



Demographic Research a free, expedited, online journal of peer-reviewed research and commentary in the population sciences published by the Max Planck Institute for Demographic Research Konrad-Zuse Str. 1, D-18057 Rostock · GERMANY www.demographic-research.org

DEMOGRAPHIC RESEARCH

**VOLUME 25, ARTICLE 1, PAGES 1-38
PUBLISHED 01 JULY 2011**

<http://www.demographic-research.org/Volumes/Vol25/1/>

DOI: 10.4054/DemRes.2011.25.1

Research Article

The future of death in America

Gary King

Samir Soneji

© 2011 Gary King & Samir Soneji.

This open-access work is published under the terms of the Creative Commons Attribution NonCommercial License 2.0 Germany, which permits use, reproduction & distribution in any medium for non-commercial purposes, provided the original author(s) and source are given credit. See <http://creativecommons.org/licenses/by-nc/2.0/de/>

Table of Contents

1	Introduction	2
2	Mortality and its patterns	3
3	Risk factors: Smoking and obesity	4
3.1	Relationship to mortality	5
3.2	Lag lengths	7
3.3	Measurement	8
3.4	Historical patterns	9
4	Forecasting methodology	11
4.1	Assessment of existing methods	11
4.2	Our approach	15
5	Mortality forecasts	15
6	Sources of uncertainty	19
7	Concluding remarks	22
8	Acknowledgements	23
	References	24
A	Appendix: Forecasting methodology	30
A.1	Overview of forecasting methodology	30
A.2	Bayesian hierarchical model	32
A.3	Search algorithm for setting priors	34
A.4	Forecasting methodology example	35

The future of death in America

Gary King¹

Samir Soneji²

Abstract

Population mortality forecasts are widely used for allocating public health expenditures, setting research priorities, and evaluating the viability of public and private pensions, and health care financing systems. In part because existing methods forecast less accurately when based on more information, most forecasts are still based on simple linear extrapolations that ignore known biological risk factors and other prior information. We adapt a Bayesian hierarchical forecasting model capable of including more known health and demographic information than has previously been possible. This leads to the first age- and sex-specific forecasts of American mortality that simultaneously incorporate, in a formal statistical model, the effects of the recent rapid increase in obesity, the steady decline in tobacco consumption, and the well known patterns of smooth mortality age profiles and time trends. Formally including new information in forecasts can matter a great deal. For example, we estimate an increase in male life expectancy at birth from 76.2 years in 2010 to 79.9 years in 2030, which is 1.8 years greater than the U.S. Social Security Administration projection and 1.5 years more than U.S. Census projection. For females, we estimate more modest gains in life expectancy at birth over the next twenty years from 80.5 years to 81.9 years, which is virtually identical to the Social Security Administration projection and 2.0 years less than U.S. Census projections. We show that these patterns are also likely to greatly affect the aging American population structure. We offer an easy-to-use approach so that researchers can include other sources of information and potentially improve on our forecasts too.

¹ Albert J. Weatherhead III University Professor, Harvard University, Institute for Quantitative Social Science, 1737 Cambridge Street, Cambridge MA 02138. E-mail: King@Harvard.Edu, <http://GKing.Harvard.Edu>.

² Assistant Professor, The Dartmouth Institute for Health Policy & Clinical Practice and The Norris Cotton Cancer Center, Dartmouth College, One Medical Center Drive, Lebanon, NH 03765. Phone: 603-653-3908, Fax: 603-653-0896, E-mail: samir.soneji@dartmouth.edu, <http://www.dartmouth.edu/~soneji>.

1. Introduction

Since 1950, U.S. life expectancy at birth has grown from 68 to 78 years. The U.S. also experienced considerable population aging resulting mainly from declining fertility and mortality. The elderly (≥ 65 years) grew from 14% of the population in 1950 to 19% in 2000, while the working age population (18 to 64 years) declined from 60% to 56%. These developments in population aging had massive implications for American life, as well as for the allocation of medical expenditures, public health efforts, research priorities, pension programs, Social Security, economic growth, and health care financing (Lee and Tuljapurkar 1997). Future American mortality patterns are of understandably widespread interest, despite their uncertainties. In this paper, we offer the first formal statistical forecasts of age and sex-specific U.S. mortality to include knowledge about common *demographic patterns* in mortality, and large and well-studied biological *risk factors*.

The early demographic work of Graunt (1662), Huygens (Boyer 1947; Vollgraff 1950), Halley (1693) and Gompertz (1825) established two ubiquitous patterns that have continued to hold up across many countries and time periods. First, age-specific mortality usually declines smoothly and gradually over time, with few sharp jumps from one year to the next. Second, time-specific mortality across age groups, known as “age profiles,” have a characteristic shape, with adjacent age groups having similar mortality rates. The vast majority of mortality forecasts to date have incorporated only the time-smoothness property; we make it possible to include as prior information (i.e., rather than either to ignore or require) both smoothness over time and age groups, when the data support it.

Smoking and obesity are two important risk factors, both with well-studied consequences for mortality and large changes over time. Smoking rates steadily declined over the last half century, and obesity rates have rapidly increased in the last thirty years. Most American mortality forecasts ignore these patterns and are based only on simple extrapolations without covariates, and so include these risk factors only indirectly. Other methods study these risk factors directly but do not include the known tendency of demographic patterns to be smooth over time and age. We directly measure and include both risk factor and known demographic patterns in the same forecasting model, which allows forecasts to take note of these factors if — and only if — the data support their inclusion.

We attempt to follow the best forecasting practices now common across many academic fields: When the exact process underlying an outcome is known, the corresponding model is specified and used. When this is not feasible, as with mortality forecasting, we rely on empirical regularities built from the best causal knowledge the health and demographic literature offers. Reliance on empirical regularities will fail in the presence of an unpredictable structural break in the data, e.g., a pandemic, new disease, novel risk factor, or unmeasured change in existing risk factors. Such events are unpredictable and add to

uncertainty (which we also measure and report) but fortunately structural breaks tend to be rare.

We forecast mortality by single year of age and sex for the next twenty-five years in the U.S. We also offer guidance about the use of our approach that will enable other scholars to produce even better forecasts than ours when new covariates or other information become available. Methods such as ours tend to be the most accurate available when the underlying data generating process is uncertain because empirical regularities. These methods should not be confused with models optimized for making causal inferences, which is a different task and requires other approaches (e.g., Ho et al. 2007; Robins 2008). Indeed, none of our results should be seen as claims about the causal effects of obesity, smoking, or any other factor.³

First, we summarize known demographic and risk factor information useful for forecasting in Sections 2 and 3. Second, we assess existing forecasting methods in Section 4.1 and our approach in Section 4.2. Third, we present our mortality forecasts in Section 5. Fourth, we discuss sources of uncertainty and possible difficulties with our approach in Section 6. Finally, we present mathematical details of our forecasting method in the appendices.

2. Mortality and its patterns

We first discuss an important issue in measuring mortality and then illustrate the venerable pattern of how mortality smoothly changes over time and age. We use patterns such as these as prior information to improve the forecasts when supported by the data.

Measurement The most commonly used indicator of mortality in demography is the *mortality rate* (usually written $n_x m_x$ for age group $[x, x + n_x)$). The mortality rate is the number of deaths in a time period and age group divided by the person-years of exposure to the risk of death. In practice, the exposure time is not known and so must be estimated based on assumptions about when deaths occur during each time interval. This is not a major issue for most demographic analyses, but it is a serious issue for forecasting. Exposure in the in-sample data (used to construct the forecast) and exposure in the the out-of-sample data (used to validate the forecast) both must be estimated. As such, validating a forecast based on $n_x m_x$ is usually impossible, since assumptions about future exposure can always be adjusted to apparently improve the forecasts.

³ Different research goals typically require entirely different specifications: For example, in estimating the causal effect of a drug given at birth on longevity, controlling for a measure of health at age 5 would be inappropriate because it would induce post-treatment bias (King and Zeng 2006); however, such a measure would greatly improve a longevity forecast and so should be included if that were the research goal.

Our solution to this problem is to use the *conditional probability of death* as our measure. This quantity is usually written as ${}_xq_x$ for age group $[x, x + n_x)$. The conditional probability of death is the number of deaths in a time period and age group divided by the population alive at the start of that time period, and so is easy to interpret. Since both components are directly observable, ${}_xq_x$ is also directly observable. As such, forecasts that withhold known mortality so it can be used for validation are vulnerable to being proven incorrect, which is required for scientific progress.

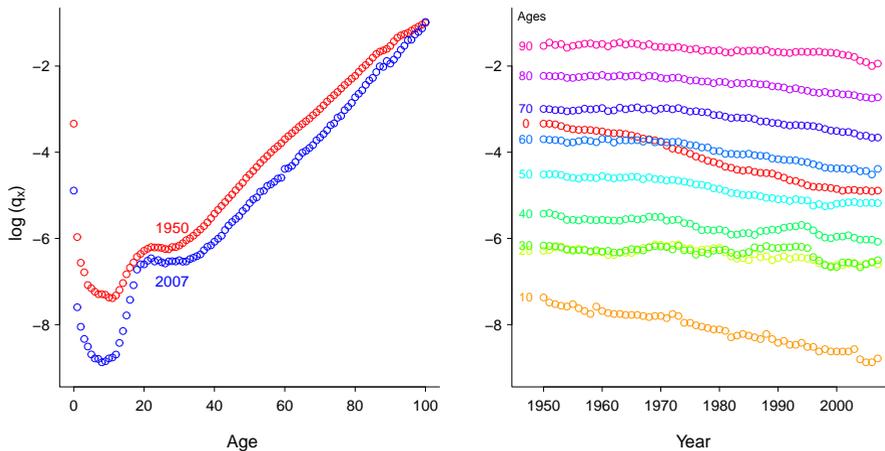
Empirical regularities Between 1950 and 2007, U.S. age-specific mortality was locally smooth and maintained an age profile similar to the pattern shown in the left panel of Figure 1. We measure ${}_xq_x$ by each single year of age from 0 to 100 (with 100 as the start of an open age interval and ${}_∞q_{100} = 1$) between 1980 and 2007. We follow the common demographic convention of dropping the left subscript when the age group is a single year of age.⁴ Smoothness within a given year can be seen by the small incremental changes in log-mortality across age groups. The consistent pattern starts with high infant mortality followed by a rapid decline until approximately age 10. Mortality then increases to a local maximum, the so-called accident hump especially prominent in males. Adult log-mortality then increases nearly linearly until an apparent deceleration starting at about age 90.

Time trends of mortality are also locally smooth and decline over time (Figure 1, right panel) consistent with patterns in most other time periods in advanced industrialized countries (Gompertz 1825; Keyfitz 1982; McNown 1992). Smoothness for a given age can be seen by small incremental changes in log-mortality over time. The pace of the decline differs slightly among ages and over some short time intervals. For example, age 0 mortality decreased fastest during the 1990s and slowed in the 2000s. The pace of mortality decline over time was quite similar among adults aged 60-80 years. A similar pace of decline among the oldest-old ages (≥ 85 years) observed in the U.S. has been well documented with high-quality demographic data in over two dozen other countries (Kannisto et al. 1994).

3. Risk factors: Smoking and obesity

In this section, we discuss evidence of tobacco consumption and obesity's effect on mortality. Covariates need not have a causal relationship to mortality to ensure predictive success, but the causal literature is certainly helpful in making choices.

⁴ To determine death counts, we use the Mortality Detail Files 1980–2007, containing information on all deaths registered on individual U.S. death certificates transmitted to the National Center for Health Statistics. We use population counts calculated by the Human Mortality Database (Human Mortality Database 2008).

Figure 1: Observed U.S. male mortality, age and time domains

Notes: The left panel shows the age profile of male log-mortality (measured by the conditional probability of death) in 1950 and 2007. The right panel shows male log-mortality for select ages observed between 1950 and 2007. Ages are listed along the left side (e.g., "60" represents age [60, 61] years).

