

When National Happiness and Suicide Mortality Diverge: Explainable Machine Learning, Anomaly Detection, and Panel Evidence from 161 Countries

Young-Chool Choi

`ycchoi@assist.ac.kr`

Seoul School of Integrated Sciences and Technologies

Research Article

Keywords: subjective well-being, suicide mortality, explainable artificial intelligence, XGBoost, anomaly detection, Korea

Posted Date: October 8th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10746095/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

When National Happiness and Suicide Mortality Diverge: Explainable Machine Learning, Anomaly Detection, and Panel Evidence from 161 Countries

Original Paper

Special Issue: Artificial Intelligence, Well-Being and Quality of Life

Anonymous manuscript

Abstract

National life evaluation and suicide mortality both inform assessments of population well-being, yet their national patterns do not always align. We treat this divergence as country-level statistical discordance, not as evidence of an individual happiness-suicide paradox. The analysis links annual Life Ladder scores, World Health Organization (WHO) age-standardized suicide rates, and World Bank indicators, yielding 2,044 country-year observations for 161 countries during 2005-2021. We combine two-way fixed-effects estimation with a country-held-out additive spline model, eXtreme Gradient Boosting (XGBoost), Isolation Forest, SHapley Additive exPlanations (SHAP), and Gaussian-mixture country profiles. To extend the descriptive record beyond 2021, WHO Mortality Database rates for 2022-2023 are examined separately from the comparable Global Health Estimates series. Within-country increases in unemployment and alcohol consumption were associated with higher suicide mortality and greater residual discordance. In country-held-out tests, XGBoost predicted residual discordance more accurately than the spline GAM ($R^2 = 0.243$ versus 0.090); alcohol consumption, urbanization, and income contributed most to its predictions. Lithuania, the Russian Federation, Kazakhstan, and the Republic of Korea had the most persistent high-happiness/high-suicide profiles under the global benchmark. Korea's position depended on the comparison group: in 2021, happiness stood at the 63rd global percentile but the 19th OECD percentile, while suicide mortality stood at the 97th global percentile and was the highest among OECD members in the analytical sample. The results favor joint monitoring of subjective well-being and preventable mortality and caution against conflating predictive explanation, causal inference, and individual risk.

Keywords: subjective well-being; suicide mortality; explainable artificial intelligence; XGBoost; anomaly detection; Korea

1. Introduction

Subjective well-being now has a visible place in public policy. National statistical offices, the Organisation for Economic Co-operation and Development (OECD), the United Nations system, and the World Happiness Report use it alongside income, employment, and health. The rationale is compelling: people's evaluations of their own lives reveal aspects of welfare that conventional economic measures overlook. Even so, a national mean cannot describe the entire distribution of well-being. Serious harm may remain concentrated in groups that barely affect the average. Suicide mortality is therefore a demanding test of whether aggregate happiness, on its own, is an adequate social dashboard.

At the individual level, the expected relationship is well established. Persistent dissatisfaction with life, mental disorder, hopelessness, social disconnection, and adverse life events are associated with suicidal behavior; longitudinal evidence also links life dissatisfaction to later suicide (Koivumaa-Honkanen et al., 2001; Turecki et al., 2019). The population-level relationship is less direct. Life evaluation and suicide mortality are drawn from different measurement systems, describe different parts of a distribution, and may respond to social and institutional change at different speeds. One is a survey mean. The other is a rare, extreme, and partly under-recorded event. A favorable national evaluation can therefore coexist with a high-risk tail that the mean represents poorly.

This coexistence is often discussed as a happiness-suicide paradox. Across European countries, Bray and Gunnell (2006) reported an inverse association between national happiness or life satisfaction and suicide, whereas Helliwell (2007) showed how social capital helps reconcile evidence from the two indicators. Daly, Oswald, Wilson, and Wu (2011) identified a different pattern: some happier places also had higher suicide rates, a finding they interpreted partly in terms of relative comparison. The paradox label is useful shorthand, but it can blur two distinctions. Associations between national means do not identify individual risk, and a country that appears happy in a global comparison may rank poorly among high-income peers. Both distinctions matter for the Korean case.

Korea is especially informative because rapid economic development and broad improvements in living conditions have not been accompanied by comparably low suicide mortality. OECD Health at a Glance 2025 reports an age-standardized rate of 23.2 deaths per 100,000 in 2023, against an OECD average of 10.7 (OECD, 2025). The World Happiness Report 2026, meanwhile, gives Korea a 2023-2025 life evaluation of 6.040 and rank 67 (Helliwell et al., 2026). These figures cannot be inserted uncritically into one panel: OECD, World Health Organization (WHO) Global Health Estimates (GHE), the WHO Mortality Database (MDB), and the World Happiness Report differ in definition,

adjustment, and reference period. They do, however, sharpen the central question: under which benchmark and structural conditions does Korea appear discordant?

Artificial-intelligence methods are useful here only when they complement population-health reasoning. eXtreme Gradient Boosting (XGBoost) accommodates thresholds and interactions without imposing one linear functional form (Chen & Guestrin, 2016). Isolation Forest identifies unusual country-years without relying on a single substantive cutoff, while SHapley Additive exPlanations (SHAP) decomposes fitted predictions for the full sample and for Korea. Each method also creates risks. Repeated country observations can leak across training and validation sets, and predictive attributions can be mistaken for causal effects. To address these concerns, the analysis combines machine learning with two-way fixed effects, country-grouped cross-validation, a temporal holdout, sensitivity tests, and explicit limits on ecological interpretation.

Three questions guide the analysis. Which countries repeatedly combine above-typical happiness with above-typical suicide mortality? To what extent is the country-level mismatch predicted by social support, inequality, unemployment, alcohol consumption, mental-health resources, population ageing, and related conditions? Finally, which country profile contains Korea, and does that classification change when the reference set moves from the world to OECD members or to countries with recent reported mortality data?

The study makes four related contributions. It replaces a broad paradox claim with two reproducible country-level outcomes: a directional standardized discordance index and a country-held-out residual mismatch. It then brings panel estimation, nonlinear prediction, anomaly detection, explainable AI, and probabilistic country profiling into a single workflow designed to prevent country leakage. The empirical analysis covers 161 countries through the latest globally comparable suicide year and adds newer reported observations without mixing incompatible statistical series. Most importantly, the results separate two features of the Korean case: its unusually high suicide mortality is robust, whereas the description of Korea as a high-happiness country depends on the comparison group.



Aggregate discordance is not an individual-level happiness–suicide paradox.

Fig. 1 Analytical framework

Note. The study examines country-level discordance and does not infer an individual-level happiness–suicide relationship.

2. Theoretical and Analytical Framework

2.1 Two indicators, different distributions

Life evaluation and suicide mortality convey different kinds of welfare information. The Cantril Ladder asks respondents to locate their current life between the worst possible life (0) and the best possible life (10) (Cantril, 1965). A national average describes the central tendency of reflective well-being, a construct distinct from affective experience (Kahneman & Deaton, 2010). Suicide mortality records something quite different: the realized endpoint of processes involving vulnerability, acute stress, access to means, social context, and the response of services. It is neither a survey measure nor a comprehensive measure of mental health. Relatively favorable average life evaluation may coexist with high suicide mortality when distress is concentrated, well-being is unequally distributed, or high-risk groups differ demographically and socially from the majority.

The key distinction is between a mean and a tail. Let H_{it} represent mean national life evaluation and S_{it} the age-standardized suicide rate in country i and year t . Even if lower individual happiness increases individual suicide risk, the observed pair (H_{it}, S_{it}) says nothing directly about the joint distribution of well-being and risk within that country. Aggregate correlations may diverge from individual correlations because they omit within-group covariance and compositional detail (Morgenstern, 1995; Robinson, 1950). For this reason, we refer to the outcomes as national

discordance and residual mismatch. The term paradox is retained only when discussing the literature, not as an individual-level causal claim.

The measurement systems introduce further differences. Survey estimates reflect sampling error, cultural response styles, and variation in field conditions. Suicide counts depend on death-registration completeness, cause-of-death certification, stigma, legal context, and the redistribution of ill-defined causes. WHO GHE uses modelling and consultation with member states to improve comparability. WHO MDB retains reported mortality, which gives later observations for some countries but less uniform coverage. Treating the two series as interchangeable would therefore create a false appearance of continuity.

2.2 Why discordance may arise

Social integration is the first candidate mechanism. Durkheim placed integration and regulation at the center of population differences in suicide (Durkheim, 1951), and contemporary well-being research consistently finds that having someone to rely on is closely related to life evaluation, including during crises (Helliwell & Putnam, 2004; Helliwell et al., 2014). Yet support need not influence the two aggregate indicators in parallel. Strong in-group ties can coexist with exclusion, status pressure, loneliness among outsiders, or reluctance to disclose distress. National social support should generally predict less discordance, although its within-country coefficient may be weak if the measure changes slowly or misses high-risk groups.

