THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF RISK MANAGEMENT

Gender Differences in U.S. Life Expectancy Forecasts: A State-Level Analysis

JUSTIN MORRISON
SPRING 2024

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Actuarial Science
with honors in Actuarial Science

Reviewed and approved* by the following:

Nan Zhu
Associate Professor of Risk Management
Thesis Supervisor

Amanda Hammell
Professor of Practice and Faculty-In-Charge of the Actuarial Science Program
Honors Adviser

* Electronic approvals are on file.

ABSTRACT

The level and trend of life expectancy are key forces of concern to both public policymakers and insurance companies; it is an indicator of the quality of life of the general population or the financial obligation to insurance policyholders. Understanding historical life expectancy trends in the U.S. already has significant literature and research. Many studies have examined some of the demographic and socioeconomic differences in life expectancy. However, forecasting the direction of mortality improvement at the state level is an area of research with untapped potential. In this thesis, the Double-Gap Life Expectancy model, a combination of simple linear regression and ARIMA time series models, is used to forecast life expectancy for gender-specific and aged-based sub-populations for the six most populous states in the U.S. This paper quantifies differences in life expectancy across these representative states and explores qualitative reasons for these differences using a correlation-based approach.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ACKNOWLEDGEMENTS

Thank you to Mandy Hammell and Nan Zhu for their support during my thesis writing process. Mandy ensured that I stayed on track with my thesis timeline. Nan helped me choose my thesis topic and gave great suggestions in finding data and organizing this paper.

I also thank the Risk Management Department for providing me with a foundation to succeed in the classroom, on SOA exams, and in my professional career. Specifically, I would like to thank Mandy Hammell, Nan Zhu, Amin Hassan Zadeh, Lisa Posey, and Dave Cather for their superior instruction in my major-related coursework.

In addition, I want to give a special shoutout to Priyangi Bulathsinhala. Priyangi's applied time series course gave me the inspiration and knowledge needed to write this thesis.

Lastly, I thank my Penn State actuarial science peers, my Prudential intern class, my friends, and my family for their unconditional support throughout the thesis writing process. I could not have completed this paper without them pushing me to do my best.

Chapter 1

Introduction and Literature Review

Definition and Overview of Life Expectancy in the U.S.

Life expectancy is the expected value of an individual's future lifetime random variable, given their current age. In theory, there are two measures of life expectancy: complete expectation of life and curtate expectation of life. The complete expectation of life is a continuous representation of life expectancy and is calculated by integrating a survival function up to a limiting age. The curtate expectation of life is a discrete representation of life expectancy (i.e., expected number of complete years lived) and is calculated by summing survival probabilities up to a limiting age. In practice, there are two methods to calculate life expectancy: period life expectancy and cohort life expectancy. Period life expectancy estimates life expectancy by considering mortality probabilities from a single-year life table without considering future changes in mortality (mortality improvement). Conversely, cohort life expectancy uses a combination of mortality probabilities from life tables for groups (cohorts) of individuals with the same birth years and projections for mortality improvement in future years (Buxton, 2023). For the methods used later in this paper, the complete expectation of life is estimated using period life expectancy.

Life expectancy is a key quality of life statistic for both public policymakers and insurance companies. Every decade, the Census Bureau releases reports and data about the population and the economy. This data includes changes and trends in life expectancy by state, gender, race, and ethnicity. Both life and health insurers should monitor and project future changes in life expectancy as this external force directly affects their operations and risk management efforts. Mortality improvement (year-over-year changes in mortality rates) is a common actuarial assumption used in developing and maintaining pricing and reserving frameworks for insurance products. For life insurers, changes in their insured population's life expectancy present two competing risks: mortality risk and longevity risk. Mortality risk results from policyholders dying earlier than anticipated while longevity risk results from policyholders living longer than anticipated. Thus, mortality risk is an issue for life insurance products since the benefit is paid upon death of the policyholder. Conversely, longevity risk is an issue for annuity products, long-term care insurance, and pensions since the benefits are contingent on the policyholder still being alive at the time of the benefit payment.

Besides product risks, life insurers must also consider how life expectancy varies between the general population and its insured population, and between different states. Life insurance underwriting, which involves a combination of questionnaires and health tests, determines if an individual should be sold life insurance, and if they are eligible, then they are categorized into an appropriate risk class based on factors such as age, gender, and smoking status. Due to how rigorous underwriting is for life insurance with a face amount (benefit) greater than $100,000, the insured population should be healthier on average than the general population. For Munich Re, a German-based international life reinsurer, this proves to be the case as its policy portfolio has more cancer-related deaths (which are difficult to predict through traditional life underwriting tests) than the general population and less cardiovascular disease-related deaths (which can be predicted through blood tests, fluid tests, and family history) than the general population (Serykh and Yang, 2019). Therefore, life insurance underwriting screens out some (but not all) high-risk (unhealthy) individuals, which should result in higher life expectancy for its customers than the general population, on average.

Even though health insurance is a short-term coverage (i.e., one-year policies that may or may not be renewed), health insurers should be aware of trends in life expectancy. By understanding these trends, health insurers can make more informed projections regarding future expected health claims. For example, the longer individuals live, the more benefits may need to be paid out, especially given higher healthcare utilization rates among the elderly population. However, for younger policyholders, longer life expectancy would result in lower costs for health insurers in the long run as preventative healthcare would temporarily increase short-term claim costs in exchange for lower long-term claim costs.

Both public policymakers and insurers potentially impact life expectancy trends, which may explain statewide differences in life expectancy. Legislation, executive orders, administrative rules, and court rules may directly or indirectly impact life expectancy in an individual state or nationwide. According to Montez et. al, 2020, states with more liberal legislation (with 'liberal' defined as "expanding state power for economic regulation and redistribution or for protecting marginalized groups, or restricting state power for punishing deviant social behavior") when it comes to tobacco, labor, immigration, and environmental policies tend to have higher life expectancy for both men and women, on average. If every state adopted the most liberal policies in these aforementioned areas, life expectancy is estimated to improve by 2.8 years for women and 2.1 years for men (Montez et al., 2020). Alternatively, the most conservative policies reduce life expectancy significantly. Life expectancy in the U.S. only marginally increased from 2010 to 2014, but if conservative policies had not been implemented in certain states, this increase would have been 25% steeper for women and 13% steeper for men

(Montez et al., 2020). In the private sector, health insurers improve survival outcomes for insured individuals versus uninsured individuals due to usage of and access to preventative health services (Woolhandler & Himmelstein, 2017). In the life insurance industry, John Hancock offers discounts and benefits (e.g., lower insurance premiums) for physically active customers based on fitness-tracking data. Thus, insurers can incentivize customers to engage in healthier habits, which improves their life expectancy.

Historical and Future Trends of U.S. Life Expectancy

Since 1940, life expectancy of the overall population in the U.S. has generally increased. Improvements in the diagnosis and treatment of cardiovascular diseases and declines in the percentage of smokers explain some of these life expectancy gains (Patkee & Strange, 2023). However, mortality improvement has slowed since the 1980s and has even stagnated or declined since the 2010s, according to Figure 1 and Table 1 (Reynolds & Dattani, 2023).

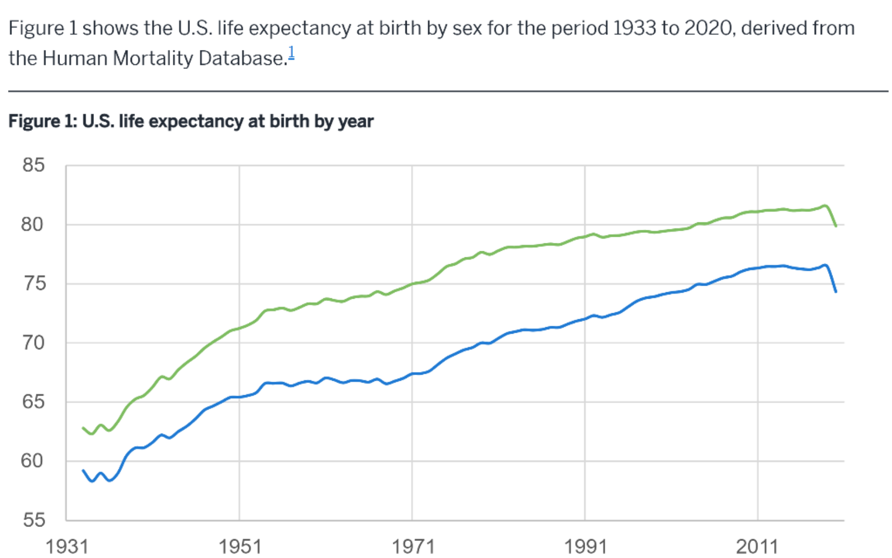

Figure 1: U.S. Life Expectancy at Birth by Year

Table 1: U.S. Average Mortality Improvement by Decade

Start Year	End Year	Male	Female
1971	1980	2.07%	2.20%
1981	1990	1.38%	1.13%
1991	2000	1.73%	0.49%
2001	2010	1.70%	1.47%
2011	2020	-1.63%	-1.17%

According to an SOA report on mortality improvement, mortality rates have not declined significantly since 2012. The only causes of death to improve during this period were cancer and heart disease due to new treatments and preventative measures (Serykh and Yang, 2019). Decades of high mortality improvement followed by recent years of little mortality improvement are unsurprising since mortality improvement is not constant; as seen in Figure 2, life expectancy improves in waves as new innovations in medicine and healthcare are discovered (Patkee & Strange, 2023). In this figure, the size of these ‘bubbles’ (representing medical innovations) corresponds to the relative mortality improvement. Thus, the more impactful a medical innovation is at reducing mortality, the bigger the size of its bubble in Figure 2. The relative impact of these medical innovations can also be tied to the steepness of changes in life expectancy in Figure 1.

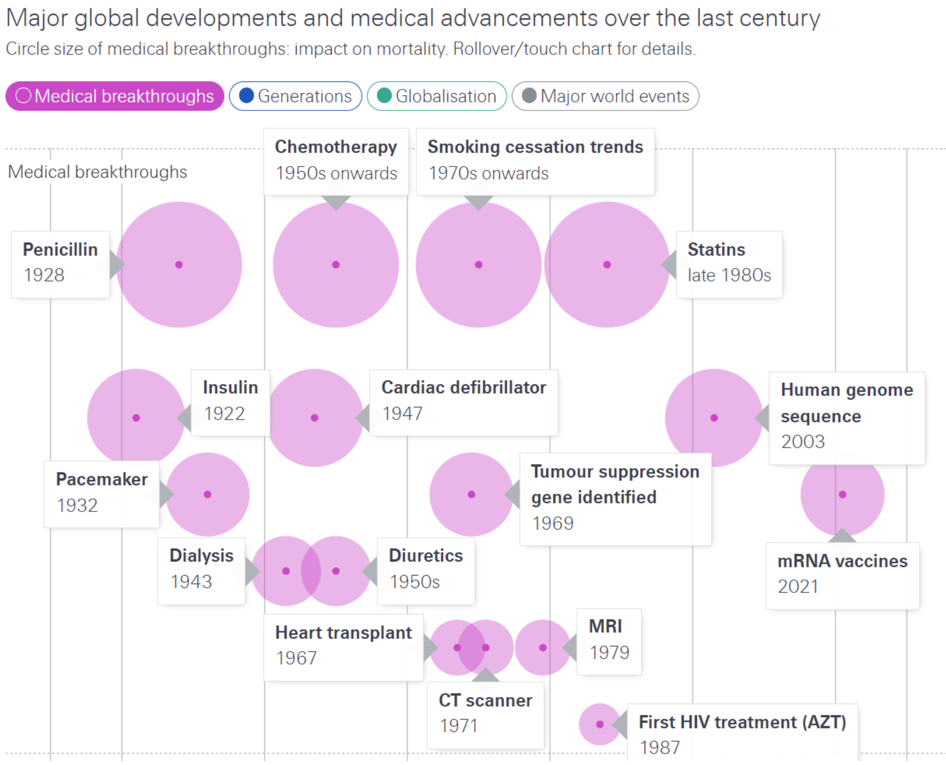

Figure 2: Major Medical Advancements over the Past Century

Since the late 2010s, some years have even resulted in negative mortality improvement. In 2017, accidental deaths due to opioids resulted in a 0.4% decrease in mortality improvement (Serykh and Yang, 2019). The next year with negative mortality improvement was 2020 due to the COVID-19 pandemic. COVID-19 disproportionately resulted in premature deaths for the elderly and individuals with pre-existing conditions, which offset decades of mortality improvement gains for the entire U.S. population, as shown in Figure 1 (Reynolds & Dattani, 2023).