3.1 Relationship to mortality

The strong link between cigarette smoking and increased risk of mortality has been well established since 1950 (Doll 1999; Parascandola 2004). Cigarette smoking increases the risk of cardiovascular disease, lung cancer, chronic obstructive pulmonary disease mortality, and nearly fifty other causes of death. After adjusting for several known mortality risk factors, Jacobs et al. (1999) found smokers' hazard of death 1.3 times greater than non-smokers and the attributable risk of smoking to be 23%. In the fifty-year followup of their seminal study, Doll et al. (2004) estimated male cigarette smokers died ten years younger than nonsmokers, on average.

Obesity, characterized by excess adipose tissue, is highly heritable (Wardle et al. 2008) and likely stems from a complex interaction of multiple genes, environmental factors, and behavior (Yang, Kelly, and He 2007). Emerging biological research is providing more information on caloric excess in humans and other species (Baur et al. 2006). In humans, excessive weight gain may disrupt metabolism through closely linked metabolic and inflammatory signaling pathways (e.g., the insulin/insulin-like growth factor-1 pathway). Excess glucose and lipids trigger metabolic dysfunction, which, in turn, triggers inflam-

matory responses. The result is a cycle of further metabolic dysfunction and inflammatory responses that contribute to cellular stress. For example, excessive oxidation of glucose or free fatty acids by mitochondria creates oxidative stress, which can damage cellular structures.

Within the scientific community, there is uncertainty on obesity's effect on mortality at the population level (Crimmins, Preston, and Cohen 2011). There is general consensus of excess age-specific mortality among the class I, II, and III obese (Fontaine et al. 2003; Olshansky et al. 2005; Stewart, Cutler, and Rosen 2009).⁵ Yet, how much the overweight experience leads to excess age-specific mortality remains an open empirical question. In their careful prospective analysis of 900,000 adults, Prospective Studies Collaboration (2009) considered the relationship between obesity and mortality among current cigarette smokers and those who never smoked regularly. Among smokers, compared to the normal weight group, the risk of mortality was higher for the underweight, approximately equal for the overweight, higher again for class I obese, and higher still for class II obese. Among non-smokers, compared to the normal weight group, the risk of mortality was no higher for the underweight or overweight groups, was higher for the class I obese, and higher still for the class II obese.

Equally uncertain is the extent to which obesity affects life expectancy. Peeters et al. (2003) found an estimated reduction in life expectancy at age forty of six years for obese males and seven years for obese females, compared to their non-obese counterparts. Fontaine et al. (2003) found an estimated decline in male life expectancy at age 50 ranging from one to seven years for those ranging from BMI 30 to 45+, respectively. A similar decline was noted in female life expectancy. Recently, Preston and Stokes (2010) found an estimated gain in life expectancy at age fifty from hypothetically redistributing the obese into optimal BMI categories between 0.6 and 1.9 years for males and 0.8 to 1.6 years for females.

Flegal et al. (2010) suggest the mortality risk associated with obesity may be changing over time. More effective medical management of obesity related complications using existing therapeutic agents may mitigate the impact on mortality and longevity for some. Emerging biological knowledge of the mechanisms and signaling pathways by which genetic and environmental factors interact to favor weight gain and disrupt metabolism is also opening the possibility of new therapeutic intervention (Baur et al. 2006). Although compliance with medical recommendations in a population that also experiences difficulty complying with dietary advice may not be very high.

A population's health is a complex multi-dimensional concept that changes over age, time, and across countries. We consider cigarette smoking and obesity, though there cer-

⁵ The World Health Organization classification of BMI: underweight ($BMI \leq 18.5$), normal weight ($18.5 \leq BMI \leq 24.9$), overweight ($25.0 \leq BMI \leq 29.9$), class I obesity ($30.0 \leq BMI \leq 34.9$), class II obesity ($35.0 \leq BMI \leq 39.9$), and class III obesity ($BMI \geq 40$).

tainly are other important risk factors. The interaction of risk factors is quite likely, as well. Freedman et al. (2006) found the mortality risk among obese smokers, even young obese smokers, far exceeded the sum of their individual risks. Considerable variation in mortality risk may also exist among individuals, even within a specific sub-population. For example, Wessel et al. (2004) note substantial differences in cardiorespiratory fitness among females at all levels of obesity, but the precise and complex interactions among physical fitness, obesity, and mortality remain largely unknown (Allison et al. 2003). Including all necessary interactions and sub-population effects to build a completely specified causal model would be an important task, but is not the subject of this paper. Our alternative strategy of forecasting based on causally informed empirical regularities is likely to be somewhat safer, at least until considerably more data become available.

3.2 Lag lengths

Contemporaneous relationships and lagged relationships with mortality are two approaches to using covariates in forecasting. Using the contemporaneous relationship between covariates and mortality is generally not a good idea for two reasons. First, we would not expect the risk factors to instantly affect mortality. Second, using the risk factors at time t to predict mortality at time t requires one to separately forecast covariates prior to forecasting mortality, since knowledge of the covariates would be required at time $t+k$ for a k step ahead forecast. This extra forecasting step is not only inconvenient and model dependent, but it also propagates considerable additional uncertainty into the ultimate mortality forecasts.

In contrast, using lagged covariates in mortality forecasting models is typically a more objective, less model-dependent, and less uncertain approach. We thus lag the covariates in a cohort-based analysis. That is, we lag the risk factors k years in time while keeping the cohort, which is k years younger, the same. We then use a one-step ahead forecasting procedure with the final observed year of covariate values to forecast k years into the future. This approach bases forecasts on risks already experienced by the population rather than on risks predicted to occur in the future.

Literature on the life course is helpful in determining the lag length because it focuses on the timing of the association between an exposure and the eventual mortality outcome (Lynch and Smith 2005). For example, Rogers et al. (2005) and Preston and Wang (2006) both found a strong effect between a cohort's earlier smoking history and its later mortality. Peace (1985) found a strong time-lagged correlation of 20 to 25 years between cigarette sales and lung cancer mortality at the population level. A time lag also exists among obesity, development of chronic conditions, and mortality. Unlike the historical decline in smoking, the rise in obesity is much more recent and we are likely to observe its consequences in the future (Sturm 2002; Guterman 2008). Baker, Olsen, and Sørensen

(2007) recently found a strong association between childhood obesity and coronary heart disease mortality between ages 25 and 60.

Thus, we keep the lag length k to approximately what we learn from the life course perspective, which is about $k = 25$ years, for both smoking and obesity, although the lag value need only be approximate given the high autocorrelation in mortality. We demonstrate this point by also considering shorter lags, $k = 5, 10, 15,$ and 20 years in Section 6. When lagging far enough back so that the covariate would not be meaningful, such as smoking rates for infants, we drop the covariate. A crucial advantage of our statistical methodology is that it enables us to use different covariates for forecasting different age groups while still borrowing strength in estimating all the cross-sections together, and smoothing over age and time appropriately.

As Soneji and King (2011) note, period effects may also strongly affect mortality, in addition to cohort effects. For example, the 1964 U.S. Surgeon General's Report on Smoking and Health (Office of the Surgeon General 1964) marked a period of intense anti-tobacco public health campaigns. The collection, now totaling over thirty reports, focused on changing behavior, attitudes, and knowledge among potential and current smokers. Unlike a cohort-specific model, which incorporates time-lagged correlation between potential covariates and mortality, a period-specific model would require the prediction of covariates for ages in the future. Also noteworthy is the current debate within the demographic research community on whether period life expectancy is biased because of changes in the timing or age of mortality (Bongaarts and Feeney 2003; Guillot 2003; Wilmoth 2005). Such changes in the timing of death may result, for example, from the considerable decline in U.S. smoking.

3.3 Measurement

We estimate historical smoking prevalence from 1955 to 2007 from national-level cross-sectional surveys, including the 1955 Current Population Survey and seventeen National Health Interview Surveys (NHIS) beginning in 1966. From 1955 to 1991, a respondent was considered a current smoker if he or she responded affirmatively to the questions "Have you smoked at least 100 cigarettes in your entire life?" and "Do you now smoke?". In 1992, the second question changed to "Do you now smoke every day, some days, or not at all?". Since 1992, a current smoker was defined as someone who has smoked at least 100 cigarettes in their life and currently smokes every day or some days. In all our analyses, survey weights calculated by the National Center for Health Statistics are incorporated so that respondents represent their population share. Non-response through refusal to answer, lack of knowledge, or inability to ascertain tobacco usage was less than 1.4% for all years of the survey.

We estimate historical obesity 1961–2007 from national-level cross-sectional surveys,

including the 1959–1962, 1971–1975, and 1976–1980 waves of the National Health and Nutrition Examination Survey and annual NHIS since 1981. A respondent is considered obese if his or her Body Mass Index (BMI) is 30 kilograms per meter squared or greater. We recognize the possible disadvantage combining information from surveys with different methodologies and a single BMI threshold for obesity (Mehta and Chang 2009). BMI itself may also have limitations as a measure of adiposity because it does not differentiate between fat mass (e.g., visceral fat) and fat-free lean mass (Bergman et al. 2006; Snijder et al. 2006). Several studies argue that a greater percentage of fat mass increases mortality risk, especially among the elderly. A greater percentage of fat-free lean mass, however, may decrease mortality risk (Heitmann et al. 2000). Yet in their systematic review of 54 cohort studies, Romero-Corral et al. (2006) note the higher BMI, the better the discriminatory power for body fat and lean mass. However, the signal from the pattern of obesity change in the U.S. appears to overwhelm any such measurement noise. Other thresholds for obesity, and other measures such as average BMI, do change the forecasts, but do not have a major impact on the results. The resulting data is a rich history of the best existing height and weight measurements over the last 47 years.

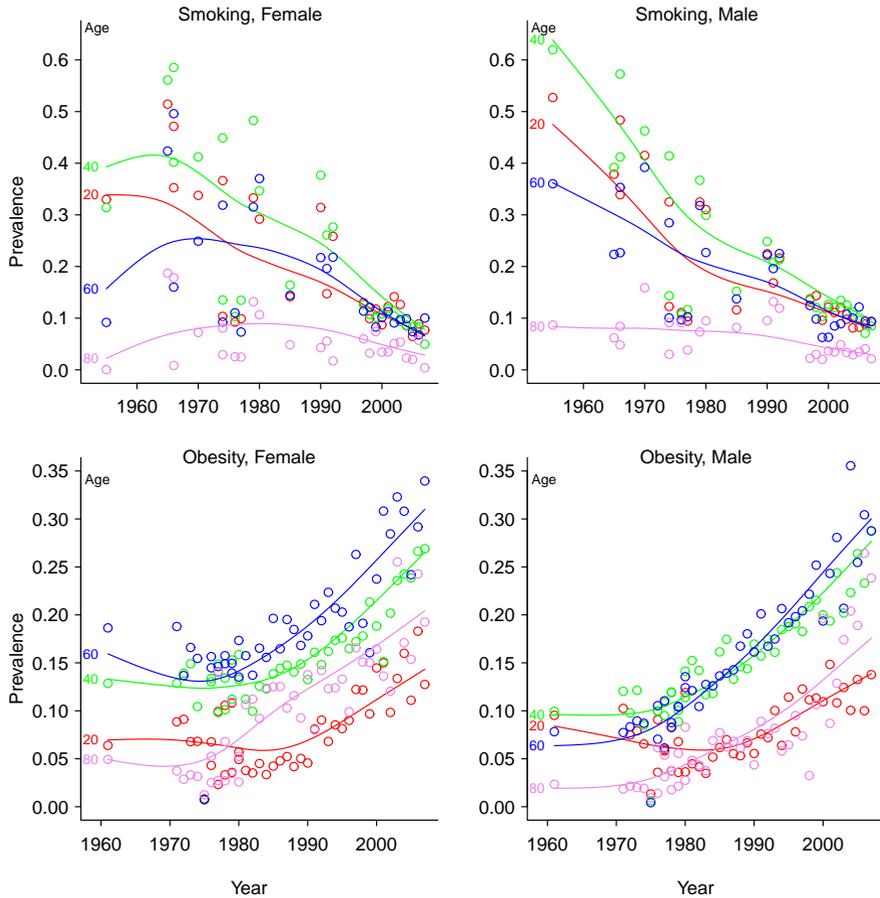
Finally, following Murray and Lopez (1996) and others, we use a linear time trend as a proxy for technological progress. This is a crude measure, but we have not found dramatic differences with other basis functions so long as our priors for demographic patterns are also included. Our preliminary experiments using patents and National Institutes of Health research funding also does not seem to have a large effect, although this is worth much further study.

