Economic Modelling 164 (2026) 107763

Contents lists available at ScienceDirect

Economic Modelling

journal homepage: www.journals.elsevier.com/economic-modelling

Worldwide well-being efficiency and its determinants: New insights from the ordered probit stochastic Frontier model[☆]

Oleg Badunenko[a,*], José M. Cordero[b], Subal C. Kumbhakar[c]

[a] Brunel Business School, Brunel University of London, UK
[b] Universidad de Extremadura, Badajoz, Spain
[c] Department of Economics, State University of New York, Binghamton, NY 13902-6000, USA

ARTICLE INFO

Dataset link: Worldwide well-being efficiency and its determinants: New insights from the Ordered Probit Stochastic Frontier Model (Original data)

JEL classification:
C13
I31
H53

Keywords:
Well-being efficiency
Stochastic Frontier analysis
Cross-country analysis
Subjective life satisfaction
Ordered probit

ABSTRACT

This paper examines how efficiently individuals and societies transform economic and social resources into subjective well-being. We propose an ordered probit stochastic frontier model that explicitly accounts for the ordinal nature of life satisfaction data while jointly estimating well-being shortfalls and their determinants within a single-stage framework. We also show why treating ordinal well-being indicators as continuous variables in conventional stochastic frontier models may lead to misleading inference and interpretation. Using World Values Survey data for more than 214,000 individuals across 74 countries, combined with macroeconomic and institutional indicators, we estimate attainable well-being levels conditional on individual and contextual characteristics. The results reveal substantial heterogeneity in well-being efficiency across individuals and countries. Institutional quality, social trust, and labor market conditions are strongly associated with well-being efficiency, while unemployment generates the largest shortfalls from attainable well-being. These findings highlight the importance of accounting for the ordinal nature of subjective well-being measures when analyzing well-being efficiency and its determinants.

1. Introduction

Traditional economic indicators like GDP and income provide only a partial view of a nation's or an individual's well-being (Fleurbaey and Blanchet, 2013; Agrawal et al., 2025). In recent decades, economists have increasingly embraced the concept of subjective well-being (SWB), measured through self-reported indicators such as happiness or life satisfaction to complement traditional objective indicators (Nikolova and Graham, 2021).[1] The growing availability of internationally comparable data on subjective well-being (e.g., Gallup World Poll, World Values Survey, or European Social Survey) has greatly contributed to the expansion of a literature aimed at understanding how individuals evaluate their lives and what factors shape these evaluations mainly in the field of economics (Dominko and Verbič, 2019; Frijters and Krekel, 2021; Helliwell et al., 2025).

A large body of empirical evidence has examined the factors associated with higher levels of subjective well-being using mainly regression analysis (Frey and Stutzer, 2013). These models typically assume that well-being measures are absolute and that individuals can accurately evaluate and report their internal states. However, some researchers suggest that this assumption may be too strong (Schimmack and Oishi, 2005; Steffel and Oppenheimer, 2009). Rather than reflecting a direct and objective readout of internal emotional states, life satisfaction measures tend to be

☆ We are grateful to the Editor, the Associate Editor, and two anonymous referees for constructive comments that substantially improved this paper.
* Corresponding author.
E-mail addresses: oleg.badunenko@brunel.ac.uk (O. Badunenko), jmcordero@unex.es (J.M. Cordero), kkar@binghamton.edu (S.C. Kumbhakar).

[1] Although these terms are closely related, they are not fully equivalent. Subjective well-being is usually understood as a broad umbrella concept that includes both evaluative and affective components. Life satisfaction refers mainly to the evaluative assessment individuals make of their lives as a whole, whereas happiness is often more closely associated with affective states and may therefore be more sensitive to short-term circumstances or emotions (see, for example, Veenhoven, 1991). In this paper, we use a measure of subjective well-being based on life satisfaction and therefore capture its evaluative dimension. For readability, we occasionally use related expressions such as well-being or happiness when referring to the broader literature, but the empirical analysis should be interpreted as referring to subjective life satisfaction.

https://doi.org/10.1016/j.econmod.2026.107763
Received 19 November 2025; Received in revised form 15 July 2026; Accepted 15 July 2026
Available online 31 July 2026

shaped by comparative evaluations, in which individuals assess their well-being relative to others or to their own expectations, and in response to specific contextual conditions or reference standards. This relativistic nature of these assessments highlights an important limitation of traditional approaches and suggests the need for alternative methods that account for individual heterogeneity in the translation of resources into well-being. In this respect, it becomes crucial to assess not only which factors are associated with the level of well-being, but also the extent to which individuals attain life satisfaction levels relative to their observed resources and circumstances. Some individuals or groups may systematically attain higher levels of life satisfaction than others, given similar observable resources and contextual conditions. This highlights the need to examine not only the determinants of well-being, but also differences in attainable well-being levels conditional on individual and contextual characteristics.

To address this question, a growing body of research has begun to apply frontier-based techniques originally developed in the fields of operations research and productivity analysis to evaluate the extent to which individuals attain their potential well-being levels given their available resources and contextual conditions as suggested by Binder and Broekel (2011, 2012). Within this framework, the concept of a well-being frontier serves as a benchmark for the maximum attainable level of life satisfaction given a set of personal and contextual characteristics. Observed shortfalls from this frontier are interpreted as inefficiencies in the transformation process.[2] Importantly, in the context of subjective well-being, these estimated shortfalls should not be interpreted exclusively as evidence of behavioral or institutional inefficiency in a strict economic sense. Rather, they represent conditional deviations from an estimated statistical benchmark, given observed individual and contextual characteristics. Such deviations may reflect multiple underlying mechanisms, including heterogeneous preferences, adaptation processes, psychological traits, reporting differences, or other unobserved factors affecting the translation of resources into subjective life satisfaction. Within this framework, the frontier approach allows researchers to estimate individual-specific well-being shortfalls relative to attainable latent life satisfaction levels and to explore why some individuals appear systematically closer to this conditional benchmark than others, as well as which individual or contextual factors are associated with these differences.

The review in the following section shows that while research has traditionally favored non-parametric approaches such as Data Envelopment Analysis (DEA), there is a growing shift toward SF models. This transition is motivated by the SF framework's superior ability to disentangle unobserved heterogeneity and random noise from structural inefficiency, while providing a robust statistical foundation for identifying the determinants of performance in the presence of random noise (Kumbhakar and Lovell, 2003).

Despite its advantages, the application of SF models to the study of subjective well-being presents a key conceptual challenge that the previous studies have overlooked. Life satisfaction is measured on an ordinal scale with a limited number of discrete categories, whereas standard SF models generally assume a continuous output variable. To overcome this shortcoming, we propose a novel methodological approach by extending the ordered probit model allowing inefficiency. Our ordered probit stochastic frontier model (OPSF) respects the ordinal nature of the well-being indicator while estimating individual inefficiency scores and their determinants. The suggested framework will be particularly useful for researchers working within the happiness economics literature, as well as in other areas involving ordinal outcomes, such as product satisfaction surveys, credit ratings, or evaluations of service quality.

We also demonstrate consequences of applying the standard SF model treating the well-being indicator as a continuous variable. Using the standard SF model with non-continuous dependent variable presents a number of problems. For example, the SF likelihood assumes a continuous density, while discrete outcomes assign probability mass, not density, to integer values. Further, identification of inefficiency in the SF model depends on the third moment (skewness) of the composite error. Discreteness introduces artificial bunching in the residual distribution that can mimic or mask genuine inefficiency-driven skewness. With a misspecified continuous likelihood on discrete data, the Hessian-based standard errors are invalid, making hypotheses testing unreliable. Finally, efficiency ratio compares actual discrete output to a non-integer benchmark, which lacks a clear economic interpretation.

Our empirical application uses data from the World Values Survey (WVS) longitudinal database (Inglehart et al., 2022), which provides information on approximately 215,000 individuals across 74 countries. We combine this micro-level data with macro-level indicators that capture national economic, social, and institutional characteristics, enabling us to assess the extent to which individuals efficiently transform available resources into well-being in very heterogeneous contexts. The inclusion of country-level variables is particularly important, as a growing body of research has shown that a substantial share of the variation in subjective well-being cannot be explained solely by individual characteristics. Instead, contextual factors related to the country of residence, such as institutional quality, cultural norms or economic conditions, also exert a significant influence. For instance, Exton et al. (2015) and Ye et al. (2015) show that cultural factors explain a significant proportion of the variance in subjective life satisfaction after controlling for individual-level variables. Similarly, Emmerling et al. (2021) highlight the key role that institutional frameworks play in shaping cross-national differences in subjective well-being across countries, particularly through mechanisms related to social protection, governance quality, and economic stability. Therefore, an additional contribution of the present study is to explore the role that these country-level factors play in explaining the variability observed in well-being efficiency in terms of marginal effects.

The remainder of this paper is structured as follows. Section 2 reviews previous studies that have applied frontier techniques to estimate well-being efficiency measures and explore the main determinants of inefficiencies. Section 3 presents the methodological framework and estimation strategy. Section 4 describes the dataset and variables used in empirical analysis. Section 5 reports and discusses the main results. Finally, the paper ends with some concluding remarks in Section 6.

2. Previous literature

Research on the determinants of subjective well-being has grown substantially in recent decades, establishing a solid field commonly referred to as the economics of happiness.[3] This literature seeks to understand how individuals evaluate their lives and what factors shape their reported levels of happiness and life satisfaction. These include a broad set of sociodemographic and social factors such as age, gender, employment status, health, education, income, social connections, institutional quality or environmental conditions (Clark, 2018; Agrawal et al., 2023; Briguglio et al., 2025). Recent contributions have also emphasized the multidimensional nature of well-being, integrating insights from psychology and behavioral economics to provide a richer understanding of how subjective perceptions and objective living conditions interact (Kimball and Willis, 2023).

[2] We borrow the terms frontier and inefficiency from stochastic frontier production models and refer to them here as well-being frontier and well-being shortfall.

[3] Extensive reviews on the 'happiness economics' literature can be found in Frey (2008), Dolan et al. (2008) and Diener et al. (2018).

Overall, literature portrays subjective well-being as a multifaceted phenomenon shaped by individual circumstances, social environments, and institutional settings.

Building on this foundation, a smaller but growing body of research has applied frontier techniques to evaluate the extent to which individuals, regions, or countries attain subjective well-being given their available resources. The central idea, rooted in the seminal contribution of Rayo and Becker (2007), is that individuals seek to maximize their levels of life satisfaction given their characteristics and resource endowments. However, observed well-being often falls below the potential maximum, suggesting that some individuals are more efficient than others in transforming resources into life satisfaction. To capture this phenomenon, studies have employed frontier-based approaches, which allow researchers to measure the shortfall between achieved and attainable well-being. These techniques provide a useful framework not only for quantifying efficiency levels but also for exploring the factors that may explain the observed differences in efficiency across individuals, regions, or countries.

A substantial body of empirical research has applied nonparametric approaches to estimate well-being efficiency at the individual level. Binder and Broekel (2012) pioneered this approach by applying the robust order-m method Cazals et al. (2002) to British panel data, uncovering notable heterogeneity in individual efficiency scores, i.e. many individuals could potentially achieve higher levels of well-being given their personal characteristics and available resources. Their second-stage regression analysis reveals that education, employment status, and social interactions significantly influence well-being efficiency. Cordero et al. (2017) extend this line of research by using a conditional order-m approach Daraio and Simar (2005) to incorporate individual and country-level contextual factors directly into the estimation for a sample of individuals from 26 OECD countries, finding that factors such as health, education, and governance quality are positively associated with well-being efficiency scores. Mamatzakis and Tsionas (2021) employed a Markov Chain Monte Carlo (MCMC) procedure to estimate efficiency measures for a longitudinal sample of British individuals, showing that personality traits explain a large share of the variation in individual well-being efficiency.

Several studies have extended the analysis of well-being efficiency using aggregated data at country or regional level, mostly using nonparametric methods. Debnath and Shankar (2014) estimate DEA-based well-being efficiency scores for 130 countries, highlighting that similar policies might affect attainment across countries differently depending on country-specific contexts. Carboni and Russu (2015) applied DEA and Malmquist productivity indices to assess well-being efficiency across 20 Italian regions from 2005 to 2011, finding that northern regions consistently outperform southern ones in terms of well-being attainment, but overall, regions did not experience significant improvements during the period analyzed. Mizobuchi (2017) also employed DEA to estimate a multidimensional well-being production function for 36 countries and shows that health-related factors explain a significant share of cross-country differences in subjective well-being. Nikolova and Popova (2021) used a robust nonparametric order-α approach (Aragon et al., 2005) combined with second-stage fixed-effects panel regressions for 91 countries and find that poor labor market conditions are associated with lower well-being efficiency, while social support, freedom, and the rule of law are positively related to efficiency scores. Sarracino and O'Connor (2022) also employed a two-stage DEA with data from 39 European countries, showing that greater productive capacity, better health outcomes, optimism, and higher employment levels are strongly associated with improved efficiency. Castellano et al. (2023) analyze OECD countries using DEA and evaluate two alternative definitions of well-being efficiency: technical efficiency, which focuses on the effective transformation of well-being indicators into outcomes, and social efficiency, which incorporates the social and environmental costs of achieving certain well-being levels by treating them as undesirable inputs. Their results show that countries with high levels of well-being are not always technically efficient in producing them, while lower-income countries tend to display the lowest levels of social efficiency. Finally, Dziechciarz (2025) applies a nonparametric DEA model to assess the efficiency with which selected EU countries transform various inputs into well-being outcomes, using data from the Eurostat database for the period 2013–2022. The results reveal a high degree of variability in efficiency across countries and over time, with a general tendency towards improving well-being efficiency.

In contrast to the predominance of studies applying nonparametric techniques, a smaller but growing strand of literature has adopted parametric or semiparametric frontier methods to estimate well-being efficiency and examine its determinants. At the individual level, Pena-López et al. (2020) use SFA to assess well-being efficiency for individuals in Colombia, Costa Rica, Mexico, and the United States. Their findings suggest that cultural referents play a key role in shaping how individuals translate available resources and circumstances into subjective well-being, with significant cross-country differences linked to the prevalence of specific conceptualizations of well-being. Similarly, Thai (2025) apply SFA to data from the World Values Survey, covering over 90,000 individuals from 64 countries. They find considerable cross-country variation in well-being efficiency and show that individuals living in countries with stronger institutions, measured in terms of government effectiveness, political stability, and voice and accountability, tend to be more efficient. At the aggregate level, Bonanno et al. (2020) use SFA to analyze the performance of Italian regions and find that northern regions tend to be more efficient than southern ones. Moreover, dependency on external funding and limited autonomy in resource allocation are associated with lower efficiency. Finally, Cordero et al. (2021) apply the StoNED method, which combines stochastic and nonparametric features and allows for the direct integration of contextual factors into the estimation process (Kuosmanen and Kortelainen, 2012), to estimate well-being efficiency across 82 countries. Their results show that social indicators, particularly the quality of governance and inequality levels, play a central role in explaining cross-country differences in efficiency estimates.

In all those studies using parametric techniques, the output variable used to represent well-being or life satisfaction is treated as continuous, despite it is commonly measured on an ordered categorical scale. This inconsistency can lead to biased efficiency estimates and reduce the reliability of the results. In this paper, we also adopt a parametric approach but explicitly account for the ordinal nature of the well-being variable to address this issue. To that end, the following section introduces a novel stochastic frontier model based on an ordered probit specification, which provides a more appropriate framework for analyzing ordinal outcomes within efficiency analysis.

3. Methodology

We noted that there are several fundamental problems with treating a discrete dependent variable as continuous in SF models. First, the likelihood function is inherently misspecified. Standard SF models are built on a continuous density for the dependent variable, whereas discrete outcomes allocate probability mass to integer realizations. Imposing a continuous likelihood on discrete data therefore yields a pseudo-likelihood, for which consistency of the estimators is not guaranteed.

Second, the distributional assumptions underlying the composite error term are violated. In particular, residuals can no longer be interpreted as draws from the assumed normal–half-normal (or similar) distribution. Identification of inefficiency in SF models typically relies on the skewness of residuals; however, with discrete outcomes, this skewness is confounded by the intrinsic discreteness of the data. The resulting "lumpiness" in residuals contaminates the separation between noise and inefficiency, leading to biased parameter estimates.

