RIETI Discussion Paper Series 25-E-019

# Reevaluating Midlife Well-being: The role of external and inherent factors

**LI, Chao**
Kyushu University

**MI, Jie**
ADB Institute

**ZHANG, Jiaxu**
Kyushu University

**KEELEY, Alexander Ryota**
Kyushu University

**SHI, Bo**
Kyushu University

**MANAGI, Shunsuke**
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 25-E-019
**February 2025**

# Reevaluating Midlife Well-Being: The role of external and inherent factors*

Chao Li[1], Jie Mi[2], Jiaxu Zhang[3], Alexander Ryota Keeley[1,3], Bo Shi[1], Shunsuke Managi[1,3,4]
[1]Urban Institute, Kyushu University, Fukuoka, Japan
[2]ADB Institute, Tokyo, Japan
[3]Department of Civil Engineering, Kyushu University, Fukuoka, Japan
[4] Research Institute of Economy, Trade and Industry

Abstract

This study delves into the complex causes of low well-being among middle-aged individuals by analyzing over 1.9 million global responses from 168 countries between 2009 and 2022. Employing an exogenous switching treatment effect model and advanced machine learning techniques, this study identifies a U-shaped relationship between age and well-being, where middle-aged individuals experience the lowest levels of well-being. The present study reveals that middle-aged individuals face significantly poorer external treatment compared with the younger and older populations, contributing to a noticeable decrease in their well-being. Conversely, older adults benefit from inherent factors that boost their well-being, illustrating a positive relationship between age and well-being at older ages. Furthermore, the widening disparity in external treatment between age groups over time is particularly pronounced for middle-aged individuals. These findings provide crucial insights for policymakers, emphasizing the need for targeted interventions that address the external challenges disproportionately faced by middle-aged individuals. By understanding and addressing these external disparities, policies can be developed to enhance overall well-being across all age groups.

Keywords: Human Well-Being, Age, Exogenous Switching Treatment Effect Model, Machine Learning, Causal Inference, External Factors, Inherent Factors

JEL classification: I31, C45, J14



*This study is conducted as a part of the Project "Institutional design to achieve a society of well-being" undertaken at the Research Institute of Economy, Trade and Industry (RIETI).
The draft of this paper was presented at the RIETI DP seminar for the paper. I would like to thank participants of the RIETI DP Seminar for their helpful comments.

## Introduction

Well-being has been gradually viewed as a key individual and societal goal, leading most recently to public policy decisions [1-4]. Accurately improving the well-being of a certain group is an effective means of enhancing the entire society. The declining birth rate phenomenon has made the aging problem increasingly serious in major countries, or even in all countries [5,6]. Lower levels of well-being among a certain age group can affect society as a whole. Therefore, obtaining more information on how well-being evolves with age becomes especially pertinent. In fact, although this topic has received considerable attention, it continues to generate extensive discussion and debate among academics, politicians, and the public [1,7]. While numerous studies have attempted to map out well-being trends throughout the lifespan, their findings, especially in empirical research, have varied, often presenting conflicting conclusions [3].

Most studies have shown a link between age and well-being [3,8-10]. Limited by the amount of data, computing power, and people's ability to understand, previous studies have usually attempted to describe the relationship between age and well-being using simple graphs, mainly linear [11,12], U-shaped [8,13], and inverted U-shaped [14]. The U-shaped relationship between age and human well-being has long been observed and is the most widely accepted. Specifically, human well-being hits the lowest point during middle age. This description is most in line with people's intuition because midlife crisis is a common phenomenon. For example, financial stress, one reason for midlife crisis, peaks at middle age, following an inverted U-shaped pattern [15,16]. However, extant studies illustrate an inverted U-shaped or linear link between age and human well-being, which is significantly inconsistent with U-shaped relations. Several studies point out that middle-aged people are relatively better in terms of energy, health status, experience, and ability, among others; therefore, they do not tend to achieve lower well-being in middle age [17,18]. Relatively, the research supporting the inverted U-shaped and linear relation is rare, compared with the evidence supporting U-shaped links. One major aim of this study is to confirm the empirical relationship between age and well-being.

Most previous studies have uniquely focused on the empirical relationship between age and human well-being, but whether this relationship is caused by inherent or external factors is rarely touched upon. Inherent factors refer to qualities or attributes that are natural, essential, and built-in to a person, fundamentally characterizing or influencing their nature without external addition. Previous studies indicate that people born in one generation are more likely to have better well-being than those born in other decades [19-21]. Additionally, health has a strong effect on human well-being, and the elderly population tends to have a poorer health status [22-24]. These factors are typical and inherent to people of a particular age group and may contribute to well-being variation in people of different age groups [1,3]. External reasons are caused by society and family, among others. For instance, financial stress affects human well-being and varies significantly with age [15,16]. If it can be determined whether the age–well-being nexus is caused by internal and/or external reasons, policymakers can develop a series of targeted, effective, and efficient strategies.

The exogenous switching treatment effect model (ESTEM) is a causal inference method based on counterfactual prediction to detect the inherent and external effects of age on human well-being. It allows for the study of how exogenous variable changes affect outcomes under different hypothetical scenarios, thus enabling the distinguishing of exogenous-variable-specific effects from other environmental or contextual factors [25,26]. Technologically, based on the ESTEM, external effects are also called treatment effects because this difference is caused by various treatments. As a simple example, the human well-being status of elderly people should change if they were treated as a young population. This difference is regarded as the treatment effect. Of course, the population of an age group cannot be treated as another age group widely in reality; therefore, the predictions are counterfactual. Moreover, inherent effects derive from base heterogeneity effects. This difference is investigated by comparing populations of different ages treated in the same age group.

Relatively lower goodness-of-fit and poor prediction ability have always been cited as the main problems in previous studies based on regression techniques. Most empirical studies neither conduct cross-validation nor distinguish between training and testing accuracies. If the study only needs to fit the age–well-being relationship, the regression results can still meet the requirements. However, the ESTEM requires models to have good predictive capabilities because counterfactual predictions are its core steps. Machine learning techniques are gaining increasing attention and can be used to produce models with better predictive performances [27,28]. In this study, we replace the linear regression model with an advanced tree-based machine learning model—an extreme gradient boosting (XGBoost) model [27]. We solve several technical problems and provide an example of conducting analyses based on the ESTEM powered by machine learning.

## Materials and Methodology

### *Materials*

#### *Survey Information*

Our study is based on the individual-level survey conducted globally by Gallup, Inc., named the Gallup World Poll (GWP). The current GWP dataset covers 18 years (from 2005 to 2022) and includes 17 waves of surveys. Notably, the first wave was conducted in 2005 and 2006, and each survey wave from the second wave onward is completed within a year. The current version dataset contains 2.594 million individual observations from 168 countries or regions. Specifically, at least 1,000 individuals are sampled from each country and each wave of the GWP survey. The GWP is the largest global dataset, mainly concentrating on human well-being, and it has been widely used in previous studies [8,29,30]. The detailed sampling method and data collection methodology for the GWP are concretely reported on Gallup's website (https://www.gallup.com/178667/gallup-world-poll-work.aspx).

Some observations were removed during the data cleaning process; thus, 1,911,212 observations were finally taken into account in the actual analysis. In the first three waves, income was not asked; thus, the first three waves were directly removed. In the other waves, if respondents did not provide answers about their income, we used the mean income of such respondents' countries in the corresponding wave to estimate the unprovided income values. The mean income is computed by averaging the other available values in the GWP survey in the corresponding country and wave. If the income question was not asked for a country in a particular wave, the data for that country were deleted in that wave. After this step, 2,172,297 observations were retained. As our dependent variable was well-being, we required respondents to have answered the well-being question. This step retained 2,141,833 observations. Previous studies indicate that disability significantly affects human well-being [31,32]; therefore, we excluded observations without available answers. Here, 2,025,803 observations were retained. Additionally, information on age, gender, and employment was strictly needed. After dropping observations with no-answer items, 1,911,212 observations were retained in our dataset. The observation counts for each country and wave are summarized in **Supplementary Materials Table S1**.

*Subjective Well-Being Measurement*

Subjective well-being (SWB) has long been regarded as a reasonable indicator of human well-being [1,33,34]. Overall life evaluation is a critical approach to measure SWB [1,35,36]: it extracts well-being from people's thoughts about the quality of their overall life. In the GWP, an 11-point Cantril ladder is employed to evaluate overall human well-being, that is, respondents imagine a ladder with 11 steps [1]. The ladder ranges from the lowest step, numbered 0, which represents the worst possible life for respondents, to the highest step, numbered as 10, which represents the best possible life for respondents. Respondents then select the step in which they believe they are currently located. The number of the selected step is the respondent's evaluation of life, which is a number ranging between 0 and 10. Because the Cantril ladder is straightforward to understand and has been widely used in previous studies [8,30], it is taken as the dependent variable in our study.

*Independent Variables*

In this study, our dataset includes 63 independent variables, namely, “Wave,” “Country,” “Household Income,” “Health Disability,” “Female Dummy,” “Age,” “Marital Status,” “Employment,” “Children Under 15,” “Feeling of Income,” “Income Level,” “Having Relatives to Rely on,” “Living Standard Changing Direction,” “Enough Food,” “Enough Shelter,” “Well Rested,” “Respected,” “Smiling,” “Interesting Things,” “Enjoyment,” “Physical Pain,” “Worry,” “Sadness,” “Stress,” “Anger,” “Satisfied with City,” “Economic Changing Direction,” “Good Time to Find Job,” “Satisfied with Public Transportation,” “Satisfied with Road,” “Satisfied with Education,” “Satisfied with Air Quality,” “Satisfied with Water Quality,” “Satisfied with Healthcare,” “Satisfied with Affordable House,” “Satisfied with Opportunity to Make Friends,” “Good Place for Ethitical Minority,” “Good Place for Gay or Lesbian,” “Good Place for Immigrants,” “Donated Recently,” “Did Volunteer Recently,” “Helped Stranger,” “Voiced Opinion to Official,” “Confidence in Local Police,” “Safety of Alone Night Walking,” “Stolen,” “Assaulted,” “Religion Importance,” “Children Respected,” “Opportunity for Children Learning,” “Women Respected,” “Satisfied with Poverty Alleviation,” “Satisfied with Environmental Efforts,” “Freedom of Choosing Life,” “Confidence in Military,” “Confidence in Judicial System,” “Confidence in National Government,” “Confidence in Financial System,” “Confidence in Election Honesty,” “Freedom of Media,” “Corruption within Business,” “Corruption within Government,” and “Approval of Leadership Performance.”

**Table 1** summarizes all variables except “Wave” and “Country.” The details of each question in the survey and their value explanations are listed in **Supplementary Materials Table S2**.

***Relationship between Well-being and Age Investigation***

Previous empirical studies have repeatedly indicated that the relationship between age and well-being is U-shaped, as the coefficient of age squared is always significant in the linear regression or other generalized linear regression [3,18]. However, no solid evidence supports this perspective, although it is usually consistent with people’s intuition and observation. In other words, the phenomenon of midlife crisis is common and has been widely noticed [37,38]. Owing to the limitations of data volume and technology, linear regression is a compromising but effective method, that is, linear regression is not good at fitting non-linear relationships. Machine learning models are designed to optimize predictive accuracy by minimizing prediction errors [27]. Additionally, machine learning models make no assumptions about the shape of relationships, thus increasing their ability to fit linear relationships [27,28].

*XGBoost and its fine-tuning*

To detect the empirical relationship between age and well-being, we first use machine learning to fit our dataset. Second, we choose a reasonable explanation approach to elucidate the machine learning methods because such models are not as straightforward as linear models [39,40]. In this study, we take the XGBoost regressor as the main method to replace the linear regression or other linear methods used in previous studies to detect the empirical relationship. Our analysis is set as a regression task, as in previous studies, because the dependent variable is an 11-point Cantril ladder well-being evaluation. We use the Shapley additive explanation (SHAP) method [41] to explain the XGBoost results.

XGBoost offers several significant advantages, and it is a decision-tree-based model [27]. Decision trees have a strong ability to handle complex, non-linear relationships between variables with relatively "rich" tabular data [28]. Furthermore, decision trees can process various data types including binary, continuous, and categorical variables. Moreover, decision trees are completely non-parametric, that is, they do not assume any specific distribution for the data [27]. However, decision trees are prone to overfitting, especially when they are grown extremely deep. Adopting ensemble methods, such as gradient boosting and random forest, and taking decision trees as base learners can significantly enhance performance. The traditional gradient boosting process is inherently sequential, making it challenging to parallelize [27]. XGBoost is an optimized gradient boosting model that supports parallel computation and even GPU acceleration. Of course, other technologies, such as CatBoost [42] and LightGBM [43], as well as random forest based on XGBoost, are also compatible with GPU acceleration. After fine-tuning those models, XGBoost performs the best in generalization, and specifically, the $R^2$ values of fine-tuned XGBoost, CatBoost, LightBoost, and random forest are 37.81%, 37.53%, 37.38%, and 36.46%, respectively. Therefore, we select XGBoost as the main model for our analyses.

