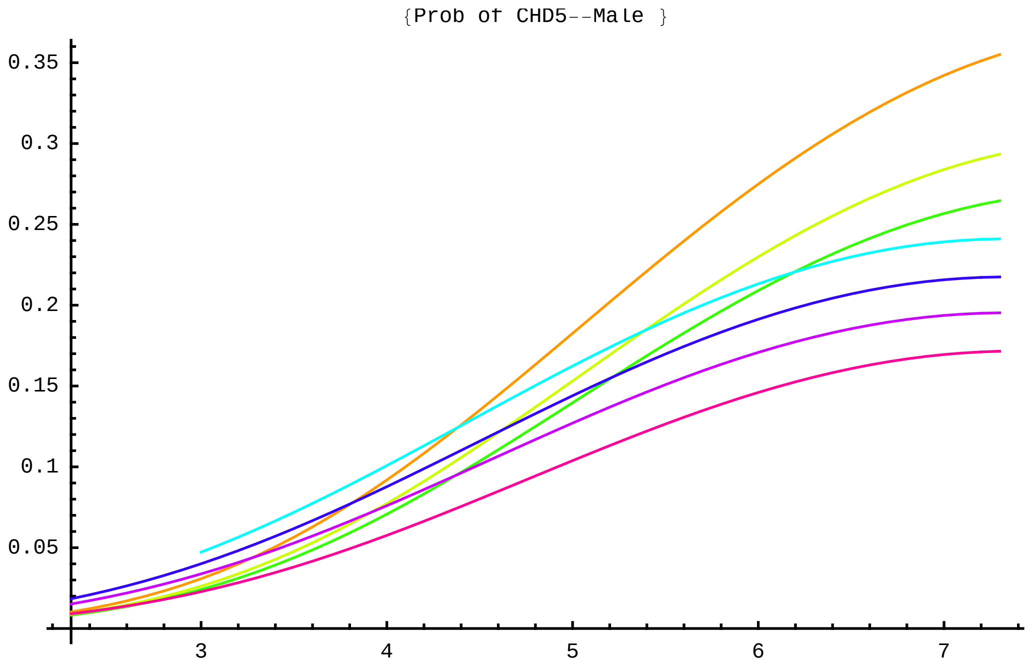
- **Table4: Parameter Estimates for the Current Annual Treatment of CHD5 Diseases**

Variable	Parameter	Estimate	Standard Error	t - value
Current smoker in var	$\varphi 1$	- .128	0.165	- 0.781
Former smoker in var	$\varphi 2$	- .113	0.184	- 0.618
Constant	$\beta 0$	- 3.925	0.417	- 9.409
Age in decades >17	$\beta 1$	0.805	0.172	4.669
Age-squared	$\beta 2$	-0.0544	0.0177	
toxc	$\beta 3$	0.022	0.00725	7.0402
toxff	$\beta 4$	0.0315	0.0208	2.435
toxfu	$\beta 5$	-0.0034	0.0215	- 0.273

(Source : 8252004 chd54.nb)

- **Figure 4: The probability of annual treatment for CHD5, given age and smoking behavior**

.



Bottom to top (evaluated at 7 decades of smoking history) :

never - smoker (red)
former - smoker, 10 yr, 1 pk / day (purple)
former - smoker, 20 yr, 1 pk / day (blue)
former - smoker, 30 yr, 1 pk / day (light blue)
current - smoker, 1 / 2 pack / day, (green)
current - smoker, 1 pack / day, (lime)
current - smoker, 2 packs / day, (tan)

Section 3: Models of Self-Reported Poor Health Status

Self-Reported poor-health status is modeled as an ordered probability model.

Health* = (β0 + β1 age + β3 toxc + β4 toxff + β5 toxfu) + ε
where ε ~ N[0, Exp (φ1 curr + φ2 form)]

where Health* is the latent index of poor health status and (in the discussion below) Health*^ is the expected value of the latent index of poor health status;

The parameters μ1 and μ2 (in Table 4) are boundary values of the latent index of self-reported poor-health status between good and fair and between fair and poor health status, respectively;

The probability of excellent, good, fair, and poor health are given by:
CDF[-Health*^],
CDF[μ1-Health*^]- CDF[-Health*^],
CDF[μ2-Health*^]- CDF[μ1-Health*^],
and 1- CDF[μ2-Health*^], respectively.

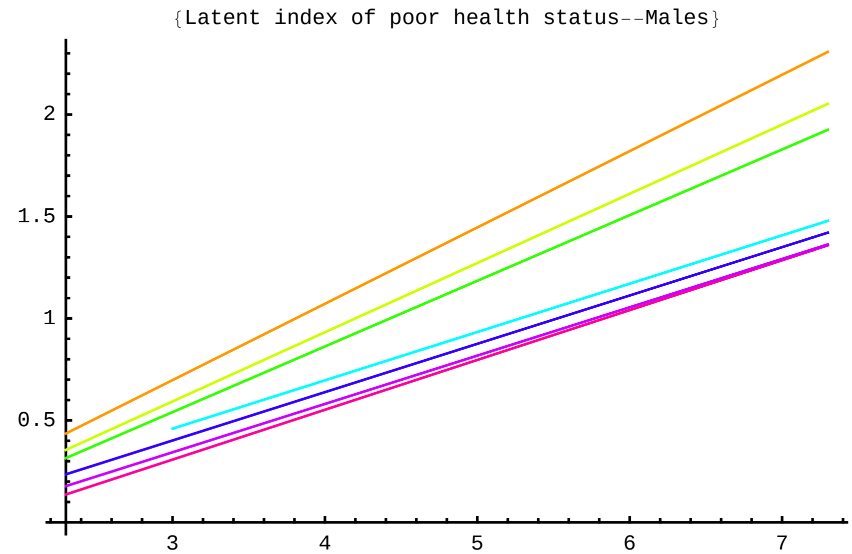
Table 5: Parameter Estimates for the Poor Health Status model

Variable	Parameter	Estimate	Standard Error	t-value
Current smoker in variance	φ1	-0.129	0.0711	-1.81
Former smoker in variance	φ2	-0.136	0.0755	-1.795
Constant	β0	-0.425	0.0581	-7.315
Age (in decades >17)	β1	0.244	0.0165	14.815
Toxc	β3	0.0349	0.00326	10.688
Toxff	β4	0.0228	0.0123	1.851
Toxfu	β5	-0.00425	0.0102	-0.415
Boundary values between good and fair	μ1	1.438	0.0416	34.542
Boundary values between fair and poor	μ2	2.452	0.0699	35.068

(Source:/home/len/mathematica files/Male poor health/7232004ph3.nb)

- **Figure 5: Latent Index of Poor Health Status as a function of age and smoking history**

