

Comenius University in Bratislava

Project Bluezones Europe

*Statistical Validation of Buettner's (2016) Assumptions
in European Context*

Stefan Vach

Statistical Analysis II

Lecturer: Ing. Jakub Szabó, PhD.

January 31, 2025

Abstract

This study investigates the social and psychological determinants of life expectancy across selected European countries, focusing on variables identified by David Buettner's Blue Zones research, including social connections, sense of purpose, stress levels, family support, caloric intake, physical activity, and moderate alcohol consumption. Using linear regression analysis, we examine both direct effects and interactions among these predictors, such as gender with family support and alcohol consumption with caloric intake, to assess their combined influence on longevity. Findings reveal distinct patterns that underscore the importance of integrating social, psychological, and lifestyle factors for understanding and promoting long life expectancy. These insights contribute to public health interventions by identifying key areas for enhancing individual and community well-being.

Introduction

This project examines the relationship between key psychological, social, and lifestyle factors and life expectancy across selected European countries. Drawing on the foundational “Blue Zones” research by David Buettner (Buettner and Skemp 2016), this study investigates the influence of social connections, sense of purpose, stress reduction, family support, caloric moderation, spirituality, physical activity, plant-based eating, and moderate alcohol consumption on longevity within a European context.

Understanding these relationships is critical, particularly in the context of Europe’s aging population. By analyzing how specific factors impact life expectancy, this research provides actionable insights that can inform public health strategies and individual choices aimed at fostering longer, healthier lives. Insights gained from this study could guide policymakers and healthcare providers in developing targeted interventions to enhance wellness and resilience among older adults, potentially alleviating healthcare challenges linked to aging.

This study integrates Buettner’s predictors with data from European countries to explore their relevance and impact in a distinct cultural and economic setting. Using robust linear regression analysis, it highlights the most significant predictors of longevity and examines interactions between them. This approach provides a nuanced understanding of how these factors collectively influence health outcomes, moving beyond isolated predictors to a holistic perspective.

By addressing the cross-cultural applicability of Buettner’s predictors, this research fills a gap in existing literature. It aims to inform academic discussions and practical applications in public health, ultimately supporting more effective approaches to improving long-term well-being in diverse European populations.

Literature Review

Dan Buettner's exploration of "Blue Zones" (Buettner and Skemp 2016) emphasizes lifestyle factors such as plant-centered diets, movement, social networks, and life purpose as crucial for longevity, drawing from observations in high-longevity regions like Okinawa and Sardinia. These findings align with (Kitsera et al., n.d.), whose study on longevity in Ukraine highlights consistent physical activity and plant-based diets as critical, suggesting that core lifestyle factors transcend cultural and geographical contexts. Both studies underscore the role of social connections and low-stress environments in promoting long life, linking the Blue Zones framework with regional studies to illustrate the universality of these determinants.

Building on these insights, Herpich et al. (Herpich, Müller-Werdan, and Norman 2022) provide a nuanced understanding of the health benefits of plant-based diets (PBDs). Their narrative review highlights lower mortality rates among vegetarians and vegans compared to individuals whose diets regularly include meat. PBDs are associated with improved metabolic and inflammatory profiles and a reduced incidence of cardiovascular diseases, all contributing to better health and longevity. However, Herpich et al. also point out limitations, such as the lower protein content of PBDs, which may reduce their health-promoting effects in adults over 65 who require higher protein intake. While the specific mechanisms of how PBDs extend life span remain unclear, the modulation of the gut microbiome and protein restriction are posited as key factors. This underscores the importance of PBDs as a multifaceted contributor to health and longevity, complementing Buettner's observations on plant-centered diets in the Blue Zones (Buettner and Skemp 2016).

Giacosa et al. (Giacosa et al. 2016) expand on the dietary determinants of longevity by highlighting the Mediterranean diet's emphasis on moderate wine consumption, abundant plant foods, olive oil, and minimal red meat intake. Their review identifies moderate red wine consumption as a potential contributor to longevity due to bioactive compounds like resveratrol, which exhibit antioxidant, anti-inflammatory, and cardioprotective properties. The "Mediterranean way of drinking," characterized by regular, moderate wine consumption during meals, reduces cardiovascular disease risk and supports healthy aging without significantly increasing cancer risk. These findings align with the broader literature on plant-based diets and longevity, reinforcing the importance of dietary patterns rich in bioactive compounds.