The training process of the XGBoost regressor for the total dataset is given as follows:

$$XGB_{tot} = \Theta(\boldsymbol{Xtrain}_{tot}, \boldsymbol{ytest}_{tot}, \boldsymbol{Hyperparameter}_{tot}) \quad (1)$$

where $XGB_{tot}$ represents the well-trained XGBoost regression model based on the total dataset, $tot$ represents the total dataset including all independent variables and observations, $\boldsymbol{Xtrain}_{tot}$ represents the independent variables of the training dataset split from the total dataset, $\boldsymbol{ytrain}_{tot}$ represents the dependent variables of the training dataset, $\boldsymbol{Hyperparameter}_{tot}$ represents a set of hyperparameters to train a high-accuracy XGBoost model, and $\Theta$ represents the training process. The split ratio between the training and test datasets is 9:1. In other words, we randomly sample 90% of the data as the training dataset, and the remaining 10% of the data is left as the test dataset.

The hyperparameter set, $\boldsymbol{Hyperparameter}_{tot}$, includes the number of trees ("n_estimators"), learning rate ("learning_rate"), the maximum depth of each tree ("max_depth"), the subsample ratio of training instance ("subsample"), the minimum loss reduction required for a new split to be added to the tree ("gamma"), the minimum sum of instance weight needed in a child leaf ("min_child_weight"), the maximum step size that a XGBoost model's weight can change at each boosting iteration ("max_delta_step"), L1 regularization term on weights ("reg_alpha"), and L2 regularization term on weights ("reg_lambda"). The abbreviations in parentheses are consistent with XGBoost's Python API to facilitate reproduction and imitation by other researchers. To distinguish them from textual words, they are always enclosed in quotation marks. We adopt the cross-validation method to search for the best hyperparameters. The 10-fold cross-validation method is employed, but owing to the limits of computing resources, we only conduct 3 folds of the 10 to balance the time consumption and stability of the searching process. The metric of the searching process is the $R^2$ of the test dataset. The $R^2$ of the test dataset is computed as follows:

$$\widehat{\boldsymbol{ytest}}_{tot} = XGB_{tot}(\boldsymbol{Xtest}_{tot}) \tag{2}$$

$$R^2_{test\ tot} = 1 - \frac{(\boldsymbol{ytest}_{tot} - \widehat{\boldsymbol{ytest}}_{tot})^2}{(\boldsymbol{ytest}_{tot} - \overline{\boldsymbol{ytest}}_{tot})^2} \tag{3}$$

where $\widehat{\boldsymbol{ytest}}_{tot}$ represents the predicted values corresponding to the input test dataset, $\boldsymbol{Xtest}_{tot}$, of the well-trained XGBoost model $XGB_{tot}$; $\overline{\boldsymbol{ytest}_{tot}}$ is the mean real value of an independent variable; and $R^2_{test\ tot}$ represents the $R^2$ of the test dataset for the model trained using the training dataset from the total dataset. By combining **Equations 1**, **2**, and **3**, it is clear that $R^2_{test\ tot}$ is highly related to hyperparameters.

We use Bayesian hyperparameter optimization to search for the best hyperparameter set [44]. Bayesian hyperparameter optimization normally encompasses four steps: initialization based on several sets of hyperparameters, surrogate function construction, selection of the next set of hyperparameters and metric estimation, and updating the surrogate function. The third and fourth steps are iterated several times to obtain a set of hyperparameters that can achieve high performance. In this study, we set the number of iterations to 20. Simply put, the input of the surrogate function is a set of hyperparameters, and the output is the estimated $R^2$ of the test dataset. The surrogate functions are continuously optimized by interactions. The ranges of these hyperparameters are as follows: 100–5000 for "n_estimators"; 0.001–0.1 for "learning_rate"; 3–16 for "max_depth"; 0.5–1.0 for "subsample"; 0.001–10 for "min_child_weight"; 0.001–10 for "max_delta_step"; 0.001–10 for "gamma"; 0.001–10 for "reg_alpha"; and 0.001–10 for "reg_lambda". We compare the Bayes hyperparameter optimization results with 20 iterations and grid searching results with more than 3,000 possible hyperparameter sets. The Bayes hyperparameter optimization results are relatively better. Although further fine-tuning the hyperparameters by grid searching has the potential to increase the performance, the time cost is too high, and the improvement might be marginal. Hence, we use Bayes hyperparameter optimization to fine-tune all the XGBoost models in this study.

*Contributions of Independent Variables to Well-being*

Tree-based ensemble models such as XGBoost are entirely non-parametric, making it challenging to interpret their results [39]. The SHAP method offers a novel and effective approach for individually estimating the contribution of each specific independent variable to the dependent variable in machine learning models [41]. The method leverages cooperative game theory and Shapley values to ensure that the contributions of the independent variables to the predictions of the complex model are fairly and evenly distributed [39,41]. Naturally, Shapley values are computed by evaluating the change in the predictions based on the well-trained machine learning model before and after adding a certain independent variable to all possible subsets of other independent variables and then averaging these marginal contributions. The contributions of each independent variable at the individual level can be expressed as follows:

$$\boldsymbol{SHAPtest}_{tot} = SHAP(XGB_{tot}, \boldsymbol{Xtest}_{tot}) \quad (4)$$

where $XGB_{tot}$ represents the XGBoost regression model trained using the training dataset split from the total dataset, $SHAP$ represents the standard SHAP algorithm, and $\boldsymbol{SHAPtest}_{tot}$ represents the SHAP values of each independent variable and observation in the test dataset. Theoretically, we can directly use the well-trained XGBoost model and SHAP algorithm to explain all observations, even though the model is overfitting. This is because the SHAP method enumerates all the subsets, including the independent variable of interest, and among hundreds of subsets, only one is the same as the input dataset with all independent variables. If the model learns the observations during the training process, the prediction performance would be significantly better than that of the general observations. To avoid this situation, we investigate the observations in the test dataset. Of course, we can use 10-fold explanations for the entire dataset to solve this issue, which is similar to 10-fold cross-validation. Specifically, we divide the total dataset into 10 folds, use 9 folds to train the XGBoost model, employ the SHAP method and the trained XGBoost to explain the remaining 1-fold observations, and iterate this process until all possible combinations are enumerated. However, the SHAP method is computationally expensive. Each test dataset has approximately 200,000 observations. With a rational setting, we need at least 100 GPU hours to complete the computation. Additionally, we only want to observe an intuitive relationship, and approximately 200,000 observations are considered sufficiently representative.

***Exogenous Switching Treatment Effect Model (ESTEM)***

To investigate the treatment effects and base heterogeneity effects for the difference in well-being among three age groups, we employ an ESTEM that estimates causal effects based on counterfactual analyses. The ESTEM divides all observations into several groups as sub-datasets by an exogenous variable, whereafter it uses each sub-dataset to train the corresponding models, takes each model to predict each sub-dataset, and compares the differences between the predicted values. Simply put, the pattern of external treatments to a specific group of people can be learned using a complex model. If the dependent variables of a certain group of observations are predicted by another model, it can be considered that this group is treated as another group externally. Of course, this is impossible in reality; hence, this method is counterfactual. In this study, we divide the entire dataset into three age groups: the young population aged 40 or less, the middle-aged population aged 40 or above but not exceeding 65, and the elderly population aged more than 65. We present three separate training processes for the young, middle-aged, and elderly populations as follows:

$$\begin{cases} XGB_{yo} = \Theta(\boldsymbol{Xtrain}_{yo}, \boldsymbol{ytest}_{yo}, \boldsymbol{Hyperparameter}_{yo}) \\ XGB_{ma} = \Theta(\boldsymbol{Xtrain}_{ma}, \boldsymbol{ytest}_{ma}, \boldsymbol{Hyperparameter}_{ma}) \\ XGB_{el} = \Theta(\boldsymbol{Xtrain}_{el}, \boldsymbol{ytest}_{el}, \boldsymbol{Hyperparameter}_{el}) \end{cases} \tag{5}$$

where $XGB_{yo}$, $XGB_{ma}$, and $XGB_{el}$ represent the well-trained XGBoost regression models based on the young, middle-aged, and elderly population datasets, respectively; $\boldsymbol{Xtrain}_{yo}$, $\boldsymbol{Xtrain}_{ma}$, and $\boldsymbol{Xtrain}_{el}$ represent the independent variables of the training dataset split from the young, middle-aged, and elderly population datasets, respectively; $\boldsymbol{ytrain}_{yo}$, $\boldsymbol{ytrain}_{ma}$, and $\boldsymbol{ytrain}_{el}$ represent the dependent variables of the three datasets; and $\boldsymbol{Hyperparameter}_{yo}$, $\boldsymbol{Hyperparameter}_{ma}$, and $\boldsymbol{Hyperparameter}_{el}$ represent three sets of hyperparameters used to train high-accuracy XGBoost models for each sub-dataset. We also adopt the cross-validation method to search for the best hyperparameter sets as the previous process as **Equations (1)**, **(2)**, and **(3)**. Additionally, it must be noted that the independent variable “age” is not included in the sub-dataset when training the models and predicting.

To assess the role of age in well-being for each age group, we evaluate the counterfactual well-being status of each age group. Based on this method, the effects of age on well-being can be disentangled by comparing the predicted well-being status under the actual and counterfactual situations. The predictions of the actual and counterfactual well-being of each age group are computed as follows:

$$\widehat{\boldsymbol{ytest}_{yo}^{yo}} = XGB_{yo}(\boldsymbol{Xtest}_{yo}) \tag{6}$$

$$\widehat{\boldsymbol{y}_{yo}^{ma}} = XGB_{yo}(\boldsymbol{X}_{ma}) \tag{7}$$

$$\widehat{\boldsymbol{y}_{yo}^{el}} = XGB_{yo}(\boldsymbol{X}_{el}) \tag{8}$$

$$\widehat{\boldsymbol{y}_{ma}^{yo}} = XGB_{ma}(\boldsymbol{X}_{yo}) \tag{9}$$

$$\widehat{\boldsymbol{ytest}_{ma}^{ma}} = XGB_{ma}(\boldsymbol{Xtest}_{ma}) \tag{10}$$

$$\widehat{\boldsymbol{y}_{ma}^{el}} = XGB_{ma}(\boldsymbol{X}_{el}) \tag{11}$$

$$\widehat{\boldsymbol{y}_{el}^{yo}} = XGB_{el}(\boldsymbol{X}_{yo}) \tag{12}$$

$$\widehat{\boldsymbol{y}_{el}^{ma}} = XGB_{el}(\boldsymbol{X}_{ma}) \tag{13}$$

$$\widehat{\boldsymbol{ytest}_{el}^{el}} = XGB_{el}(\boldsymbol{Xtest}_{el}) \tag{14}$$

where $\widehat{\boldsymbol{ytest}_{yo}^{yo}}$, $\widehat{\boldsymbol{ytest}_{ma}^{ma}}$, and $\widehat{\boldsymbol{ytest}_{el}^{el}}$ are the predicted well-being statuses of observations in test datasets of young, middle-aged, and elderly population sub-datasets estimated by well-trained corresponding models, respectively, which are the actual cases; $\widehat{\boldsymbol{y}_{yo}^{ma}}$, $\widehat{\boldsymbol{y}_{yo}^{el}}$, $\widehat{\boldsymbol{y}_{ma}^{yo}}$, $\widehat{\boldsymbol{y}_{ma}^{el}}$, $\widehat{\boldsymbol{y}_{el}^{yo}}$, and $\widehat{\boldsymbol{y}_{el}^{ma}}$ are counterfactual predicted well-being. Because we use 90% of the data in a sub-dataset to train the model for an age group, this model can be directly applied for the prediction for all data in this age group. Overfitting is the main reason for the aforementioned. Similarly, we employ a 10-fold process to predict all actual cases. Moreover, because these 10-fold predictions are from 10 rather dissimilar models, this would reduce the reliability and stability of the predictions. We conduct 10-fold prediction 10 times based on different fold division strategies. In this way, 100 different models would be obtained for each age group based on the same hyperparameters. The 10-fold 10-epoch computations can be written as follows:

$$\begin{cases} XGB_{yo@fe} = \Theta(\boldsymbol{Xtrain}_{yo@fe}, \boldsymbol{ytest}_{yo@fe}, \boldsymbol{Hyperparameter}_{yo}) \\ XGB_{ma@fe} = \Theta(\boldsymbol{Xtrain}_{ma@fe}, \boldsymbol{ytest}_{ma@fe}, \boldsymbol{Hyperparameter}_{ma}) \\ XGB_{el@fe} = \Theta(\boldsymbol{Xtrain}_{el@fe}, \boldsymbol{ytest}_{el@fe}, \boldsymbol{Hyperparameter}_{el}) \end{cases} \tag{15}$$

$$\widehat{\boldsymbol{ytest}^{yo}_{yo@fe}} = XGB_{yo-fe}(\boldsymbol{Xtest}_{yo@fe}) \tag{16}$$

