Can Changing Patterns in Smoking Prevalence among Women and Men Explain the Narrowing Gender Gap in Life Expectancy? A Cross-National Perspective*

Roland Rau† Marleen Dettmann‡

March 7, 2008

Paper Prepared for: Session 169:
" Gender and Health Risks over the Life Course" (Chair: Marilyn Sinkewicz)
PAA Annual Meeting 2008, April 17–19, New Orleans

Preliminary Results:
Please Do Not Quote Without Authors' Permission

Abstract

The gap in life expectancy between women and men first increased and then decreased during the second half of the 20th century. It has been argued that the more widespread smoking behavior of men played a major role in the widening, whereas the recent narrowing has been attributed to women's increased smoking behavior during later time periods. A recent article by Preston and Wang (2006) emphasized

*This work is supported by a NIH/NIA Grant ("Oldest-old Mortality — Demographic Models and Analysis", PO1-AG008761) and also by the Max Planck Institute for Demographic Research in Rostock, Germany

†Corresponding Author; Duke University; Duke Population Research Institute (DuPRI) & Terry Sanford Institute of Public Policy; Population, Policy, and Aging Research Center (PPARC); Box 90309; 302 Towerview Road; Durham, NC 27708-0309; USA; Email roland.rau@duke.edu

‡Robert-Koch-Institut, Abt. für Infektionsepidemiologie, Seestraße 10, 13353 Berlin, Germany, Email: DettmannM@rki.de

the impact of cohort effects in the life expectancy gap. Using data from the Human Mortality Database and IMASS Smoking Database, we analyzed both hypotheses in a comparative perspective for several countries. Our results find support for both hypotheses: a) We discovered a clear relationship in a period perspective only for moderate differences in the smoking prevalence gap. If excess mortality of men was larger than 22 percentage points, we could not detect any changes in (lagged) life expectancy. b) Previous smoking experience in a cohort seems to be a good indicator for subsequent mortality; the results for the cohort hypothesis are, however, based on considerably less data (25%) than for the period approach.

1 Introduction

Women live longer than men. The difference in life expectancy between females and males, though, did not remain constant over time. Glei and Horiuchi (2007) demonstrate that in most countries the typical difference was between two to four years for long periods of time. Beginning in the late 1940s and the 1950s, an interesting development in most developed countries has been observed: after gaining an advantage until the mid-1970s, the gap in life expectancy (at birth) between women and men decreased in almost all developed countries since then. Figure 1 (Page 15) captures this development. As indicated by the asterisks, marking the maximum gap for each country, life expectancy differed between women and men in the United States in 1975 about 7.7 years and decreased ever since. In 2004, the year with the most recent data available, the difference was slightly more than 5 years. The one exception to this general pattern is found for Japan where a peak in the difference between female and male life expectancy has been reached only in recent years.

What mechanisms can produce this development? To gain a preliminary answer and some first insight, we plotted in the upper panel of Figure 2 (Page 16) a surface over age and time depicting the gap in death rates between women and men in the United States from 1933 until 2004 for age 0 to 100. The higher life expectancy of women as shown in the lower panel of this figure, is a direct translation of lower female mortality at virtually all ages. With the exception of young adult mortality (ages 18–30), the advantage of females is decreasing since the mid-1970s at almost all ages. As indicated by the lower panel of Fig. 2, the resulting differences in life expectancy are often separated into biological reasons on the one hand and into social/behavioral reasons on the other hand.

An example for a biological survival advantage of women is the presence of two X chromosomes among females (Christensen et al., 2001). Although it is not straightforward to disentangle social/behavioral and biological causes, there seems to be a consensus that biological factors are of minor relevance. Vallin, for example, cites Pressat's calculation of an advantage of 2 years of life expectancy at birth (Vallin, 2006). In his study of Bavarian nuns and monks — to control for social and behavioral factors — Luy concludes that "biological factors at most confer a survival advantage for women of about one year in remaining life expectancy at young adult ages" (Luy, 2003, p. 668).

That leaves the majority of the gap between women and men to be explained by behavioral and social factors (see also lower part of Figure 2). Although an array of theories is offered (see for an overview, for example, Luy (2002)), a large share of the research literature devotes its attention to one of these factors in particular: *smoking*.
At least since the early 1950s, the hazardous nature of smoking has been recognized, triggering — most notably but not exclusively — (lung) cancer and diseases of the circulatory system such as stroke and myocardial infarction (among many, see for example: Doll et al. (1994), Doll et al. (2004) or McBride (1992)). For example, Doll et al. (1994) conclude after studying mortality in relation to smoking for 40 years among male British doctors: "It now seems that about half of all regular cigarette smokers will eventually be killed by their habit" (p. 901). Recent estimates of the "Centers for Disease Control and Prevention" (CDC) claim that "[e]very year, smoking kills more than 276,000 men and 142,000 women" in the United States (Centers for Disease Control and Prevention (CDC), 2006). On a global scale, the World Health Organization consideres tobacco to be the biggest risk factor in developed countries being responsible for more than 12% of 214 million DALYs (World Health Organization (WHO), 2002, p. 83)[1]

According to Doll's "Tobacco: A Medical History" (1999), the wide-spread adoption of smoking started in the 1920s and by the end of World War II, about 80% of British males were smokers. Women started smoking at a later period of time, [i]n some [...] developed countries, such as France and Spain, only in the last two or three decades women begun smoking" (p. 291).

[1]One DALY (disability-adjusted life year) is equal to the loss of one healthy life year (World Health Organization (WHO), 2002, p. 12).

Having the knowledge of the detrimental health consequences of smoking, and the coincidence of increasing smoking prevalence among men and the subsequent widening in the gender-gap in life-expectancy, it should be not surprising that researchers started to determine the contribution of the smoking differential on the life expectancy differential. Retherford (1972), for example, estimated that 47% of the female-male differences in life expectancy between ages 37 and 87 in the year 1962 were due to smoking. According to his estimates, even 75% of the increase in the difference in life expectancy for the same age-range between 1910 and 1962 was caused by tobacco. Valkonen and van Poppel (1997) show for Denmark, Finland, the Netherlands, Norway and Sweden that more than 40% of the gender-gap in life expectancy can be attributed to smoking in the years 1970–1974.[2]

Likewise in the opposite direction, it has been argued that the diminishing difference between female and male life expectancy since the mid 1970s can be traced back to the reduced differential in smoking between women and men. Nault (1997, p. 38), for example, is rather careful in his assessment: "trends in smoking and being overweight, risk factors for some leading causes of death, have mirrored the narrowing gap in male and female life expectancy." A more direct advocate is Fred C. Pampel. In 2002, p. 96, he concludes: "The results presented here do not suggest that cigarette smoking fully accounts for the sex differential in mortality between males and females; *rather, smoking fully explains the recent narrowing of the sex differential*".[3]

It has been argued, however, that the cohort perspective is important when assessing the impact of smoking on mortality. An example is the slow increase in female life expectancy in Denmark for many years. Jacobsen et al. (2002) showed that this is mainly caused by the high smoking prevalence of Danish women born between the two World Wars. Recently, Preston and Wang (2006) applied such an approach to study the impact of smoking on gender gaps in mortality in the United States. They "demonstrated that a cohort's smoking history prior to age 40 has a powerful impact on the cohort's subsequent mortality" (p. 641).

We would like to ask therefore two questions in our paper:

[2]To be precise, Valkonen and van Poppel (1997) refer to remaining life expectancy at age 35.

[3]Emphasis added by RR and MD.

Period Hypothesis If the differential development of smoking behavior between women and men is actually the main contributor to changes in the gender-gap in life expectancy, we should see a (lagged) correlation between the gender smoking prevalence gap and the gender life expectancy gap.

Cohort Hypothesis Do our data allow to detect strong cohort effects not only for the US as shown by Preston and Wang (2006) but also for other countries?

2 Data & Methods

2.1 Data Sources

We used two data sources in our analysis. Data on mortality, exposures and population counts were downloaded from the Human Mortality Database University of California, Berkeley (USA), and Max Planck Institute for Demographic Research, Rostock, (Germany) (2008). All smoking related data were obtained from the "International Mortality and Smoking Statistics" (IMASS) Database (Lee Statistics and Computing Ltd, 2006). The IMASS Database includes data on prevalence of smoking and consumption of smoking. We chose the prevalence data since they are available for longer time periods for most of the countries. In Table 1 (page 17), we give an overview for which time intervals data are available in the IMASS Database as well as in the Human Mortaliy Database.