Third, treating a discrete variable as continuous induces implicit heteroskedasticity. Because the support of the dependent variable is discrete and bounded in practice, the variance of the residuals becomes mechanically dependent on the conditional mean further undermining the reliable disentangling of inefficiency from random noise.

Finally, the Jondrow–Lovell–Materov–Schmidt (JLMS) estimator breaks down in this context. With a discrete dependent variable, residuals take on only a limited set of values, which in turn restricts the support of the conditional inefficiency estimates. As a result, firm-level efficiency scores become highly sensitive to small changes in observed outcomes, rendering efficiency rankings unstable and diminishing their economic interpretability. We suggest an approach to address these issues.

Our model builds on a well-established strand of literature in happiness economics that examines how various aspects of individuals' living and working environments affect their subjective well-being. These analyses commonly assume the existence of an unobserved latent variable y_i^*, which represents the individual's underlying level of well-being. This latent construct is typically modeled as a function of observable individual or contextual characteristics x_i, including socioeconomic status, employment conditions, health status, social relationships, and other relevant factors. Formally, the latent well-being level is specified as:

$$y_i^* = x_i\beta + \varepsilon_i \tag{1}$$

where β is a vector of coefficients capturing the marginal contribution of each factor, and ε_i is a random disturbance term accounting for unobserved influences and measurement error. While y_i^* is not directly observed, individuals report their life satisfaction on a discrete, ordinal scale, which is assumed to reflect different intervals of the underlying a continuous latent variable.

While traditional regression models focus on estimating the relationship between x_i and y_i^*, they do not account for the possibility that individuals may differ in their ability to translate individual and contextual resources into well-being. Thus, we propose an alternative framework based on stochastic frontier analysis, which allows the estimation of attainable latent life satisfaction levels conditional on observed characteristics. This is achieved by decomposing the error term ε_i into two components: a symmetric random noise term v_i, and a one-sided inefficiency term u_i that captures the conditional deviation from the maximum attainable well-being frontier. In the context of subjective well-being, this shortfall should not be interpreted exclusively as a direct measure of behavioral or institutional inefficiency, since it may also reflect heterogeneous preferences, adaptation processes, psychological traits, reporting differences, or other unobserved factors affecting how individuals translate available resources and circumstances into reported life satisfaction. Accordingly, the estimated frontier should be understood as a conditional statistical benchmark rather than as a universally attainable optimum.

The challenge is to accommodate the ordinal nature of the well-being variable. When the outcome is continuous, estimation proceeds via a standard SF model. The discrete, ordered nature of the observed outcome, however, violates the continuity assumption underlying the standard SF specification. To address this, we embed the frontier model within an ordered probit framework. The model is expressed as:

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + v_i - pu_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, N, \tag{2}$$

where y^* is a latent continuous variable. What we observe is

$$y_i = m \quad \text{if} \quad \mu_{m-1} < y_i^* \le \mu_m,$$

where $m = 1, \dots, M$, $\mu_0 = -\infty$ and $\mu_M = \infty$. The model in (2) is the proposed ordered probit stochastic frontier model (OPSF).

Set-up We assume $v_i \sim N(0, \sigma_v^2)$ with $\sigma_v^2 = 1$ for identification. The probability of observing $y_i = m$ for the OPSF model, conditional on u_i, is given by

$$\begin{aligned}
p_{im} &= \mathtt{Prob}(y_i = m) \\
&= \mathtt{Prob}(\mu_{m-1} < \boldsymbol{x}_i'\boldsymbol{\beta} + v_i - pu_i \le \mu_m) \\
&= \mathtt{Prob}(v_i \le \mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) + \mathtt{Prob}(v_i > \mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) \\
&= \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i).
\end{aligned} \tag{3}$$

where $\Phi(.)$ is the normal cumulative distribution function (cdf). The conditional distribution of y_i is given by

$$f(y_i, \boldsymbol{\mu}, \boldsymbol{x}_i'\boldsymbol{\beta} - pu_i) = \sum_{m=1}^{M} I_m(y_i) p_{im},$$

where $\boldsymbol{\mu} = \mu_1, \dots, \mu_{M-1}$ and

$$I_m(y_i) = \begin{cases} 1 & \text{if} \quad y_i = m \\ 0 & \text{otherwise.} \end{cases}$$

Given that $u \ge 0$, we assume it to follow a half-normal distribution, i.e., $u \sim N^+(0, \sigma_{u_i}^2)$, $\sigma_{u_i}^2 = \exp(\boldsymbol{z}_i\boldsymbol{\gamma})$, where $N^+(.)$ denotes positive half of a $N(0, \sigma_{u_i}^2)$ distribution and $\boldsymbol{z}_i$ is the vector of determinants of variance of u. In our specification, the vector $\boldsymbol{z}_i$ includes covariates defined both at the individual and at the country level. This distinction is crucial because some sources of inefficiency are inherently tied to personal attributes (such as age, gender, or employment status), while others reflect broader contextual conditions that shape individual opportunities and constraints (such as institutional quality, income inequality, or GDP per capita). Ignoring this hierarchical structure could obscure the relative contribution of each level, potentially leading to biased or misleading conclusions. By explicitly incorporating both levels of information, we aim to provide a more accurate and comprehensive explanation of the observed differences in well-being efficiency.

The determinants of the shortfall are modeled by making the variance of u heteroskedastic. The half-normal assumption implies that the probability density function of u_i is

$$f(u_i) = \frac{2}{\sqrt{2\pi}\sigma_{u_i}} \exp\left[-\frac{u_i^2}{2\sigma_{u_i}^2}\right],$$

and $E(u_i) = \sqrt{(2/\pi)}\,\sigma_{u_i} = \sqrt{(2/\pi)}\,\exp\{\frac{1}{2}(\boldsymbol{z}_i\boldsymbol{\gamma})\}$, which shows that the $\boldsymbol{z}$ variables affect the mean shortfall and therefore they can be viewed as determinants of well-being shortfall.

Log-likelihood function To derive the log-likelihood in a closed form, we expand (3) for the ordered *probit* case

$$p_{im} = \mathtt{Prob}(y_i = m) = \begin{cases} \mathtt{Prob}(v_i \le \mu_1 - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = 1 \\ \begin{bmatrix} \mathtt{Prob}(v_i \le \mu_m - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) \\ -\mathtt{Prob}(v_i \le \mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ 1 - \mathtt{Prob}(v_i \le \mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = M \end{cases}$$

$$= \begin{cases} \Phi(\mu_1 - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = 1 \\ \begin{bmatrix} \Phi(\mu_m - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) \\ -\Phi(\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ 1 - \Phi(\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = M. \end{cases} \tag{4}$$

Proposition 1. *The log-likelihood of the model in (4) is given by*

$$ll_i(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\gamma}) = \ln\left\{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m\right\}, \tag{5}$$

where

$$\mathcal{L}_m = \begin{cases} \Phi\left(h_1\right) + 2T\left(h_1, b\right) & \text{if } y_i = 1 \\ \Phi\left(h_m\right) + 2T\left(h_m, b\right) - \Phi\left(h_{m-1}\right) - 2T\left(h_{m-1}, b\right) & \text{if } 1 < y_i = m < M \\ \Phi\left(-h_{M-1}\right) + 2T\left(-h_{M-1}, -b\right) & \text{if } y_i = M, \end{cases}$$

T is the Owen's T function and $\zeta = 1/\sqrt{1+\sigma_u^2}$, $a_j = \mu_j - \mathbf{x}_i'\boldsymbol{\beta}$, $h_j = a_j\zeta$, and $b = p\sigma_u$.

The gradients of the log-likelihood function in (5) in closed-form are derived in Appendix B.

Inefficiency The task after estimating the parameters of the model and their marginal effects on the probability of belonging to a category is to predict well-being shortfall, i.e., we need to predict u_i. This is done from the conditional mean of u_i which requires derivation of the conditional pdf of u_i, i.e., $f(u_i|\epsilon_i)$

$$f(u_i|\epsilon_i) = \frac{f(u, \epsilon_i)}{f(\epsilon_i)} = \frac{\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}. \tag{6}$$

Note that after estimating the parameters of the model, we can compute ϵ_i from the residuals. Since the denominator in (6) does not depend on u_i, the expected value of $u_i|\epsilon_i$ is

$$E(u_i|\epsilon_i) = \frac{1}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}\int_0^\infty u_i\frac{2}{\sigma_u}\phi\left(\frac{u_i}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}du_i. \tag{7}$$

To evaluate this integral, we need to integrate a function of the following type

$$\int_0^\infty u\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \mathbf{x}_i'\boldsymbol{\beta} - pu_i)du_i. \tag{8}$$

Proposition 2. *The prediction of u_i is*

$$E(u_i|\epsilon_i) = 2\sigma_u\frac{\sum_{m=1}^{M} I_m(y_i)\mathcal{B}_m}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}, \tag{9}$$

where

$$\mathcal{B}_m = \begin{cases} b\zeta f\left(h_1\right) + \dfrac{\Phi\left(a_1\right)}{\sqrt{2\pi}} & \text{if } y_i = 1 \\ b\zeta f\left(h_m\right) + \dfrac{\Phi\left(a_m\right)}{\sqrt{2\pi}} - b\zeta f\left(h_{m-1}\right) - \dfrac{\Phi\left(a_{m-1}\right)}{\sqrt{2\pi}} & \text{if } 1 < y_i = m < M \\ b\zeta f\left(h_M\right) + \dfrac{\Phi\left(-a_M\right)}{\sqrt{2\pi}} & \text{if } y_i = M, \end{cases}$$

and

$$f(h_m) = \begin{cases} \phi\left(h_1\right)\Phi\left(-bh_1\right) & \text{if } y_i = 1 \\ \phi\left(h_m\right)\Phi\left(-bh_m\right) & \text{if } 1 < y_i = m < M \\ \phi\left(-h_{M-1}\right)\Phi\left(-bh_{M-1}\right) & \text{if } y_i = M. \end{cases} \tag{10}$$

The details of the derivation of inefficiency term are shown in Appendix B.

Marginal effects

Calculation of marginal effects is nontrivial due nonlinear nature of the model. They will depend on whether the variable is categorical or continuous. The marginal effects of environmental variables on inefficiency, efficiency and outcome variable can be calculated. Appendix B gives details of calculating marginal effects of (i) x_k on $\mathtt{Prob}(y_i = m)$, (ii) z_k on $E(u)$, (iii) z_k on efficiency, and (iv) z_k on $\mathtt{Prob}(y_i = m)$.

Table 1
Descriptive statistic of the variables included in the empirical analysis.

Variable	Type	Mean	SD	Min	Max
Life satisfaction/well–being	Output	6.57	2.44	1	10
Education level	Input	4.73	2.22	1	8
Incomes	Input	4.69	2.30	1	10
Health status	Input	3.85	0.88	1	5
Individual characteristics					
Age	Covariate	40.56	15.89	16	99
Female	Covariate	0.51	0.50	0	1
Unemployed	Covariate	0.09	0.29	0	1
Married	Covariate	0.64	0.48	0	1
Religious	Covariate	0.68	0.47	0	1
Country-level controls					
Social protection expenditure	Covariate	11.97	7.67	0.5	32.07
GDP pc	Covariate	8.71	1.32	5.61	11.43
Unemployment rate	Covariate	8.7	5.97	0.48	29.83
Gini index	Covariate	38.73	9.4	23.7	64.8
Corruption Perception Index	Covariate	43.06	21.72	12	95
Quality of governance (WGI)	Covariate	0.03	0.89	−1.75	1.84

4. Data

Data used in this study comes mainly from the World Values Survey, an extensive dataset that collects information on multiple aspects regarding social and political life in nearly 100 countries worldwide through interviewing representative national samples of individuals. Samples are drawn from the entire population of 16 years and older without imposing an upper age limit using a stratified random sampling procedure. The first wave of the survey was conducted between 1981 and 1984 and, since then, there have been six more waves, covering the period 1981 to 2022.[4] Since we were interested in having as many observations as possible, we created a pseudo-panel dataset by merging information at the individual level from different waves. This approach is particularly appropriate in the context of subjective well-being, as self-reported measures can be influenced by temporary factors or individual-specific biases. Grouping similar individuals across waves helps mitigate these issues and allows for a more reliable identification of the systematic factors associated with variations in subjective well-being outcomes. After dropping the observations due to the presence of missing data in some variables, our final sample consists of 214,292 observations from 74 different countries. Badunenko et al. (2026) provide the data and the replication package.

The WVS dataset provides information on our main variable of our interest, subjective well-being or life satisfaction. Specifically, the former is derived from individuals' responses to the following question: "*All things considered, how satisfied are you with your life as a whole these days?*". This accounts for their feelings about their lives as a whole, including both economic and non-economic factors that are difficult to measure (Frey and Stutzer, 2010).[5] Responses are based on a scale from 1, which means 'completely dissatisfied', to 10, meaning 'completely satisfied'.

In addition, the WVS contains information on individual socioeconomic and demographic characteristics that we have incorporated as inputs in our analysis. We have chosen three variables that fulfill the requirement of isotonicity or monotonicity, which have been also used in some previous studies (e.g., Binder and Broekel, 2012; Cordero et al., 2017). The first one is the level of education completed, which is grouped into eight different categories according to the total number of years of completed education, the second one is the level of income, represented by the declared relative position of the individuals in the income distribution of their country expressed in deciles (ten different levels),[6] and the third one is an indicator of the health status perceived by the individuals on a five-level scale (very poor, poor, fair, good, or very good).

As additional explanatory variables, we have selected several variables that previous literature identifies as the most common predictors of subjective well-being (e.g., Blanchflower and Oswald, 2008; Stone et al., 2018), among which we distinguish between personal characteristics, also retrieved from WVS pooled dataset, and other indicators reflecting the country's economic and social position collected from different sources. The selected covariates (control variables) at the individual level are the age (continuous variable) and four dummy variables representing the gender of the individuals (female takes the value 1), their employment status (unemployed takes the value 1), marital status (married or living together as married takes the value 1) and being religious. At the country level, we have also selected several indicators previously identified in the literature as potential factors influencing individual life satisfaction (e.g., Bjørnskov et al., 2010; O'Connor, 2017), including several economic indicators such as the proportion of public expenditure devoted to social protection, which can be interpreted as a proxy of welfare-state policies, the gross domestic product (GDP) per capita, included to account for economic development, and two variables reflecting economic inequality, the unemployment rate and the Gini index. The first indicator was collected from the World Social Security Report, while the other three variables were available in the World Bank Open Data section. Likewise, we have also included data on two variables representing institutional quality, such as the corruption perception index (CPI) developed by Transparency International and an aggregate index constructed as the mean of the six subcomponents from the Worldwide Governance Indicators (WGI) available through the World Bank (Langbein and Knack, 2010; Kaufmann et al., 2014).[7]

[4] The second-, third-, fourth-, fifth-, sixth- and seventh-wave data were collected from 1990 to 1994, from 1995 to 1998, from 1999 to 2004, from 2005 to 2009, from 2010 to 2014, and from 2017 to 2022, respectively.

[5] The dataset also provides information about the level of happiness, but we decided not to use this indicator because it could be more influenced by emotions or feelings, while life satisfaction involves a more cognitive construct (Nettle, 2005).

[6] Considering relative income is more frequent in the literature than absolute income (Clark et al., 2008).

[7] This variable has also been used in other previous empirical studies (e.g., Abdallah et al. (2008) or J. Helliwell and H.F. Huang (2008)).

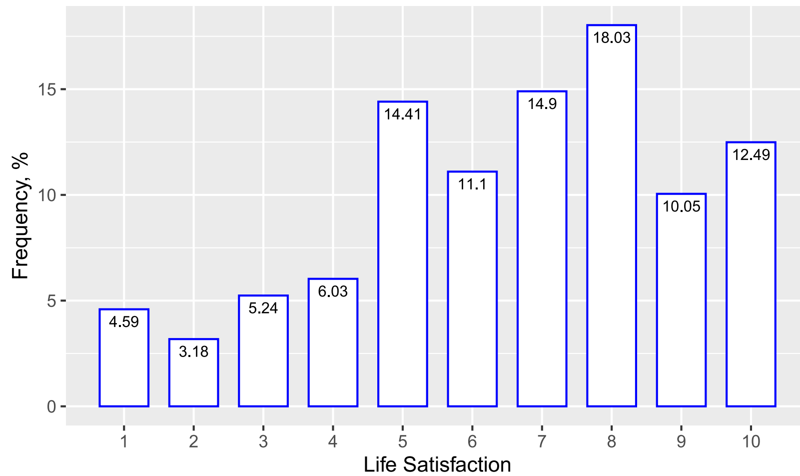


Fig. 1. The frequency of life satisfaction.