$$\widehat{\boldsymbol{y}^{ma}_{yo@fe}} = XGB_{yo@fe}(\boldsymbol{X}_{ma}) \tag{17}$$

$$\widehat{\boldsymbol{y}^{el}_{yo@fe}} = XGB_{yo@fe}(\boldsymbol{X}_{el}) \tag{18}$$

$$\widehat{\boldsymbol{y}^{yo}_{ma@fe}} = XGB_{ma@fe}(\boldsymbol{X}_{yo}) \tag{19}$$

$$\widehat{\boldsymbol{ytest}^{ma}_{ma@fe}} = XGB_{ma@fe}(\boldsymbol{Xtest}_{ma@fe}) \tag{20}$$

$$\widehat{\boldsymbol{y}^{el}_{ma@fe}} = XGB_{ma@fe}(\boldsymbol{X}_{el}) \tag{21}$$

$$\widehat{\boldsymbol{y}^{yo}_{el@fe}} = XGB_{el@fe}(\boldsymbol{X}_{yo}) \tag{22}$$

$$\widehat{\boldsymbol{y}^{ma}_{el@fe}} = XGB_{el@fe}(\boldsymbol{X}_{ma}) \tag{23}$$

$$\widehat{\boldsymbol{ytest}^{el}_{el@fe}} = XGB_{el-fe}(\boldsymbol{Xtest}_{el@fe}) \tag{24}$$

where $f$ represents the fold indicator, $e$ represents the epoch indicator, $XGB_{yo@fe}$ represents the XGBoost model trained for the young population in the $f$ fold of the $e$ epoch, $\widehat{\boldsymbol{ytest}^{yo}_{yo@fe}}$ represents the well-being predictions of the test data of the young population sub-dataset based on the corresponding model in the $f$ fold of the $e$ epoch, $\widehat{\boldsymbol{y}^{ma}_{yo@fe}}$ represents the prediction of the middle-aged population based on the model trained by the young population sub-dataset in the $f$ fold of the $e$ epoch, and the explanations of the other symbols are similar. In the 10-fold 10-epoch computations, each actual case should be estimated 10 times, and each counterfactual case should be calculated 100 times. We use the mean values of each individual for further computation:

$$\boldsymbol{apy}^{yo}_{yo} = iwm(\widehat{\boldsymbol{ytest}^{yo}_{yo@fe}}) \tag{25}$$

$$\boldsymbol{apy}^{ma}_{yo} = iwm(\widehat{\boldsymbol{y}^{ma}_{yo@fe}}) \tag{26}$$

$$\boldsymbol{apy}_{yo}^{el} = iwm(\widehat{\boldsymbol{y}_{yo@fe}^{el}}) \tag{27}$$

$$\boldsymbol{apy}_{ma}^{yo} = iwm(\widehat{\boldsymbol{y}_{ma@fe}^{yo}}) \tag{28}$$

$$\boldsymbol{apy}_{ma}^{ma} = iwm(\widehat{\boldsymbol{ytest}_{ma@fe}^{ma}}) \tag{29}$$

$$\boldsymbol{apy}_{ma}^{el} = iwm(\widehat{\boldsymbol{y}_{ma@fe}^{el}}) \tag{30}$$

$$\boldsymbol{apy}_{el}^{yo} = iwm(\widehat{\boldsymbol{y}_{el@fe}^{yo}}) \tag{31}$$

$$\boldsymbol{apy}_{el}^{ma} = iwm(\widehat{\boldsymbol{y}_{el@fe}^{ma}}) \tag{32}$$

$$\boldsymbol{apy}_{el}^{el} = iwm(\widehat{\boldsymbol{ytest}_{el@fe}^{el}}) \tag{33}$$

where $\boldsymbol{apy}_{yo}^{yo}$, $\boldsymbol{apy}_{yo}^{ma}$, $\boldsymbol{apy}_{yo}^{el}$, $\boldsymbol{apy}_{ma}^{yo}$, $\boldsymbol{apy}_{ma}^{ma}$, $\boldsymbol{apy}_{ma}^{el}$, $\boldsymbol{apy}_{el}^{yo}$, $\boldsymbol{apy}_{el}^{ma}$, and $\boldsymbol{apy}_{el}^{el}$ represent the individual-wise average predicted well-being statuses of the young, middle-aged, and elderly populations, respectively, based on models trained by the young, middle-aged, and elderly populations, and $iwm$ represents the individual-wise mean method.

The treatment effects can be estimated as follows:

$$TE_{mag1-mag2}^{pag} = \overline{\boldsymbol{apy}_{mag1}^{pag}} - \overline{\boldsymbol{apy}_{mag2}^{pag}} \tag{34}$$

where $TE_{mag1-mag2}^{pag}$ represents the treatment effect of changing from the treatment for age group $mag1$ to the treatment for age group $mag2$ on the population of age group $pag$; $\overline{\boldsymbol{apy}_{mag1}^{pag}}$ represents the average value of the predictions of the population in age group $pag$ based on the model trained by the $mag1$ population; $\overline{\boldsymbol{apy}_{mag2}^{pag}}$ represents the average value of the predictions of the population in age group $pag$ based on the model trained by the $mag2$ population; and $pag$, $mag1$, and $mag2$ represent one of the young, middle-aged, and elderly age groups, respectively. The base heterogeneity effects can be computed as follows:

$$BHE_{mag}^{pag1-pag2} = \overline{\boldsymbol{apy}_{mag}^{pag1}} - \overline{\boldsymbol{apy}_{mag}^{pag2}} \tag{35}$$

where $BHE_{mag}^{pag1-pag2}$ represents the base heterogeneity effects between age groups $pag1$ and $pag2$ based on the model trained by the population in age group $mag$; $\overline{\boldsymbol{apy}_{mag}^{pag1}}$ represents the average value of the predictions of the population in age group $pag1$ based on the model trained by the $mag$ population; $\overline{\boldsymbol{apy}_{mag}^{pag2}}$ represents the average value of the predictions of the population in age group $pag2$; and $pag1$, $pag2$, and $mag$ represent one of the age groups. To confirm the significance of the treatment effects and base heterogeneity effects, we conduct a t-test between each group of average predicted values computed using **Equations (25)–(33)**.

The temporal variations in base heterogeneity effects and treatment effects can be estimated because the wave order of the survey is included as an independent variable in the analysis. In **Equations (16)–(24)**, we need to refine input data by wave order. Specifically, to compute the base heterogeneity effects and treatment effects in a year, say 2016, we only predict the well-being status of observations surveyed in 2016. The country-level variations in treatment effects should be calculated directly based on all observations in a single country. This is to reduce the complexity of the results and make them easier to understand. The country-level treatment effects are computed as follows:

$$\boldsymbol{apy}_{mag}^{country} = \boldsymbol{apy}_{mag}^{yo@country} \cup \boldsymbol{apy}_{mag}^{ma@country} \cup \boldsymbol{apy}_{mag}^{el@country} \quad (36)$$

$$CTE_{mag1-mag2}^{country} = \overline{\boldsymbol{apy}_{mag1}^{country}} - \overline{\boldsymbol{apy}_{mag2}^{country}} \quad (37)$$

where $\boldsymbol{apy}_{mag}^{country}$ represents the union of the individual-wise average predicted well-being status of the young, middle-aged, and elderly populations in a certain country, $country$, based on the model trained by the population in age group $mag$; $CTE_{mag1-mag2}^{country}$ represents country-level treatment effects between age groups $mag1$ and $mag2$ in the country, $country$. It should be noted that country-level base heterogeneity effects are challenging to detect because the differences in the dataset sizes of the three age groups are more significant within a single country. Therefore, in this study, we do not further investigate base heterogeneity within countries.

## Results

### *Results of Intuitive Relationship between Well-being and Age*

To estimate the intuitive relationship between well-being and age, we employ two steps: first, we calibrate the best model to fit the relationship with human well-being, and second, we use the SHAP method to estimate the contribution of age to human well-being. It should be noted that the independent variables of this model include age, which differs from other models in the ESTEM.

To fine-tune the hyperparameters of XGBoost for the datasets, including the independent variable (age), we conduct Bayesian hyperparameter optimization with 20 iterations based on our cross-validation process. The best hyperparameter set includes "n_estimators" of 2,441, "learning_rate" of 0.0309, "max_depth" of 11, "subsample" of 0.653, "min_child_weight" of 0.167, "max_delta_step" of 0.382, "gamma" of 0.001, "reg_alpha" of 1.124, and "reg_lambda" of 0.007. The average test $R^2$ of the fine-tuned XGBoost is 38.96% from three single-fold $R^2$ values of 38.93%, 38.94%, and 39.02%. Its average training $R^2$ is 60.09% from three single-fold $R^2$ values of 60.10%, 60.08%, and 60.09%. Evidently, there is some overfitting. Reducing model complexity can help mitigate overfitting; however, it may also lead to a decrease in the accuracy of the model on the test data. Therefore, we keep the best hyperparameter set.

**Figure 1** illustrates the relationship between the SHAP values of age and age. The SHAP values are explained as the contributions of age to human well-being. Obviously, the middle ages contribute to human well-being the least, and the U-shape can be easily detected. Thus, the intuitive or empirical relationship between human well-being and age is consistent with that of previous studies.

### ***Well-being Differences among Age Groups and Reasons for the Differences***

#### *Well-being Differences among Age Groups*

The dataset is divided into three age groups: young, middle-aged, and elderly populations. The young population sub-dataset has 1,031,174 observations, the middle-aged population sub-dataset has 663,573 observations, and the elderly population sub-dataset has 216,465 observations. The mean SWB scores in the young, middle-aged, and elderly populations are 5.462, 5.546, and 5.728, respectively. To determine whether the differences among each age group are significant, we conduct a t-test between each group. The three t-test results between the three groups are all significant because the p-values of the three tests are smaller than 0.1%. Specifically, there are significant differences in SWB among people in these three age groups, and the mean SWB increases with age.

#### *Models for Three Age Groups' Fine-tuning and Their Performance*

We employ Bayesian hyperparameter optimization with 20 iterations based on the cross-validation process to calibrate the best hyperparameter sets for the three models for the three age groups. The average test $R^2$ of the best model for the young population is 35.23% based on three single-fold test $R^2$ values of 34.80%, 35.51%, and 35.39%, while its average train $R^2$ is 42.76% based on three single-fold train $R^2$ of 42.78%, 42.75%, and 42.74%. In this way, the overfitting status exists in the model for the young population. Therefore, 10-epoch 10-fold predictions are necessary to solve the overfitting problem. The best hyperparameter set includes "n_estimators" of 1,136, "learning_rate" of 0.0252, "max_depth" of 8, "subsample" of 0.797, "min_child_weight" of 0.033, "max_delta_step" of 3.066, "gamma" of 6.636, "reg_alpha" of 0.001, and "reg_lambda" of 0.030. The average test $R^2$ of the best model for the middle-aged population is 40.95% based on three single-fold test $R^2$ values of

41.30%, 40.59%, and 40.95%, while its average train $R^2$ is 60.31% based on three single-fold train $R^2$ of 60.31%, 60.35%, and 60.26%. The best hyperparameter set contains "n_estimators" of 3304, "learning_rate" of 0.0162, "max_depth" of 13, "subsample" of 0.896, "min_child_weight" of 1.762, "max_delta_step" of 0.100, "gamma" of 0.060, "reg_alpha" of 10, and "reg_lambda" of 0.003. Moreover, the average test $R^2$ of the best model for the elderly population was 40.46% based on three single-fold test $R^2$ values of 40.21%, 40.46%, and 40.72%, while its average train $R^2$ is 67.24% based on three single-fold train $R^2$ of 67.14%, 67.27%, and 67.30%. The best hyperparameter set contains "n_estimators" of 540, "learning_rate" of 0.0306, "max_depth" of 13, "subsample" of 0.557, "min_child_weight" of 0.002, "max_delta_step" of 0.381, "gamma" of 1.205, "reg_alpha" of 6.590, and "reg_lambda" of 0.016.

**Table 2** summarizes the test $R^2$ in 10-epoch 10-fold prediction. Each row in **Table 2** is a summary of 100 test $R^2$ values computed based on a 9:1 cross-validation. The model for a certain age group predicts the corresponding observations 10 times and predicts the observations in other age groups 100 times. From an overall view, the model's performance for young population prediction is relatively worse than that for the other age groups. The main reason for this is that the data size of the young population sub-dataset is large. The variation within this sub-dataset is more complex, which is more challenging to be completely grasped by models. However, the performance of each model is still acceptable because the XGBoost models have significantly improved the prediction ability compared with the linear models.