Physical activity also emerges as a critical determinant of longevity. Paffenbarger et al. (Paffenbarger et al. 1986) demonstrate a clear inverse relationship between physical activity levels and all-cause mortality in their longitudinal study of college alumni. Regular moderate to vigorous exercise was associated with significant re-

ductions in mortality risk, highlighting physical activity's protective effects against chronic diseases and its role in healthy aging. These findings support Buettner's observations of consistent movement as a lifestyle factor in the Blue Zones and align with kitsera_importance_nodate's emphasis on physical activity in promoting longevity across diverse contexts (Kitsera et al., n.d.).

Economic and structural factors also emerge as critical in life expectancy research. Monsef and Mehrjardi (Monsef and Mehrjardi 2015) show how economic stability and urbanization positively influence longevity, while Roffia et al. (Roffia, Bucciol, and Hashlamoun 2023) extend this argument, emphasizing the role of accessible healthcare and equitable health expenditures in improving life expectancy across OECD countries. These findings connect to Siegel et al. (Siegel, Schug, and Rieger 2022), who demonstrate that socioeconomic factors, such as education and household income, play a more significant role in longevity than healthcare access in Germany. Together, these studies suggest that while economic and social infrastructure is foundational to health, its impact is mediated by individual and regional disparities in education and resource allocation.

The role of spirituality and social cohesion emerges in studies like Dominguez et al. (Dominguez, Veronese, and Barbagallo 2024), which links spiritual practices to reduced stress and enhanced social support, and Rose (Rose 1964), which highlights the stabilizing influence of family environments and marital bonds on mental and physical health. Both align with Buettner's emphasis on purpose-driven living and faith-based communities in Blue Zones (Buettner and Skemp 2016), reinforcing the interplay between social and psychological well-being as critical predictors of life expectancy.

Interaction; alcohol consumption - caloric intake

Lifestyle choices, such as diet, physical activity, and substance use, further illustrate complex relationships with longevity. Paffenbarger and Lee (Paffenbarger and Lee 1996) demonstrate the protective role of regular physical activity against chronic diseases, complementing Rizza et al.'s (Rizza, Veronese, and Fontana 2014) findings on calorie restriction's benefits for metabolic health. These insights align with (Kitsera et al., n.d.), who identified low-calorie, nutrient-dense diets as central to longevity in Ukraine. However, when intersecting with alcohol consumption, these relationships become nuanced. Cockerham (Cockerham 1997) highlights alcohol abuse's detrimental effects in Eastern Europe, while Rizza et al. (Rizza, Veronese, and Fontana 2014) and (Buettner and Skemp 2016) suggest moderate alcohol consumption may support longevity when combined with balanced dietary habits. This complexity highlights the importance of examining interaction ef-

fects, such as the interplay between caloric intake and alcohol consumption, to understand compounded impacts on health outcomes.

Finally, the findings on family support and social networks in Rose (Rose 1964) and Siegel et al. (Siegel, Schug, and Rieger 2022) raise questions about gendered experiences of family dynamics. Women often serve as primary providers of family support, while men may benefit differently from family structures, a distinction that warrants exploration in studies of family support and longevity. Such interactions could reveal how social roles mediate the impact of family relationships on health outcomes.

By connecting these studies, a more integrated perspective emerges, demonstrating that life expectancy determinants are not isolated factors but interdependent influences shaped by socioeconomic, psychological, and lifestyle variables. This approach aligns with the research gap identified in this study, which seeks to examine the compounded effects of lifestyle and social factors on longevity.

Research Question

What is the relationship between social connections, sense of purpose, stress levels, and family support on life expectancy in selected European countries?

Hypotheses

General Hypothesis: Adherence to lifestyle factors such as a plant-based diet, social connections, sense of purpose, and moderate alcohol consumption will be positively associated with life expectancy, while higher body mass index (BMI) and immoderate alcohol consumption frequency will exhibit negative associations.

Interaction Hypothesis: The interaction between moderate alcohol consumption frequency and BMI will significantly influence life expectancy, such that the negative impact of immoderate alcohol consumption on life expectancy will be attenuated at higher BMI levels.