2.2 Data Quality

As usual, the final results of a quantitative analysis are highly dependent on reliable data. Due to its widespread use and the rare cases of critcisms, we did not check for any data quality issues in the Human Mortality Database. Although we did not have considerable previous experience with the the IMASS Database, we are confident about the data quality as we conducted several checks — all having positive results:

- The favorable review in the *International Journal of Epidemiology* (Lawlor, 2004)
- In Germany, the IMASS database cites a consumption of 6.3 cigarettes[4] per adult per day for the period 1991–1995. According to the WHO, the annual per capita consumption of cigarettes in Germany in 1995 was 2297. Divided by 365 days results

[4] All cigarettes: Manufactured cigarettes and other cigarettes

in a daily consumption of 6.29 cigarettes per day. We interpret this to be an indicator of good data quality.

- Merte Osler published data on smoking habits of Danes derived from surveys in 1953–1954, 1986–1987 and 1990–1991(Osler, 1992). In Table 2 (page 18), we compared those data to the ones in the IMASS database and concluded that they match fairly well — especially for the first wave. The larger differences for the two later time periods can probably be explained by the smaller sample sizes for 1986/87 and 1990/91.[5]

2.3 Data Preparation

2.3.1 Estimation of 1x1 Prevalence Surfaces

Data in the IMASS database are given in 5-year age-groups as well as in 5-year calendar-time categories. The first step was to "expand" those data to single years and ages to obtain smoking prevalence surfaces. We did this by distributing the 5x5 grid onto 1x1 grid elements. The second step was to estimate a smooth surface of smoking prevalence. We did this by following the approach of Currie et al. (2004) to smooth and forecast mortality rates. The model is basically a two-dimensional poisson regression where the death rates are estimated using the age- and year specific population exposures as an offset and *P*-splines[6] as regression bases. Please see Figure 3 (page 19) for an illustration of the smoothing effect on the prevalence surfaces.

2.3.2 Period Smoking and Mortality Indicators

Our *period smoking* indicator was the smoking prevalence either of women or of men in a given single year at ages 15–50. This indicator was estimated using the aforementioned 1x1 prevalence surfaces. We age-standardized our data with the sex-specific population in the United States in the year 2000 using conventional demographic methodology (Preston et al., 2001, pp. 22–28). Since many previous studies focused on the impact of smoking on life expectancy at birth, we simply chose this indicator as our *period mortality* measurement for women and men.

[5] The sample sizes were: 1953/54: 30,018; 1986/87: 4,753; 1990/91: 4,818.

[6] *P*-Splines are *B*-Splines with a penalty on the regression coefficients. See for an introduction to *P*-Splines: Eilers and Marx (1996).

2.3.3 Cohort Smoking and Mortality Indicators

The *cohort smoking* indicator for each sex was the smoking prevalence in a single birth cohort at ages 20–40. This age-range is comparable to Preston and Wang (2006).[7] We used again the same age-standardization as for the period smoking indicator.

The *cohort mortality* indicator was the mortality of each birth cohort at ages 40–65. It would have been desirable to follow cohorts until they reach higher ages. This was not possible, though, due to data limitations. Nevertheless, we believe that the impact of smoking should already be present by the age of 65: Using lung cancer deaths and mortality as an indicator for the impact of smoking,[8], you can see in Figure 4 (page 20) that lung cancer mortality peaks almost at age 65 (upper panel) and less than half of all lung cancer deaths occur after age 65.

2.4 Methods

Our main method of analysis are scatterplots: For the *period analysis* we plot the life expectancy gap against the gap in smoking prevalence between women and men. Since the detrimental impact of smoking does not affect life expectancy immediately, we plotted life expectancy with a lag of 20 years. This is in accordance with the findings of Gajalakshmi et al. (2000, p. 24): "Across the populations of the industrialized countries with a history of prolonged smoking, past consumption trends predicts current tobacco-attributable mortality remarkably with a lag of about 20 years."

For the *cohort analysis* we plot the difference in death rates ($m(Men) - m(Women)$) at ages 40–65 against the difference in the cohort smoking experience (men minus women) at ages 20–40.

Hence, this analysis is following the advice of William S. Cleveland who advocates: "Data display is critical to data analysis. Graphs allow us to explore data to see the overall pattern and to see detailed behaviour; no other approach can compete in revealing the structure of the data so thoroughly" (Cleveland, 1994, p. 5).
Detecting relations in the scatterplot has been facilitated by adding (occassionally) the

[7] Preston and Wang (2006, p. 435): "We converted these data into an estimate of the average number of years spent as a current smoker before the age of 40."

[8] It is argued that not every smoker develops lung cancer, but that almost all lung cancer cases happen to smokers. REF?

LOWESS scatterplot smoother (Cleveland, 1979). This method smoothes the data using locally weighted polynomial regression. Adding such a line to the plot will allow us to get more insight into the functional relationship between independent and dependent variable. The corresponding computer routine (Cleveland, 1981) has been implemented into the R Language (R Development Core Team, 2006) which we used for all analyses in this paper.

3 Results

Given our hypotheses are correct, we expect to see a positive relationship in each of our scatterplots, i.e. the larger the gap in smoking prevalence (period or cohort), the larger the lagged gap in life expectancy (period hypothesis) or in mortality at ages 40–65.

Period Results Figures 5–7 (pages 21–23) illustrate our results for the *period hypothesis*. The x-axis in each of the plots denotes the difference in smoking prevalence between men and women. The y-axis gives the gap in life expectancy at birth twenty years after the corresponding gap in smoking prevalence. For example, the gap in smoking prevalence in the year 1970 is plotted together with the gap in life expectancy in the year 1990. To obtain a comprehensive picture, we plotted not only each country in a separate panel (Figures 6 to 7) but also pooled together in a single figure (Fig. 5).

Our results indicate that there is no uniform pattern among the selected countries. Some of the countries show the expected monotonous relationship (see Figures 6 and 7), e.g. Austria, Canada, Finland, Germany (West), or the Netherlands. Other countries display a rather concave parabolic shape. Examples are Australia, France, Italy or Sweden. According to our plots, the US displays also such a pattern. It is caused, however, by a few data-points marking the earliest time-period where the smoking prevalence data deviate completely from the rest of the data. Only Japan differs from the expected relationship completely.

Pooling all countries (Fig. 5) exposes several interesting patterns: first, in only about one percent of all cases (7 out of 607), depicting seven years in Sweden, smoking is more common among women than among men. Across all cases, the difference in smoking is between 22 (median) and 25 (mean) percentage points. The largest gaps are observed in Japan where on average about two thirds of all men were smoking during our observation period but only 13 percent of women. The second interesting feature is exposed by

the solid black line which represents the LOWESS scatterplot smoother for all data: if we choose to split the data at the median prevalence gap value of 0.22, we can detect two different patterns: if the gap in smoking prevalence is smaller than 22 percentage points, there appears to be a rather linear relationship between the two variables. We applied a linear regression to this part of the data and found a slope of $+0.1346$: until 22 percentage points any additional increase by one point in the male-female smoking prevalence gap corresponds to a larger life expectancy gap of 0.135 years or slightly more than 1.5 months. Higher values in the smoking prevalence gap appear to have little effect on the life expectancy gap, seemingly reaching a "saturation point": our linear regression for this part of the data results in a slope estimate of -0.022 years. If Japan is excluded, the slope is even closer to zero with -0.0078.

Cohort Results The results for our cohort analysis are given in Figure 8 (page 24) where all countries are pooled together and in Figure 9 (page 25) in a separate panel for each country. Due to higher data requirements — we needed to follow each cohort for 45 years (ages 20–65) — we have less data-points in each country or even had to exclude several countries.

On the x-axis, we plotted again the gap in smoking prevalence (Male minus Female). This time, we restricted ourselves to ages 15–40. The y-axis denotes the difference in mortality at ages 40–65 in the corresponding cohort.

The general picture supports our initial hypothesis: (again) with the exception of Japan, each of our 11 countries of analysis displays a monotonous, positive relationship between the two variables. The degree of the relationship differs, though. Countries such as Sweden or Norway exhibit a rather moderate increase. Other countries like the United States or Australia show a more pronounced change in mortality for a given change in the smoking prevalence gap.

4 Summary

The second half of the twentieth century experienced an increase and a decrease in the life expectancy gap between women and men. The largest difference was observed in most countries between the mid-1970s and mid-1980s. Numerous articles attributed large shares of this change to the more widespread habit of smoking among men than among women. Likewise, the narrowing in the life expectancy difference in recent decades could

be (partly) explained by the fact that women started smoking more frequently than previously. We looked at two different angles:

Period Hypothesis: Can we detect a positive relation between the difference in smoking prevalence and the difference in life expectancy twenty years later?

Cohort Hypothesis: Preston and Wang (2006) stressed the importance of a cohort component: the smoking history of a cohort is, according to them, a strong predictor of mortality. We asked whether we can replicate their findings for the United States and see comparable patterns also in other countries.