Economic and distributive stress provides a second pathway. Unemployment can remove income, identity, routine, and social contact, and meta-analytic evidence associates long-term unemployment with greater suicide risk (Milner et al., 2013). Income inequality may intensify relative comparisons even when average income and life evaluation are favorable. Daly et al. (2011, 2013) argue that comparison with happier or more affluent others can matter, although the mechanism is context-dependent and sits alongside other aggregate happiness puzzles (Deaton & Stone, 2013). GDP, inequality, unemployment, urbanization, and ageing may also interact in nonlinear ways. Good national averages do not preclude labor-market insecurity or relative deprivation among particular groups.

A third pathway combines behavioral exposure with institutional protection. Alcohol can increase impulsivity, worsen depression, damage relationships, and make a crisis more lethal. A recent meta-analysis associates each additional litre of per-capita consumption with a 3.59% increase in population suicide mortality, while emphasizing the ecological and observational basis of the estimate (Guo et al., 2025; see also Roche et al., 2018). Health spending and the mental-health workforce may reduce risk through access and continuity of care, but resources can also expand in response to high need.

Cross-sectional resource coefficients are consequently sensitive to reverse causation and timing. Institutional protection should thus be understood within the wider social determinants of mental health, not as service capacity alone (Lund et al., 2018).

Ageing and urbanization have no unambiguous expected signs. Older age groups may face high age-specific suicide risk, but age standardization removes the mechanical effect of population composition from the dependent variable. Any remaining association may reflect cohort experience, pension security, family structure, chronic illness, or service organization (Case & Deaton, 2015). Urban settings likewise combine better access to services and opportunity with possible losses of community connection, housing pressure, and isolation. These competing processes make flexible prediction useful, provided that prediction is not presented as causal explanation.

2.3 From association to a dual analytical strategy

Explanation and prediction overlap, but they are not the same task (Shmueli, 2010). Two-way fixed-effects models ask whether changes within a country covary with changes in suicide mortality or mismatch after common year shocks are removed. They control for stable country characteristics, but have little leverage over slowly moving variables and remain exposed to time-varying confounding. XGBoost and a spline generalized additive model (GAM) ask a different question: how accurately can observed covariate patterns predict mismatch in held-out data? Their functional flexibility is useful, although fitted dependence still does not identify an intervention effect.

SHAP describes how a fitted prediction is assembled. For each observation, it expresses the prediction as a baseline plus feature-specific contributions. This property is useful when comparing Korea with the global pattern, but attribution may be shared or redistributed among correlated features. Thus, a positive SHAP value for ageing means that Korea's observed value raised the model's predicted mismatch relative to its baseline under the learned joint distribution. It does not imply that increasing the older population would itself cause a rise in suicide mortality.

The hypotheses are deliberately limited to patterns that the design can assess. H1 predicts positive associations of unemployment and alcohol consumption with suicide mortality and residual discordance. H2 anticipates that a nonlinear model will outperform the additive spline model if relevant thresholds or interactions are present. H3 expects high-happiness/high-suicide anomalies to persist in a small group of countries rather than appear only in isolated years. H4 predicts that Korea's high suicide mortality will be robust but its happiness classification will vary with the benchmark.

3. Data and Methods

3.1 Data architecture and sample

The core dataset contains 2,044 country-year observations for 161 countries between 2005 and 2021. The panel is unbalanced because the Gallup World Poll did not survey every country in every year. Life evaluation, social support, log GDP per capita, healthy life expectancy, freedom to make life choices, generosity, perceived corruption, positive affect, and negative affect come from the World Happiness Report 2023 annual analytical file (Helliwell et al., 2023). The Life Ladder is the national mean response to the 0–10 Cantril question.

Age-standardized suicide rates for both sexes come from WHO GHE 2021 (World Health Organization [WHO], 2024). This release provides internally comparable country estimates for 2000–2021 and is not directly comparable with earlier GHE vintages. Sex-specific age-standardized rates and 2021 age-band crude rates come from the WHO Global Health Observatory (WHO, 2025b). World Bank World Development Indicators (WDI) supply unemployment, the Gini coefficient, alcohol consumption per adult, the population share aged 65 and older, urbanization, and population (World Bank, 2026). Current health expenditure is the WHO Global Health Expenditure Database series distributed through the World Bank and Our World in Data. Latest available psychiatrist density comes from the WHO Global Health Observatory and Mental Health Atlas (WHO, 2025a).

The main panel ends in 2021 because that is the latest year in the globally comparable WHO GHE suicide series. To respond to the need for newer evidence, a strictly separate extension merges WHO MDB reported suicide rates for 2022 and 2023 with the World Happiness Report 2026 rolling three-year life-evaluation average ending in the corresponding year (Helliwell et al., 2026). Fifty-five countries have a 2022 match and 29 have a 2023 match. Because MDB values are reported rather than GHE-modelled and country availability is selective, this extension is descriptive. The OECD 2023 value for Korea is used only as external context. Table 1 summarizes the architecture.

Table 1. Measurement and data architecture

Construct	Operational measure	Coverage	Source	Analytical use
Happiness	Annual mean Life Ladder, 0–10	2005–2021	World Happiness Report 2023	Core exposure; discordance
Suicide mortality	Age-standardized rate, both sexes, per 100,000	2005–2021	WHO GHE 2021	Core outcome
Later suicide	Reported age-standardized rate	2022–2023; selective	WHO Mortality Database	Descriptive extension only
Social relations	Social support; positive/negative affect	Annual as observed	World Happiness Report	Predictors

Construct	Operational measure	Coverage	Source	Analytical use
Economic conditions	Log GDP, unemployment, Gini	Annual/sparse	World Happiness Report; World Bank WDI	Predictors; Gini enrichment
Behavior and demography	Alcohol, age 65+, urbanization	Annual as available	World Bank WDI / WHO	Predictors
Health system	Health spending; latest psychiatrist density	Annual / latest cross- section	WHO expenditure / observatory data	Predictor / enrichment
Korea current context	WHR rank and OECD suicide statistic	2023–2025	WHR 2026; OECD 2025	External context; not pooled

GHE and MDB are separate statistical series. The extension is not used to refit the core models.

3.2 Outcomes and discordance measures

The first outcome is $\ln(S_{it} + 0.1)$, where S_{it} is the WHO age-standardized suicide rate. The small constant is immaterial at observed rates but makes the transformation defined. Log transformation reduces right skew and lets fixed-effects coefficients be read approximately as proportional changes for small values.

The directional discordance index is constructed within each year. Let $h_{it} = z_t(H_{it})$ and $s_{it} = z_t[\ln(S_{it} + 0.1)]$, where z_t standardizes across countries observed in year t . The index $D_{it} = h_{it} + s_{it}$ is positive when happiness and suicide are jointly high relative to that year's cross-section, and negative when they are jointly low. Its absolute value captures magnitude but not direction. The sum is deliberately transparent; it is not a clinical scale.

A second target isolates suicide mortality that is unusual conditional on happiness. A cubic-spline GAM predicts log suicide from Life Ladder and year. Predictions are generated out of fold with five country-grouped folds, so no observation from a validation country is used to fit that country's expected value. Residual mismatch is $R_{it} = \ln(S_{it} + 0.1) - \hat{g}_{-i}(H_{it}, t)$. Positive R_{it} means suicide mortality is higher than the cross-national happiness–year pattern predicts. This target is used for XGBoost, the multivariable GAM comparison, SHAP, and part of the anomaly detector.

3.3 Covariates and missing data

The pre-specified predictors are social support, log GDP per capita, unemployment, Gini inequality, alcohol consumption, the share aged 65 and older, current health expenditure as a share of GDP, urban population share, freedom, perceived corruption, positive affect, negative affect, and calendar year. Social support is the principal family/community proxy because a globally harmonized annual measure of family relationship quality is unavailable. Gini coverage is 65.5%, so inequality enters a reduced-sample fixed-effects specification and is retained by XGBoost with fold-specific imputation.

Short internal gaps in WDI variables are linearly interpolated within country for a maximum of two consecutive years; no extrapolation is performed at the beginning or end of a series. Fixed-effects models use complete observations after this limited interpolation. For machine learning, missing predictors are replaced with training-fold medians and those medians are applied to the validation fold. This prevents validation information from determining imputation values. Mental-health workforce data are not interpolated across the entire panel. The latest psychiatrist density is analyzed only in a country-level 2017–2021 enrichment model with heteroskedasticity-consistent standard errors.

3.4 Statistical and machine-learning procedures

Panel models include country and year fixed effects. Predictors are standardized within each estimation sample, and standard errors are clustered by country. Three nested specifications are reported: happiness alone; a social-economic model adding social support, log GDP, unemployment, alcohol, ageing, and health expenditure; and an inequality-enriched model adding Gini. The models estimate contemporaneous associations and are not claimed to identify causal effects.

Predictive models use the country-held-out residual mismatch as the outcome. XGBoost has 450 trees, depth 3, learning rate .03, minimum child weight 8, subsampling .82, column subsampling .80, L1 penalty .15, and L2 penalty 2.0. These conservative settings limit tree complexity in a modest panel. The comparator is an additive cubic-spline generalized additive model (GAM) with five knots per predictor and a ridge penalty selected over a log-spaced grid (Hastie & Tibshirani, 1990). Five-fold cross-validation (CV), implemented with GroupKFold, assigns entire countries to folds (Roberts et al., 2017). Target construction is nested: within every outer split, the happiness–year GAM is estimated only from outer-training countries, training residuals are generated by an inner country-grouped split, and validation residuals are generated from the outer-training fit. The temporal test likewise estimates both the residual benchmark and predictive models using 2005–2017 before evaluating 2018–2021. Performance is reported using root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2) against a training-mean baseline.