Methodology

Datasets used

The study utilizes aggregated data from two European Social Survey (ESS) datasets (ESS7 and ESS11) and the 2024 Revision of World Population Prospects for life expectancy. ESS11 includes data from Austria, Switzerland, Germany, Finland, the United Kingdom, Croatia, Hungary, Ireland, Lithuania, the Netherlands, Norway, Slovenia, and Slovakia. ESS7, on the other hand, covers Austria, Belgium, Switzerland, the Czech Republic, Germany, Denmark, Estonia, Spain, Finland, France, the United Kingdom, Hungary, Ireland, Israel, Lithuania, the Netherlands, Norway, Poland, Portugal, Sweden, and Slovenia.

Data transformation

The independent variables were operationalised based on the Blue Zones framework. If indexes were created, mean of the values was used as an index value. Diet was measured through fruit and vegetable consumption (etfruit, eatveg); physical activity through sports participation (dosprt); social connections via frequency of social interactions (sclmeet) and number of confidants (inprpsc); sense of purpose through life satisfaction (stflife) and enjoyment of life (enjlf); moderate alcohol consumption by frequency (alcfreq); stress reduction by depression (fltdpr), effort (flteeff), and restless sleep (slprl); spirituality through religious attendance (rlgatnd) and prayer (pray); family support by hours helping family members (hlpfmhr); and caloric intake using Body Mass Index (BMI). The dependent variable, life expectancy, was sourced from the UN dataset for its reliability and cross-national consistency. Dataset was an aggregation of means of the index values per country.

Model Description

The statistical analysis employed a multinomial linear regression model, which is well-suited for examining the direct associations between life expectancy and a range of predictors, aligning with the approach used by (Rizza, Veronese, and Fontana 2014) to analyze interaction effects and direct influences. The model was fitted using the `lm` function from base R, and robust standard errors were computed using the `vcovHC` function from the `sandwich` package with the HC1 adjustment, which is particularly suited for small sample sizes. This approach improves the reliability of standard error estimates without requiring additional

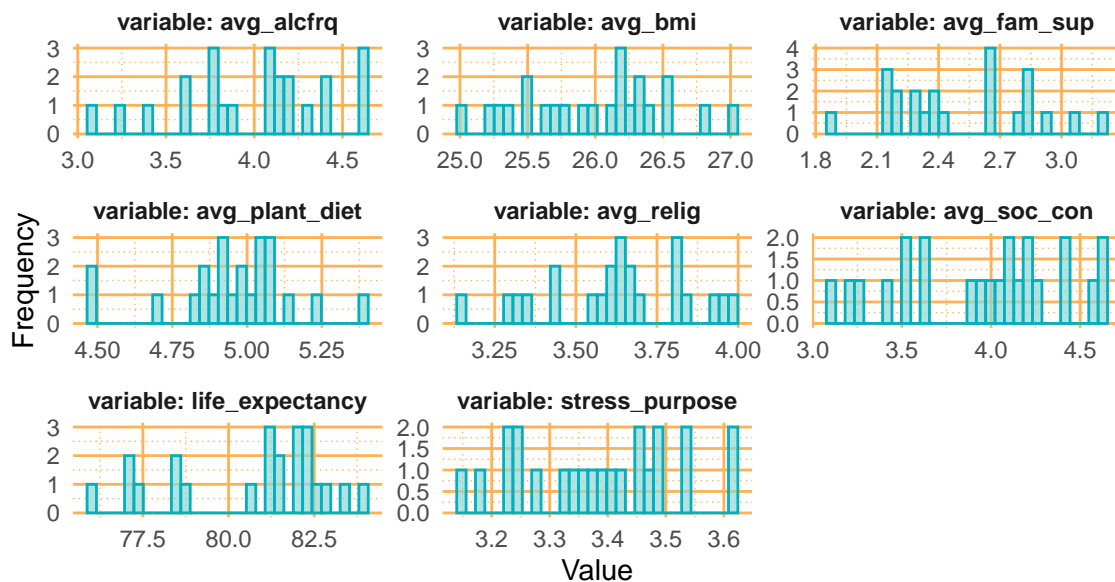
assumptions about the data structure. To enhance the model's interpretability, measures were taken to address multicollinearity among predictors. Specifically, the variable `avg_fam_sup` was excluded due to its high correlation with `avg_bmi` ($r=0.989$). Furthermore, the variables `avg_stress_red` and `avg_sen_pur`, which were strongly correlated ($r=0.854$), were combined into a composite variable, "stress_purpose," to improve model parsimony and clarity.