We conducted our analysis using data from the Human Mortality Database (University of California, Berkeley (USA), and Max Planck Institute for Demographic Research, Rostock, (Germany), 2008) and the IMASS database (Lee Statistics and Computing Ltd, 2006). When we were able to check the quality of the data from IMASS with other published information, we were surprised by the similarity. Nevertheless, we have to acknowledge that some data could be problematic and could cast a shadow on our results. For example, the first decade of US data differs completely from the rest of the US data.

In our analysis we found support in our data for both hypotheses: In the case of the *period hypothesis*, our results were two-fold: if the prevalence gap was smaller than 22 percentage points, we observed a linear increase of about 0.135 years in the life expectancy gap. If the prevalence gap was larger, though, we could not find any indication for a change in the life expectancy gap. How can we interpret this? It seems that only until a certain point, male excess smoking behavior translates to a change in the life expectancy gap. Higher differences than 22 percentage points in smoking prevalence among men than among women, seems to be unimportant for the gap in life expectancy.

We discovered a clearer picture for the *cohort hypothesis*: in all of the 11 countries of analysis we found the expected positive relationship: the higher the prevalence of smoking among men in a given cohort, the larger is also the difference in mortality of the respective cohort between women and men. Our results, covering primarily cohorts born in the 1920s and 1930s, are also in accordance with research by Jacobsen et al. (2002) who argued that a major reason for the slow increase in life expectancy among Danish women is the increased smoking behavior of the cohorts born between the two World Wars. It has to be mentioned, though, that we had 75% less data available to check this cohort hypothesis than the period hypothesis.

References

Centers for Disease Control and Prevention (CDC) (2006). Smoking & tobacco use. fact sheet. cigarette smoking-related mortality (updated september 2006). Available online at `http://www.cdc.gov/tobacco/data_statistics/Factsheets/cig_smoking_mort.htm`.

Christensen, K., K.-H. Ørstavik, and J. W. Vaupel (2001). The X chromosome and the female survival advantage: an example of the intersection between genetics, epidemiology and demography. In M. Weinstein, A. I. Hermalin, and M. A. Stoto (Eds.), *Population health and aging: strengthening the dialogue between epidemiology and demography*, Volume 954 of *Annals of the New York Academy of Sciences*, pp. 175–183.

Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association 74*(368), 829–836.

Cleveland, W. S. (1981). LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *The American Statistician 35*(1), 54.

Cleveland, W. S. (1994). *The Elements of Graphing Data*. Murray Hill, New Jersey: AT&T Bell Laboratories.

Currie, I. D., M. Durban, and P. H. Eilers (2004). Smoothing and forecasting mortality rates. *Statistical Modelling 4*, 279–298.

Doll, R. (1999). Tobacco: A Medical History. *Journal of Urban Health: Bulletin of the New York Academy of Medicine 76*(3), 289–313.

Doll, R., R. Peto, J. Boreham, and I. Sutherland (2004). Mortality in relation to smoking: 50 years' observations on male British doctors. *British Medical Journal 328*, 1519–.

Doll, R., R. Peto, K. Wheatley, R. Gray, and I. Sutherland (1994). Mortality in relation to smoking: 40 years' observations on male British doctors. *British Medical Journal 309*, 901–911.

Eilers, P. H. C. and B. D. Marx (1996). Flexible Smoothing with B-splines and Penalties. *Statistical Science 11*, 89–102.

Gajalakshmi, C., P. Jha, K. Ranson, and S. Nguyen (2000). Global patterns of smoking and smoking-attributable mortality. In P. Jha and F. Chaloupka (Eds.), *Tobacco Control in Developing Countries*, Chapter 2, pp. 12–39. Oxford: Oxford University Press.

Glei, D. A. and S. Horiuchi (2007). The narrowing sex differential in life expectancy in high-income populations: Effects of differences in the age pattern of mortality. *Population Studies 61*(2), 141–159.

Jacobsen, R., N. Keiding, and E. Lynge (2002). Long term mortality trends behind low life expectancy of Danish women. *Journal of Epidemiology and Community Health 56*, 205–208.

Lawlor, D. A. (2004). A book and website all ije readers should know about forey b, hamling j, lee p, wald n. international smoking statistics (2nd edition). a collection of historical data from 30 economically developed countries. wolfson institute of preventive medicine, oxford: Oxford medical publications, 2002, pp. 854. £110. *International Journal of Epidemiology 33*, 433–434.

Lee Statistics, P. and Computing Ltd (2006). International Mortality and Smoking Statistics, Version 3. Available online at: `http://pnlee.co.uk/imass.htm`.

Luy, M. (2002). Die geschlechtsspezifischen Sterblichkeitsunterschiede — Zeit für eine Zwischenbilanz. *Zeitschrift für Gerontologie und Geriatrie 35*, 412–429.

Luy, M. (2003). Causes of Male Excess Mortality: Insights from Cloistered Populations. *Population and Development Review 29*(4), 647–676.

McBride, P. E. (1992). The Health Consequences of Smoking. Cardiovascular Diseases. *Medical Clinics of North America 76*(2), 333–352.

Nault, F. (1997). Narrowing mortality gaps, 1978 to 1995. *Health Reports 9*(1), 35–41.

Osler, M. (1992). Smoking Habits in Denmark from 1953 to 1991: A Comparative Analysis of Results from Three Nationwide Health Surveys among Adult Danes in 1953-1954, 1986-1987 and 1990-1991. *International Journal of Epidemiology 21*(5), 862–871.

Pampel, F. C. (2002). Cigarette Use and the Narrowing Sex Differential in Mortality. *Population and Development Review 28*(1), 77–104.

Preston, S. H., P. Heuveline, and M. Guillot (2001). *Demography. Measuring and Modeling Population Processes*. Oxford, UK: Blackwell Publishers.

Preston, S. H. and H. Wang (2006). Sex mortality differences in the United States: The role of cohort smoking patterns. *Demography 32*(4), 631–646.

R Development Core Team (2006). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0.

Retherford, R. D. (1972). Tobacco Smoking and the Sex Mortality Differential. *Demography 9*(2), 203–216.

University of California, Berkeley (USA), and Max Planck Institute for Demographic Research, Rostock, (Germany) (2007). Human Mortality Database. Available at www.mortality.org.

University of California, Berkeley (USA), and Max Planck Institute for Demographic Research, Rostock, (Germany) (2008). Human Mortality Database. Available at www.mortality.org.

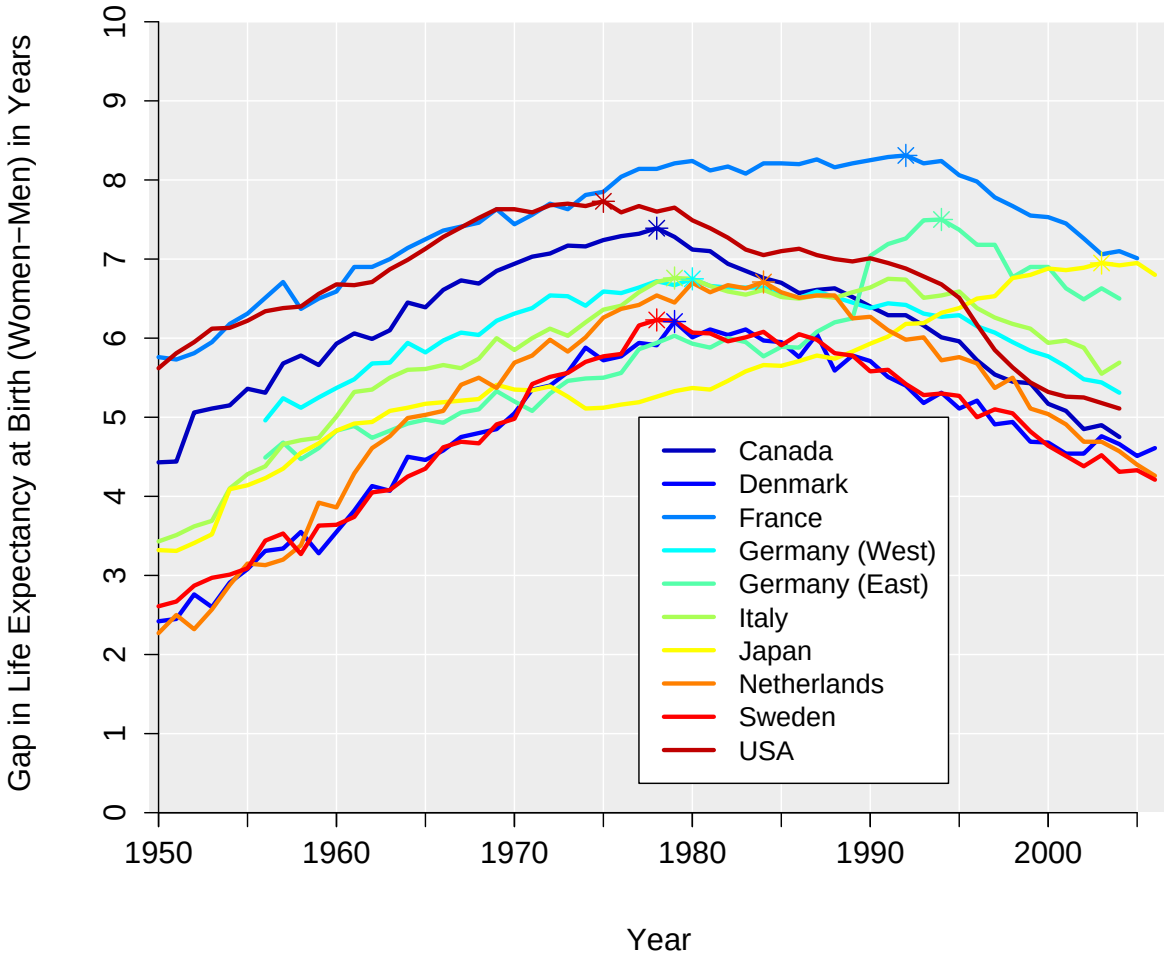
Valkonen, T. and F. W. van Poppel (1997). The contribution of smoking to sex differences in life expectancy : Four nordic countries and the Netherlands 1970–1989. *European Journal of Public Health 7*, 302–310.