In the U.S. and most other countries, female life expectancy tends to be higher than male life expectancy. In other words, there is heterogeneity in life expectancy when partitioned by gender. According to the judge in the 2011 *Test-Achats* case, speculated reasons for these gender-specific life expectancy differences include higher stress jobs, poorer diets, less exercise, and increased alcohol/drug use among men compared to women, on average. In addition, age-zero male (newborn) life expectancy is lower than age-zero female life expectancy due to the accident hump: the period when adolescent and teenage males start to engage in riskier behavior that may result in accidental death. The same phenomenon is not observed in female mortality rate trends, which results in a higher newborn life expectancy than males, on average.

Changes in life expectancy and mortality rates since the 20th century have also shown significant heterogeneity by geographic location and level of rurality. In the early to mid-20th century, age-adjusted mortality rates in urban areas were higher than those of rural areas due to increased prevalence of illness and disease that resulted from lack of public health and sanitation measures in high-density populations. However, this rural versus urban mortality differential has shifted in the opposite direction since the 1980s as rural areas do not have as much access to healthcare and treatment as urban areas, which may also explain why rural areas have not seen as many decreases in heart disease and cancer-related deaths as urban areas (Rhubart & Santos, 2023). Other causes of death that have contributed to this rural versus urban mortality disparity include COPD, lung cancer, stroke, suicide, and diabetes (Singh & Siahpush, 2014). Due to these deaths related to healthcare access and treatment differences, life expectancy is lower, on average, for rural areas compared to urban areas. Therefore, when adjusting for other confounding factors and only comparing levels of rurality, it is expected a more urbanized state would have higher levels of life expectancy than more rural states.

Furthermore, this difference has only increased over time. From 1969-1971, the difference in urban versus rural life expectancies was only 0.4 years. However, from 2005-2009, this difference increased to 2 years (Singh & Siahpush, 2014). Overall, changes in life expectancy are dependent on geographic location. From 2001 to 2014, changes in life expectancy ranged from an increase of more than 4 years to a decrease of more than 2 years when segmenting the U.S. into commuting zones (Chetty et al., 2016). Therefore, understanding changes in life expectancy across the U.S. should be better studied at the state or local level, data-permitting, due to evidence of significant heterogeneity by geographic location.

As demonstrated in Figure 3, the gap in rural versus urban mortality rates, known as the rural mortality penalty, also varies significantly by race and ethnicity (Rhubart & Santos, 2023).

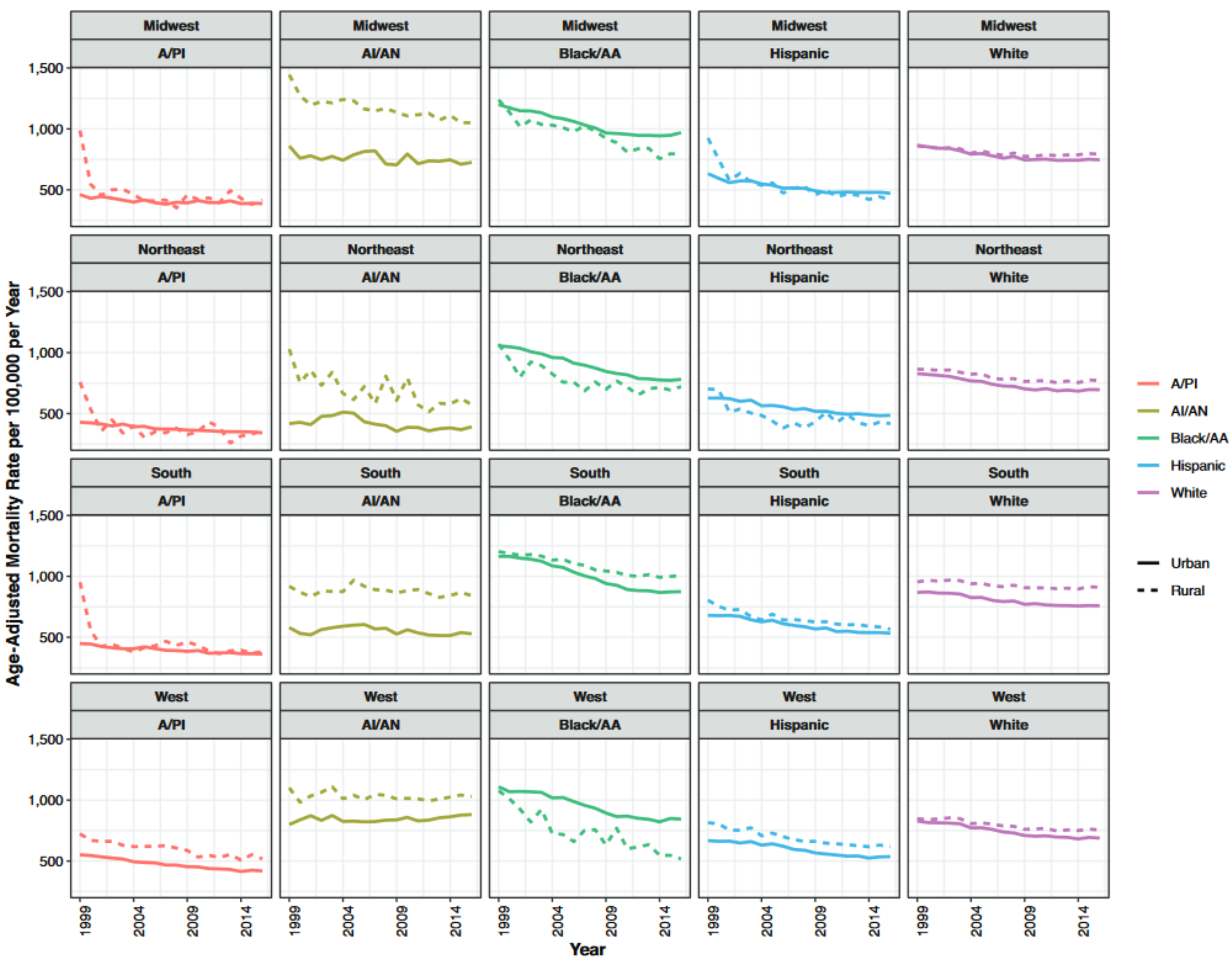

Figure 3: Mortality Rates by U.S. Census Region of Residence and Race and Ethnicity

For example, African Americans only suffer from a rural mortality penalty if they live in the South while Latinos suffer from a rural mortality penalty in both the South and West. In addition, White and Native populations suffer from a rural mortality penalty in all regions of the U.S., but Native populations show the most variation across regions (Rhubart & Santos, 2023). However, the specific causes of these different rural mortality penalties based on the intersection of region and race/ethnicity are currently unknown and are an area for future research.

In addition to geographic and racial/ethnic differences, income inequality provides another lens to understanding differences in life expectancy. Multiple studies have found that socio-economic factors explain around 60% of variation in life expectancy at the county-level in the U.S. (Li & Hyndman, 2021). From 1999 to 2014, the gap in life expectancy between the richest 1% of the U.S. population and the poorest 1% was 14.6 years for men and 10.1 years for women (Chetty et al., 2016). Similar to the rural mortality inequality, differences in life expectancy due to income inequality have also increased over time. From 2001 to 2014, the top 5% of the income distribution experienced large increases in life expectancy, 2.91 years for women and 2.34 years for men, while the bottom 5% experienced only experienced modest gains, 0.04 years for women and 0.32 years for men (Chetty et al., 2016). In the U.S., a person's income is positively correlated with life expectancy The higher an individual's income, the longer their life expectancy is, on average, across both genders and across the entire income distribution (Chetty et al., 2016).

Life expectancy in the U.S. has consistently improved over time besides during the opioid epidemic and the COVID-19 pandemic, but predicting if this positive trend will continue proves difficult as there are arguments for both continued improvement and possible stagnation or decline. Some factors that may lead to continued mortality improvement include advances in medical technology (e.g., new vaccines, treatments, and procedures), improved public health initiatives (e.g., improved public sanitation, cleaner water and air), increased automobile safety (e.g., anti-lock brakes and electronic stability control), and declines in DUI/DWI rates (Reynolds & Dattani, 2023).

However, the impact of these mortality improvements may be temporary and limited. For example, as individuals in the U.S. live longer, neurodegenerative diseases such as Alzheimer's will become a more common cause of death, so life expectancy would not increase substantially until new treatments are discovered and utilized (Patkee & Strange, 2023). In addition, lifestyle factors may also offset any healthcare-related life expectancy gains. Americans' increasingly sedentary lifestyles and unhealthy diets are expected to lead to increased rates of obesity and diabetes that could offset future mortality improvement due to medical advances (Patkee & Strange, 2023). In addition, current inequalities such as geographic location, politics, socio-economic status, and education are likely to expand, which would slow future mortality improvement (Li & Hyndman, 2021).

Alternatively, since life expectancy began to flatten since the 2010s and even declined since 2020, it can be argued that the long-term trend of increasing life expectancy is reversing. Thus, it is not unreasonable to assume a continued decline in life expectancy in the future. Potential causes of a continued decline include climate change (i.e., climate change could result in economic disruptions, tropical disease spreading, natural disasters, pollution, resource depletion, and famine), long COVID (i.e., individuals with pre-existing conditions who survived COVID-19 may suffer long-term, debilitating symptoms that lead to premature deaths), animal-human disease transmission (e.g., HIV/AIDS, bird flu, Ebola, COVID, etc.), alcohol and drug abuse, increasing economic equality (i.e., wealthier individuals have better access to medical care, sanitation, and nutrition, so increases in life expectancy are skewed toward this population), and war (e.g., global wars leading to food and medicine constraints in the supply chain) (Reynolds & Dattani, 2023).

No matter the direction of the trend of future life expectancy, predicting it is complicated by these aforementioned unknowns. No demographer, statistician, or actuary can account for future legislation, wars, pandemics, or climate change. Models may try to forecast future life expectancy, but external factors may be impossible or difficult to incorporate.

Summary of Modern Life Expectancy Forecast Models

The Lee-Carter model is one of the most predominantly used life expectancy and mortality rate trend models since the early 1990s. Often used on its own or as a baseline to compare to newer models, this applied time series method models mortality rates by first applying singular value decomposition on a matrix of historical mortality rates and then forecasting mortality rates using an Autoregressive Integrated Moving Average (ARIMA) time series model. The key assumption of the Lee-Carter model is that annual mortality improvement follows a positive linear trend (Lee & Carter, 1990).

Although the main predictor in the Lee-Carter model is age, there is not a smooth functional form when predicting mortality rates or life expectancy across different ages or years, which makes the model more difficult to interpret (Carins et al., 2008). Taking a different approach, the Carins-Blake-Dowd model includes age, period, and cohort effects and allows for mortality improvement to occur at different ages and times. However, the application of this model is very limited as it is only appropriate for near-retirement ages, so it could only be used to forecast life expectancy for retirement products such as pensions or annuities (Carins et al., 2008).