-

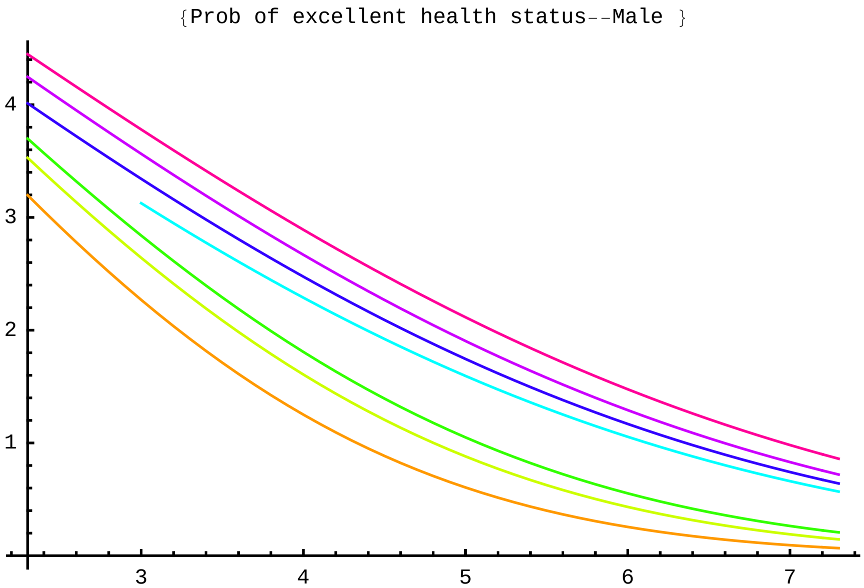


- Graphics -

```
Latent Index of Poor Health Status
Bottom to Top (evaluated 7 decades after average onset of smoking age (17)
           red = never - smoker
           purple = former - smoker, 10 years
           dark blue = former - smoker, 20 years
           light blue = former - smoker, 30 years
           green = current smoker, 1/2 pack/day
           lime = current smoker, 1 pack/day
           tan = current smoker, 2 packs/day
```

(Source: poorhealthgraphs3.nb)

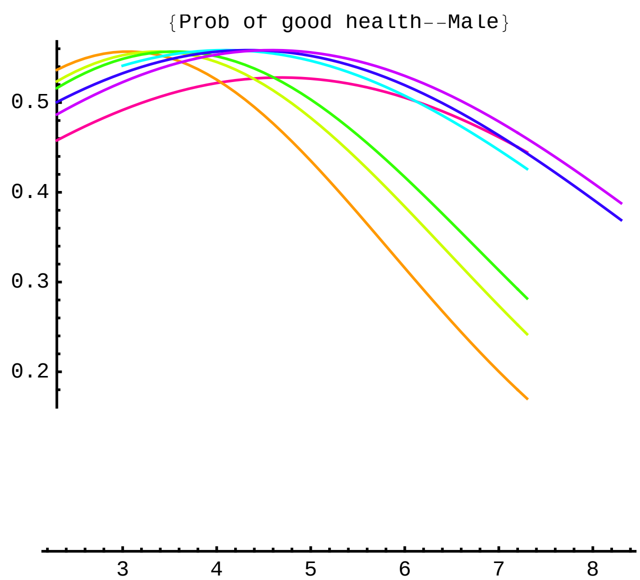
- **Figure 6A: Probability of self-reported excellent health as a function of age and smoking history   (Source: poorhealthgraphs3.nb)**



- Graphics -

```
Probability of self-reported excellent health vs decades after age 17
       red=never-smoker
       purple=former-smoker, 10 years
       dark blue=former-smoker, 20 years
       light blue=former-smoker, 30 years
       green=current smoker, 1/2 pack/day
       light green=current smoker, 1 pack/day
       orange=current smoker, 2 packs/day
```

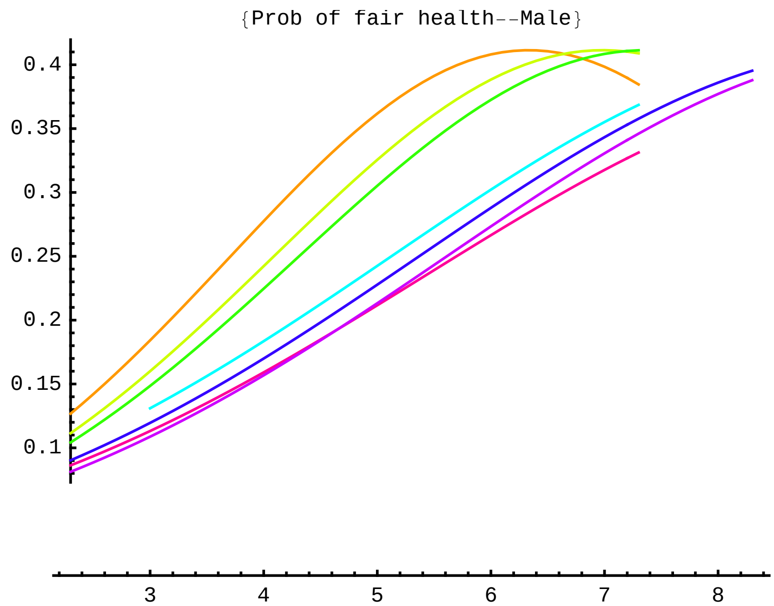
- **Figure 6B: Probability of self-reported good health as a function of age and smoking history   (Source: poorhealthgraphs3.nb)**



- Graphics -

```
Probability of self-reported good health vs decades after age 17
       red=never-smoker
       purple=former-smoker, 10 years
       dark blue=former-smoker, 20 years
       light blue=former-smoker, 30 years
       green=current smoker, 1/2 pack/day
       light green=current smoker, 1 pack/day
       orange=current smoker, 2 packs/day
```

- **Figure 6C: Probability of self-reported fair health as a function of age and smoking history (Source: poorhealthgraphs3.nb)**



- Graphics -

Probability of self-reported fair health vs decades after age 17

red=never-smoker

purple=former-smoker, 10 years

dark blue=former-smoker, 20 years

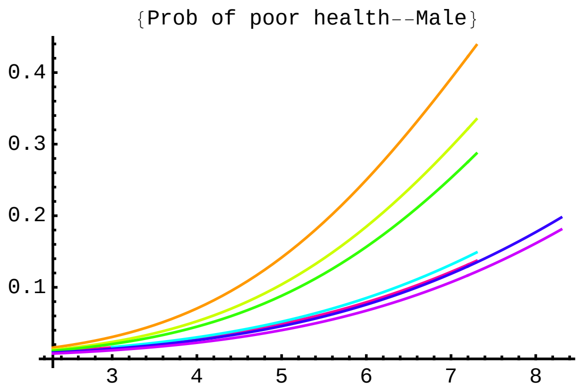
light blue=former-smoker, 30 years

green=current smoker, 1/2 pack/day

light green=current smoker, 1 pack/day

orange=current smoker, 2 packs/day

- **Figure 6D: Probability of self-reported poor health as a function of age and smoking history (Source: poorhealthgraphs3.nb)**



Probability of self-reported poor health vs decades after age 17
red=never-smoker
purple=former-smoker, 10 years
dark blue=former-smoker, 20 years
light blue=former-smoker, 30 years
green=current smoker, 1/2 pack/day
light green=current smoker, 1 pack/day
orange=current smoker, 2 packs/day

Section 4: Annual Health Care cost models as functions of dynamic smoking variables .

Parameter Estimates for Expected Medical Expenditures for Current LC5 Treatment

Since current-treatment is determined by hospital ICD-9 codes, and since all currently treated respondents have either hospital stays or have had laboratory work done, currently-treated respondents all have positive expenditures. The cost for LC5 treatment was estimated with the determinants of the logarithm of annual expenditures for respondents with LC5 treatment. Only treatment for LC5 mattered. Log expenditures equaled 7.847. The t value was 62.6. The smearing coefficients for retransformation40, calculated by smoking status, were: never-smokers' smearing coefficient=1.0138; current-smokers' smearing coefficient=3.366; and former-smokers' smearing coefficient=3.326.