*Overall Base Heterogeneity and Treatment Effects*

**Table 3** illustrates the base heterogeneity and treatment effects of age on well-being. All the three age groups obtain the lowest well-being status when they are treated as the middle-aged population. The treatment effects between young-population and middle-aged-population treatments, that is, the difference between average predictions of an age group treated as a young population and a middle-aged population, are significantly positive. This means that if a person is treated as a young person, she/he tends to have a better well-being status compared with a situation where she/he is treated as a middle-aged person. The treatment effects between the middle-aged and elderly populations are significantly negative. If a person is treated as a middle-aged population, her/his well-being status would be lower than the situation if she/he is treated as an elderly population. To summarize, the treatment for the middle-aged population significantly reduces human well-being. There is no significant difference between the young population treated as the young and elderly population. The well-being status whereby the middle-aged population is treated as the young population is higher than that of being treated as the elderly population. The well-being status that the elderly population is treated as the young population is lower than that whereby they are treated as the elderly population. Regarding the base heterogeneity effects, relatively older age groups tend to have higher well-being. All base heterogeneity effects between age groups are significant.

### *Variations of Base Heterogeneity and Treatment Effects*

#### *Temporal Variation of Base Heterogeneity and Treatment Effects*

**Figure 2** illustrates the temporal variation in the treatment effects. The treatment effects between being treated as young population and being treated as the middle-aged population in each age group are gradually decreasing temporally in terms of numbers. As the treatment effects of switching from treatments for the young population to treatments for the middle-aged population are mainly positive, people in any age group treated as the young population are prone to have higher human well-being compared with the situations where they are treated as the middle-aged population. The treatments for the young and middle-aged populations become closer. **Table 4** presents the treatment effects for each year. Additionally, the difference in any age group between the treatments for the middle-aged and those for the elderly population gradually become larger. People in any age group treated as an elderly population tend to have higher human well-being compared with the situations in which they are treated as the middle-aged population. The treatment effects between populations in any age group treated as young and elderly are gradually reversing. Before 2015, the population treated as the young population tended to achieve a higher level of human well-being compared with the population treated as the elderly population. After 2017, the treatment effects were reversed: that is, the population treated as the elderly population could obtain higher human well-being. Indeed, in terms of mean values, the human well-being of the elderly group is higher than that of the other age groups, and this difference is also increasing, as shown in **Figure 3**.

**Figure 4** depicts the temporal variation in base heterogeneity effects. The base heterogeneity effects between the young and middle-aged populations are temporally stable, without significant and clear change trends. The base heterogeneity effects between the young and elderly populations and those between the middle-aged and elderly populations become larger temporally, accompanied by a certain level of vibration. **Table 5** presents the base heterogeneity effects for each year. These temporal variations in the base heterogeneity effects are mainly caused by the change in the elderly population supported by **Figure 3**.

*Country-level Variation of Treatment Effects*

The country-level of treatment effects are summarized in **Table 6**. We have roughly divided these countries into five styles based on how harshly they treat each age group: Younger Enjoy, Older Enjoy, Middle-age Suffer, Middle-age Enjoy, and No Trend. Among the three values of the country-level treatment effects between the three age groups in a certain country, if two or more of them are not significant, then that country would be classified as No Trend, that is, no significant trend exists. If two or more values are significantly positive, the country would be classified as Younger Enjoy, that is, people who are treated as younger are prone to having a better well-being status. Conversely, if two or more values are significantly negative, the country would be classified as Older Enjoy, meaning that people who are considered relatively older tend to have better well-being. If, in a country, the treatment effects between the young and middle-aged populations are positive and those between the middle-aged and elderly populations are negative, the country should be labeled as Middle-age Suffer. In other words, if the people are treated as the middle-aged population, they tend to have the lowest well-being. If the treatment effects between the young and middle-aged populations are negative and those between the middle-aged and elderly populations are positive, the country should be labeled as Middle-age Enjoy. If people are considered middle-aged, they tend to have the highest levels of well-being. Among 165 countries, the counts of each style are listed as follows: 51 of Older Enjoy, 49 of Younger Enjoy, 48 of Middle-age Suffer, 10 of Middle-age Enjoy, and 6 of No Trend. Although Middle-age Suffer is not absolutely mainstream in these countries, treatments for the middle-aged population are basically in a relatively unfavorable state.

**Discussion**

Based on a large global dataset and cutting-edge technologies, the reasons for the low human well-being in middle-aged people are investigated. In the GWP survey, we employ the ESTEM based on machine learning technology to analyze more than 1.9 million observations from 168 countries or regions during the 2009–2022 period. This study contributes to the literature in the following ways: First, with high-accuracy machine learning models based on big data, the empirical relationship between human well-being and age is U-shaped, that is, the middle ages contribute the least to human well-being. Second, middle-aged people receive the worst external treatment, and the external treatments for the young and elderly populations are similar. Third, the base heterogeneity difference shows that older people generally experience higher levels of human well-being, both naturally and inherently. Fourth, the external treatments for young and middle-aged people are gradually becoming more stringent compared with those for the elderly. Fifth, according to the temporal variation in the base heterogeneity difference between each age group, older people inherently become more likely to achieve higher levels of human well-being temporally. This study demonstrates that external treatments may be an important factor vis-à-vis midlife crisis. Additionally, we explore the trends in treatment effects and base heterogeneity effects among several age groups. Our study provides insights into the variations in human well-being among different age groups. These findings can contribute to society by informing policies and programs aimed at improving quality of life, tailoring social services to meet the specific needs of various age groups, and fostering a better understanding of how well-being evolves with age.

The empirical relationship between age and human well-being in this study is U-shaped, which is significantly consistent with the findings of various previous studies [1,4,8,12,18]. There are two popular explanations for this phenomenon. First, midlife crises and their various accompanying symptoms are critical reasons for the lower human well-being of middle-aged people [37,38]. Specifically, in middle age, people experience more sleeping problems [45], concentration difficulties, financial stress [16], and extreme depression [38], which can significantly reduce human well-being. Second, people born in some decades are more likely to achieve higher well-being [11,19,20]. For instance, Shu et al. [20] indicated that individuals born between 1956 and 1961 faced challenges at several critical stages in their lives, including education, employment, economic stability, and social connections, and they report lower human well-being compared with other cohorts. This cohort happens to be middle-aged people. Sutin et al. [19] demonstrated that while well-being generally increased with age for everyone, cohorts that experienced the economic hardships of the early 20th century reported lower well-being compared with those born during more prosperous times. However, these explanations do not distinguish whether the causes of this U-shaped relationship are externally or internally idiosyncratic. For example, previous studies have not illustrated whether mental disorders are triggered by external pressure and treatments or caused by the vulnerable status of a certain life stage. Our study proves that external treatments for the three age groups are U-shaped. Specifically, the middle-aged population is treated relatively more harshly.

The trend of the base heterogeneity difference demonstrates that older people are prone to experiencing higher human well-being naturally and inherently. Older people have a higher age but also live longer and spend more time achieving the goals of their possible lives. On the one hand, aging consistently reduces human well-being [22], as various health problems, attitudes toward life changes, and social relationships are associated with aging [4,10,23,24]. On the other hand, the effect of time cannot be ignored. Because the SWB method used in this research is the Cantril ladder, the elderly population has more time to climb the higher steps. In this way, inherently, elderly people are prone to having a higher global human well-being evaluation. It should be noted in this study that as we only investigate global well-being based on life evaluation, the findings may be inconsistent with research using other well-being indicators, such as hedonic [4] and eudemonic well-being [46]. The base heterogeneity effects can be regarded as the effects of time. Combining the results of previous studies with our findings, the effects of time and aging are completely opposite to each other. Because they cancel each other out, the empirical result is U-shaped.

Temporally, the treatments for the young and middle-aged populations are becoming similar, and the differences in treatments between the elderly population and other age groups are gradually increasing. The base heterogeneity effects between the elderly population and the other two age groups become larger gradually. Because a cohort of people becomes an elderly population that tends to experience higher human well-being, the heterogeneity effects vary temporally. This finding supports previous studies showing gaps between generations [20]. Additionally, although the situation varies from one country to another, the treatment of the middle-aged population in most countries is relatively unfavorable. Therefore, adopting strategies that can reduce stress among middle-aged people is an important means of improving their level of well-being.

Our innovation in connection to the method is also noteworthy. First, we use tree-based machine learning methods to replace the traditional regression method, which is more suitable for understanding non-linear relationships. Second, the SHAP method can fairly distribute the contribution to each independent variable individually. It is an essential way to accurately illustrate the relationship between age and well-being at the statistical and social science levels. In fact, the accuracy of our model significantly exceeds that of previous studies based on linear regressions or similar technologies, as, normally, their $R^2$ is less than 25%, for example, Blanchflower and Piper [9] and Blanchflower and Graham [13]. Seemingly, we only slightly improved the accuracy compared with previous studies based on regression technologies. However, the $R^2$ in the previous studies is the training accuracy, that is, employing the regression method in that way is completely unable to monitor and avoid overfitting. The test $R^2$ is an effective and necessary metric for checking whether the model really grasps the relationship. If the generalization of the models in the ESTEM is poor, the counterfactual prediction would be totally unreliable. When data are abundant, using more complex machine learning models can effectively reduce the impact of this problem.

Society and governments should pay more attention to the middle-aged population. They are the backbone of social and family development; therefore, they also bear relatively more pressure. Inherently, the middle-aged population should feel achieving a better life than young people; however, their responsibilities or expectations put them into a midlife crisis. Policies should consider the needs and dilemmas of middle-aged people to achieve a sustainable society. Additionally, we are aware that the harsh treatment is gradually spreading to younger people. A life situation of the young population that becomes increasingly difficult because of external treatments can lead to a variety of problems, such as a declining birth rate, a lower marriage rate, and a relatively weak economic environment.

Although we adopt advanced technologies and the largest global dataset to probe the reasons for the lower human well-being among the middle-aged population, several limitations of this study should be noted. First, there are relatively large differences in the sample sizes among the three age groups. This results in a relatively poor generalization ability of some models. Second, some important variables, such as educational background, are not obtained in the dataset. Furthermore, most of the variables in the analysis are binary, which somewhat provides insufficient information. Third, limited by the computing ability of the hardware and the data size, we only divide the total dataset into three age groups. If we set more age groups, more interesting findings might be obtained. To enhance future studies on human well-being across different age groups, several improvements should be considered. First, the sample sizes across age groups should be balanced to enhance model performance in the ESTEM. Second, a broader range of variables should be included to provide deeper insights into the determinants of well-being. Third, more detailed variable types beyond binary options should be used to capture nuanced data effectively. Additionally, adopting more sophisticated statistical or machine learning methods could address complex datasets and reveal more intricate patterns. Finally, expanding the age categorizations and incorporating longitudinal and cross-cultural data could uncover dynamic trends and cultural influences on well-being.

## Conclusions

The research findings indicating a U-shaped relationship between age and well-being statistically and empirically, with middle-aged individuals experiencing the lowest levels of well-being, have significant implications for policy and practice. The apparent disparities in external treatment across different age groups suggest that targeted interventions are necessary to enhance well-being among the middle-aged population, who seem to receive the least favorable external conditions. Furthermore, the inherent tendency for older adults to experience higher well-being naturally suggests that policies should focus on maintaining these levels while also addressing the stricter external conditions imposed on the younger and middle-aged groups. To optimize well-being across the lifespan, policies should be dynamic and adapt to the shifting base differences among age groups, ensuring that interventions are timely and tailored to the evolving needs of each demographic. This strategy not only promotes a more equitable distribution of resources but also supports the overall goal of enhancing human well-being in an aging society.

**Acknowledgments**

This study was conducted as part of a project titled "Institutional design to achieve a society of well-being" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). The authors would like to express their deepest gratitude to the professors and researchers whose invaluable advice and thoughtful feedback have greatly enriched the quality of this research.