Clustering, survey weights, fixed / random effects: Clustering adjustments and survey weights, typically used to account for individual-level sampling probabilities and within-group correlation, are less relevant in this context, as the analysis is already based on aggregated country-level data.

Exploratory Data Analysis

The data was analyzed for distributions, variability, and outliers. Descriptive statistics indicate that most predictors exhibit normal or near-normal distributions. Stress and purpose are tightly clustered. Life expectancy has a mean of 80.7 years, ranging from 76.03 to 83.95, with minimal outliers. Visualizations, including histograms and density plots, confirmed these trends, highlighting normal distributions for most variables, while some, like family support, showed slight skewness.

Figure 1: Faceted Histograms of Predictors Values Frequencies



Results

The intercept (61.24, $p = 0.046$) represents the baseline life expectancy when all predictors are at their reference or baseline values. It is statistically significant that the average life expectancy under these conditions is not zero.

Among the predictors, social connections (avg_soc_con, Coefficient = 3.40, $p = 0.022$) were found to have a strong and significant positive association with life expectancy. A one-unit increase in the social connections score corresponds to a 3.40-year increase in life expectancy, emphasizing the substantial role of social factors in promoting longevity. In contrast, the relationship between adherence to a plant-based diet (avg_plant_diet, Coefficient = 2.07, $p = 0.300$) and life expectancy, while positive, was not statistically significant, providing limited evidence for its impact in this model.

Similarly, sense of purpose (stress_purpose, Coefficient = 0.96, $p = 0.791$) exhibited a small, positive, but non-significant association with life expectancy, suggesting minimal influence in this context. Alcohol consumption (avg_alcfrq, Coefficient = -0.54, $p = 0.602$) and religiosity (avg_relig, Coefficient = -0.81, $p = 0.689$) showed weak, non-significant negative associations, while BMI (avg_bmi, Coefficient = -0.09, $p = 0.905$) displayed an even weaker negative association, indicating minimal or no impact on life expectancy in this analysis.

The model demonstrated a reasonable fit, with a dispersion parameter of 0.9389, indicating minimal unexplained variability in the residuals and that the model accounts for the majority of the variation in life expectancy among the included predictors.

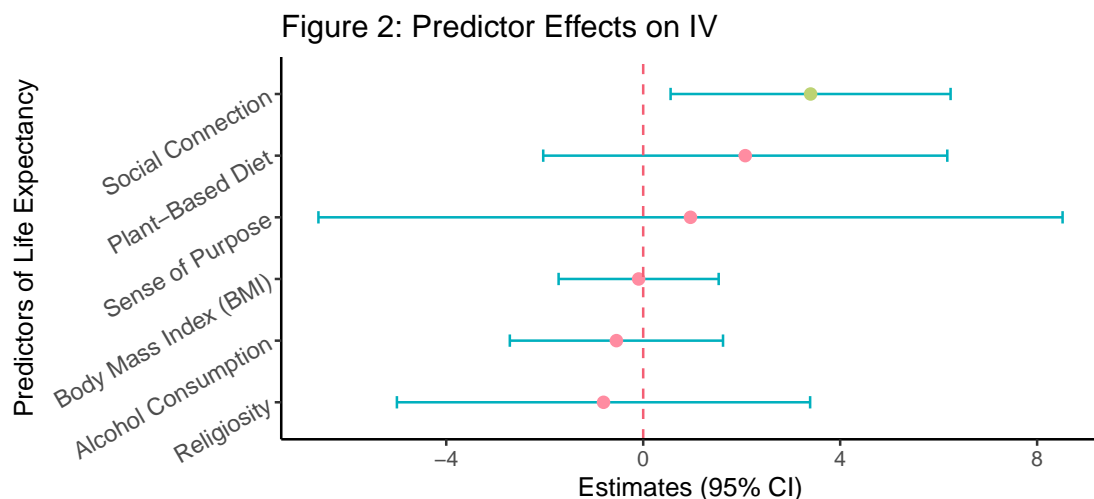


Table 1: Regression Results for Life Expectancy Model with Robust Standard Errors