Log[**expenditure**] = (β**0**) + ε

where $\varepsilon \sim$ **N**[**0**, σ^2]

{ParameterTable →

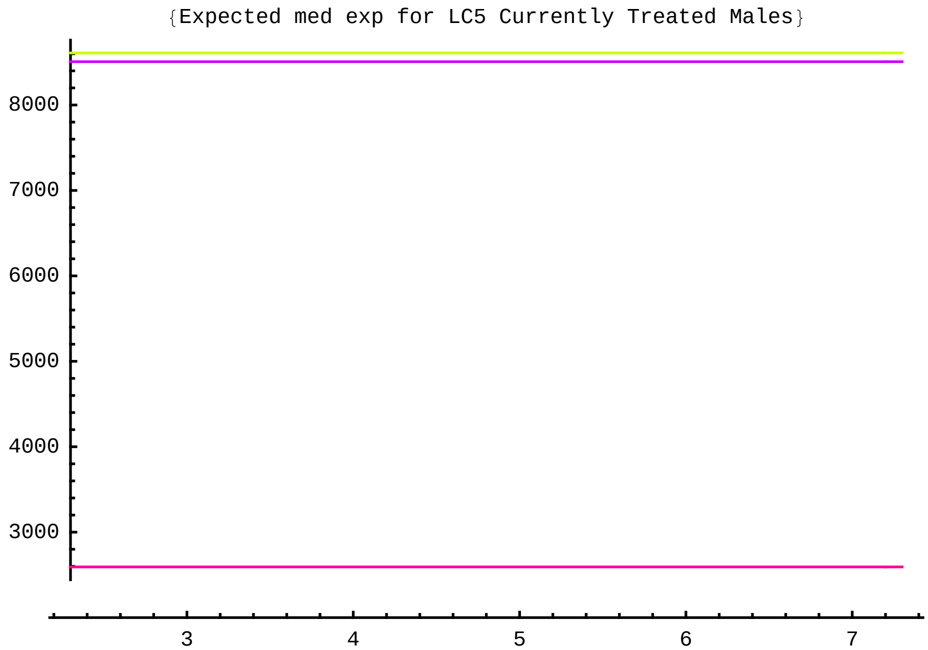
""	"Estimate"	"SE"	"TStat"	"PValue"
"Constant"	7.84688	0.12542	62.5652	0

,
RSquared → 0, AdjustedRSquared → 0, EstimatedVariance → 2.7685}

- **Table 6: Parameter Estimates of the Expected Logarithm of Medical Expenditures for Current Treatment of LC5**

Variable	Parameter	Estimate	Sta Error	t - value	Smoking Status	Ψ = Smearing Coefficient
Constant	$\beta 0$	7.847	0.125	62 .6	Never	1 .0138
					Current	3 .366
					Former	3 .326

- **Figure7A: Expected medical expenditure on males who were treated within the year for LC5 (Source:/male expenditures/7232004explc52.nb)**



- Graphics -

Bottom to top (at 6.4 decades after mean initiation time (17 years)) :
never - smoker (red)
former - smokers, (purple)
current - smokers, (light green)

Parameter Estimates for Expected Logarithm of Medical Expenditures for CHD5 Treatment

The CHD5 cost model estimated the determinants of the logarithm of annual expenditures for respondents with CHD5 treatment. Health expenditures were found to be a function of ever-smoker status and the expected latent index of poor health status. Note that the expected poor health status model, which was estimated with persons not currently treated, was here used with the people currently treated for CHD5 to obtain an expected value. As seen above, the expected latent index of poor health status is a function of tobacco exposure. The more exposure the higher the expected poor health status, the higher expected CHD5 expenditures. The never-smoker’s smearing coefficient equaled 2.918, the current-smoker’s smearing coefficient equaled 3.397, and the former-smoker’s smearing coefficient equaled 2.830.

Log[expenditure | expenditure > 0] = β0 + β1 ever + β2 healthstar + ε

{**ParameterTable →**

""	"Estimate"	"SE"	"TStat"	"PValue"
"Constant"	7.81761	0.217024	36.0219	0
"Eversmoker	−0.307497`	0.1624249	−1.8932	0.059
"Health*^	0.42511	0.199207	2.13401	0.033

,

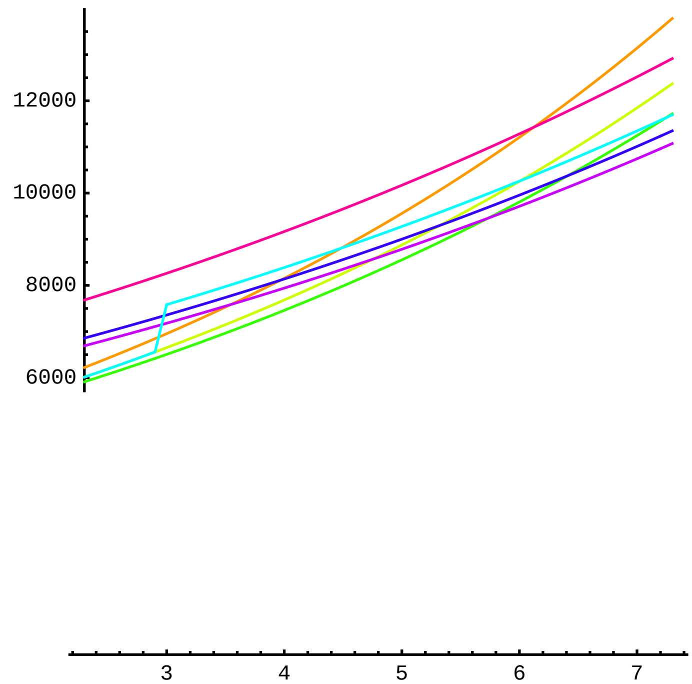
RSquared → 0.0109, AdjustedRSquared → 0.0076`, EstimatedVariance → 2.326,

Table7:Parameter Estimates for Expected Medical Expenditures for Current Treatment of CHD5

Variable	Parameter	Estimate	Sta Error	t - value	Smoking Status	Ψ = Smearing Coefficient
Constant	β0	7.8176	0.2170	36.02	Never	2.918
Eversmoker Status	β1	-0.3075	0.1624	-1.893	Current	3.3970
health*^	β2	0.4251	0.1992	2.134	Former	2.830

- **Figure 7B: Medical Expendiutre of males annually treated for CHD5 (Source:/male expenditures/7232004expCHD52.nb)**

{Expected med exp for CHD5 Currently Treated Males, curr,1/2pk/day}



- Graphics -

```
Bottom to top  (at 7. decades after mean initiation time (17 years)) :
  former - smoker, 10 yr, 1 pk/day (purple)
  former - smoker, 20 yr, 1 pk/day ( blue)
  current - smoker, 1/2 pk day  (darker green)
  former - smoker, 30 yr, 1 pk/day (light blue)
  current - smoker, 1 pk/day  (lime)
  never - smoker (red)
  current - smoker, 2 pks/day  (tan)
```

- **Parameter Estimates for Expected Medical Expenditures for men who are not currently treated for smoking caused diseases.**

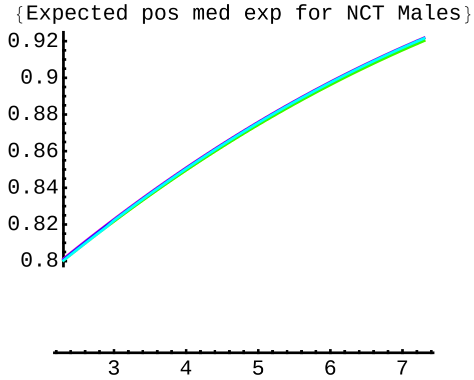
A two part model was used to estimate the health care expenditures for respondents who were not currently treated. The first part, a model of the probability of any health care expenditure, was estimated as a Probit model with exponential heterogeneous variance. The propensity was specified as a function of age, expected poor health status, and indicators of current or former smoking status. Controlling for the expected value of the latent index of poor health status, the smoking status indicators are demand parameters. Only the age and poor health status parameters were significant. The second model of the two part model is a model of the logarithm of the level of positive expenditures. It was specified like the probability model. Former-smoker status, age (measured in decades greater than 17 years of age) and poor health status contributed positively to these expenditures. The never-smoker’s smearing coefficient equaled 4.246, the current-smoker’s smearing coefficient equaled 3.594, and the former-smoker’s smearing coefficient equaled 3.289.