The summary statistics for all the variables included in our analysis are reported in Table 1. The analysis of these values allows us to note that, in line with previous studies (see, for example, Headey and Muffels, 2018; Moro-Egido et al., 2022) individuals included in the sample seem to be quite satisfied with their lives (mean value of 6.57 out of 10). Fig. 1 shows frequencies of observing individuals in life satisfaction categories. The declared level of income is significantly lower. The majority of individuals report that they are in good health and have a medium or medium-high level of education. With regard to control variables (also labeled as regressors/covariates in econometric models or contextual variables in nonparametric studies), we observe that our sample is almost evenly distributed by gender (women represent 51% of the individuals), the average age is slightly over 40 years, almost 70% of individuals declare to be religious, almost two-thirds are currently married or have a partner and less than 10% are unemployed. Finally, the most striking feature of the country-level variables is their large variability across countries in all the indicators, which is expected given the number of countries that we consider.

5. Empirical results

5.1. Regression results

Table 2 shows the estimated coefficients of the well-being frontier. All the estimated coefficients are highly significant, which implies valid statistical inference. However, these estimates cannot be interpreted as in a linear model. The regressors are predictors of probability, hence the marginal effect of regressors is the change in probability of being in a group defined in a latent way by estimated parameters $\boldsymbol{\mu}$. Due to the nonlinear nature of our model, the coefficients are not marginal effects. Appendix B.4 shows formulas for the calculation that is required to interpret the coefficients presented in Table 2.

Table 3 shows the marginal effects of continuous regressors on the probability of belonging to a life satisfaction category using formulas from Appendix B.4. Reported are the average partial effects for each category. To convert the interpretation to a percent, the marginal effects are multiplied by 100. First, note that the numbers in each column sum to zero since the probabilities sum to 1. The marginal effects in this table can be interpreted as follows. If, for example, the unemployment rate in the country increases by 1 percent, holding everything else fixed, the probability of observing a person in category 1 (the lowest) increases by 0.164%. The probability of observing an individual in higher categories gets smaller as the unemployment rate increases. For example, if it increases by 1 percent, holding everything else fixed, the probability of observing a person in category 10 (the highest) decreases by 0.382%. Thus, we find that higher unemployment rates are associated with lower levels of well-being, and that this relationship is non-linear.

The estimated associations for the Gini index and social protection expenditure differ from those commonly reported in studies examining the determinants of subjective well-being (e.g., Pacek and Radcliff, 2008; Verme, 2011). In particular, higher inequality is associated with a slightly greater probability of belonging to higher life satisfaction categories, whereas higher social protection expenditure is associated with a marginally lower probability. In any case, the estimated magnitudes are extremely small, implying a practically negligible influence on the estimated measures of well-being attainment. These findings should therefore be interpreted with considerable caution.

One possible explanation is that the present framework differs conceptually from conventional regression-based approaches commonly used in the subjective well-being literature. Most previous studies analyze how macroeconomic variables are associated with average levels of reported well-being. By contrast, our approach focuses on conditional well-being attainment, that is, the extent to which individuals achieve attainable life satisfaction levels given their observed individual and contextual characteristics. As a result, some macro-level variables may affect the estimated distance to the frontier differently from how they affect average subjective well-being levels in standard econometric models. In this context, the estimated associations may partly reflect cross-country institutional heterogeneity or differences in the way macroeconomic conditions interact with the efficiency of transforming available resources into subjective well-being across countries. The possibility that these patterns are driven by specific subsets of countries cannot be ruled out and deserves further investigation in future research using more disaggregated regional or country-group analyses.

Table 3 also reveals a negative association between the CPI and well-being. This is an expected result considering that this index is actually defined inversely, i.e., higher values reflect lower levels of corruption. Therefore, holding everything else fixed, the increase in the index (lower corruption) leads to a decrease in the probability of being in higher life satisfaction categories and increases the probability of being in lower categories, although again the values are quite small. The WGI index acts in the opposite direction and, in this case, the marginal effects do

Table 2
SF-ordered probit regression estimates. The dependent variable is life satisfaction.

Variable	Coefficient	z-value	Variable	Coefficient	z-value
factor(health)2	0.3251	9.03	factor(income)6	0.4250	41.91
factor(health)3	0.7799	22.11	factor(income)7	0.5060	46.97
factor(health)4	1.0892	30.83	factor(income)8	0.6005	49.70
factor(health)5	1.4064	39.53	factor(income)9	0.5969	37.11
factor(education)2	−0.0006	−0.06	factor(income)10	0.7275	42.48
factor(education)3	0.0040	0.36	log(gdppc)	0.1616	48.11
factor(education)4	−0.0553	−5.74	factor(wave)3	−0.3093	−15.37
factor(education)5	−0.0089	−0.79	factor(wave)4	−0.3702	−18.28
factor(education)6	−0.0674	−6.85	factor(wave)5	−0.2830	−14.22
factor(education)7	−0.0466	−3.89	factor(wave)6	−0.1665	−8.27
factor(education)8	−0.0251	−2.45	Unemployment rate	−0.0209	−44.97
factor(income)2	0.0860	8.48	Gini index	0.0154	50.66
factor(income)3	0.1244	12.74	Social protection expenditure	−0.0045	−9.18
factor(income)4	0.2581	26.70	Corruption Perception Index	−0.0071	−19.23
factor(income)5	0.3273	34.43	Quality of governance (WGI)	0.2261	26.62
$\ln \sigma^2_{u_i}$			μ		
(Intercept)	−5.3335	−20.47	μ_1	0.3063	6.23
Age	0.2156	17.73	μ_2	0.6136	12.51
Age squared	−0.0026	−17.26	μ_3	0.9592	19.58
Female	−0.3857	−10.12	μ_4	1.2538	25.61
Religious	−0.6448	−14.51	μ_5	1.7837	36.47
Unemployed	1.1912	23.17	μ_6	2.1266	43.48
Married	−1.1609	−21.31	μ_7	2.5639	52.42
			μ_8	3.1430	64.25
			μ_9	3.5702	72.96
N = 214 292, log-likelihood = −443 643.7998					

Table 3
Marginal effects on probability of belonging to a life satisfaction category.

Group	Unemployment rate	Gini index	Social protection expenditure	Corruption perception index	Quality of governance (WGI)
10	−0.382	0.280	−0.082	−0.130	4.13
9	−0.169	0.124	−0.036	−0.058	1.83
8	−0.165	0.121	−0.036	−0.056	1.79
7	−0.016	0.012	−0.003	−0.005	0.17
6	0.061	−0.045	0.013	0.021	−0.66
5	0.180	−0.133	0.039	0.062	−1.95
4	0.115	−0.084	0.025	0.039	−1.24
3	0.123	−0.090	0.027	0.042	−1.33
2	0.088	−0.065	0.019	0.030	−0.95
1	0.164	−0.121	0.035	0.056	−1.78

Notes: Reported are average marginal effects over all individuals in their respective category. The sum of marginal effects in each column equals zero.

represent relevant percentages. Thus, improvements in the quality of governance are associated with higher levels of well-being attainment. This is in line with previous evidence about the importance of this variable as a determinant of subjective well-being (Bjørnskov et al., 2008, 2010). All described effects are highly nonlinear. Consider positive value for category 1. The marginal effects remain virtually the same for categories 1 through 6, they then decrease until they reach the opposite sign by category 7 and then they fall even more whereby the magnitude doubles by category 10. The trend for the marginal effects that are negative for category 1 is the opposite.

Fig. 2 summarizes the marginal effects (multiplied by 100) associated with the dummy variables. The size and color of each marker reflect the magnitude of the marginal effect, while the exact value is reported next to the corresponding marker. As with continuous variables, the marginal effects across categories sum to zero. Consider first the role of health status. As health status improves, i.e., moves rightwards in the left panel of Fig. 2, the probability of belonging to lower life satisfaction categories decreases, whereas the probability of belonging to higher categories increases. For example, reporting a health status of 4 or 5, conditional on the remaining characteristics, is associated with reductions of 9.6 and 7.6 percentage points, respectively, in the probability of belonging to the lowest life satisfaction category. At the same time, the probability of belonging to the highest category increases by 21 and 34.6 percentage points, respectively. Overall, the results indicate a strong positive association between health status and subjective well-being attainment.

The relationship between education and subjective well-being appears considerably weaker (right panel of Fig. 2). Moving from left to right across educational categories, higher educational attainment is associated with relatively small changes in the probability of belonging to different life satisfaction categories. In contrast to the clear pattern observed for health status, the estimated marginal effects for education are more modest and less systematic across categories. Depending on the educational level considered, higher education may be associated with either small increases or small decreases in the probability of belonging to higher life satisfaction categories, while the changes in the probability of belonging to lower categories also remain limited in magnitude. Taken together, these results suggest that the association between educational attainment and subjective well-being is comparatively weak once the remaining individual and contextual characteristics are taken into account.

Income also exhibits a strong association with subjective well-being attainment (bottom panel of Fig. 2). As income levels increase (moving from left to right across income categories), the probability of belonging to higher life satisfaction categories tends to rise, whereas the probability of

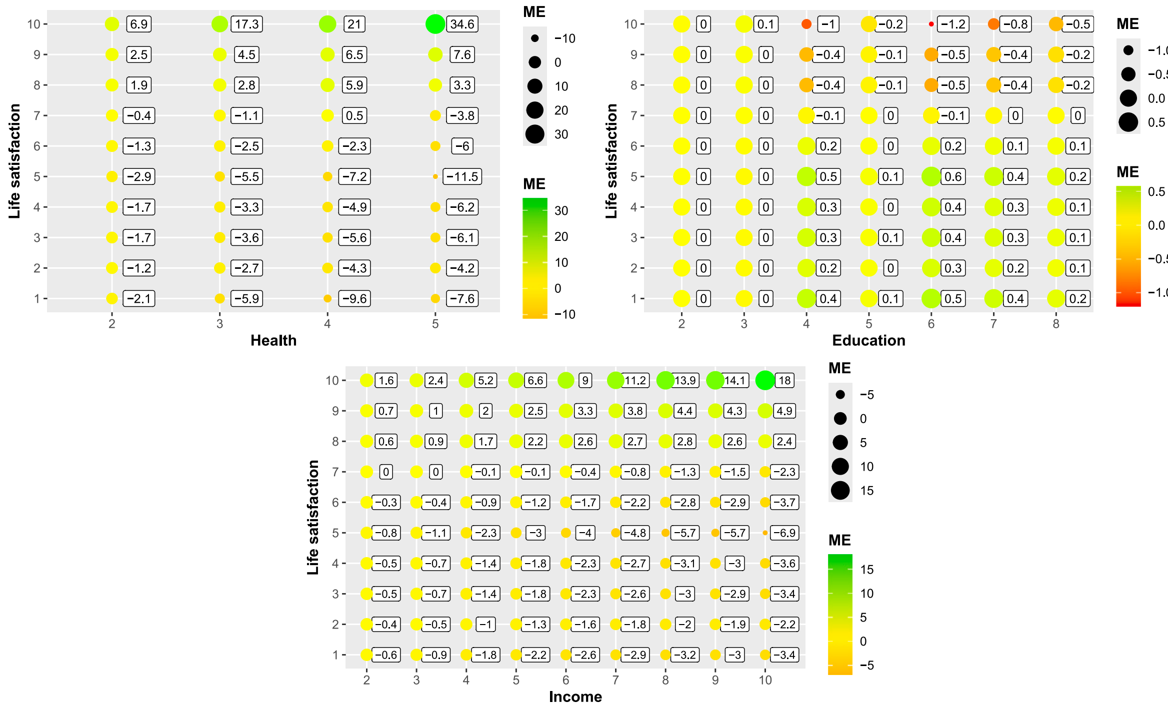


Fig. 2. Marginal effects of health and education on the probability of belonging to a life satisfaction category. Reported are average marginal effects over all individuals in their respective category. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

belonging to lower categories declines. This pattern is similar to the one previously observed for health status, although the estimated magnitudes vary across categories. When these marginal effects are considered alongside the observed distribution of life satisfaction categories presented in Fig. 1, the marginal effects of both health and education are quite substantial. For instance, approximately 12.5 percent of individuals are located in the highest life satisfaction category in the sample. Conditional on the remaining observed characteristics, belonging to the highest income category is associated with an increase in the predicted probability of belonging to this top category to around 30 percent.

5.2. Well-being attainment

To better understand well-being shortfall, discussed in Section 3, recall that for an individual i the shortfall represents the *gap* between the potential level of latent well-being (y_i^* – the maximum that the individual i can attain) and its realized level. Formally, this shortfall is defined with respect to y_i^*, the latent well-being variable introduced in Eq. (1). While y_i^* itself is unobserved, individuals report their life satisfaction on a discrete, ordinal scale which poses a challenge in both theory and application. These reported outcomes are assumed to reflect intervals of the underlying continuous latent variable y_i^*, separated by a set of cut-off points μ_m. Because the same cut-offs are used to map latent well-being into observed life satisfaction categories, an individual whose latent well-being y_i^* lies below its potential level $\bar{y}_i^*$ may report a lower life satisfaction category purely because their latent position falls into a lower interval. Using the estimated cut-off points reported in Table 2, we examine how individuals would be placed into life satisfaction categories in the absence of a shortfall. Specifically, we calculate $x_i\hat{\beta}$, which we refer to as the *frontier level* of latent well-being, and compare it with $x_i\hat{\beta} - E(u_i|\epsilon_i)$, which we label *with a shortfall*. Fig. 3 shows the estimated densities of the frontier and shortfall latent well-being measures together with the estimated cut-off points. The majority of individuals have lower latent well-being. Applying these cut-offs, we reassign individuals to life satisfaction categories under both the frontier and shortfall scenarios. We find that 52.5% of individuals (112,543) do not change their reported category (see Table 4). However, 45.8% (98,136) would move up by one satisfaction category in the absence of a shortfall, 1.64% (3512) by two categories, and 0.05% (101) by as many as three categories. When these counterfactual categories are compared with the observed levels of life satisfaction, improvements are found across the entire distribution, although the gains are more pronounced among individuals reporting lower initial levels of satisfaction. The results suggest no change for 57% of individuals in the second highest satisfaction category versus 46% in the lowest satisfaction category.

The results indicate that well-being shortfall has meaningful implications for observed distributions. While just over half of individuals would remain in the same reported life satisfaction category even in the absence of a shortfall, a substantial share would experience an upward reclassification, most frequently by one category. Importantly, these potential gains are not confined to individuals at the lower end of the distribution but occur across the entire range of reported measures. Therefore, latent well-being shortfalls are present across the entire life satisfaction distribution, implying that policies aimed at reducing such shortfalls could potentially contribute to broad improvements in subjective well-being.

Table 5 presents descriptive statistics of well-being attainment after classifying the countries into different groups according to the geographical area to which they belong (see Appendix A.1 in Appendix A). This information is complemented by the estimated densities reported in Fig. 4.

Despite the significant variations in average life satisfaction levels across different countries reported in Fig. 5, with Latin American and Caribbean nations consistently reporting the highest scores worldwide as has been documented on multiple occasions in the literature (Rojas, 2024), the results presented in Table 5 and Fig. 4 indicate that there is no discernible difference in the distribution of well-being efficiency among

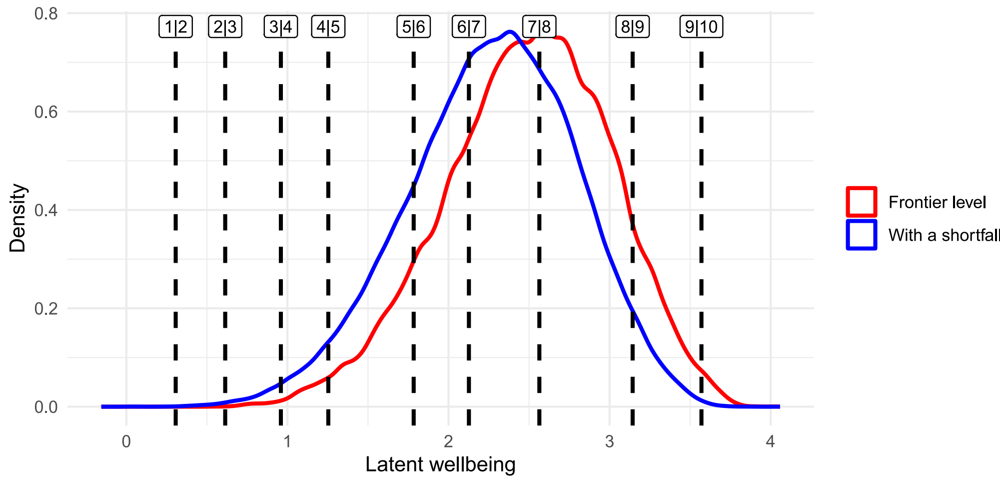


Fig. 3. Kernel densities of estimated frontier and shortfall latent well-being measures together with the estimated cut-off points.