Predictor	Coefficients				
	Estimate	Std. Error	t-Statistic	P-Value	Signif.
(Intercept)	61.242	28.280	2.190	0.044	*
avg_plant_diet	2.073	1.935	1.114	0.282	
avg_soc_con	3.400	1.341	3.133	0.006	**
stress_purpose	0.961	3.564	0.281	0.782	
avg_alcfrq	-0.544	1.021	-0.576	0.572	
avg_relig	-0.806	1.979	-0.467	0.647	
avg_bmi	-0.093	0.767	-0.125	0.902	

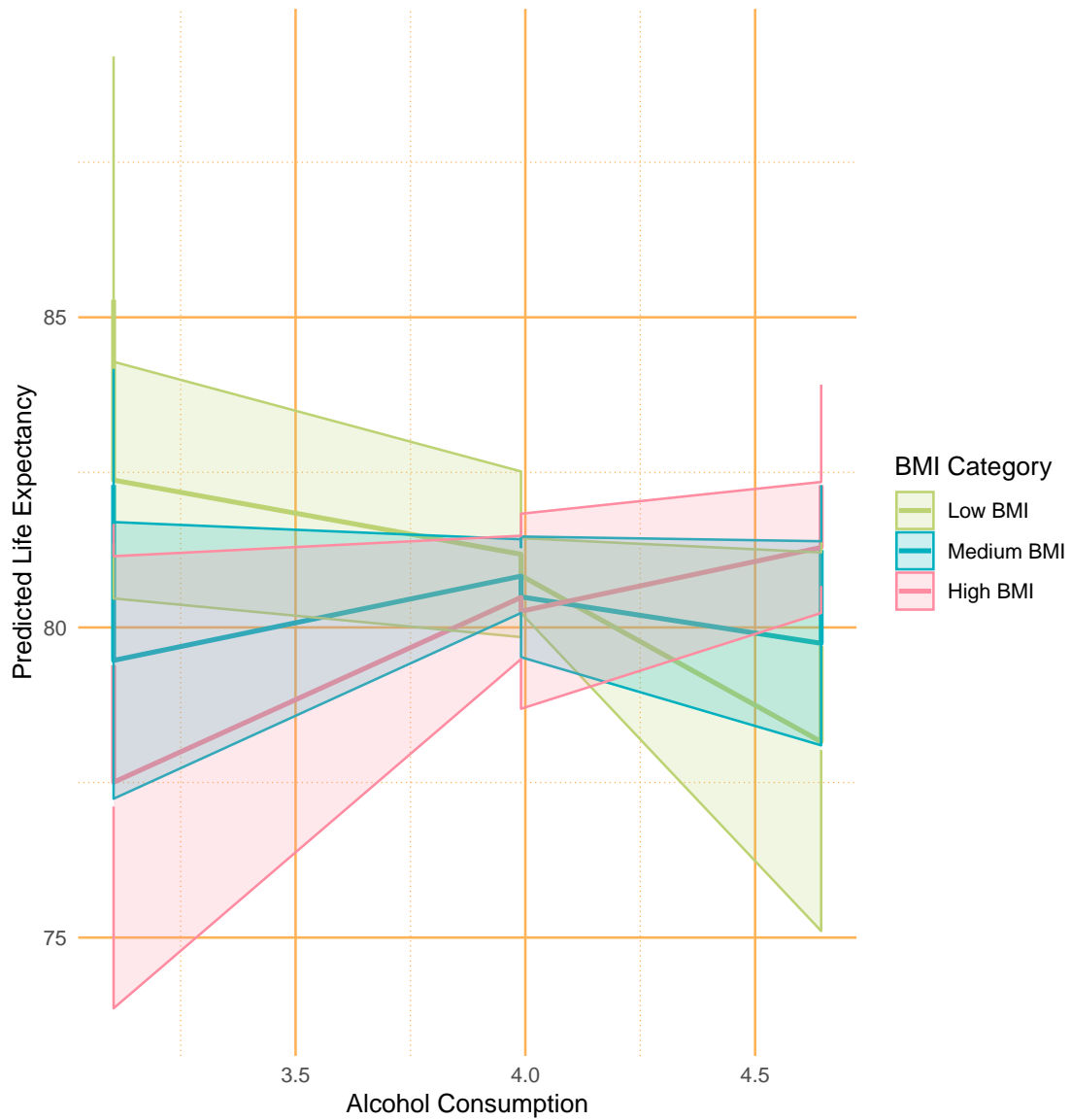
Note: Significance stars are based on p-values. Robust standard errors are computed using HC1 correction

Interactions

Interaction: alcohol consumption - BMI The interaction model shows a significant effect of alcohol consumption (avg_alcfrq) and BMI (avg_bmi) on life expectancy ($p < 0.001$), with a positive interaction coefficient (3.739). This indicates that the impact of alcohol consumption on life expectancy depends on BMI. Specifically, higher alcohol consumption decreases life expectancy for individuals with low BMI, has minimal effect for medium BMI, and is associated with increased life expectancy for high BMI. The crossing point around moderate alcohol consumption (~ 3.75) further highlights this variability.