E[Medical Expenditures] =
(CDF[NormalDistribution[0, Exp[φ1 curr + φ2 form],
β0 + β1 Age + β2 Health$^{*\wedge}$ + β3 curr + β4 form])*
Exp[γ0 + γ1 Age + γ2 Ehealthstar + γ3 curr + γ4 form])*Ψ

- **Table8a:Parameter Estimates for Probability Medical Expenditures Greater Than Zero for Not Currently Treated (NCT)**

Variable	**Parameter**	**Estimate**	**Standard Error**	**t - value**
Current smoker in variance	φ1	0.2851	1 .0849	0.263
Former smoker in variance	φ2	- 0.5178	0.851	- 0.609
Constant	β0	0.5772	0.196	2.952
Age (in decades > 17 years)	β1	0.1184	0.0867	1.366
Health$^{*\wedge}$	β2	- 0.01543	0.322	- 0.0479
Current smoker status	β3	- 0.1204	0 .432	- 0.279
Former smoker status	β4	- 0.03287	0 .420	- 0.0783

- **Figure 8: Illustration of Prob. of Medical Expenditure for NCT(Source:7232004probexpNCTtheoryadj.nb)**



- Graphics -

Note : The health effects from smoking do not influence the probability of any expenditures for persons who are NCT.

- Graphics -

Log[exp] = β0 + β1 curr + β2 form + β3 aage + β4 healthstar + ε

where ε ~ N[0, sig^2]

{ParameterTable →

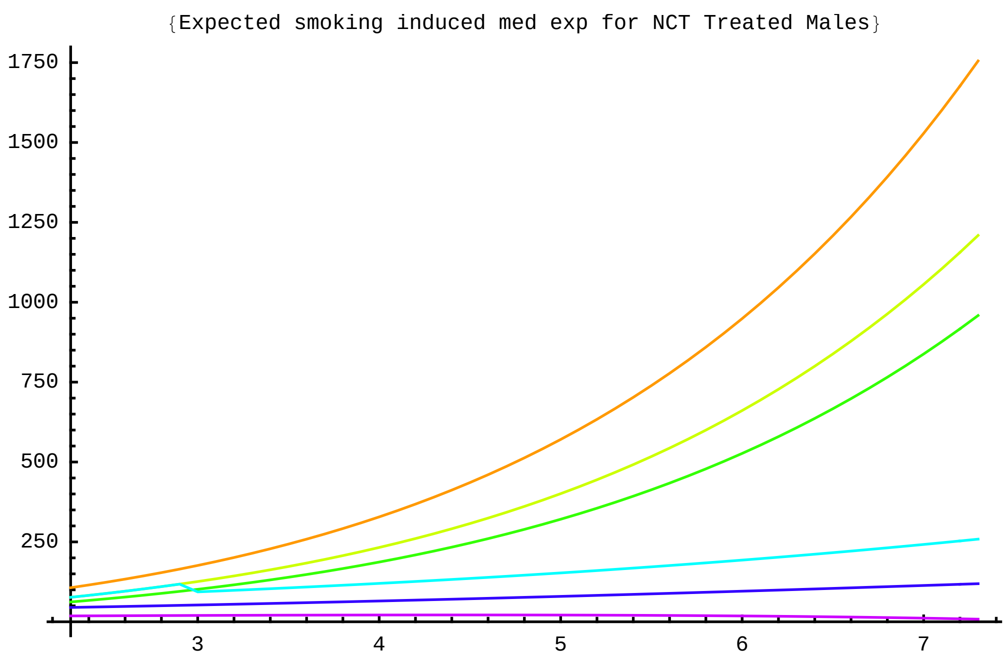
""	"Estimate"	"SE"	"TStat"	"PValue"
"Constant"	5.34852	0.18854	28.3682	0
"Current"	-0.10996	0.13019	-0.8446	0.398
"Former"	0.11245	0.06081	1.84927	0.064
"AdjAge"	0.15451	0.08675	1.78122	0.075
"Healthstar"	0.44798	0.31135	1.43885	0.150

RSquared → 0.0494, AdjustedRSquared → 0.0484, EstimatedVariance → 2.2208

- **Table8b:Parameter Estimates for Log Medical Expenditures for Not Currently Treated, Given Medical Expenditures Greater Than Zero**

Variable	Parameter	Estimate	Standard Error	t-value	Ψ= Smearing Coefficient
Never					4.246
Current					3.594
Former					3.289
Constant	γ0	5.3485	0.1885	28.37	
Age (in decades >17)	γ1	0.1545	0.1302	-0.845	
Health*^	γ2	0.4480	0.0608	1.849	
Current	γ3	-0.1100	0.0867	1.781	
Former	γ4	0.1125	0.3113	1.439	

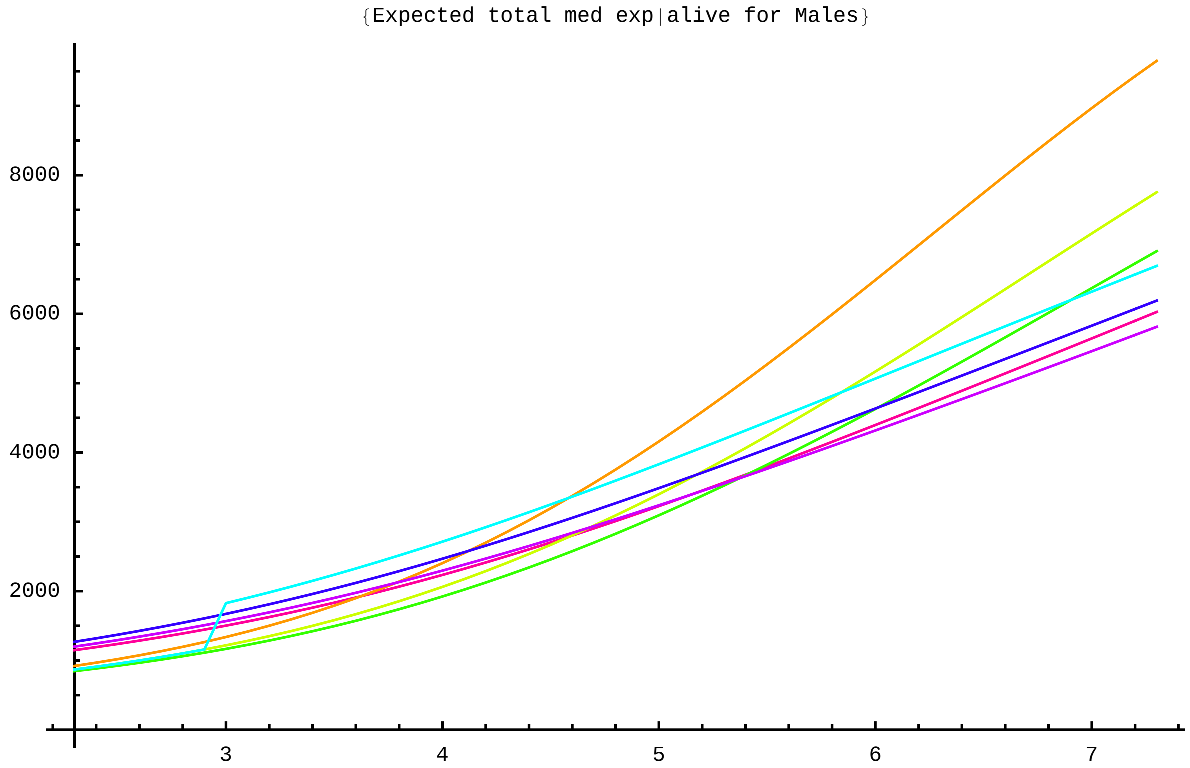
- **Illustration of results (Source : male expenditures/expectedsmokingnducedNCTexp2theoryadj.nb)**