Both the Lee-Carter and Carins-Blake-Dowd models are extrapolative projection models. These types of models use past trends (i.e., historic mortality rates) to predict future mortality rates, which is only valid assuming past trends continue into the near future. Due to factors that may change the trend of mortality improvement (e.g., medical advances), such models are not suitable for long-term life expectancy forecasts (Carins et al., 2008).

Another drawback of both the Lee-Carter and Carins-Blake-Dowd models is that they only consider a single population at a time. Thus, a comparative analysis of mortality rates or life expectancy between different populations (e.g., countries in the EU, states in the U.S., etc.) would require the assumption that their mortality and life expectancy trends are independent, which is typically an invalid assumption. For both genders, life expectancy is strongly positively correlated over time between EU countries (Pascariu et al., 2018).

The Double-Gap Life Expectancy model solves the challenges of model interpretability and the correlated nature of mortality rates and life expectancies between countries or states. This model's ease of interpretation comes from the fact that it compares each country's life expectancy to a baseline life expectancy measure. Some countries move closer to the baseline, others remain stagnant, and the last group of countries may move away from this baseline in an extreme scenario with negative mortality improvement. In addition, the Double-Gap Life Expectancy model also measures the difference in life expectancy between genders within each country to compare these differences across countries and predict the trend of this "gender gap" over time.

The Double-Gap Life Expectancy model is broken down into three parts: 1) best practice life expectancy, 2) the gap to best practice life expectancy, and 3) the gender gap (Pascariu et al., 2018). In the first modeling step, best practice life expectancy is estimated by regressing record life expectancy (the maximum life expectancy each year across the regions being compared) to the calendar year via a simple linear regression model. This step recognizes the idea that there is a "competition" between regions to improve their population's life expectancy over time, so regions trade off as the "recordholder", which explains the strong positive correlation between regions' life expectancies over time (Pascariu et al., 2018). Although this trend is linear in Pascariu et al. (2018), the trend of life expectancy for a different group of regions does not have to be linear. In step two, the gap between the best practice life expectancy and female life expectancy in each individual region is modeled via an ARIMA model, which the Lee-Carter model also utilizes in its model-fitting process. In the last step (which is optional if the forecast is only concerned with female life expectancy), the gender gap (the gap between female and male life expectancy) is modeled using a piecewise model (i.e., a two-period autoregressive (AR) model combined with a linear predictor, or a random walk ARIMA model), with the model's functional form depending on a limiting age to prevent extrapolation regarding the gender gap (Pascariu et al., 2018). However, the functional form of the gender gap in Pascariu et al. (2018) is highly specific to a dataset containing 38 countries (mostly in the EU), so for other populations, the gender gap may need to be modeled differently. Overall, the idea of the Double-Gap Life Expectancy model allows for flexibility of each model component to best fit the trends of a given dataset's comparative life expectancies.

No study has been conducted to analyze statewide differences in mortality experience or life expectancy (Li & Hyndman, 2021), so the rest of this paper focuses on applying a modified version of the Double-Gap Life Expectancy model to a dataset of the 48 mainland U.S. states to forecast life expectancy in the six most populous states for the next decade. The following chapter briefly discusses the time series datasets used to fit the life expectancy model. Chapter 3 begins by discussing data quality issues and assumptions and then transitions into fitting the three model components, which include model equations and forecasts. Chapter 4 hypothesizes external factors for differences in life expectancy trends and forecasts between the six representative states. Finally, Chapter 5 concludes with final thoughts, limitations of this life expectancy model, and areas for future research.

Chapter 2

Data Description

The datasets used to adapt the Double-Gap Life Expectancy model at the state level come from the U.S. Mortality Database (USMDB)[1]. For the 48 mainland states, gender-specific life tables that include mortality rates and life expectancy for ages between 0 and 110 and for each year from 1941 to 2020 are available. However, only life expectancy for age 0 (newborns) and age 65 (retirees) for each gender and for years between 1980 and 2019 are used in training an adapted Double-Gap Life Expectancy model. For the years 2010 to 2019, mortality data are preliminary estimates until final and complete 2020 Census data is released.

To model best practice life expectancy, one-by-one age-interval, period-year-interval historical life expectancies for the 48 mainland states' female populations are used. To model the gap to best practice life expectancy, historical life expectancies for the six most populous states' (California, Texas, Florida, New York, Pennsylvania, and Illinois) female populations are used. To model the gender gap, historical life expectancies for the same six most populous states' male populations are used.

California, Texas, Florida, New York, Pennsylvania, and Illinois are chosen as the six representative states to be modeled for a few reasons. These states are the six largest in terms of 2023 population, so trends in these states may also be experienced across the entire U.S. In addition, these states have geographic diversity as they cover four distinct regions in the U.S. (the Northeast, the South, the Midwest, and the West). Therefore, not only do these states

[1]See https://usa.mortality.org/ for U.S. state-level, gender-specific life expectancy data

represent the entire U.S. in terms of proportion of population, they also individually represent a particular region to model life expectancy at a more granular level.

Chapter 3

Double-Gap Life Expectancy Model

Best Practice Life Expectancy

Like in Pascariu et al. (2018), the best practice life expectancy for both female newborns and female retirees is fitted using simple linear regression with calendar year as the sole predictor. This model form can be summarized in Equation 1 (Pascariu et al):

Equation 1: Best Practice Life Expectancy Model

$$e_{x,t}^{bp} = \alpha_{x\,0} + \alpha_{x\,1} * t,$$

where $e_{x,t}^{bp}$, is best practice life expectancy, x is the age of the population being modeled, t is the calendar year, $\alpha_{x\,0}$ is the intercept coefficient, and $\alpha_{x\,1}$ is the slope coefficient. Note that an error term is excluded since this model equation represents a forecast, and errors are assumed to be normally distributed with a mean of zero and a constant variance of σ^2.

Before proceeding with training this model component, data quality and linear assumption checks were performed. The first data quality check performed was determining which state held the record life expectancy each year. Theoretically, the recordholder state should change as states compete to improve their female population life expectancy. The recordholder switched between six different states (Iowa, Minnesota, North Dakota, South Dakota, Connecticut, and California) for female newborns and between eight different states (South Dakota, North Dakota, Arizona, Minnesota, Connecticut, California, Florida, and New York) for female retirees.

The next data quality check performed was considering an appropriate window of historical record life expectancy data to train the linear model. Although the USMDB datasets contain life expectancy data from the years 1941 to 2020, only the years 1980 to 2019 were used to train the model. Narrowing the training set to a smaller window is for two reasons: 1) as demonstrated by multiple sources in Chapter 1, the trend of life expectancy and mortality improvement has slowed but has been relatively consistent since the 1980s, and 2) life expectancy dropped significantly in 2020 due to hundreds of thousands of deaths in the general population during the early months of the COVID-19 pandemic. The second point can be illustrated in Figures 4 and 5, as 2020 life expectancy for both female newborns and female retirees is an influential point.

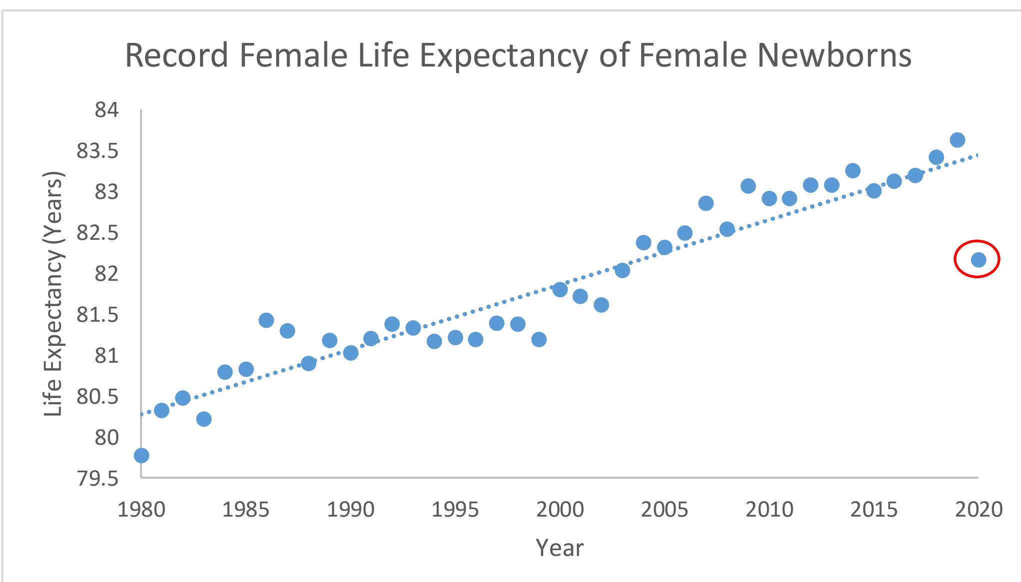

Figure 4: Record Female Life Expectancy of Female Newborns

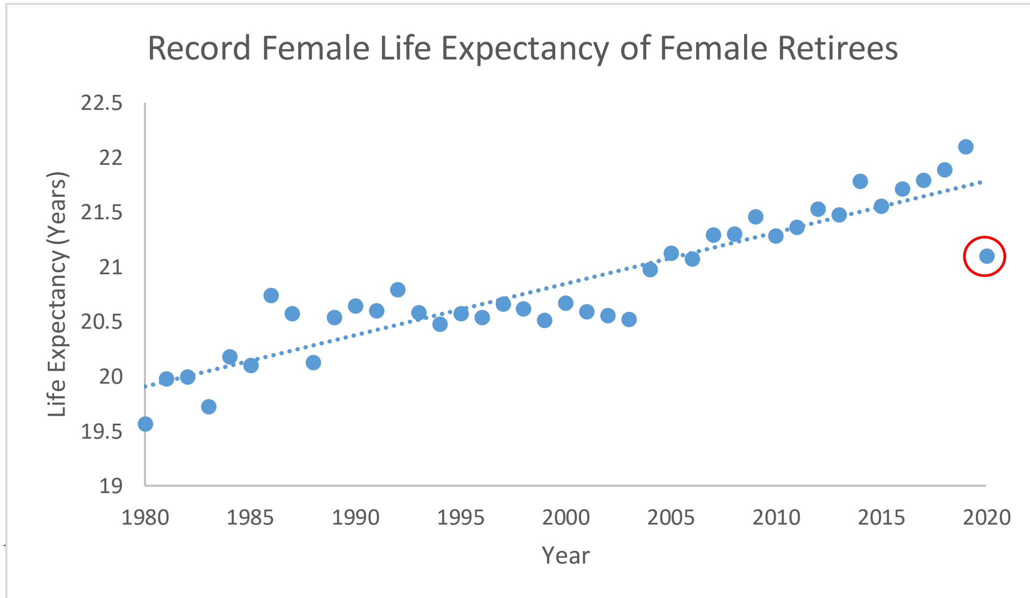

Figure 5: Record Female Life Expectancy of Female Retirees

After fitting linear models for both female newborns and female retirees and verifying the model's assumptions (independence of errors, normally distributed errors, and equal variance), best practice life expectancy was estimated for both populations from 2020-2029. The coefficients and other model summary output for best practice life expectancy are shown below in Table 2, and forecasts with prediction limits are shown below in Figures 6 and 7.