**Figures**

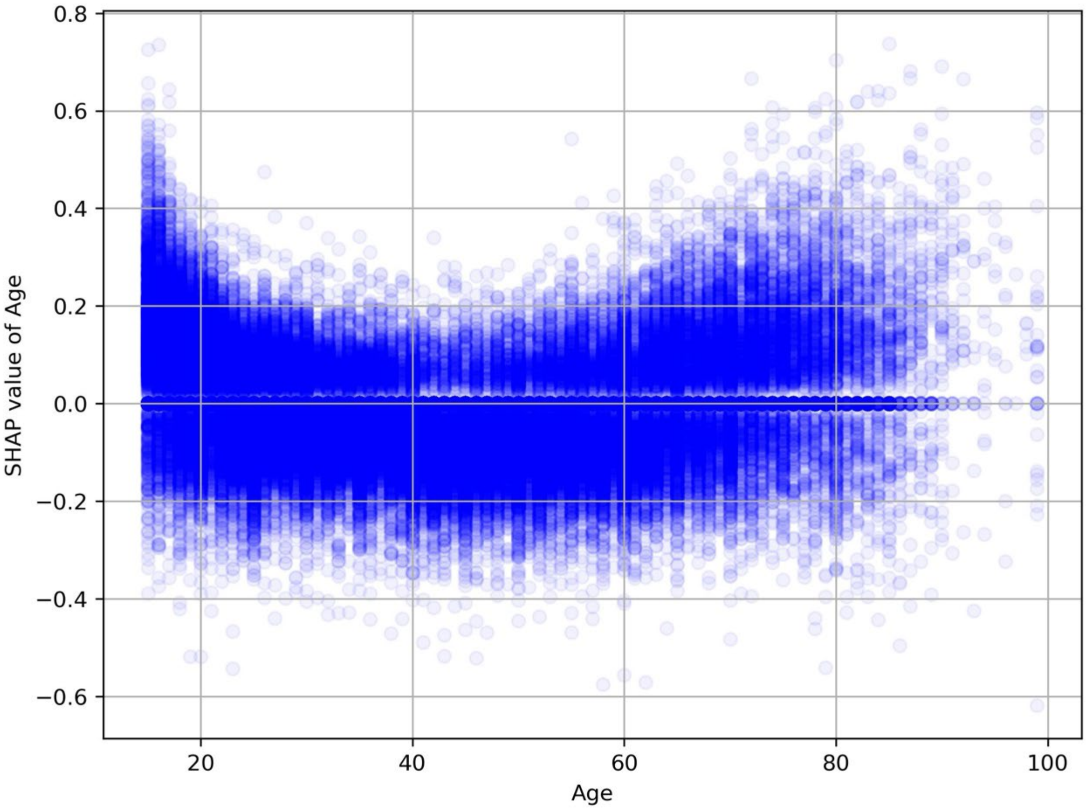


**Figure 1: The Relationship between SHAP Value of Age and Age**

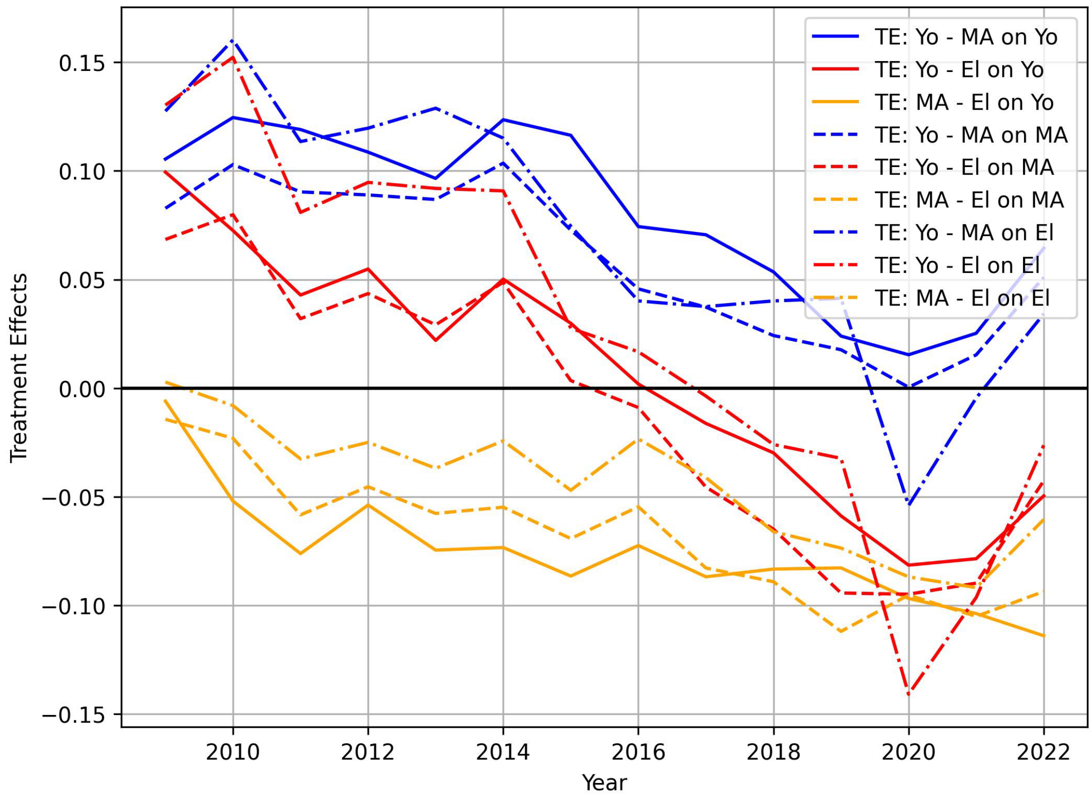


**Figure 2: Temporal Variation of Treatment Effect**

(**Note:** TE: Yo – MA on Yo represents the treatment effect between the young population treated as the young and middle-aged population. Other explanations are similar.**)**

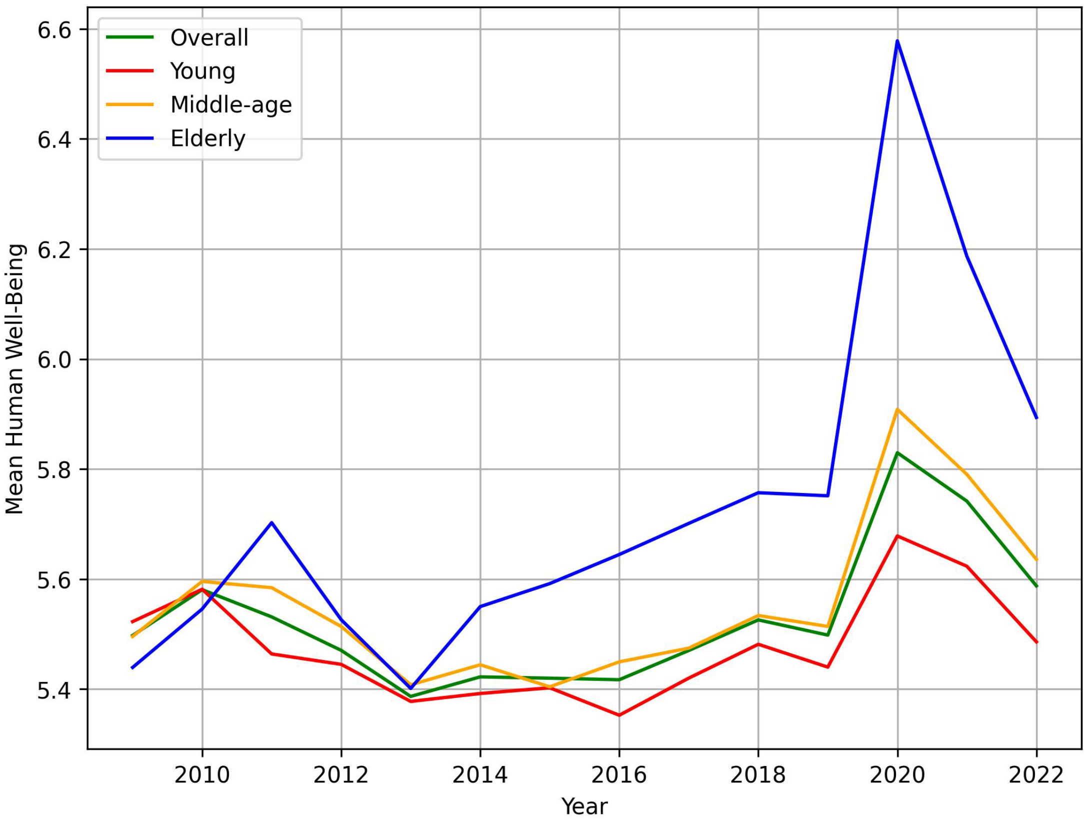


**Figure 3: The Average Human Well-Being Variation of Each Age Group**

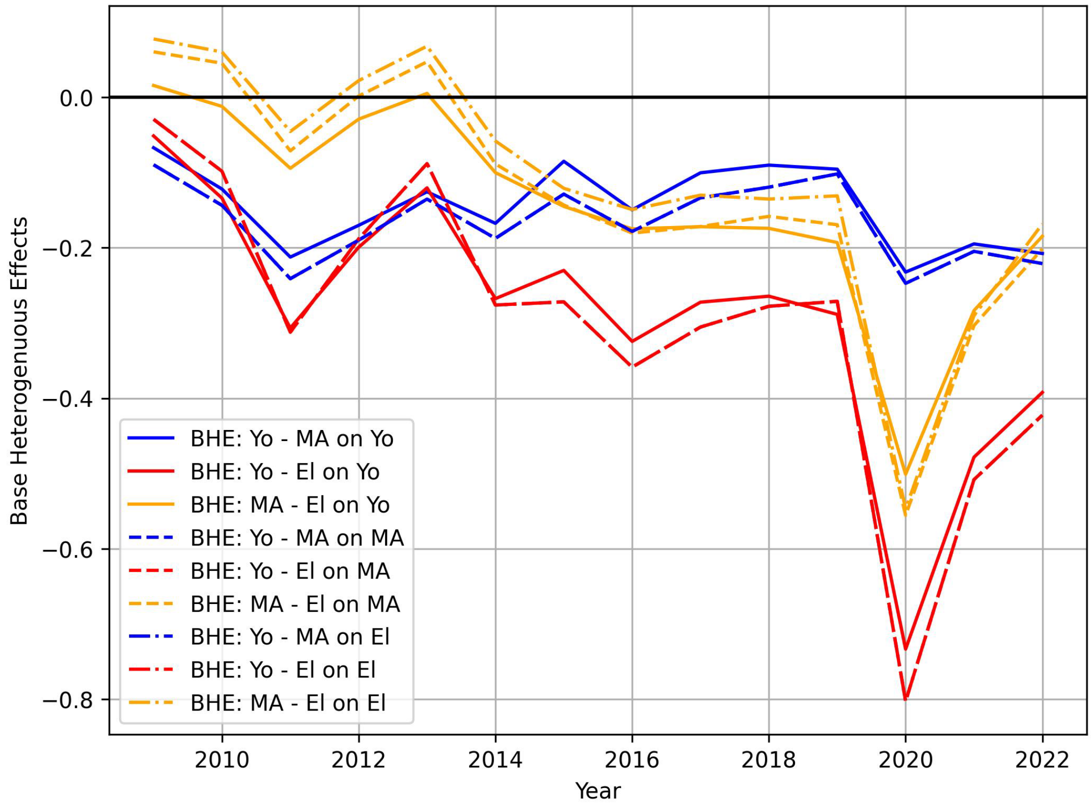


**Figure 4: Temporal Variation of Base Heterogeneity Effect**

(**Note:** BHE: Yo – MA on Yo represents the base heterogeneity effect between the young population treated as the young and middle-aged population. Other explanations are similar.)