Vallin, J. (2006). Mortality, Sex, and Gender. In G. Caselli, J. Vallin, and G. Wunsch (Eds.), *Demography. Analysis and Synthesis*, Volume II, Chapter 53, pp. 177–194. Amsterdam, NL: Elsevier.

World Health Organization (WHO) (2002). The World Health Report 2002. Technical report, World Health Organization (WHO), Geneva, Switzerland.

A Figures and Tables

Figure 1: Gap in Life Expectancy (at Birth) between Women and Men in Selected Countries, 1950–2006



The asterisks indicate the maximum gap for each country.

Source: Human Mortality Database 2007

The corresponding numerical values are tabulated in Tables 3 and 4

Figure 2: *Upper Panel*: Differences in Male-Female Death Rate in the United States, 1933–2004; *Lower Panel*: Translation of Gaps in Death Rates into Life Expectancy and its Distinction into Social/Behavioral and Biological Components

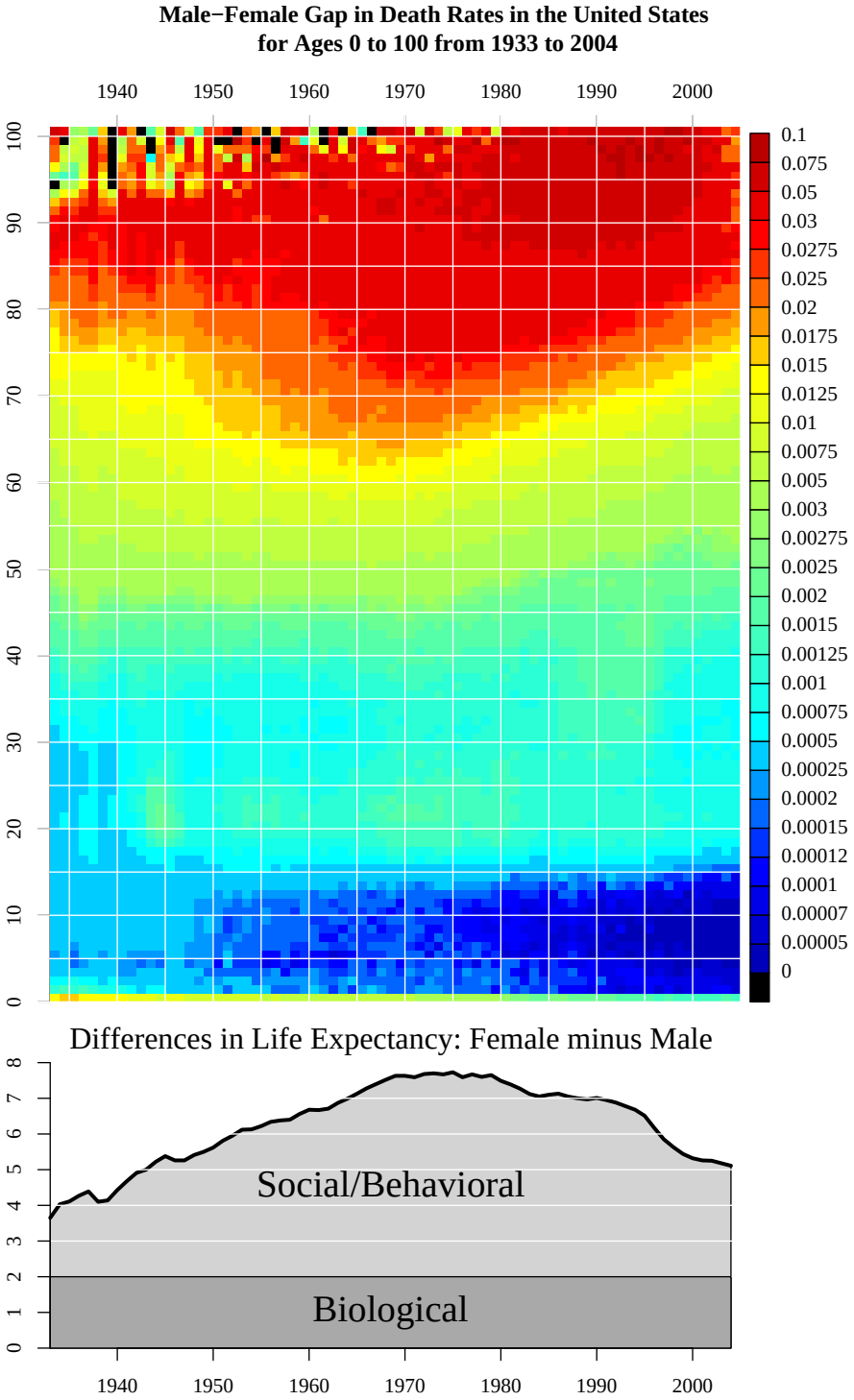


Table 1: Data Availability

Country	Abb.*	HMD Coverage† Years	IMASS Coverage‡ Years	Type§
Australia	AUS	1921-2004	1941/1945 – 2001/2005	AT
Austria	AUT	1947-2005	1951/1955 – 1991/1995	AT
Belgium	BEL	1841-2005	1951/1955 – 1991/1995	AT
Canada	CAN	1921-2004	1961/1965 – 2001/2005	TC
Denmark	DEN	1835-2006	1951/1955 – 1991/1995	AT
Finland	FIN	1878-2006	1956-1960 – 1991/1995	AT
France	FRA	1899-2005	1951/1955 – 1991/1995	AT
Germany (West)	FRG	1956-2004	1946/1950 – 1986/1990	AT
Hungary	HUN	1950-2005	1966/1970 – 1991-1995	AT
Iceland	ICE	1838-2006	1976/1980 – 1991/1995	AT
Italy	ITA	1872-2004	1946/1950 – 1991/1995	TC
Japan	JAP	1947-2006	1946/1950 – 2001/2005	AT
Netherlands	NLD	1850-2006	1956/1960 – 1991/1995	AT
New Zealand	NZL	1948-2003	1961/1965 – 1991/1995	AT
Norway	NOR	1846-2006	1951/1955 – 1991/1995	AT
Portugal	POR	1940-2006	1971/1975 – 1991/1995	AT
Spain	ESP	1908-2005	1966/1970 – 1991/1995	AT
Sweden	SWE	1751-2006	1946/1950 – 1991/1995	AT
Switzerland	CHE	1867-2006	1971/1975 – 1991/1995	AT
USA	USA	1933-2004	1931/1935 – 2001/2005	TC

∗ Abb: Abbreviation
† HMD: Human Mortality Database
‡ IMASS: International Mortality and Smoking Statistics
§ AT: Prevalence of Smoking: All Tobacco Products
TC: Prevalence of Smoking: Total Cigarettes

Table 2: Comparing the Data Quality of the IMASS Database with data published by Merete Osler

Males							
	Osler Data†				IMASS Data‡		
Age	1953-1954	1986-1987	1990-1991	Age	1951-1955	1986-1990	1991-1995
15-19	61	24	24	15-19	58.1	24.1	24.1
20-29	83	47	41	20-20	81.0	40.2	42.0
				25-29	82.0	47.1	47.1
30-39	85	53	55	30-34	83.3	51.2	49.3
				35-39	84.9	51.9	47.6
40-49	83	53	58	40-44	84.1	51.5	49.4
				45-49	83.3	51.6	52.2
50-59	79	59	44	50-54	81.7	56.5	49.8
				55-59	79.4	55.3	47.9
60-69	70	55	46	60-64	75.4	53.2	45.0
				65-69	70.2	51.5	44.9