Table 2: Best Practice Life Expectancy Model Coefficients and Summary Output

Age	α_0	α_1	RSE	R^2
0	-86.2923	0.0841	0.2715	0.9308
65	-78.3086	0.0496	0.2204	0.8766

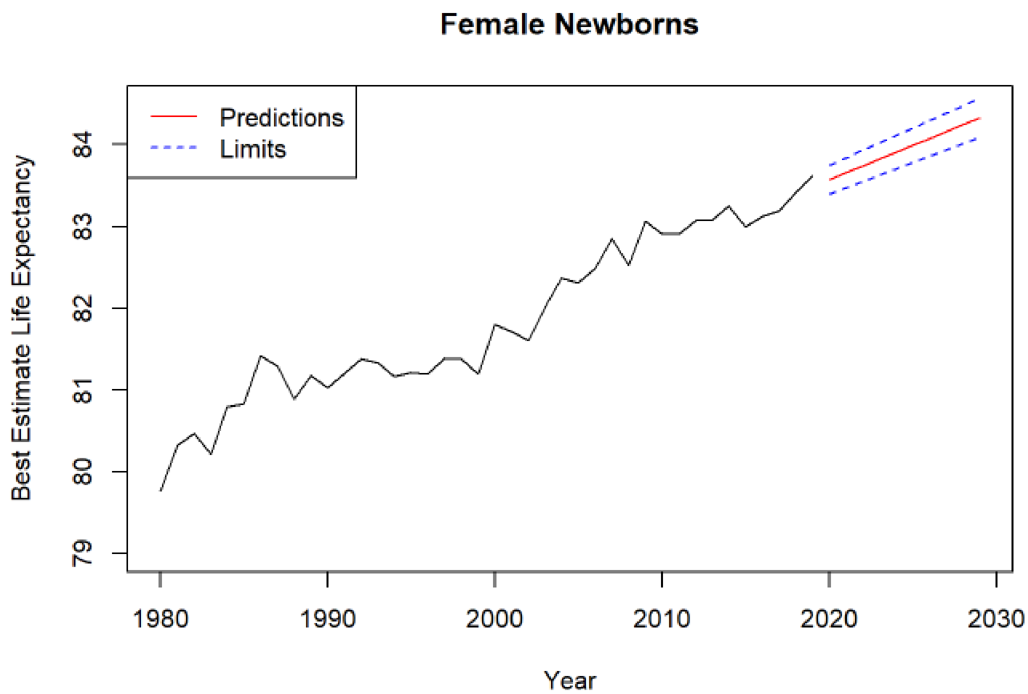

Figure 6: Best Practice Life Expectancy Forecast for Female Newborns

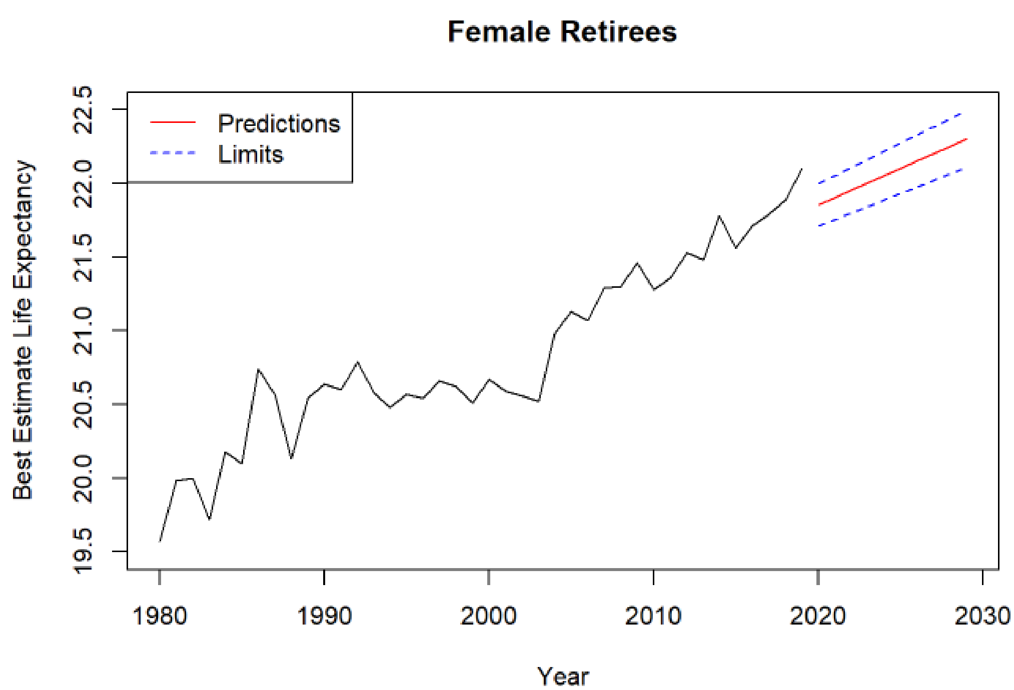

Figure 7: Best Practice Life Expectancy Forecast for Female Retirees

As shown by the forecasts and R^2 values for both female populations, best practice life expectancy predicts a very strong positive, linear trend to continue over time. However, for female newborns, annual mortality improvement (α_1) is approximately double that for female retirees. This is expected as mortality improvement for retirees will not increase as much as newborns due to a likely maximum upper bound of life expectancy at age 120 and the lack of treatments for aging and neurogenerative diseases.

Gap to Best Practice Life Expectancy

Next, the gap to best practice life expectancy for the six selected states is modeled using an ARIMA time series model like in Pascariu et al. (2018). Once the gap is calculated, it is subtracted from best practice life expectancy to forecast the trend of female life expectancy (for both newborns and retirees) for each state. The model for forecasting female life expectancy and the ARIMA(p,d,q) are both shown below in Equations 2 and 3 (Pascariu et al.).

Equation 2: Forecasting Female Life Expectancy

$$e^{f}_{k,x,t} = e^{bp}_{x,t} - D_{k,x,t}$$

, where $e^{f}_{k,x,t}$ is the state-specific forecasted female life expectancy, $D_{k,x,t}$ is the state-specific gap to best practice life expectancy, and k is the state being modeled.

Equation 3: ARIMA(p,d,q) Time Series Model

$$\nabla^d D_{k,x,t} = \mu_{k,x} + \sum_{i=1}^{p} \phi_i \nabla^d D_{k,x,t-i} + \epsilon^{(1)}_{k,x,t} + \sum_{j=1}^{q} \theta_j \epsilon^{(1)}_{k,x,t-j}$$

, where $D_{k,x,t}$ is the gap to best practice life expectancy, ∇^d indicates differencing to make the model stationary, $\mu_{k,x}$ is the average value of the time series, $\emptyset_i$'s are the autoregressive (AR) coefficients, θ_i's are the moving average (MA) coefficients, $\varepsilon^{(1)}_{k,x,t}$ are the random noise components, p is the number of AR coefficients, q is the number of MA coefficients, and d is the number of times differencing occurs to make the time series stationary.

To determine the level of differencing needed for stationarity, KPSS unit-root tests were used before and after differencing occurred. A significance threshold of 0.05 is chosen as the cutoff point to assess if the non-differenced (or differenced) data contained a unit root. This procedure continues until the KPSS test shows evidence of stationarity or until differencing occurred twice to avoid overfitting the ARIMA model. Thus, no differencing (d is zero), first-order differencing (d is one), or second-order differencing (d is two) can occur for each state.

Once the data was transformed to be stationary via differencing (or not differenced if the data was already stationary), Akaike Information Criterion (AIC) was used to select the number of model parameters (the number of AR coefficients, p, plus the number of MA coefficients, q) for the stationary component. To prevent overfitting, the search for the number of AR coefficients ranged from 0 to 5, and the search number of MA coefficients ranged from 0 to 2. Finally, with the number of p, d, and q parameters for the ARIMA model considered, the female gap to best practice life expectancy for each state was fitted via maximum likelihood estimation (MLE). The chosen model form and estimated coefficients are shown below in Table 3:

Table 3: State-Level Female Gaps to Best Practice Life Expectancy Models

State	Age	(p,d,q)	ϕ_1	ϕ_2	ϕ_3	ϕ_4	θ_1
CA	0	(0,1,0)	-	-	-	-	-
	65	(0,1,0)	-	-	-	-	-
FL	0	(0,1,0)	-	-	-	-	-
	65	(3,0,0)	0.9099	0.4202	-0.4084	-	-
IL	0	(0,1,0)	-	-	-	-	-
	65	(2,1,1)	-0.9089	-0.4223	-	-	-0.7308
NY	0	(0,1,0)	-	-	-	-	-
	65	(1,1,0)	-0.2700	-	-	-	-
PA	0	(4,1,1)	0.7026	0.1811	-0.0477	-0.2581	-
	65	(0,1,0)	-	-	-	-	-
TX	0	(2,0,1)	1.7497	-0.8238	-	-	1
	65	(3,0,1)	1.8188	-0.8216	-0.0373	-	1

After fitting 12 ARIMA(p,d,q) models (6 for female newborn state populations and 6 for female retiree state populations), similarities and differences across states can be analyzed. For female newborns, the most common model across states is a random walk model (an ARIMA(0,1,0) model). However, for female retirees, there was much more variation regarding the number of AR and MA terms, meaning that the gap to best practice life expectancy is more complex to model for female retirees than for female newborns, and this gap for female retirees behaves differently across states. In addition, all states besides Florida (female retirees only) and

Texas had a differencing parameter of one, meaning that the gap to best practice life expectancy should remain close to the 2019 recorded gap for the 10-year forecast based on the behavior of ARIMA models. Conversely, Florida (female retirees only) and Texas have a differencing parameter of zero, meaning that the gap to best practice life expectancy will move closer to the historical gap to best practice life expectancy for the 10-year forecast based on the behavior of ARMA models (which are a special case of ARIMA(p,0,q) models).

Using the state-level ARIMA and ARMA models from the model-fitting procedure described above and the model coefficients in Table 8, forecasted gaps to best practice life expectancy and historical gaps for each state are displayed below in Figures 8 and 9:

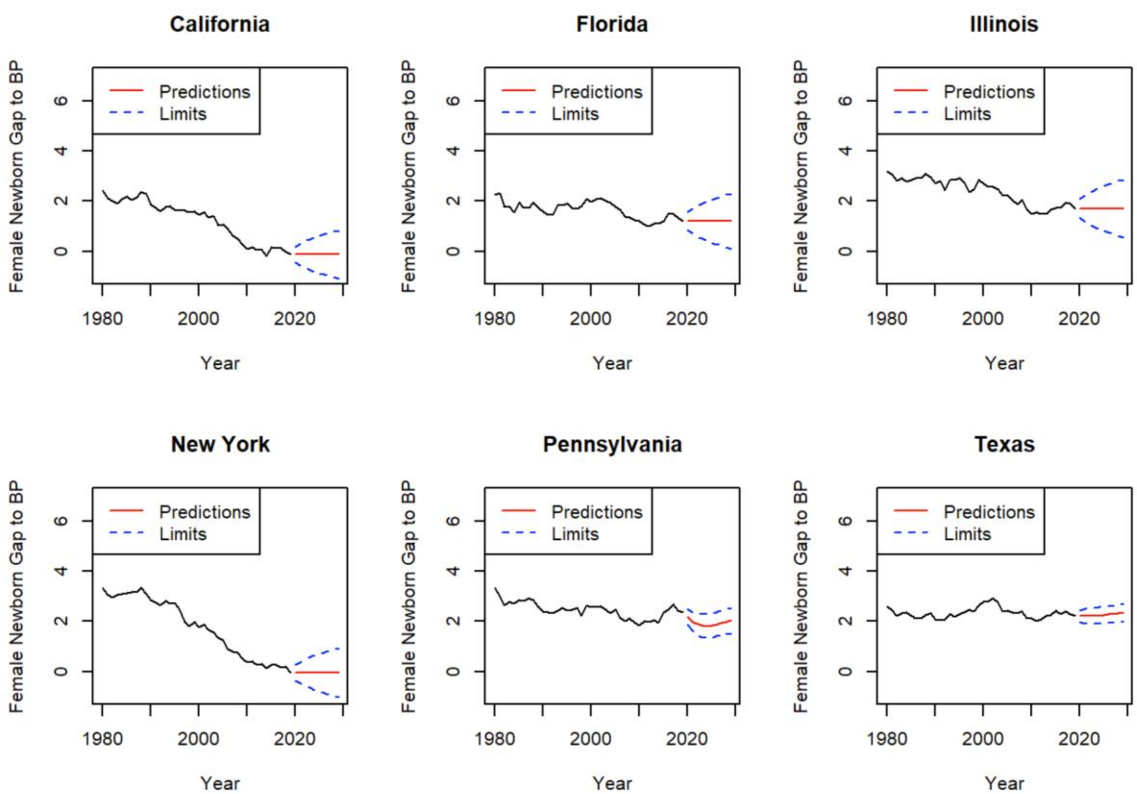

Figure 8: State-Level Female Newborn Gap to Best Practice Life Expectancy Forecasts