3.4 Historical patterns

We estimate smoking and obesity prevalence for males and females ages 0 to 100 between 1955 to 2007. As with mortality, we expect the prevalence of these risk factors to change smoothly over time, though the observed values are more noisy because they are based on random samples. We therefore apply the smoothing techniques described in Section 4.2 which also enable us to interpolate prevalence during years without sample data.

Figure 2 presents time series plots of the observed data (circles) and smoothed trajectories (lines) for smoking (top panels) and obesity (bottom panels), for females (left panels) and males (right panels), and for selected age groups (represented in different colors and labeled on the left of each line). The smoking results in the top graphs portray the broad declines across age and sex groups in smoking, with the sharpest declines occurring among younger adult males. Among forty-year old males, for example, smoking decreased from approximately 64% in 1955 to 10% in 2007. Among similarly aged females, smoking prevalence decreased from 39% in 1955 to 8% in 2007.

Figure 2: Smoking and obesity prevalence over time



Notes: Smoking (top graphs) and obesity (bottom graphs) for females (on the left) and males (on the right) for selected ages (in colors). Observed data are shown as open circles and smoothed estimates as lines. Ages are listed along the left side of each panel (e.g., "60" represents age [60, 61] years).

Both sexes experienced a dramatic increase in obesity, as measured by excess BMI, with the sharpest increases occurring among the young (Figure 2, lower panels). For example, the prevalence of obesity increased from 10% in 1961 to 28% in 2007 among forty-year old males. Similarly aged females experienced an increase from 13% in 1961 to 27% in 2007. The fastest increase has occurred since 1990 for all age groups and both sexes.

The relationship between obesity, morbidity, and mortality may change over time with advancements in medicine. Pharmacological innovations may have contributed to reductions in cardiovascular disease (CVD) risk factors over time, an important medical complication of obesity. Gregg et al. (2005) observed secular reductions in hypertension and high total cholesterol level among all BMI groups. A crucial exception is diabetes mellitus; this CVD risk factor nearly tripled among the obese from 1960 to 2000. Our linear trend will proxy for some of this and other technological progress.

4. Forecasting methodology

We discuss here issues with existing methods and then introduce our approach, which attempts to build on the existing methods and address some of the issues they raise.

4.1 Assessment of existing methods

Ideally, a mortality forecast would simultaneously incorporate the effect of the rapid increase in obesity and the steady decline in cigarette consumption, while still maintaining the long-standing patterns of smooth mortality age profiles and time trends. At the same time, these patterns only affect the forecasts if not greatly contradicted by the in-sample data. Forecasting methods to date have not been able to incorporate this combination of properties. First, by construction, purely extrapolative methods cannot include information about the direct effects of risk factors with known health effects. Second, because of various independence assumptions, existing methods that allow for the incorporation of key risk factors often yield forecasts that violate the well known demographic patterns of smooth age and time profiles.

We illustrate these problems by using the two most common existing forecasting methods: the time-series based Lee-Carter approach (Lee and Carter 1992) and least squares (LS) regression. We use 1980–2007 U.S. male mortality data to forecast through the year 2030.

Forecasts The first column of Figure 3 presents forecasts of log-mortality using the approach of Lee and Carter (1992). The results seem reasonable in the time domain, although the age profiles exhibit some local non-smoothness that increase over time. Sec-

ond, we consider a forecast of log-mortality as a linear function of time (Figure 3, second column). We write this model as:

$$E(\log(q_{a,t})) = \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t. \quad (1)$$

This forecast yields nearly identical results to those of Lee-Carter. Third, we consider a forecast of log-mortality as a linear function of time and smoking lagged by 25 years (Figure 3, third column). We write this model as:

$$E(\log(q_{a,t})) = \begin{cases} \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t, & \text{if } a < 50 \\ \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t + \beta_a^{(smoking)} \text{smoking}_{a-25,t-25}, & \text{if } a \geq 50 \end{cases} \quad (2)$$

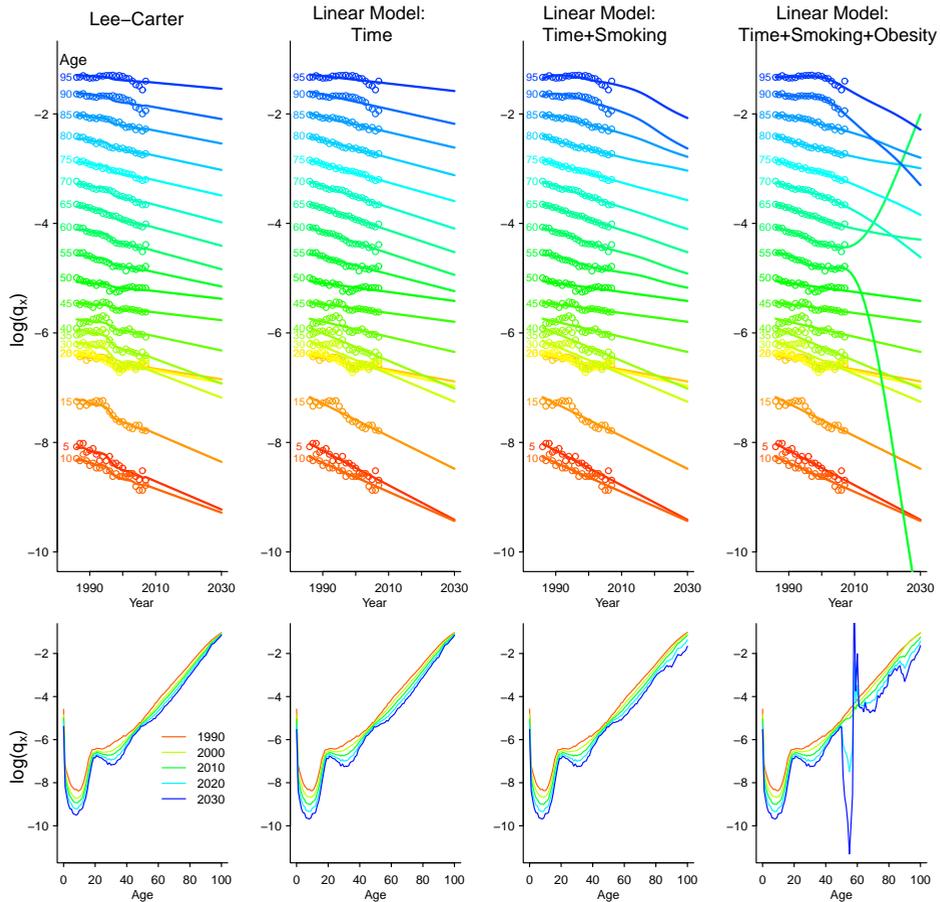
where $\text{smoking}_{a-25,t-25}$ is smoking prevalence 25 years earlier in age and 25 years earlier in time. The results here are plausible for some ages and implausible for others; the age profiles grow increasingly non-smooth. For example, mortality is low for 30-year-olds, enormously higher for 35-year-olds and almost nonexistent for 40-year-olds. This forecast is very far from any historical experience and so cannot be taken seriously. Fourth, we consider a forecast of log-mortality as a linear function of time, smoking lagged by 25 years, and obesity lagged by 25 years (Figure 3, fourth column). We write this model as:

$$E(\log(q_{a,t})) = \begin{cases} \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t, & \text{if } a < 50 \\ \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t + \beta_a^{(smoking)} \text{smoking}_{a-25,t-25} \\ \quad + \beta_a^{(obesity)} \text{obesity}_{a-25,t-25}, & \text{if } a \geq 50 \end{cases} \quad (3)$$

where $\text{smoking}_{a-25,t-25}$ is similarly defined and $\text{obesity}_{a-25,t-25}$ is obesity prevalence 25 years earlier in age and 25 years earlier in time. The dubious time series plot and highly non-smooth age profile plot are implausible, with no historical experience to support such a pattern.

We learn that the forecasts from linear regression models become increasingly unbelievable as more informative covariates are included, thereby violating some of the strongest empirical regularities in demography. Instead of indicating that medical science has been wrong about the effects of smoking and obesity, and demographers are about to be wrong about the smoothness of age and time profiles, these results reveal the inadequacy of forecasting methods based on independent linear regressions or pure extrapolation. The methods would also seem to violate the “more data is better” principle of statistical inference. Newer approaches may make it possible to improve on these results. We introduce one in the next section. (Another possibly useful method is a two-dimensional regression spline, which avoids linearity assumptions but has not yet been adapted to include priors or covariates that differ among different cross-sections; Currie, Durban, and Eilers 2004; Kirkby and Currie 2010).

Figure 3: Male all-cause log-mortality forecast: Lee-Carter, linear time trend, time+smoking, and time+smoking+obesity



Notes: Log-mortality forecasts by the Lee-Carter approach (1st column) and three least squares regressions: time (2nd column), time and smoking lagged 25 years (3rd column), and time, smoking lagged 25 years, and obesity lagged 25 years (4th column). Observed mortality are shown as open circles. Ages are listed along the left side of each panel in the upper row (e.g., “60” represents age [60, 61] years).

Validation We now show that the existing forecasting methods studied above, which produce implausible results based on historical experience, also turn out to be suboptimal when evaluated in rigorous out-of-sample tests. To illustrate these tests, we set aside the last ten years of observed historical mortality (1998-2007) as the validation period and create forecasts based only on the earlier data (1980-1997). We then calculate the average root mean square error (RMSE) of the out-of-sample data for all ages and all years in the validation period, for both male and female mortality. We repeat the same RMSE calculation for each of the four methods illustrated in Figure 3.

The results appear in the first four rows of Table 1. Out-of-sample RMSE is lowest for the models that ignore the most information — the Lee-Carter models and least squares with only time as a covariate, for both males and females. (We discuss our method and its lower RMSE in the next section.)

Table 1: Out-of-sample RMSE for alternative forecasting methods, based on a 10 year holdout period

	Male	Female
Lee-Carter	0.0140	0.0041
Least Squares Regression: Time	0.0135	0.0044
Least Squares Regression: Time+Smoking	0.0184	0.0065
Least Squares Regression: Time+Smoking+Obesity	0.0310	0.1606
Our Mortality Forecast	0.0079	0.0039

Validation is an important aspect of all mortality forecasts. We considered numerous factors in constructing validations. First, we incorporated plausible lags between mortality and the potentially informative covariates. Second, we utilized observed and measured values of potentially informative covariates, rather than estimate earlier values. For example, cigarette smoking prevalence data by age and sex is first available in the US for 1955. Third, we ensured an adequate historical period of mortality to form the basis of a validation forecast. We also ensured an adequate validation period of observed historical mortality to properly assess the accuracy and properties of the forecast. We prioritized forecasting based on observed historical covariate data and maintaining substantively reasonable lag lengths. In doing so, we utilized a shorter historical mortality period than might be otherwise possible. We explored different lengths of the validation period (observed mortality intentionally left out of the forecast) and found the ten-year window afforded sufficient data for the validation forecast and a sufficient length of time to properly assess the accuracy of forecasts. We make available all of our data and code so that others can make different choices and incorporate different tradeoffs in validation (King and Soneji 2011).