Table 4
Distribution of life satisfaction category improvements by actual category (%).

Life satisfaction category	Number of categories moved up			
	0	1	2	3
9	56.7	42.5	0.78	0.01
8	56.2	42.9	0.92	0.03
7	53.0	45.5	1.38	0.03
6	51.7	46.7	1.53	0.04
5	48.5	49.4	2.10	0.06
4	46.4	50.8	2.68	0.06
3	45.9	50.8	3.23	0.11
2	47.0	49.7	3.21	0.15
1	46.2	49.8	3.88	0.16

Notes: Columns report the percentage distribution of improvements in life satisfaction categories when moving from the shortfall to the frontier latent well-being measure. Percentages sum to 100 within each row.

Table 5
Summary statistics of well-being attainment by group.

Group	N	sd	Min	Mean	Median	Max
Africa	33 861	0.09	0.17	0.79	0.81	1.00
Asia	25 100	0.07	0.22	0.81	0.83	0.99
Eastern Europe	16 810	0.09	0.15	0.82	0.83	0.99
Latin and Central America	34 206	0.08	0.17	0.81	0.83	1.00
Middle East	14 208	0.08	0.26	0.80	0.82	0.99
OECD	25 033	0.09	0.16	0.80	0.82	1.00
Other	10 990	0.08	0.13	0.79	0.80	0.99
Post Soviet	32 304	0.09	0.18	0.81	0.83	0.99
Western Europe	21 780	0.10	0.19	0.81	0.82	0.99
Total	214 292	0.09	0.13	0.80	0.82	1.00

these groups. This result can be explained by the fact that once income, health, education, and other individual circumstances are taken into account, much of the observed heterogeneity in well-being across countries is absorbed by these factors. The remaining differences are largely shaped by universal mechanisms highlighted in the welfare economics literature, such as diminishing marginal utility of income, hedonic adaptation, and relative comparisons, which help to explain why well-being efficiency levels appear considerably more homogeneous across nations (Sarracino and O'Connor, 2022).

Fig. 6 illustrates the distribution of well-being attainment based on life satisfaction levels. The black solid curve represents the density of attainment for all individuals in the sample. The lower panel reveals that individuals with relatively low life satisfaction levels also have lower well-being attainment. Conversely, the top panel shows that those with extremely high life satisfaction levels have a very low shortfall in well-being. These results suggest that individuals reporting very low life satisfaction levels systematically exhibit larger estimated well-being shortfalls. From a descriptive perspective, this pattern may help identify population groups facing greater difficulties in transforming available resources into subjective well-being. Interestingly, this finding is broadly consistent with evidence from intervention studies showing that the largest improvements in subjective well-being are often observed among individuals starting from lower baseline levels (Tiley et al., 2024; Zhou et al., 2023).

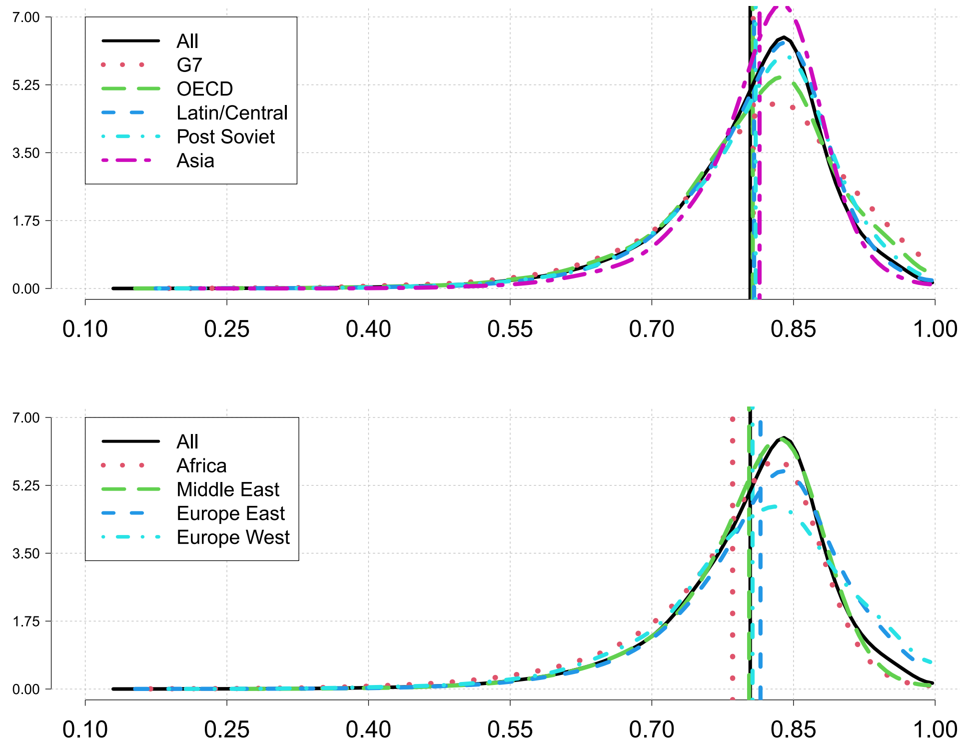


Fig. 4. Kernel densities of estimated well-being attainment by groups of countries.

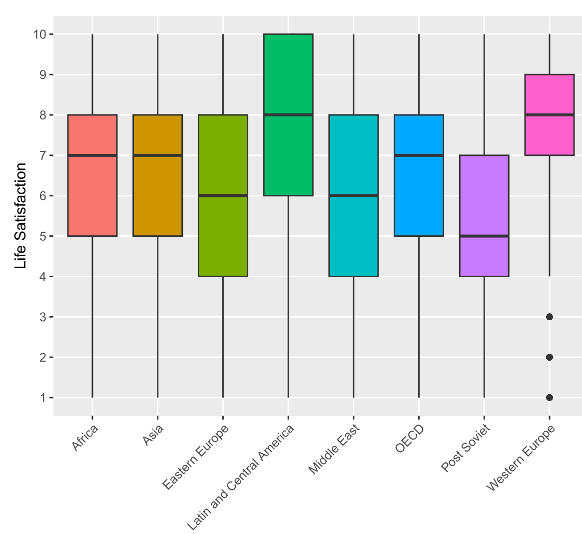


Fig. 5. Life satisfaction by group of countries.

5.3. *Treating the well-being indicator as a continuous variable*

Before turning to the discussion of other results, it is instructive to examine what happens when the standard SF model is applied to the data while treating the well-being indicator as a continuous variable. Table C.1 reports the estimated coefficients from the ordered probit stochastic frontier (OPSF), the ordered probit (OP) model that does not account for inefficiency, and the standard SF model. The frontier coefficients obtained from the OPSF and OP models are directly comparable, and a similar comparability holds for the coefficients of the inefficiency determinants between the OPSF and SF models. By contrast, the frontier coefficients of the OPSF and SF models are fundamentally not comparable. In the OPSF, the frontier coefficients affect the probability of belonging to particular well-being categories in a nonlinear way, whereas in the SF model they are interpreted as semi-elasticities of the numerical well-being category label with respect to the regressors. The latter interpretation lacks intuitive meaning. Indeed, Table C.2 demonstrates that applying the SF model to the well-being category label Y and rescaling it by adding constants of 6, 10, 20, 30, or 100 leads to frontier and inefficiency coefficients that differ markedly in magnitude and, in some cases, even in sign. Fig. C.1 further shows that such "relabeling" alters the variation of the well-being indicator, producing entirely different shapes of the efficiency distributions on both own and common scales. Note also that when the well-being category label takes values from 1 to 10 (model m0, top-left panel in Fig. C.1),

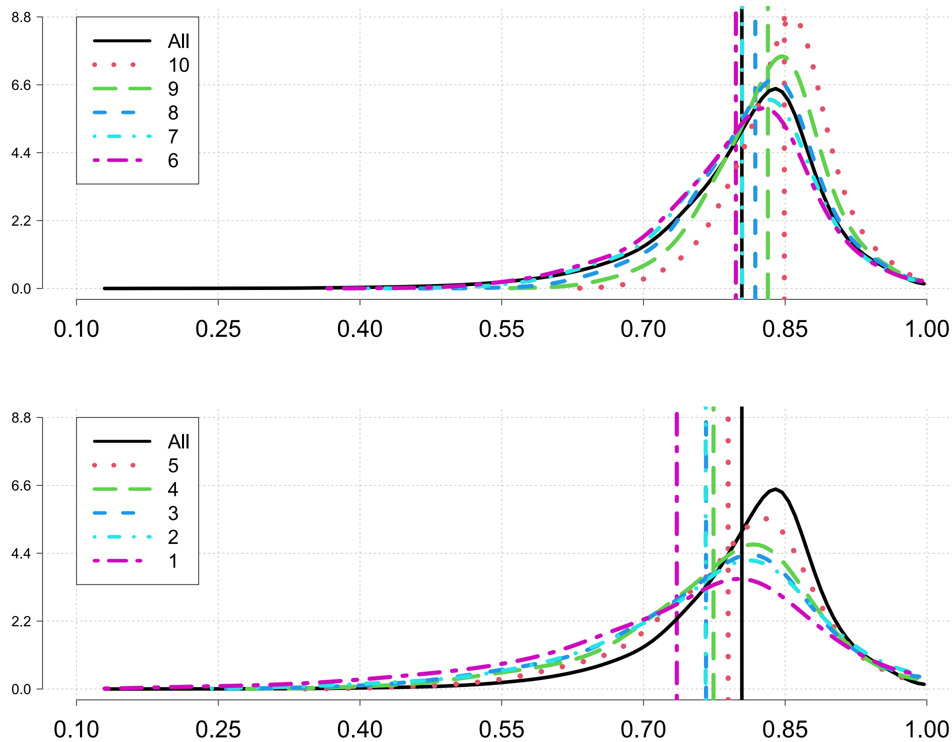


Fig. 6. Kernel densities of estimated well-being attainment by degree of life satisfaction. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Marginal effects on well-being attainment.

Life satisfaction		Marginal effect (multiplied by 100)				
Category	Frequency, %	Female	Religious	Unemployed	Married	Age
10	12.49	0.26	0.45	−0.89	0.82	−0.208
9	10.05	0.25	0.43	−0.89	0.80	−0.177
8	18.03	0.50	0.87	−1.84	1.62	−0.122
7	14.90	0.46	0.80	−1.71	1.48	−0.122
6	11.10	0.36	0.63	−1.36	1.18	−0.106
5	14.41	0.50	0.89	−1.91	1.64	−0.079
4	6.03	0.23	0.41	−0.88	0.75	−0.048
3	5.24	0.21	0.37	−0.80	0.69	−0.034
2	3.18	0.13	0.24	−0.51	0.43	−0.023
1	4.59	0.22	0.41	−0.88	0.74	4e−04

Notes: Reported are average marginal effects over all individuals in their respective life satisfaction categories. The marginal effects do not sum to zero.

the estimated efficiencies take exactly ten distinct values with an interval of 0.1 (that is, 0.1, 0.2, …, 1), implying that a categorical left-hand-side variable mechanically generates categorical efficiency estimates.

Consequently, depending on the arbitrary labeling of the ordinal outcome, the standard SF model can be manipulated to yield virtually any result in terms of frontier coefficients, inefficiency determinants, and the shape of the efficiency distribution. In contrast, the proposed OPSF model is *invariant* to the labeling of the ordinal dependent variable.

5.4. Determinants of well-being attainment and their effect on life satisfaction

Our next goal is to examine the marginal effect of z_{ki} on the well-being attainment ($\exp(-u_i)$) and probability ($\mathrm{Prob}(y_i = m)$) discussed in Section 3 and Appendix B.

5.4.1. Marginal effects of z_k on well-being attainment

We have observed quite a large shortfall in well-being attainment. Now we would like to explore possible explanatory factors behind this shortfall. Table 6 summarizes the marginal effects of determinants $\boldsymbol{z}$. The scale of attainment in this case is 0 to 100. Except for age, all determinants are dummy variables. Note, that here the column sums are not equal to 0 as was the case for the probabilities.

In general terms, our results align with previous empirical evidence based on traditional econometric techniques regarding the influence of gender, religiosity, marital status, and employment on self-reported levels of subjective well-being (Diener et al., 2018; Clark, 2018). Men experience a greater shortfall compared to women across all life satisfaction categories. The largest gender differences in well-being attainment are observed in categories 5 to 8. Conditional on the remaining characteristics, women exhibit a 0.26 higher probability of belonging to the highest life

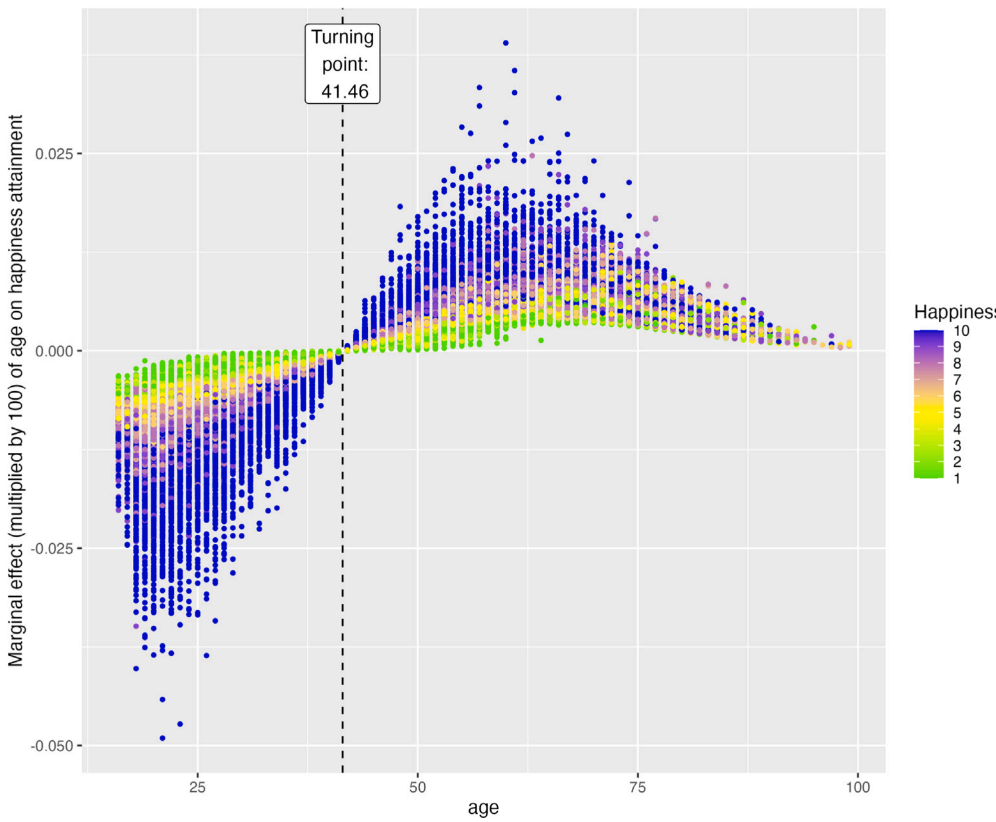


Fig. 7. Marginal effect of age on well-being attainment. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

satisfaction category (10) compared with men. Interestingly, religiosity is also associated with higher levels of subjective well-being attainment, with an estimated magnitude approximately twice as large as that associated with the gender indicator. The third dummy variable, "married", is positively associated with well-being attainment. For instance, being married is associated with a 0.82 higher probability of belonging to the highest life satisfaction category (10), representing one of the largest estimated associations among the individual characteristics considered. The magnitude of this relationship is particularly pronounced for the intermediate life satisfaction categories. In contrast, unemployment is strongly associated with lower levels of well-being attainment and with larger shortfalls relative to the estimated potential well-being frontier.