- Graphics -

```
Expected smoking attributable medical expenditures for NCT males in 1987 dollars.
 Bottom to top  (evaluated at 7 decades of smoking):
    current smoker,  2 pk/day  (orange)
    current smoker, 1 pk/day  (lime)
    current smoker, 1/2 pk/day  ( green)
    former smoker, 10yr, 1 pk/day(purple)
    former smokers,20 yr, 1pk/day (blue)
    former smoker, 30yr,1pk/day(light blue),
```

- **Expected Total Medical Expenditures, given one is alive**



- Graphics -

(Source:Expected Smoking Induced Med Exp.nb)

Evaluated at 7.2 decades after the mean smoking initiation age (17 years))
Bottom to Top:

former smoker 10 yr, 1 pk/day (purple)
never smoker (red)
former smoker, 20 yr, 1 pk/day (blue)
former smoker, 30 yr, 1 pk/day (light blue)
current smoker 1/2 pk/day (green)
current smoker 1 pk/day (light green)
current smoker 2 pks/day (orange)

Evaluation of the Economic Impact of California' s Tobacco Control Program : A Dynamic Model Approach-- Appendix 4 : The Evaluation of TCP through Simulations with Computer Experimental Design.

Leonard S. Miller

Background

We have estimates of the age specific smoking initiation and smoking cessation rates in California over the 1990-1999 period, and estimates of these rates in the US as a whole. Additionally, we have adjusted these rates so that they are equal to the California rates in 1989. These rates provide factual and counterfactual smoking behavioral relationships in California over the 1990-1999 period in the presence of and in the absence of California's Tobacco Control Program (TCP). However, in fact, the consequences of smoking, in terms of medical resources attributable to smoking, cases of smoking attributable diseases, health status, person years of life saved, and the value of years of life saved, take a life time to be revealed. So, to estimate the value generated by California's TCP we need to estimate the full life consequences to California residents operating with smoking behaviors obtained with and without California's TCP.

We estimated these results by simulating the life-time outcomes for California's 1990 residents twice, once using the observed smoking initiation and quitting rates and once using the adjusted national smoking initiation and quitting rates. We then compared the outcomes from the two simulations and attributed the estimated difference to California's Tobacco Control Program. The basic information for this evaluation is: (1) a description of the California population, a sample, describing the age and smoking behavior of its residents in the base year 1990; (2) age specific estimates of the smoking initiation and quit rates over the decade of the nineties; (3) models to estimate probabilities of relevant events--death, disease, health status, and costs (given disease status) in the simulations, and random processes to convert the probabilities into events determining the calculation of the comparative simulated outcomes. We seek to estimate the distribution of the population's simulated outcome arising from the different smoking conditions and the random processes. However, owing to the large standard errors when the TCP evaluation is based directly on the California 1990 tobacco sample, we adopted a cell-replication design to represent that sample. The purpose of this appendix is to explain how we went about obtaining results from this comparative simulation design.

The following notation will help provide structure to the argument: X_s denotes a vector of the relevant characteristics of individuals in the sample describing the California population in the base year of the simulations; Y[s] denotes an outcome over the simulation based on sample observations; w[s] denotes the number of people in the population represented by sample observation s; and Ns is the number of observations in the sample. If the sample were used as the basis for comparing the simulations, the expected value of the outcome Y[s] is estimated by Ybar[s]. The estimate of the variance of Ybar[s] is estimated by $\left(S^2\right)_{\text{Ybar}}[s]$. We can calculate Y[s], but how do we estimate Ybar[s]? We can calculate $S^2$for the population from the sample, but how do we estimate $S^2{}_{\text{Ybar}}[s]$? In more detail, then, a principal purpose of this appendix is to address these questions efficiently and by addressing questions efficiently I mean for the calculation effort expended--that is, how do we create minimum variance estimates of the distribution of outcome in the population.

A Computer Experiment

Our strategy is to reformulate the information in the sample into a designed computer experiment and to estimate the answers sought from the experiment. Then, to transfer the knowledge back to the sample, which is then used to estimate the distribution of outcome in the population (Santner, Williams, Notz, 2003). To accomplish this, we construct a design representation of the population as described by the sample. We partition the space describing the relevant population characteristics into Nc disjoint cells. Then we represent the individuals in each cell with a prototypical individual with characteristics X_i, i=1,...Nc. Since any sample member will be subjected to the random processes required by simulations, understanding the distribution of outcome resulting from these processes requires replications of each prototypical individual in each cell i. Let J[i] denote the number of replications in cell i. The term w[i] is the number of people in the population represented by cell i. It is estimated by counting up the weights attached to the sample members who would occupy cell i. $S_{\text{Ybar}}[X[i]]$ is the standard deviation of the average outcome score for cell i. It is estimated by

$$S_{\text{Ybar}}[X[i]] = \sum_{i=1}^{Nc} (Y[X[i]] - \text{Ybar}[X[i]])^2 \Big/ (J[i] (J[i] - 1)).$$

Make J[i] simulations, calculate Ybar[X[i]] and $(S^2)_{\text{Ybar}}[X[i]]$. The product of w[i] and $(S^2)_{\text{Ybar}}[X[i]]$ estimates the variance of the outcome score in the population derived from cell i.

The term $\text{Sqrt}\left[\sum_{i=1}^{Nc} w[i] \, (S^2)_{\text{Ybar}}[X[i]]\right]$ is the estimate of the total standard deviation of the outcome score in the population represented by the designed computer experiment. If C denotes the total number of simulations to be made in the experiment, one for each replication in the design, how many replications should be made in each cell so as to minimize the estimate of the variance in the distribution of outcome in the population?

Theorem: On the efficient allocation of replications in a computer experiment.
The efficient allocation of C replications across the Nc cells follows from choosing the number of replications for each cell, J[i], according to cell i's fraction of the total standard deviation of the outcome score in the population. That is,

$$J[i] = C \sqrt{w[i]} \, \hat{S}[i] \Big/ \sum_{i=1}^{Nc} \sqrt{w[i]} \, \hat{S}[i].$$

Proof:
Since each cell i is represented by a single replication's description X_i , we can suppress the dependence of outcome on characteristics. The estimate of the average and variance of the population outcome from cell i is given by equations [1] and [2]:

[1]

$$\textbf{Ybar}_{\text{in Population}} = \sum_{i=1}^{Nc} (w[i] \, \text{Ybar}[i]) \Big/ \sum_{i=1}^{Nc} w[i]$$

[2]

$$(S^2)_{\text{in Population}} = \sum_{i=1}^{Nc} w[i] \, (S^2)_{\text{Ybar}}[i]$$

where:

Ybar[i] = $\sum_{j=1}^{J[i]} Y[i, j] / J[i]$;

$S^2[i] = \sum_{j=1}^{J[i]} \frac{(Y[i,j]-Ybar[i])^2}{-1+J[i]}$.; and

$\left(S^2\right)_{Ybar}[i] = S^2[i]$ /J[i]

To simplify the exposition, assume $S^2[i]$ is estimated independent of the J[i] determination process. For example, a two step estimation of $\left(S^2\right)_{Ybar}[i]$ is made. First, with a relatively small sample, $S^2[i]$ is estimated for the purpose of understanding how replications should be allocated to cells, and then $\left(S^2\right)_{Ybar}[i]$ is estimated with the J[i] replications for the purpose of furthering the outcome analysis.