Confidence intervals suggest significant differences between BMI groups, particularly at the extremes of alcohol use, while predictions near the midpoint are more stable. These results emphasize the context-dependent relationship between alcohol and life expectancy, suggesting that public health recommendations should account for BMI. The findings also highlight potential non-linear dynamics, warranting further exploration to refine predictive accuracy.

Figure 3: Interaction Effect of Alcohol Consumption and BMI on Life Expectancy



Social connections had a significant positive impact on life expectancy, while other predictors showed weak or non-significant associations. The interaction model highlighted that the effect of alcohol consumption on life expectancy varies by BMI, emphasizing the need for context-specific public health recommendations.

Discussion on results

The final model was a linear regression fitted using the `lm` function with robust standard errors (HC1 adjustment) to ensure reliable inference while addressing potential heteroskedasticity. These adjustments allowed the analysis to accurately estimate the relationships between life expectancy and its predictors, which included plant-based diet adherence, social connections, sense of purpose, alcohol consumption frequency, religiosity, and BMI.

The results demonstrated a significant positive association between social connections (`avg_soc_con`, Coefficient = 3.40, $p = 0.022$) and life expectancy, highlighting the critical role of social factors in promoting longevity. Other predictors, such as plant-based diet adherence, sense of purpose, alcohol consumption frequency, religiosity, and BMI, did not show statistically significant associations, suggesting weaker or less stable relationships within this dataset.

Additionally, the interaction model revealed a significant interaction between alcohol consumption (`avg_alcfrq`) and BMI (`avg_bmi`) ($p < 0.001$). This finding suggests that the impact of alcohol consumption on life expectancy depends on BMI: higher BMI appears to mitigate the negative effects of alcohol consumption, with the relationship potentially reversing at higher BMI levels. The crossing of predicted lines in the interaction plot at moderate alcohol consumption (~ 3.75) underscores the complexity of these effects. These insights emphasize the importance of considering BMI when evaluating the health implications of alcohol use.

Overall, the use of robust standard errors, along with the modeling of interactions, ensured the robustness of the analysis and provided meaningful insights into the determinants of life expectancy.

This study investigated the determinants of life expectancy using a linear regression model with robust standard errors, incorporating key lifestyle, social, and psychological predictors. The analysis highlighted the significant positive impact of social connections on life expectancy, emphasizing the critical role of social factors in promoting longer, healthier lives. The interaction analysis further revealed that the impact of alcohol consumption on life expectancy varies with BMI, where higher BMI mitigates the negative effects of alcohol use and even reverses the association at higher levels. Other predictors, such as adherence to a plant-based diet, sense of purpose, religiosity, and BMI, did not exhibit statistically significant direct associations, reflecting weaker or context-specific effects.

Model refinements, including the exclusion of `avg_fam_sup` due to multicollinearity and the creation of the composite variable “`stress_purpose`,” ensured a parsimonious and interpretable framework. The use of robust standard errors improved

the reliability of the findings, particularly in the context of a small, cross-national dataset with potential heteroskedasticity concerns.

Ultimately, this study underscores the importance of addressing social and lifestyle factors to improve life expectancy, with a particular focus on fostering social connections and understanding the nuanced interplay between alcohol consumption and BMI. These findings highlight the need for future research to explore interaction effects further, as well as longitudinal studies to capture the dynamic and collective influences of these factors on longevity.