The relationship between age and well-being attainment deserves special attention, as the specification allows for a nonlinear pattern. The last column in Table 6 reports negative marginal effects, suggesting that, on average, well-being attainment tends to be lower among older individuals. The marginal effect for a continuous variable can provide more information than that of a dummy variable due to its greater variability (see the part "Marginal effects of z_k on $E(u)$" in Appendix B about nonlinear effect of continuous variables). Fig. 7 plots the estimated marginal effect of age on well-being attainment (scaled by 100) across the observed age range. Each point represents the marginal effect evaluated for each individual in our sample—which is larger than 200,000—at a given age, with color indicating their reported life satisfaction level. The vertical dashed line marks the estimated turning point (age of 41.46, considering coefficients of age and ages squared in Table 2, 0.2156/2/0.0026 = 41.46), where the marginal effect changes sign. Values above zero indicate that older individuals tend to have higher attainment, whereas values below zero suggest that older individuals tend to have lower attainment. The figure visualizes the distribution and heterogeneity of marginal effects across both age and life satisfaction levels, rather than a smoothed or averaged relationship. We observe that the marginal effects become larger as individuals move toward higher life satisfaction categories. Up to approximately 41 years of age, the marginal association between age and well-being attainment is negative, although its magnitude gradually declines. This pattern is particularly pronounced among individuals located in the higher life satisfaction categories. After the age of 42, the relationship becomes positive and increases progressively in magnitude. Around the age of 65, an additional year of age is associated with a reduction of approximately one percentage point in well-being attainment, although this association weakens at older ages. Finally, the results also indicate that, among individuals located in the lowest life satisfaction categories, age is associated with only very small changes in well-being attainment, suggesting a relatively stable pattern across the life cycle for this group.

5.4.2. *Marginal effects of z_k on* $\mathtt{Prob}(y_i = m)$

Table 7 shows the marginal effects of shortfall determinants on the probability of belonging to a life satisfaction category. Reported are the average effects in the respective category. They are multiplied by 100 for ease of interpretation. Here we can notice that, in general terms, our findings are broadly consistent with most of the available evidence on the determinants of subjective well-being summarized in Section 2. For example, concerning gender, women are 0.848 percent more likely than men to be among the highest category of life satisfaction. On the contrary, men are 0.506 percent more likely than women to be placed in the lowest category of life satisfaction. The effect of being religious is even greater, in the sense that people who consider themselves to be religious are much more likely to belong to the highest category in terms of well-being.

Table 7
Marginal effects on $\text{Prob}(y_i = m)$.

Life satisfaction		Marginal effect (multiplied by 100)				
Category	Frequency, %	Female	Religious	Unemployed	Married	Age
10	12.49	0.712	1.23	−2.49	2.25	−0.0309
9	10.05	0.329	0.57	−1.22	1.06	−0.0138
8	18.03	0.340	0.60	−1.34	1.11	−0.0126
7	14.90	0.056	0.11	−0.30	0.20	0.0004
6	11.10	−0.100	−0.16	0.30	−0.30	0.0065
5	14.41	−0.334	−0.57	1.19	−1.06	0.0163
4	6.03	−0.223	−0.39	0.83	−0.72	0.0095
3	5.24	−0.245	−0.43	0.93	−0.79	0.0094
2	3.18	−0.180	−0.32	0.69	−0.58	0.0062
1	4.59	−0.355	−0.64	1.41	−1.16	0.0091

Notes: Reported are average marginal effects over all individuals. The sum of marginal effects in each column equals zero.

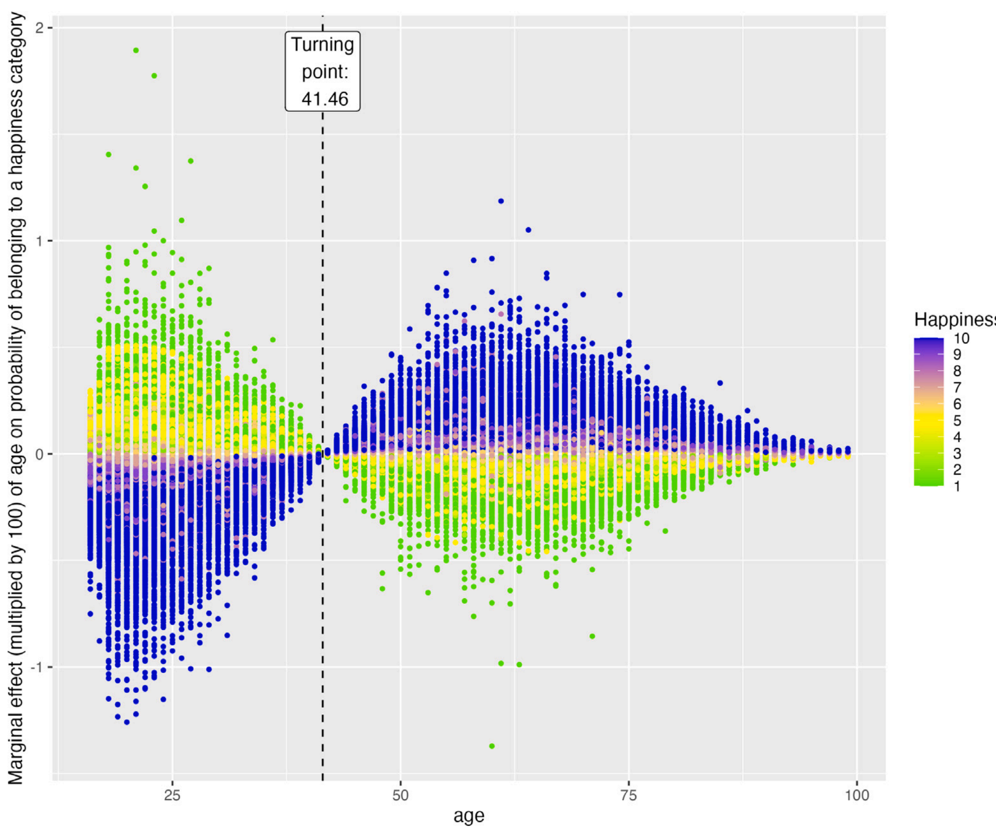


Fig. 8. Marginal effect of age on probability to be in a given life satisfaction group. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7 also suggests that marital status is strongly associated with subjective well-being. Married individuals are associated with a 2.25 percentage point higher probability of belonging to the highest life satisfaction category and a 1.06 percentage point higher probability of belonging to category 9. In addition, being married or living with a partner is associated with a 1.16 percentage point lower probability of belonging to the lowest life satisfaction category. Considering that only 4.59 percent of individuals are located in this lowest category (see column 2), the magnitude of this association also appears economically meaningful. Finally, unemployment exhibits the strongest association with subjective well-being among the individual characteristics considered. Unemployed individuals are associated with a 2.49 percentage point lower probability of belonging to the highest life satisfaction category, which represents 12.49 percent of the sample. At the lower end of the life satisfaction distribution, unemployment is associated with a 1.41 percentage point higher probability of belonging to the lowest category, whose observed frequency is 4.59 percent. The last column of Table 7 also indicates a negative association between age and well-being attainment. However, as previously discussed, this relationship is nonlinear for continuous variables. Fig. 8 shows pointwise estimates of the marginal effect of age (×100) on the probability of belonging to a life satisfaction category. Points are color-coded by reported categories. The dashed line marks the turning point at age 41.46. Positive values indicate a higher likelihood among older individuals, while negative values indicate a lower likelihood, highlighting heterogeneity rather than an averaged age profile. Up to 41 years of age, an additional year of age is associated with lower levels of well-being attainment, whereas after the age of 42, the association gradually becomes positive. This result is consistent with the well-documented U-shaped relationship between age and subjective well-being (Cheng et al., 2017).

6. Conclusion

This paper studies subjective well-being by estimating how individuals attain life satisfaction levels relative to their observed resources and personal characteristics. To operationalize this idea, we develop a new SF model tailored to ordered outcome variables by extending the standard ordered probit model that allows for inefficiency.

The proposed model is derived from a generic random utility framework, leading to a nonlinear specification in which interpretation requires additional structure. We provide closed-form expressions for the likelihood function, efficiency estimates, and marginal effects, and offer detailed guidance on how to implement the model in practice. In particular, we clarify the interpretation of frontier coefficients and show how to compute marginal effects of both continuous and categorical inputs on the outcome variable. We further demonstrate how to recover (in) efficiency measures and evaluate the role of environmental variables in shaping both (in) efficiency and observed outcomes. More broadly, the framework is applicable to a wide range of settings involving ordered categorical outcomes, such as credit ratings, consumer evaluations, or satisfaction measures across different domains. In our present application inefficiency is interpreted as short fall in well-being.

We have estimated a well-being frontier for a very large sample of almost 215,000 individuals from 74 different countries participating in different waves of the World Value Survey. Our results show that the overall average efficiency levels is around 80 percent. This value is close to those estimated in some recent studies (e.g., Mamatzakis and Tsionas, 2021; Cordero et al., 2021), which suggests that individuals in our sample still have the capacity to make better use of their available resources to achieve the well-being frontier.

Across countries, we find no meaningful differences in individuals' average well-being efficiency. This is a striking result given the substantial cross-country variation in reported well-being levels. This suggests that conditional well-being attainment patterns are relatively similar across countries once observable individual characteristics are taken into account. Among the macro-level determinants, governance quality is positively associated with well-being attainment, corroborating previous evidence highlighting the importance of institutional quality for subjective well-being (JohnF. Helliwell and Haifang Huang, 2008).

Looking at individuals' determinants, we show that women, religious people, and those being married have higher probabilities of reporting higher levels of well-being and are more likely to be among the individuals reporting higher life satisfaction levels. Moreover, the employment status turns out to be an even more relevant factor in explaining the level of well-being of individuals, which is much lower for the unemployed, while in the case of age, the relationship is U-shaped, reaching the lowest values between the ages of 40 and 45. We find that individuals reporting very low levels of life satisfaction exhibit the largest shortfalls relative to the estimated well-being frontier. In contrast, individuals located in the highest life satisfaction categories display comparatively small shortfalls. The results also indicate that, among individuals with the lowest levels of life satisfaction, age is associated with only limited variation in well-being attainment after controlling for the remaining characteristics. More generally, these findings suggest that well-being shortfalls are particularly concentrated among individuals located at the lower end of the life satisfaction distribution.

The results also provide relevant insights into the distribution of well-being shortfalls across individuals and social groups. In particular, individuals reporting lower levels of life satisfaction tend to exhibit substantially larger shortfalls relative to the estimated well-being frontier, whereas individuals located in the highest life satisfaction categories display comparatively small shortfalls. This pattern suggests that disparities in well-being attainment are not uniformly distributed across the population and highlights the potential relevance of factors such as labour market conditions, health, and institutional quality in understanding these differences. However, these findings should be interpreted with appropriate caution. Given the cross-sectional nature of the data, the estimated relationships cannot be interpreted in causal terms and should instead be understood as conditional associations. Moreover, the estimated well-being shortfalls may reflect multiple factors beyond inefficiency in a strict economic sense, including unobserved preferences, adaptation mechanisms, psychological traits, or systematic differences in reporting behavior for which it is not possible to fully control within the empirical specification.

Despite these limitations, by explicitly accounting for the ordinal nature of subjective well-being measures and jointly modeling well-being attainment and its determinants within a stochastic frontier framework, the proposed approach represents an important methodological improvement over conventional frontier models that treat subjective well-being indicators as continuous variables. More broadly, the proposed framework may also provide useful complementary information for policy analysis by shifting attention from average subjective well-being levels to differences in conditional well-being attainment across individuals and countries. In this respect, the approach may help identify groups systematically associated with larger well-being shortfalls relative to their observed endowments, thereby contributing to a more nuanced understanding of inequalities in subjective well-being beyond conventional mean-based analyses.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Sample

A.1. Groups of countries used

G7	(1) Canada, (2) France, (3) Germany, (4) Italy, (5) Japan, (6) United States, (7) Great Britain
OECD	(1) Australia, (2) Canada, (3) Chile, (4) Estonia, (5) Finland, (6) France, (7) Germany, (8) Hungary, (9) Italy, (10) Japan, (11) South Korea, (12) Latvia, (13) Lithuania, (14) Mexico, (15) Netherlands, (16) New Zealand, (17) Norway, (18) Poland, (19) Slovakia, (20) Slovenia, (21) Spain, (22) Sweden, (23) Switzerland, (24) Turkey, (25) United States, (26) Great Britain, (27) Czech Rep.
Latin Am	(1) Argentina, (2) Brazil, (3) Chile, (4) Colombia, (5) Ecuador, (6) El Salvador, (7) Guatemala, (8) Mexico, (9) Peru, (10) Puerto Rico, (11) Trinidad and Tobago, (12) Uruguay, (13) Venezuela

Post Soviet	(1) Azerbaijan, (2) Armenia, (3) Belarus, (4) Georgia, (5) Kazakhstan, (6) Kyrgyzstan, (7) Moldova, (8) Russia, (9) Ukraine, (10) Uzbekistan
Asia	(1) Bangladesh, (2) Taiwan, (3) India, (4) Indonesia, (5) Malaysia, (6) Pakistan, (7) Philippines, (8) Thailand, (9) Viet Nam
Africa	(1) Algeria, (2) Ghana, (3) Libya, (4) Mali, (5) Morocco, (6) Nigeria, (7) Rwanda, (8) South Africa, (9) Zimbabwe, (10) Tunisia, (11) Uganda, (12) Egypt, (13) Tanzania, (14) Zambia
M East	(1) Palestine, (2) Iran, (3) Iraq, (4) Jordan, (5) Lebanon, (6) Yemen
Europe E	(1) Albania, (2) Bulgaria, (3) Hungary, (4) Montenegro, (5) Poland, (6) Romania, (7) Serbia, (8) Slovakia, (9) Slovenia, (10) Macedonia, (11) Serbia and Montenegro, (12) Bosnia, (13) Czech Rep.
Europe W	(1) Finland, (2) France, (3) Germany, (4) Italy, (5) Netherlands, (6) Norway, (7) Spain, (8) Sweden, (9) Switzerland, (10) Great Britain

Appendix B. Mathematical proofs and derivations

B.1. Auxiliary lemmas

Lemma 1. *An integral of type*

$$\int_0^{\infty} \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)du_i, \tag{B.1}$$

is equal to

$$\Phi\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) + 2T\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, p\sigma_u\right), \tag{B.2}$$

where T is the Owen's T function.

Proof. We know that

$$\begin{aligned}\int \phi(\boldsymbol{x})\Phi(a+b\boldsymbol{x})\,d\boldsymbol{x} = {} & T\left(\boldsymbol{x}, \frac{a}{\boldsymbol{x}\sqrt{1+b^2}}\right) + T\left(\frac{a}{\sqrt{1+b^2}}, \frac{\boldsymbol{x}\sqrt{1+b^2}}{a}\right)\\ & - T\left(\boldsymbol{x}, \frac{a+b\boldsymbol{x}}{\boldsymbol{x}}\right) - T\left(\frac{a}{\sqrt{1+b^2}}, \frac{ab+\boldsymbol{x}(1+b^2)}{a}\right)\\ & + \Phi(\boldsymbol{x})\Phi\left(\frac{a}{\sqrt{1+b^2}}\right)\end{aligned} \tag{B.3}$$

Denote $a_2 = \dfrac{a}{\sqrt{1+b^2}}$, then (B.3) becomes

$$\underbrace{T\left(\boldsymbol{x}, \frac{a_2}{\boldsymbol{x}}\right)}_{\text{term 1}} + \underbrace{T\left(a_2, \frac{\boldsymbol{x}}{a_2}\right)}_{\text{term 2}} - \underbrace{T\left(\boldsymbol{x}, b+\frac{a}{\boldsymbol{x}}\right)}_{\text{term 3}} - \underbrace{T\left(a_2, b+\frac{\boldsymbol{x}(1+b^2)}{a}\right)}_{\text{term 4}} + \underbrace{\Phi(\boldsymbol{x})\Phi(a_2)}_{\text{term 5}}$$

replace $z = u/\sigma_u$ in (B.1), then

$$\begin{aligned}&\int_0^{\infty} \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)du_i\\ &= \int_0^{\infty} \frac{2}{\sigma_u}\phi(z)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + \sigma_u pu_i)\sigma_u du_i\\ &= 2\int_0^{\infty} \phi(z)\Phi(a+bu_i)\sigma_u du_i\end{aligned}$$

where $a = \mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}$ and $b = \sigma_u p$. Integral (B.1) is needed on the domain $[0, +\infty)$.