The final XGBoost fit is explained with Tree SHAP (Lundberg & Lee, 2017). Global importance is mean absolute SHAP. Korea-specific explanations average local contributions across its 16 observed years and also report the latest-year contribution. Because correlated features share predictive information, feature ordering is interpreted as a property of the fitted model and data, not as a ranking of intervention effects.

Isolation Forest uses year-standardized happiness, year-standardized log suicide, and residual mismatch. Six hundred trees are fitted with a pre-specified 7% contamination rate, with 5% and 10% sensitivity checks (Liu et al., 2008). Country

persistence is the fraction of observed years flagged as anomalous. A high–high share records the fraction of years with positive happiness and suicide z-scores. Their product identifies persistent anomalies in the high-happiness/high-suicide direction.

Country profiles average 2017–2021 values and require at least two observed years. Nine variables—happiness, suicide, social support, unemployment, Gini, alcohol, ageing, health expenditure, and urbanization—are median-imputed after aggregation and standardized. Gaussian mixture models with full covariance, 20 random starts, and two through six components are compared by Bayesian information criterion (BIC) (McLachlan & Peel, 2000). The retained three-component solution is descriptive rather than a claim that countries belong to immutable latent types.

Auxiliary analyses re-estimate fixed-effects models separately for male and female suicide rates, summarize 2021 crude rates by age group, restrict the panel to OECD members, World Bank high-income economies, and countries with at least ten WHO MDB annual observations, and estimate the mental-health-resource enrichment models. All analyses use random seed 20260818. The accompanying workbook (Online Resource 1) records row-level data, fold predictions, coefficients, SHAP values, anomaly scores, cluster assignments, data provenance, and SHA-256 checksums.

3.5 Computational transparency and artificial-intelligence assistance

OpenAI ChatGPT assisted with locating official sources, drafting code, organizing data, and refining the English language. All numerical statements were produced through the documented analytical workflow and checked against the processed tables and official-source files. The system is not an author and cannot assume responsibility for the research. Verification, interpretation, and approval of the submitted text remain solely with the human author(s).

4. Results

4.1 Descriptive pattern and Korea

The core panel's mean Life Ladder was 5.47 (SD = 1.12), and the mean age-standardized suicide rate was 9.60 per 100,000 (SD = 6.03). The median suicide rate was 8.81, confirming right skew. The directional discordance index ranged from -6.05 to 3.22, while residual mismatch ranged from -2.89 to 1.62. Table 2 reports the main distributions.

The 2021 cross-section shows why a bivariate average is insufficient (Figure 2). Countries occupy all four quadrants around the annual means, and the lowess curve does not impose a uniformly negative relationship. Korea had a Life Ladder of 6.113 and a GHE suicide rate of 20.575 per 100,000. This placed it at the 63rd happiness percentile and 97th

suicide percentile globally. Among OECD members, the same happiness value was at only the 19th percentile, whereas Korea had the highest suicide rate in the sample. Thus, Korea is globally a moderately high-happiness/high-suicide case but an OECD low-happiness/high-suicide case.

Korea's time series reinforces the asymmetry (Figure 3). Its GHE suicide rate declined from a peak near 29.6 in 2009 to 20.6 in 2021, yet remained far above the OECD and global unweighted means. Life evaluation fluctuated, including a high observation in 2011, but showed no commensurate sustained rise. Across 2017–2021, Korea's mean happiness was at the 62nd global percentile and 11th OECD percentile; mean suicide was at the 97th global percentile and highest in the OECD subset.

Table 2. Descriptive statistics for the core panel

Variable	N	Mean	SD	Min	Median	Max
Life Ladder	2,044	5.47	1.12	2.18	5.39	8.02
Suicide rate	2,044	9.60	6.03	0.61	8.81	36.77
Directional discordance	2,044	-0.00	1.43	-6.05	-0.00	3.22
Residual mismatch	2,044	0.00	0.74	-2.89	0.11	1.62
Social support	2,031	0.81	0.12	0.29	0.84	0.99
Unemployment (%)	2,044	7.65	5.70	0.12	5.86	35.23
Gini	1,338	36.64	7.48	23.20	35.20	62.50
Alcohol (litres)	1,895	6.04	4.31	0.00	5.78	18.34
Age 65+ (%)	2,044	8.98	6.28	0.86	6.43	29.22
Health expenditure (% GDP)	2,018	6.49	2.68	1.60	6.24	22.11

Country-year observations are unweighted. Gini has a smaller sample because it is not observed annually in many countries.

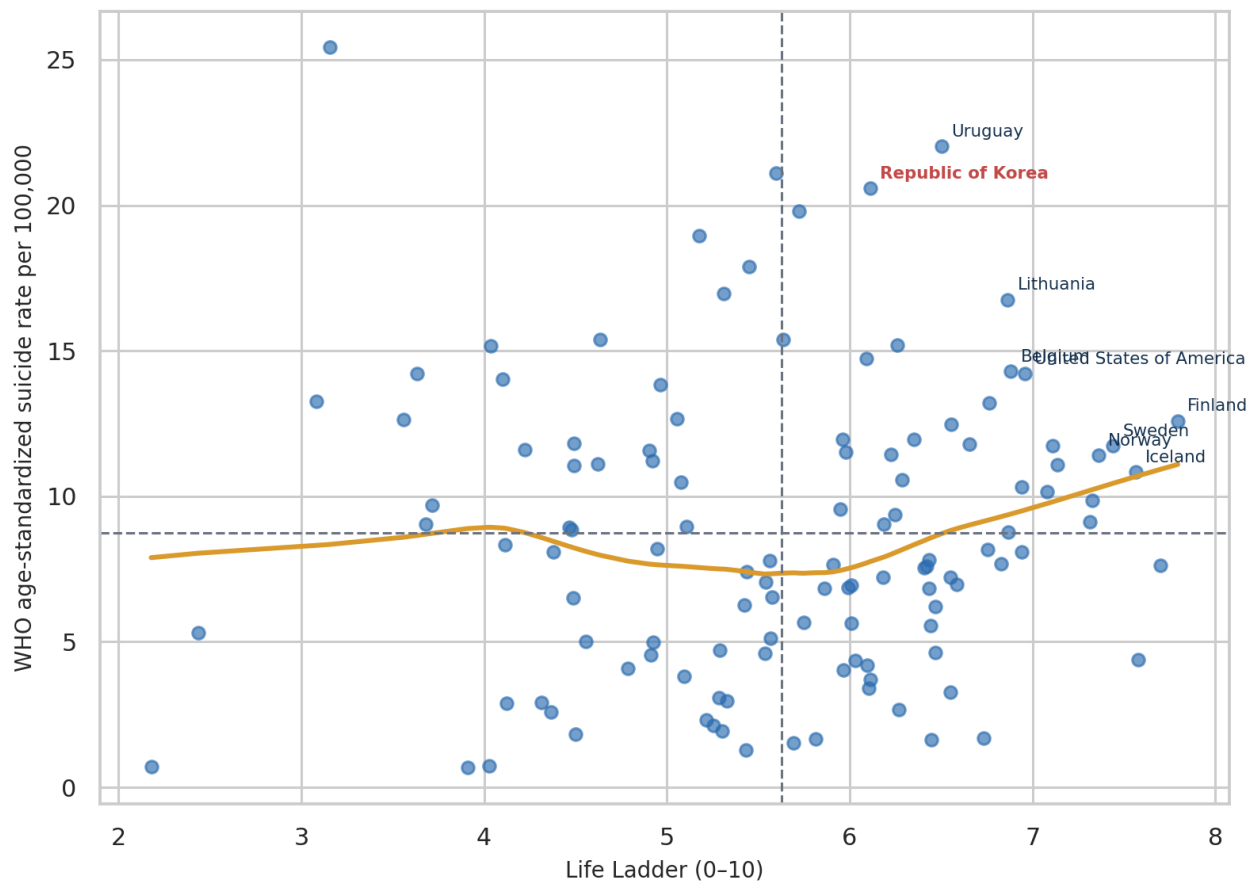


Fig. 2 Happiness and suicide mortality in 2021

Note. Dashed lines are sample means ($n = 119$ countries). The solid curve is a descriptive locally weighted smooth.