Females							
	Osler Data†				IMASS Data‡		
Age	1953-1954	1986-1987	1990-1991	Age	1953-1954	1986-1987	1990-1991
15-19	41	35	30	15-19	41.5	27.0	25.8
20-29	59	50	44	20-20	64.2	41.6	41.8
				25-29	57.2	47.6	45.3
30-39	51	50	46	30-34	54.1	48.2	47.8
				35-39	48.3	46.2	45.8
40-49	38	48	44	40-44	40.8	44.0	44.5
				45-49	35.8	41.2	42.9
50-59	31	49	49	50-54	31.6	49.1	45.9
				55-59	27.0	43.2	40.2
60-69	20	43	35	60-64	20.8	41.2	36.3
				65-69	16.7	37.1	33.7

† Osler Data: Osler (1992)

‡ IMASS Data: Lee Statistics and Computing Ltd (2006)

Figure 3: Graphical Demonstration of Smoothing. The upper panel displays the smoking prevalence of Australian men from 1941 to 2005 for ages 15 to 89 as given in the IMASS Database. The lower panel illustrates the smoothed prevalence surface using the methodology mentioned in the main text. The colors indicate the prevalence of smoking in percent. Please see also the legend on the right hand side of each panel.

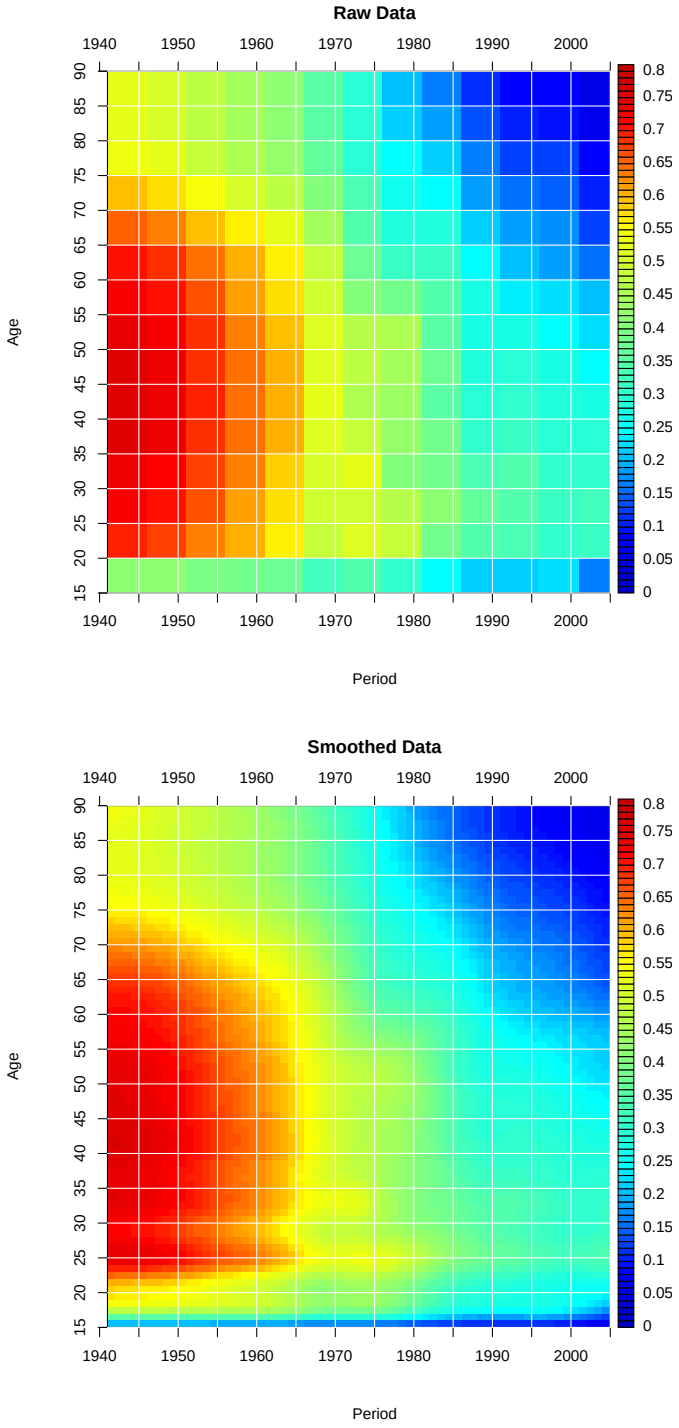
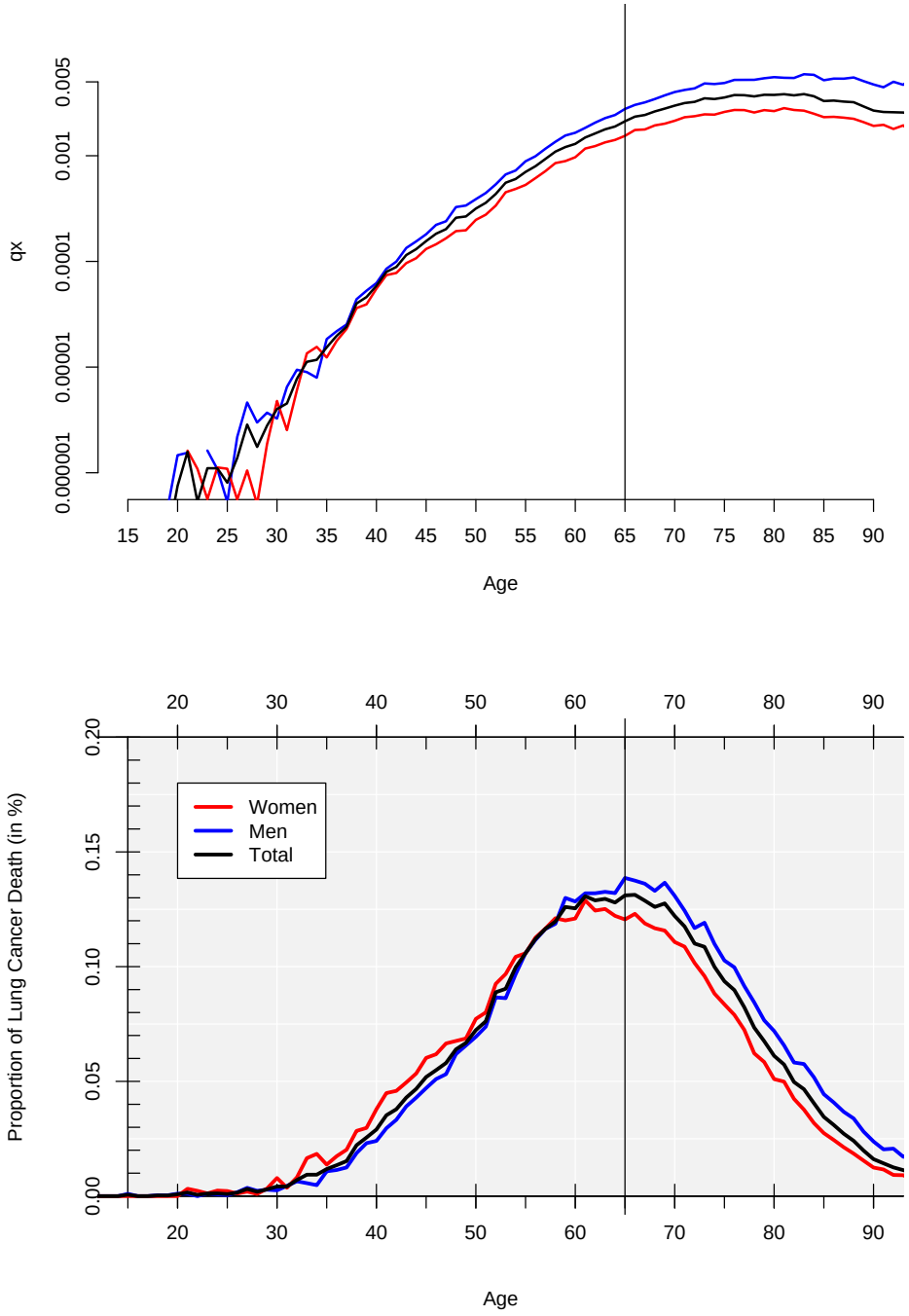


Figure 4: Development of Lung Cancer over Age in the Year 2000 in the United States for Women (red), Men (blue), and Total (black). *Upper Panel:* Probability of Dying ($q(x)$) *Lower Panel:* Proportion of Lung Cancer Deaths in Relation to All Deaths



Source: Death Data: Multiple Cause of Death Data; Population Data: Human Mortality Database; Authors' Own Estimations

Figure 5: Pooled Results for Period Hypothesis: Prevalence Gap in Smoking (Ages 15–50, Men minus Women) and the Gap in Life Expectancy at Birth (Women minus Men) twenty years later. The black solid line represents the LOWESS smoother for all data.