We will use three properties of Owen's T function to simplify the integral

$$T(H, 0) = 0$$

$$T(0, a) = \frac{1}{2\pi}arctan(a)$$

Therefore,

$$T(0, +\infty) = 0.25$$

$$T(a, \infty) = \frac{1}{2}(1 - \Phi(|a|))$$

Denote $a_2 = \dfrac{a}{\sqrt{1+b^2}}$. Consider individual terms

$$(\text{term } 1)\Big|_{\boldsymbol{x}=0}^{\boldsymbol{x}=+\infty} = \underbrace{T\left(\boldsymbol{x}, \frac{a_2}{\boldsymbol{x}}\right)}_{\text{term 1}}\Bigg|_{\boldsymbol{x}=0}^{\boldsymbol{x}=+\infty} \tag{B.4}$$

$$= T\left(+\infty, \frac{a_2}{+\infty}\right) - T\left(0, \frac{a_2}{0}\right) = 0 - 0.25 = -0.25$$

$$(\text{term } 2)\Big|_{x=0}^{x=+\infty} = \underbrace{T\left(a_2, \frac{\boldsymbol{x}}{a_2}\right)}_{\text{term } 2}\Bigg|_{x=0}^{x=+\infty} \tag{B.5}$$

$$= T\left(a_2, +\infty\right) - T\left(a_2, \frac{0}{a_2}\right) = \frac{1}{2}(1 - \Phi(a_2))$$

$$(\text{term } 3)\Big|_{x=0}^{x=+\infty} = \underbrace{-T\left(\boldsymbol{x}, b + \frac{a}{\boldsymbol{x}}\right)}_{\text{term } 3}\Bigg|_{x=0}^{x=+\infty} \tag{B.6}$$

$$= -T\left(+\infty, b + \frac{a}{+\infty}\right) + T\left(0, b + \frac{a}{0}\right) = 0.25$$

$$(\text{term } 4)\Big|_{x=0}^{x=+\infty} = \underbrace{-T\left(a_2, b + \frac{\boldsymbol{x}(1+b^2)}{a}\right)}_{\text{term } 4}\Bigg|_{x=0}^{x=+\infty} \tag{B.7}$$

$$= -T\left(a_2, +\infty\right) + T\left(a_2, b + \frac{0(1+b^2)}{a}\right)$$

$$= -\frac{1}{2}(1 - \Phi(a_2)) + T(a_2, b) \tag{B.8}$$

$$(\text{term } 5)\Big|_{x=0}^{x=+\infty} = \underbrace{\Phi(\boldsymbol{x})\Phi\left(a_2\right)}_{\text{term } 5}\Bigg|_{x=0}^{x=+\infty} \tag{B.9}$$

$$= \Phi(+\infty)\Phi\left(a_2\right) - \Phi(0)\Phi\left(a_2\right) = \Phi\left(a_2\right) - 0.5\Phi\left(a_2\right) = 0.5\Phi\left(a_2\right)$$

Adding these 5 terms yields

$$\int_0^\infty \phi(\boldsymbol{x})\Phi(a + b\boldsymbol{x})\,d\boldsymbol{x} = T(a_2, b) + 0.5\Phi\left(a_2\right) = T\left(\frac{a}{\sqrt{1+b^2}}, b\right) + 0.5\Phi\left(\frac{a}{\sqrt{1+b^2}}\right)$$

which is also the integral `10,010.6` in Owen (1980). So (B.1), then becomes

$$\int_0^\infty \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)du_i$$
$$= \int_0^\infty \frac{2}{\sigma_u}\phi(z)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + \sigma_u p z_i)\sigma_u dz_i$$
$$= 2\int_0^\infty \phi(z)\Phi(a + bz_i)\sigma_u dz_i$$
$$= 2T\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u{}^2}}, p\sigma_u\right) + \Phi\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u{}^2}}\right) \quad \square$$

Lemma 2. *An integral of type*

$$\mathcal{I} = \int_0^\infty u\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)du_i \tag{B.10}$$

is equal to

$$2\sigma_u\left[\frac{p\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(\frac{\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\frac{-p\sigma_u(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) + \frac{1}{\sqrt{2\pi}}\Phi(\mu^* - \boldsymbol{x}_i'\boldsymbol{\beta})\right]. \tag{B.11}$$

Proof. Write

$$\mathcal{I} = \int_0^\infty u\frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right)\Phi(\mu^* - x\beta + pu)\,du \tag{B.12}$$

substitute $\frac{u}{\sigma_u} = z$, then $\frac{du}{\sigma_u} = dz$ and $u = \sigma_u z$

$$\mathcal{I} = \int_0^\infty \sigma_u z\frac{2}{\sigma_u}\phi(z)\Phi(\mu^* - x\beta + p\sigma_u z)\,\sigma_u dz \tag{B.13}$$

$$2\sigma_u\int_0^\infty z\phi(z)\Phi(a + bz)\,dz, \tag{B.14}$$

where $a = \mu^* - x\beta$ and $b = p\sigma_u$. Using integral `10,011.2` in Owen (1980), we can write

$$\int_0^\infty z\phi(z)\Phi(a + bz)\,dz = \frac{b}{\sqrt{1+b^2}}\phi\left(\frac{a}{\sqrt{1+b^2}}\right)\Phi\left(\frac{-ab}{\sqrt{1+b^2}}\right) + \frac{1}{\sqrt{2\pi}}\Phi(a) \quad \square \tag{B.15}$$

B.2. Proof of propositions

Proof of Proposition 1. The probability density function of the convolution of two error components, because of their independence, is

$$f(u_i, \epsilon_i) = \frac{2}{\sigma_u}\phi\left(\frac{u}{\sigma_u}\right) \times \sum_{m=1}^{M} I_m(y_i)p_{im}, \tag{B.16}$$

where $\phi(.)$ is the standard normal probability density function. To find the density of ϵ_i, we need to integrate u_i out from $\int_0^\infty f(u_i, \epsilon_i)du_i$. Using Lemma 1, the log-likelihood is given by

$$ll_i(\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\gamma}) = \ln\left\{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m\right\}, \tag{B.17}$$

where

$$\mathcal{L}_m = \begin{cases} \Phi\left(\dfrac{\mu_1 - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) + 2T\left(\dfrac{\mu_1 - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, p\sigma_u\right) & \text{if } y_i = 1 \\ \left[\begin{array}{l} \Phi\left(\dfrac{\mu_m - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) + 2T\left(\dfrac{\mu_m - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, p\sigma_u\right) \\ -\Phi\left(\dfrac{\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) - 2T\left(\dfrac{\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, p\sigma_u\right) \end{array}\right] & \text{if } 1 < y_i = m < M \\ \Phi\left(-\dfrac{\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right) + 2T\left(-\dfrac{\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}, -p\sigma_u\right) & \text{if } y_i = M. \end{cases} \tag{B.18}$$

which can be written succinctly as

$$\mathcal{L}_m = \begin{cases} \Phi\left(h_1\right) + 2T\left(h_1, b\right) & \text{if } y_i = 1 \\ \Phi\left(h_m\right) + 2T\left(h_m, b\right) - \Phi\left(h_{m-1}\right) - 2T\left(h_{m-1}, b\right) & \text{if } 1 < y_i = m < M \\ \Phi\left(-h_{M-1}\right) + 2T\left(-h_{M-1}, -b\right) & \text{if } y_i = M, \end{cases}$$

where T is the Owen's T function and $\zeta = 1/\sqrt{1+\sigma_u^2}$, $a_j = \mu_j - \mathbf{x}_i'\boldsymbol{\beta}$, $h_j = a_j\zeta$, and $b = p\sigma_u$. □

Proof of Proposition 2. Using Lemma 2, the prediction of u_i is given by

$$E(u_i|\epsilon_i) = 2\sigma_u\frac{\sum_{m=1}^{M} I_m(y_i)\mathcal{B}_m}{\sum_{m=1}^{M} I_m(y_i)\mathcal{L}_m}, \tag{B.19}$$

where

$$\mathcal{B}_m = \begin{cases} \left[\begin{array}{l} \dfrac{p\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(\dfrac{\mu_1 - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\dfrac{-p\sigma_u(\mu_1 - \mathbf{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) \\ +\dfrac{1}{\sqrt{2\pi}}\Phi(\mu_1 - \mathbf{x}_i'\boldsymbol{\beta}) \end{array}\right] & \text{if } y_i = 1 \\ \left[\begin{array}{l} \dfrac{p\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(\dfrac{\mu_m - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\dfrac{-p\sigma_u(\mu_m - \mathbf{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) \\ +\dfrac{1}{\sqrt{2\pi}}\Phi(\mu_m - \mathbf{x}_i'\boldsymbol{\beta}) \\ -\dfrac{p\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(\dfrac{\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\dfrac{-p\sigma_u(\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) \\ -\dfrac{1}{\sqrt{2\pi}}\Phi(\mu_{m-1} - \mathbf{x}_i'\boldsymbol{\beta}) \end{array}\right] & \text{if } 1 < y_i = m < M \\ \left[\begin{array}{l} \dfrac{p\sigma_u}{\sqrt{1+\sigma_u^2}}\phi\left(-\dfrac{\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta}}{\sqrt{1+\sigma_u^2}}\right)\Phi\left(\dfrac{-p\sigma_u(\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta})}{\sqrt{1+\sigma_u^2}}\right) \\ +\dfrac{1}{\sqrt{2\pi}}\Phi(-\left(\mu_{M-1} - \mathbf{x}_i'\boldsymbol{\beta}\right)) \end{array}\right] & \text{if } y_i = M, \end{cases}$$

which can be written succinctly as

$$\mathcal{B}_m = \begin{cases} b\zeta f\left(h_1\right) + \dfrac{\Phi\left(a_1\right)}{\sqrt{2\pi}} & \text{if } y_i = 1 \\ b\zeta f\left(h_m\right) + \dfrac{\Phi\left(a_m\right)}{\sqrt{2\pi}} - b\zeta f\left(h_{m-1}\right) - \dfrac{\Phi\left(a_{m-1}\right)}{\sqrt{2\pi}} & \text{if } 1 < y_i = m < M \\ b\zeta f\left(h_M\right) + \dfrac{\Phi\left(-a_M\right)}{\sqrt{2\pi}} & \text{if } y_i = M, \end{cases}$$

where

$$f(h_m) = \begin{cases} \phi\left(h_1\right)\Phi\left(-bh_1\right) & \text{if } y_i = 1 \\ \phi\left(h_m\right)\Phi\left(-bh_m\right) & \text{if } 1 < y_i = m < M \\ \phi\left(-h_{M-1}\right)\Phi\left(-bh_{M-1}\right) & \text{if } y_i = M \quad \square \end{cases} .$$

B.3. Gradients

Noting that

$$\frac{\partial T(h,a)}{\partial \omega} = T_1'(h,a)\frac{\partial h}{\partial \omega} + T_2'(h,a)\frac{\partial a}{\partial \omega}, \tag{B.20}$$

where

$$T_1'(h,a) = \frac{\partial T(h,a)}{\partial h} = -\frac{1}{2\sqrt{2\pi}}(2\Phi(ha)-1)\exp\left(-\frac{h^2}{2}\right) = -\phi(h)(\Phi(ha)-0.5)$$

and

$$T_2'(h,a) = \frac{\partial T(h,a)}{\partial a} = \frac{1}{2\pi(1+a^2)}\exp\left(-\frac{h^2}{2}\left(1+a^2\right)\right)$$

we can establish the following proposition.

Proposition 3. *The derivatives of the log-likelihood or the log of* (5) *are given by*

$$\frac{\partial ll_i}{\partial \beta} = -\boldsymbol{x}_i \frac{\mathcal{D}_{1m}}{\mathcal{L}_m} \tag{B.21}$$

$$\frac{\partial ll_i}{\partial \mu_i} = \frac{\mathcal{D}_{1m}}{\mathcal{L}_m}, \tag{B.22}$$

where

$$\mathcal{D}_{1m} = \begin{cases} 2\zeta f(h_1) & \text{if } y_i = 1 \\ 2\zeta\left(f(h_m) - f(h_{m-1})\right) & \text{if } 1 < y_i = m < M \\ -2\zeta f(h_{M-1}) & \text{if } y_i = M \end{cases} \tag{B.23}$$

and

$$f(h_m) = \begin{cases} \phi\left(h_1\right)\Phi\left(-bh_1\right) & \text{if } y_i = 1 \\ \phi\left(h_m\right)\Phi\left(-bh_m\right) & \text{if } 1 < y_i = m < M \\ \phi\left(-h_{M-1}\right)\Phi\left(-bh_{M-1}\right) & \text{if } y_i = M \end{cases} . \tag{B.24}$$

Specifying $\sigma_{u_i}^2 = \exp\left(\boldsymbol{z}_i\boldsymbol{\gamma}\right)$,

$$\frac{\partial ll_i}{\partial \gamma} = -z_i \frac{1}{\mathcal{L}_m}\mathcal{D}_{2m}, \tag{B.25}$$

where

$$\mathcal{D}_{2m} = \begin{cases} {\sigma_u}^2\zeta^2 h_1 f(h_1) - \exp\left(-\dfrac{h_1^2}{2\zeta^2}\right)\dfrac{a\zeta^2}{2\pi} & \text{if } y_i = 1 \\ \left[\begin{array}{l} {\sigma_u}^2\zeta^2 h_m f(h_m) - \exp\left(-\dfrac{h_m^2}{2\zeta^2}\right)\dfrac{a\zeta^2}{2\pi} \\ -{\sigma_u}^2\zeta^2 h_{m-1} f(h_{m-1}) + \exp\left(-\dfrac{h_{m-1}^2}{2\zeta^2}\right)\dfrac{a\zeta^2}{2\pi} \end{array}\right] & \text{if } 1 < y_i = m < M \\ {\sigma_u}^2\zeta^2 h_{M-1} f(h_{M-1}) - \exp\left(-\dfrac{h_{M-1}^2}{2\zeta^2}\right)\dfrac{a\zeta^2}{2\pi} & \text{if } y_i = M \end{cases} \tag{B.26}$$

B.4. Marginal effects

Marginal effects of x_k on $\texttt{Prob}(y_i = m)$

The marginal effects of a variable x_k on $\texttt{Prob}(y_i = m)$ are calculated differently for continuous and discrete (binary) variables. Due to the nonlinear nature of the ordered probit model, the magnitude of the marginal effect is not constant. It depends on the values of all the determinants $\boldsymbol{z}$, the regressors $\boldsymbol{x}$, and estimated parameters $\boldsymbol{\beta}$, $\boldsymbol{\mu}$ and $\boldsymbol{\gamma}$. If x_k is continuous, we first use (3),

$$\begin{aligned}\frac{\partial \texttt{Prob}(y_i = m)}{x_k} &= \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)\right]}{\partial x_k}\\ &\approx \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i))\right]}{\partial x_k}.\end{aligned} \tag{B.27}$$

Calculation of $E(u_i)$ is due to Proposition 2. Now we can spell out the marginal effect for each group, viz.,

$$\frac{\partial \texttt{Prob}(y_i = m)}{z_k} = \begin{cases} -\dfrac{\partial(\boldsymbol{x}_i'\boldsymbol{\beta})}{\partial x_k}\phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)) & \text{if } y_i = 1\\ -\dfrac{\partial(\boldsymbol{x}_i'\boldsymbol{\beta})}{\partial x_k}\begin{bmatrix}\phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i))\\ -\phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i))\end{bmatrix} & \text{if } 1 < y_i = m < M\\ \dfrac{\partial(\boldsymbol{x}_i'\boldsymbol{\beta})}{\partial x_k}\phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)) & \text{if } y_i = M\end{cases}. \tag{B.28}$$

The values in (B.28) are calculated for all observations and then the average is reported, aka APE.