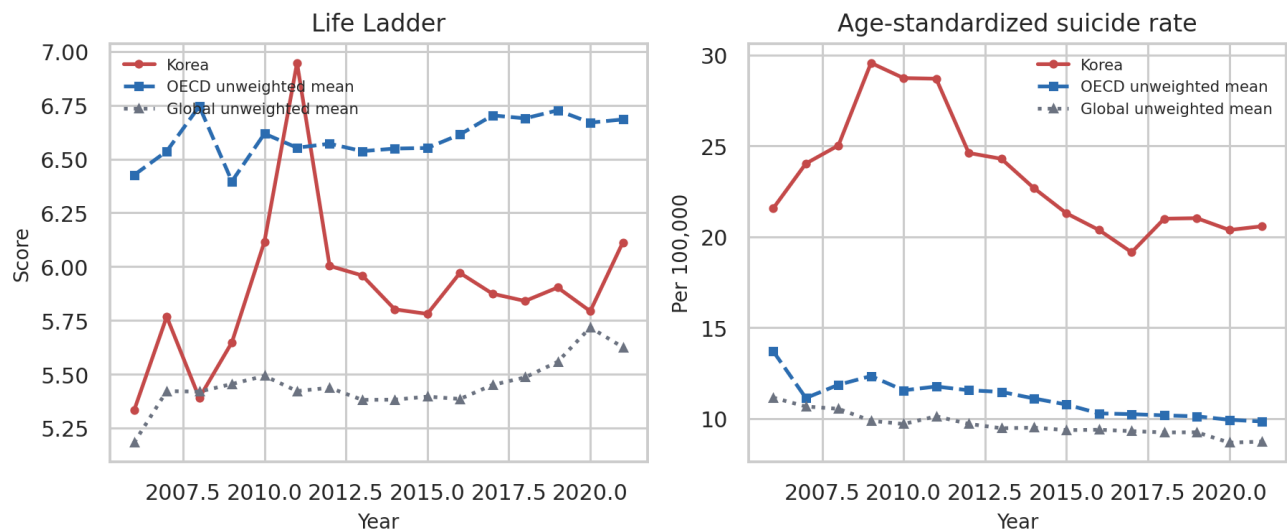


Fig. 3 Korea in global and OECD context, 2005–2021

Note. Means are unweighted across available countries; the panel is unbalanced.

4.2 Fixed-effects associations

In the bivariate fixed-effects model, a one-standard-deviation increase in Life Ladder was associated with a 0.044 decrease in log suicide ($p = .014$). That association attenuated to near zero after the social-economic covariates were added. In the full core model, unemployment ($b = 0.058$, $p = .007$) and alcohol consumption ($b = 0.100$, $p = .033$) were positively associated with log suicide. In the inequality-enriched sample, the corresponding coefficients were 0.060 ($p = .017$) and 0.176 ($p = .009$). Gini itself was not statistically distinguishable from zero ($b = -0.047$, $p = .168$).

Results for residual mismatch were similar for the two most consistent covariates. In the core specification, unemployment ($b = 0.101$, $p < .001$) and alcohol ($b = 0.133$, $p = .006$) were positive. In the Gini-enriched specification, unemployment remained positive ($b = 0.087$, $p = .001$) and alcohol increased to 0.209 ($p = .002$). Social support did not have a detectable within-country coefficient. This null does not imply that support is unimportant: it may reflect measurement error, limited within-country change, or the difference between a national average and support among high-risk individuals.

The fixed-effects findings support H1 for unemployment and alcohol, but not a simple distributive or service-resource account. Health expenditure, ageing, and Gini were not independently significant in the main within-country models. The estimates remain associational; policy changes, recession dynamics, and mortality reporting can jointly influence predictors and outcomes.

Table 3. Selected two-way fixed-effects results

Predictor	Log suicide b (SE)	p	Residual mismatch b (SE)	p
Life Ladder	-0.003 (0.013)	.824	0.057 (0.022)	.011
Social support	-0.002 (0.012)	.841	-0.004 (0.016)	.775
Unemployment	0.058 (0.022)	.007	0.101 (0.026)	< .001
Alcohol	0.100 (0.047)	.033	0.133 (0.049)	.006
Age 65+	-0.088 (0.085)	.302	-0.079 (0.099)	.421
Health expenditure	0.017 (0.022)	.441	0.035 (0.029)	.219
Gini	-0.047 (0.034)	.168	-0.033 (0.035)	.349

Predictors are standardized. Country and year fixed effects are included; standard errors are clustered by country. Gini rows come from the inequality-enriched model (N = 1,237); other rows use the core social-economic model (N = 1,859).

4.3 Predictive performance and explainable AI

XGBoost outperformed the spline GAM in the demanding country-grouped validation. Its RMSE was 0.641, MAE 0.487, and R^2 0.243; the GAM produced RMSE 0.703, MAE 0.541, and R^2 0.090. The country-held-out R^2 is modest, which is expected when predicting countries not represented during training. It nevertheless improves meaningfully on

the mean baseline and supports H2. In the 2018–2021 temporal holdout, XGBoost achieved R^2 0.603, while the GAM achieved 0.095. The higher temporal score partly reflects the availability of earlier observations from the same countries, so it should not be compared directly with the country-held-out score as if the test sets had equal difficulty.

The three largest global mean absolute SHAP values were alcohol consumption (0.221), urbanization (0.141), and log GDP (0.131). Negative affect, ageing, health expenditure, positive affect, and social support followed. The dependence plots show substantial nonlinearity rather than constant marginal effects. This helps explain why XGBoost outperformed the additive spline comparator: the fitted pattern distributes signal across thresholds and interactions that are not summarized by a single coefficient.

Korea's average SHAP profile differed from the global ranking. Alcohol made the largest positive contribution to its predicted residual mismatch (+0.280), followed by positive affect (+0.140) and ageing (+0.139). Urbanization made the largest negative contribution (−0.165), while health expenditure and social support were slightly negative. These values should not be translated into intervention effects. In particular, the positive contribution of national positive affect illustrates how an aggregate mean can coexist with a high-risk tail and how correlated features can acquire counterintuitive predictive attributions.

Table 4. Out-of-sample prediction of residual mismatch

Model	Validation	N	RMSE	MAE	R^2
XGBoost	5-fold nested country-grouped CV	2,044	0.641	0.487	0.243
Spline GAM	5-fold nested country-grouped CV	2,044	0.703	0.541	0.090
Training-mean baseline	5-fold nested country-grouped CV	2,044	0.737	0.567	-0.000
XGBoost	nested temporal holdout 2018–2021	511	0.437	0.337	0.603
Spline GAM	nested temporal holdout 2018–2021	511	0.660	0.511	0.095
Training-mean baseline	nested temporal holdout 2018–2021	511	0.695	0.522	-0.002

Countries are kept intact in grouped folds. Temporal holdout trains through 2017 and tests 2018–2021.