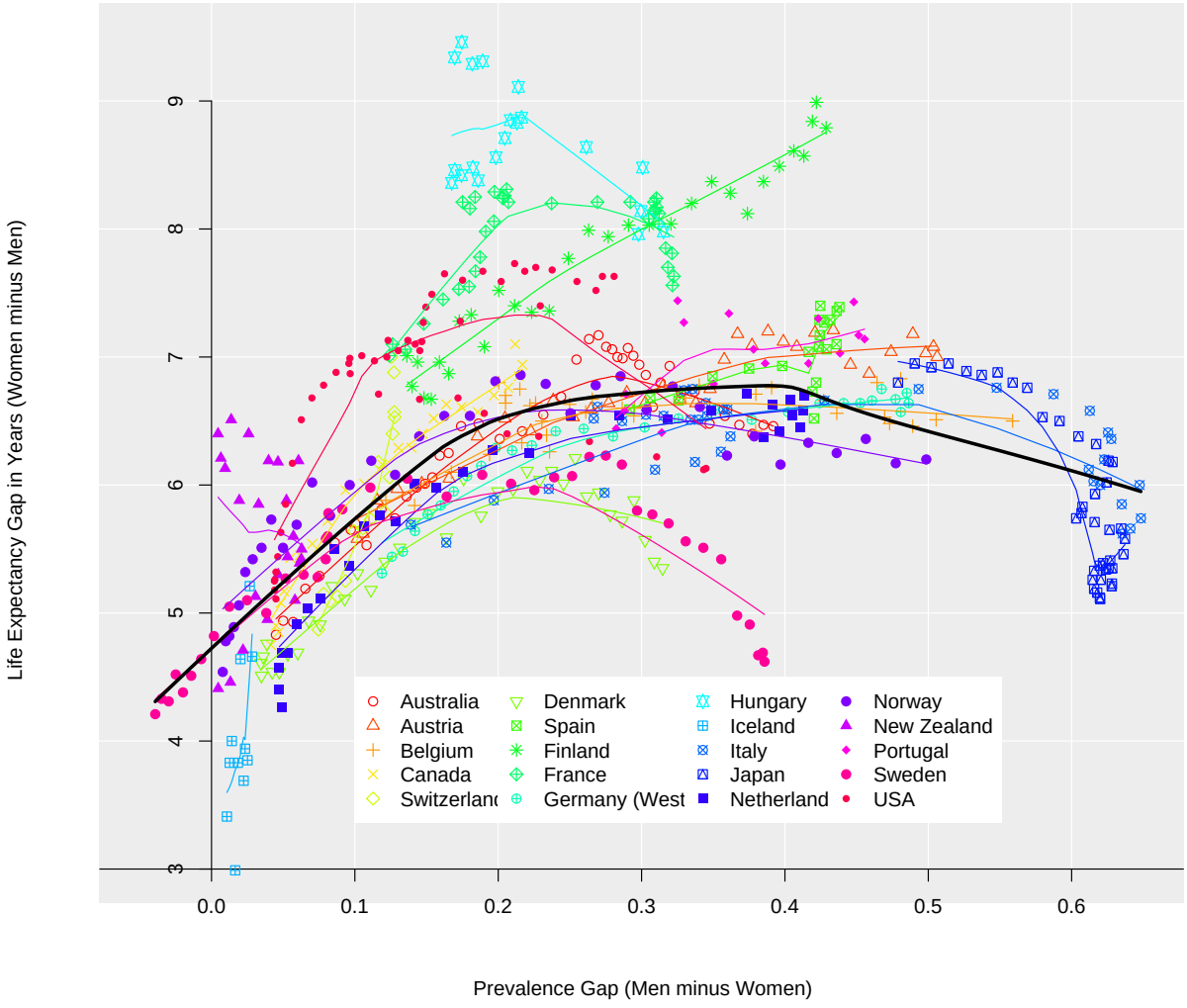


Figure 6: Results for Period Hypothesis, Selected Countries (I/II): Prevalence Gap in Smoking (Ages 15–50, Men minus Women) and the Gap in Life Expectancy at Birth (Women minus Men) twenty years later.

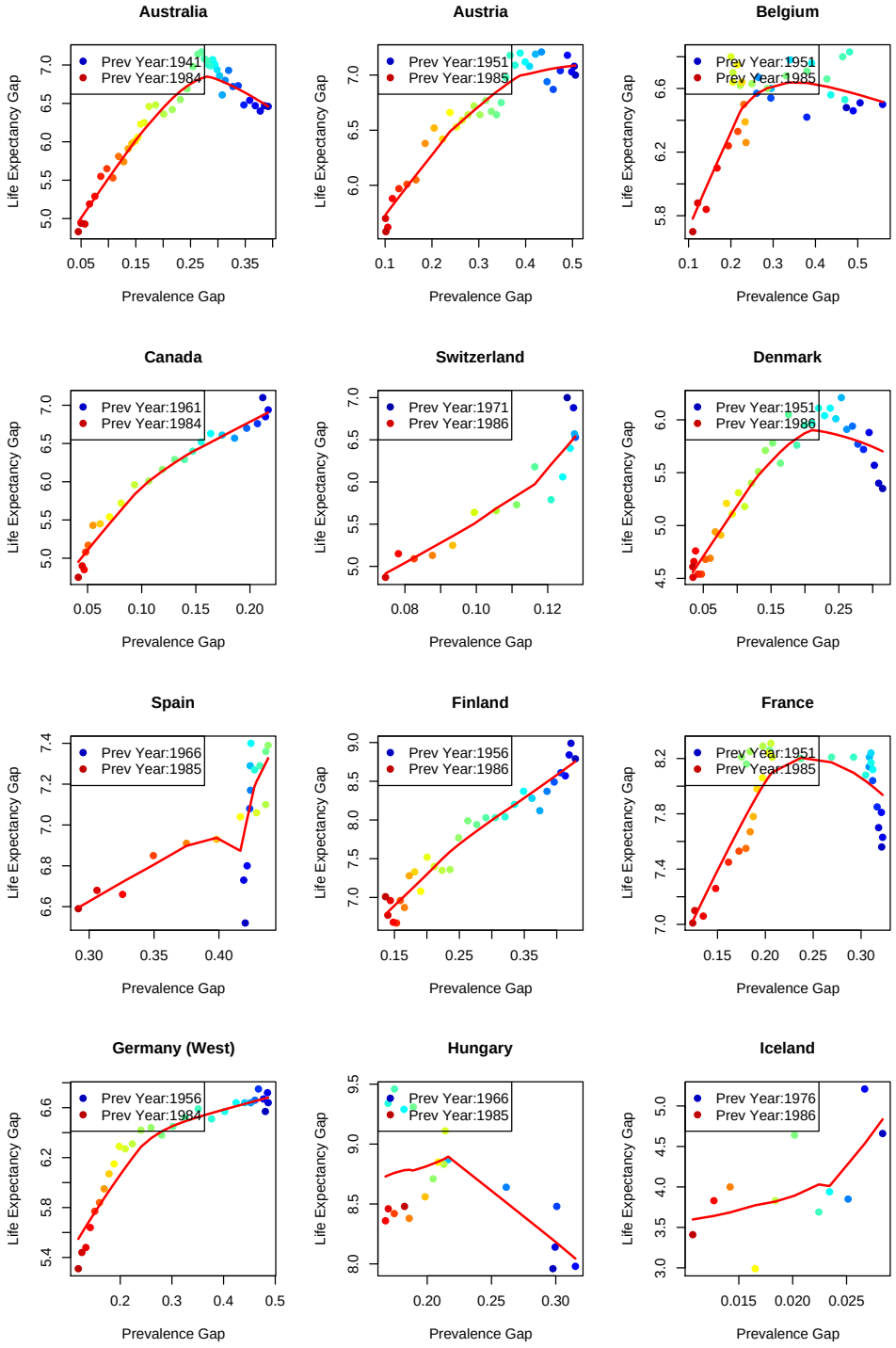


Figure 7: Results for Period Hypothesis, Selected Countries (II/II): Prevalence Gap in Smoking (Ages 15–50, Men minus Women) and the Gap in Life Expectancy at Birth (Women minus Men) twenty years later.