**Tables**

**Table 1: Data Summary**

| | mean | std | min | 25% | 50% | 75% | max |
|---|---|---|---|---|---|---|---|
| SWB | 5.521 | 2.414 | 0 | 4 | 5 | 7 | 10 |
| Household Income | 25339.301 | 925219.104 | 0.000 | 4190.875 | 10510.192 | 26055.600 | 898033649.954 |
| Health Disability | 1.754 | 0.431 | 1 | 2 | 2 | 2 | 2 |
| Female Dummy | 0.532 | 0.499 | 0 | 0 | 1 | 1 | 1 |
| Age | 41.429 | 17.552 | 15 | 27 | 39 | 54 | 99 |
| Marital Status | 2.344 | 1.710 | 1 | 1 | 2 | 2 | 8 |
| Employment | 3.646 | 2.134 | 1 | 1 | 4 | 6 | 6 |
| Children Under 15 | 0.522 | 0.500 | 0 | 0 | 1 | 1 | 1 |
| Feeling of Income | 2.323 | 0.965 | 1 | 2 | 2 | 3 | 4 |
| Income Level | 3.236 | 1.415 | 1 | 2 | 3 | 5 | 5 |
| Having Relatives to Rely on | 0.791 | 0.407 | 0 | 1 | 1 | 1 | 1 |
| Living Standard Changing Direction | 0.175 | 0.820 | -1 | -1 | 0 | 1 | 1 |
| Enough Food | 0.322 | 0.467 | 0 | 0 | 0 | 1 | 1 |
| Enough Shelter | 0.245 | 0.430 | 0 | 0 | 0 | 0 | 1 |
| Well Rested | 0.674 | 0.469 | 0 | 0 | 1 | 1 | 1 |
| Respected | 0.865 | 0.342 | 0 | 1 | 1 | 1 | 1 |
| Smiling | 0.703 | 0.457 | 0 | 0 | 1 | 1 | 1 |
| Interesting Things | 0.520 | 0.500 | 0 | 0 | 1 | 1 | 1 |
| Enjoyment | 0.689 | 0.463 | 0 | 0 | 1 | 1 | 1 |
| Physical Pain | 0.307 | 0.461 | 0 | 0 | 0 | 1 | 1 |
| Worry | 0.380 | 0.485 | 0 | 0 | 0 | 1 | 1 |
| Sadness | 0.234 | 0.423 | 0 | 0 | 0 | 0 | 1 |
| Stress | 0.329 | 0.470 | 0 | 0 | 0 | 1 | 1 |
| Anger | 0.197 | 0.398 | 0 | 0 | 0 | 0 | 1 |
| Satisfied with City | 0.758 | 0.429 | 0 | 1 | 1 | 1 | 1 |
| Economic Changing Direction | 0.054 | 0.851 | -1 | -1 | 0 | 1 | 1 |
| Good Time to Find Job | 0.337 | 0.473 | 0 | 0 | 0 | 1 | 1 |
| Satisfied with Public Transportation | 0.560 | 0.496 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Road | 0.523 | 0.499 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Education | 0.604 | 0.489 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Air Quality | 0.713 | 0.453 | 0 | 0 | 1 | 1 | 1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Satisfied with Water Quality | 0.668 | 0.471 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Healthcare | 0.544 | 0.498 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Affordable House | 0.463 | 0.499 | 0 | 0 | 0 | 1 | 1 |
| Satisfied with Opportunity to Make Friends | 0.684 | 0.465 | 0 | 0 | 1 | 1 | 1 |
| Good Place for Ethitical Minority | 0.575 | 0.494 | 0 | 0 | 1 | 1 | 1 |
| Good Place for Gay or Lesbian | 0.287 | 0.452 | 0 | 0 | 0 | 1 | 1 |
| Good Place for Immigrants | 0.569 | 0.495 | 0 | 0 | 1 | 1 | 1 |
| Donated Recently | 0.306 | 0.461 | 0 | 0 | 0 | 1 | 1 |
| Did Volunteer Recently | 0.201 | 0.401 | 0 | 0 | 0 | 0 | 1 |
| Helped Stranger | 0.495 | 0.500 | 0 | 0 | 0 | 1 | 1 |
| Voiced Opinion to Official | 0.159 | 0.366 | 0 | 0 | 0 | 0 | 1 |
| Confidence in Local Police | 0.558 | 0.497 | 0 | 0 | 1 | 1 | 1 |
| Safety of Alone Night Walking | 0.590 | 0.492 | 0 | 0 | 1 | 1 | 1 |
| Stolen | 0.139 | 0.346 | 0 | 0 | 0 | 0 | 1 |
| Assaulted | 0.047 | 0.211 | 0 | 0 | 0 | 0 | 1 |
| Religion Importance | 0.637 | 0.481 | 0 | 0 | 1 | 1 | 1 |
| Children Respected | 0.640 | 0.480 | 0 | 0 | 1 | 1 | 1 |
| Opportunity for Children Learning | 0.671 | 0.470 | 0 | 0 | 1 | 1 | 1 |
| Women Respected | 0.566 | 0.496 | 0 | 0 | 1 | 1 | 1 |
| Satisfied with Poverty Alleviation | 0.357 | 0.479 | 0 | 0 | 0 | 1 | 1 |
| Satisfied with Environmental Efforts | 0.494 | 0.500 | 0 | 0 | 0 | 1 | 1 |
| Freedom of Choosing Life | 0.723 | 0.448 | 0 | 0 | 1 | 1 | 1 |
| Confidence in Military | 0.586 | 0.493 | 0 | 0 | 1 | 1 | 1 |
| Confidence in Judicial System | 0.431 | 0.495 | 0 | 0 | 0 | 1 | 1 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Confidence in National Government | 0.420 | 0.494 | 0 | 0 | 0 | 1 | 1 |
| Confidence in Financial System | 0.531 | 0.499 | 0 | 0 | 1 | 1 | 1 |
| Confidence in Election Honesty | 0.402 | 0.490 | 0 | 0 | 0 | 1 | 1 |
| Freedom of Media | 0.543 | 0.498 | 0 | 0 | 1 | 1 | 1 |
| Corruption within Business | 0.610 | 0.488 | 0 | 0 | 1 | 1 | 1 |
| Corruption within Government | 0.581 | 0.493 | 0 | 0 | 1 | 1 | 1 |
| Approval of Leadership Performance | 0.406 | 0.491 | 0 | 0 | 0 | 1 | 1 |

**Table 2: Summary of Test $R^2$ in 10-epoch 10-fold Prediction**

| Sub-dataset | Model | Average Test $R^2$ | Maximum Test $R^2$ | Minimum Test $R^2$ | Standard Deviation | Number of Predictions for Each Obs. |
|---|---|---|---|---|---|---|
| Young Population | Young Population | 35.23% | 36.04% | 34.50% | 0.28% | 10 |
| Middle-age Population | Young Population | 39.83% | 39.86% | 39.80% | 0.02% | 100 |
| Elderly Population | Young Population | 38.74% | 38.83% | 38.60% | 0.04% | 100 |
| Young Population | Middle-age Population | 33.14% | 33.18% | 33.08% | 0.02% | 100 |
| Middle-age Population | Middle-age Population | 40.72% | 41.53% | 40.00% | 0.31% | 10 |
| Elderly Population | Middle-age Population | 40.77% | 40.81% | 40.71% | 0.02% | 100 |
| Young Population | Elderly Population | 29.79% | 29.91% | 29.66% | 0.06% | 100 |
| Middle-age Population | Elderly Population | 37.64% | 37.72% | 37.56% | 0.03% | 100 |
| Elderly Population | Elderly Population | 40.43% | 41.69% | 39.17% | 0.51% | 10 |

**Table 3: Treatment and Base Heterogeneity Effects**

| | | Treatment | | | Treatment Effect | | |
|---|---|---|---|---|---|---|---|
| | | Treated as Young Population | Treated as Middle-aged Population | Treated as Elderly Population | $TE_{Yo-MA}^{mag}$ | $TE_{MA-El}^{mag}$ | $TE_{Yo-El}^{mag}$ |
| Age Group | Young Population | 5.462 | 5.382 | 5.459 | 0.080*** | -0.077*** | 0.003 |
| | Middle-age Population | 5.605 | 5.545 | 5.615 | 0.060*** | -0.070*** | -0.010*** |
| | Elderly Population | 5.749 | 5.681 | 5.726 | 0.068*** | -0.045*** | 0.023*** |
| Base Heterogeneity Effect | $BHE_{pag}^{Yo-MA}$ | -0.143*** | -0.163*** | -0.156*** | | | |
| | $BHE_{pag}^{MA-El}$ | -0.144*** | -0.136*** | -0.111*** | | | |
| | $BHE_{pag}^{Yo-El}$ | -0.287*** | -0.299*** | -0.267*** | | | |

Note: *** p value < 1%; ** p value < 5%; * p value < 1%. Yo represents the young population, MA represents the middle-aged population, and El represents the elderly population. TE and BHE estimations follow **Equations 34** and **35**. $mag$ and $pag$ could be the young population, middle-aged population, and elderly population.

**Table 4: Temporal Variation of Treatment Effects**

| Year | $TE^{Yo}_{Yo-MA}$ | $TE^{Yo}_{Yo-El}$ | $TE^{Yo}_{MA-El}$ | $TE^{MA}_{Yo-MA}$ | $TE^{MA}_{Yo-El}$ | $TE^{MA}_{MA-El}$ | $TE^{El}_{Yo-MA}$ | $TE^{El}_{Yo-El}$ | $TE^{El}_{MA-El}$ |
|---|---|---|---|---|---|---|---|---|---|
| 2009 | 0.105*** | 0.099*** | -0.006 | 0.083*** | 0.068*** | -0.014 | 0.128*** | 0.130*** | 0.003 |
| 2010 | 0.125*** | 0.073*** | -0.052*** | 0.103*** | 0.080*** | -0.023** | 0.160*** | 0.152*** | -0.008 |
| 2011 | 0.119*** | 0.043*** | -0.076*** | 0.090*** | 0.032*** | -0.058*** | 0.113*** | 0.081*** | -0.033* |
| 2012 | 0.109*** | 0.055*** | -0.054*** | 0.089*** | 0.044*** | -0.045*** | 0.120*** | 0.095*** | -0.025 |
| 2013 | 0.097*** | 0.022*** | -0.075*** | 0.087*** | 0.029*** | -0.058*** | 0.129*** | 0.092*** | -0.037* |
| 2014 | 0.124*** | 0.050*** | -0.073*** | 0.104*** | 0.049*** | -0.055*** | 0.115*** | 0.091*** | -0.024 |
| 2015 | 0.116*** | 0.030*** | -0.087*** | 0.073*** | 0.003 | -0.069*** | 0.075*** | 0.028 | -0.047** |
| 2016 | 0.074*** | 0.002 | -0.072*** | 0.046*** | -0.009 | -0.055*** | 0.040** | 0.017 | -0.023 |
| 2017 | 0.071*** | -0.016** | -0.087*** | 0.037*** | -0.046*** | -0.083*** | 0.038** | -0.004 | -0.041** |
| 2018 | 0.054*** | -0.030*** | -0.083*** | 0.024** | -0.065*** | -0.089*** | 0.040** | -0.026 | -0.066*** |
| 2019 | 0.024*** | -0.059*** | -0.083*** | 0.018* | -0.094*** | -0.112*** | 0.041*** | -0.032** | -0.074*** |
| 2020 | 0.015** | -0.081*** | -0.097*** | 0 | -0.095*** | -0.095*** | -0.054*** | -0.141*** | -0.087*** |
| 2021 | 0.025*** | -0.079*** | -0.104*** | 0.015 | -0.090*** | -0.105*** | -0.005 | -0.096*** | -0.092*** |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2022 | 0.064*** | -0.050*** | -0.114*** | 0.051*** | -0.043*** | -0.094*** | 0.034** | -0.026 | -0.060*** |

Note: *** p value < 1%; ** p value < 5%; * p value < 1%. Yo represents the young population, MA represents the middle-aged population, and El represents the elderly population. TE estimations follow **Equation 34**.

**Table 5: Temporal Variation of Base Heterogeneous Effects**

| Year | $BHE_{Yo}^{Yo-MA}$ | $BHE_{Yo}^{Yo-El}$ | $BHE_{Yo}^{MA-El}$ | $BHE_{MA}^{Yo-MA}$ | $BHE_{MA}^{Yo-El}$ | $BHE_{MA}^{MA-El}$ | $BHE_{El}^{Yo-MA}$ | $BHE_{El}^{Yo-El}$ | $BHE_{El}^{MA-El}$ |
|---|---|---|---|---|---|---|---|---|---|
| 2009 | -0.067*** | -0.051*** | 0.016 | -0.090*** | -0.029* | 0.060*** | -0.090*** | -0.029 | 0.078*** |
| 2010 | -0.122*** | -0.134*** | -0.012 | -0.144*** | -0.098*** | 0.045*** | -0.144*** | -0.098*** | 0.060*** |
| 2011 | -0.212*** | -0.307*** | -0.094*** | -0.241*** | -0.312*** | -0.071*** | -0.241*** | -0.312*** | -0.045*** |
| 2012 | -0.170*** | -0.199*** | -0.029** | -0.190*** | -0.188*** | 0.002 | -0.190*** | -0.188*** | 0.022* |
| 2013 | -0.126*** | -0.120*** | 0.005 | -0.135*** | -0.088*** | 0.047*** | -0.135*** | -0.088*** | 0.068*** |
| 2014 | -0.168*** | -0.268*** | -0.100*** | -0.188*** | -0.276*** | -0.089*** | -0.188*** | -0.276*** | -0.058*** |
| 2015 | -0.085*** | -0.230*** | -0.145*** | -0.129*** | -0.272*** | -0.143*** | -0.129*** | -0.272*** | -0.121*** |
| 2016 | -0.150*** | -0.324*** | -0.175*** | -0.178*** | -0.359*** | -0.180*** | -0.178*** | -0.359*** | -0.149*** |
| 2017 | -0.100*** | -0.272*** | -0.172*** | -0.134*** | -0.305*** | -0.172*** | -0.134*** | -0.305*** | -0.130*** |
| 2018 | -0.090*** | -0.264*** | -0.174*** | -0.119*** | -0.278*** | -0.158*** | -0.119*** | -0.278*** | -0.135*** |
| 2019 | -0.096*** | -0.289*** | -0.193*** | -0.102*** | -0.271*** | -0.169*** | -0.102*** | -0.271*** | -0.131*** |
| 2020 | -0.232*** | -0.733*** | -0.501*** | -0.247*** | -0.803*** | -0.556*** | -0.247*** | -0.803*** | -0.547*** |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2021 | -0.195*** | -0.478*** | -0.284*** | -0.205*** | -0.508*** | -0.303*** | -0.205*** | -0.508*** | -0.290*** |
| 2022 | -0.208*** | -0.392*** | -0.185*** | -0.221*** | -0.422*** | -0.201*** | -0.221*** | -0.422*** | -0.168*** |

Note: *** p value < 1%; ** p value < 5%; * p value < 1%. Yo represents the young population, MA represents the middle-aged population, and El represents the elderly population. BHE estimations follow **Equation 35**.