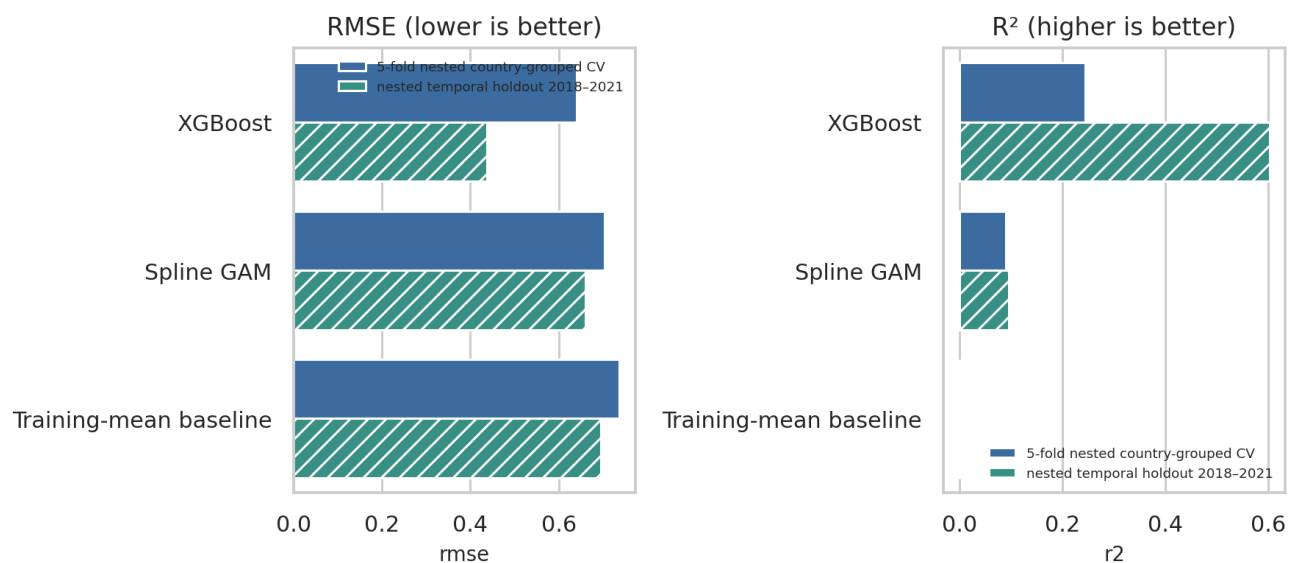


Fig. 4 Out-of-sample prediction of residual discordance

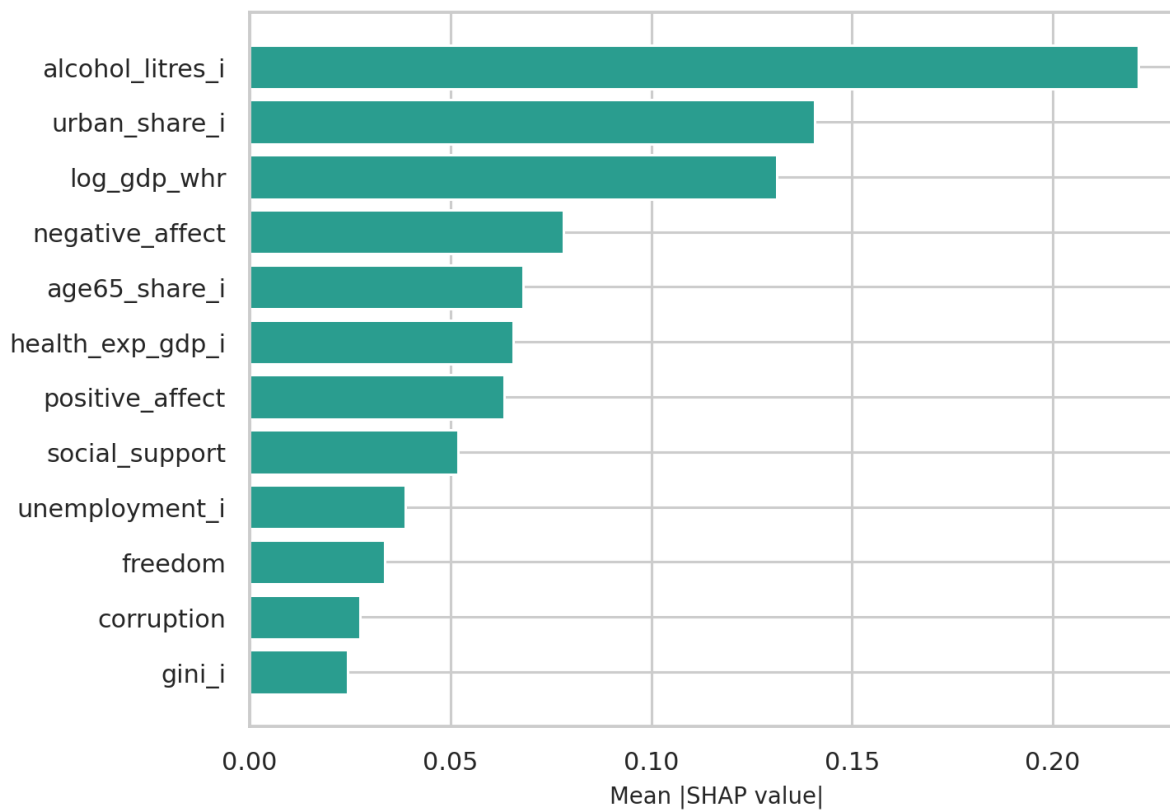


Fig. 5 Global XGBoost feature importance

Note. Mean absolute SHAP measures predictive attribution magnitude.

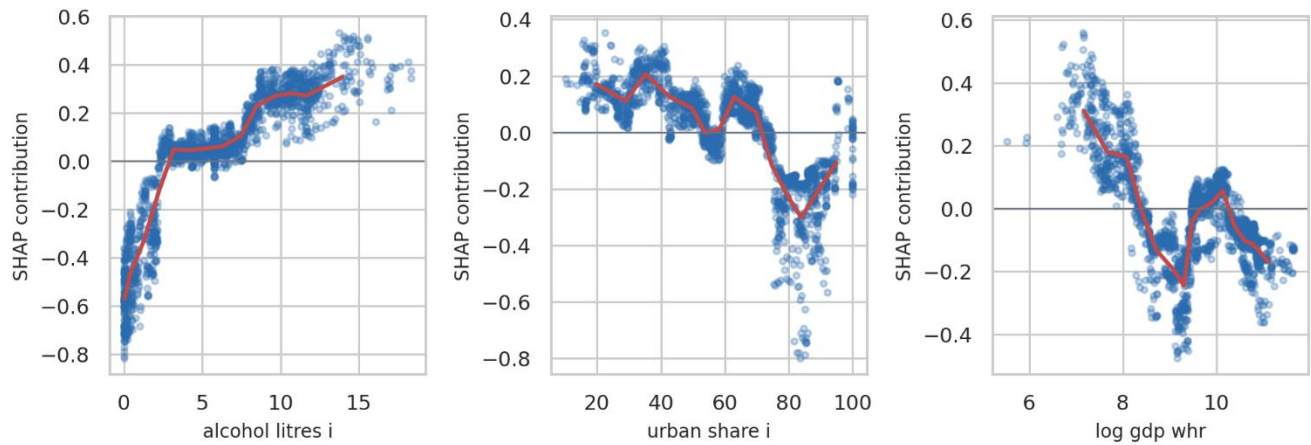


Fig. 6 Nonlinear SHAP patterns for leading predictors

Note. Solid curves are binned descriptive means of SHAP contributions, not causal response functions.

4.4 Anomalies, profiles, and Korea

The 7% Isolation-Forest specification flagged 144 country-years across 25 countries. Lithuania had the largest persistent high–high score, followed by the Russian Federation, Kazakhstan, and Korea. Korea was flagged in 12.5% of its 16 observed years and occupied the high–high quadrant in 93.8% of them. Its mean residual mismatch was 1.183 log-rate units. The anomaly result was not a single-year artifact: Korea's anomaly rate was 6.3% under 5% contamination and 18.8% under 10% contamination. Country-set overlap with the 7% specification had Jaccard similarity .880 at 5% contamination and .625 at 10%.

The three-component Gaussian mixture separated a small Southern African high-mortality cluster (Lesotho, Eswatini, and South Africa), an affluent and ageing cluster of 50 mostly European and other high-income countries, and a 90-country younger, lower-resource cluster. Korea belonged to the affluent/ageing cluster with posterior probability 1.00. Within that cluster, its 2017–2021 suicide rate was second only to Uruguay and exceeded the Russian Federation and Lithuania by the profile mean, although small differences should not be overinterpreted. Korea's mean Life Ladder was below the cluster mean of 6.376, and its social support was below the cluster mean of .894.

The profile result clarifies the policy comparison. Korea is not exceptional merely because it is old, urban, or affluent; many peers share those characteristics without comparable suicide mortality. Nor does the model support a conclusion that limited psychiatrist density alone explains the gap. In the enriched country model, psychiatrist density was not statistically significant after adjustment (core $b = 0.085$, $p = .538$; inequality-enriched $b = 0.185$, $p = .292$). Alcohol remained positive, while higher log GDP was associated with lower residual mismatch. Workforce quantity may be measured too sparsely to represent accessibility, continuity, quality, or crisis responsiveness.

Table 5. Leading persistent high-happiness/high-suicide country profiles

Country	Years	Anomaly rate	High-high share	Mean residual	Persistence score
Lithuania	16	37.5%	93.8%	1.279	0.352
Russian Federation	16	56.2%	50.0%	1.363	0.281
Kazakhstan	16	18.8%	93.8%	1.183	0.176
Republic of Korea	16	12.5%	93.8%	1.183	0.117
Belarus	14	0.0%	78.6%	1.071	0.000
Mongolia	14	0.0%	21.4%	1.012	0.000
Japan	16	0.0%	100.0%	0.844	0.000
Uruguay	16	0.0%	100.0%	0.837	0.000
Hungary	15	0.0%	40.0%	0.825	0.000
Somalia	7	0.0%	14.3%	0.808	0.000

Restricted to countries with at least five observations. Persistence score = anomaly rate × high-high share; zero scores may still have consistently positive residual mismatch.