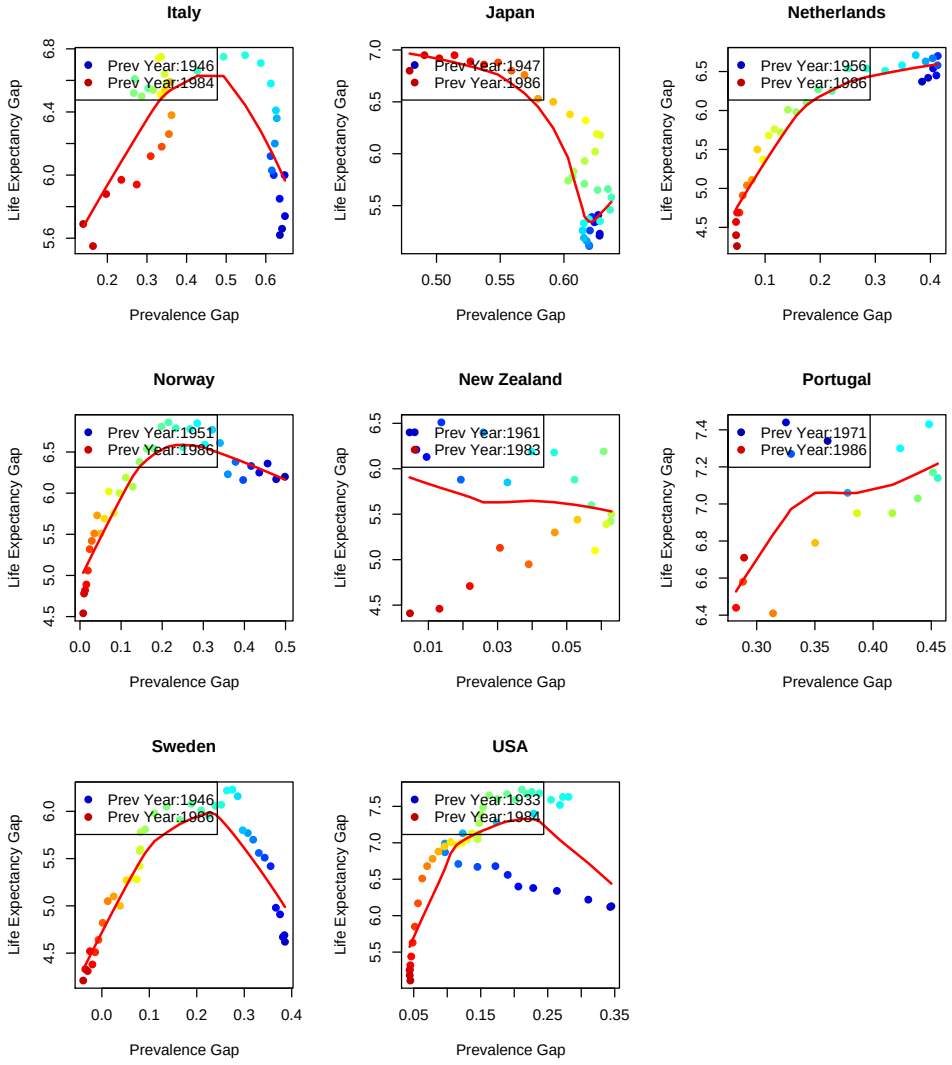


Figure 8: Pooled Results for Cohort Hypothesis: Prevalence Gap in Smoking (Ages 15–40, Men minus Women) and the Gap in Mortality (Men minus Women) in the Respective Cohorts at Ages 40–65.

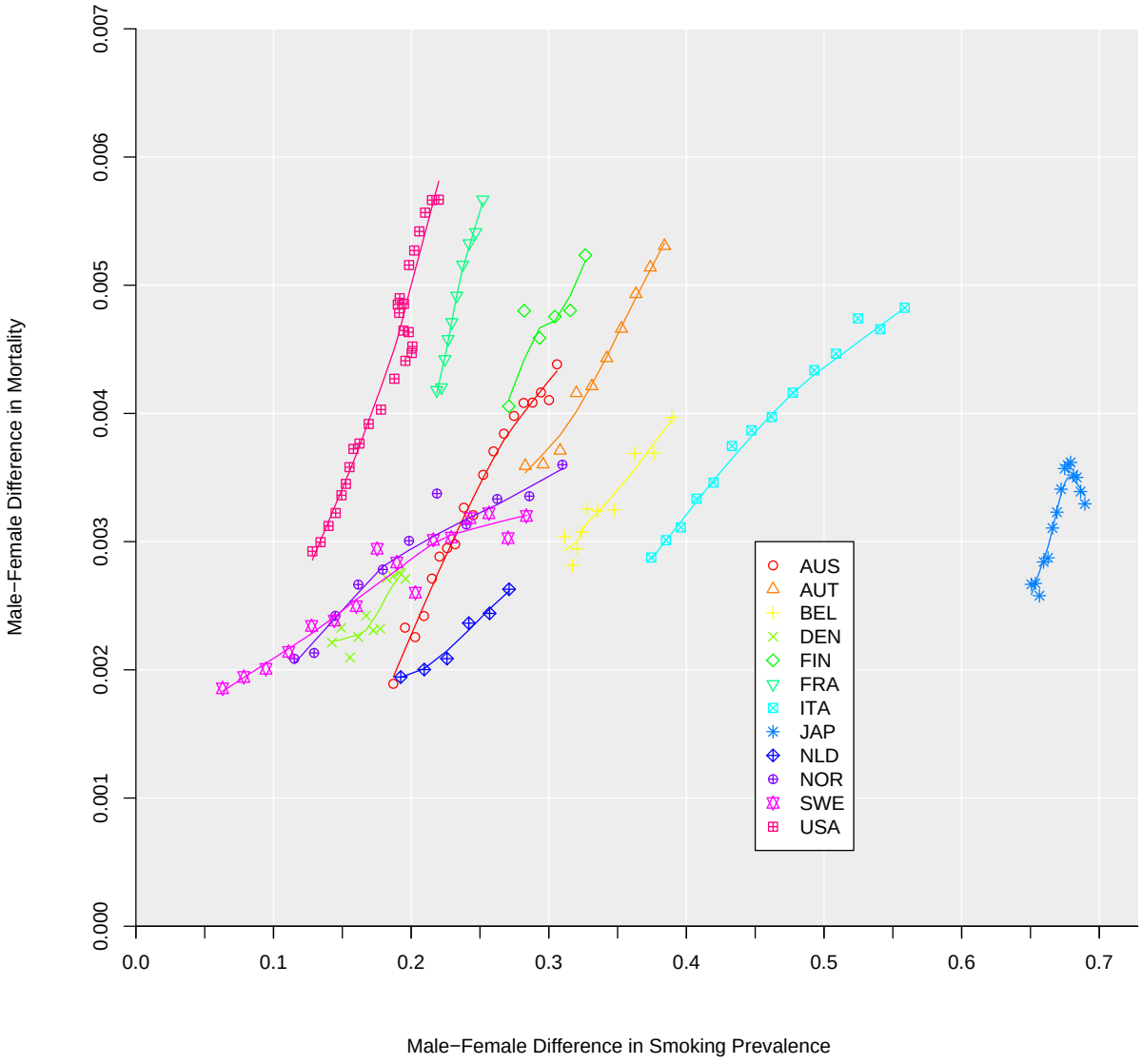
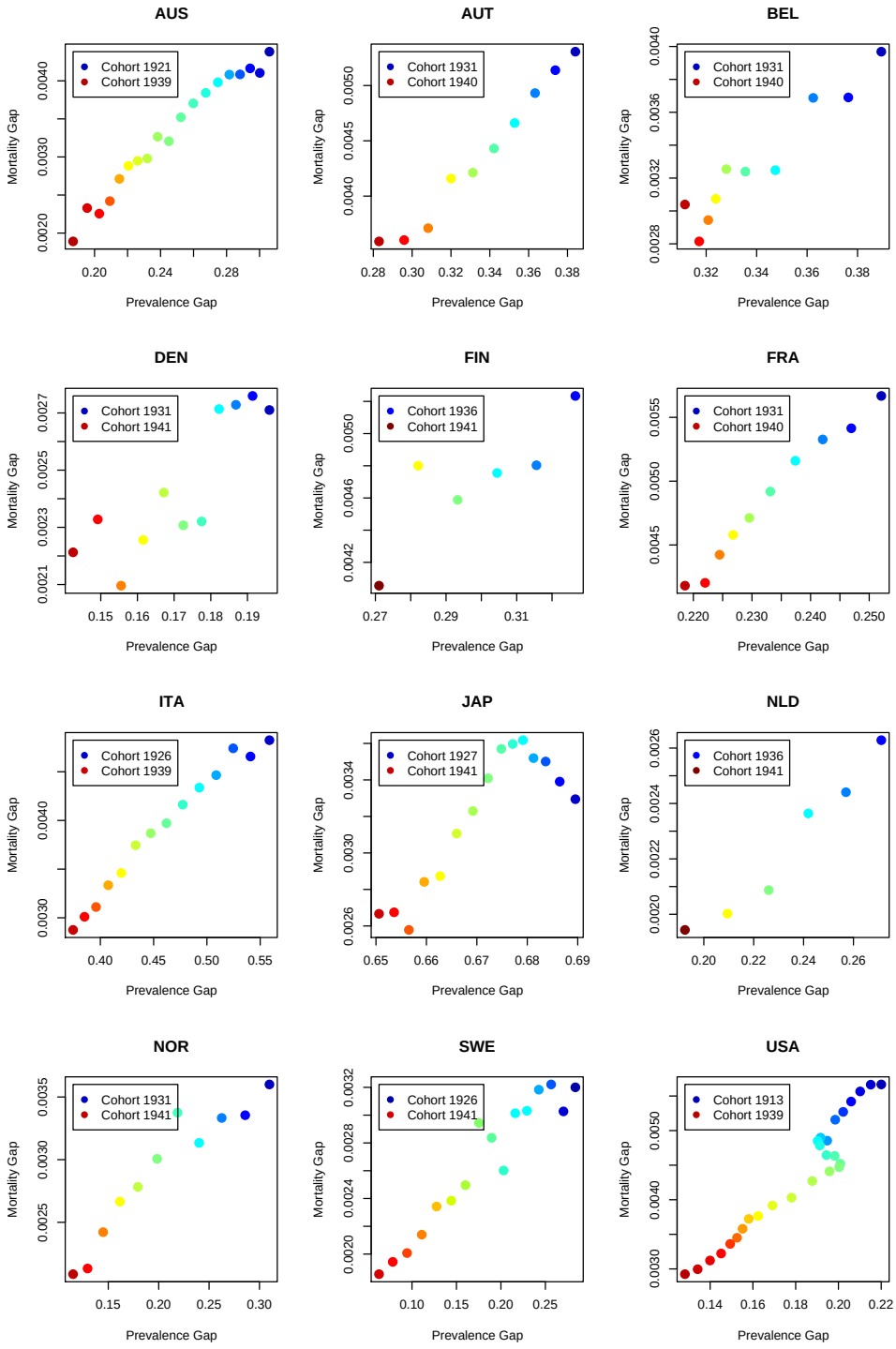