4.2 Our approach

We employ a Bayesian hierarchical modeling approach developed by Girosi and King (2008) to incorporate both the key risk factors of smoking and obesity and the long-standing demographic patterns of smooth mortality age profiles and time trends into the same forecasting approach (King and Soneji 2011). In this approach, risk factors and time are included in linear regression models as measured covariates. Demographic information on smoothness of expected mortality across age groups and time periods enters the model as Bayesian priors. The priors are not merely on difficult-to-interpret coefficients, as in classic Bayesian approaches, but instead are stated as beliefs about aspects of expected mortality such as smoothness across age groups and time periods, which is what prior demographic research has taught us. The method also estimates the set of regressions from age-specific forecasts together, rather than making implausible independence assumptions across age groups or time periods, or requiring the same covariates for all cross-sections (such as requiring tobacco consumption among infants).

The Bayesian priors thus incorporate previous empirical patterns and formalize qualitative knowledge demographers have gained over the last 350 years. The Bayesian model uses demographic and risk factor information, but is designed to down-weight or ignore it if contradicted by observed empirical patterns. The priors only have their effect on forecasts in areas where the data are weak, and the risk factors only have an effect on the forecasts if the hypothesized pattern is found in the in-sample data. We summarize the details of the statistical methods we use, along with a worked example, in the Appendix.

Girosi and King (2008, Chapters 11–13) present extensive tests of the model for numerous mortality data sets in many countries. They show that including information in this model in the way we do substantially improves out-of-sample forecasts. Thus, to these validation results, we add the final row of Table 1; this reports the out-of-sample RMSE for our forecasts, which are lower than the other for methods reported, consistent with the results of Girosi and King (2008).

5. Mortality forecasts

Using the methodology in Section 4.2 (and the Appendix), we forecast mortality by sex and single year of age 0–100 between 2008 and 2032. As discussed in Section 3.2, for ages less than 50 years, the only covariate in the model is a time trend. For ages 50 and greater, covariates include a time trend, smoking prevalence lagged 25 years in time and age, and obesity prevalence also lagged 25 years in time and age. In other words, we consider the smoking and obesity prevalence of the birth cohort 25 years earlier. We may write our model as:

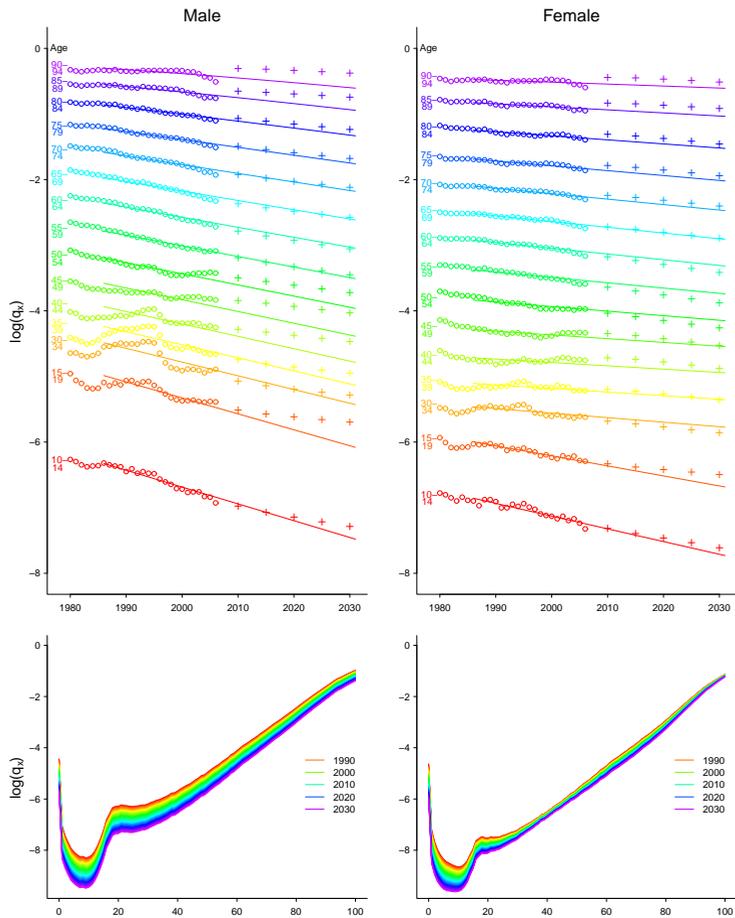
$$E(\log(q_{a,t})) = \begin{cases} \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t, & \text{if } a < 50 \\ \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t + \beta_a^{(smoking)} \text{smoking}_{a-25,t-25} \\ \quad + \beta_a^{(obesity)} \text{obesity}_{a-25,t-25}, & \text{if } a \geq 50 \end{cases} \quad (4)$$

where $\text{smoking}_{a-25,t-25}$ is smoking prevalence and $\text{obesity}_{a-25,t-25}$ is obesity prevalence 25 years earlier in age and 25 years earlier in time. The regression coefficients, β , are drawn from a Bayesian posterior distribution, as described in Section A2. The covariates are not projected or forecasted into the future. Rather we base forecasts on past risks already experienced by the population. We discuss forecasts of mortality, life expectancy, and the population age structure. We also compare our forecasts with official U.S. projections.

Future mortality In Figure 4, we present our mortality forecasts in the time and age domains and highlight key properties and differences between the sexes and among age groups. Unlike the demographically unreasonable least squares forecasts with the method given in Figure 3, our forecasts maintain common historical demographic patterns: they are smooth over time, as seen in the solid lines in the upper panels of the figure, and smooth across age groups, as shown in the bottom panels (color-coded by year).

The pace of mortality decline may be faster than officially projected, especially for males ≥ 50 years (Figure 4, top left panel). For example, we forecast a reduction of 6,320 deaths in the 80-84 year age group between 2010 and 2030 for each 100,000 males who reach age 80. In comparison, the intermediate Social Security Administration (SSA) forecast for the same age group predicts a reduction of 5,280 deaths between 2010 and 2030 per 100,000 males who reach age 80. The pace of mortality decline may also be faster for males than females, especially above age 50 years. For example, we forecast a decline in the conditional probability of death among male 80-84 year olds from 0.33 in 2010 to 0.27 in 2030. In comparison, we forecast a decline among females 80-84 years old from 0.26 in 2010 to 0.24 in 2030.

Figure 4: Male and female log-mortality over time and age groups for our model including time, smoking, obesity, and smoothness priors

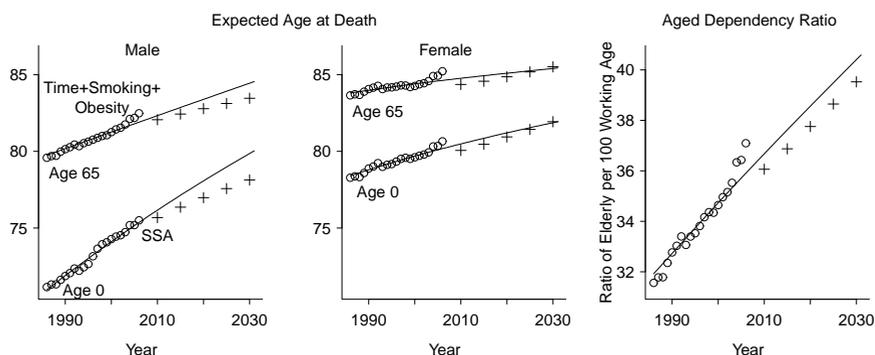


Notes: The top panel shows observed male (left) and female (right) log-conditional probability of death (in circles) along with model forecasts (solid lines) for selected ages. The bottom panels give the age profile of model forecasts between 1980 and 2030, again for males (left) and females (right), and color-coded for select years. SSA forecasts are represented by + signs. Age groups are listed along the left side of each upper panels (e.g., "60-64" represents age [60, 65) years).

Life expectancy Period life expectancy is the expected remaining life of an individual of a given age who experiences the age-specific mortality of a given year (Preston, Heuveline, and Guillot 2001). In the left and center panels of Figure 5 for males (left) and females (center), we present this statistic for ages 0 and 65, calculated from our forecast conditional probabilities of death⁶. For males, our estimates (given by solid lines) are substantially higher than SSA's projections (given by +). For males, we estimate life expectancy at birth to increase steadily from 75.6 years in 2007 to 79.9 years in 2030. In comparison, the SSA forecast in 2030 is 78.1 years. For females, we estimate virtually the same life expectancy at birth in 2030 as the SSA of 81.9 years.

We also calculate the expected age at death among those alive at age 65 (equal to the sum of 65 years and life expectancy at age 65 in a period life table) as shown in the upper lines of the left and center panels of Figure 5. For males age 65 in 2030, we estimate the expected age at death to be 84.4 years, compared to 83.4 years by the SSA. For females at the same age and year, we estimate virtually the same expected age at death as the SSA of 85.4 years.

Figure 5: Expected age at death and aged dependency ratio over time



Notes: The left and middle panels give expected age at death for males (left) and females (middle) at age 0 and 65 under the time+smoking+obesity model, as well as Social Security Administration projections (+). (In a period life table, the expected age of death equals the sum of life expectancy and age.) The right panel gives the ratio of elderly (≥ 65 years) to the working age population (between 20 and 64 years).

⁶ To close the life table, we follow the approach of Horiuchi and Coale (1982) and Preston, Heuveline, and Guillot (2001) and assume the mortality hazard above age 100 is Gompertz (log-linear) and that the population above age 100 follows a stable distribution.

Ageing population structure We calculate the ratio of the number of elderly ≥ 65 years to the number of adults between ages 20 to 64 years, which is known as the *aged dependency ratio*. For programs relying on intergenerational transfers of wealth like Social Security, larger ratios imply greater strain on the working age population to support the elderly dependent population. The right panel of Figure 5 gives the aged dependency ratio from 1980 to 2030. In 1980, there were 30.5 elderly per 100 people of working age. We estimate that the ratio rises steadily over time and faster than officially projected. By 2030, we forecast 40.6 elderly will be alive per 100 people of working age. In contrast, the SSA projects the ratio lower, 39.5 per 100 people of working age. The difference of 1.1 additional elderly per 100 people of working age is considerable when compared to historical data in the U.S. and elsewhere.

6. Sources of uncertainty

Physicians, policy makers, and public health officials make many decisions based on mortality forecasts, regardless of the uncertainties. Nevertheless, any user of these or other forecasts should be aware of how forecasts can go wrong. We may have reduced the uncertainties considerably by including more demographic and risk factor information in the forecasts, but the following unknowns remain.

First, surveys used to estimate obesity and smoking prevalence have sampling error, partially mitigated by smoothing over time and large sample sizes (11,000–116,000); they also have self-reporting error. Second, death and population counts may also be measured with error, especially for individuals without birth certificates such as elderly southern blacks. Third, growing social stigmas of smoking and obesity may lead to under-reporting of smoking behavior and weight (Ezzati et al. 2008). However, error in self-reports would not likely cause many to under-report their weight so much as to become recategorized as overweight instead of obese, and in any event will only bias our analyses to the extent that the stigma itself varies substantially over time, regardless of the degree to which the prevalence of smoking and obesity change.

A final source of uncertainty is model dependence, such as that due to choices in making covariate, lag, and prior specifications (King and Zeng 2006). The model-based uncertainty in mortality forecasting almost always far exceeds sampling uncertainty — which is why sampling error-based confidence intervals are excluded from most mortality forecasting reports, including ours. However, the near irrelevance of these traditional confidence intervals does not imply that forecasters should only study point estimates and ignore uncertainty altogether. We now offer a way to understand and formalize these sources of uncertainty in our forecasts.