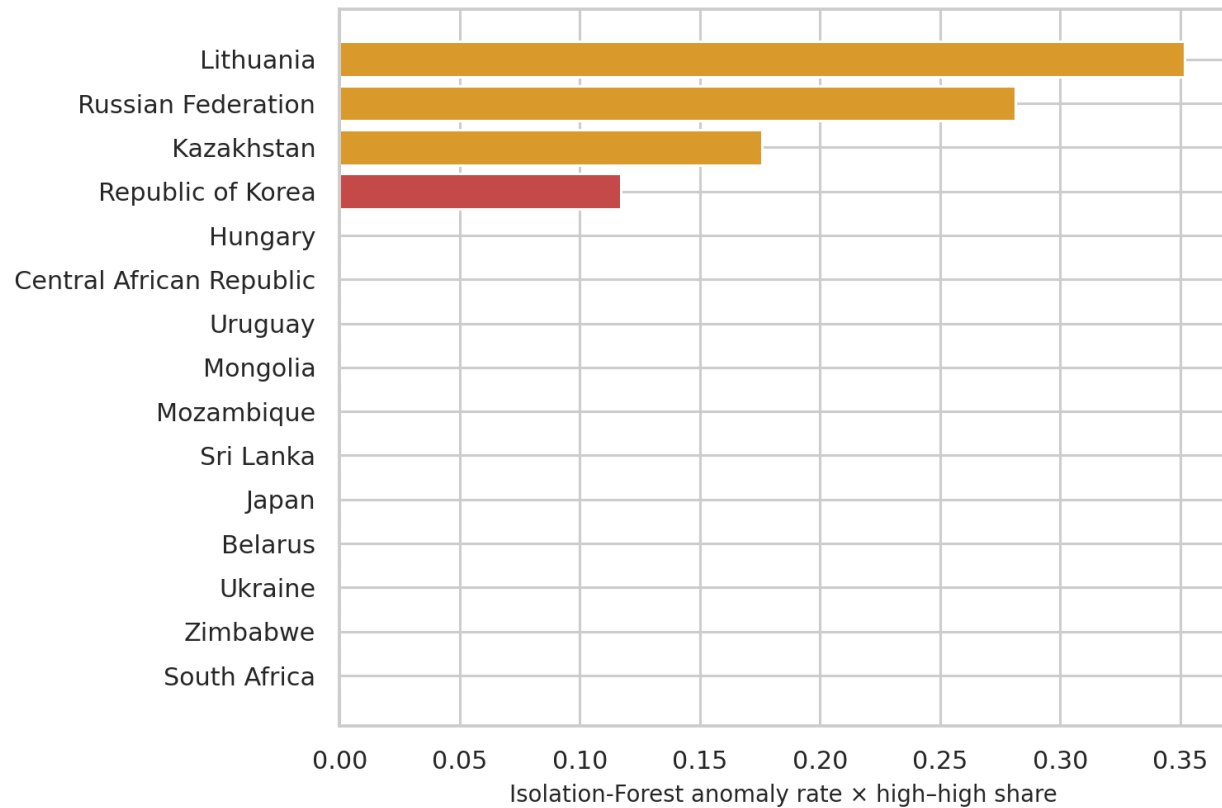


Fig. 7 Persistent high-happiness/high-suicide anomaly profiles

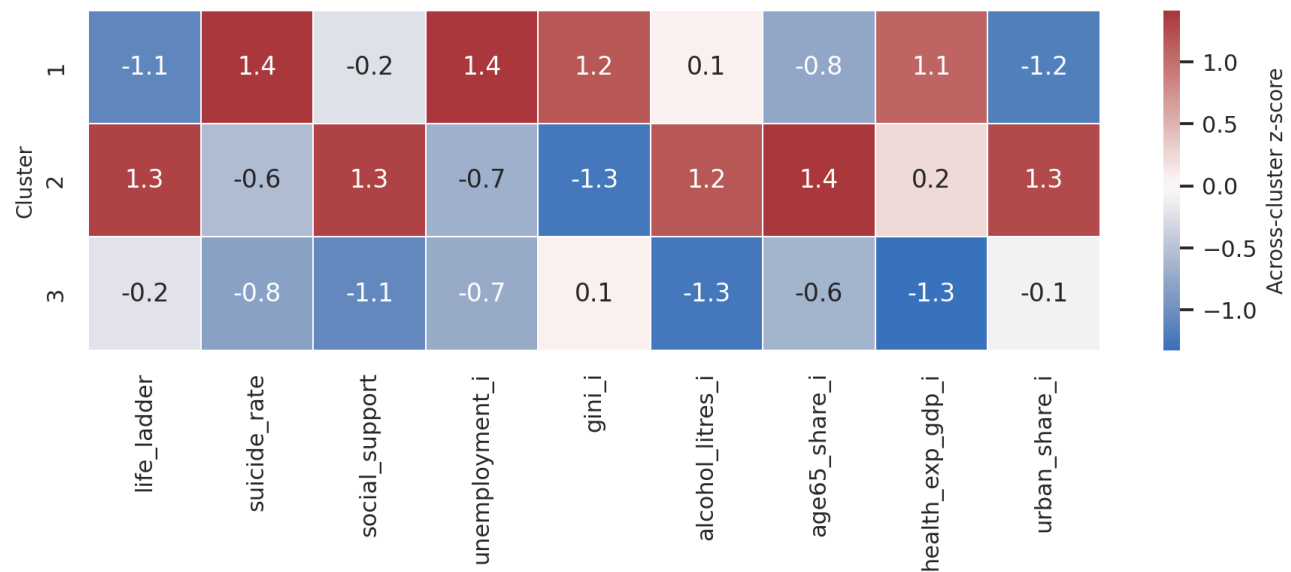


Fig. 8 Gaussian-mixture country profiles, 2017–2021

Note. Cells are standardized across the three cluster means. The three-country first cluster should be treated as a small extreme profile.

4.5 Sex, age, and post-2021 extension

In 2021, Korea's age-standardized male suicide rate was 28.156 and the female rate 13.629 per 100,000, a male-to-female ratio of 2.07. In sex-specific fixed-effects models, unemployment was positively associated with male log suicide ($b = 0.071$, $p = .001$) but not with female log suicide ($b = 0.040$, $p = .178$). The alcohol coefficients were positive but imprecise in both models. These estimates do not establish a statistically significant sex difference in coefficients, but they identify male labor-market vulnerability as a useful hypothesis for individual- or subgroup-level research.

Korea's 2021 crude suicide rate increased strongly with age: 11.27 for ages 15–19, 20.16 for ages 15–29, 27.96 for ages 30–49, 31.28 for ages 50–69, and 56.49 for ages 70 and older. Because these are crude age-specific rates and the main outcome is age-standardized, the comparison is descriptive. It nevertheless shows that a national aggregate conceals large life-course gradients.

The later-data extension adds an important qualification. Among the 55 countries with matched 2022 reported mortality and rolling happiness data, Korea's life evaluation was at the 24th percentile, while its reported suicide rate of 19.02 was at the 98th percentile. Hence it was not a high-happiness country within that selective comparison set, although the directional index remained positive because the suicide deviation was large. Twenty-nine countries have a 2023 match, but Korea is absent from WHO MDB 2023 in the downloaded release. The OECD separately reports 23.2 for Korea in 2023. The World Happiness Report 2026 places Korea at 6.040 and rank 67 using the 2023–2025 average. These updates reinforce, rather than erase, the benchmark-sensitive conclusion.

Table 6. Korea profile across analytical modules

Module	Happiness / composition	Suicide	Interpretation
2021 core, global	Life Ladder 6.113 (63rd percentile)	GHE suicide 20.575 (97th percentile)	High-high
2021 core, OECD	Life Ladder 6.113 (19th percentile)	Highest suicide rate in subset	Low-high
2017–2021 profile	Cluster 2; below cluster happiness/support means	Second-highest cluster suicide mean	Affluent/ageing outlier
2021 sex	Male 28.156; female 13.629	Male/female ratio 2.07	Sex gradient
2021 age	70+ crude rate 56.49	15–29 crude rate 20.16	Age gradient
2022 reported extension	Life evaluation 5.951 (24th percentile)	MDB suicide 19.023 (98th percentile)	Benchmark-sensitive
2023 external OECD	Not pooled	Korea 23.2; OECD mean 10.7	Current context
WHR 2026	6.040; rank 67; 2023–2025 average	No matched comparable GHE rate	Current context

Percentiles are calculated within the indicated analytical sample. OECD 2023 and WHR 2026 values are contextual and are not pooled with WHO GHE.