Figure 9: Results for Cohort Hypothesis: Prevalence Gap in Smoking (Ages 15–40, Men minus Women) and the Gap in Mortality (Men minus Women) in the Respective Cohorts at Ages 40–65.



B Additional Tables and Figures

Table 3: Gap in Life Expectancy at Birth between Women and Men in Selected Countries (1950–1980)

Year	Country									
	CAN	DEN	FRA	FRG	GDR	ITA	JAP	NLD	SWE	USA
1950	4.43	2.42	5.76			3.43	3.32	2.27	2.61	5.62
1951	4.44	2.45	5.73			3.51	3.31	2.50	2.67	5.81
1952	5.06	2.76	5.81			3.62	3.41	2.32	2.87	5.95
1953	5.11	2.60	5.95			3.69	3.52	2.57	2.97	6.12
1954	5.15	2.91	6.18			4.10	4.09	2.88	3.01	6.13
1955	5.36	3.08	6.31			4.28	4.14	3.15	3.09	6.22
1956	5.31	3.31	6.51	4.96	4.49	4.38	4.23	3.13	3.44	6.34
1957	5.68	3.34	6.71	5.24	4.68	4.66	4.35	3.20	3.53	6.38
1958	5.78	3.55	6.37	5.12	4.47	4.71	4.55	3.38	3.27	6.40
1959	5.66	3.28	6.50	5.25	4.61	4.74	4.67	3.92	3.63	6.56
1960	5.93	3.55	6.59	5.37	4.83	5.01	4.83	3.86	3.64	6.68
1961	6.06	3.82	6.90	5.48	4.89	5.32	4.92	4.29	3.74	6.67
1962	5.99	4.13	6.90	5.68	4.74	5.35	4.94	4.61	4.05	6.71
1963	6.10	4.07	7.00	5.69	4.83	5.50	5.08	4.76	4.08	6.87
1964	6.45	4.50	7.14	5.94	4.92	5.60	5.12	4.99	4.25	6.99
1965	6.39	4.46	7.25	5.82	4.97	5.61	5.17	5.03	4.35	7.13
1966	6.61	4.58	7.36	5.97	4.93	5.66	5.19	5.08	4.62	7.28
1967	6.73	4.75	7.41	6.07	5.06	5.62	5.21	5.41	4.69	7.40
1968	6.69	4.80	7.46	6.04	5.10	5.74	5.23	5.50	4.67	7.52
1969	6.85	4.85	7.63	6.22	5.33	6.00	5.41	5.37	4.91	7.63
1970	6.94	5.05	7.44	6.31	5.20	5.85	5.35	5.69	4.98	7.63
1971	7.03	5.35	7.56	6.38	5.08	6.00	5.34	5.78	5.42	7.59
1972	7.07	5.40	7.70	6.54	5.30	6.12	5.39	5.98	5.51	7.68
1973	7.17	5.57	7.63	6.53	5.46	6.03	5.26	5.83	5.56	7.70
1974	7.16	5.88	7.81	6.41	5.49	6.20	5.11	6.01	5.70	7.67
1975	7.24	5.72	7.85	6.59	5.50	6.36	5.12	6.26	5.77	7.73
1976	7.29	5.77	8.04	6.57	5.56	6.41	5.16	6.37	5.80	7.59
1977	7.32	5.94	8.14	6.64	5.86	6.58	5.19	6.42	6.16	7.67
1978	7.39	5.91	8.14	6.72	5.94	6.71	5.26	6.54	6.23	7.60
1979	7.28	6.21	8.21	6.67	6.03	6.76	5.33	6.45	6.22	7.65
1980	7.12	6.01	8.24	6.75	5.93	6.75	5.37	6.70	6.07	7.49

Table 4: Gap in Life Expectancy at Birth between Women and Men in Selected Countries (1981–2006)

Year	Country									
	CAN	DEN	FRA	FRG	GDR	ITA	JAP	NLD	SWE	USA
1981	7.10	6.11	8.12	6.66	5.88	6.66	5.35	6.58	6.06	7.39
1982	6.94	6.04	8.17	6.64	6.00	6.59	5.46	6.67	5.96	7.27
1983	6.85	6.11	8.08	6.64	5.95	6.55	5.58	6.63	6.01	7.12
1984	6.76	5.97	8.21	6.64	5.77	6.61	5.66	6.71	6.08	7.05
1985	6.70	5.95	8.21	6.57	5.89	6.52	5.65	6.58	5.91	7.10
1986	6.57	5.76	8.20	6.51	5.88	6.50	5.71	6.51	6.05	7.13
1987	6.61	6.05	8.26	6.59	6.08	6.54	5.78	6.54	5.98	7.05
1988	6.63	5.59	8.16	6.52	6.20	6.51	5.74	6.54	5.81	7.00
1989	6.52	5.78	8.21	6.45	6.25	6.58	5.83	6.25	5.78	6.97
1990	6.40	5.71	8.25	6.38	7.04	6.64	5.93	6.27	5.58	7.01
1991	6.29	5.51	8.29	6.44	7.19	6.75	6.02	6.10	5.60	6.95
1992	6.29	5.40	8.31	6.42	7.26	6.74	6.18	5.98	5.42	6.88
1993	6.16	5.18	8.21	6.31	7.49	6.51	6.19	6.01	5.28	6.78
1994	6.01	5.31	8.24	6.27	7.50	6.54	6.32	5.72	5.30	6.68
1995	5.96	5.11	8.06	6.29	7.37	6.59	6.38	5.76	5.27	6.51
1996	5.72	5.21	7.98	6.15	7.18	6.38	6.50	5.68	5.00	6.17
1997	5.54	4.91	7.78	6.07	7.18	6.26	6.53	5.37	5.10	5.85
1998	5.45	4.94	7.67	5.95	6.77	6.18	6.76	5.50	5.05	5.63
1999	5.43	4.69	7.55	5.84	6.90	6.12	6.80	5.11	4.82	5.44
2000	5.17	4.68	7.53	5.77	6.90	5.94	6.88	5.04	4.64	5.32
2001	5.08	4.54	7.45	5.64	6.63	5.97	6.86	4.91	4.51	5.26
2002	4.85	4.54	7.26	5.48	6.49	5.88	6.89	4.69	4.38	5.25
2003	4.90	4.76	7.06	5.44	6.63	5.55	6.95	4.69	4.52	5.18
2004	4.75	4.66	7.10	5.31	6.50	5.69	6.92	4.57	4.31	5.11
2005		4.51	7.01				6.95	4.40	4.33	
2006		4.61					6.80	4.26	4.21	

Source: Human Mortality Database