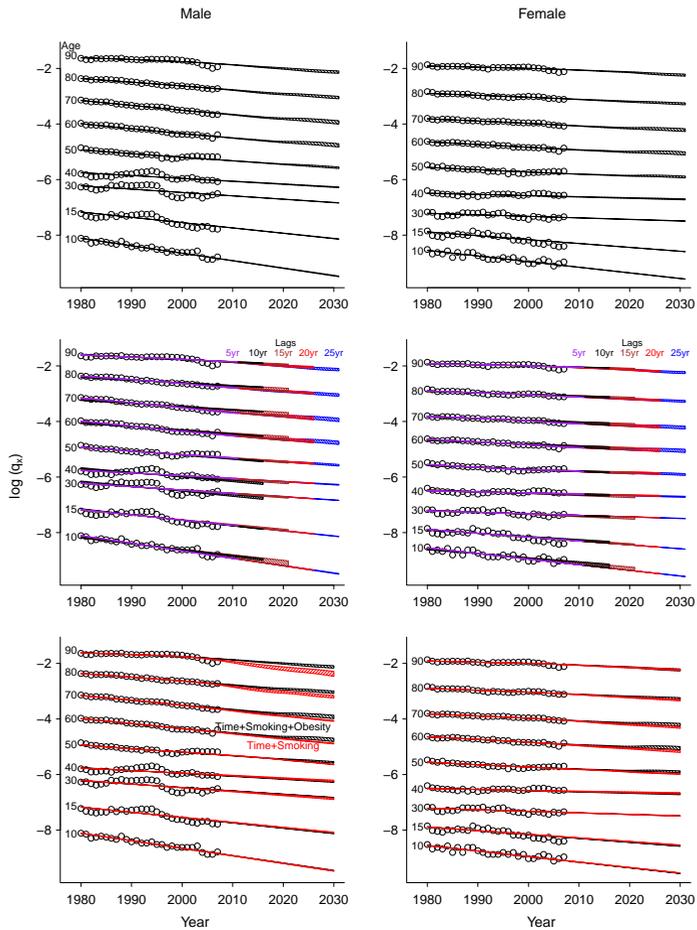
We first study uncertainty due the choice of Bayesian priors used to represent the knowledge that expected mortality is smooth over time and over age groups. We study this prior uncertainty with a version of “robust Bayesian analysis” by using a class of priors instead of only one, and with the result producing a range of forecasts instead of a single point estimate (e.g., Berger 1994; King and Zeng 2001). Thus, for each covariate specification, we use all prior specifications that pass through our search algorithm for setting priors described in Appendix A3. The result of this process is a large number of forecasts that we present as an uncertainty interval in the top panels of Figure 6. Prior uncertainty does affect the forecast, but the pattern in the forecasts remains unambiguous.

In the middle panels of Figure 6, we show the male and female forecasts time, smoking, and obesity, and with varying lag length specifications. Lag lengths are color coded and labeled in five-year intervals from 5 years at the left to 25 years at the right. Prior-based model dependence is also represented in this figure by plotting the whole forecast interval for each given lag length. The results indicate a lack of strong dependence on the lag length. Recall that, with our cohort approach, including a k -year lag of the covariates, produces a forecast k years into the future. This is why, for example, the purple-colored 5-year forecasts extends until 2012, whereas the black-colored 10-year forecasts go until 2017. Thus, to compare these two forecasts directly, we must project the 5-year forecasts an extra five years or examine where the 10-year lag forecast projects when it is only 5 years out. In most cases, each k -year forecast is consistent with the $(k + 5)$ -year forecast; slight exceptions are approximately within the prior uncertainty bounds, represented by multiple lines for each lag length.

The bottom panels in the figure portray model dependence due to the choice of covariates, with the prior uncertainty interval included as before. The red-colored forecast intervals include only time and lagged smoking, whereas the black-colored intervals include time, lagged smoking, and lagged obesity. Although this is not a causal model, it does make some sense that most of the forecast intervals that include obesity lead to higher mortality forecasts after age 50.

All these sources of uncertainty are real and represent exactly where future research is worth pursuing. But even with these uncertainty intervals and differences, the forecasts remain quite informative and most of the patterns remain intact. Although we have gone considerably further than is typical in representing these types of model-based uncertainty, this figure still omits a crucial issue that must always be kept in mind — that the future may have little to do with the past, as cures and therapies are discovered, pandemics occur, and new risk factors emerge.

Figure 6: Model-based uncertainty: Prior specification, lag length specification, covariate specification



Notes: For males (in the left column) and females (in the right column), the first row shows forecast uncertainty intervals due to changes in the prior specification, the second row due to different lag lengths, and the third row due to covariate specification (time+smoking or time+smoking+obesity, each with plausible prior specifications lagged 25 years). Observed log-conditional probabilities of mortality are shown as open circles. Ages are listed along the left side of each panel (e.g., “60” represents age [60, 61] years).

7. Concluding remarks

By including more health and demographic information than has been previously possible, we draw several important conclusions about future mortality patterns. First, by incorporating information on the steady decline in cigarette smoking prevalence and rapid increase in obesity, we forecast mortality may decline faster than officially projected, especially for males ≥ 50 years. Second, the pace of mortality decline may be faster for males than for females. Third, the impact on the demography of future populations is profound. We find faster gains in life expectancy and faster aging of the U.S. population than officially projected.

The age profiles and time trends of our mortality forecasts are smooth and maintain the same venerable demographic patterns of historical mortality in the U.S. and most other countries and time periods. By incorporating risk factors with known health effects (smoking and obesity) we are able to produce more informed mortality forecasts in a statistically sound manner.

Our findings, along with those of Olshansky (1988); Lee and Carter (1992); Lee and Miller (2001); Oeppen and Vaupel (2002); Olshansky et al. (2005); Li and Lee (2005); Wang and Preston (2009); Stewart, Cutler, and Rosen (2009); Olshansky et al. (2009) emphasize the importance of continual development and assessment of forecasting methods and the potential utility of including risk factors in forecasts. The careful inclusion of risk factors, while maintaining ubiquitous demographic properties, will allow demographers to improve the quality, accuracy, and transparency of mortality forecasts.

We must keep in mind that these forecasts are at the population level, based on time-lagged correlations, and do not necessarily imply causality at the individual level. Additional relevant covariates may further inform future forecasts, such as trends in marriage, education, and immigration. Although traditional omitted variable or confounding bias are not relevant here, as they are in estimating causal effects, one may be able to improve our forecasts with such additional information.

Future US mortality patterns may be, at least in part, affected by future mortality patterns elsewhere. As Torri and Vaupel (2011) note, this cross-national association may occur through several pathways including the transfer of scientific knowledge and innovations (e.g., pharmaceuticals, standards of care), introduction of pathogens through migration and travel, and macro-level political, social, and economic forces. Our methodology and framework may prove useful for future global forecasts of mortality that take into account relevant contextual factors and incorporates demographic commonalities and differences.

The demography of future populations in the US and abroad has direct implications upon the financing of entitlement, smoking cessation, and obesity reduction programs. Reduced smoking may translate into a net financial loss for the Social Security Adminis-

tration (Sloan et al. 2004) and Medicare (Wright 1986). Whether obesity-related mortality represents net financial gains or losses to Social Security and Medicare remain open questions. Much depends on the morbidity associated with obesity, early work disability, earnings, and mortality before and after retirement age.

8. Acknowledgements

We thank Robert Aronowitz, David Asch, Jon Bischof, Nicholas Christakis, Majid Ez-zati, Federico Girosi, Joshua Goldstein, Kenneth Hill, Kosuke Imai, Valerie Lewis, Scott Lynch, Doug Massey, Samuel Preston, Jason Schnittker, and four anonymous reviewers for helpful comments and suggestions. An earlier version of this paper was presented at the 2008 Population Association of America annual meeting and the Harvard Center for Population and Development. The authors thank the Robert Wood Johnson Foundation Health & Society Scholars program for its financial support. We also acknowledge financial support from the National Cancer Institute (RC2CA148259), National Institute of Child Health and Human Development (NIH 5T32 HD07163) and Harvard's Institute for Quantitative Social Science for research support. Authors are listed alphabetically and share authorship equally.

References

- Allison, D.B., Wang, C., Redden, D.T., Westfall, A.O., and Fontaine, K.R. (2003). Obesity and years of life lost—reply. *Journal of the American Medical Association* 289(14): 1777–1778. doi:10.1001/jama.289.14.1777-c.
- Baker, J., Olsen, L., and Sørensen, T. (2007). Childhood body-mass index and the risk of coronary heart disease in adulthood. *The New England Journal of Medicine* 357(23): 2329–2337. doi:10.1056/NEJMoa072515.
- Baur, J., Pearson, K., Price, N., Jamieson, H., Lerin, C., Kalra, A., Prabhu, V., Allard, J., Lopez-Lluch, G., Lewis, K., Pistell, P., Poosala, S., Becker, K., Boss, O., Gwinn, D., Wang, M., Ramaswamy, S., Fishbein, K., Spencer, R., Lakatta, E., Le Couteur, D., Shaw, R., Navas, P., Puigserver, P., Ingram, D., de Cabo, R., and Sinclair, D. (2006). Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* 444(16): 337–342. doi:10.1038/nature05354.
- Berger, J. (1994). An overview of robust Bayesian analysis (with discussion). *Test* 3: 5–124. doi:10.1007/BF02562676.
- Bergman, R., Kim, S., Catalano, K., Hsu, I., Chiu, J., Kabir, M., Huckling, K., and Ader, M. (2006). Why visceral fat is bad: Mechanisms of the metabolic syndrome. *Obesity* 14(Supplement 1): 16S–19S. doi:10.1038/oby.2006.277.
- Bongaarts, J. and Feeney, G. (2003). Estimating mean lifetime. *Proceedings of the National Academy of Sciences* 100(23): 13127–13133. doi:10.1073/pnas.2035060100.
- Boyer, C. (1947). Note on an early graph of statistical data (Huygens 1669). *Isis* 37(3/4): 148–149. doi:10.1086/348018.
- Coale, A.J. and Demeny, P. (1966). *Regional model life tables and stable populations*. Princeton, N.J.: Princeton University Press.
- Crimmins, E., Preston, S., and Cohen, B. (eds.) (2011). *Explaining divergent levels of longevity in High-income countries*. Washington D.C.: National Academies Press. (Panel on Understanding Divergent Trends in Longevity in High-Income Countries).
- Currie, I., Durban, M., and Eilers, P. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling* 4(4): 279–298. doi:10.1191/1471082X04st0800a.
- Department of International Economic and Social Affairs (1982). *Model Life Tables for Developing Countries*. United Nations. (Population Studies 77).
- Doll, R. (1999). Tobacco: A medical history. *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 76(3): 289–313. doi:10.1007/BF02345669.