For a discrete variable x_k, the partial effect of x_k on $\texttt{Prob}(y_i = m)$ is calculated as a change in probabilities at $x_k = 1$ and $x_k = 0$, which implies that only $\boldsymbol{x}_i'\boldsymbol{\beta}$ changes in

$$\begin{aligned}\Delta p_{im} &= \texttt{Prob}(y_i = m|z_k = 1) - \texttt{Prob}(y_i = m|z_k = 0)\\ &\approx \begin{cases}\begin{bmatrix}\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 1)\\ -\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 0)\end{bmatrix} & \text{if } y_i = 1\\ \begin{bmatrix}\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 1)\\ -\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 1)\\ -\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 0)\\ +\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 0)\end{bmatrix} & \text{if } 1 < y_i = m < M\\ \begin{bmatrix}\Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 0)\\ -\Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|x_k = 1)\end{bmatrix} & \text{if } y_i = M\end{cases}.\end{aligned} \tag{B.29}$$

Similar to the continuous case, the values in (B.29) are calculated for all observations and then the APE is reported. Note that the marginal effects in both (B.28) and (B.29) are conditioned on the data $\boldsymbol{x}$.

Marginal effects of z_k on $E(u)$

The marginal effects of a variable z_k on $E(u)$ are differently calculated for continuous and dummy variables. If the z_k is continuous, the marginal effects of a variable z_k on inefficiency are

$$\frac{\partial E(u_i)}{\partial z_k} = 2\frac{\mathcal{B}}{\mathcal{L}}d_1 + 2\sigma_u\frac{(\mathcal{B}_z\mathcal{L} - \mathcal{B}\mathcal{L}_z)}{\mathcal{L}^2}, \tag{B.30}$$

where $\mathcal{B} = \sum_{m=1}^{M} I_m(y_i)\mathcal{B}_m$, $\mathcal{L} = \sum_{m=1}^{L} I_m(y_i)\mathcal{L}_m$, and

$$d_1 = \frac{\partial\sigma_u}{\partial z_k} = \frac{\partial\exp(0.5\boldsymbol{z}\boldsymbol{\gamma})}{\partial z_k} = 0.5\sigma_u d_0.$$

In the last equation,

$$d_0 = \frac{\partial\boldsymbol{z}\boldsymbol{\gamma}}{\partial z_k}.$$

d_0 can be simply γ_k, if z_k enters $\boldsymbol{z}\boldsymbol{\gamma}$ linearly. However, if $\boldsymbol{z}\boldsymbol{\gamma} = \gamma_1 z_k + \gamma_2 z_k^2 + \cdots$, $d_0 = \gamma_1 + 2\gamma_2 z_k$. Similarly, if z_k is interacted with other variable(s), d_0 is not just a coefficient at z_k but a combination of coefficients and data. Further, $\mathcal{B}_z = \partial\mathcal{B}/\partial z_k$ is equal to

$$\mathcal{B}_z = \begin{cases} d_2(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_1) + pd_1\dfrac{\exp\left(-h_1^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} & \text{if } y_i = 1\\ \begin{bmatrix} d_2(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_m) + pd_1\dfrac{\exp\left(-h_m^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)}\\ -d_2(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_{m-1}) - pd_1\dfrac{\exp\left(-h_{m-1}^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)}\end{bmatrix} & \text{if } 1 < y_i = m < M\\ -d_2(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta})f(h_{M-1}) - pd_1\dfrac{\exp\left(-h_{M-1}^2(1+\sigma_u^2)/2\right)}{\pi(1+\sigma_u^2)} & \text{if } y_i = M\end{cases},$$

where $f(h)$ is defined above.

$$d_2 = \frac{\partial \dfrac{1}{\sqrt{1+\sigma_u^2}}}{\partial z_k} = -0.5\frac{\sigma_u^2}{\left(\sqrt{1+\sigma_u^2}\right)^3} d_0.$$

The next component $\mathcal{L}_z = \partial\mathcal{L}/\partial z_k$ is defined as

$$\mathcal{L}_z = \begin{cases} 0.5f(h_1)\left(d_3 + h_1^2 bd_2\right) + h_1 bg(h_1)d_3 & \text{if } y_i = 1 \\ \begin{bmatrix} 0.5f(h_m)\left(d_3 + h_m^2 bd_2\right) + h_m bg(h_m)d_3 \\ -0.5f(h_{m-1})\left(d_3 + h_{m-1}^2 bd_2\right) - h_{m-1}bg(h_{m-1})d_3 \end{bmatrix} & \text{if } 1 < y_i = m < M \\ -0.5f(-h_{M-1})\left(-d_3 + h_{M-1}^2 bd_2\right) - h_{M-1}bg(h_{M-1})d_3 & \text{if } y_i = M \end{cases},$$

where

$$g(h) = \phi(h)\phi(bh)$$

and

$$d_3 = \frac{\partial \dfrac{p\sigma_u}{\sqrt{1+\sigma_u^2}}}{\partial z_k} = p\frac{1}{\sqrt{1+\sigma_u^2}} d_1 + p\sigma_u d_2.$$

For a dummy variable z_k, the partial effect of z_k on inefficiency is calculated as a change in $E(u_i)$ given in (9) when at $z_k = 1$ and $z_k = 0$, which implies that the three components, $\mathcal{L}_m$, $\mathcal{B}_m$, and σ_u change.

Marginal effects of z_k on efficiency can be obtained by differentiating $\exp(-u)$

$$\frac{\partial \exp(-u)}{\partial z_k} \approx \frac{\partial \exp(-E(u))}{\partial z_k} = -\exp(-E(u))\frac{\partial E(u)}{\partial z_k} \tag{B.31}$$

Marginal effects of z_k on $\texttt{Prob}(y_i = m)$ are differently calculated for continuous and dummy variables. If the z_k is continuous, using (3), we obtain

$$\begin{aligned} \frac{\partial \texttt{Prob}(y_i = m)}{z_k} &= \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i)\right]}{\partial z_k} \\ &\approx \frac{\partial\left[\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)) - \Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i))\right]}{\partial z_k}. \end{aligned}$$

Hence,

$$\frac{\partial \texttt{Prob}(y_i = m)}{z_k} = \begin{cases} p\dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = 1 \\ p\dfrac{\partial E(u_i)}{\partial z_k}\begin{bmatrix} \phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) \\ -\phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ -p\dfrac{\partial E(u_i)}{\partial z_k}\phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pu_i) & \text{if } y_i = M \end{cases}. \tag{B.32}$$

The values in (B.32) are calculated for all observations and then the average is reported, aka APE.

For a dummy variable z_k, the partial effect of z_k on $\texttt{Prob}(y_i = m)$ is calculated as a change in probabilities when at $z_k = 1$ and $z_k = 0$, which implies that only $E(u_i)$ changes in

$$\begin{aligned} \Delta p_{im} &= \texttt{Prob}(y_i = m|z_k = 1) - \texttt{Prob}(y_i = m|z_k = 0) \\ &\approx \begin{cases} \begin{bmatrix} \Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 1) \\ -\Phi(\mu_1 - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 0) \end{bmatrix} & \text{if } y_i = 1 \\ \begin{bmatrix} \Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 1) \\ -\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 1) \\ -\Phi(\mu_m - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 0) \\ +\Phi(\mu_{m-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 0) \end{bmatrix} & \text{if } 1 < y_i = m < M \\ \begin{bmatrix} \Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 0) \\ -\Phi(\mu_{M-1} - \boldsymbol{x}_i'\boldsymbol{\beta} + pE(u_i)|z_k = 1) \end{bmatrix} & \text{if } y_i = M \end{cases}. \end{aligned} \tag{B.33}$$

Similar to the continuous case, the values in (B.33) are calculated for all observations and then the APE is reported.

Appendix C. Auxiliary tables

See Tables C.1 and C.2 and Fig. C.1

Table C.1
Different models results.

Variable	OPSF		OP		SF	
	Coef.	z	Coef.	z	Coef.	z
factor(health)2	0.325	(9.03)	0.304	(8.56)	-1.7×10^{-7}	(−0.005)
factor(health)3	0.780	(22.1)	0.734	(21.1)	2.2×10^{-6}	(0.059)
factor(health)4	1.09	(30.8)	1.02	(29.5)	3.4×10^{-6}	(0.092)
factor(health)5	1.41	(39.5)	1.33	(38.0)	4.8×10^{-6}	(0.130)
factor(education)2	-5.9×10^{-4}	(−0.059)	−0.012	(−1.27)	-3.1×10^{-8}	(−0.003)
factor(education)3	0.004	(0.356)	−0.027	(−2.46)	-2.6×10^{-7}	(−0.019)
factor(education)4	−0.055	(−5.74)	−0.094	(−9.94)	-7.8×10^{-7}	(−0.067)
factor(education)5	−0.009	(−0.789)	−0.043	(−3.95)	-2.0×10^{-7}	(−0.015)
factor(education)6	−0.067	(−6.85)	−0.104	(−10.8)	-8.5×10^{-7}	(−0.071)
factor(education)7	−0.047	(−3.89)	−0.086	(−7.36)	-8.0×10^{-7}	(−0.052)
factor(education)8	−0.025	(−2.45)	−0.068	(−6.81)	-7.6×10^{-7}	(−0.059)
factor(income)2	0.086	(8.48)	0.101	(10.1)	-1.6×10^{-7}	(−0.013)
factor(income)3	0.124	(12.7)	0.143	(14.9)	-3.0×10^{-7}	(−0.026)
factor(income)4	0.258	(26.7)	0.276	(29.2)	2.3×10^{-7}	(0.020)
factor(income)5	0.327	(34.4)	0.347	(37.3)	8.0×10^{-7}	(0.071)
factor(income)6	0.425	(41.9)	0.445	(44.9)	8.6×10^{-7}	(0.070)
factor(income)7	0.506	(47.0)	0.528	(50.1)	9.6×10^{-7}	(0.072)
factor(income)8	0.600	(49.7)	0.623	(52.7)	1.5×10^{-6}	(0.100)
factor(income)9	0.597	(37.1)	0.621	(39.3)	1.6×10^{-6}	(0.083)
factor(income)10	0.728	(42.5)	0.748	(44.6)	2.3×10^{-6}	(0.117)
log(gdppc)	0.162	(48.1)	0.157	(47.9)	8.0×10^{-7}	(0.176)
factor(wave)3	−0.309	(−15.4)	−0.321	(−16.2)	-1.5×10^{-6}	(−0.064)
factor(wave)4	−0.370	(−18.3)	−0.384	(−19.3)	-1.3×10^{-6}	(−0.058)
factor(wave)5	−0.283	(−14.2)	−0.303	(−15.5)	-1.1×10^{-6}	(−0.048)
factor(wave)6	−0.166	(−8.27)	−0.187	(−9.45)	-4.5×10^{-7}	(−0.019)
Unemployment rate	−0.021	(−45.0)	−0.022	(−48.3)	-9.0×10^{-8}	(−0.143)
Gini index	0.015	(50.7)	0.015	(49.4)	8.2×10^{-8}	(0.212)
Social protection expenditure	−0.005	(−9.18)	−0.004	(−8.40)	-7.9×10^{-8}	(−0.118)
Corruption Perception Index	−0.007	(−19.2)	−0.007	(−19.2)	-5.9×10^{-8}	(−0.122)
Quality of governance (WGI)	0.226	(26.6)	0.221	(26.6)	1.7×10^{-6}	(0.152)
(Intercept)	−5.33	(−20.5)			−1.08	(−47.9)
Age	0.216	(17.7)			0.022	(19.7)
Age squared	−0.003	(−17.3)			-1.7×10^{-4}	(−14.7)
Female	−0.386	(−10.1)			−0.007	(−1.21)
Religious	−0.645	(−14.5)			0.057	(8.53)
Unemployed	1.19	(23.2)			0.460	(43.9)
Married	−1.16	(−21.3)			−0.172	(−24.5)
μ_1	0.306	(6.23)	0.418	(8.77)		
μ_2	0.614	(12.5)	0.717	(61.7)		
μ_3	0.959	(19.6)	1.05	(117)		
μ_4	1.25	(25.6)	1.34	(160)		
μ_5	1.78	(36.5)	1.86	(354)		
μ_6	2.13	(43.5)	2.19	(358)		
μ_7	2.56	(52.4)	2.62	(504)		
μ_8	3.14	(64.2)	3.19	(687)		
μ_9	3.57	(73.0)	3.60	(564)		
(Intercept)					2.30	(42 018)
lnVARv					−22.2	(−55.0)

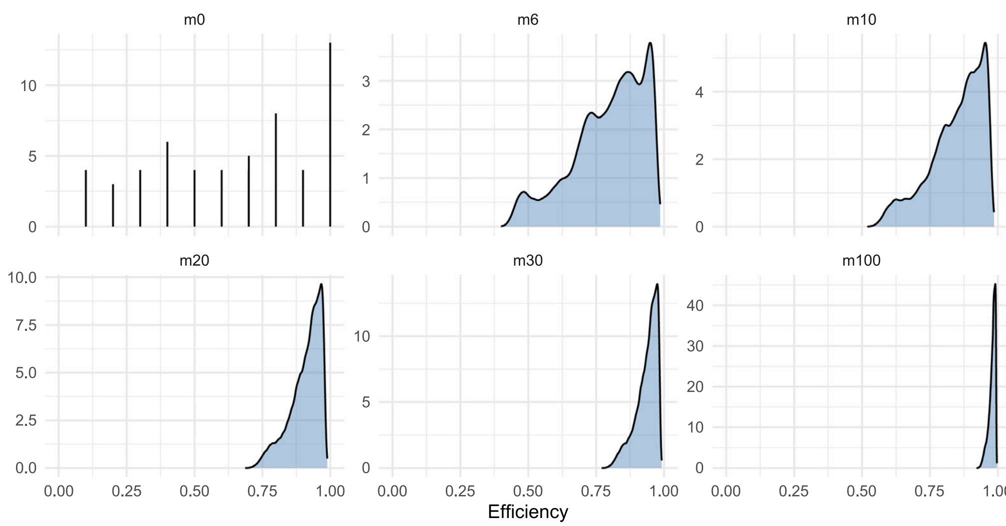
Table C.2
Estimation results of the standard SF model. *p*-value in parentheses.