**Table 6: Country-level Treatment Effects of Age**

| Country | Country Name | $CTE_{Yo-MA}^{Country}$ | $CTE_{MA-El}^{Country}$ | $CTE_{Yo-El}^{Country}$ | Style |
|---|---|---|---|---|---|
| AFG | Afghanistan | 0.026* | 0.185*** | 0.211*** | Younger Enjoy |
| AGO | Angola | 0.211*** | 0.576*** | 0.787*** | Younger Enjoy |
| ALB | Albania | 0.231*** | -0.105*** | 0.126*** | Middle-age Suffer |
| ARE | United Arab Emirates | 0.009 | -0.385*** | -0.376*** | Older Enjoy |
| ARG | Argentina | 0.271*** | 0.002 | 0.273*** | Younger Enjoy |
| ARM | Armenia | 0.201*** | 0.044*** | 0.245*** | Younger Enjoy |
| AUS | Australia | -0.311*** | -0.224*** | -0.535*** | Older Enjoy |
| AUT | Austria | 0.038*** | -0.037*** | 0.002 | Middle-age Suffer |
| AZE | Azerbaijan | 0.127*** | 0.117*** | 0.245*** | Younger Enjoy |
| BDI | Burundi | -0.022 | 0.077*** | 0.055*** | Younger Enjoy |
| BEL | Belgium | 0.028*** | -0.016 | 0.012 | No Trend |
| BEN | Benin | 0.017 | -0.115*** | -0.097*** | Older Enjoy |
| BFA | Burkina Faso | -0.049*** | -0.189*** | -0.237*** | Older Enjoy |
| BGD | Bangladesh | -0.086*** | -0.165*** | -0.250*** | Older Enjoy |
| BGR | Bulgaria | 0.090*** | 0.070*** | 0.161*** | Younger Enjoy |
| BHR | Bahrain | 0.035** | -0.196*** | -0.161*** | Middle-age Suffer |
| BIH | Bosnia and Herzegovina | 0.298*** | 0.129*** | 0.427*** | Younger Enjoy |
| BLR | Belarus | 0.135*** | 0.213*** | 0.349*** | Younger Enjoy |
| BLZ | Belize | -0.075 | -0.171*** | -0.246*** | Older Enjoy |
| BOL | Bolivia | 0.197*** | -0.301*** | -0.104*** | Middle-age Suffer |
| BRA | Brazil | 0.094*** | -0.258*** | -0.164*** | Middle-age Suffer |
| BTN | Bhutan | 0.02 | -0.030** | -0.01 | No Trend |
| BWA | Botswana | 0.257*** | 0.123*** | 0.380*** | Younger Enjoy |
| CAF | Central African Republic | 0.058*** | -0.100*** | -0.042** | Middle-age Suffer |
| CAN | Canada | -0.259*** | -0.163*** | -0.423*** | Older Enjoy |
| CHE | Switzerland | -0.040*** | -0.076*** | -0.116*** | Older Enjoy |
| CHL | Chile | 0.283*** | 0.095*** | 0.378*** | Younger Enjoy |
| CHN | China | 0.009 | -0.332*** | -0.323*** | Older Enjoy |
| CIV | Ivory Coast | 0.019 | -0.157*** | -0.138*** | Older Enjoy |
| CMR | Cameroon | 0.174*** | -0.061*** | 0.113*** | Middle-age Suffer |
| COD | Democratic Republic of the Congo | 0.070*** | -0.222*** | -0.152*** | Middle-age Suffer |

| | | | | | |
|---|---|---|---|---|---|
| COG | Republic of the Congo | 0.011 | 0.108*** | 0.119*** | Younger Enjoy |
| COL | Colombia | 0.072*** | -0.025* | 0.046*** | Middle-age Suffer |
| COM | Comoros | 0.007 | 0.01 | 0.017 | No Trend |
| CRI | Costa Rica | 0.031** | -0.054*** | -0.023* | Middle-age Suffer |
| CYP | Cyprus | 0.178*** | -0.119*** | 0.059*** | Middle-age Suffer |
| CZE | Czechia | 0.212*** | 0.170*** | 0.382*** | Younger Enjoy |
| DEU | Germany | 0.122*** | 0.025*** | 0.146*** | Younger Enjoy |
| DNK | Denmark | -0.072*** | -0.253*** | -0.325*** | Older Enjoy |
| DOM | Dominican Republic | 0.412*** | -0.357*** | 0.055*** | Middle-age Suffer |
| DZA | Algeria | 0.002 | 0.057*** | 0.059*** | Younger Enjoy |
| ECU | Ecuador | 0.341*** | 0.098*** | 0.439*** | Younger Enjoy |
| EGY | Egypt | 0.112*** | -0.151*** | -0.039*** | Middle-age Suffer |
| ESP | Spain | 0.098*** | 0.059*** | 0.157*** | Younger Enjoy |
| EST | Estonia | 0.147*** | 0.006 | 0.152*** | Younger Enjoy |
| ETH | Ethiopia | 0.001 | -0.130*** | -0.129*** | Older Enjoy |
| FIN | Finland | -0.019* | 0.086*** | 0.067*** | Middle-age Enjoy |
| FRA | France | 0.092*** | -0.006 | 0.087*** | Younger Enjoy |
| GAB | Gabon | 0.184*** | 0.241*** | 0.425*** | Younger Enjoy |
| GBR | United Kingdom | -0.027*** | -0.136*** | -0.163*** | Older Enjoy |
| GEO | Georgia | 0.171*** | -0.029* | 0.141*** | Middle-age Suffer |
| GHA | Ghana | -0.014 | -0.011 | -0.025* | No Trend |
| GIN | Guinea | -0.062*** | -0.097*** | -0.159*** | Older Enjoy |
| GMB | Gambia | -0.084*** | 0.397*** | 0.312*** | Middle-age Enjoy |
| GRC | Greece | 0.375*** | 0.119*** | 0.494*** | Younger Enjoy |
| GTM | Guatemala | 0.127*** | -0.102*** | 0.025* | Middle-age Suffer |
| HKG | Hong Kong | -0.113*** | 0.024 | -0.089*** | Older Enjoy |
| HND | Honduras | 0.263*** | -0.139*** | 0.124*** | Middle-age Suffer |
| HRV | Croatia | 0.260*** | 0.257*** | 0.517*** | Younger Enjoy |
| HTI | Haiti | 0.101*** | 0.001 | 0.102*** | Younger Enjoy |
| HUN | Hungary | 0.233*** | 0.052*** | 0.286*** | Younger Enjoy |

| | | | | | |
|---|---|---|---|---|---|
| IDN | Indonesia | 0.045*** | -0.346*** | -0.301*** | Middle-age Suffer |
| IND | India | -0.012* | -0.087*** | -0.099*** | Older Enjoy |
| IRL | Ireland | 0.049*** | -0.214*** | -0.165*** | Middle-age Suffer |
| IRN | Iran | 0.294*** | -0.053*** | 0.241*** | Middle-age Suffer |
| IRQ | Iraq | 0.019 | -0.042*** | -0.024* | Older Enjoy |
| ISL | Iceland | -0.044*** | -0.058*** | -0.102*** | Older Enjoy |
| ISR | Israel | 0.177*** | 0.048*** | 0.226*** | Younger Enjoy |
| ITA | Italy | 0.189*** | 0.078*** | 0.268*** | Younger Enjoy |
| JAM | Jamaica | -0.081*** | -0.244*** | -0.325*** | Older Enjoy |
| JOR | Jordan | 0.094*** | -0.106*** | -0.011 | Middle-age Suffer |
| JPN | Japan | -0.149*** | 0.019 | -0.131*** | Older Enjoy |
| KAZ | Kazakhstan | -0.073*** | -0.188*** | -0.261*** | Older Enjoy |
| KEN | Kenya | 0.122*** | 0.143*** | 0.265*** | Younger Enjoy |
| KGZ | Kyrgyzstan | 0.017 | -0.185*** | -0.169*** | Older Enjoy |
| KHM | Cambodia | -0.160*** | -0.294*** | -0.454*** | Older Enjoy |
| KOR | South Korea | 0.354*** | -0.036** | 0.319*** | Middle-age Suffer |
| KWT | Kuwait | -0.089*** | -0.395*** | -0.484*** | Older Enjoy |
| LAO | Laos | -0.160*** | -0.075*** | -0.236*** | Older Enjoy |
| LBN | Lebanon | 0.266*** | -0.068*** | 0.198*** | Middle-age Suffer |
| LBR | Liberia | 0.01 | -0.041* | -0.031 | No Trend |
| LBY | Libya | 0.007 | -0.426*** | -0.418*** | Older Enjoy |
| LKA | Sri Lanka | -0.136*** | -0.148*** | -0.284*** | Older Enjoy |
| LSO | Lesotho | 0.392*** | -0.026 | 0.366*** | Younger Enjoy |
| LTU | Lithuania | 0.153*** | 0.063*** | 0.215*** | Younger Enjoy |
| LUX | Luxembourg | 0.064*** | -0.054*** | 0.01 | Middle-age Suffer |
| LVA | Latvia | 0.139*** | -0.02 | 0.120*** | Younger Enjoy |
| MAR | Morocco | -0.081*** | -0.163*** | -0.245*** | Older Enjoy |
| MDA | Moldova | 0.230*** | 0.219*** | 0.449*** | Younger Enjoy |
| MDG | Madagascar | -0.01 | -0.104*** | -0.114*** | Older Enjoy |
| MDV | Maldives | 0.082* | -0.299*** | -0.217*** | Middle-age Suffer |
| MEX | Mexico | 0.075*** | 0.236*** | 0.311*** | Younger Enjoy |
| MKD | North Macedonia | 0.251*** | -0.029* | 0.222*** | Middle-age Suffer |

| | | | | | |
|---|---|---|---|---|---|
| MLI | Mali | -0.041*** | 0.095*** | 0.054*** | Middle-age Enjoy |
| MLT | Malta | 0.115*** | -0.086*** | 0.029** | Middle-age Suffer |
| MMR | Myanmar | -0.302*** | -0.339*** | -0.641*** | Older Enjoy |
| MNE | Montenegro | 0.276*** | -0.062*** | 0.214*** | Middle-age Suffer |
| MNG | Mongolia | -0.240*** | -0.123*** | -0.363*** | Older Enjoy |
| MOZ | Mozambique | 0.056** | 0.170*** | 0.226*** | Younger Enjoy |
| MRT | Mauritania | -0.125*** | -0.054* | -0.179*** | Older Enjoy |
| MUS | Mauritius | -0.055*** | -0.376*** | -0.431*** | Older Enjoy |
| MWI | Malawi | 0.259*** | -0.166*** | 0.093*** | Middle-age Suffer |
| MYS | Malaysia | 0.049*** | -0.032*** | 0.017 | Middle-age Suffer |
| NAM | Namibia | 0.062*** | 0.104*** | 0.166*** | Younger Enjoy |
| NER | Niger | -0.043*** | -0.156*** | -0.199*** | Older Enjoy |
| NGA | Nigeria | -0.044*** | 0.224*** | 0.181*** | Middle-age Enjoy |
| NIC | Nicaragua | 0.353*** | -0.286*** | 0.067*** | Middle-age Suffer |
| NLD | Netherlands | -0.001 | -0.086*** | -0.087*** | Older Enjoy |
| NOR | Norway | -0.050*** | -0.139*** | -0.190*** | Older Enjoy |
| NPL | Nepal | 0.035*** | -0.258*** | -0.223*** | Middle-age Suffer |
| NZL | New Zealand | -0.193*** | -0.301*** | -0.494*** | Older Enjoy |
| PAK | Pakistan | -0.042*** | 0.150*** | 0.108*** | Middle-age Enjoy |
| PAN | Panama | 0.149*** | 0.001 | 0.150*** | Younger Enjoy |
| PER | Peru | 0.026** | 0.092*** | 0.118*** | Younger Enjoy |
| PHL | Philippines | 0.063*** | -0.295*** | -0.232*** | Middle-age Suffer |
| POL | Poland | 0.163*** | -0.102*** | 0.061*** | Middle-age Suffer |
| PRI | Puerto Rico | -0.044 | 0.238*** | 0.195*** | Younger Enjoy |
| PRT | Portugal | 0.238*** | 0.038** | 0.276*** | Younger Enjoy |
| PRY | Paraguay | 0.269*** | -0.039*** | 0.230*** | Middle-age Suffer |
| PSE | West Bank and Gaza | 0.235*** | -0.063*** | 0.172*** | Middle-age Suffer |
| QAT | Qatar | -0.159*** | 0.029 | -0.130*** | Older Enjoy |