- Doll, R., Peto, R., Boreham, J., and Sutherland, I. (2004). Mortality in relation to smoking: 50 years' observations on male British doctors. *British Medical Journal* 328(7455): 1519–1527. doi:10.1136/bmj.38142.554479.AE.
- Ezzati, M., Friedman, A.B., Kulkarni, S., and Murray, C. (2008). The reversal of fortunes: Trends in county mortality and cross-county mortality disparities in the United States. *PLoS Medicine* 5(5): 0557–0568. (e66). doi:10.1371/journal.pmed.0050119.
- Flegal, K., Carroll, M., Ogden, C., and Curtin, L. (2010). Prevalence and trends in obesity among U.S. adults, 1999-2008. *The Journal of the American Medical Association* 303(3): 235–241. doi:10.1001/jama.2009.2014.
- Fontaine, K.R., Redden, D.T., Wang, C., Westfall, A.O., and Allison, D.B. (2003). Years of life lost due to obesity. *Journal of the American Medical Association* 289(2): 187–193. doi:10.1001/jama.289.2.187.
- Freedman, M., Sigurdson, A., Rajaraman, P., Doody, M., Linet, M., and Ron, E. (2006). The mortality risk of smoking and obesity combined. *American Journal of Preventive Medicine* 31(5): 355–362. doi:10.1016/j.amepre.2006.07.022.
- Giroso, F. and King, G. (2008). *Demographic forecasting*. Princeton: Princeton University Press. <http://gking.harvard.edu/files/smooth/>.
- Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions* 27: 513–585. doi:10.1098/rstl.1825.0026.
- Graunt, J. (1662). *Natural and political observations mentioned in a following index, and made upon the bills of mortality*. London: John Martyn and James Allestry.
- Gregg, E., Cheng, Y., Cadwell, B., Imperatore, G., Williams, D., Flegal, K., Narayan, V., and Williamson, D. (2005). Secular trends in cardiovascular disease risk factors according to body mass index in U.S. adults. *Journal of the American Medical Association* 293(15): 1868–1874. doi:10.1001/jama.293.15.1868.
- Guillot, M. (2003). The cross-sectional average length of life (cal): A cross-sectional mortality measure that reflects the experience of cohorts. *Population Studies* 57(1): 41–54. doi:10.1080/0032472032000061712.
- Gutterman, S. (2008). Human behavior: An impediment to future mortality improvement, a focus on obesity and related matters. Society of Actuaries. (Living to 100 and Beyond Symposium).
- Halley, E. (1693). An estimate of the degrees of mortality of mankind, drawn from curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain

- the price of annuities upon lives. *Philosophical Transactions of the Royal Society of London* 17: 596–610. doi:10.1098/rstl.1693.0007.
- Heitmann, B.L., Erikson, H., Ellsinger, B., Mikkelsen, K.L., and Larsson, B. (2000). Mortality associated with body fat, fat-free mass and body mass index among 60-year-old Swedish men – a 22-year follow-up. The study of men born in 1913. *International Journal of Obesity* 24(1): 33–37. doi:10.1038/sj.ijo.0801082.
- Ho, D., Imai, K., King, G., and Stuart, E. (2007). Matching as nonparametric pre-processing for reducing model dependence in parametric causal inference. *Political Analysis* 15(3): 199–236. <http://gking.harvard.edu/files/abs/matchp-abs.shtml>. doi:10.1093/pan/mpi013.
- Horiuchi, S. and Coale, A. (1982). A simple equation for estimating the expectation of life at old ages. *Population Studies* 36(2): 317–326. doi:10.2307/2174203.
- Human Mortality Database (2008). University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). <http://www.mortality.org> (April 7, 2008).
- Jacobs, D.R.J., Adachi, H., Mulder, I., Kromhout, D., Menotti, A., Nissinen, A., and Blackburn, H. (1999). Cigarette smoking and mortality risk: Twenty-five-year follow-up of the seven countries study. *Archives of Internal Medicine* 159(7): 733–740. doi:10.1001/archinte.159.7.733.
- Kannisto, V., Lauritsen, J., Thatcher, A.R., and Vaupel, J. (1994). Reductions in mortality at advanced ages: Several decades of evidence from 27 countries. *Population and Development Review* 20(4): 793–810. doi:10.2307/2137662.
- Keyfitz, N. (1982). Choice of function for mortality analysis: Effective forecasting depends on a minimum parameter representation. *Theoretical Population Biology* 21(3): 239–252. doi:10.1016/0040-5809(82)90022-3.
- King, G. and Soneji, S. (2011). Replication data for: The future of death in America. IQSS Dataverse Network. (Version V7). <http://hdl.handle.net/1902.1/16178>.
- King, G. and Zeng, L. (2001). Explaining rare events in international relations. *International Organization* 55(3): 693–715. <http://gking.harvard.edu/files/abs/baby0s-abs.shtml>. doi:10.1162/00208180152507597.
- King, G. and Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis* 14(2): 131–159. <http://gking.harvard.edu/files/abs/counterft-abs.shtml>. doi:10.1093/pan/mpj004.
- Kirkby, J. and Currie, I. (2010). Smooth models of mortality with period shocks. *Statis-*

- tical Modelling* 10(2): 177–196. doi:[10.1177/1471082X0801000204](https://doi.org/10.1177/1471082X0801000204).
- Lee, R. and Tuljapurkar, S. (1997). Death and taxes: Longer life, consumption, and social security. *Demography* 34(1): 67–81. doi:[10.2307/2061660](https://doi.org/10.2307/2061660).
- Lee, R.D. and Carter, L.R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association* 87(419): 659–671. doi:[10.2307/2290201](https://doi.org/10.2307/2290201).
- Lee, R.D. and Miller, T. (2001). Evaluating the performance of the Lee-Carter method for forecasting mortality. *Demography* 38(4): 537–549. doi:[10.1353/dem.2001.0036](https://doi.org/10.1353/dem.2001.0036).
- Li, N. and Lee, R.D. (2005). Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method. *Demography* 42(3): 575–594. doi:[10.1353/dem.2005.0021](https://doi.org/10.1353/dem.2005.0021).
- Lynch, J. and Smith, G.D. (2005). A life course approach to chronic disease epidemiology. *Annual Review of Public Health* 26(1): 1–35. doi:[10.1146/annurev.publhealth.26.021304.144505](https://doi.org/10.1146/annurev.publhealth.26.021304.144505).
- McNown, R. (1992). Modeling and forecasting U.S. mortality: Comment. *Journal of the American Statistical Association* 87(419): 671–672. doi:[10.2307/2290202](https://doi.org/10.2307/2290202).
- Mehta, N. and Chang, V. (2009). Mortality attributable to obesity among middle-aged adults in the United States. *Demography* 46(4): 851–872. doi:[10.1353/dem.0.0077](https://doi.org/10.1353/dem.0.0077).
- Murray, C.L.J. and Lopez, A.D. (eds.) (1996). *The Global Burden of Disease*. Harvard University Press and WHO.
- Oeppen, J. and Vaupel, J. (2002). Broken limits to life expectancy. *Science* 296(5570): 1029–1031. doi:[10.1126/science.1069675](https://doi.org/10.1126/science.1069675).
- Office of the Surgeon General (1964). Smoking and health, report of the advisory committee to the surgeon general of the public health service. United States Public Health Service. (Public Health Service Publication 1103).
- Olshansky, S.J. (1988). On forecasting mortality. *The Milbank Quarterly* 66(3): 482–530. doi:[10.2307/3349966](https://doi.org/10.2307/3349966).
- Olshansky, S.J., Goldman, D., Zheng, Y., and Rowe, J. (2009). Aging in America in the twenty-first century: Demographic forecasts from the MacArthur Foundation Research Network on an aging society. *Milbank Quarterly* 87(4): 842–862. doi:[10.1111/j.1468-0009.2009.00581.x](https://doi.org/10.1111/j.1468-0009.2009.00581.x).
- Olshansky, S.J., Passaro, D., Hershov, R., Layden, J., Carnes, B., Brody, J., Hayflick, L., Butler, R., Allison, D., and Ludwig, D. (2005). A potential decline in life expectancy in the United States in the 21st century. *The New England Journal of Medicine* 352(11):

- 1138–1145. doi:10.1056/NEJMs043743.
- Parascandola, M. (2004). Skepticism, statistical methods, and the cigarette: A historical analysis of a methodological debate. *Perspectives in Biology and Medicine* 47(2): 244–261. doi:10.1353/pbm.2004.0032.
- Peace, L.R. (1985). A time correlation between cigarette smoking and lung cancer. *The Statistician* 34(4): 371–381. doi:10.2307/2987825.
- Peeters, A., Barendregt, J., Willekens, F., Mackenbach, J., Al Mamun, A., and Bonneux, L. (2003). Obesity in adulthood and its consequences for life expectancy: A life-table analysis. *Annals of Internal Medicine* 138(1): 24–32.
- Preston, S. and Stokes, A. (2010). Is the high level of obesity in the United States related to its low life expectancy? Population Studies Center, University of Pennsylvania. (Working Paper 2010-08).
- Preston, S. and Wang, H. (2006). Sex mortality differences in the United States: The role of cohort smoking patterns. *Demography* 43(4): 631–646. doi:10.1353/dem.2006.0037.
- Preston, S.H., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Oxford, England: Blackwell.
- Prospective Studies Collaboration (2009). Body-mass index and cause-specific mortality in 900 000 adults: Collaborative analyses of 57 prospective studies. *The Lancet* 373(9669): 1083–1096. doi:10.1016/S0140-6736(09)60318-4.
- Robins, J.M. (2008). Causal models for estimating the effects of weight gain on mortality. *International Journal of Obesity* 32: S15–S41. doi:10.1038/ijo.2008.83.
- Rogers, R., Hummer, R., Krueger, P., and Pampel, F. (2005). Mortality attributable to cigarette smoking in the United States. *Population and Development Review* 31(2): 259–292. doi:10.1111/j.1728-4457.2005.00065.x.
- Romero-Corral, A., Montori, V., Somers, V., Korinek, J., Thomas, R., Allison, T., Mookadam, F., and Lopez-Jimenez, F. (2006). Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: A systematic review of cohort studies. *The Lancet* 368(9536): 666–678. doi:10.1016/S0140-6736(06)69251-9.
- Sloan, F., Ostermann, J., Picone, G., Conover, C., and Taylor, D. (2004). *The price of smoking*. Cambridge, Mass.: The MIT Press.
- Snijder, M.B., van Dam, R.M., Visser, M., and Seidell, J.C. (2006). What aspects of body

- fat are particularly hazardous and how do we measure them? *International Journal of Epidemiology* 35(1): 83–92. doi:10.1093/ije/dyi253.
- Soneji, S. and King, G. (2011). Statistical security for social security. *Demography* (forthcoming). <http://gking.harvard.edu/files/abs/ssc-abs.shtml>.
- Stewart, S.T., Cutler, D.M., and Rosen, A.B. (2009). Forecasting the effects of obesity and smoking on U.S. life expectancy. *The New England Journal of Medicine* 361(23): 2252. doi:10.1056/NEJMs0900459.
- Sturm, R. (2002). The effects of obesity, smoking, and drinking on medical problems and cost. *Health Affairs* 21(2): 245–253. doi:10.1377/hlthaff.21.2.245.
- Torri, T. and Vaupel, J.W. (2011). Forecasting life expectancy in an international context. *International Journal of Forecasting* 27. (Forthcoming).
- Vollgraff, J.A. (ed.) (1950). *Oeuvres complètes de Christiaan Huygens. Publiées par la Société hollandaise des sciences*. La Haye: M. Nijhoff.
- Wang, H. and Preston, S. (2009). Forecasting United States mortality using cohort smoking histories. *Proceedings of the National Academy of Sciences* 106(2): 393–398. doi:10.1073/pnas.0811809106.
- Wardle, J., Carnell, S., Haworth, C., and Plomin, R. (2008). Evidence for a strong genetic influence on childhood adiposity despite the force of the obesogenic environment. *American Journal of Clinical Nutrition* 87(2): 398–404.
- Wessel, T.R., Arant, C.B., Olson, M.B., Johnson, B.D., Reis, S.E., Sharaf, B.L., Shaw, L.J., Handberg, E., Sopko, G., Kelsey, S.F., Pepine, C.J., Bairey, M., and Noel, C. (2004). Relationship of physical fitness vs body mass index with coronary artery disease and cardiovascular events in women. *Journal of the American Medical Association* 292(10): 1179–1187. doi:10.1001/jama.292.10.1179.
- Wilmoth, J. (2005). On the relationship between period and cohort mortality. *Demographic Research* 13(11): 231–280. doi:10.4054/DemRes.2005.13.11.
- Wright, V.B. (1986). Will quitting smoking help medicare solve its financial problems? *Inquiry* 23(1): 76–82.
- Yang, W., Kelly, T., and He, J. (2007). Genetic epidemiology of obesity. *Epidemiologic Reviews* 29(1): 49–61. doi:10.1093/epirev/mxm004.