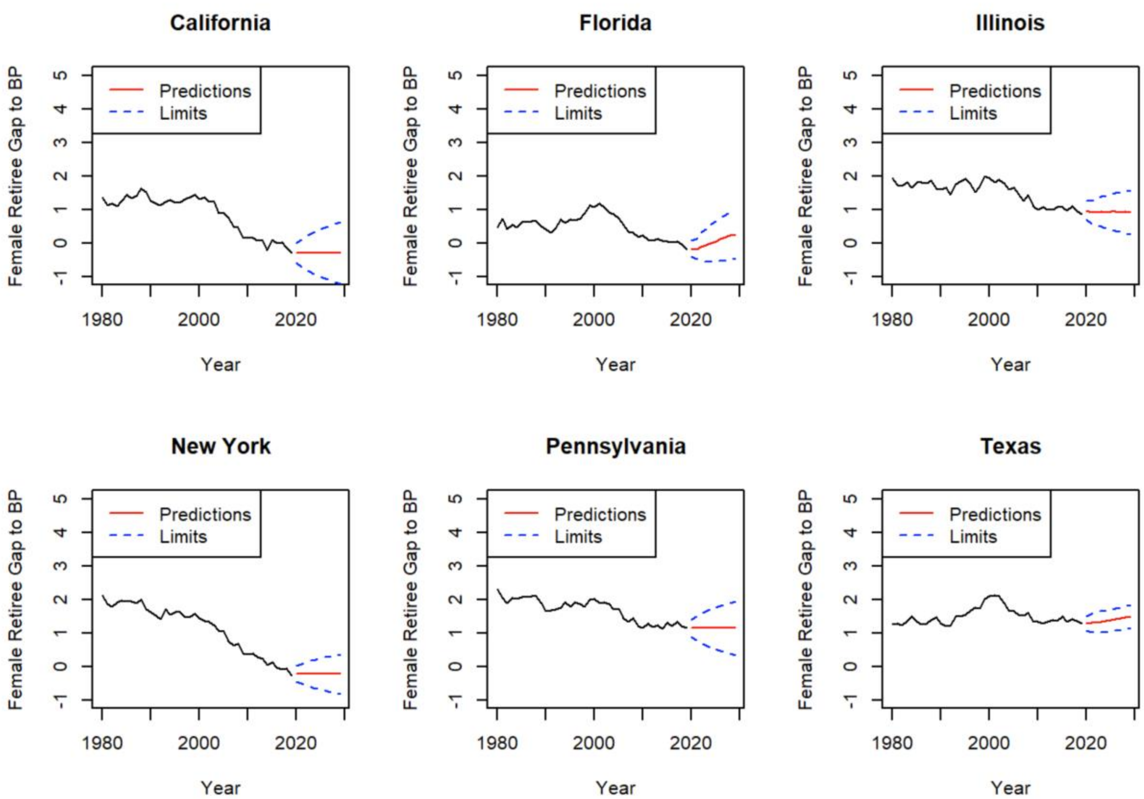

Figure 9: State-Level Female Retiree Gap to Best Practice Life Expectancy Forecasts

Figure 8 shows that the gap to best practice life expectancy for female newborns for each of the six states is relatively stable. The forecast period centers around the historical gap for the last recorded year in 2019. In addition, Figure 8 shows that California and New York are the only two states that have seen their gaps decrease since 1980. This fact is unsurprising as both states have been competing since the 2010s as the recordholder for best practice life expectancy, so their forecasted gaps should approach zero. For the other four states, their gaps remain around two years.

While Figure 8 shows similarities between the gap to best practice life expectancy for female newborns across states, Figure 9 shows differences between the gap to best practice life expectancy for female retirees across states. Notably, both Florida and Texas show an increasing gap over the forecast period. This positive increasing trend can be explained in Table 3 as both gap models are ARMA models, which means the previous trends of a decreasing gap for Florida and a constant gap for Texas will cease to continue into the forecast period due to ARMA forecasts tending to return to their stationary mean values over time. For retirees in New York and California, their gaps decrease and approach zero as these states are also competing to be the recordholder for best practice life expectancy for this age demographic. For the remaining two states of Pennsylvania and Illinois, their gaps remain stable around a value of one year.

Gender Gap

The third and final component of the Double Gap Life Expectancy model is the gap between female and male life expectancy. This component is subtracted from forecasted female life expectancy to predict male life expectancy, and this model form is summarized below in Equation 4 (Pascariu et al., 2018).

Equation 4: Forecasting Male Life Expectancy

$$e^{m}_{k,x,t} = e^{f}_{k,x,t} - G_{k,x,t}$$

, where $e^{m}_{k,x,t}$ is the state-specific forecasted male life expectancy, and $G_{k,x,t}$ is the state-specific gap to sex-based gap in life expectancy,

Unlike in Pascariu et al. (2018), the gender gap is estimated using an ARIMA model. This time series model is used to estimate both the gap to best practice life expectancy and the gender gap since it is simpler and does not assume specific future trends in life expectancy forecasts. In the original Double Gap Life Expectancy model, Pascariu et al. (2018) assume that the gender gap will stop widening and then will eventually narrow until the gap is constant. However, they use a forecast period that is over three decades, and in this paper, the forecast period is only a decade, so it would be less likely to observe convergence and narrowing of the gender gap in such a short forecast period.

The same model-fitting process to estimate the state-level gaps to best practice life expectancy was applied to estimate state-level gender gaps. KPSS unit-root tests were used to determine stationarity by assessing if non-differenced (differenced) data contained a unit root. After differencing (or not differencing) the data, AIC was used to select the number of model parameters, and the search ranged from 0 to 5 for AR coefficients and from 0 to 2 for MA coefficients. The only exception to this was for the male newborn gender gap for Texas as the stationary ARMA component of the ARIMA model was estimated to be of the form ARMA(5,2), which meant the search for number of parameters needed to be widened.

Once the number of p, d, and q parameters for the ARIMA model were found, the gender gap for each state was fitted via (MLE). The chosen model form and estimated coefficients are shown below in Table 4:

Table 4: State-Level Gender Gap Models

State	Age	(p,d,q)	$\emptyset_1$	$\emptyset_2$	$\emptyset_3$	θ_1	θ_2	θ_3	θ_4
CA	0	(1,1,0)	0.3953	-	-	-	-	-	-
	65	(0,2,1)	-	-	-	0.8891	-	-	-
FL	0	(1,1,0)	0.2549	-	-	-	-	-	-
	65	(0,1,0)	-	-	-	-	-	-	-
IL	0	(0,1,0)	-	-	-	-	-	-	-
	65	(2,1,1)	0.3108	0.4733	-	0.6321	-	-	-
NY	0	(1,1,0)	0.5954	-	-	-	-	-	-
	65	(2,1,2)	-0.0281	-0.7718	-	0.2771	-1	-	-
PA	0	(0,1,0)	-	-	-	-	-	-	-
	65	(3,1,0)	-0.4171	0.1906	0.3636	-	-	-	-
TX	0	(3,1,4)	0.2762	-0.3942	0.2898	0.0716	-0.7236	0.2118	-0.8642
	65	(2,1,0)	-0.1284	0.3158	-	-	-	-	-

Compared to the ARIMA model estimation for the state-level gaps to best practice life expectancy, there appears to be more variation in model form for state-level gender gaps. Most of the best practice life expectancy models were random walks, but only three of the gender gap models are random walks. This means that the gaps between female and male life expectancy are more complex to model than the female gap to best practice life expectancy since the gender gap ARIMA models have more than one model parameter to be estimated.

Another notable observation is that every gender gap ARIMA model has a differencing factor of at least one (California was the only state to have a differencing factor of two) while some best practice life expectancy ARIMA models had no differencing at all. Thus, 10-year forecasts for gender gaps should be relatively close to their 2019 historical values, and only California's will increase since a differencing factor of two implies linear behavior in forecasts. The gender gap forecasting behavior for the six states is shown below in Figures 10 and 11.