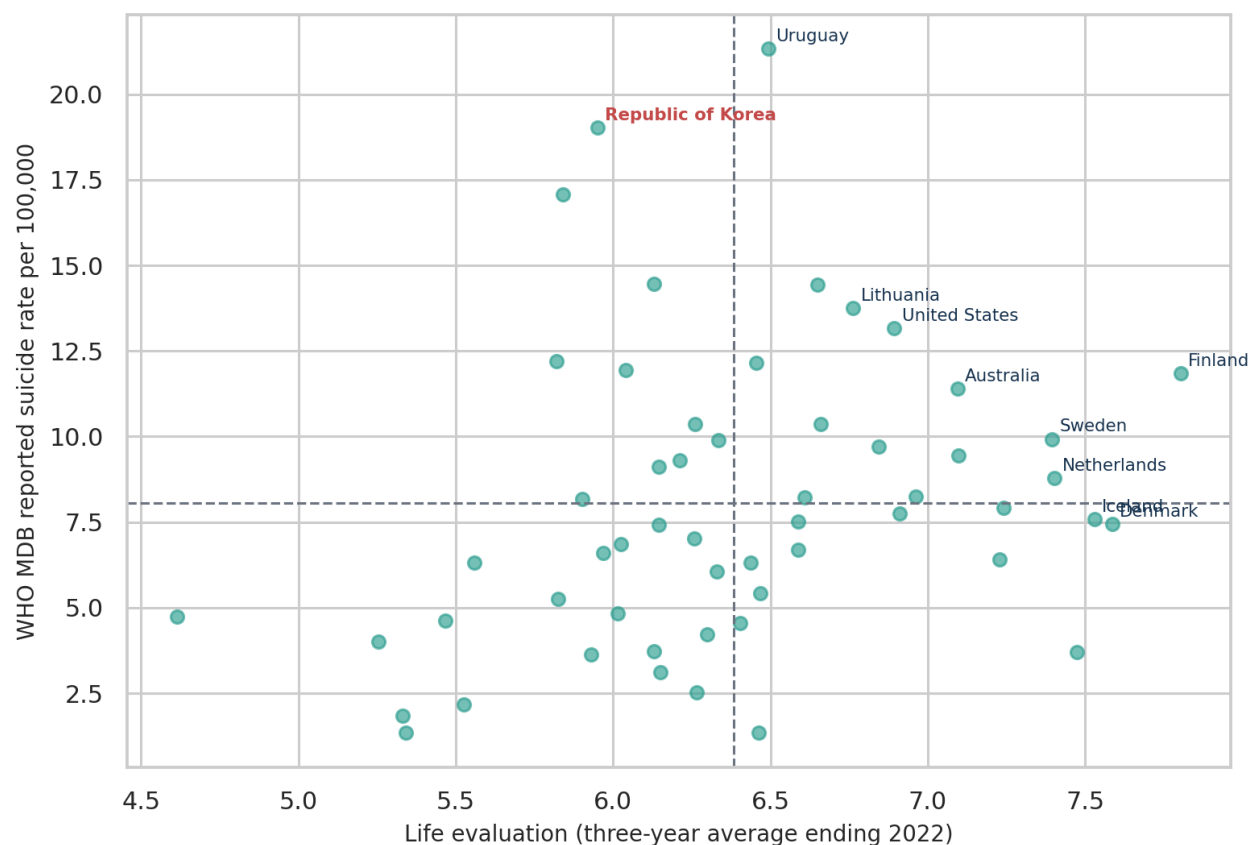


Fig. 9 Reported-data extension for the latest common year including Korea

Note. Life evaluation is a rolling three-year average; suicide is a WHO MDB reported rate (n = 55 countries in 2022). The selective sample is not directly comparable to the global GHE panel.

5. Discussion and Conclusion

5.1 Main findings

The evidence does not support the simple proposition that happier countries necessarily have lower suicide mortality. Nor does it establish a paradox at the individual level. Life evaluation and suicide mortality describe different parts of the population distribution, arise from different measurement systems, and respond to partly different social processes. A small group of high-happiness/high-suicide configurations persisted across the 161-country panel. Korea was among the clearest cases under the global benchmark, but it did not qualify as a high-happiness case within the OECD. What remains stable is Korea's exceptionally high suicide mortality; the happiness label changes with the reference group.

The methods illuminate separate parts of this result. Fixed-effects estimates connect within-country increases in unemployment and alcohol consumption with higher suicide mortality and mismatch. XGBoost shows that nonlinear and interactive patterns add predictive information out of sample. Isolation Forest locates persistent unusual configurations without reducing anomaly status to membership in one quadrant. SHAP then opens the fitted model to inspection, showing how alcohol, ageing, affect, and other structural features contributed to Korea's predicted mismatch. Finally, the Gaussian mixture places Korea among affluent, ageing countries and reveals how far it sits above most members of that profile.

Taken as a whole, the results are compatible with several pathways, although they do not prove any one of them. Labor-market disruption can bear heavily on vulnerable groups while leaving the national happiness mean almost unchanged. Alcohol may increase both underlying risk and the lethality of an acute crisis; the positive fixed-effects and SHAP results are consistent with the recent population-level meta-analysis (Guo et al., 2025). Social support may improve average well-being without reaching isolated older people, precarious workers, or individuals who hesitate to disclose distress. Similarly, spending and workforce counts say little about timely access, treatment continuity, gatekeeper coverage, or post-crisis coordination.

5.2 Implications for Korea and well-being policy

For Korea, the implication is not to abandon happiness policy, but to stop using average happiness as a sufficient proxy for safety or mental health. A useful joint dashboard would report mean life evaluation, its lower tail or dispersion, age- and sex-specific suicide mortality, and service-process measures such as timely access and continuity after self-harm.

The international sources used here cover the first element and part of the third. Identifying where low well-being and suicide risk overlap will require national microdata linked, under appropriate safeguards, to administrative records.

The unemployment finding points toward stronger income and employment protection during transitions, especially for working-age men. Korea's steep older-age gradient also directs attention to pension adequacy, chronic-disease management, bereavement support, social participation, and access outside major cities. These are policy hypotheses drawn from an aggregate pattern, not program-specific effect estimates. Alcohol offers a more directly testable population-policy lever because taxation, availability, marketing restrictions, brief intervention, and treatment can be evaluated using established designs. Even so, alcohol policy belongs within a comprehensive prevention strategy rather than a single-cause account.

The sensitivity to benchmarks also matters for public communication. Describing Korea as 'happy but suicidal' is memorable, yet it exaggerates the happiness component when the comparison is restricted to OECD peers or the later reported-data sample. A more faithful summary is that Korea's mean life evaluation is moderate globally and low among affluent peers, while suicide mortality is extremely high in both settings. This wording keeps average quality of life and concentrated risk in view without implying that cheerful individuals are more likely to die by suicide.

5.3 Implications for explainable AI in quality-of-life research

The methodological lesson extends beyond suicide research. Explainable AI can contribute to quality-of-life studies when the validation design matches the intended generalization. A random split of country-years would place the same country's institutional signature in both training and test data and would overstate performance. Nested country-grouped validation instead produced the more demanding R^2 of 0.243. The nested temporal holdout, which retained earlier observations from test countries, produced $R^2 = 0.603$. Both estimates answer legitimate but different questions and should not be treated as interchangeable measures of accuracy.

Model explanations must also be tied to a precisely defined target. In this study, SHAP explains residual mismatch generated out of country rather than raw suicide mortality alone; a positive contribution is therefore conditional on the happiness-year benchmark. The causal boundary is equally important. SHAP allocates fitted dependence among features, but it cannot resolve confounding, reverse causation, measurement error, or correlation among predictors. Pairing SHAP with fixed effects makes those differences easier to see; it does not make them disappear.

Used in this way, AI helps identify configurations that standard dashboards may miss, tests nonlinear prediction under realistic validation, and indicates where closer human investigation is warranted. It does not decide which policy follows

from the evidence. In high-stakes well-being research, transparency requires an account of both what the model detects and what the available data leave unresolved.

5.4 Limitations and future research

The findings should be read with several limitations in mind. Most importantly, the analysis is ecological. It cannot show that happier individuals face greater suicide risk, identify the people who died, or estimate the within-country overlap between low well-being and mortality risk. Those questions require individual longitudinal studies or linked administrative data. Suicide also cannot stand in for mental health as a whole; it is a rare and consequential outcome shaped by clinical, social, behavioral, cultural, and measurement factors.

International suicide statistics remain uncertain as well. GHE improves comparability through modelling, but its estimates depend on source quality and may have wide uncertainty intervals. MDB supplies newer reported values for a selective set of countries. Restricting the sensitivity sample to countries with at least ten reported years reduces the problem of absent reporting, but it is not a formal assessment of completeness. A stronger future analysis would incorporate WHO uncertainty draws and direct measures of death-registration quality.

The happiness series has its own constraints. National means contain sampling error, and many countries are not observed annually. The annual WHR analytical file ends in 2022, whereas the freely distributed 2026 file reports overlapping three-year averages. For that reason, the post-2021 extension cannot reproduce the annual structure of the main panel. New globally comparable GHE estimates and access to annual Gallup microdata would allow a more consistent update.

Several predictors are necessarily approximate. The social-support item does not measure family conflict, loneliness, or participation in community life. Psychiatrist density is observed in different years across countries and omits affordability, primary-care integration, and service quality. Gini coverage is sparse, and the underlying concept may refer to income in some countries and consumption in others. Limited interpolation and separate enrichment models narrow these weaknesses but cannot remove them.

The machine-learning results also depend on design choices. We used one pre-specified XGBoost configuration rather than an extensive nested hyperparameter search. This limits tuning variance and researcher discretion, although future comparisons could consider other algorithms and monotonic constraints. Isolation-Forest contamination is another choice rather than an empirical constant. The sensitivity tests show that Korea's flagging rate varies with the threshold, even though its elevated residual does not.

Finally, contemporaneous fixed-effects estimates do not determine temporal order. Longer official panels could support distributed lags, while policy discontinuities, synthetic controls, or other quasi-experimental designs could examine changes in alcohol regulation, employment protection, or service access. The most direct test of the proposed mechanisms would link individual life evaluations and local conditions to subsequent mortality risk in a multilevel design.

5.5 Conclusion

National happiness and suicide mortality are best interpreted together, not reduced to a single indicator. In the 2005-2021 panel, unemployment and alcohol consumption were the most consistent within-country correlates of excess suicide mortality relative to happiness. XGBoost also captured nonlinear structural patterns that the additive GAM missed. Korea repeatedly appeared as a global high-happiness/high-suicide anomaly, although its happiness was low among OECD peers and in the later reported-data comparison. The central conclusion is therefore not a paradox about happy people. It is a problem of measurement and governance: favorable averages can coexist with severe, concentrated, and preventable harm. Explainable AI is valuable when it brings that coexistence into view while remaining anchored in careful benchmarking, valid cross-validation, public-health theory, and ecological restraint.

Statements and Declarations

Funding: No external funding was received for this study.

Competing interests: The author(s) declare no competing interests.

Ethics approval: Not applicable. The study uses publicly available country-level aggregate data and contains no individual-level records.