Limitations

This study has several limitations that should be acknowledged. First, the relatively small sample size of 23 countries limits the statistical power and generalizability of the findings. With such a small number of observations, there is an increased risk of Type II errors, meaning that smaller or subtler effects may go undetected. Additionally, the cross-sectional design of the study precludes the establishment of causality, as the relationships observed reflect a single point in time rather than dynamic changes over time. While robust standard errors were used to address potential heteroskedasticity, and adjustments were made for potential multicollinearity and clustering, these measures may not fully account for the heterogeneity in how life expectancy determinants vary across cultural and economic contexts. Furthermore, the reliance on self-reported measures for variables such as diet, alcohol consumption, and social connections introduces the possibility of response bias, which may affect the precision of the results. These limitations underscore the need for caution in interpreting the findings and point to the importance of future research utilizing larger, longitudinal datasets to validate and expand upon these results.

References

- Dan Buettner and Sam Skemp. Blue Zones: Lessons From the World's Longest Lived. *American Journal of Lifestyle Medicine*, 10(5):318–321, September 2016.
- William C. Cockerham. The Social Determinants of the Decline of Life Expectancy in Russia and Eastern Europe: A Lifestyle Explanation. *Journal of Health and Social Behavior*, 38(2):117, June 1997.
- Ligia J. Dominguez, Nicola Veronese, and Mario Barbagallo. The link between spirituality and longevity. *Aging Clinical and Experimental Research*, 36(1):32, February 2024.
- Attilio Giacosa, Roberto Barale, Luigi Bavaresco, Milena Anna Faliva, Vincenzo Gerbi, Carlo La Vecchia, Eva Negri, Annalisa Opizzi, Simone Perna, Mario Pezzotti, and Mariangela Rondanelli. Mediterranean Way of Drinking and Longevity. *Critical Reviews in Food Science and Nutrition*, 56(4):635–640, March 2016.
- Catrin Herpich, Ursula Müller-Werdan, and Kristina Norman. Role of plant-based diets in promoting health and longevity. *Maturitas*, 165:47–51, November 2022.
- N Kitsera, R Kozovyi, L Kovalchuk, N Kozan, R Han, I Kozova, and N Cherniuk. The Importance of Socio-Psychological Factors and Lifestyle in Shaping Health and Life Expectancy among Longlivers from Precarpathian Region (West Ukraine). *Ivano-Frankivsk National Medical University, Ivano-Frankivsk, Ukraine*.
- A. Monsef and A. S. Mehrjardi. Determinants of Life Expectancy: A Panel Data Approach. *Asian Economic and Financial Review*, 2015.
- United Nations. World Population Prospects 2024: Dataset. *Department of Economic and Social Affairs, Population Division*, 2024.
- Ralph S. Paffenbarger, Robert Hyde, Alvin L. Wing, and Chung-cheng Hsieh. Physical Activity, All-Cause Mortality, and Longevity of College Alumni. *New England Journal of Medicine*, 314(10):605–613, March 1986.
- Ralph S. Paffenbarger and I-Min Lee. Physical Activity and Fitness for Health and Longevity. *Research Quarterly for Exercise and Sport*, 67(sup3):S–11–S–28, September 1996.
- Wanda Rizza, Nicola Veronese, and Luigi Fontana. What are the roles of calorie restriction and diet quality in promoting healthy longevity? *Ageing Research Reviews*, 13:38–45, January 2014.
- Paolo Roffia, Alessandro Buccioli, and Sara Hashlamoun. Determinants of life expectancy at birth: a longitudinal study on OECD countries. *International Journal of Health Economics and Management*, 23(2):189–212, June 2023.
- C. L. Rose. Social Factors in Longevity. *The Gerontologist*, 4(1):27–37, March 1964.
- Achim Siegel, Jonas F. Schug, and Monika A. Rieger. Social Determinants of Remaining Life Expectancy at Age 60: A District-Level Analysis in Germany. *International Journal of Environmental Research and Public Health*, 19(3):1530, January 2022.

Appendix

Model diagnostics

The Breusch-Pagan test ($BP = 8.72$, $df = 6$, $p = 0.18$) showed no evidence of heteroscedasticity, confirming the assumption of constant variance. The residuals vs. fitted plot and Q-Q plot indicated random distribution and sufficient normality, despite minor tail deviations typical for small samples ($n = 23$). High correlation between `avg_fam_sup` and `avg_bmi` ($r = 0.989$) resulted in the removal of `avg_fam_sup`, and strong correlation between `avg_stress_red` and `avg_sen_pur` ($r = 0.854$) led to the creation of the composite variable “`stress_purpose`” to enhance model stability.

Figure 4: Linear Model Assumption Tests Plots