The Lagrangian of the variance minimization expresses the objective function, the variance arising in the Nc cells, subject to the conditions that the sum of the replications in all the cells equals the number of calculations C and that the sum of the weights in the cells equals the population size. The Lagrangian is as follows:

$$\mathfrak{L} = \sum_{i=1}^{Nc} \frac{w[i]\ S^2[i]}{J[i]} + \lambda \left(-C + \sum_{i=1}^{Nc} J[i]\right) + \mu \left(-P + \sum_{i=1}^{Nc} w[i]\right)$$

The optimization problem is to minimize $\mathfrak{L}$ over the choice of the set J[i].

Taking the partial derivative of the Lagrangian with respect to J[i], and setting it to zero yields the first of the first order conditions for the estimate of the variance of outcome,

$$[3] \quad \lambda - \frac{w[i]\ S[i]^2}{J[i]^2} = 0$$

Taking the partial derivative of the Lagrangian with respect to the first constraint, λ, and setting it to zero yields the second of the first order conditions for the estimate of the variance of outcome,

[4] $-C + \sum_{i=1}^{Nc} J[i] = 0$

and taking the partial derivative of the Lagrangian with respect to the second constraint, μ, and setting it to zero yields the third of the first order conditions for the estimate of the variance of outcome,

[5] $-P + \sum_{i=1}^{Nc} w[i] = 0$.

Equation [3] actually represents Nc first order conditions of the form

[6] $\lambda = \frac{S^2[i]\ w[i]}{J[i]^2}$,

which all have the following solution for the optimal number of replications, J[i],

[7] $\text{Solve}\left[J[i]^2 - (1/\lambda)\, S^2[i]\, w[i] == 0, J[i]\right]$

$$\left\{\left\{J[i] \to -\frac{\sqrt{S^2[i]}\,\sqrt{w[i]}}{\sqrt{\lambda}}\right\}, \left\{J[i] \to \frac{\sqrt{S^2[i]}\,\sqrt{w[i]}}{\sqrt{\lambda}}\right\}\right\}$$

Accepting the positive valued solution with a positive square root yields equation [8],

[8] $J[i] = \frac{\sqrt{w[i]}\ S[i]}{\sqrt{\lambda}}$.

Now incorporate the second of the first order conditions, equation [4], $\sum_{i=1}^{Nc} J[i] = C$, into the analysis. Substituting the solution for J[i] into equation [4], obtains equation [9],

[9] $$-C + \sum_{i=1}^{Nc} \frac{\sqrt{w[i]}\ S[i]}{\sqrt{\lambda}} == 0$$

which can be solved for $\sqrt{\lambda}$,

[10]

$$\sqrt{\lambda} = \sum_{i=1}^{Nc} \frac{\sqrt{w[i]}\ S[i]}{C}$$

Substitute this solution for $\sqrt{\lambda}$ back into the solution for J[i] (equation [8]) and we have proved our theorem,

[11] $J[i] = C\left(\sqrt{w[i]}\ S[i] \Big/ \sum_{i=1}^{Nc} \sqrt{w[i]}\ S[i]\right)$.

- **An algorithm to determine the optimal number of representations in a cell.**

The estimation of J[i] requires estimates of w[i] and $\hat{S}$[i]. By adding up the weights of every one in the sample represented by cell i we estimate w[i]. That is,

[12] $w[i] = \sum_{j=1}^{J[i]} w[s|s \in i]$.

Take a reasonable, but small number of replications for every cell. Perhaps 30. Conduct the simulations for each cell and from the resulting outcome measures, estimate S[i] with the use of equations [1] and [2].

Based on these estimates for w[i] and S[i], for every i, compute $\sqrt{w[i]}\ S[i]$ and $\sum_{i=1}^{Nc} \sqrt{w[i]}\ S[i]$, and then the fraction of the contribution of cell i to the standard deviation in the total outcome, fJ[i], is given by equation [13].

[13] $fJ[i] = \frac{\sqrt{w[i]}\ S[i]}{\sum_{i=1}^{Nc} \sqrt{w[i]}\ S[i]}$.

Having decided to make C calculations and hence requiring C replications, the number for cell i is simply the product of fJ[i] and C,

[14] J[i] = fJ[i] C.

- **How many calculations, C, should be made?**

Let us assume that at the end of the analysis we desire a coefficient of variation (σ/μ) to have an estimated value (S/Ybar) equal to α. The coefficient of variation, estimated by $\left(\left(S^2\right)_{\text{in Population}}\right)^{1/2} / \mathbf{Ybar}_{\text{in Population}}$, where these terms are given by equations [2] and [1], respectively.

From estimates of equation [2],

[15] $\left(\hat{S}^2\right)_{\text{in Population}} = \sum_{i=1}^{Nc} \left(w[i]\ S^2[i] \,/\, J[i]\right)$,

substitute in the value of J[i] from equation [11]]. The variance in the population is given by equation [16],

[16] $\left(S^2\right)_{\text{in Population}} = \sum_{i=1}^{Nc} \left(\left(w[i]\ S^2[i]\right) \Big/ \frac{\sqrt{w[i]}\ S[i]}{\sum_{i=1}^{Nc} \sqrt{w[i]}\ S[i]}\ C\right)$.

Simplify, and then solve for C.

$$C = \left(\sum_{i=1}^{Nc} \left(\sqrt{w[i]}\ S[i]\right)\right)^2 \Big/ \left(S^2\right)_{\text{in Population}}$$

The coefficient of variation, denoted by α, is estimated by $\left(\hat{S}\right)_{\text{in Population}}$ /Ybar $_{\text{in population}}$

the [17] $C = \left(\sum_{i=1}^{Nc} \left(\sqrt{w[i]}\ S[i]\right)^2 \Big/ \left(\alpha^2\ \text{Ybar}_{\text{in Population}}{}^2\right)\right.$

An algorithm to determine the required number of calculations.

The w[i] values are data and the initial estimates of $\hat{S}$[i] and Ybar[i] are obtained from the initial experiment. Employing equation [1] yields an estimate of $\mathbf{Ybar}_{\mathbf{in\ Population}}$ (= $\sum_{j=1}^{J[i]}$ [w[i] Ybar[i]). For a given value of α, C is calculated from equation [17] and distributed among the Nc cells according to equation [14].

- **An analysis of the designed computer experiment. A transformation of the experiment information into sample knowledge.**

At this point assume the computer experiment has been conducted and we have obtained a vector of average outcomes for the cells, Ybar[i], and a vector of the standard deviations in outcome for the cells, $S_{\text{Ybar}}[i]\}$. The task now is to transform these statistics about the computer experiment into knowledge about the sample that can be used to estimate knowledge about the population.