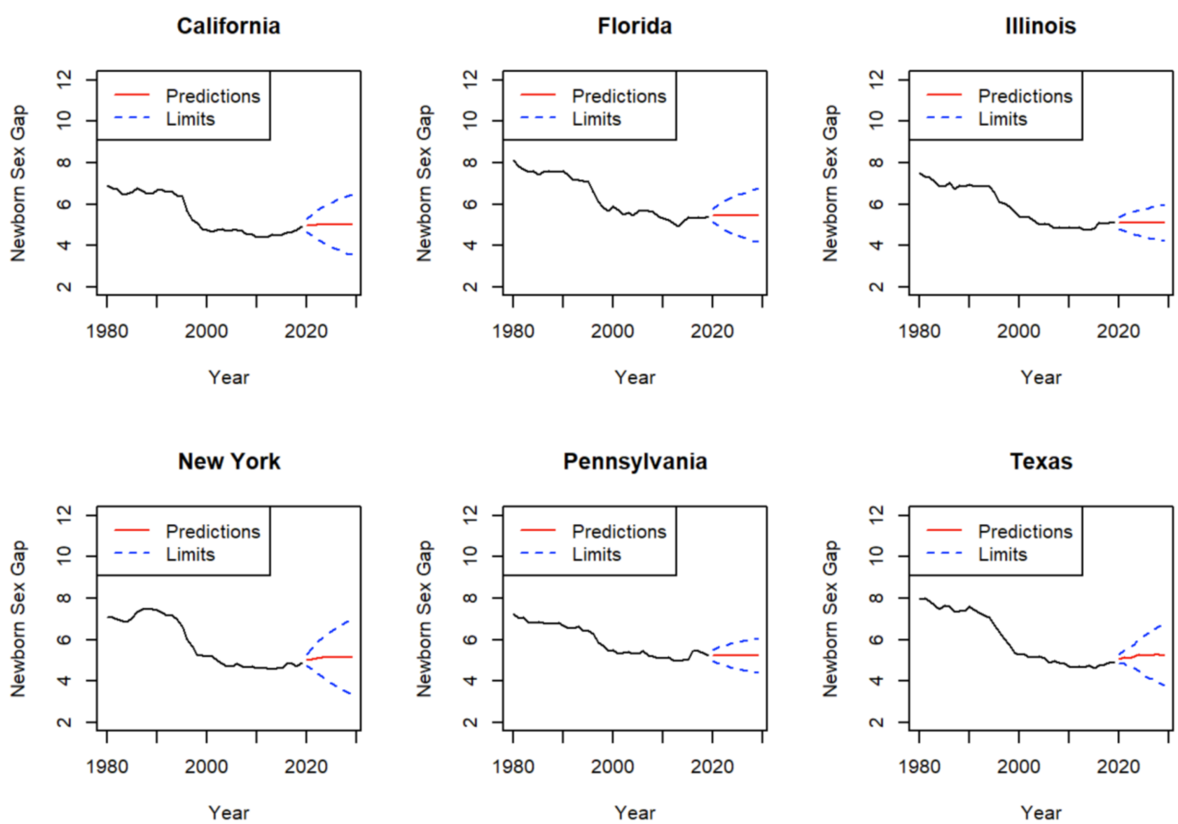

Figure 10: State-Level Male Newborn Gender Gap Forecasts

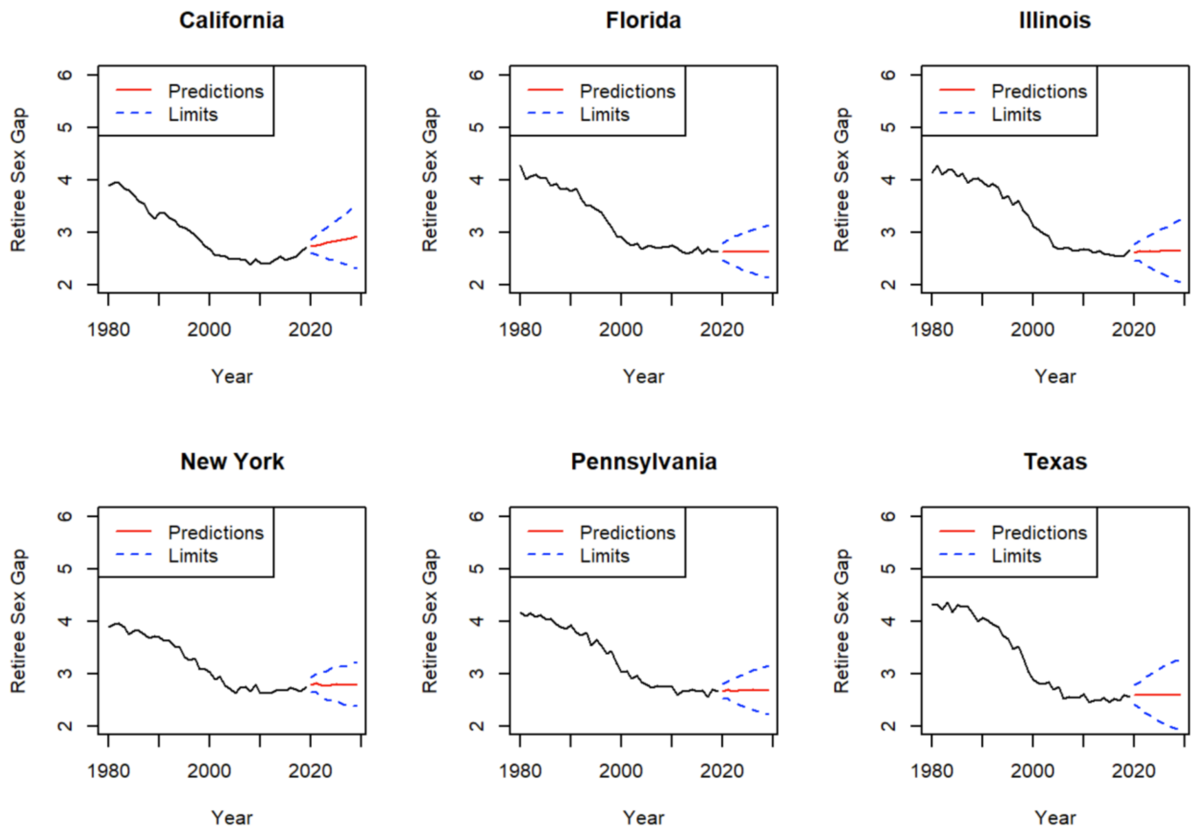

Figure 11: State-Level Male Retiree Gender Gap Forecasts

Figure 10 shows that the newborn gender gap does not show much variation across the six states. Every state since 1980 has seen their respective newborn gender gaps decrease from around 8 years to now around 5 years, on average. The biggest decrease for each state appears to have occurred around the year 2000, but not much progress has been made to narrow the newborn gender gap since, so the ARIMA model forecasts expects the newborn gender gap to remain constant between 4 to 6 years depending on the state besides some fluctuation for a few years early in the forecast window.

Similarly, Figure 11 projects the retiree gender gap to remain constant at between 2 to 3 years for each state. However, the one exception to this rule is California, as its retiree gender gap shows increasing behavior since 2010, and this increasing trend over time is expected to continue. This increasing behavior is also explained by the fact that the gender gap model for Californian male retirees contains a differencing factor of two.

Drawing conclusions based on Figures 8, 9, 10, and 11, variation in male life expectancy forecasts across states are derived mainly from the gap to best practice life expectancy rather than the gender gap. In other words, gender gaps behave very similarly over time across states while gaps to best practice life expectancy behave very differently over time across states.

Chapter 4

Analyzing Statewide Differences

Comparing Statewide Life Expectancy Forecasts

With models fitted for best practice life expectancy, state-level gaps to best practice, and state-level gender gaps, these forecasts can be combined in a linear fashion via Equations 2 and 4 to predict female and male future life expectancies, respectively, for the chosen six states to allow for state-level comparative analysis. Figures 12 and 13 show the 10-year forecasts for female newborns and female retirees, respectively.

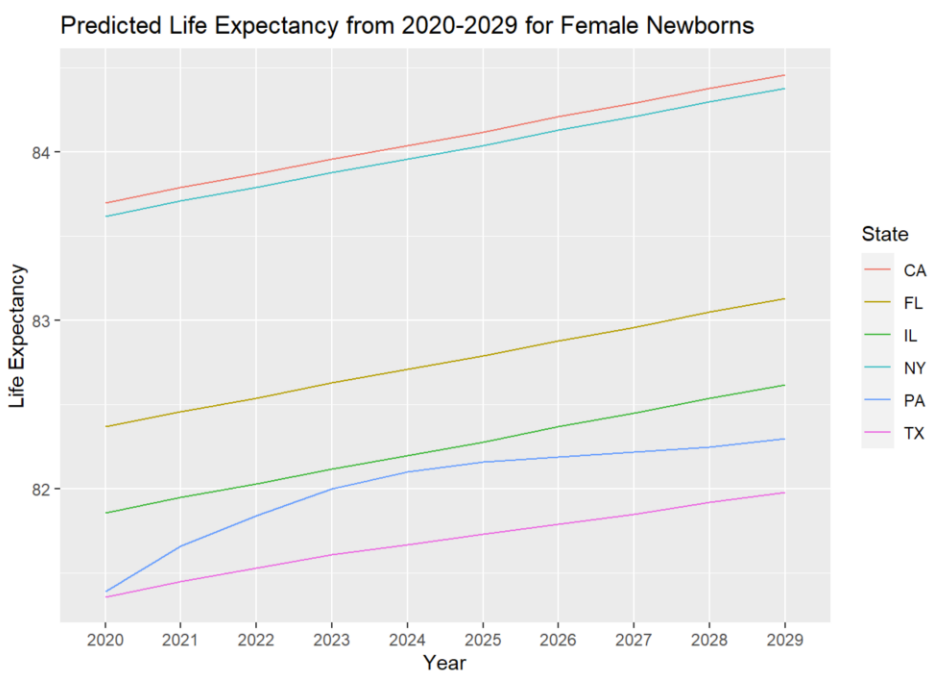

Figure 12: State-Level Forecasts for Female Newborn Life Expectancy

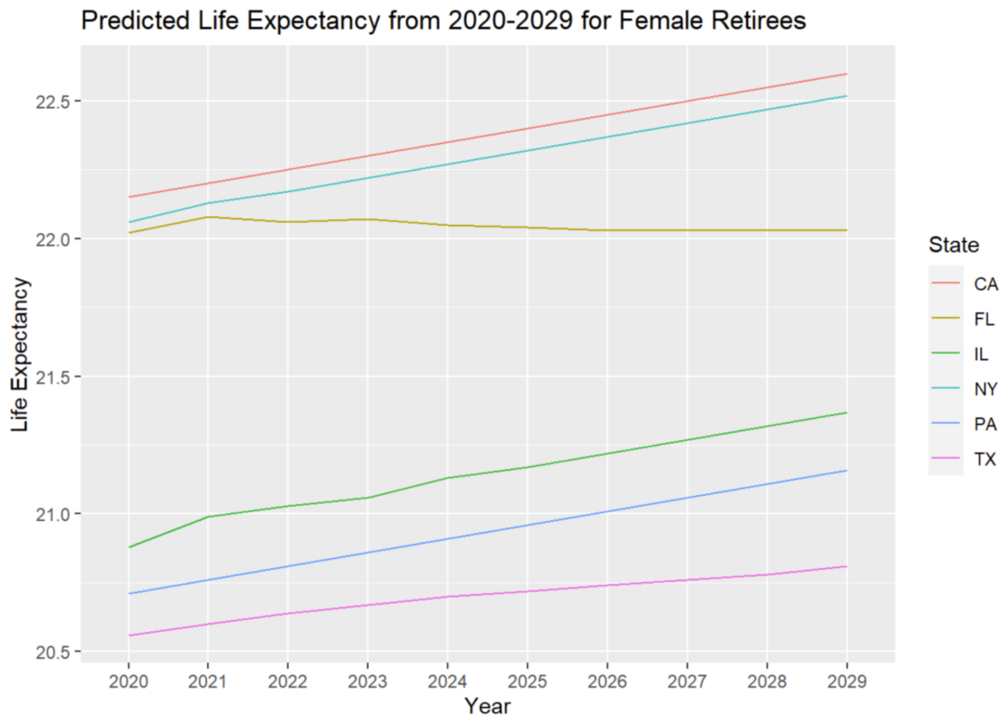

Figure 13: State-Level Forecasts for Female Retiree Life Expectancy

Unsurprisingly, California and New York dominate the other states for both female age demographics since they have both been the most recent recordholders of life expectancy among the 48 mainland states, and the best practice trend forecast expects this recent trend to continue into the next decade. For the female newborn population specifically, the forecast for all states but Pennsylvania takes an approximate linear form since these states' gaps to best practice life expectancy are modeled via a random walk, meaning that these forecasted gaps are constant (i.e., the recorded gap in 2019 is expected to be maintained for the next decade). For Pennsylvania only, life expectancy for female newborns is expected to improve rapidly from 2020 to 2023, but then for every year onward, it will level off and only increase marginally.

For the female retiree population, the behavior for California and New York is similar to that of the female newborn population in terms of maintaining the record life expectancy, but the other four states behave differently. For example, Florida's retiree life expectancy is almost on par with California and New York's in 2021, but it levels off and even begins to decrease slightly. Afterwards, this negative mortality improvement forecast for Florida is explained by the fact that its gap to best practice life expectancy is modeled by an ARMA process, which means that forecasts will tend towards the mean. In other words, Florida's past mortality improvement significantly deviated from its expected average and will be corrected during the forecast period.

For both female populations, the order of forecasted life expectancy from 2020 to 2029 in descending order is California, New York, Florida, Illinois, Pennsylvania, and Texas. This ordering does not change at all during this forecast period, although the gap between certain states (e.g., Texas and Pennsylvania for female newborns in 2020, Florida and New York for female retirees in 2020, Pennsylvania and Illinois for female newborns in 2024, etc.) approach zero in certain prediction windows.

Next, 10-year forecasts for male newborns and male retirees are displayed below in Figures 14 and 15. These forecasts are similar to the female forecasts for their respective age demographics, but they also account for an additional gender-specific gender gap, so there is an additional layer of model complexity in forecasting life expectancy for males compared to females.