Consent to participate: Not applicable.

Consent for publication: Not applicable.

Data availability: Online Resource 1 contains the processed analytical panel, source inventory, checksums, model outputs, SHAP values, anomaly scores, cluster assignments, and reproducibility specifications. Source data are available from the World Health Organization, World Happiness Report, and World Bank at the URLs listed in the supplement.

Code availability: The analytical workflow and complete parameter settings are documented in Online Resource 1. A clean script can be deposited in a public repository upon acceptance without changing the reported results.

Materials availability: Not applicable beyond the publicly available source data and the analytical materials supplied in Online Resource 1.

Author contributions: Author-identifying contribution statements are withheld for double-anonymous review and will be provided after acceptance.

Supplementary Information

Online Resource 1. Analytical data, model outputs, source inventory, checksums, and reproducibility specifications for the cross-national happiness–suicide analysis (Excel workbook).

References

- Bray, I., & Gunnell, D. (2006). Suicide rates, life satisfaction and happiness as markers for population mental health. *Social Psychiatry and Psychiatric Epidemiology*, 41(5), 333–337. <https://doi.org/10.1007/s00127-006-0049-z>
- Cantril, H. (1965). *The pattern of human concerns*. Rutgers University Press.
- Case, A., & Deaton, A. (2015). Suicide, age, and well-being: An empirical investigation. In D. A. Wise (Ed.), *Insights in the economics of aging* (pp. 307–334). University of Chicago Press.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Daly, M. C., Oswald, A. J., Wilson, D., & Wu, S. (2011). Dark contrasts: The paradox of high rates of suicide in happy places. *Journal of Economic Behavior & Organization*, 80(3), 435–442. <https://doi.org/10.1016/j.jebo.2011.04.007>
- Daly, M. C., Wilson, D. J., & Johnson, N. J. (2013). Relative status and well-being: Evidence from U.S. suicide deaths. *Review of Economics and Statistics*, 95(5), 1480–1500. https://doi.org/10.1162/REST_a_00355
- Deaton, A., & Stone, A. A. (2013). Two happiness puzzles. *American Economic Review*, 103(3), 591–597. <https://doi.org/10.1257/aer.103.3.591>
- Durkheim, E. (1951). *Suicide: A study in sociology* (J. A. Spaulding & G. Simpson, Trans.). Free Press. (Original work published 1897)
- Guo, K., Jiang, H., Shield, K. D., Spithoff, S., & Lange, S. (2025). Alcohol consumption per capita and suicide: A meta-analysis. *JAMA Network Open*, 8(9), e2533129. <https://doi.org/10.1001/jamanetworkopen.2025.33129>
- Hastie, T. J., & Tibshirani, R. J. (1990). *Generalized additive models*. Chapman & Hall.
- Helliwell, J. F. (2007). Well-being and social capital: Does suicide pose a puzzle? *Social Indicators Research*, 81(3), 455–496. <https://doi.org/10.1007/s11205-006-0022-y>
- Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. *Philosophical Transactions of the Royal Society B*, 359(1449), 1435–1446. <https://doi.org/10.1098/rstb.2004.1522>
- Helliwell, J. F., Huang, H., & Wang, S. (2014). Social capital and well-being in times of crisis. *Journal of Happiness Studies*, 15, 145–162. <https://doi.org/10.1007/s10902-013-9441-z>
- Helliwell, J. F., Layard, R., Sachs, J. D., Aknin, L. B., De Neve, J.-E., & Wang, S. (Eds.). (2023). *World Happiness Report 2023*. Sustainable Development Solutions Network.
- Helliwell, J. F., Layard, R., Sachs, J. D., De Neve, J.-E., Aknin, L. B., & Wang, S. (Eds.). (2026). *World Happiness Report 2026*. University of Oxford, Wellbeing Research Centre.
- Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences*, 107(38), 16489–16493. <https://doi.org/10.1073/pnas.1011492107>

- Koivumaa-Honkanen, H., Honkanen, R., Viinamäki, H., Heikkilä, K., Kaprio, J., & Koskenvuo, M. (2001). Life dissatisfaction and suicide: A 20-year follow-up study. *American Journal of Psychiatry*, 158(3), 433–439. <https://doi.org/10.1176/appi.ajp.158.3.433>
- Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422. <https://doi.org/10.1109/ICDM.2008.17>
- Lund, C., Brooke-Sumner, C., Baingana, F., Baron, E. C., Breuer, E., Chandra, P., Haushofer, J., Herrman, H., Jordans, M., Kieling, C., Medina-Mora, M. E., Morgan, E., Omigbodun, O., Tol, W., Patel, V., & Saxena, S. (2018). Social determinants of mental disorders and the Sustainable Development Goals: A systematic review of reviews. *The Lancet Psychiatry*, 5(4), 357–369. [https://doi.org/10.1016/S2215-0366\(18\)30060-9](https://doi.org/10.1016/S2215-0366(18)30060-9)
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. Wiley.
- Milner, A., Page, A., & LaMontagne, A. D. (2013). Long-term unemployment and suicide: A systematic review and meta-analysis. *PLOS ONE*, 8(1), e51333. <https://doi.org/10.1371/journal.pone.0051333>
- Morgenstern, H. (1995). Ecologic studies in epidemiology: Concepts, principles, and methods. *Annual Review of Public Health*, 16, 61–81. <https://doi.org/10.1146/annurev.pu.16.050195.000425>
- OECD. (2025). *Health at a Glance 2025: OECD indicators*. OECD Publishing. <https://doi.org/10.1787/8f9e3f98-en>
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Aroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15(3), 351–357. <https://doi.org/10.2307/2087176>
- Roche, S. P., Rogers, M. L., & Pridemore, W. A. (2018). A cross-national study of the population-level association between alcohol consumption and suicide rates. *Drug and Alcohol Dependence*, 188, 16–23. <https://doi.org/10.1016/j.drugalcdep.2018.02.036>
- Shmueli, G. (2010). To explain or to predict? *Statistical Science*, 25(3), 289–310. <https://doi.org/10.1214/10-STS330>
- Turecki, G., Brent, D. A., Gunnell, D., O'Connor, R. C., Oquendo, M. A., Pirkis, J., & Stanley, B. H. (2019). Suicide and suicide risk. *Nature Reviews Disease Primers*, 5, Article 74. <https://doi.org/10.1038/s41572-019-0121-0>
- World Bank. (2026). *World Development Indicators*. <https://databank.worldbank.org/source/world-development-indicators>
- World Health Organization. (2024). *Global Health Estimates 2021: Deaths by cause, age, sex, by country and by region, 2000–2021*. World Health Organization.
- World Health Organization. (2025a). *Mental Health Atlas 2024*. World Health Organization.
- World Health Organization. (2025b). *Suicide worldwide in 2021: Global health estimates*. World Health Organization.

Appendix A. Additional design and robustness information

A1. Benchmark and series rules

The main conclusions use the 2005–2021 GHE panel. The 2022–2023 extension is reported because users and policymakers require current context, but its reported WHO MDB rates are not substitutions for missing GHE years. The extension also uses rolling three-year happiness averages. Any manuscript revision should preserve this separation unless WHO releases a new globally comparable GHE series covering the later years.

Table A1. Anomaly-threshold sensitivity

Contamination	Flagged country-years	Countries ever flagged	Korea anomaly rate	Jaccard vs. 7%
5%	103	22	6.2%	0.880
7%	144	25	12.5%	1.000
10%	205	40	18.8%	0.625

Country-set Jaccard similarity compares countries with at least one flagged year.

Table A2. Fixed-effects sensitivity for unemployment and alcohol

Sample	Predictor	b (SE)	p	Countries	N
All countries	Unemployment	0.107 (0.026)	< .001	158	1,861
All countries	Alcohol	0.124 (0.048)	.010	158	1,861
OECD members	Unemployment	0.043 (0.018)	.021	38	523
OECD members	Alcohol	0.065 (0.046)	.163	38	523
World Bank high income	Unemployment	0.071 (0.020)	< .001	51	655
World Bank high income	Alcohol	0.078 (0.072)	.277	51	655
WHO MDB ≥ 10 annual observations	Unemployment	0.088 (0.026)	< .001	83	1,041
WHO MDB ≥ 10 annual observations	Alcohol	0.149 (0.069)	.030	83	1,041

Outcome is residual mismatch. Predictors are standardized within each sample.

A2. Reproducibility checklist

- Annual happiness file and all official-source downloads are identified by URL, retrieval date, and SHA-256 checksum in the workbook.
- Country names are converted to ISO3 before merging; only country-years appearing in both WHR and WHO GHE enter the core panel.
- WDI interpolation is limited to two consecutive internal years and is never used to extend a series beyond observed endpoints.
- Residual mismatch is generated out of country; XGBoost and GAM assessment use nested country-grouped target construction and validation.
- All stochastic algorithms use seed 20260818.
- SHAP values are stored at row level, and Korea-specific values can be audited against the core panel.
- The published manuscript should archive the analysis script together with the workbook before final acceptance.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [ESM1HappinessSuicideAnalysisData.xlsx](#)