A Appendix: Forecasting methodology

In this appendix, we summarize the forecasting methodology described in Girosi and King (2008) (Section A2), explain our extension of the Girosi-King approach to facilitate prior specification (Section A3), and give an empirical example (Section A4).

A1 Overview of forecasting methodology

Here we provide a concise overview of our forecasting model. We begin with observed conditional probabilities of death over age and time. As a baseline, we first start with a simple least squares regression where the dependent variable is the logarithm of the conditional probability of death for a given age, and the independent variable is time. Second, we add potentially informative covariates, namely a cohort's previous smoking patterns. Third, we discuss the advantages and disadvantages of the linear regression framework. Finally, we present a solution to these disadvantages.

First, consider a simple linear regression of the logarithm of the conditional probability of death as a function of year. For simplicity, we focus on the the age [75, 76) years. We may write this regression model as: $E(\log(q_{75,t})) = \beta^{(0)} + \beta^{(1)}\text{year}_t$, where $q_{75,t}$ is the conditional probability of death for age 75 years and time t , $\beta^{(0)}$ is the intercept parameter and $\beta^{(year)}$ is the slope parameter for the covariate year. The resulting forecasts are often plausible, especially of all-cause mortality in low-mortality countries with a short forecasting window.

Second, consider a multiple linear regression model of the logarithm of the conditional probability of death as a function of year and a cohort's smoking patterns 25 years earlier. The model is similar to simple linear regression and offers the promise of incorporating additional covariates. We may write this regression model as: $E\log(q_{75,t}) = \beta^{(0)} + \beta^{(year)}\text{year}_t + \beta^{(smoking)}\text{smoking}_{50,t-25}$, where $\beta^{(0)}$ and $\beta^{(year)}$ are similarly defined, and $\beta^{(smoking)}$ is the slope parameter for the covariate smoking. The covariate smoking ("lagged smoking") conveys the cohort's earlier smoking patterns when it was 50 years of age 25 years ago.

Third, linear regression is a useful framework for demographic forecasting. With an appropriate transformation (in our case the natural logarithm), conditional probabilities of death may be modeled as a linear function of year. We can also include potentially informative covariates, either a 'cohort effect' (e.g., a cohort's earlier smoking patterns) or a 'period effect' (e.g., the 1964 US Surgeon General's Report on Smoking and Health (Office of the Surgeon General 1964)). We can further specify covariates for some age groups and not for others, depending on the context. For example, we only consider lagged smoking for ages 50 years and older. In addition to easily incorporating covariates, we can also assess the plausibility of resulting forecasts, as is done in Figure 3. We may

write the regression for all ages concisely as:

$$E(\log(q_{a,t})) = \begin{cases} \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t, & \text{if } a < 50 \\ \beta_a^{(0)} + \beta_a^{(year)} \text{year}_t + \beta_a^{(smoking)} \text{smoking}_{a-25,t-25}, & \text{if } a \geq 50 \end{cases} \quad (5)$$

where $\text{smoking}_{a-25,t-25}$ is smoking prevalence 25 years earlier in age and 25 years earlier in time.

A multiple linear regression that includes time and lagged smoking is a plausible model specification. Yet most demographers and population biologists would not expect the resulting forecasts—adult mortality that does not increase monotonically and age profiles that do not maintain the quintessential all-cause shape. The problem is not the demographic knowledge that was used to specify the model. Rather, the problem is the model itself. Specifically, individual multiple regression models treat each cross-section of age-specific mortality separately. Therein lies the research gap—how to incorporate potentially informative covariates while still maintaining plausible and smooth mortality forecasts.

Fourth, we propose a solution to this statistical problem that is based on the intuitive appeal of multiple linear regression and maintains reasonable demographic properties. Using the framework developed by Girosi and King (2008), we jointly estimate the multiple regression models for each age. The only difference between standard multiple linear regression and our model is how we estimate the regression coefficients. We follow a Bayesian perspective and draw the regression coefficients from posterior distributions. We describe more details in Section A2, and the model is fully developed and rigorously evaluated in Girosi and King (2008). The Bayesian methodology enables demographers to specify if and how to smooth mortality across age, time, and cohort through the use of smoothness functionals and smoothness parameters. Demographers are able to tune smoothness parameters that control:

1. Smoothness over time: how much an age-specific mortality rate changes over time,
2. Smoothness over age: how much age-specific mortality rates change over time between neighboring ages,
3. Smoothness over age and time: how much the pace of change in an age-specific mortality rate over time varies the pace of change in a neighboring age-specific mortality rate over time.

For example a demographer might believe in smoothness over time and state that age 75 mortality in 2020 will be a similar value as age 75 mortality in 2019. A demographer might also believe in smoothness over age and state that age 75 mortality in 2020 will be slightly higher than age 74 mortality in 2020 and slightly lower than age 76 mortality in 2020. Finally, a demographer might believe in smoothness over age and time and state that if age 75 mortality is declining at a certain pace over time, age 74 and age 76 will

decline at a similar, though not necessarily identical, pace. The key innovation is that demographic experts state their prior beliefs on the expected value of mortality, of which a great deal is known.

The smoothing parameters can be tuned to extreme levels and become equivalent to linear regression. For example, if we fully tune out all smoothness parameters, each cross-section of age-specific mortality would be modeled independently, yielding a forecast equivalent to multiple linear regression. Another example would be if we fully tune on smoothness over time and fully tune out smoothness over age and smoothness over age and time. The resulting forecast would be equivalent to simple linear regression and ignore any covariate other than time. Of course, a better forecast would be one that tunes the smoothness parameters in a more nuanced manner, incorporating covariate effects and maintaining likely demographic patterns. We consider a class of smoothness parameters and determine if and how to specify covariates. Common to all regression methods, a covariate will only affect the out-of-sample prediction (i.e., the forecast) if, and only if, there is an empirical relationship in the observed historical data. A group of experts may share some common beliefs on basic demographic properties and differ on other patterns. We are able to incorporate this range of expert beliefs by considering a range of smoothness parameters and covariate specifications. This flexibility in modeling forms the basis of what is formally known as ‘robust Bayesian analysis’.

A2 Bayesian hierarchical model

Here we summarize the forecasting methodology developed by Girosi and King (2008). Let \mathcal{A} be a set of ages and \mathcal{T} be a set of years that define the age and time mortality window. Let n_a represent the width (years) of an age interval starting at age $a \in \mathcal{A}$. Let ${}_{n_a}D_{a,t}$ be the number of deaths between ages a and $a + n_a$ in time t . Let $P_{a,t}$ be the population at exact age a in time t . Then, the conditional probability of death is defined as ${}_{n_a}q_{a,t} = {}_{n_a}D_{a,t}/P_{a,t}$ for all $a \in \mathcal{A}$ and $t \in \mathcal{T}$. For brevity, we drop the left subscript representing the width of the age interval, n_a .

Consider first a single cross-section of age-specific mortality observed over time and modeled individually with this separate linear-normal (least-squares) regression:

$$\log(q_{a,t}) \sim N(\mu_{a,t}, \sigma_a^2) \quad (6)$$

$$\mu_{a,t} = \mathbf{Z}_{\mathbf{a},t} \beta_a, \quad (7)$$

where $\mu_{a,t} \equiv E(q_{a,t})$ and $q_{a,t}$ is assumed to be independent over time after conditioning on \mathbf{Z} , where \mathbf{Z} is a matrix of covariates that may include lags of $\log(q_{a,t})$. This model is a multiple linear regression model written in more compact matrix notation. Suppose we assume that there is at least one death in all age-time groups. Then under these assump-

tions, the likelihood function for the model is:

$$\mathcal{P}(q|\beta_a, \sigma_a^2) \propto \prod_i \sigma_a^{-T} \exp\left(\frac{-1}{2\sigma_a^2} \sum_i (\log(q_{a,t}) - \mathbf{Z}_{\mathbf{a},t}\beta_a)^2\right). \quad (8)$$

This simple linear-normal model forms the lowest level structure of the hierarchical Bayesian approach developed by Girosi and King (2008). As such, the coefficients, β_a , and standard deviations, σ_i , are random variables with their own prior distributions. We denote the prior distribution for the standard deviation, σ_i , as $\mathcal{P}(\sigma)$. The prior distribution for the coefficients, β_a , which usually depend on one or more hyperparameters, θ , is denoted by $\mathcal{P}(\beta|\theta)$. Girosi and King (2008) specify the priors $\mathcal{P}(\sigma)$ and $\mathcal{P}(\theta)$ as a Gamma distribution for computational simplicity. The prior $\mathcal{P}(\beta|\theta)$ is treated as informative and is the main way that this approach differs from independent linear regressions. Using the likelihood function specified in Equation 8 and assuming that σ is prior independent of β and θ , the posterior distribution of β , σ , and θ conditional on the data is,

$$\mathcal{P}(\beta, \sigma, \theta|q) \propto \mathcal{P}(q|\beta, \sigma)[\mathcal{P}(\beta|\theta)\mathcal{P}(\theta)\mathcal{P}(\sigma)], \quad (9)$$

where the prior $\mathcal{P}(\beta, \sigma, \theta) \equiv \mathcal{P}(\beta|\theta)\mathcal{P}(\theta)\mathcal{P}(\sigma)$. Once the prior densities have been specified, we summarize the posterior density of β with its mean,

$$\beta^{Bayes} \equiv \int \beta \mathcal{P}(\beta, \sigma, \theta|q) d\beta d\theta d\sigma. \quad (10)$$

The variability around the mean represents one source of uncertainty (discussed in 6). As Girosi and King (2008) note, by choosing a suitable prior density for the coefficients, β , we can summarize and formalize prior demographic knowledge that shows how the coefficients are related to each other and how information is shared among cross-sections of age-specific mortality over time. Furthermore, if the prior for the coefficients is specified appropriately, the information content of the estimates of β will increase, leading to more informative and accurate forecasts.

Unfortunately, these coefficients are never observed and so the claim that anyone has prior knowledge about them is dubious. Recall that these coefficients on smoking and obesity are based on population aggregates and so do not represent the causal effects commonly estimated at the individual level. In addition, if we know that adjacent age groups have similar mortality levels, this does *not* mean that they have similar coefficients. In fact, if the covariates are not smooth, then the coefficients must also not be smooth in order to produce smooth mortality over the age groups.

Fortunately, over more than three centuries, demographers have gathered a great deal of information about mortality in numerous time periods and geographic regions, which

we can conveniently use to set priors on expected mortality, instead of focusing on the coefficients. We follow Girosi and King (2008) and employ a two-step strategy to derive a prior density on the regression coefficients, β_a . First, the prior is specified on the expected value of log-mortality. Second, this information is translated on the coefficients so standard Bayesian computation and estimation strategies can be used. In the first step, expert knowledge and information is translated into a set of L statements about the properties of μ , the expectation of log-mortality. The l th statement, for example, is denoted as $H_l[\mu]$ for $l \in [1, \dots, L]$. The statements are also known as smoothness functionals, which are then put in a probabilistic form. For example, a normal probability density prior for μ might be,

$$\mathcal{P}(\mu|\theta) \propto \left(-0.5 \sum_l \theta_l H_l[\mu] \right) \equiv \exp(-0.5H[\mu, \theta]), \quad \mu \in \mathbb{R}^{N_{\mathcal{T}} \times N_{\mathcal{A}}}, \quad (11)$$

where $N_{\mathcal{T}}$ and $N_{\mathcal{A}}$ are the cardinalities of sets \mathcal{T} and \mathcal{A} , respectively. In the second step, merely by substituting in, the prior density is transformed in terms of the coefficients β as,

$$\mathcal{P}(\beta|\theta) \propto \exp(-0.5H^\mu[\beta, \theta]), \quad (12)$$

where $H^\mu[\beta, \theta] \equiv H[\mathbf{Z}\beta, \theta]$. This works because the subspace of \mathbf{Z} is invertible (with a constant Jacobian), which represents the support of the prior $\mu - \mathbf{Z}\beta$. An attractive consequence of this procedure is that we only need to specify what we are willing to assume since the resulting prior is improper. This helps code both what is known ex ante — such as smoothness of expected log-mortality over age and time, as well as the shape of the mortality age profile — and what is not known — such as the level of expected mortality at any one time — to which the prior is indifferent so we can let the data speak more loudly.