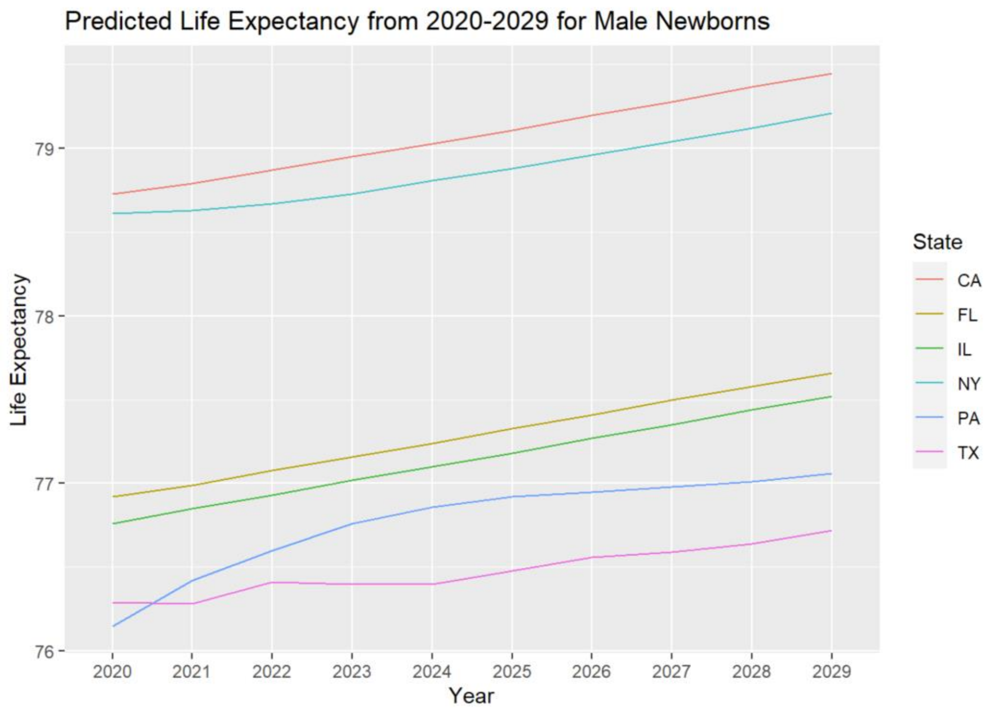

Figure 14: State-Level Forecasts for Male Newborn Life Expectancy

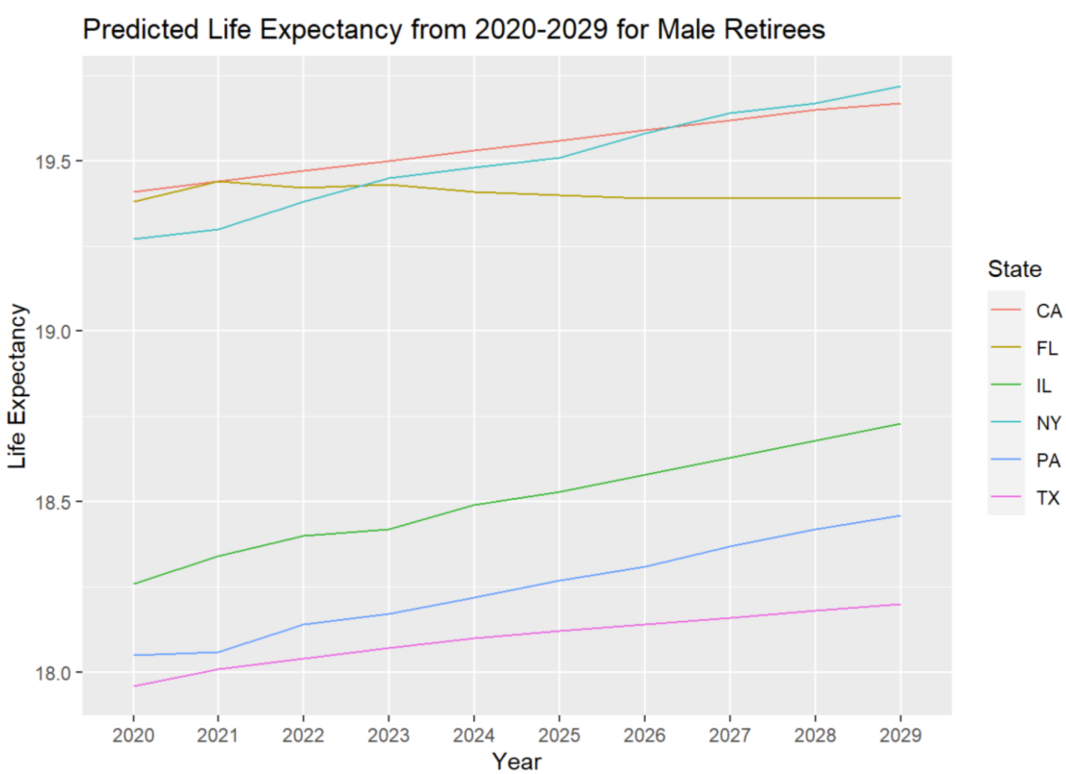

Figure 15: State-Level Forecasts for Male Retiree Life Expectancy

Compared to the female population state-level forecasts, the male population state-level forecasts behave similarly in terms of life expectancy rankings and state-level mortality improvement, but the ranking is subject to change during the forecast period. For male newborns, Pennsylvania is expected to pass Texas in forecasted life expectancy by 2021. For male retirees, the competition to be the recordholder of life expectancy proves to be very dynamic and complex. Florida ties with California as the recordholder by 2021 but then shows negative mortality improvement, leading New York to pass it in the ranking of largest life expectancy by 2025. By 2027, New York is expected to pass California and become the recordholder of male retiree life expectancy. The dynamic nature of the competition for recordholder of male life expectancy for each age demographic seems to derive from the fact that the gender gap is modeled more complexly than a random walk (i.e., forecasting differenced life expectancy involves regression on both past life expectancy differences and past random noise) and some states (e.g., California) having a differencing factor of two.

Drivers of Life Expectancy Differences across States

After discussing life expectancy forecasting differences across the six states, the next step is to understand the drivers of these differences. As referenced in Chapter 1, some of these drivers include political party control and state legislation, urbanization, poverty rate, ethnic and racial composition, quality of healthcare, and health insurance coverage rate. In Table 5, these six predictors of life expectancy are measured for each of the representative states. Comparative state rankings for each metric are shown in parentheses.

Table 5: State Rankings of Life Expectancy Metrics

State	2022 Male Newborn Life Expectancy	Partisan Lean	Urbanization Index	Poverty Rate	Racial and Diversity Index	Healthcare Ranking	Percent of Population Uninsured
CA	78.87 *(1)*	D+24 *(1)*	12.19 *(2)*	12.2% *(4)*	69.7% *(1)*	6 *(1)*	6.5% *(4)*
FL	77.08 *(3)*	R+5 *(5)*	11.46 *(4)*	12.7% *(3)*	64.1% *(4)*	27 *(4)*	11.2% *(2)*
IL	76.93 *(4)*	D+13 *(3)*	11.62 *(3)*	11.9% *(5)*	60.3% *(5)*	28 *(5)*	6.6% *(3)*
NY	78.67 *(2)*	D+22 *(2)*	12.56 *(1)*	14.3% *(1)*	65.8% *(3)*	9 *(2)*	4.9% *(6)*
PA	76.60 *(5)*	R+1 *(4)*	11.15 *(6)*	11.8% *(6)*	44% *(6)*	10 *(3)*	5.3% *(5)*
TX	76.40 *(6)*	R+17 *(6)*	11.17 *(5)*	14.0% *(2)*	67% *(2)*	32 *(6)*	16.6% *(1)*

FiveThirtyEight's Partisan lean measures "the average difference between how a state votes and how the country votes overall" (Rakich, 2020). A political lean of "D" indicates that a state votes more Democratic than the U.S. overall while a political lean of "R" indicates that a state votes more Republican than the U.S. overall. The "+" indicates by how many percentage points the state's voters deviate from the national average. As proposed by Montez et al. (2020), it is expected a state with more liberal legislation should have a higher life expectancy than a state with more conservative legislation, so a valid hypothesis is that the higher a state's political lean skews to democratic, the higher its life expectancy, on average. This hypothesis is justified in Table 5 as New York and California are among the most Democratic states and have the highest life expectancy forecasts while Pennsylvania and Texas are among the more Republican states and have the lowest life expectancy forecasts. However, Florida proves to be an exception to this trend as it is more conservative than Pennsylvania and Illinois but has a higher life expectancy forecast than both states.

FiveThirtyEight also has another metric called urbanization index, which is calculated as the weighted average of the natural log of average number of people living within 5-miles of a census tract, for each census tract within a state (Rakich, 2020). The higher the urbanization index, the more concentrated (urban) a state's population density is. As discussed by Rhubart & Santos (2023), there is a growing urban-rural gap in life expectancy that favors more urban areas and punishes more rural areas, so it is expected that more urban states should have higher life expectancies than more rural states. In Table 5, urbanization index almost aligns perfectly with the ranking of life expectancy forecasts besides one-rank swaps in position.

The poverty rate measures the percent of the population living under the federally defined poverty line, and for this paper, 2022 state poverty rates are used (DePietro, 2023). Chetty et al. (2016) claim that on average, wealthier individuals live longer than poorer individuals, but at the state level, it cannot be claimed that states with higher poverty rates should have lower life expectancies than states with lower poverty rates due to extrapolating individual statistics to statewide statistics. State poverty rates and life expectancy forecasts do not show any sort of positive or negative association in Table 5 based on comparing life expectancy forecast rankings to state poverty rate rankings.

The racial and diversity index comes from the 2020 Census data and measures how likely, on average, two randomly chosen individuals from the same states are of different racial/ethnic backgrounds (“Exploring Age Groups in the 2020 Census”, 2023). Rhubart & Santos (2023) note that the variation in life expectancy across states can be explained by the intersection of race/ethnicity and geographic location, and the rural penalty is stronger in certain regions compared to others. Thus, there should be an association between racial/ethnic diversity and life expectancy across states. Rankings in Table 5 for diversity index and life expectancy forecasts appear to line up for the most part except for Texas, which has a much lower life expectancy than its diversity index predicts. This most likely can be explained by it being the most conservative state in the group of six representative states.

State healthcare differences are measured by the U.S. News & World Report 2022 Healthcare Rankings, which combine three rankings (healthcare quality, healthcare access, and public health) into one composite ranking (“Healthcare”, 2023). The higher a state’s ranking (i.e., closer to one), the better its healthcare is rated. Thus, there should be a positive association between healthcare quality and forecasted life expectancy, as a higher ranking indicates higher healthcare, which implies individuals in that state should live longer, on average. As expected, California and New York are among the highest in both healthcare ranking and forecasted life expectancy while Texas is the lowest in both. However, Pennsylvania is an exception to this phenomenon as it has the third highest healthcare rating among the representative states but the second lowest life expectancy forecast. Thus, other factors besides healthcare differences must explain why Pennsylvania has a comparatively lower life expectancy.

The last life expectancy predictor investigated is the percentage of the population uninsured. This metric also uses 2022 estimates from the Census Bureau ("Percentage of Population Without Health Insurance Coverage by State: 2021 and 2022", 2023). States with higher uninsured rates should have lower life expectancy than states with lower uninsured rates because individuals residing in states with higher uninsured rates cannot afford or access all their necessary healthcare (e.g., preventative medicine, prescriptions, etc.). Therefore, the association between the uninsured rate and forecasted life expectancy should be negative. This is mostly the case in Table 5 but with some rankings not perfectly aligning.

In addition to Table 5, Pearson correlation was calculated for each life expectancy metric, which are shown in descending order of their correlation's absolute value in Table 6. However, this correlation measure may be biased as only 6 states out of all 50 are included in this calculation since this paper only considers the 6 most populous states as the representative states for which life expectancy forecasts were fitted.