We relate the statistics from the computer experiment with a multiplicative heteroscedastic regression model. This is the specification examined in depth by Harvey (1976), but our formulation is different because Harvey had no estimates of the variance of Ybar[i] and we do. And, accordingly, our results will differ from his. The model has the form specified by equation [18],

[18] Ybar[i] = X[i] $\beta + \epsilon[i]$

E[ϵ[i] ϵ[i]'] =$\sigma^2[i]$ =
= Exp[Z[i]γ]=Exp[γ0] Exp[Z1[i] γ1] ...Exp[Zp[i]γp]
= $\sigma_0{}^2[i]$ Exp[Z1[i] γ1] ...Exp[Zp[i]γp]
= $\sigma_0{}^2[i]$ Exp[Z*[i] γ*]

for all i, where:
X[i] is a row vector, 1 x P, of the P descriptive characteristics of the prototypical member of cell i,
β is a vector of length P,
γ is a vector of length K, and
ϵ[i] is a random variable indicating the difference between cell [i]'s average outcome and the cell i's expected outcome, given its characteristics, X[i].

If Mϵ is a vector of the cell's random variables, MYbar is a vector of the cell average outcomes, MX is a Nc by P matrix of the cell characteristics, relevant to describing Ybar and MZ is a Nc by K matrix of the cell characteristics relevant to describing the variance in Ybar, then the expected value vector, variance-covariance matrix, and estimate of the variance-covariance matrix is given by equations [19a], [19b], and [19c] respectively,

[19a] E[Mϵ]=0 ; and

[19b] E[MϵMϵ']=

Exp[$Z[i=1]\gamma$], 0 , ... , 0
0 , Exp[$Z[i=2]\gamma$], ... , 0

Σ =[ ].
0, 0 , ... , Exp[Z[i = Nc] γ]

[19c] $M\left(S^2\right)_{\text{Ybar}}$ = Ψ =

$\left(S^2\right)_{\text{Ybar}}$[X[i=1]], 0 , ... , 0
0 ,$\left(S^2\right)_{\text{Ybar}}$[X[i=2]], ... , 0
= [ ].
0 , 0, ... , $\left(S^2\right)_{\text{Ybar}}$[X[i=Nc]]

Since estimates of the variance are known, generalized least squares provides minimum variance estimates b of β.

[20] b = (X[i]' Ψ^{-1}X[i]) (X[i]' Ψ^{-1}Y[i]).

We turn now to the estimation of the γ coefficients. Based on the estimates b of β, an observed error in the model for Ybar[i] is given by

[21] e[i] = Ybar[i] - X[i] b.

The logarithm of the square of this observed error is the estimate of the variance for an observation, which by the postulated multiplicative heteroscedastic model, is

[22] Log[$e[i]^2$] = Z[i] γ + ν[i].

In our case, we have an estimate of this variance, so equation [21] can be expressed as equation [22],

[23] Log[$\left(S^2\right)_{\text{Ybar}}$[i]] = Z[i] γ + ν[i].

Let c denote the least squares estimator of γ. c is given by equation [24],

[24] c = $(Z' Z)^{-1}$Z' Log[$\left(S^2\right)_{\text{Ybar}}$[i]].

We now examine the characteristics of this estimator. Our analysis is similar to that of Harvey (1976), though somewhat simpler because the dependent variable, Log[$\left(S^2\right)_{\text{Ybar}}$[i]], is observed. From equation [17], Z[i] γ = Log $\sigma^2[i]$. After substituting into equation [23], and solving for the error term, we have equation [25],

[25] $v[\mathrm{i}] = \mathrm{Log}[\,(S^2)_{\mathrm{Ybar}}[\mathrm{~i~}]\,] - \ln \sigma^2[i] = \mathrm{Log}[\,(S^2)_{\mathrm{Ybar}}[\mathrm{~i~}]\,/\,\sigma^2[i]]$

Under the assumption that the deviations from the means of a cell are Normal, $\mathrm{Log}[\,(S^2)_{\mathrm{Ybar}}[\mathrm{~i~}]\,/\,\sigma^2[i]]$ is distributed as the natural logarithm of a Chi-Squared distribution with J[i] degrees of freedom divided by J[i] (recall J[i] are the number of observations in cell [i]) the error term v[i] is so distributed.

The expected value of the Logarithm of a Chi-Squared with one degree of freedom equals -1.27036.

```
Integrate[Log[x] PDF[ChiSquareDistribution[1], x], {x, 0, Infinity}]
```

```
-EulerGamma - Log[2]
```

```
N[%]
```

```
-1.27036
```

By subtracting the expected value of the error from the constant, c0, and from the error v[i], we obtain,

$$[26]\quad E[c] = \gamma + (Z'Z)^{-1} Z' E[\mathrm{Log}[v[i]]]$$
$$= \gamma + (Z'Z)^{-1} Z' (-1.27036)$$

which implies

$$[27]\quad E[\tilde{c}] = [c - 1.27036\, i] =$$
$$= \gamma + (Z'Z)^{-1} Z' E[\mathrm{Log}[v[i] - 1.27036]]$$
$$= \gamma$$

where $i = \{1,0,...,0\}'$, of length P.

The variance of a random variable distributed as the logarithm of a Chi-Squared with one degree of freedom has a value equal to $\frac{\pi^2}{2}$, which equals 4.9348.

```
[28] Integrate[
  (Log[x] - (-EulerGamma - Log[2]))^2 PDF[ChiSquareDistribution[1], x], {x, 0, Infinity}]
```

$$\frac{\pi^2}{2}$$

```
N[%]
```

```
4.9348
```

The variance of $\tilde{c}$ is given by equation [29],

$$[29]\quad V[c] = V\left[(Z'Z)^{-1} Z'(Z\gamma + v[i])\right]$$

$$= (Z'Z)^{-1} (Z'\, V[v[i]] Z)\, (Z'Z)^{-1}$$

$$= (Z'Z)^{-1}\, 4.9348$$

- **Applying the cell analysis to the sample**

We turn now to use the estimated models to estimate the mean outcome and its variance, given the sample data.. Let X[s] and Z[s] describe the relevant model characteristics with a sample member s. We wish to forecast the distribution of the average outcome for sample member s and its variance. The average outcome forecast for sample member s is given by equation [30],

[30] $\hat{\mathrm{Y}}\,\mathrm{bar}[s] = X[s]\,\mathrm{b} + \in[s]$,

where ϵ[s] is the forecast error. The Gauss-Markov theorem insures that the minimum variance linear unbiased estimator of the forecasted average outcome for sample member s , $\hat{Y}$bar[s], is given by equation [31],

[31] $\hat{\hat{\mathrm{Y}}}\,\mathrm{bar}[s] = X[s]\,\mathrm{b}$.

Let S^2_{forcast} denote an estimate of the variance in the forecast error. Based on the cell data, we estimate the variance of the forecast error with equation [32],

[32] $\hat{S}^2_{\text{forcast error}} = \sum_{i=1}^{Nc} \left(-\mathrm{b}\,\mathrm{X}[i] + \hat{\mathrm{Y}}\,\mathrm{bar}[i]\right)^2$/(Nc-1).