A3 Search algorithm for setting priors

We now discuss a search algorithm we developed to facilitate choosing the specific values of the priors, given a choice of smoothness functional and the methodology described in Appendix A2. Suppose, given our empirical evidence, we wish to smooth over age and time. We consider the following prior that includes two smoothness functionals,

$$\begin{aligned} H[\mu, \theta_{age}, \theta_{time}] \equiv & \frac{\theta_{age}}{N_{\mathcal{A}} N_{\mathcal{T}}} \int_0^{N_{\mathcal{T}}} \int_0^{N_{\mathcal{A}}} \left(\frac{d^2}{da^2} (\mu(a, t) - \bar{\mu}(a)) \right)^2 da dt \\ & + \frac{\theta_{time}}{N_{\mathcal{A}} N_{\mathcal{T}}} \int_0^{N_{\mathcal{T}}} \int_0^{N_{\mathcal{A}}} \left(\frac{d^2}{da^2} \mu(a, t) \right)^2 da dt \end{aligned} \quad (13)$$

where $N_{\mathcal{A}}$ and $N_{\mathcal{T}}$ are similarly defined as they were in Equation 11 for age and time. Given this prior distribution, we only need to choose the smoothness parameters θ_{age} and θ_{time} .

As Girosi and King (2008) note, the value of these smoothness parameters determines how much weight is put on the prior as compared to the data in the estimation. The value of the smoothness parameter also determines how smooth the forecast will be. Instead of specifying the value of θ_{age} and θ_{time} , we specify the value of the standard deviation of the priors, σ_{age} and σ_{time} , which are equivalent but more easily interpretable. A small value of σ_{age} , for example, would impose more weight on the prior and impose more smoothing over age compared to a relatively larger value of σ_{age} . The same is true for values of σ_{time} and smoothing over time.

We consider a grid of σ_{age} and σ_{time} values. In practice, the grid is either evenly spaced on the true scale or evenly spaced on the logarithmic scale. Each point on the grid represents a separate model and typical grid would contain several hundred points. For each point on the grid, we set aside the most recent years of historical observation as a validation period (\mathcal{T}^*) and forecast. For each forecast, we also calculate four summary measures. Our first summary measure is the prediction error (measured by the root mean square error) for all age groups and years in the validation period. Second, we measure the arc lengths of the age profiles in the validation period once the mean age profile from the validation period is removed. Third, we measure the arc length of each time profile's deviation from its own trend line. We consider various degrees of polynomials to which the time profiles are smoothed. Fourth, we remove the constant from all the time profiles (so that each has mean zero) and measure the arc length of their deviations from the mean time profile.

We derive an objective function based on the four summary measures. For brevity, we use the abbreviation PE for prediction error, Age AL for age arc length, Time AL for time arc length, and Trend Dev for trend deviation. The objective function is,

$$f(\text{PE, Age AL, Time AL, Trend Dev}) = w_1 \sqrt{\text{PE}} + w_2(\text{Age AL}) + w_3(\text{Time AL}) + w_4(\text{Trend Dev}), \quad (14)$$

where \vec{w} is chosen by the user and $\sum_{i=1}^4 w_i = 1$. We find the $(\sigma_{age}, \sigma_{time})$ that minimizes the objective function and this choice of priors represents our 'best' forecast.

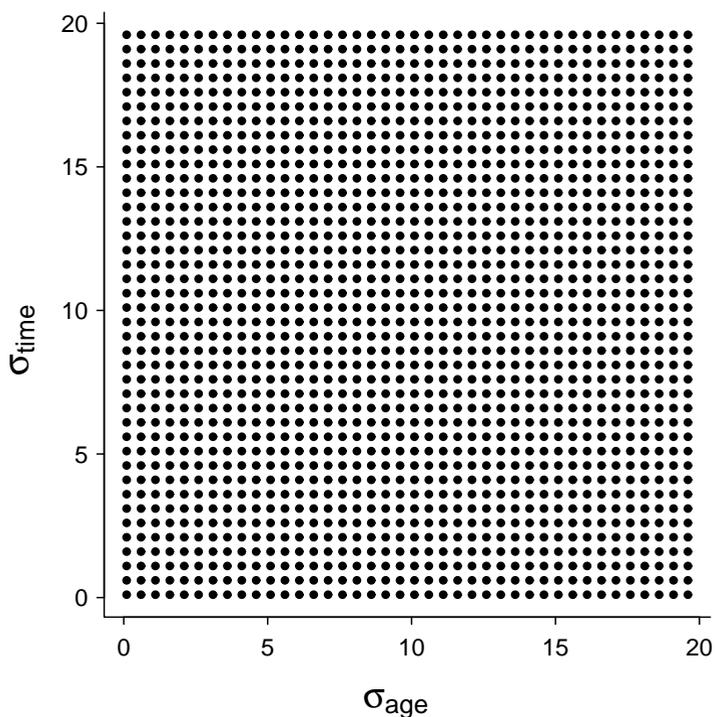
A4 Forecasting methodology example

Finally, we offer a simple worked example of male mortality including time, lagged smoking, and lagged obesity as covariates. Suppose we wish to incorporate our belief that age-specific log-mortality is smooth over time. We also believe nearby ages will share

similar, though not identical, patterns over time. We represent these beliefs as prior information through a smoothness functional. The goal of the search algorithm is to select a sufficiently large weight for the prior that yields a forecast with similar patterns across time for nearby ages. Yet, the weight is not too large such that the age-specific patterns are perfectly identical over time. And if the data contradict the prior sufficiently, the prior is automatically down-weighted or ignored in the final forecasts.

We begin with a two-dimension grid of σ_{age} and σ_{time} values. For example, we chose candidate σ_{age} and σ_{time} values evenly spaced on the logarithmic scale between $\log(0.1)$ and $\log(20)$. The resulting sigma grid is shown in Figure 7.

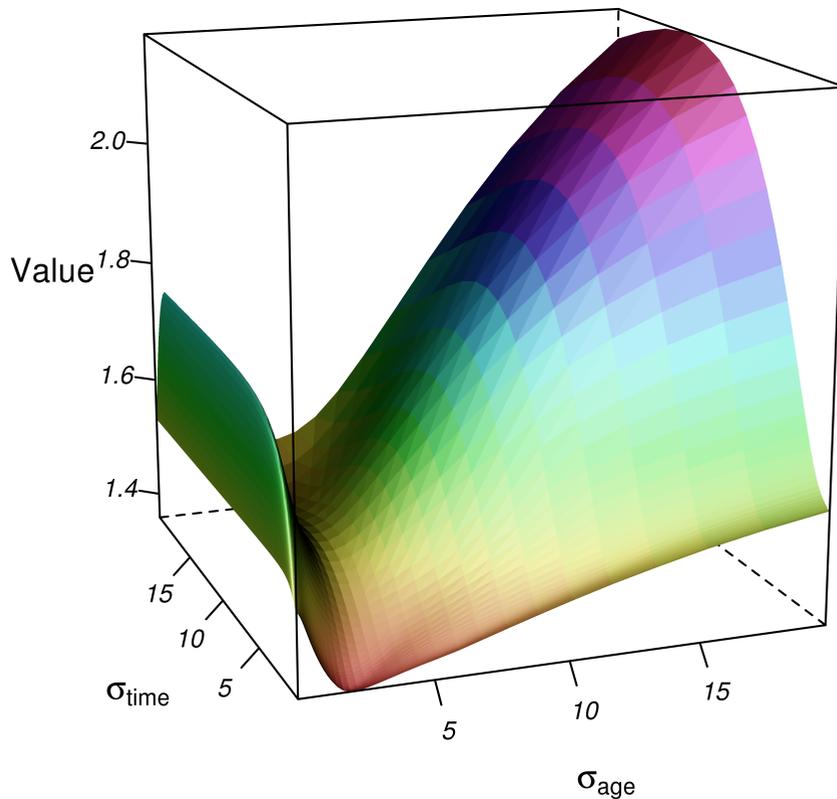
Figure 7: **Sigma grid**



Note: Each $(\sigma_{age}, \sigma_{time})$ point represents a separate model.

We chose the weight vector $\vec{w} = (0.4, 0.2, 0.2, 0.2)$. Given this weight vector, the prediction error component of the objective function receives the greatest weight, although the age arc length, time arc length, and trend deviation components also receive substantial weight. The objective function surface is plotted in Figure 8. The sigma combination $(\sigma_{age}, \sigma_{time})$ that yields the lowest objective function is (1.99, 0.10). The forecast based on this prior distribution represents our ‘best’ forecast.

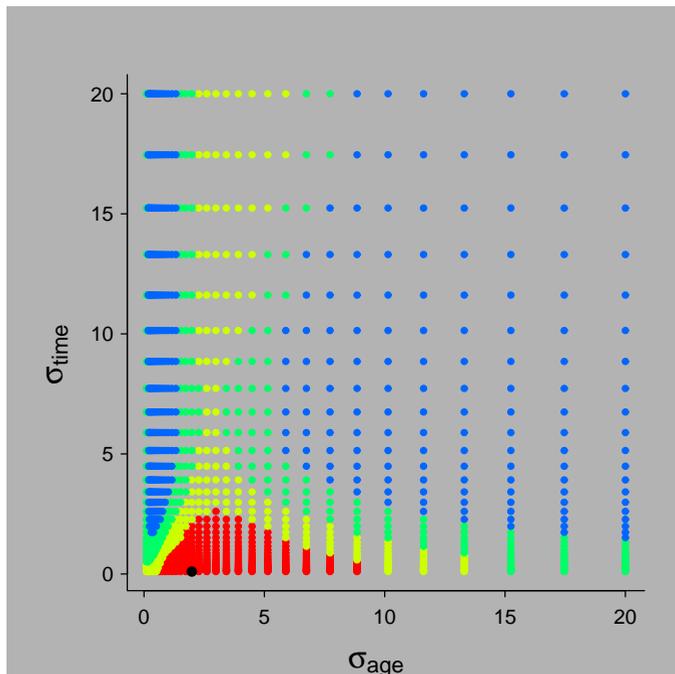
Figure 8: Objective function



Note: Three-dimensional plot of σ_{age} , σ_{time} , and corresponding objective function values.

The objective function surface is also used to examine model-based uncertainty based on prior specification and form uncertainty intervals for the forecast. In Figure 9, we replot the sigma grid and color each sigma combination by its objective function value percentile. For example, the red colored $(\sigma_{age}, \sigma_{time})$ values represent the highest quartile of objective function values. The yellow colored, green colored, and blue colored sigma combinations represent the second, third, and fourth highest quartiles of objective function values, respectively. The black colored sigma combination represents the ‘best’ forecast. A user may select the percentile cutoff (e.g., 50%). Forecasts resulting from sigma combinations with objective functions below the percentile cutoff represent the uncertainty interval.

Figure 9: Sigma grid with objective function quartiles



Notes: Sigma grid color-coded to represent the quartile of objective function values. Sigma combinations with the lowest objective function quartile are shown in red, second lowest in yellow, second highest in green, and highest in blue. The sigma combination minimizing the objective function is shown as a black dot.