Table 6: Correlations of Life Expectancy Predictions and Metrics

Life Expectancy Metric	**Correlation**
Urbanization Index	0.9490
Political Lean	0.8665
Healthcare Ranking	0.7192
Racial and Diversity Index	0.5308
Percent of Population Uninsured	-0.5027
Poverty Rate	0.0180

The hypotheses related to each metric as discussed in the preceding paragraphs and measured in Table 5 via comparative rankings are supported via the correlations in Table 6. Urbanization index and political lean both demonstrate very strong positive correlations with life expectancy forecasts, healthcare ranking demonstrates moderately strong positive correlation with life expectancy forecasts, and uninsured rate demonstrates moderately negative correlation with life expectancy forecasts. Thus, the direction of the association between life expectancy forecasts and life expectancy predictors is as expected. In addition, the poverty index shows a very weak correlation with forecasted life expectancy. This may be due to the percentage of impoverished individuals within each state being relatively similar.

Chapter 5

Conclusion

Summary

Life expectancy is of key importance for both public policymakers and insurance companies. For the former, life expectancy is indicative of the general health and well-being of their constituent populations. For the latter, the trend of future life expectancy is an important assumption in pricing and reserving. Public policymakers and insurers both can impact life expectancy through different mechanisms (legislation for the former and incentives for the latter).

In the U.S., the trend of life expectancy has generally increased over time. However, this mortality improvement began to decline in the 1980s and has stagnated in the 2010s, mainly due to the opioid epidemic and the COVID-19 pandemic. Mortality improvement is not linear and comes in waves due to healthcare advancements. This increase is not uniform, though, and varies by geographic location, gender, income level, and race/ethnicity.

Time series applications to mortality and life expectancy models are relatively new. The Lee-Carter model is the most widely used life expectancy forecasting method still, but it is difficult to interpret and does not allow for population comparison. The Double-Gap Life Expectancy model solves both issues by establishing a linear baseline for the annual recordholder population of life expectancy and then comparing each population to this baseline using an ARIMA model.

This paper applies a modified version of the Double-Gap Life Expectancy model to the six largest U.S. states by population to forecast life expectancy for both newborn and retiree subpopulations, split by gender, for the next decade. The model consists of three parts: 1) best practice life expectancy (modeled via simple linear regression), 2) the gap to best estimate life expectancy (estimated by an ARIMA time series model), and the life expectancy gender gap (also estimated by an ARIMA time series model).

Once the 24 models for each combination of age, gender, and state were constructed, an analysis of the 10-year forecasts between each state was conducted. California and New York dominated the other states due to being the most recent recordholders of life expectancy. Besides Florida retirees, all states' subpopulations demonstrated a largely linear increasing trend of future life expectancy. However, the ranking of states in life expectancy did not stay constant during the entire forecast period as some states passed each other (e.g., Pennsylvania overtaking Texas in male newborn life expectancy and New York overtaking both California and Florida in male retiree life expectancy).

After building these state-level life expectancy forecast models, a correlation-based analysis was conducted to hypothesize reasons for forecasting differences. Urbanization levels, state politics, and healthcare quality appear to be strongly correlated to the state-level life expectancy forecasts. In addition, racial and ethnic diversity and the percent of a state's population without health insurance appear to be moderately correlated to state-level life expectancy forecasts while the poverty rate shows barely any correlation.

Model Limitations and Areas for Future Research

As previously discussed in Chapters 1 and 2, the Double-Gap Life Expectancy model forecasts life expectancy using period life expectancy. Historical gender-specific, state-level life expectancies from the Census Bureau, which are used as explanatory variables in this paper's model, are calculated from single-year mortality tables and thus do not consider future mortality improvement. Therefore, it is likely that the state-level forecasts of life expectancy will be underestimated. Future research could extend this paper by adding a mortality improvement component to this modified Double-Gap Life Expectancy model.

In addition, the Double-Gap Life Expectancy model does not consider positive or negative shock scenarios that will drastically impact life expectancy across all states. For example, the impact of a groundbreaking medical discovery or a global pandemic are not factored into this paper's life expectancy predictions. While these scenarios are probabilistically low and difficult to predict, future research could try to incorporate extreme changes in mortality improvement into its model.

Another restriction of the Double Gap-Life Expectancy model is that it only considers gender for explaining one of the gaps in life expectancy. However, there is no reason that other demographic variables could not be considered for being modeled via a gap (e.g., race/ethnicity, income, education, etc.). Future research could expand upon the Double-Gap Life Expectancy model to explore these other gaps to quantify the impact of demographic and socioeconomic differences (besides gender) and their correlation with life expectancy.

This paper only considers six representative states rather than every mainland U.S. state. Thus, it is inappropriate to extrapolate this paper's results entirely to individual states not modeled or the entire U.S. Future research could be more robust and replicate the modeling process of this paper and apply it to all 48 mainland states. In addition, this model could be expanded to neighboring countries' states and provinces (e.g., Canada and Mexico) as well.

BIBLIOGRAPHY

Buxton, J. (2023, January 19). *Period and cohort life expectancy explained—Office for National Statistics*. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/methodologies/periodandcohortlifeexpectancyexplained

Cairns, A., Blake, D., & Dowd, K. (2008). *The Cairns-Blake-Dowd Model*.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The Association Between Income and Life Expectancy in the United States, 2001-2014. *JAMA*, *315*(16), 1750–1766. https://doi.org/10.1001/jama.2016.4226

DePietro, A. (2023, November 9). *2023 Poverty Rate By State, Per The Latest Census Data*. Forbes. https://www.forbes.com/sites/andrewdepietro/2023/11/09/2023-poverty-rate-by-state-per-the-latest-census-data/

Exploring Age Groups in the 2020 Census. (2023, May 23). Census.Gov. https://www.census.gov/library/visualizations/interactive/exploring-age-groups-in-the-2020-census.html

Lee, R. D., & Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association*, *87*(419), 659–671. https://doi.org/10.1080/01621459.1992.10475265

Li, H., & Hyndman, R. J. (2021). Assessing mortality inequality in the U.S.: What can be said about the future? *Insurance: Mathematics and Economics*, *99*, 152–162. https://doi.org/10.1016/j.insmatheco.2021.03.014

Montez, J. K., Beckfield, J., Cooney, J. K., Grumbach, J. M., Hayward, M. D., Koytak, H. Z., Woolf, S. H., & Zajacova, A. (2020). US State Policies, Politics, and Life Expectancy.

The Milbank Quarterly, *98*(3), 668–699. https://doi.org/10.1111/1468-0009.12469

Pascariu, M. D., Canudas-Romo, V., & Vaupel, J. W. (2018). The double-gap life expectancy forecasting model. *Insurance: Mathematics and Economics*, *78*, 339–350. https://doi.org/10.1016/j.insmatheco.2017.09.011

Patkee, P., & Strange, A. (2023, May 22). *The future of life expectancy: Forecasting long-term mortality improvement trends for insurance | Swiss Re*. https://www.swissre.com/institute/research/topics-and-risk-dialogues/health-and-longevity/future-life-expectancy-long-term-mortality-improvement-insurance.html

Percentage of Population Without Health Insurance Coverage by State: 2021 and 2022. (2023, September 14). Census.Gov. https://www.census.gov/library/visualizations/interactive/percentage-without-health-insurance-coverage-by-state-2021-2022.html

Rakich, N. (2020, April 14). How Urban Or Rural Is Your State? And What Does That Mean For The 2020 Election? *FiveThirtyEight*. https://fivethirtyeight.com/features/how-urban-or-rural-is-your-state-and-what-does-that-mean-for-the-2020-election/

Reynolds, C., & Aatman, D. (2023, March 31). *What if mortality stops improving?* https://www.milliman.com/en/insight/what-if-mortality-stops-improving

Rhubart, D., & Santos, A. (2023). Research Note Showing That the Rural Mortality Penalty Varies by Region, Race, and Ethnicity in the United States, 1999–2016. *Demography*, *60*(6), 1699–1709. https://doi.org/10.1215/00703370-11078239

Serykh, N., & Yang, A. (2019). *Analyzing Individual Life Insurance Mortality Trends—Cause of Death Impact | Munich Re Life US*. https://www.munichre.com/us-life/en/perspectives/mortality-studies/analyzing-ind-life-ins-mortality-trends-cause-of-death.html

Singh, G. K., & Siahpush, M. (2014). Widening Rural–Urban Disparities in Life Expectancy, U.S., 1969–2009. *American Journal of Preventive Medicine*, *46*(2), e19–e29. https://doi.org/10.1016/j.amepre.2013.10.017

These U.S. States Have the Best Healthcare. (2023). US News & World Report. https://www.usnews.com/news/best-states/rankings/health-care

Woolhandler, S., & Himmelstein, D. U. (2017). The Relationship of Health Insurance and Mortality: Is Lack of Insurance Deadly? *Annals of Internal Medicine*, *167*(6), 424–431. https://doi.org/10.7326/M17-1403

ACADEMIC VITA

Justin Morrison

EDUCATION

The Pennsylvania State University **University Park, PA**
Intended Bachelor of Science in Actuarial Science *August 2020 – May 2024*
Schreyer Honors College, Smeal College of Business
Anticipated Minors: Statistics, Mathematics, and Information Systems Management

AWARDS / DISTINCTIONS

- **Actuarial Exams Passed**: Exam P (Probability), Exam FM (Financial Mathematics), Exam SRM (Statistics for Risk Modeling), Exam FAM (Fundamentals of Actuarial Mathematics), and Exam PA (Predictive Analytics)
- **Undergraduate Scholastic Awards**: President Walker Award, President Sparks Award, Evan Pugh Scholar Award
- **Scholarships**: Recipient of the Schreyer Academic Excellence Scholarship
- **2023 Society of Actuaries Student Research Case Study Challenge**: Semi-Finalist

WORK EXPERIENCE

Prudential Financial – *Actuarial Intern* *June 2023 – August 2023*
- Created an Excel lookup table of accruals for 80 longevity risk transfer deals and coded VBA user-defined functions to reduce monthly auditing from a few days to a few hours
- Developed a Power BI dashboard to visualize quarterly trends in UK mortality, death reporting lag time, and mortality gains/losses for actuarial and finance senior leaders
- Researched and reported on the strategy behind developing a multi-year guaranteed annuity product in relation to Prudential's ideals (customer obsession and risk smart)

Prudential Financial – *Actuarial Intern* *June 2022 – August 2022*
- Tested and documented change requests from underwriters to ensure that updates to Excel-based Factor tools were implemented correctly for group insurance pricing
- Wrote VBA code to automate the importation of 35 plan design assumption tables into SQL databases, reducing user importation time from 12 hours to a few minutes
- Analyzed and presented on how current financial market and macroeconomic conditions affect the life insurance industry and business units within Prudential

LEADERSHIP / ACTIVITIES

Actuarial Science Club – *Director of Corporate Outreach* *January 2023 – December 2024*
- Organized the 2023 Fall Actuarial Science Career Fair, where I collaborated with 30 insurers and consulting companies in hosting information sessions and interviews
- Prepared actuarial students for the Career Fair by hosting a career development night event and scheduling one-on-one resume reviews

Risk Management Department – *Teaching Assistant* *August 2023 – December 2023*
- Held weekly office hours to aid students in learning financial mathematics and long-term actuarial mathematics
- Provided timely grades and feedback on weekly homework to 55 students
- Led review sessions for students to review challenging concepts and problems

TECHNICAL SKILLS

Advanced Skills: Excel, PowerPoint, and Word
Intermediate Skills: Python, R, Access, SQL, Power BI, and JIRA