Variable	Dependent variable (Y is the satisfaction)					
	log(Y)	log(Y+6)	log(Y+10)	log(Y+20)	log(Y+30)	log(Y+100)
Intercept	2.3	2.5	2.8	3.2	3.5	4.6
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(health)2	−1.74e−07	0.015	0.021	0.017	0.014	0.0052
	(0.996)	(0.003)	(0.000)	(0.000)	(0.000)	(0.000)
factor(health)3	2.17e−06	0.064	0.066	0.05	0.038	0.014
	(0.953)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(health)4	3.37e−06	0.092	0.093	0.069	0.053	0.02
	(0.927)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(health)5	4.80e−06	0.13	0.12	0.09	0.069	0.025
	(0.897)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(education)2	−3.07e−08	−0.003	−0.0025	−0.0013	−8.00e−04	−2.00e−04
	(0.998)	(0.022)	(0.025)	(0.069)	(0.116)	(0.267)
factor(education)3	−2.61e−07	−0.0061	−0.0044	−0.0021	−0.0013	−3.00e−04
	(0.985)	(0.000)	(0.001)	(0.012)	(0.040)	(0.217)
factor(education)4	−7.78e−07	−0.014	−0.011	−0.0068	−0.0047	−0.0015
	(0.947)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(education)5	−2.01e−07	−0.0065	−0.0051	−0.0028	−0.0018	−5.00e−04
	(0.988)	(0.000)	(0.000)	(0.001)	(0.003)	(0.021)
factor(education)6	−8.54e−07	−0.018	−0.014	−0.0085	−0.0059	−0.0019
	(0.944)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(education)7	−7.97e−07	−0.017	−0.013	−0.007	−0.0048	−0.0014
	(0.958)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(education)8	−7.57e−07	−0.016	−0.011	−0.0063	−0.0042	−0.0012
	(0.953)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)2	−1.59e−07	−0.0027	−2.00e−04	0.0013	0.0015	8.00e−04
	(0.989)	(0.048)	(0.858)	(0.073)	(0.008)	(0.000)
factor(income)3	−3.01e−07	−0.0038	−0.00e+00	0.0021	0.0023	0.0012
	(0.979)	(0.005)	(0.982)	(0.003)	(0.000)	(0.000)
factor(income)4	2.27e−07	0.008	0.011	0.011	0.0089	0.0037
	(0.984)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)5	7.95e−07	0.018	0.019	0.016	0.013	0.005
	(0.943)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)6	8.61e−07	0.022	0.025	0.021	0.017	0.0065
	(0.944)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)7	9.59e−07	0.028	0.032	0.026	0.02	0.008
	(0.943)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)8	1.48e−06	0.039	0.041	0.032	0.025	0.0096
	(0.920)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)9	1.62e−06	0.043	0.043	0.033	0.026	0.0097
	(0.934)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(income)10	2.26e−06	0.056	0.054	0.04	0.031	0.012
	(0.906)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log(gdppc)	8.04e−07	0.018	0.016	0.011	0.0086	0.0031
	(0.860)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(wave)3	−1.47e−06	−0.034	−0.031	−0.022	−0.016	−0.0059
	(0.949)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(wave)4	−1.34e−06	−0.033	−0.032	−0.023	−0.018	−0.0065
	(0.954)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(wave)5	−1.11e−06	−0.033	−0.03	−0.021	−0.015	−0.0054
	(0.961)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
factor(wave)6	−4.45e−07	−0.019	−0.018	−0.013	−0.0093	−0.0033
	(0.985)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Unemployment rate	−8.97e−08	−0.002	−0.0018	−0.0013	−0.001	−4.00e−04
	(0.886)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Gini index	8.21e−08	0.0017	0.0015	0.0011	8.00e−04	3.00e−04
	(0.832)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Social protection expenditure	−7.88e−08	−0.0012	−9.00e−04	−5.00e−04	−4.00e−04	−1.00e−04
	(0.906)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Corruption Perception Index	−5.95e−08	−0.001	−9.00e−04	−6.00e−04	−4.00e−04	−1.00e−04
	(0.903)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Quality of governance (WGI)	1.72e−06	0.028	0.024	0.016	0.012	0.0042
	(0.879)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
lnVARv	-22	−5.7	−5.8	−6.3	−6.8	−8.6
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Intercept	−1.1	−2.7	−3.4	−4.5	−5.3	−7.6
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age	0.022	0.02	0.021	0.023	0.024	0.027
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age squared	−2.00e−04	−2.00e−04	−2.00e−04	−3.00e−04	−3.00e−04	−3.00e−04
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

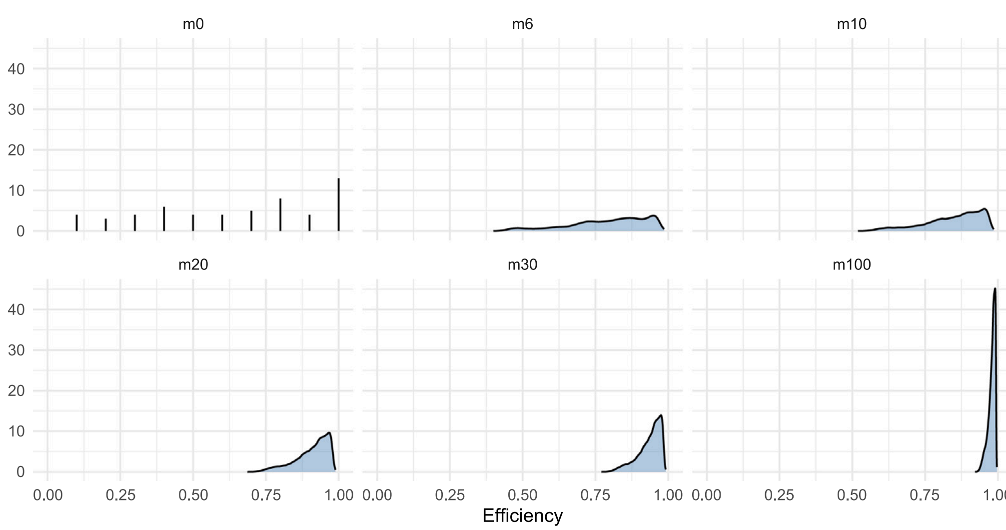
(*continued on next page*)

Table C.2 (*continued*).

Female	−0.0074	−0.036	−0.044	−0.054	−0.059	−0.07
	(0.227)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Religious	0.057	0.019	0.0047	−0.013	−0.022	−0.04
	(0.000)	(0.005)	(0.515)	(0.081)	(0.005)	(0.000)
Unemployed	0.46	0.36	0.36	0.35	0.35	0.36
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	−0.17	−0.16	−0.17	−0.19	−0.2	−0.22
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
sigma(v)	1.48e−05	0.057	0.056	0.043	0.034	0.013
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log-likelihood	−95 285.88	63 950.74	126 475.26	230 650.15	300 185.79	530 926.04



(a) Densities on own scales.



(b) Densities on the common scale.

Fig. C.1. Densities of efficiencies from models in Table C.2.

Data availability

Worldwide well-being efficiency and its determinants: New insights from the Ordered Probit Stochastic Frontier Model (Original data) (Mendeley Data)

References

Abdallah, S., Thompson, S., Marks, N., 2008. Estimating worldwide life satisfaction. Ecol. Econom. 65 (1), 35–47.

Agrawal, Sanjeev, Sharma, Neeraj, Agrawal, Rajesh, Iazzolino, Giulio, Bruni, Maria Elena, 2025. Beyond GDP: A triple helix framework for a happiness economy. Soc. Indic. Res. 179, 67–94.

Agrawal, Sanjeev, Sharma, Neeraj, Bruni, Maria Elena, Iazzolino, Giulio, 2023. Happiness economics: Discovering future research trends through a systematic literature review. J. Clean. Prod. 416, 137860.

Aragon, Yves, Daouia, Abdelaati, Thomas-Agnan, Christine, 2005. Nonparametric frontier estimation: A conditional quantile-based approach. Econometric Theory 21 (2), 358–389.

Badunenko, Oleg, Cordero, José M., Kumbhakar, Subal C., 2026. Worldwide well-being efficiency and its determinants: New insights from the ordered probit stochastic frontier model. Mendeley Data V3.

Binder, Martin, Broekel, Tom, 2011. Applying a non-parametric efficiency analysis to measure conversion efficiency in Great Britain. J. Hum. Dev. Capab. 12 (2), 257–281.

Binder, Martin, Broekel, Tom, 2012. Happiness no matter the cost? An examination on how efficiently individuals reach their happiness levels. J. Happiness Stud. 13 (4), 621–645.

Bjørnskov, Christian, Dreher, Axel, Fischer, Justina A.V., 2008. Cross-country determinants of life satisfaction: Exploring different determinants across groups in society. Soc. Choice Welf. (ISSN: 1432-217X) 30 (1), 119–173.

Bjørnskov, Christian, Dreher, Axel, Fischer, Justina A.V., 2010. Formal institutions and subjective well-being: Revisiting the cross-country evidence. Eur. J. Political Econ. 26 (4), 419–430.

Blanchflower, David G., Oswald, Andrew J., 2008. Is well-being U-shaped over the life cycle? Soc. Sci. Med. 66 (8), 1733–1749.

Bonanno, Graziella, D'Orio, Gianpiero, Lombardo, Rosaria, 2020. Generating well-being and efficiency: Evidence from Italy. Econ. Anal. Policy 65, 262–275.

Briguglio, Marie, Czap, Natalia, Laffan, Kate, 2025. Wellbeing and Policy: Evidence for Action. Taylor & Francis.

Carboni, Oliviero Antonio, Russu, Paolo, 2015. Assessing regional wellbeing in Italy: An application of Malmquist–DEA and self-organizing map neural clustering. Soc. Indic. Res. 122 (3), 677–700.

Castellano, Rosalia, De Bernardo, Giuseppe, Punzo, Gianluigi, 2023. Well-being in OECD countries: An assessment of technical and social efficiency using data envelopment analysis. Int. Rev. Econ. 70 (2), 141–176.

Cazals, Catherine, Florens, Jean-Pierre, Simar, Leopold, 2002. Nonparametric frontier estimation: A robust approach. J. Econometrics 106 (1), 1–25.

Cheng, Terence C., Powdthavee, Nattavudh, Oswald, Andrew J., 2017. Longitudinal evidence for a midlife nadir in human well-being: Results from four data sets. Econ. J. 127 (599), 126–142.

Clark, Andrew E., 2018. Four decades of the economics of happiness: Where next? Rev. Income Wealth 64 (2), 245–269.

Clark, A.E., Frijters, P., Shields, M.A., 2008. Relative income, happiness, and utility: An explanation for the Easterlin paradox and other puzzles. J. Econ. Lit. 46 (1), 95–144.

Cordero, Jose Manuel, Polo, Cristina, Salinas-Jiménez, Javier, 2021. Subjective well-being and heterogeneous contexts: A cross-national study using semi-nonparametric frontier methods. J. Happiness Stud. 22 (2), 867–886.

Cordero, José Manuel, Salinas-Jiménez, Javier, Salinas-Jiménez, M Mar, 2017. Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis. European J. Oper. Res. 256 (2), 663–672.

Daraio, Cinzia, Simar, Leopold, 2005. Introducing environmental variables in nonparametric frontier models: A probabilistic approach. J. Prod. Anal. 24 (1), 93–121.

Debnath, Ramesh Mohan, Shankar, Ravi, 2014. Does good governance enhance happiness: A cross nation study. Soc. Indic. Res. 116 (1), 235–253.

Diener, Ed, Oishi, Shigehiro, Tay, Louis, 2018. Handbook of Well-Being. DEF Publishers, Salt Lake City, UT.

Dolan, Paul, Peasgood, Tessa, White, Mathew, 2008. Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being. J. Econ. Psychol. 29 (1), 94–122.

Dominko, Mitja, Verbič, Miroslav, 2019. The economics of subjective well-being: A bibliometric analysis. J. Happiness Stud. 20 (6), 1973–1994.

Dziechciarz, Józef, 2025. Application of the DEA method for the assessment of the efficiency of achieving well-being in EU countries in the years 2013–2022. Sustainability 17 (2), 6375.

Emmerling, Johannes, Navarro, Pablo, Sisco, Matthew R., 2021. Subjective well-being at the macro level—Empirics and future scenarios. Soc. Indic. Res. 157 (3), 899–928.

Exton, Conal, Smith, Carrie, Vandendriessche, Delphine, 2015. Comparing Happiness Across the World. OECD Statistics Directorate Working Paper 62, OECD Publishing, Paris.

Fleurbaey, Marc, Blanchet, Didier, 2013. Beyond GDP: Measuring Welfare and Assessing Sustainability. Oxford University Press.

Frey, B.S., 2008. Happiness: A Revolution in Economics. The MIT Press, Cambridge, Massachusetts.

Frey, B., Stutzer, A., 2010. Happiness and Economics: How the Economy and Institutions Affect Human Well-Being. Princeton University Press. Princeton, New Jersey.

Frey, Bruno S., Stutzer, Alois (Eds.), 2013. Recent Developments in the Economics of Happiness. Edward Elgar, Cheltenham.

Frijters, Paul, Krekel, Christian, 2021. A Handbook for Wellbeing Policy-Making: History, Theory, Measurement, Implementation, and Examples. Oxford University Press.

Headey, Bruce, Muffels, Ruud, 2018. A theory of life satisfaction dynamics: Stability, change and volatility in 25-year life trajectories in Germany. Soc. Indic. Res. 140 (2), 837–866.

Helliwell, J., Huang, H.F., 2008. How's your government? International evidence linking good government and well-being. Br. J. Political Sci. 38, 595–619.

Helliwell, John F., Huang, Haifang, 2008. How's your government? International evidence linking good government and well-being. Br. J. Political Sci. 38 (4), 595–619.

Helliwell, John F., Layard, Richard, Sachs, Jeffrey D., De Neve, Jan-Emmanuel, Aknin, Lara B., Wang, Shun (Eds.), 2025. World Happiness Report 2025. University of Oxford, Wellbeing Research Centre.

Inglehart, Ronald, Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., Puranen, B., 2022. World values survey: All rounds - country-pooled datafile.

Kaufmann, D., Kraay, A., Mastruzzi, M., 2014. Worldwide Governance Indicators. World Bank.

Kimball, Miles S., Willis, Robert J., 2023. Utility and Happiness. Working Paper w31707, National Bureau of Economic Research.

Kumbhakar, Subal C., Lovell, C. A. Knox, 2003. Stochastic Frontier Analysis. Cambridge University Press.

Kuosmanen, T., Kortelainen, M., 2012. Stochastic non-smooth envelopment of data: Semi-parametric frontier estimation subject to shape constraints. J. Prod. Anal. 38 (1), 11–28.

Langbein, L., Knack, S., 2010. The worldwide governance indicators: Six, one, or none? J. Dev. Stud. 46, 350–370.

Mamatzakis, E.C., Tsionas, M.G., 2021. Making inference of British household's happiness efficiency: A Bayesian latent model. European J. Oper. Res. 294 (1), 312–326.

Mizobuchi, Hideyuki, 2017. Measuring socio-economic factors and sensitivity of happiness. J. Happiness Stud. 18 (2), 463–504.

Moro-Egido, Ana I., Navarro, Miguel, Sánchez, Antonio, 2022. Changes in subjective well-being over time: Economic and social resources do matter. J. Happiness Stud. 23 (5), 2009–2038.

Nettle, D., 2005. Happiness: The Science Behind Your Smile. Oxford University Press, Oxford.

Nikolova, Milena, Graham, Carol, 2021. The economics of happiness. In: Zimmermann, Klaus F. (Ed.), Handbook of Labor, Human Resources and Population Economics. Springer International Publishing, pp. 1–33.

Nikolova, Milena, Popova, Olga, 2021. Sometimes your best just ain't good enough: The worldwide evidence on subjective well-being efficiency. BE J. Econ. Anal. Policy 21 (1), 83–114.

O'Connor, K.J., 2017. Happiness and welfare state policy around the world. Rev. Behav. Econ. 4 (4), 397–420.

Owen, D.B., 1980. A table of normal integrals. Comm. Statist. Simulation Comput.

Pacek, Alexander C., Radcliff, Benjamin, 2008. Welfare policy and subjective well-being across nations: An individual-level assessment.

Pena-López, Atilano, Rungo, Paolo, López-Bermúdez, Beatriz, 2020. The "Efficiency" effect of conceptual referents on the generation of happiness: A cross-national analysis. J. Happiness Stud. 22 (6), 2457–2483.

Rayo, Luis, Becker, Gary S., 2007. Evolutionary efficiency and happiness. J. Political Econ. 115 (2), 302–337.

Rojas, Mariano, 2024. The joint enjoyment of life. Explaining high happiness in Latin America. J. Happiness Stud. 25 (7), 100.

Sarracino, Francesco, O'Connor, Kelsey J., 2022. A measure of well-being efficiency based on the world happiness report. Int. Prod. Monit. 43, 10–40.

Schimmack, Ulrich, Oishi, Shigehiro, 2005. The influence of chronically and temporally accessible information on life satisfaction judgments. J. Pers. Soc. Psychol. 89, 395–406.

Steffel, Mary, Oppenheimer, Daniel M., 2009. Happy by what standard? The role of interpersonal and intrapersonal comparisons in ratings of happiness. Soc. Indic. Res. 92 (1), 69–79.

Stone, Arthur A., Schneider, Stefan, Krueger, Alan, Schwartz, Joseph E., Deaton, Angus, 2018. Experiential wellbeing data from the American Time Use Survey: Comparisons with other methods and analytic illustrations with age and income. Soc. Indic. Res. 136 (1), 359–378.

Thai, Cao Hao An, 2025. Using stochastic frontier analysis to measure the satisfactory efficiency of countries. Ho Chi Minh City Open Univ. J. Sci. – Econ. Bus. Adm. 15 (3), 93–106.
Tiley, Kirsty, Crellin, Ruby, Domun, Tulsi, Harkness, Francesca, Blodgett, Joanna M., 2024. Effectiveness of 234 interventions to improve life satisfaction: A rapid systematic review. Soc. Sci. Med. 117662.
Veenhoven, R., 1991. Is happiness relative? Soc. Indic. Res. 24 (1), 1–34.
Verme, Paolo, 2011. Life satisfaction and income inequality. Rev. Income Wealth 57 (1), 111–127.
Ye, Dingding, Ng, Yew-Kwang, Lian, Yujun, 2015. Culture and happiness. Soc. Indic. Res. 123 (2), 519–547.
Zhou, Yong, Huang, Xianzhong, Shen, Yue, Tian, Linlin, 2023. Does targeted poverty alleviation policy lead to happy life? Evidence from rural China. China Econ. Rev. 81, 102037.