The estimate of the variance of the mean of a sample observation, given X[s], is based on the models estimated and the sample information. Let $(\hat{\hat{S}}^2)_{\text{Ybar}}$[s] represent the estimate of the variance of a forcasted value of Ybar based on sample data. This variance should be the sum of the estimated variance of Ybar, given sample data, plus the variance of the forecast error. The expression for $(\hat{\hat{S}}^2)_{\text{Ybar}}$[s], simply combines the predicted value of the estimate of the variance of Ybar and the average forecast error calculated with equation [32],

[33] E[$(\hat{\hat{S}}^2)_{\text{Ybar}}$[s]]= E[Exp[Z[s] γ + v[s]] + $\hat{S}^2_{\text{forcast error}}$

= E[Exp[Z[s] $\tilde{c}$]] + $\hat{S}^2_{\text{forcast error}}$ =

- **An example of a model specification**

To illustrate the ideas developed above, we offer an example of a specification for the model presented by equation [18]. The specification of X[i] might be described by equation [34] :

[34] $X[i] = \{1, X1[i], X1[i]^2, X1[i]^3, X22[i], X23[i], X3[i], X4[i], X5[i]\};$

$\beta = \{\beta 0, \beta 11, \beta 12, \beta 13, \beta 22, \beta 23, \beta 3, \beta 4, \beta 5\}';$

and the individual X elements are defined as follows:
X1=age;
X2={1/0} according as observation is an {ever-smoker in 1990/otherwise};
X22={1/0} according as observation is a {former smoker in 1990/otherwise};
X23={1/0} according as observation is a {current smoker in 1990/otherwise};
X3=start age of a current or former smoker in 1990;
X4=quit age of a former smoker in 1990;
X5=cigarettes smoked per day,Modulo 1/2.

and the specification of the Z might be:

[35] Z[i]= {1, X1[i], X1[i], X2[i], X4[i], X5[i]}

and accordingly, $\gamma = \{\gamma 0, \gamma 11, \gamma 12, \gamma 2, \gamma 4, \gamma 5\}'$.

- **The Evaluation of the Program with sample knowledge.**

Having estimated the average outcome and its variance for every sample member, the TCP program is then evaluated. The outcome in the population is estimated from the sample data with equation [36],

[36] $Y_{\text{in Population}} = \sum_{s=1}^{Ns} \hat{\hat{Y}}\text{bar}[s]\, w[s],$

and the variance of the outcome in the population is estimated from the sample data with equation [37],

[37] $\left(S^2\right)_{\text{in Population}} = \sum_{s=1}^{Ns} w[s]^2 \left(\hat{\hat{S}}^2\right)_{\text{Ybar}}[s].$

In fact, since we have many different outcomes of interest, a particular one must be choosen to determine the design structure. We choose years of life saved and the design criteria.

- **Optimum allocation of simulation calculations.**

1. Allocate 30 replications to each cell, generate the simulation outcomes, and estimate the average and variance of these cell outcomes, Ybar[i] and $S^2[i]$, for each cell.

2. Compute w[i] for each cell.

3. Estimate the set {Ji} of replications per cell.

The following algorithm is then implemented to obtain a simulation result.

- **1. Model Ybar[i]**
- **2. Model $(S)_{\mathrm{Ybar}}[i]$**
- **3. Estimate the average outcomes in the sample, based on cell data.**
- **4. Estimate the variance of the average outcomes in the sample, based on cell data.**
- **5. Estimate the average outcomes in the population, based on sample data.**
- **6. Estimate the variance of the average outcomes in the population, based on sample data.**
- **7. Report the point and interval estimates of the outcomes attributable to TCP.**

- **Implementation of the cell-replication design.**

In the study at hand, a cell is described with five values in 1990:
X1=age, range 1 to 90 ;
X2=smoking status 1,2,3;
X3=start age 11 to 22 ;
X4=quit age 20 to 90 ; and
X5=smk packs/day 0 to5/2, in units of 1/2 packs per day, 0-1/2, 1/2-1,...

Cells were determined by age and smoking status, and then were further described by average age of smoking initiation (smkage), average age of smoking cessation (quitage), and average number of cigarettes smoked per day (smkperdy). In the table listing the cells, which follows directly, frequency is the number of individuals in the sample in the cell, weight is the number of individuals in the California 1990 population in the by the cell, and replications are the number of identically described individuals in a cell that are used in the analysis to follow. The 55 cells are described as follows:

age	status	smkage	quitage	smkperdy	frequency	weight	replications
2	1	0	0	0	2042	882747	2327
5	1	0	0	0	1831	807553	2314
8	1	0	0	0	1877	799532	2266
11	1	0	0	0	1579	695679	2351
11	3	8	0	2	1	88	500
14	1	0	0	0	1274	604482	2021
14	3	11.7863	0	6.039	30	10886	500
17	1	0	0	0	862	630586	2072
17	3	13.884	0	11.3733	223	109865	778
17	2	14.603	17.736	12.7459	36	22767	500
20	1	0	0	0	329	406278	1668
20	3	15.6778	0	14.4162	307	179183	970
20	2	15.3473	14.3752	19.2006	116	82487	615
23	1	0	0	0	279	413696	1718
23	3	16.6516	0	12.476	299	153686	1004
23	2	15.8488	21.5762	13.7742	154	113557	793
26	1	0	0	0	305	560850	1995
26	3	16.785	0	14.8342	333	191420	1165
26	2	16.3209	22.9694	13.3254	173	133264	820
30	1	0	0	0	484	727392	2490
30	3	16.9693	0	15.6027	654	402773	1768
30	2	16.3804	26.1504	14.4243	388	314757	1424
35	1	0	0	0	402	598588	2216
35	3	17.4031	0	17.9864	647	359521	1772
35	2	17.0151	28.7542	13.9436	438	366977	1738
40	1	0	0	0	295	467274	2176
40	3	17.4678	0	18.5935	630	313846	1806
40	2	17.1184	31.151	11.4854	461	362637	1798
45	1	0	0	0	211	288589	1771
45	3	17.3644	0	20.0742	466	220500	1721
45	2	17.5803	33.4554	15.2068	441	338298	1883
50	1	0	0	0	141	258740	1699
50	3	16.4539	0	20.5198	353	198828	1600
50	2	17.288	39.0741	15.0388	384	336349	1962
55	1	0	0	0	99	195249	1369
55	3	17.1051	0	24.8793	248	125179	1204
55	2	16.6765	40.8493	16.5595	286	276852	1691

60	1	0	0	0	91	145895	1104
60	3	16.8282	0	21.7766	216	113109	1010
60	2	16.9499	44.1983	15.0165	260	230949	1429
65	1	0	0	0	63	106838	838
65	3	15.7986	0	24.5115	161	98611	848
65	2	17.8541	47.9391	15.3725	269	292991	1462
70	1	0	0	0	58	93696	691
70	3	16.9293	0	19.8263	121	60467	550
70	2	17.7366	48.4007	14.4596	213	222179	1029
75	1	0	0	0	49	92327	500
75	3	17.8749	0	15.4282	50	27482	500
75	2	16.9098	50.4745	16.7271	132	166162	702
80	1	0	0	0	21	39627	500
80	3	14.716	0	13.9163	17	9883	500
80	2	17.0443	49.9471	18.5443	74	100345	500
85	1	0	0	0	15	21600	500
85	3	15.469	0	13.571	8	5548	500
85	2	17.2158	49.0652	13.2088	20	34877	500

I ran factual and counter-factual simulations on each replication in each cell. The number of replications, last column in the table above, is the maximum of 500 and the fraction of the proportion of the total standard error estimated to be contributed by a cell. A total of 72128 replications were estimated so as to yield a coefficient of variation equal to 0.3. The replication calculations were based on a preliminary analysis with 500 replications per cell.

Six different algorithms were used to evaluation estimates of the effect of TCP. For each of these algorithms, the cell mean and cell standard error was computed for the factual and counter factual simulations. A cell's outcome was calculated as the difference between the mean of the cell's outcomes in the factual simulation and the mean of the cell's outcome in the counter-factual simulation. The program result was the sum of the population-weighted mean cell differences. The program standard error was the square root of the average of the sum of the population-weighted cell variances.

Bibliography

Santner, TJ, B.J.Williams, and W.I.Notz, 2003. The Design and Analysis of Computer Experiments.Springer.New York.