| | | | | | |
|---|---|---|---|---|---|
| ROU | Romania | 0.316*** | -0.033* | 0.283*** | Middle-age Suffer |
| RUS | Russia | 0.194*** | 0.109*** | 0.303*** | Younger Enjoy |
| RWA | Rwanda | 0.028** | 0.104*** | 0.131*** | Younger Enjoy |
| SAU | Saudi Arabia | -0.241*** | -0.236*** | -0.477*** | Older Enjoy |
| SDN | Sudan | -0.001 | 0.030* | 0.029* | Younger Enjoy |
| SEN | Senegal | 0.078*** | -0.176*** | -0.097*** | Middle-age Suffer |
| SGP | Singapore | -0.056*** | -0.092*** | -0.148*** | Older Enjoy |
| SLE | Sierra Leone | -0.196*** | 0.425*** | 0.229*** | Middle-age Enjoy |
| SLV | El Salvador | 0.334*** | -0.166*** | 0.167*** | Middle-age Suffer |
| SOM | Somalia | 0.386*** | 0.071*** | 0.457*** | Younger Enjoy |
| SRB | Serbia | 0.295*** | 0.021 | 0.317*** | Younger Enjoy |
| SSD | South Sudan | 0.194*** | -0.342*** | -0.148*** | Middle-age Suffer |
| SUR | Suriname | -0.112** | 0.103*** | -0.009 | Middle-age Enjoy |
| SVK | Slovakia | 0.166*** | 0.129*** | 0.295*** | Younger Enjoy |
| SVN | Slovenia | 0.366*** | -0.083*** | 0.284*** | Middle-age Suffer |
| SWE | Sweden | -0.200*** | -0.109*** | -0.309*** | Older Enjoy |
| SWZ | Eswatini | 0.330*** | -0.140*** | 0.190*** | Middle-age Suffer |
| SYR | Syria | -0.103*** | 0.192*** | 0.089*** | Middle-age Enjoy |
| TCD | Chad | -0.013 | -0.622*** | -0.635*** | Older Enjoy |
| TGO | Togo | -0.106*** | 0.089*** | -0.017 | Middle-age Enjoy |
| THA | Thailand | 0.158*** | -0.242*** | -0.084*** | Middle-age Suffer |
| TJK | Tajikistan | 0.082*** | 0.098*** | 0.180*** | Younger Enjoy |
| TKM | Turkmenistan | -0.066*** | 0 | -0.066*** | Older Enjoy |
| TTO | Trinidad and Tobago | -0.448*** | 0.073 | -0.375*** | Older Enjoy |
| TUN | Tunisia | 0.012 | -0.171*** | -0.159*** | Older Enjoy |
| TUR | Turkey | 0.063*** | -0.109*** | -0.047*** | Middle-age Suffer |
| TWN | Taiwan | -0.019 | -0.106*** | -0.124*** | Older Enjoy |
| TZA | Tanzania | -0.138*** | -0.060*** | -0.197*** | Older Enjoy |
| UGA | Uganda | 0.012 | 0.105*** | 0.117*** | Younger Enjoy |

| | | | | | |
|---|---|---|---|---|---|
| UKR | Ukraine | 0.242*** | 0.173*** | 0.415*** | Younger Enjoy |
| URY | Uruguay | 0.194*** | -0.091*** | 0.102*** | Middle-age Suffer |
| USA | United States of America | -0.125*** | -0.262*** | -0.387*** | Older Enjoy |
| UZB | Uzbekistan | -0.272*** | -0.669*** | -0.941*** | Older Enjoy |
| VEN | Venezuela | 0.137*** | 0.047** | 0.184*** | Younger Enjoy |
| VNM | Vietnam | 0.043*** | -0.154*** | -0.111*** | Middle-age Suffer |
| XKX | Kosovo | 0.188*** | -0.136*** | 0.051*** | Middle-age Suffer |
| XNC | | 0.045** | -0.453*** | -0.408*** | Middle-age Suffer |
| XNK | | -0.02 | -0.01 | -0.03 | No Trend |
| XSR | | 0.213*** | 0.058*** | 0.271*** | Younger Enjoy |
| YEM | Yemen | 0.178*** | -0.081*** | 0.098*** | Middle-age Suffer |
| ZAF | South Africa | -0.031** | -0.178*** | -0.209*** | Older Enjoy |
| ZMB | Zambia | -0.091*** | 0.238*** | 0.147*** | Middle-age Enjoy |
| ZWE | Zimbabwe | 0.004 | -0.163*** | -0.159*** | Older Enjoy |

Note: *** p value < 1%; ** p value < 5%; * p value < 1%. Yo represents the young population, MA represents the middle-aged population, and El represents the elderly population. Country-level treatment effects estimations follow **Equation 37**.

## References


1 Diener, E., Oishi, S. & Tay, L. Advances in subjective well-being research. *Nature Human Behaviour* **2**, 253-260 (2018). https://doi.org:10.1038/s41562-018-0307-6

2 Diener, E. & Chan, M. Y. Happy People Live Longer: Subjective Well-Being Contributes to Health and Longevity. *Applied Psychology: Health and Well-Being* **3**, 1-43 (2011). https://doi.org:10.1111/j.1758-0854.2010.01045.x

3 López Ulloa, B. F., Møller, V. & Sousa-Poza, A. How Does Subjective Well-Being Evolve with Age? A Literature Review. *Journal of Population Ageing* **6**, 227-246 (2013). https://doi.org:10.1007/s12062-013-9085-0

4 Stone, A. A., Schwartz, J. E., Broderick, J. E. & Deaton, A. A snapshot of the age distribution of psychological well-being in the United States. *Proceedings of the National Academy of Sciences* **107**, 9985-9990 (2010). https://doi.org:10.1073/pnas.1003744107

5 Margolis, R. & Myrskylä, M. A Global Perspective on Happiness and Fertility. *Population and Development Review* **37**, 29-56 (2011). https://doi.org:10.1111/j.1728-4457.2011.00389.x

6 Nomaguchi, K. & Milkie, M. A. Parenthood and Well-Being: A Decade in Review. *Journal of Marriage and Family* **82**, 198-223 (2020). https://doi.org:https://doi.org/10.1111/jomf.12646

7 Wu, W., Wu, H., Wu, X., Gu, J. & Qi, X. A Meta-Analysis of Life Satisfaction's Association with Cognitive Reappraisal and Expressive Suppression: The Influences of Age, Gender, and Cultural Values. *J. Happiness Stud.* **25**, 41 (2024).

8 Blanchflower, D. G. Is happiness U-shaped everywhere? Age and subjective well-being in 145 countries. *Journal of Population Economics* **34**, 575-624 (2021). https://doi.org:10.1007/s00148-020-00797-z

9 Blanchflower, D. G. & Piper, A. There is a mid-life low in well-being in Germany. *Economics Letters* **214**, 110430 (2022). https://doi.org:https://doi.org/10.1016/j.econlet.2022.110430

10 Steptoe, A., Deaton, A. & Stone, A. A. Subjective wellbeing, health, and ageing. *The Lancet* **385**, 640-648 (2015). https://doi.org:10.1016/s0140-6736(13)61489-0

11 Van Landeghem, B. A test for the convexity of human well-being over the life cycle: Longitudinal evidence from a 20-year panel. *Journal of Economic Behavior & Organization* **81**, 571-582 (2012). https://doi.org:10.1016/j.jebo.2011.08.001

12 Frijters, P. & Beatton, T. The mystery of the U-shaped relationship between happiness and age. *Journal of Economic Behavior & Organization* **82**, 525-542 (2012). https://doi.org:https://doi.org/10.1016/j.jebo.2012.03.008

13 Blanchflower, D. G. & Graham, C. L. The Mid-Life Dip in Well-Being: a Critique. *Soc. Indic. Res.* **161**, 287-344 (2022). https://doi.org:10.1007/s11205-021-02773-w

14 Easterlin, R. A. Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology* **27**, 463-482 (2006). https://doi.org:https://doi.org/10.1016/j.joep.2006.05.002

15 García Mata, O. At what age do Mexicans suffer the most financial stress? *Journal of Economics, Finance and Administrative Science* (2023). https://doi.org:10.1108/jefas-04-2023-0087

16 Plagnol, A. C. Financial satisfaction over the life course: The influence of assets and liabilities. *Journal of Economic Psychology* **32**, 45-64 (2011). https://doi.org:https://doi.org/10.1016/j.joep.2010.10.006

17 Hansen, T. & Blekesaune, M. The age and well-being "paradox": a longitudinal and multidimensional reconsideration. *European Journal of Ageing* **19**, 1277-1286 (2022). https://doi.org:10.1007/s10433-022-00709-y

18 Blanchflower, D. G. & Oswald, A. J. Is well-being U-shaped over the life cycle? *Social Science & Medicine* **66**, 1733-1749 (2008). https://doi.org:10.1016/j.socscimed.2008.01.030

19 Sutin, A. R. *et al.* The Effect of Birth Cohort on Well-Being. *Psychol. Sci.* **24**, 379-385 (2013). https://doi.org:10.1177/0956797612459658

20 Shu, X., Chen, J. & Zhu, Y. Changing times and subjective well-being in urban China 2003–2013: An age-period-cohort approach. *Chinese Journal of Sociology* **9**, 321-354 (2023). https://doi.org:10.1177/2057150x231180022

21 Gerstorf, D. *et al.* Cohort differences in adult-life trajectories of internal and external control beliefs: A tale of more and better maintained internal control and fewer external constraints. *Psychology and Aging* **34**, 1090 (2019).

22 Li, C. & Managi, S. Income raises human well-being indefinitely, but age consistently slashes it. *Scientific Reports* **13** (2023). https://doi.org:10.1038/s41598-023-33235-7

23 Bamidis, P. D. *et al.* A review of physical and cognitive interventions in aging. *Neuroscience & Biobehavioral Reviews* **44**, 206-220 (2014). https://doi.org:https://doi.org/10.1016/j.neubiorev.2014.03.019

24 Luu, J. & Palczewski, K. Human aging and disease: Lessons from age-related macular degeneration. *Proceedings of the National Academy of Sciences* **115**, 2866-2872 (2018). https://doi.org:10.1073/pnas.1721033115

25 Kassie, M., Ndiritu, S. W. & Stage, J. What Determines Gender Inequality in Household Food Security in Kenya? Application of Exogenous Switching Treatment Regression. *World Development* **56**, 153-171 (2014). https://doi.org:https://doi.org/10.1016/j.worlddev.2013.10.025

26 Balcilar, M., Gupta, R. & Miller, S. M. Regime switching model of US crude oil and stock market prices: 1859 to 2013. *Energy Economics* **49**, 317-327 (2015). https://doi.org:https://doi.org/10.1016/j.eneco.2015.01.026

27 Chen, T. & Guestrin, C. in *22nd SIGKDD Conference on Knowledge Discovery and Data Mining* (ACM).

28 Bentéjac, C., Csörgő, A. & Martínez-Muñoz, G. A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review* **54**, 1937-1967 (2021). https://doi.org:10.1007/s10462-020-09896-5

29 Diener, E., Tay, L. & Oishi, S. Rising Income and the Subjective Well-Being of Nations. *J. Pers. Soc. Psychol.* **104**, 267-276 (2013). https://doi.org:10.1037/a0030487

30 Jebb, A. T., Tay, L., Diener, E. & Oishi, S. Happiness, income satiation and turning points around the world. *Nature Human Behaviour* **2**, 33-38 (2018). https://doi.org:10.1038/s41562-017-0277-0

31 Kobau, R., Sniezek, J., Zack, M. M., Lucas, R. E. & Burns, A. Well-Being Assessment: An Evaluation of Well-Being Scales for Public Health and Population Estimates of Well-Being among US Adults. *Applied Psychology: Health and Well-Being* **2**, 272-297 (2010). https://doi.org:10.1111/j.1758-0854.2010.01035.x

32 Fredrickson, B. L. *et al.* A functional genomic perspective on human well-being. *Proceedings of the National Academy of Sciences* **110**, 13684-13689 (2013). https://doi.org:10.1073/pnas.1305419110

33 Oswald, A. J. & Wu, S. Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the USA. *Science* **327**, 576-579 (2010). https://doi.org:10.1126/science.1180606

34 Diener, E. Subjective well-being. *Psychol. Bull.* **95**, 542-575 (1984). https://doi.org:10.1037/0033-2909.95.3.542

35 Helliwell, J. F. & Aknin, L. B. Expanding the social science of happiness. *Nature Human Behaviour* **2**, 248-252 (2018). https://doi.org:10.1038/s41562-018-0308-5

36 Kahneman, D. & Deaton, A. High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences* **107**, 16489-16493 (2010). https://doi.org:10.1073/pnas.1011492107

37 Lachman, M. E. Mind the gap in the middle: A call to study midlife. *Research in human development* **12**, 327-334 (2015).

38 Giuntella, O. *et al.* The Midlife Crisis. *Economica* **90**, 65-110 (2023). https://doi.org:10.1111/ecca.12452

39 Molnar, C. *Interpretable machine learning*. (Lulu. com, 2020).

40 Molnar, C., Casalicchio, G. & Bischl, B. 417-431 (Springer International Publishing, 2020).

41 Lundberg, S. M. *et al.* From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence* **2**, 56-67 (2020). https://doi.org:10.1038/s42256-019-0138-9

42 Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V. & Gulin, A. in *Advances in neural information processing systems* Vol. 31 (2018).

43 Ke, G. *et al.* in *Advances in neural information processing systems* Vol. 30 (2017).

44 Turner, R. *et al.* in *NeurIPS 2020 Competition and Demonstration Track.*3-26 (PMLR).

45 Krueger, P. M. & Friedman, E. M. Sleep Duration in the United States: A Cross-sectional Population-based Study. *American Journal of Epidemiology* **169**, 1052-1063 (2009). https://doi.org:10.1093/aje/kwp023

46 Bussière, C., Sirven, N. & Tessier, P. Does ageing alter the contribution of health to subjective well-being? *Social Science & Medicine* **268**, 113456 (2021). https://doi.org:https://doi.org/10.1016/j.socscimed.2020.113456